

Texto e gênero na pesquisa em processamento linguístico: pertinência e implicações metodológicas

*Text and genre in linguistic processing research:
relevance and methodological implications*

Rosângela Barros da SILVA 

Universidade Federal de Sergipe
São Cristóvão, Brasil
rosangelabsilva@academico.ufs.br

Resumo: Avaliar a pertinência dos conceitos de texto e gênero em estudos que envolvem o processamento linguístico de classes gramaticais no português brasileiro requer acioná-los como parte da escolha metodológica, estabelecendo critérios de suporte para a análise do *corpus*, a seleção dos dados e o delineamento experimental da pesquisa. O uso linguístico de advérbios terminados em “-mente” ocorre em estruturas sentenciais específicas, cuja interpretação pode ser modulada por características textuais e de gênero. Este trabalho envolve estudos de processamento linguístico e variação morfossintática. A perspectiva teórica adotada para essa discussão baseia-se na Linguística Sistêmico-Funcional (Halliday; Hasan, 1989), especialmente nas relações entre linguagem, contexto e aprendizagem. O conceito de gênero presente nas características do *corpus* de fala, constituído por entrevistas semiestruturadas com perfis sociais e linguísticos distintos, tem como estatuto metodológico orientar a coleta e a organização dos dados experimentais. Este artigo objetiva demonstrar como a análise de *corpus* de fala, fundamentada nos conceitos de texto e gênero, fornece subsídios empíricos para o delineamento futuro de experimentos com rastreamento ocular. A análise de *corpus* visa identificar padrões de variação morfossintática que orientarão a construção de estímulos experimentais. Os resultados apresentam critérios metodológicos baseados na articulação entre linguística textual e funcionalismo para estudos de processamento linguístico em situação de uso real da língua.

Palavras-chave: texto; gênero textual; processamento linguístico; cognição; advérbios em “-mente”.

Abstract: Assessing the relevance of text and genre concepts in studies involving the linguistic processing of grammatical classes in Brazilian Portuguese requires mobilizing them as part of the methodological framework, establishing guiding criteria for corpus analysis, data selection, and the experimental design of the research. The linguistic use of adverbs ending in -ly (-mente) occurs in specific sentential structures whose interpretation may be modulated by textual and genre-related features. This study integrates research on linguistic processing and morphosyntactic variation. The theoretical perspective adopted draws on Systemic Functional Linguistics (Halliday; Hasan, 1989), particularly the relations between language, context, and learning. The concept of genre, as reflected in the characteristics of the spoken corpus, comprising semi-structured interviews with diverse social and linguistic profiles, serves a methodological role in guiding data collection and experimental organization. This article aims to demonstrate how spoken corpus analysis, grounded in text and genre concepts, provides empirical support for the future design of eye-tracking experiments. The corpus analysis seeks to identify morphosyntactic variation patterns that will guide the construction of experimental stimuli. The results present methodological criteria based on the articulation between textual linguistics and functionalism for linguistic processing studies in real-life language use situations.

Keywords: text; textual genre; linguistic processing; cognition; adverbs ending in -ly (-mente).

1 INTRODUÇÃO

Diferentes situações de interação comunicativa mobilizam distintas formas de organização linguística, sejam elas verbais ou visuais. Na dinâmica da língua em uso, elementos como texto, gramática e gênero não são categorias isoladas, mas interdependentes na produção e interpretação da linguagem. Uma conversa espontânea da vida cotidiana aciona a gramática da língua oral, que se distingue da gramática da língua escrita mobilizada durante a leitura de um artigo acadêmico, por exemplo. Nessas experiências comunicativas, os recursos empregados na organização de cada forma textual refletem as variedades e finalidades da língua em uso e influem no processamento linguístico, tanto da fala quanto da leitura. É um processo que mostra as nuances das regularidades gramaticais que emergem das práticas sociais dos usuários da língua.

Essa é uma das perspectivas sobre a inter-relação entre a linguística de texto, o funcionalismo e os estudos de processamento linguístico, que se articulam a partir de um interesse comum pelos estudos da linguagem em uso. A linguística de texto fornece instrumentos analíticos para descrever

como os textos se estruturam em contextos específicos, considerando princípios como coesão, coerência e progressão temática. O funcionalismo, especialmente em abordagens como a Gramática Sistêmico-Funcional (Halliday, 2014, 2010; Halliday; Hasan, 1989), compreende a língua como um recurso semiótico orientado por funções sociais, explicando os fenômenos linguísticos com base em suas motivações comunicativas e contextuais. Assim, “as palavras que são trocadas nesses contextos obtêm seu significado das atividades nas quais estão inseridas, que, por sua vez, são atividades sociais com agentes e objetivos sociais específicos” (Halliday; Hasan, 1989, p. 5).

Halliday (1989, 2014), ao tratar da arquitetura funcional da linguagem humana e sua relação com texto e gramática, afirma que, ao falar ou escrever, as pessoas produzem textos. São esses textos que mobilizam a atenção e a interpretação dos ouvintes e leitores. O termo “texto” diz respeito a qualquer manifestação da linguagem, em qualquer modalidade, desde que faça sentido para alguém que conheça a língua em uso no contexto da comunicação. O texto, nessa perspectiva, é definido como a linguagem em funcionamento no contexto da situação comunicativa, um processo de significação situado em contexto (Halliday; Hasan, 1989). Essa concepção converge com aspectos dos estudos sobre processamento linguístico que investigam os mecanismos cognitivos que são ativados por falantes, ouvintes e leitores em situação de fala ou leitura em tempo real.

