

## EFEITOS DE SENTIDO DE UMA DISCUSSÃO RACIALIZADA NO INSTAGRAM: UMA ANÁLISE BASEADA NA ANÁLISE DO DISCURSO E NA LINGUÍSTICA DE CORPUS

*EFFECTS OF MEANING ON A RACIALIZED DISCUSSION ON INSTAGRAM: AN ANALYSIS BASED ON DISCOURSE ANALYSIS AND CORPUS LINGUISTICS*

Carlos Guedes Pinto Júnior  0000-0003-2174-1310  
Faculdade de Letras - Universidade Federal de Minas Gerais  
carlosguedespj@gmail.com

Mateus Miranda  0000-0003-2575-8769  
Faculdade de Letras - Universidade Federal de Minas Gerais  
mateusesm2@gmail.com

*Recebido em 31 de dezembro de 2022*

*Aceito em 13 de fevereiro de 2023*

**Resumo:** Este trabalho, sustentado pelos princípios da Análise do Discurso e metodologicamente executado por técnicas provenientes da Linguística de Corpus, tem como objetivo investigar aspectos do discurso online. Mais especificamente, utiliza-se uma abordagem de corpus em um continuum quantitativo-qualitativo para observar os efeitos de sentido da discussão racial na rede social *Instagram*. Como ponto de partida, foi compilado um corpus de comentários oriundos de uma postagem viralizada na rede. Em um segundo momento, o corpus foi processado pelo software *Sketch Engine* e técnicas da Linguística de Corpus (lista de frequência, concordanciador, linhas de concordância) foram aplicadas. Este tipo de análise possibilitou mapear os itens lexicais e identificar colocados mais frequentes no discurso dos internautas a fim de iluminar análises qualitativas. Os resultados apontam para dizeres que perpassam a esfera racial, mais especificamente, o lugar de fala sobre raça e racismo na interação entre os participantes da postagem. A análise inicial mostra que para os participantes falar sobre questões raciais deve partir de um sujeito que vivencia esses dilemas em seu cotidiano.

**Palavras-chave:** Análise do Discurso. Linguística de Corpus. Lugar de fala. Racismo. Efeitos de Sentido.

**Abstract:** This paper, based on the principles of Discourse Analysis and methodologically executed by Corpus Linguistics techniques, aims to investigate aspects of online discourse. More specifically, a corpus approach is used in a quantitative-qualitative continuum to observe the effects of meaning of racial discussion on the social network Instagram. First, a corpus of comments from a viral post on the network was compiled. Then, the corpus was processed by Sketch Engine software, and Corpus Linguistics techniques (word frequency list, concordancer, concordance lines) were applied. This analysis made it possible to map the lexical items and identify the most frequent collocates in the internet users' discourse to shed light on qualitative analyses. The results point to sayings that permeate the racial sphere, more specifically, the place of speech about race and racism in the interaction between the participants of the post. The initial analysis shows that for participants to talk about racial issues, they must have experienced these dilemmas in their daily lives.

**Keywords:** Discourse Analysis. Corpus Linguistics. Place of Speech. Racism. Effects of Meaning.

## 1 Introdução<sup>1</sup>

As mídias sociais, que compõem o ciberespaço, permitem que os usuários demonstrem “suas emoções de forma mais livre, tanto as positivas quanto as negativas; as pessoas se mostram mais nos espaços *Web* e os conflitos se tornam mais visíveis, dando lugar a embates polêmicos” (CABRAL; LIMA, 2017, p. 88). Como apontam Cabral e Lima (2017, 2018), situações de conflito não são incomuns e permeiam o discurso na internet. A viralização de conteúdos, que acontece quando algum post se espalha rapidamente no ambiente online a partir de compartilhamentos em massa, pode ter repercussão positiva ou negativa. Quando repercutem negativamente, eles tendem a receber “manifestações linguísticas da violência verbal por parte dos usuários” (CABRAL; LIMA, 2017, p. 86), parte de um “discurso que visa a criticar o *post*” (CABRAL, LIMA, 2017, p. 86).

As redes sociais fornecem dados reais da língua em uso<sup>2</sup>. O material linguístico oriundo destas plataformas pode ser compilado e analisado em maior escala. Sugere-se, para a coleta de dados retirados um determinado contexto *online*, estabelecer critérios para que sejam representativos do discurso que se deseja estudar (PAGE *et al.*, 2014). Nesta esteira, a Linguística de Corpus (LC) é uma área que ganhou força e tem se desenvolvido rapidamente a partir do advento dos computadores e da internet. Como o nome indica, este é um campo que analisa a linguagem por meio de corpora (singular corpus) e ferramentas computacionais. Segundo Crystal (2008, p. 117), corpus refere-se a “uma grande coleção de dados linguísticos [...] que pode ser usada como ponto de partida da descrição linguística ou como meio de verificação de hipóteses sobre uma língua”. Procura-se, na compilação de um corpus, buscar a representatividade ou uma “amostra justa” (SAMPSON; McCARTHY, 2004) dos dados.

A LC é considerada uma área interdisciplinar, já que pode dialogar com outras disciplinas (BAKER, 2006). No entanto, segundo Baker (2006), não é tarefa fácil construir pontes entre áreas distintas, como, por exemplo, a LC e a Análise do Discurso (AD). A primeira dificuldade está na construção de corpora, pois a AD considera investigar menores quantidades de dados (BAKER, 2006). Assim, “quanto mais específico for o uso da linguagem, menos necessidade haverá de coletar milhões de palavras de dados” (BAKER, 2006, p. 31). A segunda, é que a LC e a AD tendem a utilizar, respectivamente, métodos quantitativos e qualitativos<sup>3</sup>.

