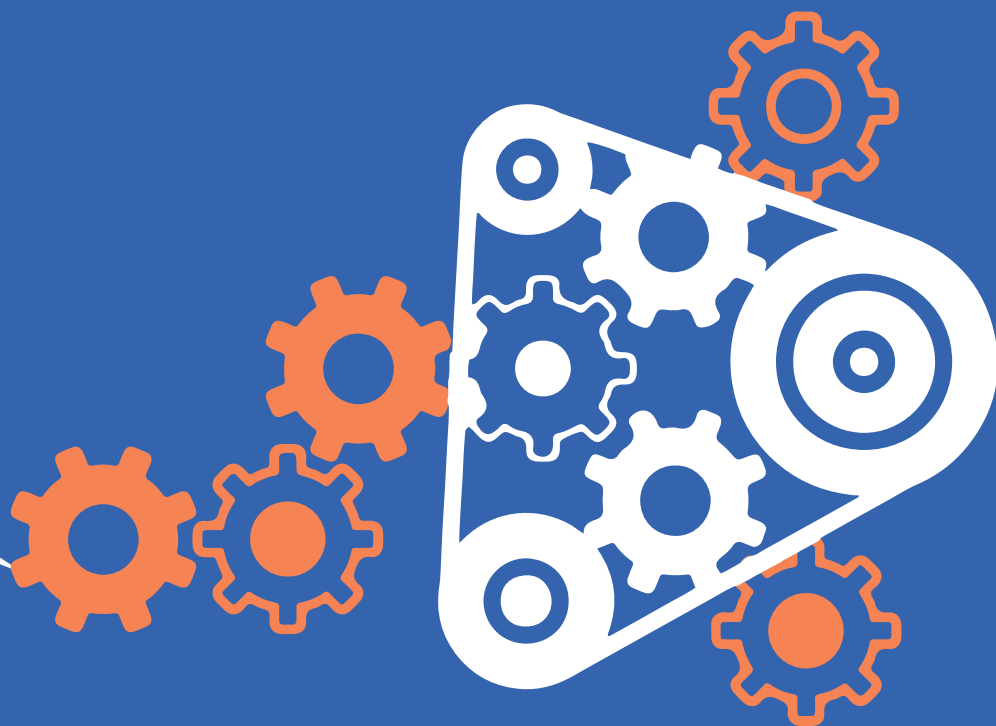


# METODOLOGIA DA PESQUISA FUNCIONALISTA

IVO DA COSTA DO ROSÁRIO (ORG.)



# METODOLOGIA DA PESQUISA FUNCIONALISTA

**Ivo da Costa do Rosário**  
(Organizador)



FUNDAÇÃO UNIVERSIDADE FEDERAL DE RONDÔNIA

Reitora **Marcele Regina Nogueira Pereira**  
Vice-Reitor **José Juliano Cedaro**



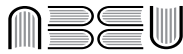
EDITORA DA UNIVERSIDADE FEDERAL DE RONDÔNIA

Editora-chefe: **Franciéle Carneiro Garcês da Sila**  
Bibliotecária: **Rejane Sales de Lima Paula**

CONSELHO EDITORIAL

Presidente **Lou-Ann Kleppa**  
**Carlos Alexandre Barros Trubiliano**  
**Cristiane Marina Teixeira Girard**  
**Gean Carla Silva Sganderla**  
**Geane Valesca da Cunha Klein**  
**Heloisa Helena Siqueira Correia**  
**Júlio César Schweickardt**  
**Márcio Secco**  
**Oswaldo Copertino Duarte**  
**Pedro Ivo Silveira Andretta**  
**Xênia de Castro Barbosa**

Editora Filiada



Associação Brasileira  
das Editoras Universitárias

Edufro - Editora da Universidade Federal de Rondônia  
BR 364, Km 9,5  
Campus Unir  
76801-059 - Porto Velho - RO  
Tel.: (69) 2182-2175  
[www.edufro.unir.br](http://www.edufro.unir.br)  
[edufro@unir.br](mailto:edufro@unir.br)

# METODOLOGIA DA PESQUISA FUNCIONALISTA

**Ivo da Costa do Rosário**  
(Organizador)



Porto Velho - RO

© 2023 by Ivo da Costa do Rosário (Organizador)  
Esta obra é publicada sob a Licença Creative Commons  
Atribuição-Não Comercial 4.0 Internacional.



**Editora-chefe:**  
**Franciéle Carneiro Garcês da Sila**

**Bibliotecária:**  
**Rejane Sales de Lima Paula**

**Capa:**  
**Camila Vieira Vilarim de Sá**

**Revisão:**  
**Alzimar Rodrigues Ramalho**

**Projeto gráfico:**  
**Edufro - Editora da Universidade Federal de Rondônia**

**Diagramação:**  
**Alcindo Donizeti Boffi**

**Impressão e acabamento:**  
**Seike & Monteiro Editora**

**Aprovado no Edital 2021/EDUFRO**

**Esta obra foi avaliada por pareceristas no sistema de duplo-cego**

Dados Internacionais de Catalogação na Publicação  
Fundação Universidade Federal de Rondônia  
Ficha catalográfica elaborada pela Editora da Universidade Federal de Rondônia

---

F981 Fundação Universidade Federal de Rondônia.

Metodologia da pesquisa funcionalista. / Ivo da Costa do Rosário (Organizador). –Porto Velho, RO, Edufro, 2023.  
222 p.: il.

ISBN: 978-85-7764-105-5 (digital)  
ISBN: 978-85-7764-109-3 (impresso)

1. Linguística. 2. Funcionalismo. 3. Análise linguística da pesquisa. I. Rosário, Ivo da Costa do. II. UNIR. III. Título.

CDU 81-11

---

Bibliotecária Rejane Sales de Lima Paula - CRB 11/903

# Sumário

- 7      **Prefácio**  
Ivo da Costa do Rosário (UFF/CNPq/Faperj)  
Monclar Guimarães Lopes (UFF)
- 15     **Linguística Funcional Centrada no Uso: caracterização teórico-metodológica e aplicação prática**  
Maria Angélica Furtado da Cunha (UFRN/CNPq)  
Edvaldo Balduino Bispo (UFRN/CNPq)
- 37     **Metodologia da pesquisa sincrônica**  
Monclar Guimarães Lopes (UFF)  
Ivo da Costa do Rosário (UFF/CNPq/FAPERJ)
- 57     **Metodologia de pesquisa diacrônica**  
Mariangela Rios de Oliveira (UFF/UFOP/CNPq/Faperj)  
Flávia Saboya da Luz Rosa (UFF)
- 77     **Construcionalidade e mudança na sincronia**  
Ivo da Costa do Rosário (UFF/CNPq/FAPERJ)  
Monclar Guimarães Lopes (UFF)
- 103    **Análise Colostrucional**  
Karen Sampaio Braga Alonso (UFRJ/CNPq)  
Diego Leite de Oliveira (UFRJ)
- 121    **Uso do software *Antconc* na análise de dados do uso**  
Patrícia Fabiane Amaral da Cunha Lacerda (UFJF)  
Lauriê Ferreira Martins Dall’Orto (UFJF)
- 137    **Metodologia para a análise prosódica no tratamento de dados do uso**  
Lauriê Ferreira Martins Dall’Orto (UFJF)  
Patrícia Fabiane Amaral da Cunha Lacerda (UFJF)

**151 Metodologia no tratamento da variação construcional**

Marcia dos Santos Machado Vieira (UFRJ/CNPq/Faperj)

Marcos Luiz Wiedemer (UERJ/Faperj)

Pâmela Fagundes Travassos (UFRJ, Université de Lille)

**175 Contribuições da abordagem experimental para o tratamento de dados do uso**

Eduardo Kenedy (UFF / CNPq)

Ana Cláudia Machado dos Santos (UFF)

**214 Informações sobre os autores**

# Prefácio

Por natureza, as teorias de base funcionalista são indutivas. Isso significa que os seus pressupostos teóricos SEMPRE devem representar generalizações a que chegamos a partir de uma análise sistemática e extensiva a partir de dados do uso. Sob esse ponto de vista, podemos afirmar categoricamente: se os dados contrariam uma das generalizações da teoria, ou o dado não foi bem analisado ou, ainda, a teoria não apresenta plena adequação explanatória e, por isso, precisa ser revista e/ou reformulada.

Embora o aspecto indutivo a que estão subordinadas as teorias funcionalistas pareça algo (bastante) desafiador, ele apresenta um lado bom: à medida que as descrições linguísticas das línguas naturais vão avançando, vamos alcançando uma teoria cada vez mais robusta e eficiente do ponto de vista teórico-analítico.

A Linguística Funcional Centrada no Uso (LFCU) – ou Linguística Cognitivo-Funcional – (cf. Cunha et al., 2013; Traugott e Trousdale, 2013; Hilpert, 2014; Oliveira e Rosário, 2015; Diessel, 2019; entre outros), modelo funcional adotado nesta obra, representa uma nova fase do Funcionalismo Norte-Americano, a qual, em nosso ponto de vista, apresenta maior rigor formal e metodológico à descrição gramatical das línguas naturais.

Para a LFCU, toda língua natural é vista como uma rede de construções, isto é, um conjunto bastante extensivo – e produtivo – de pareamentos indissociáveis de forma e função. De um lado, a noção de produtividade está associada ao constante processo de (re)categorização linguística. À medida que a língua vai sendo posta em uso, novas construções emergem e outras se reconfiguram, assumindo novas funções linguísticas. De outro, a noção de pareamento indissociável de forma e função está associada à ideia de



que uma determinada forma está simbolicamente ligada a uma função na língua, de modo que tanto uma forma  $x$  implique uma função  $y$  quanto uma função  $y$  implique uma forma  $x$ . Cabe frisar também que, nesse modelo, a forma é representada por propriedades fonológicas e morfossintáticas; já a função, por propriedades semânticas, pragmáticas e discursivo-funcionais.

O tratamento multidimensional da língua, presente na abordagem funcionalista, não se dá somente em virtude da pluralidade das propriedades formais e funcionais observáveis nas construções linguísticas, como também no próprio tratamento dado ao objeto. Como sabemos, um mesmo objeto pode ser descrito sob diferentes ângulos, decisão essa a ser tomada com base nos objetivos de pesquisa traçados pelo analista. Um pesquisador  $x$ , por exemplo, pode intentar descrever os diferentes usos do verbo *ir* no português e flagrará em sua pesquisa, dentre outros pontos, ocorrências em que *ir* atua como 1) verbo pleno de um sintagma verbal; 2) verbo auxiliar de uma perífrase verbal. Um pesquisador  $y$ , por sua vez, pode buscar descrever a trajetória de mudança do verbo *ir*, isto é, investigar se um dos usos supracitados é originário do outro em sincronias anteriores do português. Além disso, um pesquisador  $z$  pode buscar investigar a variação existente entre a construção de perífrase de futuro com o verbo *ir* (por exemplo, *vou comprar*) e a construção do futuro do presente (por exemplo, *comprarei*).

Como podemos notar, os objetivos de pesquisa vão implicar análises distintas e, paralelamente, vão exigir metodologias distintas. Inclusive, há uma crítica muito pertinente e bem fundamentada por Martelotta (2009) ao questionar o emprego extensivo e sistemático de metodologia quantitativa, nos moldes da sociolinguística variacionista, aos estudos funcionalistas norte-americanos<sup>1</sup>, já que as duas correntes têm finalidades distintas. Afinal, o método deve estar articulado aos objetivos almejados pela pesquisa. É exatamente por esse motivo que este livro procura agregar diferentes abordagens metodológicas usualmente empregadas em pesquisas desenvol-

---

<sup>1</sup> Vale ressaltar que o referido texto do Martelotta (2009) é um dos primeiros estudos de referência sobre metodologia em pesquisa funcionalista. O texto original está disponível no seguinte endereço: [https://bit.ly/deg\\_mario\\_metodologia](https://bit.ly/deg_mario_metodologia).

vidas sob a perspectiva da LFCU. Apesar de haver diferentes obras de referência voltadas para a descrição dos pressupostos teóricos do modelo, faltava uma obra como esta, com o fim didático de demonstrar a aplicação prática de diferentes abordagens metodológicas desenvolvidas nas pesquisas em LFCU.

Ao longo do livro, organizado em nove capítulos, os autores apresentam diferentes perspectivas metodológicas para o tratamento dos dados em LFCU, algumas delas em interface com outras linhas de pesquisa, como a Linguística de *Corpus*, a Linguística Experimental e a Sociolinguística Variacionista. Todos os capítulos contam com dados empíricos de pesquisas realizadas pelos próprios autores ou por eles orientadas. Com isso, busca-se ilustrar, de forma didática, a aplicação de cada um dos métodos.

No primeiro capítulo, *Linguística Funcional Centrada no Uso: caracterização teórico-metodológica e aplicação prática*, Maria Angélica Furtado da Cunha e Edvaldo Balduino Bispo descrevem e ilustram conceitos básicos da LFCU, como iconicidade, marcação, informatividade e plano discursivo, bem como processos cognitivos, projeções metafóricas e metonímicas, construção e seus fatores (composicionalidade, esquematicidade e produtividade). Além disso, mostram a aplicação prática desses conceitos por meio de um estudo descritivo da expressão do grau no português do Brasil.

No segundo capítulo, *Metodologia da pesquisa sincrônica*, Monclar Guimarães Lopes e Ivo da Costa do Rosário tratam do emprego do método misto para a descrição de dados sincrônicos sob a perspectiva da LFCU. Para isso, mostram uma série de procedimentos com que o analista deve se preocupar na elaboração de sua pesquisa, tais como: a) escolha do objeto de pesquisa; b) seleção ou constituição de *corpus*; c) revisão de literatura; d) análise piloto; e) levantamento das hipóteses e planejamento dos fatores de análise; f) análise sistemática, tabulação das ocorrências e apresentação dos dados. Por fim, como ilustração, aplicam esses procedimentos no planejamento da investigação da construção conectora [fora isso] no português.

No terceiro capítulo, *Metodologia da pesquisa diacrônica*, Mariangela Rios de Oliveira e Flávia Saboya da Luz Rosa demonstram como o ana-

lista pode lidar metodologicamente com a pesquisa histórica na LFCU a partir da análise interpretativa dos contextos de uso e de sua gradiência ao longo da trajetória da língua. Propõem ao modelo de análise indicado por Diewald (2006) e Diewald e Smirnova (2012) a inclusão dos nanopassos de mudança para os objetos que necessitam de descrição intracontextual. Como exemplo, empregam o modelo proposto na análise da emergência do marcador discursivo refrador-argumentativo [ $Indut_R Afix_{Loc}$ ] no português, que licencia microconstruções como *espera aí*, *perai*, *espera lá*, entre outras.

No quarto capítulo, *Construcionalidade e mudança na sincronia*, Ivo da Costa do Rosário e Monclar Guimarães Lopes reelaboram o conceito de construcionalidade (Rosário e Lopes, 2019), com o objetivo de acentuar mais fortemente a interface sincronia x diacronia. Trata-se de um modelo que tem como meta a investigação da origem e da trajetória histórica das construções linguísticas com base em dados sincrônicos, sob a defesa de que “as mudanças são sempre manifestadas na variação sincrônica, e as mudanças do passado são comumente refletidas nas alternâncias sincrônicas” (Hoffmann e Trousdale, 2011, p. 12). Como ilustração de aplicação do modelo, os autores reinterpretam resultados de duas pesquisas recentemente desenvolvidas sob a perspectiva da LFCU à luz da construcionalidade.

No quinto capítulo, *Análise colostrucional*, Karen Sampaio Braga Alonso e Diego Leite de Oliveira apresentam um conjunto de métodos conhecidos nos modelos linguísticos baseados no uso. Ao todo, apresentam três modelos de análise: a) análise colexêmica simples; b) análise colexêmica distintiva; c) análise colexêmica covariacional. Cada um deles apresenta objetivos relativamente distintos e auxilia o pesquisador a determinar as relações de preenchimento de *slots* de uma dada construção (cf. Diessel, 2019). Como exemplo, os autores apresentam dados de construções licenciadas pelo subesquema quantificador [um N de N], como [um monte de N], [uma montanha de N] e [uma enxurrada de N].

No sexto capítulo, *Uso do software Antconc na análise de dados do uso*, Patrícia Fabiane Amaral da Cunha Lacerda e Lauriê Ferreira Mar-

tins Dall’Orto apresentam os procedimentos para a operacionalização do referido *software*, projetado para a análise linguística de *corpora*. Para as estudiosas, trata-se de uma ferramenta bastante útil para a seleção e interpretação dos dados em pesquisas de natureza funcionalista. O foco do texto está na descrição das funções destinadas ao processamento de dados em *corpora*, a saber: i) wordlist; ii) concordance; iii) Concordance Plot; iv) File View; v) Clusters/N-Grams; vi) Collocates; e vii) Keywords.

No sétimo capítulo, *Metodologia para a análise prosódica no tratamento de dados do uso*, Lauriê Ferreira Martins Dall’Orto e Patrícia Fabiane Amaral da Cunha Lacerda apresentam o uso do *software* PRAAT, que permite a análise acústica de dados gravados da modalidade oral. De acordo com as pesquisadoras, o emprego desse *software*, a depender do objeto de pesquisa, permite uma descrição mais adequada das propriedades fonológicas das construções, aspectos linguísticos normalmente menos considerados nas investigações em LFCU. Como exemplo, mostram quatro padrões construcionais complexos com o adjetivo *real*, cuja função está associada à presença ou não de *break* acústico durante sua realização na fala.

No oitavo capítulo, *Metodologia no tratamento da variação construcional*, Márcia dos Santos Machado Vieira, Marcos Luiz Wiedemer e Pâmela Fagundes Travassos apresentam uma abordagem multimetodológica e multidimensional para a investigação da variação construcional, desenvolvida com base nas propostas de Leino & Östman (2005), Cappelle (2006) e Perek (2015). Segundo os autores, trata-se de um modelo que propõe estabelecer a correlação entre atributos/valores associados às aloconstruções como parâmetros/fatores condicionantes dos usos de cada construção, que podem ser calculados em termos de atração/coerção representativa/esquemática de cada construção. Como ilustração, apresentam os resultados da investigação de duas variantes da construção com verbo-suporte: DAR + (uma) X-(z)inho[o,a] e DAR + (uma) X-adela.

