

INFODEMIA NA ERA DA COVID-19: UMA ANÁLISE DOS DISCURSOS ANTIVACINA NAS REDES SOCIAIS

INFODEMIC IN THE COVID-19 ERA: AN ANALYSIS OF ANTI-VACCINE DISCOURSES ON SOCIAL MEDIA

Mirella WHITEMAN

mirellawhiteman@gmail.com

Pontifícia Universidade Católica de São Paulo, São Paulo, Brasil

Resumo: Na última década, os movimentos antivacina ganharam força, impulsionados pelas redes sociais. O ressurgimento de doenças como sarampo e poliomielite tornou-se uma preocupação global, afetando também o Brasil. Desde 2015, observa-se uma queda na cobertura do Programa Nacional de Imunizações (PNI), tendência agravada pela pandemia de Covid-19. Durante o período pandêmico, as vacinas e sua eficácia tornaram-se temas centrais no discurso público. Esse destaque foi intensificado pelo impacto direto das plataformas digitais e suas sofisticadas redes de informação, muito utilizadas por movimentos contrários a vacinação para disseminar a hesitação em relação à imunização. Este estudo analisa os discursos antivacina na plataforma Twitter (atual X) no Brasil, utilizando a Análise Multidimensional Lexical (Berber Sardinha, 2016, 2017, 2020, 2021) para examinar um corpus de 16.841 tweets. Após etiquetagem, lematização e análise fatorial, identificaram-se três dimensões discursivas: (1) Alerta sobre riscos da dose de reforço em jovens versus Resistência a medidas compulsórias de combate à Covid-19, (2) Teorias político-conspiratórias sobre a vacina versus Evidências pseudocientíficas contra a vacina e (3) Defesa do tratamento precoce em detrimento da vacina versus Relatos pessoais sobre efeitos colaterais da vacina. Esta pesquisa contribui para a compreensão dos discursos antivacina no Brasil nas redes sociais, oferecendo evidências empíricas para o enfrentamento da hesitação vacinal e desinformação.

Palavras-chave: Vacina; Linguística de Corpus; Análise Multidimensional Lexical; Análise do Discurso.

Abstract: Over the past decade, anti-vaccine movements have gained momentum, fueled by social media. The resurgence of diseases such as measles and polio has become a global concern, also affecting Brazil. Since 2015, a decline in coverage of the National Immunization Program has been observed, a trend exacerbated by the Covid-19 pandemic. During this period, vaccines and their efficacy became central topics in public discourse, a focus that was intensified by the direct influence of digital platforms and their sophisticated information networks, which anti-vaccine movements have extensively used to spread vaccine hesitancy. This study analyzes anti-vaccine discourses on Twitter (now X) in Brazil, using Lexical Multidimensional Analysis (Berber Sardinha, 2016, 2017, 2020, 2021). to examine a corpus of 16,841 tweets. After tagging, lemmatization, and factor analysis, three discursive dimensions were identified: (1) Warnings about booster dose risks in young people versus Resistance to compulsory measures against Covid-19, (2) Political-conspiracy theories versus Pseudoscientific evidence against vaccines, and (3) Advocacy for early treatment versus Personal accounts of vaccine side effects. This research contributes to the understanding of anti-vaccine discourses in Brazil on social media, providing empirical evidence for addressing vaccine hesitancy and misinformation.

Keywords: Vaccine; Corpus Linguistics; Lexical Multidimensional Analysis; Discourse Analysis.

INTRODUÇÃO

O final de 2019 marcou o início de uma crise global sem precedentes. Em 31 de dezembro, autoridades chinesas notificaram a Organização Mundial da Saúde (OMS) sobre um surto de pneumonia de origem desconhecida em Wuhan. Este evento desencadeou uma série de acontecimentos que culminariam na pandemia de Covid-19 (Haraki, 2021), que surgiu acompanhada por um fenômeno paralelo: a infodemia.

A infodemia refere-se à proliferação excessiva de informações, muitas vezes não verificadas, associadas a um tema específico. O termo foi inicialmente introduzido pelo cientista político David Rothkopf durante a pandemia de SARS, em 2003, referindo-se a um fenômeno em que “fatos, combinados com medo, especulação e boatos, são amplificados e rapidamente disseminados globalmente por meio das tecnologias de informação modernas” (Rothkopf, 2003). Este fenômeno, amplificado pelas plataformas digitais, dificulta a identificação de fontes confiáveis e pode gerar desconfiança nas autoridades, comprometendo as respostas de saúde pública (OMS, 2020).

As redes sociais, com seu alcance global e instantâneo, tornaram-se terreno fértil para a propagação de teorias conspiratórias e desinformação sobre a Covid-19 e vacinas (Taylor, 2020). O *Twitter* (atual X), em particular, destacou-se como plataforma central neste contexto, especialmente no Brasil, onde foi adotado como canal oficial de comunicação pelo então presidente Jair Bolsonaro durante a pandemia. Ademais, o *Twitter* oferece um ambiente que recompensa a comunicação curta, agressiva e descontextualizada (Haigh; Haigh, 2020), tornando-o a arena ideal para a circulação de desinformação.

A rápida disseminação de informações não verificadas e teorias conspiratórias nas redes sociais contribuiu para um clima de incerteza e desconfiança. Isso se refletiu na hesitação vacinal, fenômeno reconhecido pela OMS como uma das dez principais ameaças à saúde global (OPAS, 2019). Durante surtos anteriores, como os de ebola e cólera, teorias da conspiração resultaram em agressões e até assassinatos de profissionais de saúde (Cohn, 2017).

A pandemia catalisou um esforço global para o desenvolvimento de vacinas. Apesar do rigor científico mantido nos ensaios clínicos, a velocidade do processo gerou desconfiança em parte da população. Esta desconfiança foi alimentada por narrativas antivacina que ganharam força nas redes sociais, criando um paradoxo: enquanto a ciência oferecia uma solução para a crise sanitária, uma parcela significativa da sociedade resistia à vacinação.

Diante deste cenário, torna-se crucial compreender os discursos que circularam nas plataformas digitais sobre as vacinas no período pandêmico. Esta pesquisa propõe-se a analisar os discursos emergentes em torno do movimento antivacina no Brasil, focando especificamente no *Twitter*, no período de 2020 a 2022.

Estudos anteriores envolvendo *big data* concentraram-se predominantemente em aspectos como sentimentos e tópicos (e.g., Jang et al., 2021; Wicke & Bolognesi, 2020), sem direcionar a atenção para os discursos. Por outro lado, pesquisas em Análise do Discurso durante a pandemia trabalharam com volumes menores de dados (e.g., Recuero & Stumpf, 2021; Jones, 2021). Neste estudo, ancorado na Linguística de Corpus, será adotada a Análise Multidimensional Lexical (Berber Sardinha, 2016, 2017, 2020, 2021; Fitzsimmons-Doolan, 2014), abordagem que possibilita investigar discursos em *corpora* de grande escala. Dessa maneira, será possível mapear as dimensões lexicais e os principais discursos relacionados ao movimento antivacina.

As perguntas centrais que norteiam este estudo são: 1) Quais são as dimensões lexicais dos discursos em torno do movimento antivacina? 2) Quais os principais discursos associados a essas dimensões lexicais?

Ao responder a estas questões, busca-se contribuir para uma compreensão mais profunda dos fatores que influenciam a percepção coletiva no Brasil sobre vacinação e saúde pública em tempos de crise. A análise desses discursos nas redes sociais é fundamental para entender como as informações são configuradas e disseminadas, influenciando percepções e decisões em relação à vacinação em um contexto de pandemia global.