Isto posto, a LC propõe o uso de ferramentas ou softwares que são capazes de identificar e quantificar diferentes aspectos da linguagem. As análises automáticas tornam possível o que seria inviável fazer manualmente (McENERY; HARDIE, 2012) e metodologias quantitativas podem fornecer “preciosas informações para outras análises posteriores, ainda mais porque o desenvolvimento de ferramentas técnicas cada vez mais eficientes [...] permitem, atualmente, proceder rapidamente a tais levantamentos” (CHARAUDEAU, 2011, p. 17). Nesse contexto, a figura do pesquisador é fundamental para a interpretação dos dados extraídos automaticamente (BAKER, 2006; CHARAUDEAU, 2011), pois cabe a ele “dar sentido aos padrões de linguagem que são encontrados dentro de um corpus, postulando razões para sua existência ou procurando mais evidências para apoiar hipóteses” (BAKER, 2006, p. 18).

Baker (2006, p. 13) também explica que “uma das maneiras mais importantes pelas quais os discursos circulam e se fortalecem na sociedade é por meio do uso da linguagem e a

<sup>1</sup> Agradecemos aos(as) pareceristas anônimos(as) pelas valiosas contribuições.

<sup>2</sup> Cf. Cabral e Lima (2017) e Cabral e Lima (2018) para estudos sobre o *Facebook*; Oliveira e Carneiro (2018) e Oliveira, Carneiro e Cunha (2021) para estudos sobre o *Twitter*.

<sup>3</sup> Cf. Baker (2006) para uma discussão detalhada sobre as diferenças entre as duas áreas.

tarefa dos analistas do discurso é descobrir como a linguagem é empregada, muitas vezes de maneiras bastante sutis, para revelar os discursos subjacentes”. Com base no exposto acima, um dos objetivos desta proposta foi investigar, considerando aspectos metodológicos da LC e AD, grandes quantidades de amostras do discurso digital. Mais especificamente, busca-se analisar possíveis efeitos de sentido que emergem de um corpus da rede social *Instagram*, unindo metodologias quantitativas e qualitativas para a interpretação dos dados. Assim, o presente artigo está estruturado da seguinte maneira: na seção 2, algumas noções decorrentes da AD e da LC são discutidas. O percurso metodológico utilizado é introduzido na seção 3. A seção 4 abarca a análise dos dados. Por fim, a seção 5 apresenta as considerações finais.

## 2 Referencial teórico

### 2.1 Discurso e Efeitos de sentido

Começamos, então, pela noção de discurso. Pêcheux (1988, p. 16) entende o discurso “como estrutura e como acontecimento”. O acontecimento, na perspectiva pecheutiana, é “o ponto de encontro de uma atualidade e uma memória” (1988, p. 17). Assim, o discurso acontece por meio do intradiscurso e do interdiscurso. O intradiscurso é o fio condutor do discurso; da ordem da linearidade da linguagem o intradiscurso se refere ao que o sujeito enuncia em um determinado momento, o que ele disse anteriormente e dirá depois. O interdiscurso que, por sua vez, determina o intradiscurso, está relacionado às memórias implícitas, ou seja, o “já dito”, que perpassa o discurso e é da ordem do inconsciente.

Parafraseando Pêcheux, Orlandi (2013) define discurso como o “efeito de sentido entre locutores”. Ou seja, o discurso não é simplesmente uma comunicação linear entre emissor, receptor, código, referente e mensagem. O discurso é o lugar onde o emissor e o receptor atuam de forma simultânea em um processo para construir significados (ORLANDI, 2013, p. 21). A autora propõe que esse momento de construção de significado deve ser compreendido como discurso e não como mensagem. Orlandi (2003) aponta, ainda, que o discurso não é um simples momento de transmissão de informações, pois é nele que acontece a relação entre sujeitos, na qual os sentidos são construídos e atravessados por outros sujeitos, produzindo novos sentidos. O discurso é, portanto, “o lugar onde ocorrem processos de identificação do sujeito, de argumentação, de construção da realidade etc.” (ORLANDI, 2013, p. 21).

Segundo Pêcheux (1997, p. 160), “as palavras, expressões, proposições, etc. mudam de sentido segundo as posições sustentadas por aqueles que as empregam”. Dessa afirmação é possível depreender que o sentido “não existe em si mesmo”, ele necessita do discurso e dos locutores para se manifestar. Portanto, os sentidos das palavras mudam de acordo com o uso que delas é feito a partir dos locutores, não sendo o sentido algo linear. Para Orlandi (2013, p. 64), quando o analista de discurso interpreta um dizer, seu papel

[...] não é atribuir sentidos, mas expor-se à opacidade do texto (ainda Pêcheux), ou, como tenho proposto (Orlandi, 1987), é compreender, ou seja, explicitar o modo como um objeto simbólico produz sentidos, o que resulta em saber que o sentido sempre pode ser outro. (ORLANDI, 2013, p. 64)

Portanto, este trabalho pretende analisar os possíveis efeitos e sentidos que ecoam a partir de uma interação entre usuários em um post viralizado na rede social *Instagram*. Para esta análise, considera-se que os comentários dos sujeitos participantes da discussão emergem a partir de outros discursos presentes no interdiscurso, tendo em vista que o discurso é opaco

e permeado por outros dizeres. Isso significa que, a partir da relação com o outro, esses dizeres (re)produzem efeitos de sentidos que vão além do controle dos sujeitos. Ao enunciar, os participantes não dizem somente aquilo que *desejam* dizer, mas seus discursos também apontam para dizeres anteriores, como exposto a seguir em relação à formação discursiva.