Os processos envolvidos nas dinâmicas comunicativas acionam vários níveis da língua, que ocorrem de maneira muito rápida e exigem esforço cognitivo para processar os fenômenos da linguagem em um nível mais amplo. A compreensão, então, não se restringe ao sentido lexical das palavras, mas emerge da articulação entre múltiplas camadas de informação linguística, fonológica, morfológica, semântica, sintática e pragmática, que se integram de modo não linear nem isolado. Trata-se de um processamento linguístico dinâmico e interativo que reflete a adaptabilidade e a eficiência do sistema cognitivo humano (Freitag, 2025, p. 23).

Assim, ao integrar a linguística de texto e o funcionalismo, em diálogo com os estudos de processamento linguístico que se debruçam a estudar a organização das informações da linguagem e os contextos de uso, torna-se

possível adentrar um caminho para compreender não apenas o que os usuários da língua produzem, mas como produzem e por quê, considerando que as escolhas gramaticais são um sistema de escolhas dinâmicas condicionado por gêneros, textos e situações de uso na interação.

O componente cognitivo no processamento da linguagem e a análise da posição de advérbios terminados em “-mente”, em função de sua ativação no fluxo do discurso em tempo real, constituem um fenômeno que envolve expectativas cognitivas na situação de interação e de contexto. De acordo com Halliday (1989, p. 5), embora a aprendizagem envolva também expectativas não verbais, especialmente no contexto escolar, ela se realiza por meio da linguagem e as expectativas linguísticas são determinantes para seu sucesso.

É nessa inter-relação que se evidencia a pertinência de vincular o fenômeno ao contexto situacional e aos processos de aprendizagem mediados pela linguagem. Nessa perspectiva de articulação entre diferentes áreas teóricas, o estudo aqui apresentado faz parte de atividades ligadas à minha pesquisa de doutorado e discute a interface entre Linguística de Texto e Funcionalismo na análise de advérbios terminados em “-mente”. A investigação parte da pressuposição de que os advérbios terminados em “-mente”, embora classificados pela Nomenclatura Gramatical Brasileira (NGB) como termos acessórios da oração (Brasil, 1959), influem no processamento linguístico da estrutura sentencial, particularmente no desempenho da leitura. Essa influência pode estar associada à mobilidade sintática dessa classe gramatical e à posição de destaque que pode ocupar na oração, impactando a variação morfossintática e o processamento linguístico em tempo real.

O comportamento sintático-semântico envolve uma interação entre os advérbios e seu escopo, entendido como o alcance interpretativo que exercem sobre outros elementos da oração. No caso específico dos advérbios terminados em “-mente”, o escopo pode variar conforme a posição sintática e influenciar semanticamente diferentes níveis da estrutura sentencial, seja um sintagma, o predicado ou a oração inteira, podendo gerar múltiplas interpretações. Durante o processamento da linguagem oral, o falante precisa solucionar ambiguidades por meio de escolhas gramaticais, como a ordem dos constituintes e a atribuição de

funções semânticas. Em situações de leitura, esse mesmo processo exige que o leitor ajuste o escopo dos advérbios para interpretar adequadamente a estrutura e o significado do enunciado. Para demonstrar a mobilidade sintática desses advérbios, Castilho (2014, p. 550) apresenta o seguinte exemplo:

- (1) a. **Realmente** você vê que aqui você passa melhor.
- b. Você **realmente** vê que aqui você passa melhor.
- c. Você vê **realmente** que aqui você passa melhor.
- d. Você vê que **realmente** aqui você passa melhor.
- e. Você vê que aqui **realmente** você passa melhor.
- f. Você vê que aqui você **realmente** passa melhor.
- g. Você vê que aqui você passa **realmente** melhor.
- h. Você vê que aqui você passa melhor, **realmente**.

O advérbio “realmente” dispõe de certa mobilidade ao longo da oração. É desse exemplo que o objeto central da pesquisa é o fenômeno microestrutural da variação posicional e interpretativa de advérbios terminados em “-mente”. A metodologia desse estudo envolve a análise de *corpus* de fala, constituído por entrevistas sociolinguísticas do banco de dados Falares Sergipanos (Freitag, 2013), e de textos de material didático do ensino básico de uma escola pública do estado de Sergipe. Desse modo, os movimentos da interseção entre Linguística de Texto e Funcionalismo na pesquisa de processamento linguístico surgem ao empregar os conceitos de texto e gênero como ferramentas metodológicas; não como objeto central da investigação.

Abordagens funcionalistas contemporâneas integram categorias como texto e gênero na concepção do desenho metodológico e analítico em investigações sobre o funcionamento real da linguagem para análise de processamento linguístico e variação morfossintática. Sob pressupostos teóricos alinhados à Gramática Sistemico-Funcional (GSF), essas abordagens se destacam pelas inter-relações de uso da linguagem, contextos de comunicação e processos cognitivos envolvidos na interação.