FUNDAMENTAÇÃO TEÓRICA

Este estudo emprega os princípios da Linguística de Corpus e Análise Multidimensional, para examinar a disseminação de discursos nas plataformas de mídia social. A pesquisa visa preencher uma lacuna significativa na pesquisa linguística contemporânea, onde a investigação em larga escala e sistemática dos discursos em ambientes digitais permanece um território relativamente inexplorado.

Linguística de Corpus

A Linguística de Corpus (LC) é uma abordagem que estuda a língua em uso e se baseia na análise sistemática de grandes volumes de dados linguísticos autênticos, permitindo a observação de padrões e variação em contextos reais de comunicação (Conrad, 2011). Com o desenvolvimento de tecnologias computacionais, a LC consolidou-se como uma metodologia que combina elementos automatizados e intervenção humana, tornando possível a análise de aspectos linguísticos complexos. De acordo com Berber Sardinha (2004, p. 3), a LC se ocupa da coleta e exploração de *corpora* — conjuntos de dados linguísticos textuais criteriosamente coletados — com o objetivo de representar variedades específicas da língua, utilizando evidências empíricas extraídas por computador.

Essa abordagem parte do pressuposto de que a língua é um sistema probabilístico no qual as combinações de palavras não ocorrem ao acaso, sendo possível identificar correlações entre traços linguísticos e contextos situacionais (Berber Sardinha, 2004). Conforme Berber Sardinha e Moreira (2023), a LC utiliza coleções de textos em diversos modos semióticos (falados, escritos, sonoros ou

visuais) para descrever situações de uso das linguagens e compreender a relação entre os recursos de expressão e seus contextos de produção. A LC tem contribuído significativamente para pesquisas em ensino de línguas, análise de gênero, tradução e linguística forense (Berber Sardinha, 2009), ao fornecer uma nova perspectiva baseada em dados empíricos e na análise da linguagem real em uso.

Análise Multidimensional e Análise Multidimensional Lexical

A Análise Multidimensional (AMD) constitui o arcabouço teórico e metodológico desta pesquisa, sendo aplicada aqui em sua vertente Lexical (Berber Sardinha, 2016, 2017, 2020, 2021). A AMD foi inicialmente desenvolvida por Douglas Biber na década de 1980 (Biber, 1988) para descrever a variação entre registros sob uma perspectiva funcional. Biber explorou a variação de registros em várias línguas, como inglês, espanhol e coreano (Biber, 1995), além de contextos específicos, como o meio acadêmico (Biber, 2006).

No contexto da AMD, o termo 'registro' refere-se a qualquer variedade textual associada a situações específicas de comunicação ou a objetivos comunicacionais (Biber, 1995). Os traços linguísticos que compõem os textos variam de maneira sistemática com base nos contextos e funções comunicativas em que são produzidos, o que significa que essa variação não é aleatória. A AMD possibilita a análise empírica das características linguísticas que coocorrem em registros específicos e a comparação de variação entre registros com base nos padrões de coocorrência desses traços. Esses atributos são agrupados em dimensões, que representam conjuntos de características linguísticas identificadas por meio de análise fatorial e interpretadas em termos de funções comunicativas comuns (Biber, 1995).

A partir do modelo multidimensional de Biber (1988), Berber Sardinha (2014) propôs uma nova abordagem que se concentra na análise discursiva, substituindo variáveis gramaticais por variáveis lexicais. Para tanto, o autor investigou o uso dos adjetivos *American* e *Brazilian*, bem como suas colocações – palavras que ocorrem próximas ao núcleo – para identificar os parâmetros de representação de identidade nacional e cultural através dos quais os EUA e o Brasil são retratados nas produções textuais em inglês a partir do século XIX disponíveis no Google Books (Delfino, 2021). Sob a perspectiva lexical, a AMD utilizada no estudo considerou como variáveis apenas as palavras de conteúdo ou combinações de palavras para a identificação das dimensões de variação.

Na vertente Lexical da AMD (Berber Sardinha, 2016, 2017, 2020, 2021), o objetivo é mapear parâmetros de variação lexical em *corpora*, que podem indicar temas e discursos subjacentes. Assim como a AMD Funcional, a versão Lexical utiliza análise estatística multivariada, especificamente a análise fatorial, para identificar variáveis latentes que operam abaixo do nível perceptível. Tais variáveis são agrupadas em dimensões de variação discursiva, compostas por conjuntos de itens lexicais que tendem a ocorrer juntos em diferentes tipos de textos (Berber Sardinha; Moreira, 2023).

Embora a AMD Funcional e a AMD Lexical compartilhem princípios metodológicos, a abordagem Lexical seleciona variáveis com base em palavras lematizadas (formas canônicas), e sua lista de variáveis é aberta, diferentemente da versão funcional, que se limita às categorias gramaticais identificadas por softwares de etiquetagem (Araújo; Berber Sardinha; Delfino, 2018).

Com base em Berber Sardinha (2016, 2017, 2020, 2021), o passo a passo para a realização da AMD Lexical nesta pesquisa envolve:

1. Identificação e contagem das palavras;
2. Normalização das frequências das variáveis lexicais;
3. Extração fatorial inicial não rotacionada baseada nas frequências normalizadas para identificar os fatores a serem utilizados;
4. *Scree plot*: definição do número de fatores para análise por meio de um gráfico de análise de sedimentação;
5. Eliminação das variáveis lexicais com comunalidades menores que 0,15 (Cf. Biber, 2006, p. 183);
6. Extração fatorial final rotacionada contendo o número de fatores estabelecidos para análise;
7. Cálculo da quantidade de variação compartilhada pelos fatores extraídos;
8. Checagem da variância dos fatores;
9. Cálculo dos escores de fator de cada texto;
10. Interpretação dos fatores em termos de seus discursos subjacentes por meio da observação dos textos, registros e variáveis.

A AMD Lexical se mostra particularmente versátil para análises baseadas em léxico, sendo útil na identificação de construções discursivas e de estruturas ideológicas subjacentes. Essa abordagem é convergente com metodologias assistidas por *corpora* na análise do discurso, enfatizando a importância das escolhas lexicais como indicadores de projetos discursivos mais amplos (Berber Sardinha; Fitzsimmons-Doolan, no prelo).

Discurso

Estudos anteriores têm abordado a infodemia predominantemente sob a perspectiva da quantidade excessiva e da rápida disseminação de textos verbais. No entanto, sugerimos aqui que a infodemia possui um alcance mais amplo do que apenas o excesso de informações, sendo, na verdade, caracterizada por uma proliferação de discursos.

Esta pesquisa se baseia na premissa de que os discursos são "maneiras de enxergar o mundo, de construir objetos e conceitos de certos modos, de representar a realidade" (Baker & McEnery, 2015, p. 5). Além disso, reconhece que os discursos "enquadram certos problemas, [distinguindo] certos aspectos da situação em vez de outros" (Hajer, 1993, p. 45). Ao analisar os discursos antivacina, pretende-se examinar não apenas o conteúdo das mensagens, mas também as visões de mundo, ideologias e representações que indexam valores, intenções e modos de agir (Foucault & Faubion, 2000; Van Dijk, 2008).