## 2.2 Formação discursiva

A noção de formação discursiva, inicialmente apresentada por Foucault (1995, p. 35) e desenvolvida por Pêcheux (1988) e Orlandi (2013), é importante para a análise dos comentários realizados pelos sujeitos na rede social. Para Pêcheux (1995, p. 143), as formações discursivas são elementos relacionados diretamente às formações ideológicas, pois refletem dizeres que apontam para uma ideologia dominante.

Em diálogo com Pêcheux, Orlandi (2013, p. 42) sugere que os sentidos são “determinado[s] por posições ideológicas e não existem sozinhos” e “são formados pela interação social” (p. 42). Os sentidos das palavras mudam de acordo com o uso que é feito delas a partir dos locutores. Citando Foucault (1984), a autora define formação discursiva “como aquilo que numa formação ideológica dada — ou seja, a partir de uma posição dada em uma conjuntura sócio-histórica dada — determina o que pode e deve ser dito” (ORLANDI, 2013, p. 43). Isso é, as formações discursivas apontam para as formações ideológicas presentes no repertório discursivo dos sujeitos que, por sua vez, apontam para diferentes efeitos e sentidos. Assim, a análise dos comentários permite entender como o discurso dos participantes é construído. Segundo Orlandi (2013, p. 43), “as palavras não têm um sentido nelas mesmas, elas derivam seus sentidos das formações discursivas em que se inscrevem”. Ou seja, é possível, por meio do discurso, perceber as formações discursivas presentes em determinados dizeres e suas relações com discursos anteriores.

Para a autora, “os sentidos sempre são determinados ideologicamente. Não há sentido que não o seja.” (2013, p. 43). Uma mesma palavra pode ter um significado diferente pelo fato de estar inserida em uma formação discursiva distinta. Orlandi (2013, p. 43) discute que “as palavras falam com outras palavras. Toda palavra é sempre parte de um discurso. E todo discurso se delinea na relação com outros: dizeres presentes e dizeres que alojam na memória.” Portanto, para Orlandi (2013) é no funcionamento do discurso que podemos perceber os diferentes sentidos das palavras na formação discursiva.

## 2.3 Da Análise do Discurso à Linguística de Corpus

A língua é percebida pela AD como o lugar onde o discurso se manifesta e os sujeitos produzem sentidos (ORLANDI, 2013). O discurso não é compreendido somente como a interação entre emissor e receptor, o discurso é “efeito de sentido entre locutores” (ORLANDI, 2013, p. 21). Portanto, o discurso é a base de trabalho da AD, pois, por meio dele, os sentidos são produzidos e as ideologias se manifestam. Orlandi (2013, p. 77) argumenta que o analista de discurso necessita compreender o funcionamento do discurso como central para a análise, para isso, ele deve examinar “os processos e mecanismos de constituição dos sentidos e sujeitos”.

Para a análise de dados, apresentamos resumidamente, os procedimentos propostos por Pêcheux (1988):

- (a) Primeiro, estudaremos as palavras dos discursos: quais são os seus termos constituintes, adjetivos, substantivos, verbos e advérbios, até que se compreenda a construção das frases.
- (b) Em seguida, cada frase deve ser dividida em proposições. Isso exige operações linguísticas para restabelecimento de ordem, reagrupamento de termos e explicitação de proposições latentes.
- (c) As unidades mínimas são possíveis em função dos mecanismos de produção dos discursos, pois estes envolvem a “repetição do idêntico”, de formas diferenciadas. Ou seja, tentaremos buscar, por trás de variações, a unidade que dá sentido ao conjunto.
- (d) Por fim, buscamos elaborar a análise, considerando a produção social do texto como geradora de seu sentido.

Em diálogo com Pêcheux, Orlandi (2009, p. 77) propõe as seguintes etapas de análise:

**Quadro 1:** Etapas de análise.

<b>1ª etapa</b>	Passagem da	superfície Linguística para o	texto (Discurso)
<b>2ª etapa</b>	Passagem do	objeto discursivo para o	formação discursiva
<b>3ª etapa</b>		processo discursivo	formação ideológica

**Fonte:** Adaptado de Orlandi (2013, p. 77).

Biber *et al.* (2007) apontam que a AD e a LC possuem métodos investigativos diferentes. De um lado, como exposto acima, o foco qualitativo da AD em textos individuais preocupa-se com a organização textual e com as estruturas para “além da frase” (BIBER *et al.* 2007, p. 2). De outro, a LC considera um conjunto de textos que compõem o corpus e busca quantificar padrões que ocorrem com frequência. No entanto, os autores concordam que os estudos de corpora podem ser vistos como um tipo de análise do discurso ao considerarmos que eles descrevem formas linguísticas em determinado contexto. Assim, as análises podem contemplar o léxico, descrito a partir de colocações mais recorrentes, além da variação gramatical no que tange a estruturas sintáticas (BIBER *et al.*, 2007). Isto posto, é possível, por meio da LC, realizar um levantamento de dados, de forma a quantificar (BAKER, 2006) procedimentos de análise propostos pela AD (PÊCHEUX, 1988; ORLANDI, 2005, 2013). A partir destas considerações, introduzimos, na seção seguinte, a metodologia do estudo, que utiliza a LC como suporte metodológico para a investigação de possíveis efeitos de sentido em um corpus de comentários do *Instagram*.