Nessa perspectiva, o texto é compreendido como uma unidade semiótica funcional e o gênero como uma configuração de elementos textuais que reflete padrões socialmente compartilhados (HALLIDAY; HASAN, 1989). Em diálogo com Castilho (2014, p. 544), a investigação está ancorada na hipótese de que uma maneira de compreender o

comportamento sintático e a interpretação dos advérbios em “-mente” é examinando as funções sentenciais dessa classe, sua colocação nos enunciados e os argumentos que selecionam (se os selecionarem). Busca-se, assim, identificar as regularidades observáveis nos seus usos reais da língua e em situações comunicativas específicas.

1.1 Objetivos e delimitação do estudo

Este artigo tem como objetivo avaliar a pertinência dos conceitos de texto e gênero como categorias metodológicas em estudos de processamento linguístico, demonstrando sua aplicação na análise de advérbios terminados em “-mente” em *corpus* de fala do português brasileiro. Especificamente, busca-se: (i) identificar padrões de variação posicional e de escopo desses advérbios em entrevistas sociolinguísticas; (ii) estabelecer critérios de frequência e familiaridade lexical a partir da análise de *corpus*; e (iii) discutir as implicações desses achados para o delineamento de experimentos futuros com rastreamento ocular.

Como recorte da pesquisa de doutorado em andamento, este artigo concentra-se na primeira etapa metodológica: a análise de *corpus* de fala que estabelece a base empírica para a construção de tarefas experimentais em etapas subsequentes e complementares. Por meio desses experimentos, será investigado o processamento linguístico durante a leitura de estudantes do ensino fundamental, médio e superior. Assim, o presente estudo não apresenta resultados de rastreamento ocular, mas discute como os achados do *corpus* fundamentarão o desenho metodológico das futuras investigações. Para orientar essa investigação, formulam-se as seguintes perguntas:

- a) a pesquisa de advérbios terminados em “-mente” aciona o conceito de texto ou de gênero alinhado à perspectiva textual?
- b) qual é o estatuto operativo no contexto dessa pesquisa: objeto de estudo ou categoria metodológica?

2 LINGÜÍSTICA DE TEXTO, GÊNERO E VARIAÇÃO MORFOSSINTÁTICA

Esta seção apresenta a fundamentação teórica sobre o uso dos conceitos de texto e gênero como categorias metodológicas na análise de fenômenos de variação morfofossintática. Parte-se da concepção sistêmico-funcional de linguagem (Halliday; Hasan, 1989; Halliday, 2014) para articular as relações entre contexto situacional, escolhas gramaticais e processamento linguístico, demonstrando a pertinência dessas categorias para estudos sobre advérbios em “-mente”. O texto é caracterizado como uma unidade de significado em uso vinculado ao contexto situacional. Halliday e Matthiessen (2014, p. 51) ampliam essa concepção ao definirem que o texto constitui a forma de dados para análise linguística, ou seja, o *corpus* linguístico: uma configuração dinâmica de elementos estruturais e semânticos organizados no discurso no momento da sua produção.

Essas abordagens permitem compreender como fenômenos linguísticos variam em realização e interpretação, conforme o gênero textual e o contexto social em que circulam. Como observa Neves (2006), a descrição gramatical deve ser sensível aos usos reais da língua e à diversidade dos gêneros e dos contextos comunicativos, o que converge com a proposta funcional de compreender a linguagem como prática social. Nesse processo, para responder às perguntas de pesquisa formuladas, é necessário compreender como o conceito de gênero opera na investigação sobre advérbios terminados em “-mente”.

À luz das abordagens teóricas discutidas, o tratamento dessas classes gramaticais permite que a resposta à primeira pergunta seja afirmativa, ainda que parcialmente. O conceito de gênero se apresenta como categoria metodológica, constituída pelo *corpus* de fala como base empírica inicial da pesquisa. Com a função de orientar a coleta e a organização dos dados, que são instâncias do gênero, atua também como categoria de análise para investigar a variação morfofossintática em contexto.

O conceito de texto permite pensar sobre a funcionalidade das unidades linguísticas em contextos situacionais, enquanto o gênero *corpus* de fala ajuda a entender como unidades linguísticas variam sistematicamente em contextos comunicativos diversos. A esse respeito, Bawarshi e Reiff (2010, p. 29) argumentam que, embora o trabalho sobre

gênero na linguística de *corpus* ainda não tenha tido grande impacto na teoria de gênero, tem muito a contribuir para a pesquisa e o ensino de gênero, para a natureza da tipologia e para a mudança linguística. Os debates acerca da categorização de gênero em linguística de *corpus* podem ajudar a esclarecer como os estudos de gêneros baseados em *corpus* fornecem informações sobre como e por que os gêneros mudam. Essa perspectiva é relevante à investigação sobre os advérbios, pois a análise do *corpus* de entrevistas sociolinguísticas é a fonte empírica para a busca de padrões e variações linguísticas em textos falados e escritos.