METODOLOGIA

Corpus

O *corpus* é composto por postagens em língua portuguesa coletadas na plataforma *Twitter* entre os anos de 2020 e 2022. Para estruturá-lo, o primeiro passo foi definir os critérios de busca adequados para delimitar o escopo da pesquisa. As postagens foram extraídas utilizando termos específicos como "vacina", "pfizer", "astrazeneca", "coronavac", "janssen", "dose" e "vachina", entre outros. A coleta foi realizada com o uso da ferramenta *Twarc*, que consiste em uma biblioteca para *Python* e uma linha de comando capaz de obter e arquivar dados em formato JSON a partir da API do *Twitter*.

A construção do *corpus* foi baseada nos critérios propostos por Berber Sardinha (2004): (1) uso de textos autênticos que refletem a linguagem natural; (2) elaborados por falantes em situações reais de comunicação; (3) seleção cuidadosa para assegurar a representatividade conforme as metas do estudo.

Visão geral do processamento do *corpus*

Após a coleta das postagens, o primeiro passo é selecionar, a partir de uma lista inicial, as palavras que poderão ser consideradas como palavras-chave. Esse processo envolve a escolha de palavras de conteúdo, como substantivos, verbos, adjetivos e advérbios, que servirão de base para a análise na AMD Lexical. A lista resultante é composta por termos candidatos a palavras-chave, que posteriormente serão analisados usando o método estatístico *Log-Likelihood*, que permite verificar se a frequência observada de uma palavra no *corpus* é maior do que a frequência esperada no *corpus* de referência. As palavras que apresentarem maior valor de chavidade (*keyness*) serão consideradas palavras-chave. Em seguida, são analisadas as correlações dessas palavras para selecionar aquelas com maior valor de associação. Na última etapa, aplica-se um filtro de comunalidade, que mede a proporção da variância de cada variável explicada pelos fatores. A comunalidade define um valor de corte, eliminando termos que possuam um peso fatorial inferior a 0,3. Por fim, o Princípio da Parcimônia é empregado para garantir que apenas os elementos essenciais permaneçam na análise, assegurando que a precisão não seja comprometida.

Seleção de variáveis com base em palavras-chave

Palavras-chave são termos cuja frequência se destaca em comparação com um *corpus* de referência, sugerindo conteúdos de interesse específico, como proposto por Scott (2000). Na AMD Lexical, as categorias incluem substantivos, verbos, adjetivos e advérbios. Essa distinção revela os tópicos predominantes e destaca variações significativas entre diferentes textos ou *corpora*. Para identificar tais palavras no *corpus*, utilizou-se um script em *Shell*, destacando as palavras com maior chavidade em relação ao *corpus* de referência. Esse processo é fundamentado no teste estatístico *Log-Likelihood*, que quantifica a relevância das palavras.

Seleção de variáveis por correlação

Após a seleção inicial das variáveis por palavras-chave, é construída uma matriz de correlação para identificar as variáveis que exibem fortes associações. O objetivo é eliminar correlações de baixo valor estatístico e otimizar a etapa de análise fatorial subsequente. As correlações são determinadas pela frequência de coocorrência de duas palavras em um mesmo texto, comparada à frequência de não coocorrência. A direção da correlação (positiva ou negativa) é avaliada para cada par de termos. Por exemplo, uma correlação positiva entre "tratamento" e "precoce" indica que ambos os termos ocorrem juntos frequentemente, enquanto uma correlação negativa sugere que a presença de um termo implica a ausência do outro. Independentemente da direção, o foco está nas correlações mais altas, sejam associações ou dissociações. Ao final dessa etapa, o *corpus* está pronto para ser analisado.

Composição do corpus

O Quadro 1 apresenta a composição final do *corpus*, que contém um total de 16.841 postagens e 167.197 palavras. Com base em critérios de classificação de *corpora*, o *corpus* é composto por textos em português brasileiro (monolíngue), escritos por usuários supostamente brasileiros, e coletados na rede social *Twitter*. As postagens foram coletadas entre março de 2020 e novembro de 2022 e armazenadas no formato *.txt*, codificadas em UTF-8.

Quadro 1 – Composição do corpus

	Número total
Postagens	16.841
Palavras (<i>tokens</i>)	167.197

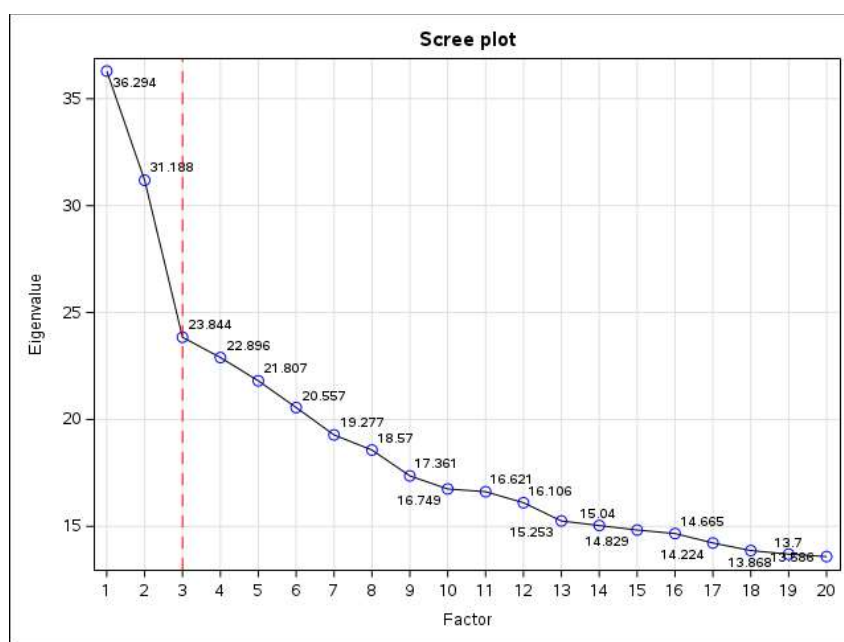
Fonte: Elaborado pela autora.

Análise Multidimensional Lexical

A metodologia da AMD Lexical é estruturada em três etapas principais: (1) o pré-processamento do *corpus*, que abrange a etiquetagem, normalização e seleção das variáveis (traços linguísticos); (2) a análise fatorial inicial (não rotacionada), que define o número de fatores a serem extraídos com base no gráfico de escarpa (*scree plot*); e (3) a análise fatorial rotacionada, que ajusta o número final de fatores, exclui variáveis de baixa carga, calcula escores e rotula as dimensões por meio de análises interpretativas.

Para etiquetar o *corpus*, foi utilizado o *TreeTagger* (Schmid, 2013) para a língua portuguesa. Em seguida, verificações manuais garantiram a correta identificação de substantivos, verbos, adjetivos, emojis e hashtags. Os lemas dessas categorias, juntamente com emojis e hashtags, foram organizados em planilhas *csv*. A análise fatorial inicial gerou o gráfico de sedimentação, representado no Gráfico 1:

Gráfico 1 – Scree Plot da Extração Fatorial



Fonte: Elaborado pela autora.

Após análise dos *eigenvalues* (autovalores), diferentes números de fatores foram considerados, variando entre duas e quatro soluções. A extração de três fatores foi a mais interpretável, como indicado pela linha tracejada no gráfico, e adotada como solução final. A extração fatorial foi realizada com a rotação *Promax*, considerando correlações entre fatores. O resultado identificou seis fatores, discutidos na seção seguinte.