No último capítulo, *Contribuições da abordagem experimental para o tratamento de dados do uso*, Eduardo Kenedy e Ana Cláudia Machado dos Santos apresentam a operacionalização da pesquisa experimental em

Linguística, uma abordagem metodológica que vem sendo cada vez mais valorizada e aplicada nas pesquisas desenvolvidas em perspectiva construcional. Ao longo do capítulo, os autores tratam, dentre outros pontos, dos diferentes modelos de experimento, bem como de sua elaboração, execução e interpretação dos resultados. Como ilustração, apresentam três experimentos aplicados na análise de marcadores discursivos do português formados pelo esquema VLoc, como, por exemplo, *olha lá, olha aí, espera lá, vamos lá* etc.

Por fim, gostaríamos de ressaltar o que já deve ter ficado evidente na leitura deste prefácio: os textos que compõem esta obra, redigidos em uma linguagem didática e acessível por especialistas brasileiros de renome na perspectiva da Linguística Funcional Centrada no Uso, são referências de grande valia tanto para o planejamento quanto para a execução metodológica de pesquisas que visem à descrição gramatical com base em dados empíricos do uso linguístico. Neste sentido, temos a certeza de que sua leitura será de grande auxílio tanto ao pesquisador iniciante quanto ao mais experiente. Desejamos a todos uma boa leitura!

Ivo da Costa do Rosário (UFF/CNPq/Faperj)  
Monclar Guimarães Lopes (UFF)

## Referências

- CAPPELLE, B. Particle placement and the case for 'allostructions'. In: SCHÖNEFELD, D. (ed.) **Constructions All Over: Case Studies and Theoretical Implications**, Special volume of *Constructions* SV1- 7/2006, 2006. p. 01-28. Disponível em: <<http://www.constructions-online.de/articles/specvol1>>. Acesso em: 04 mar. 2019.
- CUNHA, M. A. F.; BISPO, E. B.; SILVA, J. R. Linguística Funcional Centrada no Uso: conceitos básicos e categorias analíticas. In: CEZARIO, M. M.; CUNHA, M. A. F. (orgs.). **Linguística Centrada no Uso: uma homenagem a Mário Martelotta**. Rio de Janeiro: Mauad-Faperj, 2013. p. 13-39.
- DIESEL, H. **The Grammar Network: How Linguistic Structure is Shaped by Language Use**. New York: Cambridge University Press, 2019.
- DIEWALD, G. Context types in grammaticalization as constructions. **Constructions**. Düsseldorf, 2006. Disponível em: <[www.constructions-online.de/009-4-63860](http://www.constructions-online.de/009-4-63860)>. Acesso em: 20 fev. 2023.
- DIEWALD, G.; SMIRNOVA, E. Paradigmatic integration: the fourth stage in an expanded grammaticalization scenario. In: DAVIDSE, K.; BREBAN T.; BREMS, L.; MORTELMANS, T. (eds.). **Grammaticalization and Language Change: New reflections**. Amsterdam: John Benjamins, 2012. p. 111-133.
- HILPERT, M. **Construction Grammar and its application to English**. Edinburgh: Edinburgh University Press, 2014.
- HOFFMANN, T.; TROUSDALE, G. **The Oxford Handbook of Construction Grammar**. New York: Oxford University Press, 2011.
- LEINO, J.; ÖSTMAN, J-O. Constructions and variability. In: FRIED, M.; BOAS, H. (eds.) **Grammatical Constructions: Back to the roots**. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2005. p. 191-213.
- MARTELOTTA, M. E. Funcionalismo e metodologia quantitativa. In: OLIVEIRA, M. R.; ROSÁRIO, I. C. (Org.). **Pesquisa em linguística funcional: convergências e divergências**. Rio de Janeiro: Leo Christiano Editorial, 2009. p. 1-20.
- OLIVEIRA, M. R.; ROSÁRIO, I. C. **Linguística Centrada no Uso**. Rio de Janeiro: Lamparina, FAPERJ, 2015.
- PEREK, F. **Argument Structure in Usage-Based Construction Grammar**. Amsterdam: John Benjamins, 2015.

ROSÁRIO, I. C.; LOPES, M. G. Construcionalidade: uma proposta de aplicação sincrônica. **Revista Soletras**, v. 37, p.83-102, 2019.

TRAUGOTT, E. C.; TROUSDALE, G. **Constructionalization and Constructional Changes**. New York: Oxford University Press, 2013.

# 1. Linguística Funcional Centrada no Uso: caracterização teórico-metodológica e aplicação prática

Maria Angélica Furtado da Cunha (UFRN/CNPq)  
Edvaldo Balduino Bispo (UFRN/CNPq)

## Considerações iniciais

Na Linguística dos anos 1960, estabelece-se, de forma mais velada ou mais acirrada, uma polarização entre a linguística formal e a linguística funcional. A primeira é representada pela já enfraquecida Linguística Estrutural e a assim denominada *mainstream* Gramática Gerativa, idealizada por Noam Chomsky (1957, 1959). A segunda é constituída, principalmente, por pesquisadores da Costa Oeste dos Estados Unidos, por Michael Halliday, no Reino Unido, e por Simon Dik, na Holanda. Em consequência dessa oposição, pode-se dizer que, no século XX, os estudos linguísticos concentram-se em dois grandes eixos: o formalista e o funcionalista. O que distingue esses dois polos é, sobretudo, a concepção de língua e os objetivos da investigação de cada um deles. Para os estudiosos que se alinham ao polo formalista, o interesse central é a análise da forma linguística, enquanto a função é deixada de lado; para os linguistas que se filiam ao polo funcionalista, por sua vez, o foco está na investigação da função que as formas da língua desempenham em situações de interação comunicativa. Vale dizer que diversos autores compartilham essa visão bipartida da Linguística, como Dirven e Fried (1987), Schiffrin (1994), Kato (1998) e Martelotta e Kenedy (2015), entre outros.

Dada essa visão, muitos textos funcionalistas, sejam capítulos, livros ou artigos, assumem abertamente uma posição contrária à Gramática Gerativa. Um exemplo disso pode ser visto no capítulo introdutório do



livro *On understanding grammar* (Givón, 1979), obra seminal que cria as raízes do funcionalismo norte-americano moderno. Nele, o linguista critica duramente o paradigma da Gramática Gerativo-Transformacional, classificado por ele como “um modelo formal dos dados” (2012 [1979], p. 18), tanto em termos de teoria quanto de metodologia, e apresenta seu modelo funcionalista de análise linguística, desenvolvido no capítulo 2 em diante.

Na mesma linha, no capítulo introdutório do livro *An Introduction to functional grammar*, Halliday (1985) argumenta que a oposição fundamental que separa as correntes teóricas é a divisão entre aquelas que são primariamente sintagmáticas, fundamentadas na lógica e na filosofia – as gramáticas formais, compreendendo o estruturalismo e o gerativismo – e as que são primariamente paradigmáticas – as funcionalistas, respaldadas na retórica e na etnografia. Passa, então, a discorrer sobre as falhas do modelo chomskyano. É oportuno mencionar, entretanto, que se atribui a Chomsky (1957, 1959) o começo da era moderna da linguística, tendo sido ele parte central da Revolução Cognitiva dos anos 1960 (Tomasello, 1998).

No Brasil, essa polarização pode ser ilustrada pelo debate empreendido entre os funcionalistas Sebastião Votre e Anthony Naro e o gerativista Milton do Nascimento. Em seu artigo, com referência à hipótese de Votre e Naro (1989, p. 169-170) de que “do uso da língua origina-se a forma da língua”, Nascimento (1990, p. 83) propõe um “diálogo construtivo entre linguistas brasileiros que conduzem suas pesquisas, utilizando-se de diferentes quadros teóricos”. Convida, então, os leitores a “repensar a tese de que a ‘abordagem funcionalista’ se contrapõe à ‘abordagem formalista’” (ibidem, p. 87). O linguista finaliza seu artigo defendendo uma articulação dos “modelos de análise linguística que privilegiam as funções com aqueles que privilegiam a forma da linguagem”, levando em conta as especificidades de cada modelo (ibidem, p. 97).

No final do século XX, o psicolinguista Michael Tomasello (1998) organiza um livro em cuja introdução recapitula a situação da Linguística naquele século e rejeita a posição chomskyana de autonomia da sintaxe. Lista uma família de abordagens linguísticas que compartilham essa rejeição e que recebem a denominação de Linguística Cognitiva ou Linguística

Funcional. A propositura de Tomasello é que essas duas abordagens se unam em um paradigma científico coerente, tendo em vista que concordam que a língua não é um “órgão mental” autônomo, como quer Chomsky, mas sim um mosaico complexo de atividades cognitivas e sociais diretamente integrado com o resto da psicologia humana. Surge, assim, a Linguística Cognitivo-Funcional, que, como o próprio nome diz, integra tanto linguistas cognitivistas como funcionalistas.

Toda essa contextualização histórica tem por objetivo situar a Linguística Funcional Centrada no Uso (LFCU), tendência funcionalista moderna que os pesquisadores do grupo de estudos Discurso & Gramática aplicam em suas investigações<sup>1</sup>. Esse modelo de análise linguística se alimenta dos postulados e das categorias analíticas tanto da Linguística Funcional Clássica/norte-americana quanto da Gramática de Construções. A primeira tem como principais representantes Talmy Givón, Paul Hopper, Sandra Thompson, Wallace Chafe, Joan Bybee, ao passo que a segunda é representada, principalmente, por George Lakoff, Ronald Langacker, Adele Goldberg e William Croft, para citar só alguns (Tomasello, 1998, 2003; Martelotta, 2011; Cunha; Bispo; Silva, 2013). Essas duas correntes compartilham vários pressupostos teórico-metodológicos, como a rejeição à autonomia da sintaxe, a incorporação da semântica e da pragmática às análises, a não distinção estrita entre léxico e gramática, a relação estreita entre a estrutura das línguas e o uso que os falantes fazem delas nos contextos reais de comunicação, o entendimento de que os dados para a análise linguística são enunciados que ocorrem no discurso natural, por exemplo.

### **Caracterização teórica**

De acordo com a LFCU, as línguas são moldadas pela interação complexa de princípios cognitivos e comunicativos. Uma vez que as línguas naturais são muito semelhantes no que diz respeito às relações gramaticais

---

<sup>1</sup> A Linguística Funcional Centrada no Uso aproxima-se, em termos teóricos, metodológicos e epistemológicos, ao que Bybee (2015, 2016 [2010]) denomina *Usage-based Linguistics*. De acordo com essa autora, a teoria baseada no uso se desenvolveu diretamente do funcionalismo norte-americano e, em certo sentido, é apenas um novo nome para ele.

que exibem, como sujeito, objeto direto, objeto indireto etc., essas semelhanças podem ser vistas como o resultado desses princípios cognitivos e funcionais. Nesse sentido, a proposição central da LFCU é que a gramática de qualquer língua resulta da regularização ou rotinização de estratégias discursivas recorrentes (Givón, 2012 [1979]; Bybee, 2016 [2010]), de modo que gramática e uso são interdependentes. A língua é entendida como um sistema adaptativo complexo, uma estrutura plástica, emergente (Du Bois, 1985; Hopper, 1987; Bybee, 2016 [2010]), que se amolda aos contextos em que ela é usada. Convivem, ao mesmo tempo, padrões mais ou menos regulares e outros que surgem em virtude de necessidades cognitivas e/ou comunicativas (Givón, 2001; Bybee, 2016 [2010]). O sistema linguístico é, pois, dinâmico, já que surge da adaptação das habilidades cognitivas humanas a eventos de comunicação específicos e se desenvolve com base na repetição desses eventos. De fato, para Bybee (2016 [2010], p. 35), “as circunstâncias do uso impactam a representação cognitiva da língua”, o que reforça sua proposta anterior de que a gramática deve ser pensada como uma “organização cognitiva da experiência linguística” (Bybee, 2006, p. 730). Em suma, a aparente regularidade e instabilidade da língua são motivadas e modeladas pelas práticas discursivas dos falantes no cotidiano social (Cunha; Tavares, 2016).

Nesse sentido, podemos falar em variação e gradiência dos elementos linguísticos: num viés sincrônico, o uso constante da língua pelos falantes cria variação, como se dá com as formas *nós* e *a gente*, para expressão da primeira pessoa do plural, e *tu* e *você*, para codificar a segunda pessoa do discurso no português brasileiro. Numa perspectiva diacrônica, a gradiência implicada na variação pode levar à mudança, que envolve gradualidade e na qual uma das formas alternativas pode cair em desuso. No caso dessas formas pronominais, há variação estável, tendo em vista que elas coexistem na sincronia atual. Desse modo, a gradiência se refere ao fato de que muitas categorias da língua (e da gramática) não podem ser facilmente distinguidas devido à variação que há entre unidades de uma mesma categoria (em diferentes níveis) e em função da mudança que ocorre ao longo do tempo, de modo gradual, em pequenos passos, movendo um elemento em um contínuo

de uma categoria à outra. Como exemplo de mudança categorial, podemos citar o uso do verbo *ter* como pleno, com sentido referencial (*Mais de 200 anos depois, 'napoleônico' ainda é sinônimo de megalômano: aquele que tem delírios de grandeza*<sup>2</sup>), e seu uso como auxiliar em tempos compostos, com funcionamento gramatical (*E se Napoleão tivesse vencido?*). No estudo do surgimento, da variação e da mudança dos elementos linguísticos, a LFCU considera motivações comunicativas e cognitivas, uma vez que postula uma relação estreita entre a estrutura das línguas e o uso que os falantes fazem delas nos contextos reais de comunicação.

Assim, o uso da língua é central para a LFCU, de sorte que os fenômenos sob investigação são analisados em relação às funções semântico-cognitivas e discursivo-pragmáticas que eles desempenham na comunicação. Essas funções, por sua vez, influenciam a organização do sistema linguístico, contingenciando os usos da língua em práticas interacionais efetivas.

Nesse cenário, entende-se que fenômenos de natureza cognitiva têm papel importante na compreensão da linguagem. Esses fenômenos dizem respeito ao “modo como nossa mente interage com o mundo que nos cerca, bem como os processos que permeiam essa interação.” (Martelotta; Palomanes, 2008, p. 177). A LFCU assume, com Bybee (2016 [2010]), que a estrutura linguística deriva da aplicação repetida de processos cognitivos de domínio geral<sup>3</sup>, os quais não se restringem à linguagem, mas atuam também em outros aspectos da cognição, tais como visão, raciocínio matemático e habilidades musicais, por exemplo. Nessa direção, os padrões linguísticos são parte da nossa capacidade de domínio geral para categorizar, estabelecer relações e operar em nível tanto local quanto global.

Com a incorporação de uma perspectiva construcional a investigações funcionalistas, a gramática passa a ser compreendida também como um conjunto de esquemas simbólicos (construções), pareamentos de forma-função (Goldberg, 1995), utilizados na produção e organização

---

<sup>2</sup> Disponível em: <https://super.abril.com.br/historia/e-se-napoleao-tivesse-vencido/>. Acesso em: 22 jun. 2021.

<sup>3</sup> Bybee elenca *categorização, memória rica, chunking, analogia, associação transmodal*.

de discurso coerente. “Configura-se em categorias morfossintáticas rotinizadas, exibindo padrões funcionais mais regulares e formas alternativas em processo de regularização, motivada por fatores cognitivo-interacionais.” (Cunha; Bispo; Silva, 2013, p. 20)

As construções são armazenadas na mente do falante com base em enunciados reais, por meio do processo de categorização de instâncias que ocorrem frequentemente no uso interacional da língua. A interpretação de que a gramática é composta por construções (Goldberg, 2006) acarreta o entendimento de que a relação entre forma e função é básica e inerente a qualquer descrição gramatical (Östman, 2005).

Para alguns construcionistas, a construção varia em extensão, indo do morfema (-o, desinência da 1ª pessoa do singular do presente do indicativo, por exemplo) ao texto (o *slogan Tomou Doril, a dor sumiu*), incluindo vocábulo (*voz*), sintagmas (*mão de vaca*) e orações simples (*Como é o processo seletivo dos astronautas?*<sup>4</sup>) e estruturas oracionais complexas (*EUA aprovam novo remédio contra Alzheimer – mas não há consenso sobre sua eficácia*<sup>5</sup>).

Conforme Croft (2001), a arquitetura da construção envolve duas dimensões: a da forma, que compreende propriedades fonológicas, morfológicas e sintáticas; e a do sentido (função), que abarca propriedades semânticas, pragmáticas e discursivo-funcionais. A segunda dimensão inclui todos os aspectos interacionais que circunstanciam o uso linguístico. Essas dimensões estão relacionadas por elos de correspondência simbólica.

A relação entre forma e função, aliás, é pedra de toque às investigações funcionalistas e representa a base do princípio de iconicidade, definida como a relação motivada entre conteúdo e expressão (Givón, 1984; Haiman, 1985). Evidências translinguísticas demonstram que acréscimos de conteúdo semântico têm correspondência na codificação formal (Givón, 1995; Croft, 2001). Assim, por exemplo, formas derivadas, que carregam mais conteúdo que as primitivas, são, em geral, maiores do que estas (amigo

---

<sup>4</sup> Disponível em: <https://super.abril.com.br/>. Acesso em: 10 jun. 2021.

<sup>5</sup> Disponível em: <https://super.abril.com.br/>. Acesso em: 10 jun. 2021.

> amigável; esperar > esperançoso). Conforme Givón (1984), a iconicidade pode ser compreendida em termos dos subprincípios de quantidade, de integração/proximidade e de ordenação.

Relacionado à iconicidade está o princípio de marcação, o qual diz respeito à presença de uma propriedade em um determinado membro de uma categoria, considerado marcado, em contraste com outro(s) membro(s), que não a exibe(m) ou a exibe(m) em menor grau. A iconicidade pode ser aferida, conforme Givón (1990), por meio de três parâmetros: complexidade estrutural, complexidade cognitiva e distribuição de frequência. Envolve gradiência e dependência contextual, daí um padrão linguístico poder ser marcado, não marcado, mais marcado, menos marcado; marcado em um dado contexto, mas não em outro. É o que nos mostram, por exemplo, Bispo e Cunha (2019), com o *continuum* de marcação para as estratégias de relativização.