### 3 Percorso metodológico

Nesta seção, traçamos o percurso metodológico adotado no estudo, detalhando (a) a rede social escolhida, (b) o contexto dos dados, (c) a construção do corpus, (d) o software e as funções utilizadas para exploração do material.

#### (a) A rede social: *Instagram*

O *Instagram* é uma das maiores redes sociais do mundo. Estima-se que em 2019 a rede tenha atingido 1 bilhão de usuários (PALMA *et al.*, 2020). A plataforma permite que os usuários criem um perfil, público ou privado, para compartilhar posts (fotos, vídeos etc.) em

uma linha do tempo. A partir disso, as pessoas que são adicionadas ao perfil podem visualizar os posts e interagir com eles por meio de *likes* ou comentários. É possível ainda marcar amigos em determinados posts como forma de compartilhar conteúdo. Os dados deste estudo foram retirados a partir de uma publicação no *Instagram* que viralizou e ganhou grande repercussão negativa entre os usuários da plataforma, como apresentado no próximo item.

### (b) O contexto dos dados

Lília Schwarcz é escritora, antropóloga, historiadora e professora da Universidade de São Paulo. Escreveu obras importantes, como “Brasil: uma biografia”, “Lima Barreto triste visionário”, “O autoritarismo brasileiro”, entre outras. Contribuiu também como colunista para vários jornais importantes do Brasil. A fim de divulgar seu trabalho, a antropóloga possui uma conta pública no *Instagram* e realiza posts com frequência. Em 2020, Schwarcz teceu uma crítica em sua coluna na Folha de São Paulo<sup>4</sup> sobre um álbum visual lançado pela cantora pop estadunidense Beyoncé para a nova versão do filme O Rei Leão. O artigo foi transformado em *post* para seu perfil do *Instagram*<sup>5</sup> (Figura 1).

Figura 1: Crítica de Lília no *Instagram*.



Fonte: Página pessoal da professora no *Instagram*.

<sup>4</sup> Disponível em <<https://www1.folha.uol.com.br/ilustrada/2020/08/filme-de-beyonce-erra-ao-glamorizar-negritude-com-estampa-de-oncinha.shtml#erramos>>. Acesso em: 15 out. 2022.

<sup>5</sup> Texto integral da postagem no *Instagram*: “Como tudo que Beyoncé faz, seu novo álbum visual, Black is king, chega causando polêmica e trazendo muito barulho. Ele se baseia no projeto The Lion King: the gift, álbum de 2019, lançado conjuntamente com o filme da Disney. Nele, a cantora e compositora retoma a história clássica de Hamlet, personagem icônico de Shakespeare, mas a ambienta em algum lugar perdido do continente africano. O Hamlet de Shakespeare se passa na Dinamarca e conta a história do príncipe que tem como missão vingar a morte de seu pai, o rei, executado pelo próprio irmão, Cláudio. Traição, incesto e loucura são temas fortes da trama e da própria humanidade, de uma forma geral. Já a versão da Disney é ambientada na África e tem como personagem principal uma matilha de leões – os “reis dos animais”. No enredo, o filho Simba, herdeiro do trono, instado pelo irmão invejoso, desobedece ao pai e, não propositadamente, acaba sendo o pivô da morte dele e de um golpe de Estado. O tema retoma a culpa edipiana do filho que não conseguindo vingar ou salvar o pai, perde seu prumo na vida e esquece sua história. Já Beyoncé, evoca mais uma vez a tragédia de Hamlet, mas inverte a mão da narrativa. Simba vira um menino negro que procura por suas raízes para conseguir sobreviver no mundo racista norte-americano de 2020. Só não era necessário estereotipar dessa maneira uma África isolada e perdida no mundo com muitos leopardos e oncinhas. Melhor Beyoncé sair da sua sala de estar e tomar mais ar de realidade”. (Matéria completa na Ilustrada, Folha de. S. Paulo).

No entanto, os seguidores de Lília Schwarcz começaram a mostrar certo descontentamento, por meio de uma avalanche de mensagens agressivas. O *post* viralizou nas redes e gerou um grande debate em relação ao tema. Wilson Gomes, professor titular de Comunicação da Universidade Federal da Bahia, em sua coluna na Folha de São Paulo, saiu em defesa de Schwarcz com o artigo “O cancelamento da antropóloga branca e a pauta identitária<sup>6</sup>”. Gomes destaca a violência verbal do linchamento virtual da professora, como “cala a boca”, “racista”, “a antropóloga branca não sabe o seu lugar”, e aponta que “se, pelo menos, ainda fosse justa a indignação, por ter a Lilia publicado um texto racista ou ofensivo, ainda assim ficaria envergonhado pelos termos do cancelamento, mas compreenderia. O pior de tudo é que não, não há nada de errado com o artigo usado como desculpa para linchar” (GOMES, 2020).

### (c) A construção do corpus de estudo

A coleta dos comentários do post foi feita em agosto de 2020, ao mesmo tempo em que os usuários do *Instagram* reagiam à publicação da antropóloga. Foram extraídos, de forma aleatória, 2.990 comentários, totalizando 111.457 *tokens* (número total de itens) e 94.049 *types* (formas lexicais únicas). Os textos passaram por um processo de limpeza para a remoção de itens indesejados como as fotos e os nomes dos usuários. Posteriormente, cada comentário foi anonimizado, recebendo os códigos <CoM>:

<CoM> Não acredito que esse texto até agora não foi apagado!

Por fim, os dados foram salvos em arquivo *txt*. para processamento pelo software de análise.