Para avaliar a relação dos textos com os fatores, utilizou-se a Análise de Variância (ANOVA), que permite identificar os textos mais representativos para cada fator. Até esse ponto, a análise é quantitativa. A partir desses resultados, o estudo avança para um enfoque qualitativo, investigando como as variáveis selecionadas aparecem nas linhas de concordância dos textos. Com base nessa interpretação, as dimensões predominantes no *corpus* foram identificadas.

Após definir as dimensões, atribuiu-se um rótulo a cada uma, a fim de refletir os principais discursos indicados pelas variáveis cocorrentes. As etapas subsequentes detalham o processo de rotulação e interpretação dos discursos identificados.

RESULTADOS

Análise das dimensões

Com base em variáveis lexicais, o objetivo é investigar a variação discursiva presente no *corpus* por meio da metodologia da AMD Lexical, conforme estabelecido por Berber Sardinha (2016, 2017, 2020, 2021). A análise busca explorar como os discursos antívacina são organizados em termos de arranjos dimensionais no *corpus*.

No Quadro 2, são apresentadas as três dimensões encontradas, todas com polos positivos e negativos, e suas respectivas denominações:

Quadro 2 – Rótulos das dimensões discursivas

Dimensão	Rótulo curto	Rótulo longo
1	Alerta sobre riscos da dose de reforço em jovens versus Resistência a medidas compulsórias de combate à Covid-19	Alerta sobre riscos cardíacos, males súbitos e falecimentos após doses de reforço, especialmente entre jovens e crianças, considerando a vacina como experimental versus Questionamento e resistência à imposição de medidas e decretos compulsórios de combate ao avanço da Covid-19
2	Teorias político-conspiratórias sobre a vacina versus Evidências pseudocientíficas contra a vacina	Alegações sobre a origem do vírus em laboratório chinês e interesses de políticos e governos ligados a grandes grupos farmacêuticos versus Evidências de estudos pseudocientíficos publicados sobre os efeitos adversos das vacinas
3	Defesa do tratamento precoce em detrimento da vacina versus Relatos pessoais sobre efeitos colaterais da vacina	Defesa do tratamento precoce frente à letalidade do vírus e à falta de estudos sobre a eficácia e a segurança em longo prazo da vacina emergencial versus Relatos sobre amigos e terceiros que sofreram com efeitos colaterais graves – dores de cabeça, paralisia, problemas no rim, nas pernas e nos pés – e até fatais após a vacinação

Fonte: Elaborado pela autora.

Neste ponto, é relevante enfatizar que as dimensões discursivas englobam uma variedade de significados inter-relacionados. Por isso, a atribuição de rótulos às dimensões requer uma elaboração criteriosa para refletir essa complexidade de forma abrangente. Contudo, para fins de clareza no desenvolvimento do texto, utilizam-se rótulos abreviados, que são empregados nos títulos das seções subsequentes. Os rótulos completos, conforme apresentados no quadro, oferecem uma descrição mais detalhada e interpretativa das dimensões analisadas. Ambos os tipos de rótulos (curtos e completos) representam os mesmos discursos subjacentes, diferenciando-se apenas na extensão, e não no conteúdo.

Por fim, vale destacar que todos os tweets citados nas próximas seções estão entre aqueles que apresentaram os maiores escores em cada dimensão, tanto no polo positivo quanto no negativo. Em outras palavras, são enunciados que apresentam alta carga de variáveis lexicais (grifadas em cada um deles) características de cada dimensão identificada.

Dimensão 1 – Alerta sobre riscos da dose de reforço em jovens versus Resistência a medidas compulsórias de combate à Covid-19

Quadro 3 – Variáveis da Dimensão 1

Polo	Variáveis lexicais
Positivo (65 variáveis)	reforço (0,95635), fonte (0,94996), infarto (0,93656), semana (0,86775), indicar (0,86442), cardíaco (0,86312), falecer (0,83739), publicar (0,83478), pesquisa (0,82977), experimento (0,79311), sofrer (0,76066), jovem (0,74540), pós (0,72900), facebook (0,71919), dia (0,66741), dose (0,65392), súbito (0,65113), alertar (0,63157), criança (0,62101), alerta (0,60826), mês (0,60387), morte (0,58873), fulminante (0,57859), org (0,57492), mal (0,55586), doença (0,55347), morrer (0,54771), estudo (0,54304), saudável (0,52843), miocardite (0,52303), grave (0,49075), global (0,48340), bvsalud (0,47183), risco (0,45471), causar (0,45057), literat (0,44981), óbito (0,44769), primeiro (0,44696), último (0,42515), autoimune (0,42477), ano (0,42149), mrna (0,41604), inocular (0,40904), evento (0,39380), internar (0,38734), articles (0,38307), syringe_e (0,37561), contrair (0,37057), caso

	(0,36688), dar (0,35343), novo (0,34916), comprovante (0,34781), adverso (0,34257), causa (0,34164), mulher (0,34140), imunidade (0,33854), vida (0,33504), sistema (0,33086), amigo (0,32808), pessoa (0,32791), exame (0,32190), vacinal (0,31053), maioria (0,30991), existir (0,30724), hepatite (0,30286)
Negativo (52 variáveis)	enquete (-0,64958), proposta (-0,60021), imposição (-0,54932), audiência (-0,48156), consciência (-0,47040), termo (-0,46066), gel (-0,44092), Covid19 (-0,44058), zinco (-0,43595), serviço (-0,42600), mercado (-0,42056), tutelar (-0,41945), correiopaulista (-0,41739), dama (-0,41707), piada (-0,41382), álcool (-0,41347), pleno (-0,40453), questionamento (-0,39936), ferro (-0,39698), decreto (-0,39473), combate (-0,39455), restaurante (-0,38842), retirar (-0,38422), lista (-0,38409), conflito (-0,38236), calça (-0,37299), CPI (-0,37061), farmacêutica (-0,36656), distanciamento (-0,36022), comprimido (-0,35825), rebanho (-0,35384), desserviço (-0,35063), jornaldacidadeonline (-0,34456), liberdade (-0,34310), brasil (-0,34096), vice (-0,33947), RNA (-0,33676), senador (-0,33456), conselho (-0,32672), compulsório (-0,32620), suposto (-0,32219), avanço (-0,32124), compra (-0,31953), senso (-0,31852), comprovação (-0,31752), respeitar (-0,31745), confiar (-0,31494), apertar (-0,31342), world (-0,30772), adotar (-0,30588), expressar (-0,30095), vachinas (-0,30063)

Fonte: Elaborado pela autora.

No polo positivo desta dimensão, observa-se a predominância de mensagens que alertam sobre supostos riscos associados às vacinas contra Covid-19. Estas postagens enfocam principalmente problemas cardíacos, mortes súbitas e efeitos adversos em jovens e crianças. As narrativas frequentemente incluem relatos de "infartos fulminantes" em jovens após receberem doses de reforço, caracterizando as vacinas como "experimentais" e questionando sua segurança, como mostra o tweet abaixo:

“Crianças e Adolescentes sofrendo infartos Fulminantes pouco tempo depois dos reforços . . . As pessoas não querem admitir a verdade. Poucos jovens faleceram da moléstia, muitos jovens falecendo de infartos fulminantes após as doses. Despertem! ISSO NUNCA FOI NORMAL!
<https://t.co/e68plYhO98>”

Essa retórica antivacina está em consonância com um dos tópicos centrais apontados por Skafle et al. (2022) em sua análise sobre a circulação de informações incorretas relacionadas às vacinas contra a Covid-19 nas redes sociais, que é a propagação de desinformação médica. De acordo com os autores, o temor em relação a potenciais efeitos colaterais constitui uma preocupação predominante, afetando de maneira negativa a aceitação e adesão às vacinas.