Outros constructos teóricos considerados na relação entre forma e função são a informatividade e o plano discursivo. A primeira está relacionada ao conteúdo informacional que os interlocutores compartilham, ou supõem compartilhar, no momento da interação verbal. Os estudos clássicos sobre informatividade voltam-se ao estatuto informacional de referentes nominais, classificando-os em *novvo*, *evocado* (dado ou velho), *inferível* e *disponível* (Prince, 1981). Essas categorias correspondem à avaliação que o falante faz do estado de ativação dos elementos referenciais na memória do interlocutor (Chafe, 1976; Cunha; Costa; Cezario, 2015). A informatividade também está relacionada ao modo como o conteúdo é veiculado numa prática interacional, de modo que o falante procura dosar esse conteúdo com base em suas suposições acerca do conhecimento do interlocutor, além de monitorá-lo com vistas ao alcance de determinados objetivos comunicativos.

A noção de plano discursivo refere-se à organização estrutural do texto e compreende as dimensões de figura e fundo, cuja formulação original se deve à Gestalt. Essas dimensões relacionam-se à percepção e à cognição: as entidades que aparecem em primeiro plano (ou seja, as mais salientes) são percebidas com mais nitidez e facilidade, enquanto as que se

encontram fora de destaque são menos aparentes e perceptíveis. Em termos de discurso, essa distinção equivale à oposição entre informações centrais e periféricas. Essa categoria tem sido utilizada, em investigações funcionalistas, para tratar da distribuição de informações no texto, correlacionando aspectos discursivos a propriedades gramaticais (Hopper; Thompson, 1980; Cunha, 1996; Martelotta, 1998; Lemos, 2020).

Conforme já referido, a organização linguística também resulta da atuação de processos cognitivos, a exemplo de categorização. Dela decorre a prototipicidade, conceito caro à LFCU por revelar o caráter gradiente das categorias gramaticais. A prototipicidade é, segundo Rosch (1973), possivelmente uma consequência de propriedades inerentes da percepção humana, como a saliência cognitiva. O representante prototípico de uma categoria reúne os traços recorrentes de que se compõe essa categoria, enquanto os outros elementos são classificados considerando as características mais próximas e as mais distantes em relação a esse representante. Essa perspectiva não discreta permite o tratamento escalar de padrões gramaticais, no sentido de identificar alguns mais representativos e outros mais marginais, como é o caso dos termos *de* e *durante*, respectivamente, para o grupo de preposições em português.

Outros processos de natureza cognitiva também considerados pela LFCU, sobretudo em casos de variação e de mudança linguísticas, são a metáfora e a metonímia. Trata-se de operações cognitivas que envolvem mapeamentos, quer entre domínios, quer intradominiais. Na metáfora, ocorre mapeamento em que determinadas noções de um domínio são projetadas em outro (por exemplo, tempo é concebido em termos de espaço, daí expressões como “de hoje em diante” e “cinco anos atrás”). Desse modo, um conceito é formulado em termos de outro pelo fato de compartilharem alguma(s) correspondência(s) conceitual(is) (Lakoff; Johnson, 1999). Já a metonímia constitui um mapeamento dentro de um mesmo domínio conceitual, por meio do qual se consegue chegar a uma entidade com base em outra via contiguidade (Lakoff; Turner, 1989), conforme se dá com *calça* e *jeans* em “*No meu dia a dia, como estou bastante na ponte aérea, sou mais básica, do jeans e das camisetas, e invisto em peças curinga, que amassam*

*pouco*<sup>6</sup>. As relações de contiguidade em que se baseiam os vários tipos de metonímia são diversas, incluindo não apenas o sentido espacial, mas também o temporal, o causal, o conceptual.

Sob o olhar construcionista, a relação entre forma e função é tratada em termos da propriedade de composicionalidade (Traugott; Trousdale, 2013), que diz respeito ao grau de transparência na relação entre codificação e conteúdo. Nesse sentido, uma construção é composicional se o seu significado resulta da soma do significado dos itens que a compõem. De modo inverso, uma construção é não composicional se o significado do todo não corresponde à soma do significado das partes. Assim é que, por exemplo, em *tirar a comida da mesa*, há convergência entre o sentido do todo (remoção de alimento de um lugar para outro) e a junção do significado dos elementos que integram essa expressão. Já em *tirar a barriga da miséria*, não existe convergência entre o sentido da expressão (desfrutar intensamente de algo depois de muito tempo de privação) e a soma do significado dos itens constituintes. Cabe registrar que a composicionalidade é gradiente, de modo de que há construções composicionais, menos composicionais e não composicionais, revelando, assim, maior ou menor transparência ou mesmo opacidade completa na relação forma-sentido (Cunha; Bispo, 2019).

A construção é também examinada com base nas propriedades de esquematicidade e produtividade (Langacker, 2008; Traugott; Trousdale, 2013). *Esquematicidade* refere-se ao fato de a construção servir como um modelo abstrato, virtual que captura padrões de uso. Sob essa perspectiva, as construções podem ser totalmente não especificadas (ou abertas), como o esquema oracional intransitivo SN V (*O garoto nasceu*); parcialmente especificadas, como se pode ver na sequência *fazer* SN (*fazer bolo*); totalmente especificadas (ou idiossincráticas), tal como a expressão *dar com os burros n'água*.

A produtividade está relacionada à possibilidade de uma construção licenciar outras menos esquemáticas, isto é, tem a ver com o grau em

---

<sup>6</sup> Disponível em: <https://www.corpusdoportugues.org/now/>. Acesso em: 11 jun. 2021.



que uma construção pode se expandir. É o que ocorre, por exemplo, com a construção adjetival X-*vel*, cuja base lexical prototípica é derivada de verbo transitivo, como *reciclar* > *reciclável*; *remover* > *removível*. Esse esquema de formação morfológica se tornou produtivo de modo a sancionar a emergência de adjetivos com base em verbos não transitivos, a exemplo de *durar* > *durável*; *falir* > *falível*, e até mesmo em substantivos, conforme vemos em *pote* > *potável*; *via* > *viável* (Jovem; Silva, 2017).

Cabe destacar que essas propriedades são gradientes, dado que uma construção pode variar de totalmente esquemática a completamente preenchida, pode ser altamente produtiva, ter grau mediano de produtividade ou mesmo ter produtividade zero. Assim, por exemplo, a construção proverbial *Casa de ferreiro, espeto de pau* é plenamente preenchida e apresenta grau zero de produtividade, tendo em vista que licencia *type* único; a construção *tirar SN<sub>1</sub> de SN<sub>2</sub>*, por sua vez, apresenta esquematicidade parcial (há dois elementos fixos e dois *slots* a ser preenchidos), e produtividade mediana, podendo instanciar vários *types*, a exemplo de *tirar o pó da estante*, *tirar leite de pedra*, *tirar o pai da força*.

### **Caracterização metodológica**

Ao assumir que as formas linguísticas são motivadas por fatores de natureza vária (estruturais, sociocomunicativos, cognitivos e históricos), a LFCU tem o compromisso primeiro de lidar com dados da língua em situações efetivas de interação verbal. Esses fatores, em conjunto, atuam de modo diverso nos diferentes contextos de comunicação, complementando-se em uns casos e anulando-se em outros. Isso implica a adoção de uma metodologia que considere não apenas a interdependência desses fatores, mas sua atuação contextualmente diferenciada.

Os pesquisadores orientados pela LFCU buscam identificar e avaliar fatores de natureza semântico-cognitiva e discursivo-pragmática que regulam as manifestações do fenômeno investigado, atentos para as restrições de natureza formal que estimulam ou bloqueiam a regularização desse fenômeno.

Metodologicamente, valemo-nos tanto do raciocínio indutivo quanto do dedutivo, o que caracteriza, conforme Givón (1995), o método abduutivo de análise. O viés indutivo reside na consideração de instâncias particulares do fenômeno sob investigação para, a partir delas, chegar a determinadas generalizações. O caráter dedutivo decorre da ancoragem em uma base teórica consolidada, com postulados e premissas consistentes e fortemente referendadas na empiria.

Em termos de abordagem do fenômeno estudado, utilizamos uma análise que conjuga fatores qualitativos e quantitativos, que possam evidenciar tendências. O enfoque qualitativo diz respeito ao caráter descritivo e interpretativo da análise e ao viés indutivo, baseado na observação das amostras coletadas. Já a dimensão quantitativa refere-se à natureza mensurável do material empírico tomado como amostra. Para tanto, quantificamos, em termos absolutos e percentuais, a recorrência dos fatores linguísticos selecionados para a análise.

Quanto aos aspectos investigados, podemos citar a aferição da frequência de uso, a configuração das estruturas linguísticas no contexto linguístico e, no contexto comunicativo, os processos sociointeracionais e cognitivos subjacentes aos usos linguísticos. A frequência de uso de uma determinada construção leva a seu estabelecimento no repertório do falante e faz dela uma unidade de processamento, o que implica que o falante explora recursos gramaticais disponíveis para atingir seus objetivos comunicativos. Mas é importante não esquecer que o discurso exhibe padrões recorrentes que extrapolam o que é predizível pelas regras gramaticais, e a explicação para a existência desses padrões deve ser procurada no âmbito da cognição e da comunicação.

Testamos as hipóteses a respeito de aspectos sincrônicos e diacrônicos com dados de textos reais (orais, escritos e/ou multimodais). Nosso propósito é descrever e interpretar os fatos linguísticos com base nas funções que eles desempenham nos diversos contextos de uso da língua, integrando sincronia e diacronia. Interessa-nos identificar diferentes motivações funcionais e avaliar o efeito de cada uma delas na configuração concreta do fenômeno sob estudo.

## Aplicação da teoria: análise do grau

Seguindo Silva (2014), analisamos o grau como um domínio funcional ao qual se associam diferentes formas de expressão. Para essa análise, utilizamos as noções de iconicidade, marcação, informatividade, projeções metafóricas e metonímicas. Consideremos as ocorrências a seguir:

- (1) F1- é  *muito chato*  saber que está sendo gravado viu...  
F2- daqui a uns dez  *minutinhos*  você esquece...  
F1- [risos]  *pior*  que a gente num pode  *esquecer demais*  não... senão fala  *muita besteira* ...  
F2- e lá em casa... é  *tanto nome*  que sai... eu não tenho o que fazer aí eu vou...  
(Cunha, 2011, p. 1)

Nessa amostra, vemos diversas manifestações do grau, atribuído a conteúdos referenciais ( *dez minutinhos* ,  *muita besteira* ,  *tanto nome* ,  *tanta coisa engraçada* ), verbal ( *esquecer demais* ) e qualificativos ( *muito chato* ,  *pior* ). Esses usos estão associados a valores semânticos específicos: intensivo ( *muito chato* ,  *pior* ,  *esquecer demais* ) e quantitativo ( *dez minutinhos* ,  *muita besteira* ,  *tanto nome* ). Outros valores semânticos, conforme Silva (2014), são: dimensivo ( *pedras grandes* ,  *cicatriz muito grossa* ), hierárquico/posicional ( *gente de alto nível* ), avaliativo (Foi um  *jogão* ), e afetivo ( *minha*  fica fazendo guerra). Alguns desses tipos podem ocorrer de forma sobreposta.

Nos processos de graduação, podemos observar a relação icônica entre conteúdo e expressão em termos do subprincípio de quantidade, como em:

- (2) (...) É você olhar no espelho  
Se sentir um  *grandessíssimo*  idiota...  
(Seixas, 1973).

A ideia codificada por *idiota* é intensificada pelo adjetivo *grande*. Essa intensificação é reforçada pelo sufixo *-íssimo*, duplicado em *grandesíssimo*. Assim, o aumento da forma corresponde ao encarecimento do conteúdo veiculado. Esse acréscimo de forma e de conteúdo concorre para tornar a expressão *um grandesíssimo idiota* mais marcada que *um grande idiota*, tanto em termos estruturais (inclusão e redobro do sufixo) quanto em termos semântico-cognitivos (incremento da intensidade).

Em relação à informatividade, ao graduar um dado conteúdo, seja ele referencial, eventual, qualificativo ou circunstancial, o falante lhe agrega um traço conceitual, fornecendo, assim, um detalhe (ou uma informação) a mais sobre esse conteúdo. É o que se pode ver no excerto a seguir.

- (3) Depois de sagrar-se a artista que *mais vendeu* discos no ano (1 milhão, *mais do que* Ivete Sangalo e Luan Santana juntos), a cantora sertaneja Paula Fernandes, 27, pôde, enfim, comprar uma casa onde fincará raízes com o pai, ex-lavrador, e a mãe, dona de casa. “Mudamos 24 vezes. Sempre que o aluguel *aumentava*, tínhamos de ir para outro lugar”, diz a mineira, que foi alçada ao sucesso depois de um – ela não gosta de falar – *namorico* com Victor, da dupla com Leo, e de uma paquera – *gosta menos ainda* – com Roberto Carlos, seu par no comentado dueto em seu último show na Globo. Os *pernões* e a *cinturinha* tiveram algo a ver com o *estrandoso sucesso*? “As pessoas gostam do meu *jeitinho* brejeiro”, informa. (*Veja*, 17/08/2011, p. 105).

Nessa amostra, *mais* e *mais do que* representam um acréscimo informacional em relação a *vendeu*; *aumentava* dá a entender que o *aluguel* ficava *mais caro*; *-ico* informa que foi um *namoro de pouco tempo*; *menos ainda* acrescenta sentido a *gosta*; *-ões* e *-inha* são um detalhe a mais a *perna* e *cintura*, respectivamente; *estrandoso* agrega uma característica intensiva a *sucesso*; *-inho* é um “comentário” afetivo acrescentado a *jeito*. Portanto, em todos esses casos, as propriedades graduadoras adicionaram informação por menorizada aos conteúdos a que se vinculam, conferindo-lhes contornos notionais mais específicos.

Na atribuição de grau estão implicados aspectos cognitivos, notadamente relacionados a transferências conceituais de domínios ancorados na experiência física/objetiva para o domínio de conteúdos mais abstratos e subjetivos ou ainda dentro de um mesmo domínio. Essas transferências conceituais se dão por processos metafóricos e metonímicos. Analisemos as ocorrências a seguir.

- (4) ... mas à frente teria *uma mata* densa... *bem mais alta*... uma floresta...  
(*Corpus D&G/Natal*, p. 122).
- (5) ... é a única maneira que a justiça tem para livrar a nação de bandidos de *alta periculosidade*... (*Corpus D&G/Natal*, p. 387).
- (6) ... os velhos estão *morrendo*... assim... assim... está morrendo por quê? muitos deles ficam em fila de aposentado... (*Corpus D&G/RJ*, p. 6);
- (7) ... aí ele puxou a irmã dela... a irmã dela *morrendo de medo*... (*Corpus D&G/RJ*, p. 42).

Em (4), o grau dimensivo em *alta* inscreve-se num domínio conceitual mais ancorado na experiência concreta, tendo em vista exprimir uma propriedade atribuída a uma entidade cuja concepção é interpretada como tangível (a *mata*), podendo, assim, ser perceptualmente mensurada (a altura das árvores). Porém, em (5), a noção de *alto* projetou-se para o domínio conceitual abstrato, referindo-se à intensidade do referente mencionado (*periculosidade*), significando, relativamente, *perigosíssimos*.

Tal projeção se deve ao fato de *alto* possuir, em seu conteúdo básico, as noções semânticas de *grandeza* (i.e., tamanho extenso em termos verticais) e *posição/nível acima do normal*. Essa base conceitual torna possível sua projeção em algo em *nível elevado/excessivo* (a *periculosidade*, em (5)). Essas projeções metafóricas fundamentam-se no esquema conceitual INTENSIDADE É EXTENSÃO/ POSIÇÃO VERTICAL, possivelmente decorrente de um esquema cognitivo mais básico: INTENSIDADE É QUANTIDADE (Lakoff, 1987; Taylor, 1992).

Consideremos, agora, a atuação de projeções metonímico-metafóricas do grau em relação aos conceitos de *morrer* e de *dor*, respectivamente.

Observemos que, em (6), *morrendo* significa *deixar de viver/falecer*, portanto, em acepção denotativamente biofísica, mais “concreta”. Já em (7), *morrendo* é utilizado como elemento intensificador/hiperbólico de *medo* (*com medo excessivo*), estando, assim, em sentido figurado e abstrato. Essa abstratização emerge da contiguidade semântica entre causa e efeito: a morte é uma experiência biofísica negativa, resultante de algum fator causal suficientemente agressivo/forte para provocá-la.

Analisando usos intensivos de estruturas do tipo *morto de cansado*, *cansou de esbravejar*, *explodir de tão gorda* e *se acabar de tanto rir*, ilustradas de (8) a (11), Silva e Bispo (2021) identificaram um padrão construcional que licencia tais casos.

- (8) Suado e **morto de cansado**, Wellington chega à Assembleia de bicicleta. (Disponível em: <https://blogdominard.com.br/2019/02/suado-e-morto-de-cansado-wellington-chega-assembleia-de-bicicleta/>. Acesso em: 18 dez. 2019).
- (9) Tentou sair da cova e não conseguiu. Levantou o olhar para cima e viu que sozinho não conseguiria sair. Gritou. Ninguém atendeu. Gritou mais forte. Ninguém veio. **Enrouqueceu de gritar, cansou de esbravejar**, desistiu com a noite. (Disponível em: <https://www.recantodasletras.com.br/fabulas/4327185>. Acesso em: 18 dez. 2019).
- (10) Dona Redonda se despede do remake “Saramandaia” depois de **explodir de tão gorda!** Vera Holtz costumava levar 4 horas para se caracterizar como a personagem e aparentar 250kg. (Disponível em: [https://www.purepeople.com.br/noticia/vera-holtz-da-adeus-a-silhueta-de-dona-redonda-veja-quem-ja-engordou-pela-arte\\_a6117/1](https://www.purepeople.com.br/noticia/vera-holtz-da-adeus-a-silhueta-de-dona-redonda-veja-quem-ja-engordou-pela-arte_a6117/1). Acesso em: 18 dez. 2019).
- (11) e conta até com uma introdução de Adam Scott, que interpreta Trevor no seriado. Você pode assistir os erros de gravação no vídeo destacado acima – e pode também **se acabar de tanto rir** (tenho certeza que vai). (Disponível em: <https://www.corpusdoportugues.org/now/>. Acesso em: 17 jun. 2021).