### (d) O software: *Sketch Engine*

O *Sketch Engine*<sup>7</sup> (KILGARRIFF; RYCHLÝ, 2003) é um programa para a construção e análise de corpora. Há vários corpora disponíveis para uso imediato, mas a plataforma ainda permite a construção de corpora a partir de textos de websites ou por meio do upload dos próprios arquivos. Utilizamos o *Sketch Engine* para a exploração dos dados a partir de algumas funções de exploração de corpora, como (1) lista de frequência e (2) concordanciador e linhas de concordância, detalhadas a seguir.

#### **Função (1):** Gerador de lista de frequência

A lista de frequência reúne “as palavras que são apresentadas a partir das mais frequentes até as menos frequentes” (TAGNIN, 2011, p. 359). As listas de frequência, são, geralmente, o ponto de partida para a análise de grandes quantidades de dados. Viana (2011, p. 47) destaca que “esse procedimento revela inúmeros aspectos sobre o material incluído no corpus e igualmente gera hipóteses investigativas, que precisarão ser complementadas com análises mais detalhadas”. Dessa forma, o nosso objetivo aqui é verificar se a lista de itens mais frequentes é capaz de indicar características, passíveis de investigação, provenientes do discurso dos usuários.

---

<sup>6</sup> Disponível em <<https://www1.folha.uol.com.br/ilustrissima/2020/08/o-cancelamento-da-antropologa-branca-e-a-pauta-identitaria.shtml>>. Acesso em: 15 out. 2022.

<sup>7</sup> Disponível em <<https://www.sketchengine.eu/>>. Acesso em: 10 out. 2022.

**Função (2):** Concordanciador e linhas de concordância

O concordanciador “extrai todas as ocorrências de uma palavra de busca num corpus juntamente com o seu cotexto, apresentando-as na forma de uma concordância” (TAGNIN, 2011, p. 358). O concordanciador permite que o pesquisador faça observações mais qualitativas (VIANA, 2011) pelo cotexto<sup>8</sup> de uma determinada busca. Assim, as linhas de concordância centralizam as ocorrências do *nódulo*, ou palavra de busca, em seu cotexto.

**4 Análise dos dados e discussão**

**4.1 O que o percurso metodológico pode revelar sobre o discurso dos internautas?**

A tabela a seguir detalha os 10 itens mais frequentes do corpus de estudo extraídos pelo gerador de lista de frequência.

**Tabela 1:** Lista de frequência.

N	Item	Freq	N	Item	Freq
1	que	3474	6	não	2032
2	de	3390	7	é	1753
3	a	2814	8	uma	1420
4	e	2745	9	da	1052
5	o	2116	10	você	1059

**Fonte:** Elaborado pelos autores (2022).

Os 10 itens mais frequentes da lista parecem não revelar muito sobre o discurso dos usuários. No entanto, a ocorrência do pronome pessoal *você* nos chama a atenção e é um item candidato à análise. Adolphs e Lin (2011) confirmam que as listas de palavras podem apontar diferenças entre a fala e a escrita. Ao comparar os 10 itens mais frequentes no *Limerick Corpus of Irish English*<sup>9</sup> (LCIE) e no *British National Corpus*<sup>10</sup> (BNC written), mostram que no LCIE *I* (eu), *you* (você) e *yeah* (sim) são marcas frequentes da oralidade e, por isso, não aparecem entre os 10 mais frequentes do BNC escrito (ADOLPHS, LIN, 2011). Isto posto, decidimos investigar *você* e seu cotexto por meio do concordanciador, em uma análise mais qualitativa (Fig. 2).

<sup>8</sup> “Na Linguística de Corpus emprega-se o termo ‘cotexto’ para se referir ao ambiente linguístico, ou seja, aos itens que são utilizados à esquerda e à direita de uma determinada palavra de busca” (VIANA, 2011, p. 71).

<sup>9</sup> O *Limerick Corpus of Irish English* é um corpus de 1 milhão de palavras de inglês irlandês. Os 10 itens mais frequentes no LCIE são: *the, I, and, you, to, it, a, that, of, yeah* (ADOLPHS, LIN, 2011, p. 600).

<sup>10</sup> O *British National Corpus* é um corpus de 100 milhões de palavras (90% escrito e 10% oral). Os itens mais frequentes no BNC são: *the, of, and, a, in, to, is, was, it, for* (ADOLPHS, LIN, 2011, p. 600).

**Figura 2:** Linhas de concordância para “*ocê*”.

não fazer.	À propósito, quem é <b>você</b> mesmo?	Arrogância branca é foda!
Olhei a foto e pensei.. branca.. logo	<b>você</b> não entende nada	Branco fazendo bran
A crítica precisa ser feita.	<b>Você</b> está completamente equivocada com esse seu	
ndo bola fora e sendo cobrada	<b>você</b> não tem propriedade e não sou eu que estou fa	
alou bosta	<b>Você</b> precisa rever o filme.	Está erradíssima.
sível	<b>você</b> não sabe	Não fale que Beyoncé não sa
eu sei que pode ser difícil mas	<b>você</b> não tem lugar de fala	descansa, militan
Talvez pq não seja seu lugar de fala e	<b>você</b> jamais deveria estar questionando o trabalho d	
Meu Deus que vergonha!	<b>Você</b> não tem propriedade então deixa pra quem pod	

Fonte: Sketch Engine (2022).