No polo negativo, os tweets mostram a resistência à imposição da vacinação obrigatória contra a Covid-19, como demonstrado no exemplo:

“CONTRA A OBRIGATORIEDADE DA VACINA CHINESA: ENQUETE DO PL 4506 / 2020 - BIA KICIS Proposta retira vacinação compulsória da lista de medidas de combate à Covid-19
<https://t.co/q4oENq6WzK>”

Esta visão é fundamentalmente enraizada na defesa das liberdades individuais contra a intromissão estatal, um embate cujas justificativas modernas encontram elos discursivos com grupos antivacina que se opuseram às políticas de imunização obrigatória desde o século XIX. Conforme analisado por Wolfe e Sharp (2002), os *British Vaccination Acts*, promulgados entre 1840 e 1867, foram os precursores de uma ampliação dos poderes governamentais sobre as liberdades civis em prol da saúde pública, desencadeando imediata resistência. Protestos à vacinação obrigatória

também ocorreram no início do século XX no Brasil. Em 1904, o Rio de Janeiro vivenciou a Revolta da Vacina, um levante popular contrário à imposição da imunização contra a varíola pelo governo (Sevcenko, 1993). Assim, no polo negativo, o discurso se fundamenta em uma perspectiva individualista, que prioriza o indivíduo acima de compromissos sociais compartilhados.

Dimensão 2 – Teorias político-conspiratórias sobre a vacina versus Evidências pseudocientíficas contra a vacina

Quadro 4 – Variáveis da Dimensão 2

Polo	Variáveis lexicais
Positivo (61 variáveis)	chinês (0,76142), povo (0,74171), comprar (0,71011), governador (0,66846), pagar (0,64473), vender (0,64198), querer (0,63828), laboratório (0,62968), vacina (0,61471), salvar (0,61241), político (0,59485), mundo (0,59394), dar (0,57604), bilhão (0,57162), via (0,56834), cloroquina (0,56221), cabaia (0,56107), garantir (0,55709), ver (0,54319), gente (0,53357), remédio (0,53299), ficar (0,53152), liberdade (0,52182), governo (0,51455), bom (0,50492), esperar (0,49807), verdade (0,49322), perder (0,47940), saber (0,47189), país (0,46664), eficácia (0,45501), colateral (0,44921), dizer (0,44382), aprovar (0,44003), interesse (0,43366), teste (0,43163), pegar (0,43088), efeito (0,42412), tratamento (0,41759), vida (0,41642), obrigatório (0,41486), precoce (0,40288), criar (0,40283), compra (0,40057), lockdown (0,38837), campanha (0,38781), aplicar (0,38636), usar (0,38266), primeiro (0,38253), prefeito (0,37771), pandemia (0,36924), vírus (0,36890), especial (0,36429), ivermectina (0,35185), causa (0,34768), imunização (0,33777), público (0,33559), covid (0,32837), segundo (0,32553), emergencial (0,32380), farmacêutico (0,31582)
Negativo (38 variáveis)	nih (-0,87872), nlm (-0,87872), ncbi (-0,87872), literat (-0,83751), bsvalud (-0,82339), articles (-0,78500), hepatite (-0,59766), pós (-0,57290), fonte (-0,55686), global (-0,51415), célula (0,50665), facebook (-0,50445), mrna (-0,49270), cdc (-0,48555), publicar (-0,48306), emitir (-0,46946), experimento (-0,44070), mielite (-0,43035), enquete (-0,40923), atleta (-0,40743), transverso (-0,40650), futebol (-0,40494), alerta (-0,39513), correiopaulista (-0,38708), pesquisa (-0,36311), fulminante (-0,35375), diagnóstico (-0,35348), farmacêutica (-0,34409), rna (-0,33583), infecção (-0,33560), imposição (-0,33243), autoimune (-0,33120), indicar (-0,32435), dama (-0,32138), proposta (-0,31880), world (-0,31081), álcool (-0,30730), postagem (-0,30132)

Fonte: Elaborado pela autora.

No polo positivo da Dimensão 2, os tweets buscam semear desconfiança e promover a ideia de que existe uma conspiração para ocultar os reais efeitos dos imunizantes. Insinuam a existência de uma colusão entre autoridades, indústria farmacêutica e mídia com o propósito de manipular a verdade para atender a interesses próprios, como se vê no tweet:

“**Bom dia Gente! Um vírus mutante criado em laboratório? Escapa de uma segurança nível 4? Vacina? Interesses farmacêuticos? China tentando dominar as criptomoedas? Eu não digo é nada Laboratório chinês** volta a ser cogitado como origem do coronavírus veja. abril. com. br / blog / mundialis ... via @VEJA”

Conforme observado por Douglas et al. (2019) em sua análise abrangente de teorias conspiratórias, há diversas razões pelas quais uma pessoa pode aderir a elas, que variam desde traços de personalidade até a satisfação de necessidades sociais complexas. Nas redes sociais, essas teorias se disseminam rapidamente devido à natureza interconectada das plataformas e à ação dos algoritmos, criando bolhas de filtro que reforçam as crenças individuais e isolam os usuários de perspectivas alternativas (Boulianne, 2020).

No polo negativo, observa-se como o movimento antivacina utiliza, em certa medida, pesquisas científicas para fundamentar suas posições. No entanto, como mostra o exemplo abaixo, essa apropriação tende a ser enviesada e seletiva, ignorando o consenso científico estabelecido e enfatizando resultados isolados que parecem apoiar sua narrativa.

“Existem estudos, **publicados** pela própria OMS, relacionando a doença de #anthonyjohnson à #vacinas de **mRNA** . Confirmam os links: pesquisa. **bvsalud.org** / global-literat ... **ncbi.nlm.nih.gov** / pmc / **articles** / P ...”

Esta prática de utilizar estudos científicos de forma enviesada para contestar a vacinação não é nova. Conforme observado por Guttinger (2019), os avanços recentes na pesquisa do microbioma humano, por exemplo, têm sido alvo de interpretações e manipulações por movimentos antivacina. Esses grupos se apropriam de hipóteses emergentes ou descobertas preliminares e as distorcem ou generalizam para além do escopo original do estudo.