Nessas ocorrências, são intensificadas as seguintes noções: *cansado*, em (8); *gritar* e *esbravejar*, em (9); *gorda*, em (10); *rir*, em (11). A intensidade dessas noções é codificada, respectivamente, pelos termos *morto*, *enrouqueceu*, *cansou*, *tão* e *tanto*. Em (8) e (9), a mobilização dos termos para veicular a noção intensiva se deve ao fato de expressarem um estado (quase) extremo resultante do conteúdo intensificado. Assim é que, em (8), a ideia de morte representa uma consequência máxima em relação a medo; a rouquidão e o cansaço, em (9), são tomados como resultado excessivo das ações de gritar e esbravejar. Nos casos de (10) e (11), a noção intensiva é reforçada pelas expressões de valor resultativo *explodir* e *se acabar*.

Esses casos instanciam um padrão esquemático que pode ser formalizado como  $[X_{\text{RESULT}} \text{ de (Adv) } Y_{\text{CAUSA}}]_{\text{INTENS}}$ , segundo sustentam Silva e Bispo (2021). As marcações com subscritos indicam que o primeiro elemento ( $X_{\text{RESULT}}$ ) codifica o resultado do que é expresso em Y, a causa ( $Y_{\text{CAUSA}}$ ). O elemento adverbial (Adv) entre parênteses assinala que esse termo pode ou não ocorrer. Considerando as ocorrências de (8) a (11), temos o seguinte: morte como resultado de intenso cansaço (8); rouquidão decorrente de grito e cansaço resultante de esbravejamento (9); explosão como consequência de muita gordura (10); acabamento (destruição) resultante de muito riso (11). Também, nesses casos, intensifica-se determinado conteúdo (no caso, presente em Y), seja via resultado expresso em X (*morto*, *enrouqueceu* e *cansou*, para os casos (8-9)), seja por meio de elemento adverbial (*tão* e *tanto*, em (10-11)), reforçado pelo estado resultante codificado em X (*explodir* e *se acabar*).

Levando em conta os elementos que constituem a construção intensiva  $[X_{\text{RESULT}} \text{ de (Adv) } Y_{\text{CAUSA}}]_{\text{INTENS}}$ , notamos que existe apenas um elemento fixo, no caso a preposição *de*. Há um termo opcional, identificado pela categoria advérbio (Adv), no caso, *tanto(a)/tão*, que assinala a intensidade do termo subsequente, e dois *slots* (X e Y), os quais podem ser preenchidos por certa variedade de itens. A presença ou não de Adv distingue dois padrões subesquemáticos:  $[X_{\text{RESULT}} \text{ de } Y_{\text{CAUSA}}]_{\text{INTENS}}$  e  $[X_{\text{RESULT}} \text{ de Adv } Y_{\text{CAUSA}}]_{\text{INTENS}}$ . Dada a possibilidade vária de preenchimento dos *slots*, esses subesquemas podem licenciar distintos padrões microestruturais, conforme a natureza dos itens a ocupar tais *slots*: sintagma verbal (SV) –

*explodir de tão gorda* – ou sintagma adjetival (SA) – *morto de cansado* – na posição X; sintagma nominal (SN), SV ou SA – *morrendo de medo, enrouqueceu de (tanto) gritar* ou *morto de cansado*, respectivamente – na posição Y. Desse modo, atesta-se o caráter esquemático dessa construção, devido à existência de apenas um elemento fixo, e também sua produtividade, no sentido de licenciar dois padrões subesquemáticos e uma variedade de microconstruções, em função da diversidade de elementos que podem preencher seus *slots* (X e Y). Sob essa ótica, a construção  $[X_{\text{RESULT}} \text{ de (Adv) } Y_{\text{CAUSA}}]_{\text{INTENS}}$  pode ser considerada parcialmente especificada, nos termos de Goldberg (2003) e Fried (2015).

A combinação dos elementos que integram essa construção forma um todo semântico-sintático cujo significado não resulta plenamente dos significados de suas partes componentes. Ainda que a presença opcional de advérbio indicie o valor intensivo da construção, a consideração do significado dos demais constituintes não leva necessariamente ao entendimento geral da noção intensiva veiculada.

Em (11), por exemplo, o encarecimento à ideia expressa pelo verbo *rir* resulta, em parte, da contribuição do significado do advérbio *tanto*; e, em parte, do sentido de excesso construído intersubjetivamente por meio da explicitação de uma consequência exagerada (possibilidade de o interlocutor *se acabar*). Já em (8), o valor intensivo aplicado a determinado sentimento (*cansaço*, no caso) não é resultado do cálculo dos significados dos itens que integram a ocorrência *morto de cansado*, mas é construído, na interação, por meio da associação entre um estado extremo resultante (*morte*) e aquilo que, em excesso, pode ser (tomado como) sua causa (*cansaço*). Assim sendo, podemos considerar que o padrão  $[X_{\text{RESULT}} \text{ de Adv } Y_{\text{CAUSA}}]$  apresenta baixo grau de composicionalidade, nos termos de Traugott e Trousdale (2013), e que o padrão  $[X_{\text{RESULT}} \text{ de } Y_{\text{CAUSA}}]$  é totalmente não composicional. Isso porque o primeiro contém ao menos um marcador explícito de intensidade – o Adv –, o que não ocorre com o segundo.



## Considerações finais

Neste capítulo, apresentamos a abordagem denominada *Linguística Funcional Centrada no Uso*. Para tanto, traçamos, inicialmente, o percurso histórico desse modelo, discutindo a oposição entre Linguística Formal e Linguística Funcional e a união entre esta última e a Linguística Cognitiva, conforme a proposta de Tomasello (1998). Mostramos que a LFCU resulta da articulação entre a Linguística Funcional norte-americana e a Gramática de Construções. É um modelo que compreende a língua e a gramática como sistemas em constante movimento, refletido nos processos de variação e de mudança linguísticas.

Descrevemos e ilustramos conceitos básicos da LFCU, como iconicidade, marcação, informatividade e plano discursivo, além de processos cognitivos que permeiam a língua em uso, como categorização, prototipicidade, projeções metafóricas e metonímicas. Com foco na construção, unidade básica da língua, discorremos sobre as propriedades de composicionalidade, esquematicidade e produtividade. Ademais, caracterizamos, em linhas gerais, os procedimentos metodológicos norteadores das análises que seguem esse enquadre teórico.

Por fim, aplicamos os princípios e conceitos operacionais tratados neste capítulo à análise do grau no português do Brasil, considerando motivações discursivo-pragmáticas e semântico-cognitivas implicadas nas manifestações desse fenômeno linguístico.

Para estudos complementares sobre a LFCU, sugerimos a leitura de Cunha, Bispo e Silva (2013, 2018), Cunha e Bispo (2013), Cunha, Costa e Cezario (2015) e Rosário e Oliveira (2016).

## Referências

BISPO, E. B.; CUNHA, Maria Angélica Furtado da. A subordinação adjetiva. In: CASTILHO, Ataliba Teixeira de. **História do Português Brasileiro: Mudança sintática das construções: perspectiva funcionalista**. São Paulo: Contexto, 2019. v.5, p. 132-169.

BYBEE, J. From usage to grammar: the mind's response to repetition. **Language**, v. 82, p. 711-733, 2006. Disponível em: < <https://www.jstor.org/stable/4490266> >. Acesso em: 06 maio 2021.

- BYBEE, J. **Language change**. Cambridge: Cambridge University Press, 2015.
- BYBEE, J. **Língua, uso e cognição**. Tradução de Maria Angélica Furtado da Cunha. São Paulo: Cortez, 2016 [2010].
- CHAFE, W. Givenness, contrastiveness, definiteness, subjects and point of view. In: LI, C. (Ed.). **Subject and topic**. New York: Academic Press, 1976. p. 25-55.
- CHOMSKY, N. **Syntactic structures**. Mouton: The Hague, 1957.
- CHOMSKY, N. A review of B. Skinner's "Verbal Behavior". **Language**, v. 35, p. 26-58, 1959.
- CROFT, W. **Radical construction grammar: syntactic theory in typological perspective**. Oxford: Oxford University Press, 2001.
- CUNHA, M. A. F. da Transitividade e passiva. **Revista de Estudos da Linguagem**, v. 1, n. 4, p.43-66, 1996.
- CUNHA, M. A. F. da (Org.). **Corpus Discurso & Gramática** – a língua falada e escrita na cidade do Natal. Natal: EDUFRRN, 1998.
- CUNHA, M. A. F. da **Banco Conversacional de Natal**. Natal: EDUFRRN, 2011.
- CUNHA, M. A. F. da; BISPO, E. B. Pressupostos teórico-metodológicos e categorias analíticas da linguística funcional centrada no uso. **Revista do GELNE**, v. 15, n. 1/2, p. 53-78, 2013.
- CUNHA, M. A.F. da; COSTA, M. A.; CEZARIO, M. M. Pressupostos teóricos fundamentais. In: CUNHA, M. A. F. da; OLIVEIRA, M. R.; MARTELOTTA, M. E. (Orgs.). **Linguística Funcional: teoria e prática**. São Paulo: Parábola, 2015. p. 21-47.
- CUNHA, M. A.F. da; BISPO, E. B.; SILVA, J. R. Linguística funcional centrada no uso: conceitos básicos e categorias analíticas. In: CEZARIO, M. M.; CUNHA, M. A.F. da (Org.). **Linguística centrada no uso: uma homenagem a Mário Martelotta**. Rio de Janeiro: Mauad X/FAPERJ, 2013. p. 13-39.
- CUNHA, M. A.F. da; TAVARES, M. A. **Funcionalismo e ensino de gramática**. Natal: EDUFRRN, 2016.
- CUNHA, M. A.F. da; BISPO, E. B.; SILVA, J. R. (Orgs.). **Variação e mudança em perspectiva construcional**. Natal: EDUFRRN, 2018.
- CUNHA M. A. F. da; BISPO, E. B. Pra quem, é bacalhau basta: da opacidade e produtividade das construções. **Solettras**, n. 37, v. 1, p.103-116, 2019.
- DIRVEN, R; FRIED, V. **Functionalism in Linguistics**. Amsterdam: John Benjamins, 1987.

DU BOIS, J. Competing motivations. In: HAIMAN, J. (Ed.). **Iconicity in syntax**. Amsterdam: John Benjamins, 1985. p. 343-366.

GIVÓN, T. **On understanding grammar**. New York: Academic Press, 1979.

GIVÓN, T. **Syntax**: a functional-typological introduction. Amsterdam: John Benjamins, 1984.v.1

GIVÓN, T. **Syntax**: a functional-typological introduction. Amsterdam: John Benjamins, 1990. v.2

GIVÓN, T. **Functionalism and grammar**. Amsterdam: John Benjamins, 1995.

GIVÓN, T. **Syntax**: An introduction. Amsterdam: John Benjamins, 2001. v.1

GIVÓN, T. **A compreensão da gramática**. Tradução de Maria Angélica Furtado da Cunha, Mário Eduardo Martelotta e Filipe Albani. São Paulo: Cortez; Natal: EDUFRRN, 2012 [1979].

FRIED, M. Construction grammar. In: KISS, T.; ALEXIADOU, A. (Eds.). **Syntax - theory and analysis**: an international handbook. Handbooks of Linguistics and Communication Science. Berlin: Mouton de Gruyter, 2015. p. 974-1003.

GOLDBERG, A. **A construction grammar approach to argument structure**. Chicago: University of Chicago Press, 1995.

GOLDBERG, A. Constructions: a new theoretical approach to language. **Trends in Cognitive Sciences**, v. 7, n. 5, p. 219-224, 2003.

GOLDBERG, A. **Constructions at work**: the nature of generalization in language. Oxford: OUP, 2006.

HALLIDAY, M. A. K. **An introduction to Functional Grammar**. London: Edward Arnold, 1985.

HAIMAN, J. **Natural syntax**: iconicity and erosion. Cambridge: Cambridge University Press, 1985.

HOPPER, P.J. Emergent grammar. **Berkeley Linguistic Society**, v. 13, p. 139-157, 1987.

HOPPER, P.J.; THOMPSON, S. A. Transitivity in grammar and discourse. **Language**, v. 56, n. 2, p. 251-299, 1980.

JOVEM, M.; SILVA, J. R. A rede construcional dos adjetivos formados por -vel no português. **Revista Odisséia**, v. 2, p. 3-18, 2017.

KATO, M. A. As formas de funcionalismo na sintaxe. **D.E.L.T.A.**, v. 14, n. esp., p. 145-168, 1998.

- LAKOFF, G. **Women, fire, and dangerous things**: what categories reveal about the mind. Chicago: University of Chicago Press, 1987.
- LAKOFF, G.; JOHNSON, M. **Philosophy in the flesh**. New York: Basic Books, 1999.
- LAKOFF, G.; TURNER, M. **More than cool reason**: a field guide to poetic metaphor. Chicago: University of Chicago Press, 1989.
- LANGACKER, R. **Cognitive grammar**: a basic introduction. New York: Oxford University Press, 2008.
- LEMONS, M. C. L. **Plano discursivo em perspectiva funcional**: mecanismos textual-discursivos e pragmáticos em artigos de opinião. Dissertação (Mestrado em Estudos da Linguagem). UFRN/PPgEL, Natal-RN, 2020.
- MARTELOTTA, M. E. **Figura e fundo**: Uma proposta prática de análise. Rio de Janeiro: UFRJ, 1998. (Reprod.)
- MARTELOTTA, M. E. **Mudança linguística**: uma abordagem baseada no uso. São Paulo: Cortez, 2011.
- MARTELOTTA, M. E.; PALOMANES, R. Linguística cognitiva. In: MARTELOTTA, M. E. (Org.). **Manual de linguística**. São Paulo: Contexto, 2008. p. 177-192.
- MARTELOTTA, M. E.; KENEDY, E. A visão funcionalista da linguagem no século XX. In: CUNHA, M. A. F. da.; OLIVEIRA, M. R.; MARTELOTTA, M. E. (Orgs.). **Linguística Funcional**: teoria e prática. São Paulo: Parábola, 2015. p. 11-20.
- NASCIMENTO, M. Teoria gramatical e “Mecanismos funcionais do uso da língua”. **D.E.L.T.A.**, v. 6, n. 1, p. 83-98, 1990.
- ÖSTMAN, J-O. Construction discourse: A prolegomenon. In: ÖSTMAN, J-O.; FRIED, M. (Eds.). **Construction grammars**: cognitive grounding and theoretical extensions. Amsterdam: John Benjamins, 2005. p. 121-144.
- PRINCE, E. Toward a taxonomy of given-new information. In: COLE, P. (org.). **Radical pragmatics**. New York: Academic Press, 1981.
- ROSARIO, I. C.; OLIVEIRA, M. R. Funcionalismo e abordagem construcional da gramática. **ALFA**, v. 60, n. 2, p. 233-259, 2016.
- ROSCH, E. Natural categories. **Cognitive Psychology**, v. 4, p. 328-350, 1973.
- SCHIFFRIN, D. **Approaches to discourse**. Oxford: Blackwell, 1994.
- SEIXAS, R. Ouro de tolo. In: SEIXAS, R. **Krig-há, Bandolo!** (álbum). Philips Records: São Paulo, 1973.

SILVA, J. R. **O grau em perspectiva**: uma abordagem centrada no uso. São Paulo: Cortez, 2014.

SILVA, J. R.; BISPO, E. B. Morto de inveja: a construção [X<sub>RESULT</sub> de (Adv) Y<sub>CAUSA</sub>]<sub>INTENS.</sub>  
In: BISPO, E. B.; SILVA, J. R.; SOUZA, M. M. **Pesquisas funcionalistas**: da versão clássica à perspectiva centrada no uso – uma homenagem a Maria Angélica Furtado a Cunha. Natal: EDUFRN, 2021. p. 189-234.

TAYLOR, J. R. **Linguistic categorization**: prototypes in linguistic theory. Great Britain: Laredan Paperbacks, 1992.

TOMASELLO, M. (Ed.) **The new psychology of language**. New Jersey: Lawrence Erlbaum, 1998.

TOMASELLO, M. (Ed.) **The new psychology of language**. New Jersey: Lawrence Erlbaum, 2003. v.2

TRAUGOTT, E. C.; TROUSDALE, G. **Constructionalization and constructional changes**. Oxford: Oxford University Press, 2013.

VOTRE, S. J.; NARO, A. J. Mecanismos funcionais do uso da língua. **D.E.L.T.A.**, v. 5, n. 2, p. 169-184, 1989.

## 2. Metodologia da pesquisa sincrônica

Monclar Guimarães Lopes (UFF)  
Ivo da Costa do Rosário (UFF/CNPq/FAPERJ)

### Considerações iniciais

O conceito de sincronia remonta a uma famosa dicotomia saussureana<sup>1</sup>: sincronia *vs* diacronia. À primeira, cabe a descrição da língua em um ponto específico no tempo; à segunda, a descrição da língua através do tempo. Ao priorizar a sincronia, Saussure buscava defender que a língua poderia ser estudada como uma realidade autônoma em seu estágio atual, sem que fosse necessária a análise de seus processos evolutivos, até mesmo porque os dados sincrônicos, via de regra, constituem a única realidade a que os usuários naturais da língua têm acesso.

Essa dicotomia foi de extrema importância para a constituição da Linguística como ciência e foi amplamente empregada na descrição da maior parte das gramáticas a que temos acesso hoje. No entanto, uma vez que os estudos sincrônicos empreendidos sob orientação estruturalista buscavam a descrição do sistema linguístico (*langue*) sem considerar os usos concretos realizados pelos falantes (*parole*), priorizava-se a estabilidade em detrimento da variação e da mudança em curso, dois fenômenos sincronicamente observáveis.