2.1 Texto e gênero: da fala para o *corpus*

Na perspectiva funcionalista, a variação linguística não constitui um fenômeno aleatório, mas reflete adaptações sistemáticas da língua às necessidades comunicativas dos falantes. Halliday (2014) enfatiza que as escolhas linguísticas são motivadas funcionalmente, respondendo às demandas do contexto situacional e às intenções comunicativas dos usuários da língua. Essa abordagem é fundamental para compreender como os advérbios terminados em “-mente” operam na estrutura textual e como sua variação posicional reflete papéis funcionais nos enunciados.

No *corpus* constituído por 64 entrevistas sociolinguísticas com estudantes universitários, observa-se que esses elementos linguísticos não apenas modificam verbos, adjetivos ou outros advérbios, mas apresentam comportamento sintático regular e funções mais amplas. A frequência de uso, a mobilidade sintática e a variação de escopo dão pistas do processamento linguístico da fala em interação em tempo real. Essa motivação linguística é abordada por Halliday (2014) e conecta-se à formulação de Castilho (2014), que ressalta o papel dos advérbios como operadores de sentido com escopo variável.

A mobilidade do comportamento sintático mostra aspectos do funcionamento gramatical sensíveis ao contexto e à estrutura do enunciado. Do ponto de vista semântico, “os advérbios são palavras predicativas, vale dizer, operadores que transferem para seu escopo propriedades semânticas de que elas não dispunham” (Castilho, 2014, p. 551). Em uma breve inspeção dos dados linguísticos, encontram-se os seguintes fragmentos extraídos do *corpus*:

- (2) o setor agrário é muito importante, mas não traz desenvolvimento, **infelizmente** não abre oportunidade pra outras / outras carreiras, né?
- (3) [...] é, só se eles fossem **realmente**.
- (4) É, você **realmente** gosta de morar lá no Bugio, então? / Eu gosto, mas é porque eu não saio.

Nessas ocorrências, os advérbios “infelizmente” e “realmente” ocupam diferentes posições sintáticas e de escopo. Em (2), “infelizmente” ocorre em posição inicial, pré-verbal, com escopo sentencial, marcando a avaliação negativa do falante em relação ao contexto apresentado sobre as limitações do setor agrário. Em (3), “realmente” ocorre em posição final, pós-verbal, com escopo de valor implícito, funcionando como pista de intensificador na construção condicional (“se eles fossem realmente”). Em (4), o advérbio “realmente” ocorre em posição medial, entre o sujeito e o verbo, mas a relação entre classe e escopo pode desencadear significados distintos, de afirmação ou de interrogação. Esse é um breve exemplo da mobilidade funcional desse elemento linguístico, um fenômeno amplamente descrito e analisado na literatura dos estudos do português brasileiro.

Pesquisas com amostras de dados de fala podem fornecer pistas sobre o processamento e a variação linguística. O *corpus*, como ferramenta metodológica para identificar padrões de frequência, posicionamento e co-ocorrência de advérbios, possibilita a análise quali-quantitativa do fenômeno. Nesse exercício de análise linguística, ao descrever a posição de advérbios terminados em “-mente” em relação ao seu escopo, observando o comportamento dos dados no *corpus*, tem-se a resposta da segunda pergunta (Qual é o estatuto operativo no contexto dessa pesquisa?). O estatuto operativo dos conceitos de texto e de gênero na pesquisa sobre advérbios em “-mente” é, principalmente, o de categoria metodológica.

Quando utilizado como recurso metodológico, e articulado às perspectivas teóricas de gêneros textuais, linguística de texto e pesquisas alinhadas ao polo funcional (Halliday; Hasan, 1989; Neves, 2006, Castanheira, 2022), o *corpus* de fala é uma ferramenta que pode fornecer informações

empíricas sobre a relação entre a estrutura linguística, o contexto social e o processamento das categorias morfossintáticas no processo mental.

2.2 Processamento linguístico: do *corpus* às implicações experimentais

Partindo da análise dos dados de fala e avançando para as investigações que poderão capturar a atenção visual dos leitores, duas das variáveis analisadas na pesquisa são: i) a posição sintática dos advérbios terminados em “-mente”; ii) o escopo desses advérbios na estrutura da oração e os efeitos dessas variáveis no processamento linguístico. Inicialmente, essas variáveis são investigadas no *corpus* de fala constituído por entrevistas sociolinguísticas e, posteriormente, em gêneros textuais de material pedagógico. O objetivo é articular a dimensão microestrutural do fenômeno aos condicionamentos discursivos e sociais do uso linguístico. Como defendido por Castilho (2014), a análise da linguagem em uso, especialmente no nível microestrutural, exige atenção às escolhas gramaticais e sua articulação com o gênero e com os objetivos comunicativos dos interlocutores.