O concordanciador mostra o nóculo, ou palavra de busca, em vermelho, e o ambiente linguístico em que a palavra está inserida. As linhas de concordância acima (Fig. 2) são ocorrências aleatórias. Viana (2011, p. 74) enfatiza que “a leitura nesse caso deve ser iniciada justamente pela palavra em posição central, verificando as palavras empregadas à esquerda e à direita (ou vice-versa)”. A partir disso, destacamos, na Fig. 2, duas características do cotexto se destacam. Observamos, à direita, a alta frequência do item lexical *não* seguindo o nóculo *você*, formando o padrão *você + não*, que acompanha verbos como *entende*, *sabe* e *tem*. Dessa forma, *você não tem* é o padrão mais saliente, ocorrendo três vezes. O’Keeffe *et al.* (2007) destacam a função interativa do pronome *você*, utilizado como uma forma direta de abordar o outro (CRYSTAL, 2006, 2011). Além disso, pode representar uma marca da transposição do discurso oral para o escrito, característica bastante comum nas mensagens compartilhadas no ambiente online, que tendem a se aproximar do discurso oral (CRYSTAL, 2006, 2011). Assim, podemos verificar o alto grau de interação entre os usuários da rede e o post de Lília, por meio de mensagens destinadas a ela.

A partir da descoberta descrita acima, decidimos delimitar o enfoque no padrão linguístico *você não* e realizar um levantamento das formas lexicais, ou colocados, mais frequentes à sua direita [fórmula *você não \**] (Tab. 2).

**Tabela 2:** Colocados à direita do padrão *você não*.

Padrão	Colocado	Freq
você não +	tem	123
	entendeu	21
	sabe	11
	conhece	8
	pode	5

Fonte: Elaborado pelos autores (2022).

Observamos, na Tab. 1, que o colocado<sup>11</sup> mais frequente à direita de *você não é tem*, com 123 ocorrências, seguido por ‘entendeu’ (21), ‘sabe’ (11), ‘conhece’ (8) e ‘pode’ (5). Todos os colocados apresentados neste trabalho apresentaram valores >2 para o escore T e >3 para o MI, indicando que os itens lexicais se atraem, ou seja, há forte ligação entre o termo de busca e os colocados. Orlandi (2009, p. 73) considera o texto como unidade empírica representativa de um começo, meio e fim. Ou seja, por meio dele é possível analisar traços ideológicos e da constituição do sujeito. No corpus analisado, a partir da proposição mais frequente ‘você não tem’, podemos elaborar as seguintes proposições (você ⇒ eu; você não tem ⇒ eu tenho) tendo como produção os comentários no post do *Instagram*.

Devido ao alto número de ocorrências de ‘tem’, decidimos, utilizando o mesmo método, verificar quais itens lexicais acompanham o padrão ‘você não tem’ [fórmula *você não tem* \*] (Tab. 3). Identificamos que o colocado mais frequente à direita de ‘você não tem’ é ‘lugar’, com 30 ocorrências, seguido por ‘local’ (21), ‘direito’ (20) e ‘propriedade’ (13).

**Tabela 3:** Colocados à direita do padrão *você não tem*.

Padrão	Colocado	Freq
você não tem +	lugar	30
	local	21
	direito	20
	propriedade	13

**Fonte:** Elaborado pelos autores (2022).

‘Lugar’ e ‘local’, substantivos masculinos sinônimos<sup>12</sup>, são os colocados mais frequentes à direita do padrão ‘você não tem’. As figuras 3 e 4 ilustram, respectivamente, as 5 primeiras ocorrências para os padrões ‘você não tem lugar’ e ‘você não tem local’. Mais uma vez, ao observarmos as linhas de concordância, notamos que ‘de fala’ (em destaque nas figuras 3 e 4), é um item recorrente à direita dos padrões. Após uma leitura completa de todas as linhas de concordância (51 linhas, considerando ‘lugar’ e ‘local’ juntos) para verificar o ambiente à direita dos padrões, verificamos que ‘de fala’ é uma ocorrência unânime. Dessa

<sup>11</sup> Para Viana (2011, p. 40) ‘colocação’ refere-se “ao uso estatisticamente significativo de determinada palavra em companhia de outra”. Dessa forma, é qualquer elemento que co-ocorra de forma significativa com uma palavra de busca” (TAGNIN, 2011, p. 358). Sugere-se, para a classificação dos colocados, testes estatísticos como o escore T e informação mútua (MI), entre outros. Uma discussão completa das questões estatísticas está além do escopo deste trabalho. De acordo com TAGNIN (2011, p. 360), o escore T é uma “medida de significância que mede a certeza de que há alguma associação entre as palavras, ou seja, a frequência com que as palavras co-ocorrem elimina a possibilidade de essa co-ocorrência resultar de mero acaso. Para ser computada como uma ‘colocação’, costuma-se considerar um escore mínimo de dois”. A Informação mútua é uma “medida de significância que mede a força de associação entre as palavras e privilegia palavras menos frequentes. Como ocorrências pouco frequentes podem resultar de usos idiossincráticos, é praxe considerar-se ‘colocação’ apenas co-ocorrências com um escore mínimo de três” (TAGNIN, 2011, p. 360).

<sup>12</sup> Para confirmar a similaridade entre os itens lexicais, buscamos a definição lexicográfica, além de consultar a relação de sinonímia no Corpus do Português (<https://www.corpusdoportugues.org/>) por meio da expressão regular [=\*] aplicada a cada uma delas: [=lugar] e [=local]. Assim, não fazemos distinção entre as formas *você não tem lugar* e *você não tem local*.

forma, ‘você não tem [lugar/local] de fala’<sup>13</sup> é um agrupamento lexical relevante no discurso dos usuários.