Para o Funcionalismo, a gramática é uma estrutura emergente (Hopper, 1991), constantemente suscetível à variação e à mudança pelas vicissitudes do discurso. Por isso, o estudo sincrônico da variação e da mudança é altamente relevante, de modo que boa parte das pesquisas linguísticas desenvolvidas nesse modelo buscam descrever o que comumente chamamos de gradiência linguística. Esse fenômeno, por sua vez, é verificável quando um elemento ou

---

<sup>1</sup> cf. Saussure, 2008 [1996].

uma sequência de elementos é interpretada como uma estrutura polifuncional e/ou polissêmica.

Uma rápida ilustração desse fenômeno pode ser dada pela sequência *com isso* no português. Quanto à polifuncionalidade, pode ser a) complemento oblíquo: (01) *eu cisme*ei* **com isso***; b) adjunto adverbial: (02) *eu escrevi a carta **com isso***; c) conector: (03) *estamos com concurso previsto para delegado e policial. **Com isso**, nós começamos a recompor a capacidade de esses policiais de operar*; (04) *O piloto foi rápido nos primeiros giros e, com isso, não demorou a atacar seus adversários*. Quanto à polissemia, *com isso* pode ter ideia de tema em (01), instrumento em (02), conclusão em (03), consequência em (04), entre outros valores semânticos.

As conclusões a que chegamos na descrição da polifuncionalidade e da polissemia de *com isso*, parcialmente exploradas acima, só foram possíveis em virtude de uma metodologia de análise adequada, que busca descrever a língua a partir de uma extensa quantidade de dados empíricos do uso linguístico. Neste capítulo, veremos como empregar essa metodologia – conhecida como método misto ou quali-quantitativo (cf. Lacerda, 2016) – na descrição das construções linguísticas<sup>2</sup>.

## Tema central do capítulo

O método misto – ou quali-quantitativo – é empregado em diversas outras ciências além da Linguística. Caracteriza-se pelo “equacionamento entre a metodologia qualitativa e a quantitativa” (Lacerda, 2016, p. 85). No campo da Linguística, o aspecto qualitativo se manifesta na análise interpretativa de cada ocorrência da construção em estudo, submetida a uma série de critérios previamente estabelecidos, que recobrem propriedades tanto formais (fonológicas e morfossintáticas) quanto funcionais (semânticas, pragmáticas e discursivo-funcionais). Por sua vez, o aspecto quantitativo, de

---

<sup>2</sup> Vale aqui uma importante observação: o método misto é bastante caro à pesquisa funcionalista sincrônica, mas não é a única metodologia possível. Ele é facilmente articulável com outros métodos, alguns deles descritos em outros capítulos desta obra, como, por exemplo, a análise colostrucional, a investigação da construcionalidade, o uso de softwares estatísticos, a abordagem experimental etc.

um lado, manifesta-se no levantamento da produtividade dessas construções em termos de frequência *type* e *token*<sup>3</sup> (cf. Bybee e Thompson, 1997); de outro, na possibilidade de distinguir o que é propriedade intrínseca da construção – isto é, generalizável, já que é observada extensivamente nos dados – daquilo que é uma característica idiossincrática de um ou mais dados particulares.

É importante frisar que o emprego do método misto deve levar em consideração a natureza do objeto e do *corpus* (ou dos *corpora*) selecionado. Por esse motivo, é imprescindível que, antes de aplicá-lo, o analista leve em consideração uma série de passos/procedimentos, que exploraremos mais detidamente nesta seção, a saber: a) escolha do objeto de pesquisa; b) seleção ou constituição de *corpus* (ou *corpora*); c) revisão de literatura; d) análise piloto; e) levantamento das hipóteses e planejamento dos fatores de análise; f) análise sistemática, tabulação das ocorrências e apresentação dos dados.

#### a) Escolha do objeto de pesquisa

As pesquisas em gramática raramente partem do zero. Comumente, temos como ponto de partida estudos que foram desenvolvidos por outros pesquisadores, mesmo que em uma abordagem distinta da Linguística Funcional Centrada no Uso. Sendo assim, uma vez que uma pesquisa deve ser pensada a partir de um problema concreto de pesquisa, deve-se priorizar objetos<sup>4</sup> que:

- *Contradigam, em algum nível, as descrições gramaticais vigentes:*

**Exemplo:** as gramáticas do português dizem que alguns adjetivos, quando antecedem o substantivo, assumem valor mais abstrato, conotativo. Sendo assim, o adjetivo *grande*, quando antecede um

---

<sup>3</sup> Por frequência *type*, compreendemos os diferentes padrões que instanciam uma construção mais esquemática. Por exemplo: no português, há uma construção quantificadora formada pelo subesquema [um N de] (cf. Alonso, Oliveira e Fumaux, 2019), que apresenta *types* variados, como: *um monte de (livros)*, *um bocado de (comida)*, *um mar de (gente)*. Por frequência *token*, entendemos a quantidade concreta de ocorrências de cada *type* no *corpus* investigado.

<sup>4</sup> Neste texto, damos apenas algumas sugestões de objetos que julgamos relevantes à pesquisa. Vale ressaltar, no entanto, que essas categorias não são exaustivas, já que a LFCU é uma abordagem que nos permite descrever quaisquer objetos de natureza gramatical ou lexical.



substantivo, significa “grandioso” (por exemplo, grande homem), mas quando o sucede significa “corpulento” (homem grande). Trata-se de uma regra generalizável? Veja, por exemplo, uma ocorrência que contraria a regra, já que o sentido é o oposto do que prevê a generalização:

(01) Já escreveu mais de 200 músicas, sempre com o Porto no coração ao pé da boca. Rui Veloso, **homem grande** da música portuguesa, não gosta de que lhe chamem mestre nem lhe interessa ser referência. (Disponível em: <https://www.jn.pt/artes/rui-veloso-no-porto-aprendi-tudo-o-que-faco-10362015.html> - Acesso em 15 de maio de 2021).

- *Sejam de uso marginal e/ou ainda não tenham sido descritos (ou pouco descritos) na literatura linguística:*

**Exemplo:** a construção [com isso], brevemente descrita no final das considerações iniciais deste capítulo, é um conector (supra) oracional bastante convencional do português. Na base de dados *Now* do *Corpus do Português*, por exemplo, há 94.911 ocorrências da sequência *com isso*, dentre as quais 59.430 aparecem no contexto da conexão. Não obstante, esse conector não está descrito na literatura linguística nem nos compêndios gramaticais.

- *Apresentem esquemas e subesquemas construcionais que sejam altamente vinculados e pouco composicionais (unidades pré-fabricadas):*

Há uma série de esquemas e subesquemas construcionais que permitem a seleção variada de um ou outro elemento, mas cujo sentido é, em certa medida, pouco composicional (ou não composicional) na língua. A construção correlata aditiva, por exemplo, é formada por um par de correlatores, cujas partes podem ser variáveis. De acordo com Rosário (2012), um de seus *types* mais frequentes é “não só... mas também...”, como o que ocorre no exemplo “ela não só comprou uma moto, mas também um carro novo”. Trata-se de uma construção não totalmente composicional porque o sentido

de adição não advém do significado dos seus elementos constitutivos. Com exceção do “também”, os elementos “não”, “só” e “mas” não veiculam, isoladamente, adição na língua.

### **b) Seleção ou constituição de *corpus* (ou *corpora*)**

Após a escolha do objeto, o analista deve selecionar ou constituir um *corpus*. De um lado, o pesquisador pode escolher um *corpus* de acesso livre já constituído, como, por exemplo, *O Corpus do Português*<sup>5</sup>, *Norma Urbana Linguística Culta*<sup>6</sup>, *Corpus Discurso & Gramática*<sup>7</sup>, *Corpus PEUL*<sup>8</sup> etc. De outro, pode constituir o próprio *corpus* a partir de uma seleção de textos reais que leve em consideração, dentre outros aspectos, a modalidade linguística, os gêneros textuais e o grau de formalidade – ou, ainda, a(s) variedade(s) linguística(s) pretendida(s).

Cabe frisar que o *corpus* escolhido ou constituído está intimamente associado ao tipo de objeto que se pretende estudar. Se o analista busca descrever um conector lógico-semântico ou discursivo-argumentativo, por exemplo, deve selecionar, respectivamente, de modo prioritário, textos de natureza expositiva ou argumentativa. Se busca descrever um marcador conversacional empregado para a manutenção do turno, deve selecionar textos orais formados majoritariamente por sequências dialogais. Ainda, se o fenômeno que estuda apresenta baixa frequência de uso, deve selecionar *corpora* variados ou constituir um *corpus* mais extenso quanto ao número de palavras.

### **c) Revisão de literatura**

Conforme dissemos anteriormente neste texto, dificilmente iniciamos uma pesquisa do zero. Via de regra, os objetos que escolhemos já foram descritos ou mencionados por outros autores, em perspectivas teóri-

---

<sup>5</sup> Disponível em: <<http://www.corpusdoportugues.org>> – Acesso em 13 de maio de 2021.

<sup>6</sup> Disponível em: <<http://fale.ufal.br/projeto/nurcdigital>> - Acesso em 13 de maio de 2021.

<sup>7</sup> Disponível em: <<http://deg.uff.br/corpus-dg/>> - Acesso em 13 de maio de 2021.

<sup>8</sup> Disponível em: <<http://psicod.org/ufrij-projeto-peul.htm>>1 - Acesso em 13 de maio de 2021.

cas variadas. Dessa maneira, é bastante recomendável que o analista inicie seus estudos por uma revisão de literatura ampla e variada sobre o objeto investigado. Uma boa estratégia é partir da tradição gramatical – *o que dizem os gramáticos de referência sobre o objeto em estudo ou sobre a categoria em que o objeto se encontra?* – e depois incluir estudos linguísticos em fontes de referência – como capítulos de livro e artigos publicados em periódicos qualificados<sup>9</sup>, priorizando publicações mais atuais, na medida em que costumam trazer uma visão mais recente sobre o objeto de estudo<sup>10</sup>, além de dissertações de mestrado, teses de doutorado e outras publicações de cunho acadêmico.

Na etapa de elaboração da Revisão da Literatura, é importante que o analista parta dos estudos mais clássicos e tradicionais até as informações mais recentes acerca do tema. Assim, uma apresentação cronológica dos estudos realizados pode ser uma estratégia didática valiosa para os leitores. Entretanto, sem dúvida, vale destacar que essa não é a única forma de organizar esta seção.

Uma vez que a LFCU trabalha com a abordagem construcional da gramática, que busca descrever as propriedades formais (fonológicas e morfossintáticas) e funcionais (semânticas, pragmáticas e discursivo-funcionais) das construções linguísticas, recomendamos que, durante a revisão de literatura, o analista já observe o que os outros pesquisadores disseram a respeito dessas propriedades (ou de parte delas). Isso será bastante útil na fase de análise piloto.

---

<sup>9</sup> No Brasil, a produção científica qualificada é certificada pelo sistema *Qualis* da CAPES. É possível ao analista verificar na internet o ISSN/ISBN do periódico/livro consultado para consulta acerca de sua qualificação. No que tange aos periódicos, recomendamos que as fontes de pesquisa se incluam entre os índices A e B. Em relação aos livros, indicam-se os índices L1 e L2.

<sup>10</sup> Vale ressaltar que “priorizar publicações recentes” não significa excluir as antigas. Há muita referência antiga que continua altamente relevante e, por isso, é recorrentemente citada na literatura linguística. A priorização se dá pelo fato de que há uma tendência natural de as novas pesquisas trazerem contribuições novas para a descrição de um objeto, até mesmo porque também partiram de estudos anteriores – isto é, também fizeram uma boa revisão de literatura. Além disso, há obras e autores clássicos em determinadas áreas que necessitam ser citados.

#### **d) Análise piloto**

A análise piloto é constituída pela seleção e interpretação de uma quantidade inicial de dados. O objetivo principal dessa etapa da pesquisa é o de permitir ao analista aventar hipóteses e estabelecer os critérios de análise, com base em duas leituras: (i) a relação entre as propriedades levantadas na revisão de literatura e os dados empíricos de uso, isto é, as conclusões a que os outros estudiosos chegaram sobre o objeto são pertinentes ao uso real da língua? São generalizáveis, a ponto de captar os diferentes usos? Tais conclusões deram conta de captar, de maneira holística, os aspectos formais (fonológicos e morfossintáticos) e funcionais (semânticos, pragmáticos e discursivo-funcionais) das construções? (ii) a relação entre os dados empíricos do uso linguístico e as generalizações teóricas de base cognitivo-funcional, isto é, em que medida as investigações cognitivistas e funcionalistas nos ajudam a escrever a polissemia e/ou a polifuncionalidade do objeto em estudo?

Sobre essa última leitura, cabe ressaltar que os estudos funcionalistas preveem algumas generalizações extensivamente observáveis nas línguas naturais em termos de gradiência semântica e/ou funcional. Isso se dá em virtude da compreensão de como atuam os mecanismos cognitivos que impulsionam a variação e a mudança linguísticas, conhecidos na literatura da LFCU como neoanálise e analogização.

Antes de falarmos diretamente desses dois mecanismos, cabe ressaltar que compreendemos a variação e a mudança como resultado do uso linguístico. Há uma forte relação entre pragmática, rotinização dos usos e reconfiguração semântica e funcional (cf. Bybee, 2010; Diessel, 2019). Sob esse ponto de vista, quando uma construção é empregada em um novo contexto (não previsto), há uma pressão pragmática que impulsiona o pensamento metafórico ou metonímico, até mesmo porque os seres humanos tendem a cooperar comunicativamente uns com os outros, buscando atribuir significado àquilo que é dito pelos falantes (cf. Grice, 1982).

O pensamento metafórico é compreendido pela transferência entre um domínio mais concreto para outro mais abstrato, como ocorre,

por exemplo, na palavra “cabeça” na frase “ele é o cabeça do grupo”, cujo novo sentido emerge da relação entre a função e importância da cabeça e do corpo, bem como entre a função e a importância de um líder para um grupo. O raciocínio metonímico nasce de uma relação de contiguidade, mais especificamente, da interação entre elementos em um mesmo contexto linguístico. O elemento “mas” contido no conector complexo “mas também”, por exemplo, embora originalmente adversativo, assume valor aditivo em interação com o “também”, um elemento de função aditiva – e.g.: *ele comprou legumes, mas também comprou frutas*. Sob esse ponto de vista, a recorrência de elementos em contextos não prototípicos afeta sua representação na rede linguística e, conseqüentemente, acarreta sua recategorização.

A neoanálise (classicamente conhecida como reanálise) é vista como uma série de micropassos que levam à reconfiguração semântica e/ou funcional. Inicialmente, o ouvinte percebe a existência de um *mismatch* (um “desencontro”) entre os sentidos originais da construção e seu uso efetivo em um contexto atípico (cf. Traugott e Trousdale, 2013; Diewald, 2012), o que implica uma nova interpretação local para a atribuição de sentidos. Caso esses novos contextos continuem emergindo na língua, suscitando sentidos análogos, podem resultar na convencionalização, isto é, na associação de um novo sentido à forma na rede linguística. Ainda, quando esse novo sentido assume funções mais procedurais, essa representação afeta também a categorização gramatical. Contextos recorrentes de uso que nos levam a interpretar *com isso* como uma relação consecutiva e conclusiva – como mostramos nas considerações iniciais deste texto, por exemplo –, levam-nos também a categorizar essa construção como um conector do português.

A analogização é vista como o emprego de uma regra dedutiva. De um lado, pode ocorrer por meio da atração de elementos funcionais análogos. Como exemplo, podemos citar o verbo *desaparecer*, que, embora originalmente intransitivo, começou a ser empregado em construções transitivas – *Para ajudar Aécio, mídia desaparece com aeroporto dado a sua família*<sup>11</sup>.

---

<sup>11</sup> Dado disponível em: <<https://www.esmaelmorais.com.br/2014/07/para-ajudar-aecio-midia-desaparece-com-o-aeroporto-dado-a-sua-familia/>>. Acesso em 10 de junho de 2021.

Os estudos diacrônicos (cf. Lopes, 2015) sugerem que esse uso surgiu no século XIX e que, no século XX, atraiu, por semelhança, o verbo *sumir* – por exemplo, *para eu não viajar, os meninos sumiram com as chaves do carro* (cf. Lopes; Menezes, 2018). De outro, a analogização pode ocorrer por meio da produtividade dos esquemas construcionais. Como exemplo, podemos citar o esquema [Xmente] (cf. Campos, Cezario e Alonso, 2017), que serve como uma regra abstrata para a criação de advérbios de modo a partir de bases adjetivas: *normalmente, cotidianamente, sabiamente, interessantemente* etc.

Dito isso, na análise piloto, a interpretação dos dados com base nos aspectos apreendidos de uma boa revisão de literatura e das generalizações funcionalistas deve ser capaz de nos guiar tanto para a formulação das hipóteses quanto na seleção dos fatores de análise que serão empregados

### e) Levantamento das hipóteses e planejamento dos fatores de análise

Na pesquisa sincrônica em LFCU, as hipóteses podem ser vistas como uma previsão de generalização dos usos e das propriedades formais e funcionais das construções linguísticas, formuladas na análise piloto. Como ilustração, formulamos uma hipótese sobre o conector [com isso] a partir do dado abaixo:

- (02) A primeira prova do fim de semana foi disputada no fim da tarde de sábado. Sérgio, após uma largada, pulou para o sexto lugar, mas, ainda, na primeira curva, voltou ao sétimo posto, ao evitar um toque no concorrente à sua frente. Em uma prova muito bem planejada, o piloto foi rápido nos primeiros giros e, **com isso**, não demorou a atacar seus adversários. (Disponível em: <[http://www.jornaldasavassi.com.br/ver\\_noticia/7211-Sergio-sette-leva-Brasil-ao-alto-do-podio-na-austria.html](http://www.jornaldasavassi.com.br/ver_noticia/7211-Sergio-sette-leva-Brasil-ao-alto-do-podio-na-austria.html)> - Acesso em 10 de outubro de 2020.)

Em (02), *com isso* atua como um conector e estabelece tanto a coesão sequencial (uma relação de consequência) quanto a coesão referencial (já que o pronome demonstrativo *isso* encapsula a oração anterior, que está sublinhada no texto). Hipótese 1: [com isso]<sub>conector</sub> é um mecanismo de coesão híbrida, já que atua tanto na coesão referencial quanto na sequencial.