A análise do *corpus* realizada até o momento indica padrões de posição e variabilidade de escopo dos advérbios terminados em “-mente”. Os resultados preliminares dão suporte à importância de uma abordagem sociossemiótica para a análise da estrutura gramatical e das propriedades linguísticas intrínsecas ao processamento da linguagem. Nessa direção, vale destacar a afirmação de Halliday (1989) sobre como:

Nós tendemos a pensar na aprendizagem exclusivamente como um processo cognitivo e negligenciar seus aspectos linguísticos. O que estamos tentando fazer aqui é interpretar a aprendizagem como um processo linguístico, dando alguns passos preliminares em direção a uma teoria linguística da aprendizagem que vem complementar os modelos cognitivos estabelecidos. Isso deve permitir que professores e outras pessoas envolvidas explorem o valor e o papel crítico da linguagem na educação e apreciem o quanto as crianças dependem da linguagem para poderem aprender. (Halliday, 1989, p. 49)¹

¹ We tend to think of learning exclusively as a cognitive process, and to neglect its linguistic aspects. What we are attempting to do here is to interpret learning as a linguistic process, taking some tentative steps towards a linguistic theory of learning that would complement the established cognitive models. This should enable teachers and others concerned to explore the value and critical role of a language in education and to appreciate how deeply children depend on language in order to be able to learn. (Halliday, 1989, p. 49).

A compreensão da linguagem como processo cognitivo fundamenta a relevância de investigar a influência de elementos da estrutura linguística no processamento textual e na aprendizagem. Interpretar o aprendizado como um processo linguístico é etapa necessária para o entendimento dos mecanismos de processamento e de aplicações pedagógicas que considerem essa dimensão da aprendizagem. Sob essa perspectiva, a metodologia utilizada para investigar o processamento linguístico e a variação morfossintática no *corpus* de fala e em textos de material didático visa subsidiar a construção posterior de experimentos com técnica de rastreamento ocular (*eye tracking*).

Valer-se de interfaces entre abordagens teóricas e metodológicas possibilita articular as características conceituais de textos e gêneros e a pertinência de inter-relacioná-los na pesquisa linguística. Nesse contexto, as etapas de coleta e análise de dados consideram a natureza multifacetada dos fenômenos linguísticos em situações de uso real da língua, e a investigação do processamento linguístico integra os estudos com *corpus* de língua falada e escrita às análises experimentais. A escolha de tipo textual e de gênero em pesquisas sobre processamento constitui uma decisão metodológica que determina a validade dos resultados e sua aplicabilidade a contextos de comunicação autênticos.

3 PROCEDIMENTOS METODOLÓGICOS: ANÁLISE E VALOR EMPÍRICO

Estudos sobre variação morfossintática consistem, entre outros aspectos, em analisar e descrever o comportamento sintático de elementos linguísticos. A presente pesquisa utiliza uma abordagem empírico-descritiva baseada em análise de *corpus* para investigar o comportamento de advérbios terminados em “-mente” em contextos reais de uso. O objetivo é relacionar a frequência, a familiaridade lexical e os padrões de posicionamento desses advérbios, a partir de um *corpus* de entrevistas sociolinguísticas transcritas. A análise visa fornecer subsídios para o delineamento de experimentos de leitura com rastreamento ocular, com foco no papel desses elementos no processamento linguístico.

3.1 Corpus e caracterização da amostra

Os dados analisados foram extraídos da mostra *Deslocamentos 2019* do banco de dados Falares Sergipanos (Freitag, 2013). A amostra é composta por 64 entrevistas sociolinguísticas transcritas realizadas com estudantes da Universidade Federal de Sergipe, estratificados segundo três variáveis: i) tipo de deslocamento, em termos de mobilidade dos estudantes ao *campus*, que abrange alunos da região próxima à universidade, estudantes provenientes de cidades do interior de Sergipe ou de outros estados; ii) tempo de ingresso no curso superior dividido em início, do primeiro ao quarto período, e final, a partir do quinto período; e iii) sexo/gênero, categorizado como masculino ou feminino (Sousa, 2023, p. 58).

As entrevistas seguiram um modelo de roteiro semiestruturado (Freitag, 2014), totalizando 447.668 *tokens* (ocorrências) e 14.833 *types* (tipos lexicais). O gênero entrevista sociolinguística caracteriza-se por interações dialógicas em contexto controlado, constituindo base empírica adequada para estudos de frequência de uso, familiaridade lexical e comportamento sintático de fenômenos linguísticos em situação real de comunicação oral. A Tabela 1 abaixo sintetiza a caracterização do global do *corpus*.

Tabela 1 — Caracterização do *corpus* de análise

Característica do <i>corpus</i>	Valor
Total de entrevistas transcritas (files .txt)	64
Total de palavras (tokens)	447.668
Total de tipos (types)	14.833
Total de lemas (lemmas)	13.585
Total de palavras terminadas em “-mente” (extração automática)	2.163
Total de advérbios terminados em “-mente” (após filtragem manual)	2.125
Total de formas diferentes de advérbios terminados em “-mente”	146
Tipos de advérbios mais frequentes (acima de 100 ocorrências) selecionados para o estudo	5
Tipos de advérbios menos frequentes (apenas uma ocorrência) selecionados para o estudo	5

Fonte: elaborada pela autora.