**Figura 3:** Linhas de concordância para *você não tem lugar*.

Details	Left context	KWIC	Right context
1	doc#0 quivocada, não acreditei que li isso vindo de você.</s><s>Nossa que ridícula..	<b>você não tem lugar</b>	de fala.. só disse asneiras fia</s><s>Ansiosa pelo dia que vamos cobrar dos g
2	doc#0 dade escravocrata é... Quem puxa aos seus não denegera...</s><s>Fia é que	<b>você não tem lugar</b>	de fala, entendeu?</s><s>Nunca mais fale sobre Beyoncé.</s><s>Bjs</s><s>
3	doc#0 esse texto @lillaschwarcz ou reedita?</s><s>Discordo de cada vígula.</s><s>	<b>Você não tem lugar</b>	de fala pra falar sobre o que uma diva negra deve fazer, querida!</s><s>isso q
4	doc#0 >.</s><s>Você não tem ideia do quanto o tralho de Bey é importante.</s><s>	<b>Você não tem lugar</b>	de fala para falar o que ela deve fazer.</s><s>Calada é uma poeta!</s><s>Do
5	doc#0 ue que vc escreveu, parabéns</s><s>Lamentável, Lilian!</s><s>PQPQPQPQ	<b>você não tem lugar</b>	de fala aqui então apaga e pare de passar vergonha mulher!</s><s>Na vdd e

Fonte: Sketch Engine (2022).

**Figura 4:** Linhas de concordância para *você não tem local*.

Details	Left context	KWIC	Right context
1	doc#0 :s>Xiiuuu, vc não sabe de nada do povo negro.</s><s>Não perde seu tempo...	<b>você não tem local</b>	de fala nenhum para tratar do assunto.</s><s>Mas é que você não tem local d
2	doc#0 .. você não tem local de fala nenhum para tratar do assunto.</s><s>Mas é que	<b>você não tem local</b>	de fala, né?</s><s>O guincho que veio rebocar minha preguiça da discussão r
3	doc#0 atéria sobre isso rs</s><s>Comprou seu lugar de fala no débito? ah esqueci..	<b>você não tem local</b>	de fala pra dar sua opinião irrelevante</s><s>A mesma branca que faz uma p
4	doc#0 o isso errado</s><s>A braquitude não reconhece seus limites mesmo.</s><s>	<b>Você não tem local</b>	de fala aqui!</s><s>Quem você pensa que é?</s><s>Infelizmente não é o cas
5	doc#0 <s>Vaza daqui fia.. mas você não tem noção né.. e também só lembrando que	<b>você não tem local</b>	de fala quando o assunto é gente preta tá...</s><s>fia caia a boca vai</s><s>

Fonte: Sketch Engine (2022).

A partir disso, podemos investigar quais efeitos de sentido estão sendo produzidos e quais propriedades estão sendo constituídas (atribuídas) a professora. Os itens lexicais mais frequentes à direita do padrão ‘você não tem’, como ‘local’, ‘direito’, ‘propriedade’ funcionam como unidades parafrásticas da palavra ‘lugar’, a mais frequente, pois destacam a questão racial levantada pelos seguidores da antropóloga. Os efeitos de sentido produzidos pelos comentários dos usuários revelam uma formação ideológica que não é qualquer pessoa que pode falar sobre raça e racismo e, de certa forma, a professora é colocada em lugar de racista.

O objetivo deste estudo não é delimitar ou discutir o conceito de ‘lugar de fala’, contudo, é importante destacar que o termo é utilizado de forma ampla e com grande propriedade em discussões na internet. Ribeiro (2019) argumenta que ocorre um esvaziamento de conceitos importantes em discussões *online* em função de uma urgência em “criar polêmica vazia” (p. 27). No interdiscurso dos internautas é possível perceber que para se falar de raça e racismo o enunciador necessita fazer parte de algum grupo que sofra desse preconceito. No nosso caso, a professora é uma mulher cis, branca e da classe alta brasileira. Apesar de ser professora historiadora e antropóloga, uma autoridade em sua área, os efeitos dos dizeres apontam para um não-pertencimento por parte da autora em relação à questão racial no entendimento dos internautas.

Como observado, a LC auxilia nos procedimentos propostos por Pêcheux (1988) e Orlandi (2005, 2013). A princípio, foi possível identificar as palavras mais frequentes para entender o percurso de análise, extrair os colocados até o dizer mais frequente. Em seguida, por meio da análise da repetição no corpus, foi possível perceber o uso parafrástico de alguns vocábulos para a palavra ‘lugar’. Por fim, por meio do estudo dos possíveis efeitos de sentido, observa-se que a questão racial está relacionada à produção social do texto, pois somente é possível perceber esta relação ao analisar o discurso em seu contexto.

<sup>13</sup> As abreviações são bastante comuns na escrita online. Por isso, verificamos a ocorrência de “vc” como abreviação do pronome “você” e encontramos 480 ocorrências. Procedemos com as mesmas etapas metodológicas e encontramos apenas duas ocorrências para “vc não tem lugar de fala”.

## 5 Considerações finais

Este trabalho buscou observar os efeitos de sentido da discussão racial em um post da rede social *Instagram*. Para tanto, a Análise do Discurso e a Linguística de Corpus, duas áreas consideradas metodologicamente distintas, apesar do mútuo interesse na investigação da língua em uso, são integradas. Como exposto, a Linguística de Corpus pode funcionar como instrumento organizacional dos dados, a partir de técnicas executadas por meio de ferramentas computacionais para analistas do discurso. Dessa forma, o pesquisador pode compilar um corpus maior que seja representativo do seu objeto de estudo. Além disso, os princípios e etapas de análise propostos pela Análise do Discurso podem ser contemplados e realizados de forma mais rígida pelo analista.