Tal hipótese, por sua vez, permite-nos o planejamento de um dos fatores da análise, a saber: verificar, em cada ocorrência, qual é o processo coesivo estabelecido pelo conector, se referencial, sequencial, ou ainda, híbrido.

#### **f) Análise sistemática, tabulação das ocorrências e apresentação dos dados**

Depois de selecionados os fatores de análise, o analista deve proceder a uma investigação mais extensiva de dados, com o objetivo de confirmar (ou não) suas hipóteses iniciais de pesquisa. Nessa etapa, ele deve submeter cada ocorrência a todos os fatores, de forma sistemática. É importante também que construa um sistema para a tabulação desses resultados. Na análise de [com isso]<sub>conector</sub>, por exemplo, Lopes e Moura (2021) viram que a coesão híbrida é uma propriedade observável em todas as ocorrências em que *com isso* atua no domínio da conexão.

Analisadas todas as ocorrências, cabe ao analista organizar o modo como apresentará seus resultados à comunidade científica. Isso costuma ocorrer por meio de gêneros acadêmicos variados – tanto orais quanto escritos – em que o pesquisador apresenta o objeto de estudo, suas hipóteses e objetivos, o tratamento teórico-metodológico e seus resultados. Sobre esse último aspecto, cabe ressaltar que ele é o mais importante e, via de regra, o mais extenso e complexo. Por isso, além de ocupar um maior número de páginas, também costuma dividir-se em subseções variadas, de acordo com os objetivos e hipóteses da pesquisa. Como exemplo, podemos citar, mais uma vez, a construção [com isso]<sub>conector</sub>. Das quatro subseções dedicadas à análise, Lopes e Moura (2021) destinam uma delas à coesão híbrida, uma propriedade intrínseca à construção, dada a sua relevância para a caracterização do objeto.

#### **Exemplo de aplicação**

No português, *fora isso* atua na conexão paratática. Na base de dados *Now do Corpus do Português*, por exemplo, a sequência *fora isso* apresenta 2.956 ocorrências, dentre as quais 2.024 ocorrências pertencem ao domínio

da conexão – 243 conectores oracionais e 1.781 conectores interperíodo ou interparágrafo<sup>12</sup>.

A despeito de *fora isso* ser bastante frequente na função de conector – o que atesta sua convencionalidade no português –, não encontramos menção a esse tipo de uso nos compêndios gramaticais e na literatura linguística, o que torna sua investigação relevante à descrição linguística do português. Como ilustração, seguem três ocorrências:

- (03) Mas Southgate diz que, embora a premiação vá aumentar ainda mais a relevância da participação inglesa na Rússia, ela não é uma prioridade. – Quero que o Harry seja o artilheiro, mas **fora isso**, a questão do prêmio individual não interessa. (Disponível em: <https://www.terra.com.br/esportes/lance/kane-ou-lukaku-um-deles-levar-a-chuteira-de-ouro-da-copa-2018> - Acesso em 12/04/2021).
- (04) O economista e gestor do Observatório de Informações Municipais, François Bremaeker, afirma que “só municípios com mais de 200 mil habitantes” conseguem, ainda que com dificuldades, recursos próprios. “Você pode ter um município turístico, como Araxá (Alto Paranaíba) e Poços de Caldas (Sul de Minas), que têm uma renda. **Fora isso**, só os municípios com mais de 200 mil habitantes têm renda tributária razoável”, disse o economista – Disponível em: <https://www.otempo.com.br/politica/minas-gerais-tem-330-cidades-que-nao-se-sustentam-sozinhas> - Acesso em 31/03/2021).
- (05) O governo federal alega que, na última década, 3,9 milhões de moradias foram construídas com recursos do Minha Casa Minha Vida, representando investimentos de R\$ 425 bilhões. De fato, o MCMV tem sido praticamente a única iniciativa nesse campo. Mas, na verdade, o programa tem se prestado mais à publicidade do que a uma real política de habitação. Basta recordar o uso do programa pelas campanhas políticas de Lula e Dilma ao Planalto.

---

<sup>12</sup> O levantamento de frequência de *fora isso* na função de conector oracional e supraoracional – interperíodo e interparágrafo – se deu pelo controle das propriedades morfossintáticas no sistema de buscas eletrônicas. O conector é sempre margeado por pausa: antecedido por ponto ou vírgula e seguido de vírgula.



**Fora isso**, o Minha Casa Minha Vida repete os erros do passado, como a construção de grandes conjuntos em áreas isoladas e com infraestrutura deficiente. A qualidade dos projetos também deixa a desejar (Disponível em: <https://oglobo.globo.com/opiniao/os-pesadelos-que-habitam-minha-casa-minha-vida> - Acesso em 12/04/2021).

Nas três ocorrências anteriores, *fora isso* atua, respectivamente, na conexão de orações, períodos e parágrafos. É empregado na parataxe<sup>13</sup> e veicula noções semânticas de contraste em (03) e (04) e de adição em (05), em contextos expositivos ou argumentativos. A noção de contraste é mais composicional e básica, pois herda um sentido mais abstrato do advérbio “fora”, empregado, em alguns contextos de uso, com o valor de exclusão<sup>14</sup>. Em (03), *fora isso* reforça ainda mais o contraste apresentado pelo “mas” – ao qual está justaposto. Imprime mais foco e peso argumentativo à informação contida na oração subsequente: “a questão do prêmio individual não me interessa”. Em (04), a noção de contraste emerge tanto do sentido abstrato de “fora” quanto da oposição semântico-pragmática entre as unidades discursivas conectadas por *fora isso*: municípios pequenos normalmente não se mantêm com recursos próprios. Araxá e Poços de Caldas são uma exceção.

Já em (05), *fora isso* apresenta um valor semântico mais abstrato, na medida em que a semântica de adição não é composicional. A noção de adição emerge do *frame* semântico (cf. Fillmore, 1985; Diessel, 2019), haja vista que as informações veiculadas pelas unidades discursivas conectadas por *fora isso* caminham para uma mesma direção: tanto antes quanto depois do conector, apresentam-se críticas ao projeto habitacional *Minha Casa Minha Vida*.

Um outro aspecto relevante nas ocorrências é o fato de o pronome demonstrativo *isso* atuar no encapsulamento de informação precedente. Na

---

<sup>13</sup> Vale ressaltar que a parataxe tem assumido um papel mais amplo na literatura linguística desde o final do século XX. Antes tida como um sinônimo de *coordenação*, hoje há estudiosos que consideram que esse processo também pode abarcar simultaneamente a *coordenação* e a *justaposição* (cf. Quirk et al., 1985; López Garcia, 1999; Duarte, 2003).

<sup>14</sup> Na revisão de literatura, observamos que Castilho (2014) e Neves (2000) reconhecem a existência de advérbios de exclusão, como *exclusive*, *menos*, *exceto*, *salvante*, *tirante*. No entanto, não mencionam *fora*.

literatura linguística, sabemos que os pronomes demonstrativos podem atuar na dêixis exofórica ou endofórica. Nas ocorrências, *isso* remete endoforicamente a porções precedentes do texto. Por essa razão, podemos argumentar que *fora isso* é um conector que atua nos dois processos coesivos: o advérbio estabelece a coesão sequencial, promovendo uma relação lógico-semântica (contraste ou adição), e o pronome demonstrativo, a coesão referencial, por meio do encapsulamento.

Além do contexto da conexão, *fora isso* também é encontrado em contextos em que *fora* e *isso* não estão vinculados, mas apenas justapostos, cada qual com uma função distinta, como podemos observar em (06), a seguir, em que *fora* é um advérbio de lugar, fazendo referência a países estrangeiros, e *isso* encapsula a oração destacada:

(06) É importante porque vai oferecer a mais investidores uma alternativa, sem se preocupar com a parte tecnológica e como se faz custódia. É uma forma de facilitar e trazer o maior número de investidores. Lá **fora isso** já existe bastante. Aqui eu lamento que tenha tanta restrição regulatória. É um paradoxo, porque enquanto tem pelo menos 1,5 milhão de pessoas cadastradas em exchanges no Brasil, o máximo que um fundo pode investir em criptomoedas é 20% e tem que ser um fundo lá fora. (Disponível em: <https://exame.abril.com.br/mercados/a-pior-fase-do-bitcoin-pode-ter-ficado-para-tras-diz-fernando-ulrich/> - Acesso em 15/04/2021).

A análise piloto nos possibilitou a formulação de hipóteses iniciais, dentre as quais destacamos:

- a) [*fora isso*]<sub>conector</sub> é um conector paratático híbrido, que estabelece coesão sequencial e referencial. O elemento *fora* assume valores semânticos de contraste e adição. O termo *isso* encapsula porções precedentes do texto;
- b) O valor contrastivo é mais básico, dado o significado do advérbio *fora*, que também é empregado abstratamente com o valor

de exclusão em alguns contextos. O valor aditivo é mais abstrato, menos composicional e também menos frequente. Emerge do contexto a partir de relações metonímicas e do *frame* semântico;

- c) [fora isso]<sub>conector</sub> estabelece articulação entre orações, períodos e parágrafos. Há uma relação entre a posição do conector e a porção de texto encapsulada pelo pronome demonstrativo “isso”. Posição oracional costuma envolver um escopo mais reduzido de remissão textual; posição interperíodo, um escopo intermediário; posição interparágrafo, um escopo mais amplo.
- d) [fora isso]<sub>conector</sub> atua em contextos argumentativos e é de natureza intersubjetiva.

Com base na análise piloto, nas hipóteses iniciais e nas generalizações teóricas de base cognitivo-funcional, planejamos os fatores de análise, tendo em mente que o objetivo da pesquisa em LFCU é descrever os diferentes usos das construções linguísticas e buscar generalizações quanto às propriedades da forma e de função:

### **Plano da forma:**

- a) Propriedades fonológicas: [fora isso]<sub>conector</sub> pode ser concebido como um vocábulo fonológico (cf. Câmara Jr., 1975), um *chunk*, sempre margeado por pausa?
- b) Propriedades morfológicas: [fora isso] pode exercer funções morfológicas, além da conexão?
- c) Propriedades sintáticas: em que posições o conector se apresenta no texto? Que unidades estruturais conecta (orações, períodos e parágrafos)? Há uma relação entre a unidade estrutural em que a construção se apresenta e a porção do texto encapsulada?

## Plano da função:

- d) Propriedades semânticas: quais são as funções semânticas de [fora isso]<sub>conector</sub>: contraste? Adição? Ou algum outro?
- e) Propriedades pragmáticas: de que forma os elementos co(n) textuais colaboram para a identificação da polifuncionalidade (conector ou adjunto adverbial) e da polissemia (contraste e adição) da construção [fora isso]<sub>conector</sub>? Quais são as contribuições do *frame* semântico para a interpretação do valor semântico? A construção é de natureza intersubjetiva?
- f) Propriedades discursivo-funcionais: em todas as ocorrências de [fora isso]<sub>conector</sub>, percebemos a promoção de uma coesão tanto referencial quanto sequencial? Ocorre em outras sequências textuais, além da argumentativa?

Em sequência, partimos para a análise sistemática dos dados. Para isso, foi feita uma análise interpretativa de 100 ocorrências<sup>15</sup> da base de dados *Now*, do *Corpus do Português*. Abaixo, após a tabela, apresentamos algumas conclusões quanto às propriedades coesivas e semânticas de [fora isso]<sub>conector</sub>:

**Tabela 1.** Propriedades semânticas e coesivas de [fora isso]<sub>conector</sub>.

	<b>Adição</b>	<b>Contraste</b>
<b>Conector oracional</b>	3	5
<b>Conector interperíodo</b>	33	36
<b>Conector interparágrafo</b>	12	11
<b>TOTAL</b>	48	52

Fonte: Autoria própria.

Como podemos observar, na Tabela 1, é possível chegar a uma generalização quanto à propriedade coesiva de [fora isso]<sub>conector</sub>. No que tange

<sup>15</sup> Agradecemos a Priscilla Vieira de Biasi Cordeiro pelo levantamento e análise dos dados.

aos valores semânticos, embora o valor de contraste seja mais básico, há uma distribuição uniforme entre os valores de contraste e adição<sup>16</sup>. Além disso, a conexão supraoracional – interperíodo e interparágrafo – é a mais frequente, o que mostra que [fora isso]<sub>conector</sub> atua mais no nível do texto. Outra informação importante, que não consta na Tabela, é que, em todas as ocorrências de [fora isso]<sub>conector</sub>, ele desempenha um mecanismo de coesão híbrida, isto é, estabelece um processo coesivo tanto referencial quanto sequencial.

Além das conclusões apresentadas numericamente na Tabela 1, há outras generalizações a que chegamos por meio da análise das 100 ocorrências, dentre as quais, destacamos duas:

a) *Há uma correlação entre a posição do conector e a quantidade de texto encapsulada pelo pronome demonstrativo “isso”:*

Embora não seja uma regra, há uma forte tendência de que o pronome demonstrativo “isso” encapsule orações quando [fora isso]<sub>conector</sub> está em posição interoracional; encapsule períodos quando em posição interperíodo; parágrafos, quando em posição interparágrafo. É o que ocorre, por exemplo, nas ocorrências de (07) a (09), em que, respectivamente, *isso* remete a oração anterior (07), a período anterior (08) e a parágrafo anterior (09):

(07) “É uma grande perda. A tabelinha Pelé-Coutinho fez o Brasil ficar mais conhecido no mundo todo. Tenho certeza que um dia faremos tabelinha no céu. Minhas condolências à família”, disse o Rei do Futebol.

Pepe, outro ex-companheiro de Coutinho, também comentou sobre a morte do ex-companheiro. “Ele estava com problemas, mas **fora isso** continuava sendo aquele moço alegre”. (Disponível em: <http://www.reporterdiario.com.br/noticia/2641223/maior-par->

---

<sup>16</sup> Embora esse não seja o objetivo da pesquisa, essa distribuição pode servir como hipótese para uma nova investigação, de base diacrônica. É provável que, no passado, os primeiros usos no domínio da conexão tenham sido contrastivos e que os usos aditivos tenham se tornado mais frequentes com o tempo. Inclusive, além do conector paratático [fora isso], há o hipotático [fora que], frequentemente empregado em contextos de adição, o que parece evidenciar a dessemantização do advérbio *fora* ao longo do tempo.

[ceiro-da-carreira-de-pele-coutinho-morre-aos-75-anos/](#) - Acesso em 06/04/2021).

(08) O único e isolado momento aonde o texto acerta o alvo acontece quando, sem exageros e usando de artifício visual, vemos os críticos aprisionados em uma embarcação, desesperados para sair. Aqui, percebe-se a intenção com clareza, e a busca pela quebra que é capaz de gerar humor. **Fora isso**, lamentavelmente nada mais conseguiu exprimir qualquer tipo de valor. (Disponível em: <http://observatoriodocinema.uol.com.br/criticas/criticas-de-filmes/2019/03/critica-chorar-de-rir> - Acesso em 05/04/2021).

(09) Os protestos deste domingo podem significar ao final um teste de popularidade de Bolsonaro após quase cinco meses de gestão. Na última semana, pesquisa Ibope encomendada pela Confederação Nacional da Indústria (CNI) apontou que o presidente é aprovado por 35% dos brasileiros. Essa foi a pior avaliação, em 24 anos, do início de mandato de um governo.

**Fora isso**, nos últimos dias, o Executivo teve que enfrentar manifestações com milhares de pessoas nas ruas de todo o país contra os cortes na área da educação. (Disponível em: <https://www.otempo.com.br/politica/manifestantes-se-reunem-na-praca-da-liberdade-em-ato-pro-bolsonaro> - Acesso em 31/03/2021).

b) [*fora isso*]<sub>conector</sub> atua em sequências argumentativas e expositivas:

Na maior parte dos dados, [*fora isso*]<sub>conector</sub> ocorre em sequências argumentativas. No entanto, há também ocorrências em sequências expositivas, como em (10), em que há apenas o relato de um estado-de-coisas, sem a expressão explícita de um ponto de vista:

(10) Segundo a secretaria da comissão, 115 deputados já se inscreveram para falar. Membros têm direito a 15 minutos de discurso e não membros, 10 minutos, tempo que, somado, chega a durar 24 horas. O tempo ainda pode ser alterado, já que são permitidas inscrições até o primeiro orador iniciar suas declarações, o que deve ocorrer na terça-feira da próxima semana. **Fora isso**, há previsão de fala

para líderes, que podem discursar por tempo proporcional ao tamanho de suas bancadas. (Disponível em: <https://www.terra.com.br/noticias/brasil/somado-tempo-de-oradores-na-comissao-da-reforma> - Acesso em 31/03/2021).

Em virtude do espaço de que dispomos, não fizemos, aqui, uma análise exaustiva dos usos e das propriedades construcionais de [fora isso]<sub>conector</sub>. No entanto, acreditamos que a descrição realizada serviu à função pretendida neste capítulo. Para finalizar esta seção, cabe frisar que, uma vez concluída a análise, cabe ao pesquisador elaborar uma apresentação didática de seus dados para a comunidade científica. Para isso, deve partir de seus objetivos e hipóteses de pesquisa. Para a apresentação de [fora isso]<sub>conector</sub>, por exemplo, uma boa sugestão seria a divisão em quatro seções, quais sejam: a) [fora isso]<sub>conector</sub> na relação de contraste; b) [fora isso]<sub>conector</sub> na relação de adição; c) Níveis de articulação de [fora isso]<sub>conector</sub> e sua relação com o escopo remissivo do pronome demonstrativo *isso*; d) propriedades construcionais de [fora isso]<sub>conector</sub>.

### **Considerações finais**

Neste capítulo, buscamos mostrar como empregar o método misto na descrição de dados sincrônicos com base nos pressupostos teóricos da LFCU, em pesquisas voltadas para a investigação da gradiência – isto é, da polissemia e/ou da polifuncionalidade das construções linguísticas. Para isso, mostramos uma série de procedimentos/passos com que o analista deve se preocupar na elaboração de sua pesquisa, a saber: a) escolha do objeto de pesquisa; b) seleção ou constituição de *corpus* (ou *corpora*); c) revisão de literatura; d) análise piloto; e) levantamento das hipóteses e planejamento dos fatores de análise; f) análise sistemática, tabulação das ocorrências e apresentação dos dados.