3.2 Ferramentas e procedimentos de análise

A análise foi conduzida com auxílio do software #LancsBox 6.0 (Brezina; Weill-Tessier; Mcenery, 2021), utilizando a ferramenta KWIC (*keyword in context*) para extração das ocorrências de advérbios em “-

mente” e a ferramenta *Words* para cálculo de frequências e dispersão lexical. O tratamento para anotação dos dados foi realizado manualmente em planilha do Microsoft Excel, ajustadas para anotação morfossintática alinhada às etiquetas do sistema *Universal Dependencies (UD)* (De Marneffe; Manning; Nivre; Zeman, 2021)². Com relação a esse processo, Sousa (2023) explica que:

Por meio de linguagem de programação ou de ferramentas de busca em *corpora*, é possível fazer um levantamento por fenômenos linguísticos variáveis no *corpus* de trabalho utilizando apenas as etiquetas (Sousa, 2023, p. 18).

A ocorrência de advérbios como “realmente” (331 ocorrências), “principalmente” (235 ocorrências) e “geralmente” (195 ocorrências) no *corpus* de fala sugere que o gênero discursivo se torna espaço de mobilização da seleção lexical e dos padrões de frequência. Em relação à entrevista sociolinguística, trata-se de um contexto interacional específico que fornece dados empíricos para estudos de frequência de uso, familiaridade lexical e comportamento sintático do fenômeno linguístico da pesquisa. As implicações para a construção de estímulos experimentais residem no fato de que a familiaridade com determinadas formas linguísticas é condicionada pelo tipo de exposição textual do falante. A análise desses advérbios em uso na fala do português brasileiro permite verificar se as características textuais e de gênero influenciam os padrões de uso e, por conseguinte, o processamento linguístico da variação morfossintática.

3.3 Exemplos de ocorrência em contexto (KWIC)

Uma amostra aleatória de ocorrências de advérbios terminados em “-mente” foi extraída pelo software LancsBox. Na ferramenta KWIC (*keyword in context*), delimitou-se um parâmetro de contexto de análise de 15 palavras à esquerda e 15 palavras à direita do fenômeno estudado. Essa etapa

² *Universal Dependencies (UD)* é uma estrutura para anotação morfossintática em diferentes línguas naturais, desenvolvida por mais de 600 colaboradores com mais de 200 *corpus* sintaticamente anotados em mais de 150 idiomas de acesso aberto em <https://universaldependencies.org>

proporciona analisar o uso dos advérbios terminados em “-mente”, a respectiva posição sintática e o contexto de uso.

Quadro 1 — Conjunto aleatório de linhas de concordância para *.*mente\$* no *Corpus*

Nome do arquivo	Esquerda	Nó (termo de busca)	Direita
51ent.XXX-XXXXXX2018__desl. IV_início_dou.ms.21_sp.txt	muito na ficção científica algo	totalmente	distópico mas que trata de
04ent.XXX- XXXXXX 2018__desl. I_final_cla.fs.21_sp.txt	gente acaba conhecendo várias pessoas	basicamente	todo mundo de Laranjeiras vem
49ent.XXX- XXXXXX 2018__desl. IV_início_ade.ms.30_sp.txt	mas da pós graduação e	certamente	também se há algum funcionário
33ent.XXX- XXXXXX 2018__desl. III_final_apa.ms.23_sp.txt	que eu o que eu	realmente	queria e o que eu
19ent.XXX- XXXXXX 2018__desl. II_início_cla.fs.19_sp.txt	se eles iam deixar né?	provavelmente	porque eu sou de maior
13ent.XXX- XXXXXX 2018__desl. I_início_jea.ms.17_sp.txt	Aracaju cidade onde cê mora	atualmente?	Nossa Senhora do Socorro bairro?

Fonte: extraído do relatório automático do LancsBox 6.0.

Após a extração dos advérbios terminados em “-mente” do *corpus* global, são analisadas suas posições nas orações, sua relação com o verbo (pré-verbal, pós-verbal, medial) e as classes gramaticais no contexto imediato à esquerda e à direita desses elementos. Essa etapa é conduzida manualmente em planilhas exportadas do LancsBox e filtradas no Excel. Delimitar a busca dos advérbios por categorias de frequência e familiaridade lexical permite encontrar padrões de comportamento que podem influenciar o processamento durante a leitura.

A escolha metodológica por dados de língua falada, organizados em entrevistas transcritas, assegura maior validade ecológica à análise, além de evidenciar a pertinência dos conceitos de texto e gênero como categorias estruturantes para a observação linguística. As análises qualitativas são realizadas por meio de anotação manual de cada ocorrência e, com base nos resultados da análise da variação morfossintática, é elaborado o delineamento dos estímulos para posteriores experimentos de leitura com rastreamento ocular e análise da compreensão textual.

3.4 Implicações para estudos com rastreamento ocular e desempenho leitor

A integração entre a análise de *corpus* e experimentos com a técnica de rastreamento ocular representa um avanço metodológico, especialmente em estudos de variação morfossintática de classes gramaticais no português brasileiro. As variações posicionais identificadas por meio da análise contextual no *corpus* fornecem parâmetros empíricos para a formulação de hipóteses sobre custos de processamento em leitura. A metodologia apresentada permite capturar os padrões de colocação dos termos pertencentes ao discurso oral e a variabilidade sintática dos fenômenos linguísticos. Essa abordagem mostra-se ecologicamente válida e pode ser contrastada com estudos que utilizam frases frequentemente assinaladas da língua escrita.