Os resultados apontam para dizeres que perpassam a esfera racial, mais especificamente, o ‘lugar de fala’ sobre raça e racismo na interação entre os participantes do post. A análise inicial mostra que, para os participantes, falar sobre questões raciais deve partir de um sujeito que vivencia esses dilemas em seu cotidiano. Verificamos, em diálogo com Ribeiro (2019), que a questão da discussão sobre o ‘lugar de fala’ é recorrente no ambiente online, mas é um conceito de difícil definição. Portanto, o estudo nos leva a refletir sobre a seguinte questão: o que é lugar de fala e quem tem lugar/local de fala para discorrer sobre o racismo? Espera-se, assim, que este estudo instigue novas discussões sobre raça e racismo na esfera acadêmica.

## Referências

ADOLPHS, S.; LIN, P. Corpus Linguistics. In: SIMPSON, J. (Org.). *The Routledge Handbook of Applied Linguistics*. New York: Routledge, 2011, p. 597-610.

BAKER, P. *Using corpora in discourse analysis*. London/New York: Continuum, 2006.

BIBER, D.; CONNOR, U.; UPTON, T. *Discourse on the Move: Using Corpus Analysis to Describe Discourse Structure*. Amsterdam: John Benjamins Publishing, 2007.

CABRAL, A. L. T.; LIMA, N. V. Argumentação e polêmica nas redes sociais: o papel de violência verbal. *Signo*, Santa Cruz do Sul, v. 42, n. 73, p. 86-97, jan/abril 2017.

CABRAL, A. L. T.; LIMA, N. V. Interações conflituosas e violência verbal nas redes sociais: polêmica em comentários no Facebook. *Revista (Con)textos Linguísticos*, Vitória, v. 12, n. 22, p. 39-58, set. 2018.

CHARAUDEAU, P. Dize-me qual é teu corpus, eu te direi qual é a tua problemática. *Diadorim*, Rio de Janeiro, v. 10, p. 01-23, 2011.

CRYSTAL, D. *Internet Linguistics. A Student Guide*. New York: Routledge, 2011.

CRYSTAL, D. *Language and the Internet*. Cambridge: Cambridge University Press, 2006.

CRYSTAL, D. *A Dictionary of Linguistics and Phonetics*. Oxford: Blackwell Publishing, 2008

KILGARRIFF, A.; RYCHLÝ, P. *Sketch Engine*. East Sussex: Lexical Computing Limited, 2003. Disponível em: <http://www.sketchengine.eu>. Acesso em: 10 dez. 2022.

McENERY, T.; HARDIE, A. *Corpus Linguistics: Method, theory and practice*. Cambridge: Cambridge University Press, 2012.

O'KEEFFE, A.; McCARTHY, M.; & CARTER, R. *From corpus to classroom: Language use and language teaching*. Cambridge, UK: Cambridge University Press, 2007.

OLIVEIRA, A. L. A. M.; CARNEIRO, M. M. Sobre o potencial semântico pragmático das hashtags In: CUNHA, G. X.; OLIVEIRA, A. L. A. M. (Org.). *Múltiplas perspectivas do trabalho de face nos estudos da linguagem*. Belo Horizonte: FALE/POSLIN/UFMG, 2018, p. 207-226.

OLIVEIRA, A. L. A. M.; CARNEIRO, M. M.; CUNHA, G. X. O que os pobres fazem para sobreviver. (Im)Polidez e classismo no Twitter brasileiro. *Scripta*, Belo Horizonte, v. 25, n. 53, p. 562-585, jun. 2021.

ORLANDI, E. P. *Discurso e texto: formulação e circulação de sentidos*. 2. ed. Campinas: Pontes, 2005.

ORLANDI, E. P. *Análise de discurso: princípios e procedimentos*. 12. ed. Campinas: Pontes, 2013.

ORLANDI, E. P. *Discurso em análise: sujeito, sentido e ideologia*. 3. ed. Campinas: Pontes, 2017.

PAGE, R.; BARTON, D.; UNGER, J.; ZAPPAVIGNA, M. *Researching Language and Social Media: A Student Guide*. London; New York: Routledge, 2014.

PALMA, A.; ROJAS, F.; TRUCCO, D. O uso de redes sociais por crianças e adolescentes no Brasil, no Chile, na Costa Rica e no Uruguai. *TIC Kids Online Brasil 2019*, São Paulo, p. 109-116, out. 2020.

PÊCHEUX, M. *Semântica e discurso – uma crítica à afirmação do óbvio*. Trad. Eni Pulcinelli Orlandi et al. Campinas: Editora da Unicamp, 1988.

RIBEIRO, D. *O que é lugar de fala?*. Belo Horizonte: Letramento, 2019.

SAMPSON, G.; McCARTHY, D. *Corpus Linguistics: Reading in a Widening Discipline*. New York: Continuum, 2004.

TAGNIN, S. Corpora on-line. In: VIANA, V.; TAGNIN, S. E. O. (Orgs.). *Corpora no ensino de línguas estrangeiras*. São Paulo: Hub Editorial, 2011, p. 137-171.

VIANA, V. Linguística de Corpus: Conceitos, técnicas & análises. In: VIANA, V.; TAGNIN, S. E. O. (Orgs.). *Corpora no ensino de línguas estrangeiras*. São Paulo: Hub Editorial, 2011, p. 25-95.