Como ilustração, apresentamos parte do planejamento e da análise da construção [fora isso]<sub>conector</sub>. A justificativa para a escolha do objeto foi o fato de ainda não estar descrito na literatura linguística, embora se trate de um conector bastante convencionalizado no português, conforme aponta

sua frequência de ocorrência na base de dados *Now do Corpus do Português*. Esperamos que a descrição das etapas da investigação, somada à ilustração de uma pesquisa em desenvolvimento, seja útil ao leitor para o planejamento de seu trabalho.

## Referências bibliográficas

- ALONSO, K. S. B.; OLIVEIRA, D. L.; FUMAUX, N. C. A. Construções binominais quantitativas em perspectiva distintiva: uma análise colostrucional. **Revista Odisseia**, v. 4, n. Esp., p. 173-193, 2019.
- BYBEE, J. **Language, Usage and Cognition**. New York: Cambridge University Press, 2010.
- BYBEE, J; THOMPSON, S. Three Frequency Effects in Syntax. **Berkeley Linguistics Society**, v. 23, p. 378-388, 1997.
- CÂMARA Jr., J. M. **História e Estrutura da Língua Portuguesa**. Rio de Janeiro: Padrão Editora, 1975.
- CAMPOS, J. L.; CEZARIO, M. M.; ALONSO, K. S. B. Formação da construção Xmente. **Revista Delta**, São Paulo, v. 33, no 1, p.133-158, 2017.
- CASTILHO, Ataliba Teixeira de. **Gramática do Português Brasileiro**. São Paulo: Contexto, 2014.
- DIESEL, H. **The Grammar Network**. How linguistic structure is shaped by language use. New York: Cambridge University Press, 2019.
- DIEWALD, Gabriele. A model of relevant types of contexts in grammaticalization. In: Wischer, Ilse & Diewald, Gabriele (eds.). **New reflections on grammaticalization**. Amsterdam/Philadelphia: John Benjamins, 2012. p. 104-120.
- DUARTE, I. Aspectos linguísticos da organização textual. In: MATEUS et al. **Gramática da língua portuguesa**. Lisboa: Caminho, 2003. p. 85-123.
- FILLMORE, C. J. Frame Semantics. In: THE LINGUISTIC Society of Korea (Ed.). **Linguistics in the morning calm**. Seoul? Hanshin, 1982. p. 111-138.
- FILLMORE, C. Frames and the semantics of understanding. **Quaderni di Semantica**, v. 6, p.222-254, 1985.
- GRICE, P. Lógica e conversação. In: DASCAL, M (Org.). **Fundamentos metodológicos da linguística**, v. IV, Campinas, [s.n.], 1982. p.81-103.



HOPPER, J. On some principles of grammaticalization. In: TRAUGOTT, E. C.; HEINE, B. **Approaches to Grammaticalization**. Focus on Theoretical and Methodological Issues. Amsterdam/Philadelphia: John Benjamins, 1991. p. 157-170.

LACERDA, P. F. A. C. O papel do método misto na análise de processos de mudança em uma abordagem construcional: reflexões e propostas. **Revista Linguística/Revista do Programa de Pós-Graduação em Linguística da Universidade Federal do Rio de Janeiro**, Volume Espacial, p. 83-101, 2016.

LOPES, M. G. **Transitivização de *desaparecer* sob uma ótica discursivo-funcional**. 178f. Tese (Doutorado em Estudos de Linguagem) - Universidade Federal Fluminense, Instituto de Letras, 2015.

LOPES, M.G; MENEZES, V. M. C. A formação do subesquema argumental causativo no português brasileiro. **Revista Confluência**, n. 54, p. 90-112, 2018.

LOPES, M. G; MOURA, S. C. As construções conectoras [com isso] e [como se não bastasse (x)] na promoção da coesão híbrida. **Revista Soletras**, n. 41, p. 189-215, 2021.

LÓPEZ GARCIA, A. Relaciones paratáticas y hipotáticas. In: BOSQUE, I; DEMONTE, V (Orgs). **Gramática Descriptiva de la Lengua Española**. Madrid: Espasa, 1999. p. 3508-3547.

QUIRK, R; GREENBAUM, S; LEECH, G; SVARTVIK, J. **A Comprehensive Grammar of English**. Londres: Longman, 1985.

NEVES, M. H. M. **Gramática de Usos**. São Paulo: Editora da Unesp, 2000.

ROSÁRIO, I. C. **Construções correlatas aditivas em perspectiva funcional**. 250f. Tese (Doutorado em Letras) - Universidade Federal Fluminense, Instituto de Letras, 2012.

SAUSSURE, F. **Curso de Linguística Geral**. 30. ed. São Paulo: Cultrix, 2008.

TRAUGOTT, E. C.; TROUSDALE, G. **Constructionalization and Constructional Changes**. New York: Oxford University Press, 2013.

# 3. Metodologia de pesquisa diacrônica

Mariangela Rios de Oliveira (UFF/UFOP/CNPq/Faperj)  
Flávia Saboya da Luz Rosa (UFF)

## Considerações iniciais

Uma das perguntas de pesquisa mais importantes na área dos estudos funcionalistas, que guia grande parte das investigações nessa área, desde os estudos clássicos sobre gramaticalização, é a seguinte: por que as línguas mudam? De que maneira formas de dizer inovadoras ou específicas de um grupo particular, restritas a um ambiente contextual marcado, tornam-se padrões convencionais, forjando novos elementos na gramática das línguas e novos membros de paradigmas gramaticais?

A pesquisa diacrônica nos ajuda a explicar os padrões da estrutura linguística, permitindo compreender como surgem e se convencionalizam. Observamos também que os processos cognitivos que são acionados quando falante e ouvinte se comunicam são as principais causas da mudança, e por isso podemos explicar por que todas as línguas mudam de maneira semelhante.

Conforme destaca Bybee (2015), a mudança linguística é tomada tanto como parte integrante da língua quanto como consequência natural do uso interativo. Assim, mudanças são naturais – nem boas, nem ruins. Essa constatação se distingue de uma certa visão disseminada pelo senso comum, que costuma fazer a defesa de que as mudanças “diminuem” ou destroem a língua. Na verdade, é muito difícil resistir a um padrão que está sendo usado ao nosso redor o tempo todo, portanto, quando uma mudança ganha aderência, é difícil freá-la.

Nesse sentido, de acordo com a linguista estadunidense Bybee (2010; 2015) e, no Brasil, com Martelotta (2011), estudar a língua em perspectiva funcionalista é estar atento à sua natureza dinâmica. Para tanto, é preciso levantar, descrever e analisar contextos de uso com foco em suas motivações distintas, que são de ordem estrutural, cognitiva e pragmático-discursiva,

observando-se também a recorrência desses contextos. Na pesquisa diacrônica de viés funcionalista, portanto, conforme referido no capítulo anterior, também adotamos o método misto, que se define pelo “equacionamento entre a metodologia qualitativa e a quantitativa” (Lacerda, 2016, p. 85). Olhar a trajetória dos usos linguísticos implica levar em conta tanto o ambiente linguístico e extralinguístico em que tais usos ocorrem, quanto considerar a frequência desses contextos e seu impacto para a construcionalização gramatical.

É a partir dessa metodologia mista, por exemplo, que Paula (2021) capta os micropassos de mudança que levaram, no português, padrões adverbiais, como (1) *a casa dele fica **para lá da igreja*** ou (2) *a reunião durou **para lá de 4 horas***, a mudarem para estruturas intensificadoras de quantidade, como em (3) *esse vestido custou **para lá de três mil reais***, e daí para estruturas intensificadoras de grau, como em (4) *a festa foi **para lá de animada***.

Por conta desse olhar mais amplo sobre os usos linguísticos, neste capítulo, de modo complementar ao capítulo 2, vamos nos voltar para a metodologia de pesquisa diacrônica. Nosso objetivo é demonstrar como é possível flagrar, na trajetória histórica dos usos linguísticos, via micropassos de mudança contextual, nos termos de Diewald e Smirnova (2012), e mecanismos de analogização, como definidos em Traugott e Trousdale (2013), os processos metafóricos e metonímicos que conduzem à fixação de padrões convencionais de dizer. Esses processos, replicados no conjunto de usuários da língua, acabam por fixar, nos termos de Traugott (2021), uma nova associação simbólica de forma e conteúdo, ou seja, uma *construcionalização*, um pareamento inédito e inovador que se integra à gramática.

Na pesquisa da construcionalização, devemos estar atentos a mudanças construcionais prévias na trajetória dos usos linguísticos, ou seja, alterações que afetam um dos eixos da construção somente – o da forma ou o do conteúdo, no que nomeamos como *mudanças pré-construcionalização*. De outra parte, na abordagem histórica, nos interessa também tratar das mudanças ocorridas após a construcionalização, quando o novo pareamento convencionalizado se integra a uma categoria da gramática da língua, podendo ocorrer, na sequência, expansão de novos usos a partir do esquema criado, via analogização. Esse segundo tipo de alteração da forma ou do conteúdo é nomeado de *mudança pós-construcionalização*.

## Tema central do capítulo

Neste capítulo, dedicamo-nos ao mapeamento das etapas de mudança de usos linguísticos ao longo do tempo, contemplando a descrição de tipos de contexto de mudança gramatical propostos por pesquisadores da área, em especial, Diewald e Smirnova (2012). Considerando a amplitude semântica encerrada pela palavra *contexto* e suas diferentes interpretações na literatura, faz-se necessário, primeiramente, definir o termo tal como será por nós assumido neste capítulo para, então, apontarmos nossa perspectiva de análise contextual e nos ocuparmos da compatibilização de alguns modelos de estudo propostos.

Definição: Dentre as acepções apresentadas por Crystal (2008, p. 109), assumimos que o termo *contexto* “é usado para se referir a todos os fatores que sistematicamente determinam a forma, o significado ou a adequação dos enunciados”. Tal definição, portanto, abrange o que se tem chamado de *cotexto*: “partes específicas de um enunciado (ou texto) próximas ou adjacentes à unidade<sup>1</sup> que é o foco da atenção” (ibidem, p. 108). Uma visão mais holística compreende ainda o chamado *contexto extralinguístico* ou *situacional*, que se refere a “aspectos do mundo não linguístico relacionados a unidades linguísticas empregadas de modo sistemático” (ibidem, p. 109). *Lato sensu*, contexto situacional abarca todo o “histórico não linguístico de um texto ou enunciado, incluindo a situação imediata em que é usado, a consciência por parte do falante e do ouvinte do que foi dito previamente e qualquer crença ou pressuposição externa relevante” (ibidem, p. 109).

Perspectiva de análise: Assim como declarado por Oliveira (2015), entendemos que o tratamento global da forma e do conteúdo<sup>2</sup> construcional,

---

<sup>1</sup> É importante frisar que, na perspectiva construcionista, a unidade linguística, que é o objeto de análise, pode apresentar dimensões estruturais variadas, desde morfemas a padrões frasais complexos (Goldberg, 1995), como já foi apresentado no capítulo 1 desta obra.

<sup>2</sup> Traugott e Trousdale (2013) utilizam a expressão “*form-meaning pairings*”, que tem sido traduzida por *pareamentos forma-significado*. Há ainda quem defenda a ideia de *forma-função* ou *forma-sentido*. Assim como em Rosa (2019), optamos pelo termo *forma-conteúdo*, por conta de o considerarmos mais abrangente e possibilitar especificações referentes à *função, significado e sentido*, conforme proposta de Oliveira e Arena (2019).

presente nas pesquisas da LFCU<sup>3</sup>, deva ser aplicado nos estudos contextuais da mudança linguística em viés diacrônico. Portanto, com base na estrutura simbólica construcional de Croft (2001), propomos que o *contexto* seja considerado em dupla e correlacionada dimensão: contexto de forma (fonética, morfológica e sintática) e contexto de conteúdo (semântico, pragmático e discursivo).

Modelos de estudo: Segundo Diewald (2006), a noção de *tipos de contexto* é empregada para descrever os sucessivos estágios diacrônicos associados aos processos de mudança gramatical<sup>4</sup>. Sendo assim, uma nova função gramatical não surge de modo homogêneo em todos os seus usos linguísticos, mas sua origem está ligada a contextos específicos. A autora apresenta três tipos de contextos: o atípico, o crítico e o isolado. Heine (2002), ao elencar alguns dos principais fatores associados ao desenvolvimento de novos significados gramaticais, interessa-se, em especial, pelo papel do contexto, distinguindo três tipos: contexto ponte (de transição), contexto de mudança e contexto de convencionalização. Traugott (2021) afirma que seu foco se concentra na importância da rotinização, isto é, na replicação de contextos na mudança. Nesse artigo, a autora apresenta o que considera os dois principais estágios de mudança do objeto analisado, estágio 1 (uso lexical) e estágio 2 (uso discursivo), e opta por omitir os passos intermediários.

Considerando a definição de *contexto* e a perspectiva de análise assumidas neste e em muitos trabalhos do D&G<sup>5</sup> sobre mudança linguística diacrônica, temos adotado em nossa metodologia analítica a proposta tipológica de contextos de Diewald e Smirnova (2012). Contudo, embora tratem do uso lexical em seu artigo, as autoras não se referem propriamente

---

<sup>3</sup> O nome Linguística Funcional Centrada no Uso é uma adaptação do Grupo de Estudos Discurso e Gramática (D&G) da expressão Linguística Baseada no Uso, apresentada por Bybee (2010).

<sup>4</sup> Diewald (2002, 2006) refere-se aos processos de mudança gramatical sob a perspectiva da gramaticalização. Neste capítulo, adotamos o prisma da construcionalização gramatical (Traugott; Trousdale, 2013). Por entendermos que tal abordagem teórica abarca a gramaticalização, tratando-se, portanto, de modelos compatíveis, nos valem da mesma taxonomia dos tipos de contexto propostos pela autora.

<sup>5</sup> O Grupo de Estudos Discurso e Gramática – D&G – tem sedes na UFF, UFRJ e UFRN. Mais informações podem ser obtidas em <https://discursoegramaticablog.wordpress.com>. Acesso em: 18 maio 2023.

a um contexto de origem da mudança. Heine (2002), apesar de utilizar a expressão *significado fonte* e mencionar o estágio inicial em que tal significado se enquadra, tampouco o contabiliza dentre os contextos de alteração linguística por ele descritos. Além disso, Heine (2002) concentra seu olhar para a mudança semântica, sem abranger as questões estruturais. Traugott (2021), sob a ótica global da forma e do conteúdo, inclui o estágio inicial dentre o que considera os dois mais importantes da trajetória de mudança, porém não sistematiza os tipos de contexto intermediários, o que nos leva a adotar prevalentemente a proposta de Diewald e Smirnova (2012) para dar conta de nossas pesquisas.

Por entendermos, assim como Traugott (2021), que o estágio de origem apresenta importantes fatores suscetíveis a relações metonímicas e metafóricas e, portanto, propiciadores de novas formas e/ou novos conteúdos, acrescentamos aos estágios propostos por Diewald e Smirnova (2012) o que chamamos de *contexto fonte*. Desse modo, numa trajetória de mudança de um elemento A (fonte) para um elemento B (alvo), não só apontamos o estado formal e conteudístico de A, como também observamos os fatores contextuais que condicionam tal estado. A relevância do contexto fonte pode ser verificada em Rosa (2019), em que são observadas, desde esse estágio inicial, mudanças pormenorizadas relativas à forma ou ao conteúdo, nomeadas de *nanopassos* de mudança, marcando a gradualidade intracontextual. A esse respeito, corrobora Heine (2002, p. 83):

“[...] o processo de A a B é contínuo, envolvendo uma infinidade de estágios intermediários, e que sem uma compreensão mais abrangente da natureza desses estágios intermediários, as generalizações sobre o processo geral permanecem preliminares”.

No Quadro 1, apresentamos a compatibilização praticada nas pesquisas do D&G, em que se inclui o chamado *contexto fonte* aos contextos de mudança propostos por Diewald e Smirnova (2012), além do refinamento analítico desses contextos desenvolvido por Rosa (2019). Adiante, expomos a descrição de cada um dos estágios de análise adotados em nossa metodologia de pesquisa diacrônica:

**Quadro 1.** Compatibilização e refinamento dos estágios de mudança construcional.

Autores	Estágios de mudança: Micropassos				Estágio de paradigmati- zação
D&G: compatibilização com base em Diewald e Smirnova (2012)	Estágio 1: Contexto Fonte	Estágio 2: Contexto Atípico	Estágio 3: Contexto Crítico	Estágio 4: Contexto Isolado	Estágio 5:  Paradigmati- zação
Rosa (2019): refinamento analítico a partir de Diewald e Smirnova (2012)	<b>Nanopassos</b>				
	Graus F1, F2, F3...	Graus A1, A2, A3...	Graus C1, C2, C3...	Graus I1, I2, I3...	

Fonte: Autoral, baseado em Diewald e Smirnova (2012) e Rosa (2019).

- a) Contexto fonte<sup>6</sup>: Definimos como contexto fonte o estágio em que os elementos linguísticos em foco apresentam maior composicionalidade, ou seja, maior transparência entre os aspectos da forma e do conteúdo. Desse modo, o entendimento do significado de cada item individual decodifica o significado do todo. Nesse estágio, o conteúdo semântico-pragmático está vinculado ao âmbito biossocial, referenciando aquilo que é mais concreto. O contexto fonte é compatível com o estágio 1 apresentado por Traugott (2021) e, *grosso modo*, pode ser também associado ao estágio inicial mencionado por Heine (2002).
- b) Contexto atípico (Diewald, 2002; 2006): Nesse estágio, as precondições do processo para a construcionalização gramatical se desenvolvem. A unidade lexical em questão mostra uma expansão inespecífica de sua distribuição nos contextos em que não havia sido usada anteriormente, isto é, em *contextos atípicos*. Nos contex-

<sup>6</sup> Diewald e Smirnova (2012) consideram o contexto atípico como primeiro estágio de mudança. Teixeira (2015), cuja tese de doutorado tem grande relevância para a pesquisa do D&G sobre marcadores discursivos formados por advérbios locativos, opta por associar o contexto fonte ao que chama de estágio zero.

tos atípicos, observa-se o uso de construções existentes que – por meio de implicações conversacionais – aparecem em combinações incomuns e, ao mesmo tempo, podem ser facilmente interpretadas devido à sua estrutura composicional.