Dimensão 3 – Defesa do tratamento precoce em detrimento da vacina versus Relatos pessoais sobre efeitos colaterais da vacina

Quadro 5 – Variáveis da Dimensão 3

Polo	Variáveis lexicais
Positivo (55 variáveis)	letalidade (0,93945), índice (0,88251), estatística (0,88111), segurança (0,80301), infecção (0,73291), prazo (0,70034), assegurar (0,68924), estudo (0,65232), longo (0,64329), impedir (0,61274), risco (0,61259), transmissão (0,59666), taxa (0,59352), lockdown (0,57503), precoce (0,56670), vírus (0,52658), representar (0,52654), imunidade (0,52510), tratamento (0,51383), máscara (0,50804), saudável (0,49770), sanitário (0,49498), usar (0,49150), eficácia (0,48119), benefício (0,46997), lógico (0,45338), doença (0,45005), médio (0,44215), massa (0,43768), variante (0,42629), pandemia (0,42290), social (0,42220), jovem (0,42181), bom (0,38722), custo (0,38240), científico (0,38085), mentar (0,37752), isolamento (0,37651), criança (0,37113), aprovar (0,36883), emergencial (0,36556), público (0,35527), garantir (0,35513), bilhão (0,35233), natural (0,34869), interesse (0,34267), caráter (0,33719), morte (0,32596), obrigatoriedade (0,31939), maioria (0,31886), forçar (0,31611), perigoso (0,31418), transmitir (0,31083), faixa (0,30887), paciente (0,30236)
Negativo (32 variáveis)	dor (-0,70703), cabeça (-0,70414), perna (-0,58522), ataque (-0,58404), cardíaco (-0,57438), paralisia (-0,57396), amiga (-0,57255), pé (-0,56451), resfriar (-0,54646), trombone (-0,53905), exame (-0,52451), rim (-0,52205), brasil (-0,49335), facial (-0,48978), pele (-0,46964), amigo (-0,46641), falecer (-0,43905), diagnóstico (-0,42023), apertar (-0,40594), sangue (-0,40509), infarto (-0,40205), semana (-0,39123), calça (-0,38275), miocardite (-0,37461), verdade (-0,37454), lista (-0,35126), sofrer (-0,34109), segundo (-0,33326), gripe (-0,32318), pegar (-0,31796), comprimido (-0,31445), news (-0,31229)


Fonte: Elaborado pela autora.

A ênfase na defesa do chamado "tratamento precoce" define o polo positivo da Dimensão 3. O tweet abaixo mostra como seus apoiadores argumentam que esse tratamento é composto de remédios reconhecidamente seguros e com longa utilização na prática médica tradicional:

“**Tratamento precoce** com remédios utilizados há décadas e baixo **risco** de efeitos colaterais - não pode. . . não tem **estudos científicos** suficientes; Vacina com baixo **índice** de **eficácia** e testes insuficientes - não devemos ser tão **científicos** nesse momento; Tá bom . . . vai lá sabichão!”

Este fenômeno de promoção de medicamentos ou substâncias não comprovadas como curas milagrosas durante crises de saúde é chamado de "messianismo farmacêutico" por Lasco e Yu (2021). Segundo os autores, medicamentos podem se tornar veículos de ideologias que alteram narrativas sociais de acordo com as circunstâncias políticas.

O polo negativo da Dimensão 3 é centrado na disseminação de relatos pessoais descrevendo supostos efeitos colaterais adversos após a vacinação contra a Covid-19. Conforme demonstra o exemplo, a abordagem dessas narrativas é minuciosa e íntima:

“A pequena Cleuza é uma grande guerreira, enfrentando vários **diagnósticos** de eventos adversos, depois que recebeu 2 doses da . As reações começaram com **dores** de **cabeça** e pelo corpo, câimbras e formigamentos nas **pernas**, acompanhada de inchaço e vermelhidão, visão embaçada. <https://t.co/pbCTZolgNB>”

Segundo Semino et al. (2023), evidências empíricas apontam para uma vantagem persuasiva de relatos pessoais sobre a apresentação de informações (dados e estatísticas) em contextos de comunicação de saúde, pois as narrativas diminuem a contra-argumentação e a resistência à persuasão.

A análise dos resultados proporcionou a identificação dos quatro principais discursos associados às três dimensões lexicais encontradas neste estudo, como destacado no Quadro 6:

Quadro 6 – Discursos das dimensões

DIM 1	Polo positivo	Discurso do medo
	Polo negativo	Discurso da liberdade individual
DIM 2	Polo positivo	Discurso da desconfiança de figuras de autoridade
	Polo negativo	Discurso a favor da pseudociência
DIM 3	Polo positivo	Discurso a favor da pseudociência
	Polo negativo	Discurso do medo

Fonte: Elaborado pela autora.

Estratégias Discursivas

O exame dos discursos requer uma análise das práticas sociais que deles derivam. Neste contexto, as estratégias discursivas podem ser compreendidas como métodos sistemáticos de uso da linguagem para alcançar fins sociais, políticos ou psicológicos (Wodak, 2005). Estas estratégias estão intimamente relacionadas às técnicas de persuasão através do discurso, buscando influenciar as convicções dos indivíduos (Vaara, 2010).

Uma estratégia discursiva manifesta-se como um padrão recorrente nas formas de expressão, revelando um uso específico da linguagem para conduzir alguém a uma determinada crença ou ação. Assim, o estudo dessas estratégias persuasivas constitui um elemento crucial para uma compreensão mais profunda dos discursos, sua estrutura, disseminação e impacto na formação ou transformação do conhecimento sobre diversos temas.

As técnicas persuasivas empregadas nos discursos contrários à vacinação são diversificadas, englobando abordagens tanto emocionais quanto racionais, com o intuito de moldar atitudes e comportamentos. Seu objetivo é influenciar não apenas as percepções, mas também as decisões dos indivíduos em relação à vacinação e às medidas de combate à pandemia, utilizando múltiplas abordagens e táticas interligadas.

A primeira dimensão abrange, em seu polo positivo, esforços para provocar temores e oposição à dose de reforço da vacina, com particular ênfase na população jovem. Nesse contexto, são disseminadas informações sobre supostos riscos à saúde, como dados de incidentes cardiovasculares e falecimentos, visando instalar um clima de apreensão em relação à vacinação. Simultaneamente, enfatiza-se o caráter "experimental" da vacina para gerar dúvidas quanto à sua segurança e eficácia, questionando a credibilidade da pesquisa científica por trás de seu desenvolvimento. No polo negativo desta dimensão, a estratégia principal consiste em invocar os princípios de liberdade e autonomia individual, associando as medidas obrigatórias de controle da Covid-19 a uma alegada perda dessas liberdades pessoais. Paralelamente, busca-se fomentar a desconfiança nas autoridades governamentais, retratando tais medidas como imposições arbitrárias do poder público.

O polo positivo da segunda dimensão visa propagar a desconfiança nas instituições governamentais e nas corporações, especialmente as farmacêuticas, através da disseminação de narrativas conspiratórias sobre a suposta origem intencional do vírus e de relações controversas entre políticos e a indústria. Insinua-se a existência de um plano maior ou uma conspiração para manipular e controlar a população por meio da vacinação, apelando ao temor das pessoas de serem dominadas. Nesse cenário, estabelece-se uma dicotomia entre a população comum e as elites políticas e econômicas supostamente por trás dessa alegada trama conspiratória.

Por outro lado, no polo negativo, procura-se embasar as objeções à vacina através de estudos e "evidências" pseudocientíficas, desprovidas da devida validação pela comunidade científica, numa tentativa de conferir legitimidade aos argumentos contrários. Destacam-se supostos efeitos adversos da vacina documentados em estudos enviesados ou de metodologia questionável, visando dissuadir os indivíduos da vacinação. Além disso, ressaltam-se falhas, inconsistências ou lacunas propositalmente inseridas em pesquisas sobre a eficácia das vacinas, com o intuito de semear dúvidas quanto à sua real capacidade protetora.

Na terceira dimensão, as estratégias persuasivas empregadas no polo positivo apelam ao raciocínio e à precaução, argumentando ser mais prudente e seguro utilizar tratamentos já conhecidos do que confiar em uma vacina recente cujos efeitos a longo prazo ainda não foram completamente elucidados. Critica-se a ausência de estudos e dados abrangentes sobre a segurança da vacina em um horizonte temporal mais amplo, fomentando hesitações e incertezas acerca de seus potenciais riscos futuros. Adicionalmente, enfatiza-se a gravidade e os perigos do vírus para justificar o suposto uso de tratamentos precoces alternativos à vacina, questionando a real necessidade da vacinação.