A técnica de rastreamento ocular (*eye-tracking*, em inglês) é um método experimental utilizado para registrar, em tempo real, os movimentos dos olhos durante a leitura. Essa técnica fornece dados precisos sobre as fixações, sacadas e regressões oculares dos participantes. Os dados permitem analisar como os leitores processam as informações linguísticas do texto e identificar os pontos de maior esforço cognitivo (Rayner, 1998). No contexto da pesquisa com advérbios terminados em “-mente”, utiliza-se o rastreamento ocular para investigar como a posição sintática desses advérbios afeta o processamento linguístico durante a leitura. Ao comparar advérbios de alta e baixa frequência lexical em diferentes posições oracionais (inicial, medial e final), essa metodologia permite verificar quais variações morfossintáticas geram maior ou menor custo de processamento.

A técnica de rastreamento ocular contribui para validar, em ambiente experimental, as evidências empíricas previamente levantadas na análise de *corpus* de fala. Com maior robustez para a generalização dos resultados experimentais, o estudo demonstra que texto e gênero não se limitam a contextos de ocorrência dos fenômenos linguísticos. Eles se constituem em variáveis metodológicas que orientam a interseção entre análise de *corpus* e estudos experimentais. As especificidades textuais e discursivas apresentam contribuições importantes para o avanço das pesquisas sobre o processamento da linguagem, oferecendo resultados

robustos e ecologicamente válidos para a compreensão dos mecanismos cognitivos subjacentes ao uso da língua.

Presumir que advérbios menos familiares no *corpus* oral podem ser altamente familiares em outros gêneros textuais permite reconhecer a necessidade de *corpora* mais diversificados para estudos sobre processamento lexical no desempenho da leitura. Pesquisas com resultados robustos geram estratégias para lidar com construções que exigem maior esforço cognitivo. Logo, conhecer os custos de processamento associados às variações morfossintáticas permite antecipar o conhecimento das dificuldades específicas de estudantes durante as atividades de leitura.

Os padrões identificados na análise de *corpus* fornecerão parâmetros para o delineamento de experimentos com rastreamento ocular, incluindo a seleção de advérbios de alta e de baixa frequência, a construção de frases e textos com diferentes posições sintáticas, a formulação de hipóteses sobre custos de processamento e a elaboração de estímulos ecologicamente válidos baseados em usos reais.

4 CONSIDERAÇÕES FINAIS

As discussões apresentadas neste artigo mostram que, embora o foco da pesquisa esteja no nível microestrutural da linguagem, o processamento dos advérbios terminados em “-mente” é sensível aos contextos textuais e aos gêneros em que esse fenômeno ocorre. Os conceitos de texto e gênero são pertinentes para compreender o processamento desses advérbios, permitindo uma abordagem contextualizada dos fatores que interferem na produção e na interpretação linguística. As abordagens conceituais sobre texto e gênero, partindo da perspectiva sistêmico-funcional (Halliday; Hasan, 1989) e desenvolvidas na literatura funcionalista (Meurer; Bonini; Motta-Roth, 2005), possibilitam uma compreensão mais ampla e contextualizada dos fatores que interferem na produção e na interpretação linguísticas desses elementos.

Ao selecionar *corpora* constituídos por textos e gêneros pedagógicos e por entrevistas sociolinguísticas para analisar os advérbios terminados em “-mente”, o estudo não se limita a uma abordagem formal ou experimental, mas considera o processamento linguístico baseado no uso. Essa

perspectiva metodológica permite compreender como o contexto comunicativo condiciona os padrões de fenômenos linguísticos. Dessa forma, o estudo alinha-se às discussões contemporâneas dos programas de pós-graduação em Estudos Linguísticos e contribui tanto para os estudos sobre processamento linguístico quanto para o ensino de gramática no âmbito escolar.

Os padrões de mobilidade sintática identificados neste estudo convergem com as observações de Ilari (2009) sobre o comportamento dos advérbios terminados em “-mente” no português falado, especialmente no que diz respeito às posições que eles tendem a ocupar na oração. Quando aplicados a verbos, a posição preferencial é a pós-verbal. As discussões de Castilho (2014) sobre o escopo variável desses advérbios complementam essa perspectiva, fornecendo subsídios teóricos que se articulam aos achados preliminares da pesquisa, sobretudo quanto ao papel dos advérbios como operadores semânticos. Dessa forma, o diálogo entre esses autores ampara a proposta metodológica que articula análise de *corpus* e experimentos de rastreamento ocular, permitindo investigar empiricamente os custos de processamento decorrentes de variações morfossintáticas.

Além de contribuições teóricas, o estudo aponta para desdobramentos pedagógicos no ensino de gramática e leitura. A análise da mobilidade e do escopo dos advérbios em “-mente” pode ser incorporada em atividades de sala de aula voltadas à identificação de ambiguidade, à leitura crítica de enunciados e à compreensão da coesão e coerência textuais. Categorias como escopo, modalidade e posição sintática devem ser analisadas em sua interação com as condições reais de produção textual e com os gêneros discursivos, principalmente quando o objetivo é entender o processamento em situações reais de uso em diferentes linguagens.

Cabe ressaltar, porém, que o estudo apresenta limitações a serem consideradas. Primeiro, a análise realizada até o momento se restringiu a um único *corpus*, o que limita a generalização de padrões de frequência, familiaridade lexical e mobilidade sintática. Segundo, por corresponder a uma atividade específica de uma pesquisa em andamento, o artigo não apresenta resultados experimentais de rastreamento ocular, concentrando-se apenas na análise de *corpus* que servirá de base empírica aos

experimentos. Terceiro, embora a amostra de 64 entrevistas seja robusta para a análise morfossintática, prevê-se a incorporação de outras duas amostras do mesmo Banco de Dados Falares Sergipanos (Freitag, 2013) para verificar se os padrões já identificados se mantêm. Por fim, a anotação manual dos dados em *corpora* de maior dimensão poderá fortalecer o estudo e subsidiar as tarefas experimentais de rastreamento ocular.

Apesar das limitações apontadas, a interface entre estudos funcionalistas, linguística de texto e investigação experimental representa um avanço metodológico significativo, contribuindo para os estudos atuais de processamento e variação morfossintática. Nesse contexto, “uma das faces contemporâneas da posição funcionalista tem desenvolvido indagações sobre a interface discurso/gramática, buscando-se identificar o empacotamento gramatical dos segmentos textuais e dos processos de sua constituição” (Castilho, 2014, p. 138). Essa perspectiva converge com este estudo, que combina análise de corpus, fundamentação sistêmico-funcional e experimentos psicolinguísticos para compreender as implicações cognitivas de variações morfossintáticas em contextos reais de uso, oferecendo subsídios para aplicações pedagógicas futuras.

Agradecimentos:

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

REFERÊNCIAS

BAWARSHI, A. S.; REIFF, M. J. **Genre: An introduction to history, theory, research, and pedagogy**. Parlor Press LLC, 2010. Disponível em: <https://wac.colostate.edu/books/referenceguides/bawarshi-reiff/>. Acesso em: 24 jun. 2025.

BRASIL. **Nomenclatura Gramatical Brasileira**. Rio de Janeiro: Ministério da Educação e Cultura, 1959. Disponível em: <https://docs.ufpr.br/~borges/publicacoes/notaveis/NGB.pdf>. Acesso em: 26 jun. 2025.

BREZINA, V.; WEILL-TESSIER, P.; MCENERY, A. **#LancsBox v. 6.0** [software]. 2021. Disponível em: <http://corpora.lancs.ac.uk/lancsbox>. Acesso em: 23 abr. 2025.

CASTILHO, A. **Nova gramática do português brasileiro**. São Paulo: Contexto, 2014.

DE MARNEFFE, M.-C.; MANNING, C. D.; NIVRE, J.; ZEMAN, D. Universal Dependencies. **Computational Linguistics**, v. 47, n. 2, p. 255–308, 2021. Disponível em: <https://direct.mit.edu/coli/article/47/2/255/98516/Universal-Dependencies>. Acesso em: 30 ago. 2025.

FREITAG, R. M. K. **Variação linguística: diversidade e cotidiano**. São Paulo: Contexto, 2025.

FREITAG, R. M. K. **Metodologia de coleta em manipulação de dados em sociolinguística**. São Paulo: Blucher, 2014.

FREITAG, R. M. K. Banco de dados falares sergipanos. **Working Papers em Linguística**, v. 14, n. 2, p. 156-164, 2013.

HALLIDAY, M. A. K.; MATTHIESSEN, M. I. M. **Halliday's introduction to functional grammar**. 4. ed. New York: Routledge, 2014.

HALLIDAY, M. A. K. **Text, discourse and information: a systemic-functional overview**. Paper presented at Tongji University. November, 2010.

HALLIDAY, M.; HASAN, R. **Language, context, and text: aspects of language in a social-semiotic perspective**. Oxford: University Press, 1989.

ILARI, R. A categoria advérbio na gramática do português falado. **ALFA: Revista de Linguística**, São Paulo, v. 51, n. 1, 2009. Disponível em <<https://periodicos.fclar.unesp.br/alfa/article/view/1430>>. Acesso em: 27 jul. 2024.

MEURER, J. L.; BONINI, A.; MOTTA-ROTH, D. (org.). **Gêneros - teorias, métodos, debates**. São Paulo: Parábola, 2005.

NEVES, M. H. M. **Texto e gramática**. São Paulo: Contexto, 2006.

RAYNER, K. Eye movements in reading and information processing: 20 years of research. **Psychological Bulletin**, v. 124, n. 3, p. 372–422, 1998. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/9849112/>. Acesso em: 30 jul. 2025.

SOUSA, M. D. A. F. **Protocolo para anotação linguística e gerenciamento de amostras sociolinguísticas**: o caso da amostra Deslocamentos 2019. 2023. 147 f. Tese (Doutorado em Letras) – Universidade Federal de Sergipe, São Cristóvão, 2023. Disponível em: <https://ri.ufs.br/handle/riufs/18363>. Acesso em: 30 jul. 2025.

SILVA, Rosangela Barros. Texto e gênero na pesquisa em processamento linguístico: pertinência e implicações metodológicas. **Entrepalavras**, Fortaleza, v. 15, e96485, 2025. DOI: 10.36517/ep15.96485