No polo negativo, as estratégias persuasivas baseiam-se na utilização de relatos e depoimentos pessoais de supostas reações adversas, com o propósito de estabelecer uma conexão emocional com o público e aumentar a credibilidade geral da mensagem antivacina. Espera-se que os indivíduos se identifiquem com as histórias pessoais relatadas, amplificando, assim, o impacto emocional da mensagem e gerando maior empatia com a causa contrária à vacinação. Há também uma tendência a enfatizar efeitos colaterais graves supostamente relacionados à vacina, com o objetivo de intensificar o medo e a hesitação em relação a ela entre a população.

Podemos, então, sintetizar as estratégias persuasivas identificadas nesta análise da seguinte forma:

- Dimensão 1: utiliza o medo (risco à saúde) e o apelo à liberdade individual.
- Dimensão 2: promove a desconfiança (teorias conspiratórias) e recorre à autoridade científica (evidências).
- Dimensão 3: fundamenta-se na precaução (tratamento precoce) e na empatia (relatos pessoais).

A análise das estratégias persuasivas empregadas nos discursos antivacina revela uma elaborada articulação de narrativas e táticas discursivas organizadas com o propósito de moldar crenças e incitar ações específicas na população. Portanto, tais estratégias convergem no objetivo central de persuadir indivíduos a adotarem posturas de hesitação, desconfiança e resistência em relação às vacinas e políticas de enfrentamento à pandemia.

CONCLUSÃO

A análise do discurso antivacina nas mídias sociais, particularmente no *Twitter*, durante a pandemia de Covid-19, revela uma intrincada teia de narrativas que se reforçam mutuamente. Os quatro discursos predominantes identificados - medo, desconfiança de figuras de autoridade, defesa da pseudociência e liberdade individual - não apenas coexistem, mas também se entrelaçam de maneira coesa, criando uma narrativa argumentativa robusta por parte do movimento antivacina.

A natureza dialógica das dimensões identificadas pela AMD Lexical torna-se evidente na forma como os discursos interagem e se complementam. Embora cada dimensão seja caracterizada por um discurso principal, é possível observar a presença de outros discursos no mesmo espaço conceitual. Um exemplo ilustrativo é o seguinte tweet, extraído do polo positivo da Dimensão 1:

“Mais **jovens falecendo** por infartos e **males Súbitos**. É importante lembrar que a **pesquisa** científica da @CMAJ, **publicada** nesta **semana**, **indicou** q o **risco** d problemas **cardíacos** é d 100 p / cada 100mil **dozes**. É duro saber que estes **morreram** p / q alguns se sentissem seguros. . . <https://t.co/xFSmndViej>”

Neste enunciado, o discurso do medo dos efeitos colaterais da vacina é predominante. No entanto, também se nota a presença do discurso a favor da pseudociência na citação de estudos potencialmente enviesados, bem como o discurso de desconfiança na insinuação de que certos grupos se beneficiam das supostas mortes. Esta confluência de discursos em um único tweet exemplifica a característica central da variação multidimensional lexical descrita por Berber Sardinha (2021): cada texto é simultaneamente moldado pela incidência dos diversos discursos representados pelas dimensões.

A AMD Lexical demonstrou sua capacidade de captar nuances discursivas ao identificar diferentes articulações de um mesmo fenômeno em dimensões distintas. Isso fica evidente na abordagem dos efeitos colaterais, presentes tanto no polo positivo da Dimensão 1 quanto no polo negativo da Dimensão 3. Embora ambos os agrupamentos de tweets compartilhem o discurso central do medo dos efeitos colaterais da vacina, a AMD Lexical consegue distinguir as sutilezas na forma como esse discurso é articulado em cada contexto.

No cenário específico do *Twitter* durante a pandemia de Covid-19, o discurso do medo assumiu formas variadas, desde alertas sobre supostas reações adversas das vacinas até relatos pessoais detalhando experiências negativas pós-vacinação. Essa versatilidade na articulação do discurso demonstra como o movimento antivacina adapta sua narrativa para atingir públicos com diferentes níveis de adesão.

Esta habilidade de discernir matizes discursivos é particularmente valiosa na análise de grandes volumes de dados textuais, superando as limitações de outras abordagens, como a Modelagem de Tópicos. Enquanto esta última poderia apenas identificar o tópico "efeito colateral" no *corpus*, a AMD Lexical revela as diferentes formas pelas quais esse discurso é mobilizado e articulado em enunciados concretos.

Em suma, este estudo expande o escopo e a profundidade da aplicação da AMD Lexical, demonstrando sua relevância para a análise de discursos contemporâneos em plataformas digitais. A capacidade de integração com outras abordagens teóricas e metodológicas, aliada à sua habilidade de captar sutilezas discursivas em grandes *corpora*, posiciona a AMD Lexical como uma ferramenta valiosa para pesquisadores que buscam compreender as complexidades do discurso em ambientes digitais dinâmicos.

É importante também destacar as limitações inerentes a este estudo. A concentração exclusiva na plataforma *Twitter* pode ter deixado de lado discursos relevantes presentes em outras redes

sociais, como debates sobre autonomia parental ou questões de influência ocidental em comunidades tradicionais. Essa restrição sugere a necessidade de pesquisas futuras que abranjam um espectro mais amplo do ecossistema digital, visando uma compreensão mais completa do fenômeno. O foco temporal no período pandêmico, embora pertinente, também limita a análise da evolução histórica dos discursos antivacina no Brasil. Investigações adicionais poderiam explorar diferentes períodos, proporcionando uma visão mais abrangente das mudanças discursivas ao longo do tempo.

Reconhecer essas limitações não apenas reforça a importância de pesquisas contínuas sobre o tema, mas também aponta caminhos para estudos complementares. Tais esforços poderão preencher as lacunas identificadas, enriquecendo nossa compreensão dos complexos fenômenos discursivos e ideológicos associados à resistência à vacinação em diferentes contextos e plataformas.

Por fim, o presente estudo espera fornecer subsídios baseados em evidências empíricas para apoiar a formulação de estratégias comunicacionais destinadas a combater o fenômeno da hesitação vacinal.

REFERÊNCIAS

- ARAÚJO, R. F. de; BERBER SARDINHA, T.; DELFINO, M. C. N. Revista Brasileira de Linguística Aplicada: multidimensões temáticas. In: FINATOO, M. J. B. *et al.* (ed.). **Linguística de corpus: perspectivas**. Porto Alegre: Editora da UFRGS e do Instituto de Letras, 2018. p. 93-125
- BAKER, P.; MCENERY, T. Introduction. In: BAKER, P.; MCENERY, T. (ed.). **Corpora and discourse studies: Integrating discourse and corpora**. Palgrave Macmillan: Basingstoke, 2015. p. 1-20.
- BERBER SARDINHA, T. **Linguística de Corpus**. Manole: Barueri, 2004.
- BERBER SARDINHA, T. **Pesquisa em Linguística de Corpus com WordSmith Tools**. Campinas: Mercado de Letras, 2009.
- BERBER SARDINHA, T. Corpus linguistics and historiography: Finding the major discourses in the first 50 years of TESOL Quarterly. **Journal of Research Design & Statistics in Linguistics & Communication Science**, v. 7, n. 1, p. 69, 2016.
- BERBER SARDINHA, T. Lexical priming: Advances and applications. In: BERBER SARDINHA, T. **Lexical priming and register variation**. Amsterdam: John Benjamins Publishing Company, 2017. p. 190-230.
- BERBER SARDINHA, T. A historical characterisation of American and Brazilian cultures based on lexical representations. **Corpora**, v. 15, n. 2, 2020.
- BERBER SARDINHA, T. Discourse of academia from a multi-dimensional perspective. In: FRIGINAL, E.; HARDY, J. (ed.). **The Routledge Handbook of Corpus Approaches to Discourse Analysis**. Abingdon: Routledge. 2021. p. 298-328.
- BERBER SARDINHA, T.; MOREIRA, M. M. F. P. **Deus, Pátria e família: os discursos bolsonaristas na rede social Twitter**. 2023. 23 f. Relatório de Trabalho de Conclusão de Curso (Graduação em Linguística Aplicada e Estudos da Linguagem), Pontifícia Universidade Católica de São Paulo, São Paulo, 2023.
- BERBER SARDINHA, T.; FITZSIMMONS-DOOLAN, S. **Lexical Multidimensional Analysis**. Cambridge Elements in Corpus Linguistics. (no prelo).
- BIBER, D. **Variation across speech and writing**. Cambridge: Cambridge University Press, 1988.
- BIBER, D. **Dimensions of Register Variation – A Cross-Linguistic Comparison**. Cambridge: Cambridge University Press, 1995.

- BIBER, D. **University language**: A corpus-based study of spoken and written registers. Amsterdam: John Benjamins, 2006.
- BOULIANNE S.; KOC-MICHALSKA, K.; BIMBER, B. Right-wing populism, social media and echo chambers in western democracies. **New Media & Society**, v. 22, n. 4, p. 683-699, 2020.
- COHN, S. K. Cholera revolts: a class struggle we may not like. **Social History**, v. 42, n. 2, p. 162-180, 2017.
- CONRAD S. Variation in corpora and its pedagogical implications. In V. Viana, S. Zyngier, & G. Barnbrook (Eds.), **Perspectives on Corpus Linguistics**. Philadelphia: John Benjamins, pp.47-62, 2011.
- DELFINO, M. C. Análise multidimensional: os números na linguística. **Cadernos de Linguística**, v. 2, n. 4, e474, 2021.
- DOUGLAS, K. M. et al. Understanding conspiracy theories. **Political Psychology**, v. 40, n. 1, p. 3-35, 2019.
- FITZSIMMONS-DOOLAN, S. Using lexical variables to identify language ideologies in a policy corpus. **Corpora**, v. 9, n. 1, p. 57-82, 2014.
- FOUCAULT, M.; FAUBION, J. (ed.). **The Essential Works of Michel Foucault 1954–1988**: Power. New York: New York Press, 2000.
- GUTTINGER, S. The anti-vaccination debate and the microbiome: How paradigm shifts in the life sciences create new challenges for the vaccination debate. **EMBO Reports**, v. 20, n. 3, e47709, 2019.
- HAIGH, M.; HAIGH, T. Fighting and framing fake news. In: BAINES, P.; O'SHAUGHNESSY, N.; SNOW, N. (ed.). **The Sage Handbook of Propaganda**. Thousand Oaks: Sage, 2020. p. 303-323.
- HAJER, M. Discourse coalitions and the institutionalization of practice. In: FISCHER, F.; FORESTER, J. (ed.). **The argumentative turn in policy analysis and planning**. Durham: Duke University Press, 1993. p. 43-76.
- HARAKI, C. A. C. Estratégias adotadas na América do Sul para a gestão da infodemia da COVID-19. **Revista panamericana de salud publica**, v. 45, e43, 2021.
- JANG, H. et al. Tracking COVID-19 Discourse on Twitter in North America: Infodemiology Study Using Topic Modeling and Aspect-Based Sentiment Analysis. **Journal of Medical Internet Research**, v. 23, n. 2, e25431, 2021.
- JONES, R. (ed.). **Viral Discourse**. Cambridge: Cambridge University Press, 2021.
- LASCO, G.; YU, V. G. Pharmaceutical messianism and the COVID-19 pandemic. **Social Science & Medicine**, v. 292, 114567, 2021.
- ORGANIZAÇÃO MUNDIAL DA SAÚDE – OMS. **Entenda a infodemia e a desinformação na luta contra a COVID-19**. [S. l.]: Organización Panamericana de la Salud, 2020. Disponível em: <https://bit.ly/3sANrxh>. Acesso em: 21 mar. 2024.
- ORGANIZAÇÃO PAN-AMERICANA DA SAÚDE – OPAS. Dez ameaças à saúde que a OMS combaterá em 2019. **OPAS**, [s. d.]. Disponível em: <https://www.paho.org/pt/noticias/17-1-2019-dez-ameacas-saude-que-oms-combatera-em-2019>. Acesso em: 01 out. 2024.
- RECUERO, R.; STUMPF, E. M. Características do Discurso Desinformativo no Twitter: Estudo do discurso antivacinas do COVID-19. In: CAIADO, R. V. R.; LEFFA, V. J. (org.). **Linguagem: tecnologia ensino**. Campinas: Pontes Editores, 2021. p. 111-137.
- ROTHKOP, D. J. When the Buzz Bites Back. **The Washington Post**, 11 maio 2003. Disponível em: <https://www.washingtonpost.com/archive/opinions/2003/05/11/when-the-buzz-bites-back/bc8cd84f-cab6-4648-bf58-0277261af6cd/>. Acesso em: 01 out. 2024.
- SCHMID, H. Probabilistic Part-of-Speech Tagging Using Decision Trees. **New Methods in Language Processing**, n. 154, p. 1-9, 2013.
- SCOTT, M. Focusing on the text and its key words. In: International conference on teaching and language corpora, 3., 2000, Frankfurt. **Proceedings** [...]. Frankfurt: Peter Lang, 2000.

SEMINO, E. et al. Narratives, Information and Manifestations of Resistance to Persuasion in Online Discussions of HPV Vaccination. **Health Communication**, v. 21, p. 1-12, 2023.

SEVCENKO, N. A **Revolta da Vacina**. São Paulo: Editora Scipione, 1993.

SKAFLE, I. et al. Misinformation About COVID-19 Vaccines on Social Media: Rapid Review. **Journal of Medical Internet Research**, v. 24, n. 8, e37367, 2022.

TAYLOR, S. The Psychology of Pandemics. **Annual Review of Clinical Psychology**, v. 18, p. 581-609, 2020.

VAN DIJK, T. A. **Discourse and power**. Basingstoke: Palgrave Macmillan, 2008.

VAARA, E. Critical discourse analysis as methodology in strategy as practice research. In: GOLSORKHI, D. et al. (ed.). **Cambridge handbook of strategy as practice**. Cambridge: Cambridge University Press, 2010. p. 217-229.

WICKE, P.; BOLOGNESI, M. M. Framing COVID-19: How we conceptualize and discuss the pandemic on Twitter. **PLoS One**, v. 15, n. 9, e0240010, 2020.

WODAK, R. **Feminist critical discourse analysis: new perspectives for interdisciplinary gender studies**. Atenas: [S. n.], 2005.

WOLFE, R. M.; SHARP, L. K. Anti-vaccinationists past and present. **BMJ**, v. 325, n. 7361, p. 430-432, 2002.