- c) Contexto crítico (Diewald, 2002; 2006): Esse estágio marca o efetivo desencadeamento do processo que resulta em construcionalização gramatical. Ele está relacionado ao surgimento de um tipo muito específico de contexto, que é chamado *crítico*. É caracterizado por múltiplas ambiguidades estruturais e semânticas e, assim, propicia muitas alternativas de interpretação, entre elas o novo significado gramatical. Além disso, esse tipo de contexto tende a desaparecer no desenvolvimento posterior à construcionalização gramatical do elemento linguístico ou porque sua estrutura é perdida (ex.: por expansão para outros contextos) ou porque ela se desenvolve em uma estrutura não ambígua e, então, perde sua característica definidora, isto é, sua ambiguidade múltipla.
- d) Contexto isolado (Diewald, 2002; 2006): Esse estágio mostra a consolidação do processo de construcionalização, em que o novo significado gramatical é isolado, separado do mais antigo, que é mais lexical. Esse desenvolvimento pressupõe a existência de contextos isolados para cada leitura, isto é, contextos linguísticos específicos que favorecem a interpretação de um para a exclusão de outro. As construções em contextos isolados assemelham-se apenas parcialmente a outras construções existentes e apresentam uma correspondência forma-conteúdo única. Tais construções apresentam composicionalidade reduzida, isto é, no mínimo, alguma parte da sua correspondência forma-conteúdo tem de ser tratada holisticamente, e não pode ser derivada em sua totalidade de outras construções ou de uma combinação de outras construções.
- e) Paradigmatização (Diewald; Smirnova, 2012): É o processo de estabilização de um novo signo gramatical que o transforma em parte integral de um paradigma. Isso significa que, enquanto os estágios iniciais descrevem a separação do novo significado da sua



fonte, o último estágio refere-se a um processo em que o novo signo perde o seu *status* independente (sua autonomia) à medida que se associa a outros membros do paradigma, assim como com a categoria paradigmática como um todo. Por um lado, o novo signo construcionalizado passa a ser confrontado com membros em oposição/distinção do mesmo paradigma e, por outro lado, é gradualmente associado a um significado gramatical mais abstrato que serve como um denominador comum para todo o paradigma.

- f) Nanopassos (Rosa, 2019): São mudanças pormenorizadas (referentes à forma ou ao conteúdo) de arranjos linguísticos ou microconstruções enquadrados em subfases mais ou menos avançadas em um mesmo contexto mais amplo, como definido nos moldes de Diewald (2002, 2006). Tais nuances nas mudanças podem estar relacionadas à diminuição da composicionalidade, ao aumento de complexidade sintática, à diminuição de vocábulos fonológicos, ao aumento da abstratização semântico-pragmática, entre outros. Os nanopassos marcam a gradualidade intracontextual.

Na próxima seção, apresentamos exemplos de aplicação analítica baseada na classificação contextual anteriormente descrita. Indicamos alguns dados históricos, em que consta o objeto de pesquisa, de modo a ilustrar cada micro ou nanopasso de mudança a ser examinado. Os *corpora* utilizados para coleta de dados, *Corpus* do Português<sup>7</sup> e *Corpus* Tycho Brahe<sup>8</sup>, contêm textos produzidos entre os séculos XIII e XX.

### **Exemplo de aplicação**

No intuito de ilustrar a aplicação contextual descrita na seção anterior, apresentamos a trajetória de mudança de construções formadas pelo

---

<sup>7</sup> Disponível em: <<https://www.corpusdoportugues.org/>>. Acesso em: 25 de junho de 2021.

<sup>8</sup> Disponível em: <<http://www.tycho.iel.unicamp.br/~tycho/corpus/>>. Acesso em: 25 de junho de 2021.

verbo *esperar*<sup>9</sup> (elemento indutor-refreador<sup>10</sup>) e por advérbio locativo (que ganha status de afixoide<sup>11</sup>), investigadas por Rosa (2019)<sup>12</sup>. Na tabela a seguir, apresentamos a frequência de cada uma das formas encontradas (considerando as variações morfofonêmicas) nos contextos fonte, atípico, crítico e isolado. Conforme afirmado anteriormente, na pesquisa diacrônica de viés funcionalista, adotamos o método misto, equacionando metodologia qualitativa e quantitativa. Contudo, para o cumprimento do objetivo específico deste capítulo, nos concentramos, nas próximas linhas, na análise qualitativa dos dados.

**Tabela 1:** Quantitativo de *espera Loc/Afix* nos contextos de mudança

Contextos	Fonte			Atípico		Crítico		Isolado		Subtotal			Total	
	Séculos	XVIII	XIX	XX	XIX	XX	XIX	XX	XIX	XX	XVIII	XIX		
<i>espera aí</i>	-	10	07	-	05	-	-	-	18	-	10	30	40	56
<i>pera aí</i>	-	-	-	-	-	-	-	-	02	-	-	02	02	
<i>perai</i>	-	-	-	-	-	-	-	-	14	-	-	14	14	
<i>espera aqui</i>	02	04	01	-	-	-	-	-	-	02	04	01	07	
<i>espera lá</i>	-	-	-	04	05	01	-	03	04	-	08	09	17	
Subtotal	02	14	08	04	10	01	-	03	38	02	22	56		
Total	24			14		01		41					80	

Fonte: Rosa, 2019.

<sup>9</sup> Foram coletados arranjos formados pelas formas homônimas do verbo *esperar* na 2ª pessoa do singular do imperativo afirmativo (2ª pessoa do discurso: tu) e na 3ª pessoa do singular do presente do indicativo (2ª pessoa do discurso: você). Essa última forma verbal costuma ser empregada em muitas variedades do português brasileiro como substituta da forma *espere* (imperativo afirmativo, 2ª pessoa do discurso: você).

<sup>10</sup> São considerados *indutor-refreadores* os verbos, as frases nominais ou elementos de outras classes gramaticais que denotem ordem, comando, conselho, sugestão, convite, solicitação ou súplica com subfunção relacionada à contenção de ação, repressão, diminuição de intensidade, abrandamento etc. (Cf. Rosa, 2019)

<sup>11</sup> *Afixoide* é uma categoria gradiente situada entre elementos lexicais e procedurais, de constituição leve no que se refere tanto ao conteúdo (significados mais abstratos) quanto à forma (extensão modesta, que favorece integração morfofonêmica a subpartes nucleares de construções complexas). (Cf. Booij, 2013).

<sup>12</sup> Em Rosa (2019), foram investigados os marcadores discursivos refreador-argumentativos formados por elementos indutor-refreadores e afixoide de origem adverbial locativa, tais como *alto lá*, *calma aí*, *calma lá*, *espera aí*, *espera lá*, *segura aí*, *segura lá*, *aguenta aí*, *aguenta lá* e *para aí*.

- **Contexto fonte de grau 1**

Nos *corpora* de pesquisa, encontramos, desde o século XVIII, registros da sequência *espera Loc* exprimindo o significado mais concreto do verbo e do pronome locativo: pedido/ordem de que se aguarde algo/alguém em certo período de tempo, em determinado lugar. No contexto fonte de grau 1, o verbo *esperar* pode ocorrer em posição adjacente a quaisquer advérbios locativos empregados na língua portuguesa (aqui, cá, aí, ali, lá, acolá), tal como no exemplo a seguir:

- (1) Sahem Xanto, Periandro, e Efopo, que ficará como escondido. / Xanto: Efopo, **espera aqui** detraz desta cortina. /Efopo: He muy boa fala vaga! /Xanto: Amada Euripedes, tardey muito? /Euripedes: Iffo he costume antigo: donde vem a esteas horas, tamanhaõ? / (Tycho Brahe: Teatro cômico português, de António José da Silva, 1759).

Em (1), o personagem Xanto pede a Esopo que fique escondido atrás da cortina, aguardando até segunda ordem: “Efopo, espera aqui detraz desta cortina”. A continuidade da leitura do texto integral permite-nos observar que, de fato, Esopo aguarda durante um tempo a nova instrução, no local em que lhe foi indicado. O locativo *aqui* faz referência catafórica a “detrás desta cortina”, indicando local próximo ao enunciador no ato do discurso.

- **Contexto fonte de grau 2**

A partir do século XIX, observamos o frequente emprego de uma sequência particular em que constam os seguintes componentes assim ordenados: verbo *esperar*, pronome locativo e um quantificador de tempo (QT): *espera Loc QT*. Tais elementos QT são, em geral, representados por advérbios, locuções adverbiais, expressões partitivas com elemento temporal elíptico e até orações que expressem intervalos temporais. Conforme apresentado no exemplo a seguir, ao empregar o arranjo *espera Loc QT*, o enunciador, de modo geral, destaca o curto período de tempo durante o

qual o enunciatário deve aguardar. Nesse contexto, ainda é possível que o verbo ocorra em contiguidade a quaisquer advérbios locativos do português, embora os dados apontem prevalência das formações com *aí*.

- (2) E a pessoa para quem me chamam está em grande risco de vida..  
- Está, sim senhor, se não fosse isso, eu não viria incomodar o senhor padre.. - Nesse caso.. não há remédio senão acudir-lhe.. Mora muito longe a pessoa, a quem tenho de confessar.. - Não senhor; é mesmo na povoação; o senhor padre pode ir a pé; é lá no fim da vila, mas não é muito longe. - Visto isso, filho, **espera aí** um momento para ires comigo, e me guiares até lá. O padre depois de desculpar-se para com suas visitas, informando-as do motivo urgente e indeclinável, que o obrigava a retirar-se, tomou o seu bastão e o seu chapéu triangular, desceu a escada, e saiu em companhia do rapazinho que o viera chamar. (*Corpus do Português: O Seminarista*, de Bernardo Guimarães, séc. XIX).

No contexto fonte de grau 2, apesar de os pronomes locativos ainda indicarem o local em que o interlocutor deve aguardar, a informação por eles expressa não é especificada como ocorre no contexto anterior (Ex. (1): “espera aqui **detrás desta cortina**”), parecendo estar em segundo plano. Conforme exemplificado no trecho (2), em “espera aí **um momento**”, que pode ser parafraseado por “Aguarde onde você se encontra durante breve fração de tempo”, o foco recai sobre o elemento temporal. Tal transferência de foco, que se torna a base para o contexto seguinte, é permitida por conta da semântica do verbo, já que a ação de esperar algo ou alguém se dá no espaço e no tempo. Nesse tipo de contexto, flagram-se os primeiros passos da trajetória de mudança espaço-tempo-discurso.

#### • Contexto atípico de grau 1

Observamos, a partir do século XIX, o emprego da sequência *espera Loc* com maior vinculação sintático-semântica, em que o pronome locativo já não exprime referenciação espacial, passando a exercer papel de afixoide.

No contexto atípico de grau 1, verificamos ocorrência de uma construção a que denominamos *indutora de expectativa sobre atividades do mundo biossocial* (IEB), pois, ao empregá-la, o enunciador induz o enunciatário a aguardar um acontecimento que se tem como certo ou provável em tempo não distante.

- (3) [...] como em um pesadelo, bateu palmas, e deu um grito, antes um uivo horroroso, com os braços em tremor convulsivo estendidos para o céu. Quirino assustado olhou rapidamente para aquele lado; mas depois que reconheceu Jupira: - Está satisfeita? - bradou de longe mostrando a faca ensangüentada, e apontando para o fundo da canoa, onde jazia o cadáver de Carlito estrebuchando e vomitando sangue. - Bravo! bravo. muito bem! gritou a cabocla, com um sorriso de infernal ironia. - Agora venha! venha depressa receber o prêmio.. - **Espera lá** ainda, minha Jupira; preciso dar sepultura a este desgraçado... Falando assim, Quirino desatava da cintura uma forte e comprida cinta de duas voltas, que trazia de propósito, destinada a atar ao pescoço de Carlito a pedra, que pusera na canoa, e atirá-lo ao fundo do rio. (*Corpus* do Português: Histórias e Tradições da Província de Minas Gerais, de Bernardo Guimarães, 1872).

No fragmento (3), embora “Espera lá ainda” corresponda à mesma sequência morfossintática observada no contexto fonte de grau dois – *espera Loc QT* –, já não há referenciação espacial por parte do pronome locativo. A espera da personagem Jupira dar-se-ia no lugar em que se encontrava (do ponto de vista do enunciador: *aí*), e não em local distante de ambos os envolvidos no ato comunicativo (*lá*). Sendo assim, ao dizer “Espera lá ainda, minha Jupira” após o convite para receber sua premiação, o enunciador pede à interlocutora que aguarde mais um pouco para, então, lhe dar a recompensa. *Lá* exerce, portanto, papel de afixoide que, junto ao termo nuclear *espera*, indica futuridade próxima: pedido de breve postergação da prometida retribuição.

Nesse estágio, há alteração de forma, com a vinculação e fixação sintática dos afixoides locativos *aí* e *lá* à subparte *espera*, além de certa

permanência do conteúdo semântico nuclear: “pedido de que se aguarde algo”. No contexto anterior, a contiguidade sintática dos quantificadores de tempo à sequência *espera Loc* e o foco na informação sobre a duração do aguardo de algo ou alguém parecem ter propiciado uma espécie de esvaziamento semântico locativo e incorporação semântica temporal.

- **Contexto atípico de grau 2**

No segundo grau do estágio atípico, verificamos, a partir do século XIX, o uso da construção a que denominamos *indutora de refreamento de atividade do mundo biossocial (IRB)*, pois, ao empregá-la, o enunciador induz seu interlocutor ao refreamento (interrupção, abrandamento, impedimento etc.) de uma ação realizada em âmbito biossocial.

(4) [...] pôs-lhe a mão no braço por segurança. - Largue-me! Por força ninguém me leva. Era o momento em que passavam dois cavaleiros. Um deles ouvindo aquela voz, esbarrou o animal: - O que é isto, Sr. Lucas? - Manuel!... Traído, amigo, traído! O gaúcho reservou a explicação para mais tarde. - Deixem o homem, disse ele para os dois guardas. - E quem é você? - Eu já lhes digo! replicou Manuel passando a mão ao punho da faca. O outro cavaleiro adiantou-se: - **Espera lá**, rapaz. Firmando-se nos estribos e tomando o tom do comando disse para os guardas: - Permanentes, este homem está solto. - O coronel! murmuraram os guardas. Era com efeito Bento Gonçalves que chegava da sua estância. Os guardas se retiraram cabisbaixos. (*Corpus do Português: O Gaúcho*, de José de Alencar, 1870).

Por meio da leitura do fragmento (4), verificamos a instanciação da construção *espera lá* para abrandar os ânimos e, até mesmo, impedir que Manuel, de faca em punho, ferisse um dos guardas. No exemplo (4), o uso de *espera lá* não exprime um pedido de que o interlocutor aguarde em local distante do ato comunicativo, pois não há apontamento espacial. O significado do todo passa a ser uma ordem/pedido de refreamento,

culminando em não ação, de modo que é possível parafrasear o trecho destacado por “Não faça nada, rapaz”.

É interessante pensarmos que *aguardar algo ou alguém* (significado observado nos estágios iniciais) pressupõe, *grosso modo*, certa inatividade ou permanência de estado, até a chegada do que se espera. Sendo assim, por extensão metonímica, a construção do contexto atípico de grau 2 passa a expressar refreamento de atividade do mundo biossocial.

- **Contexto crítico**

O contexto crítico é marcado por ambiguidade, e entre as possibilidades de interpretação está o novo significado. Via de regra, é um contexto pouco produtivo, por conta da fase de transição que representa, tendendo ao desaparecimento após a construcionalização, conforme exposto na seção anterior. Justamente devido à baixa produtividade, não houve verificação de nanopassos intracontextuais.

- (5) Natividade (Muito calmo): - Não sejas tolo... não te quero mal... (Dando-lhe uma nota) Aqui tens cinco bagarotes. José (Admirado): - Não percebo... Natividade: - Solta um grande grito... Assim como se te estivessem matando! José: - Está doido? Natividade: - Solta um grito! (Lembrando-se) Ah! **espera lá!** (Dá-lhe um pontapé. José solta um grito e foge pelo fundo) Pronto! (*Corpus* do Português: O Califa da Rua do Sabão, de Artur Azevedo, 1880).

O uso de *espera lá* no fragmento (5) pode suscitar minimamente duas interpretações analíticas. Na primeira delas, observando que a construção antecede o pontapé aplicado a José, entendemos tratar-se de construção indutora de expectativa sobre o mundo biossocial. Sob essa ótica, a expressão pode ter sido empregada para anunciar a ação ou o movimento que ocorreria em tempo próximo: o chute. Nesse caso, a construção é lida com as mesmas características descritas no contexto atípico de grau 1. Uma segunda leitura possibilita o entendimento de que a expressão *espera lá* tenha sido empregada para refrear a própria enunciação: a ordem “Solta

um grito!”. Ao dar-se conta de que o som pretendido não seria emitido naturalmente por José, Natividade tem uma espécie de *insight*, externado por meio da interjeição *Ah!*, seguida da construção refrecedor-enunciativa *espera lá!*. A função de refrreamento enunciativo será tratada adiante no contexto isolado de grau 2.

- **Contexto isolado de grau 1**

São verificadas desde o século XX, no contexto isolado de grau 1, as microconstruções que exprimem *indução de expectativa sobre a enunciação* (IEE). Tal significado, agora circunscrito ao âmbito discursivo, conforme o exemplo a seguir, deriva do contexto atípico de grau 1, em que as expressões relacionam-se a atividades do mundo biossocial.

(6) [...] E o que aconteceu com o Roberto Jefferson? Ele começou a ver que estavam batendo muito. E estava ficando sozinho. O PT fazendo papel de vestal. Não, não temos que tergiversar, vamos duro, e tal. Ah, é assim? Então **espera aí!** E entregou o esquemão. Agora, se for procurar de onde vem o dinheiro... O dinheiro não vem da plantação de amendoim da chácara do Lula. Vem dos beneficiários da política econômica. (*Corpus do Português: fala de Roberto Requião, séc. XX*).

No fragmento (6), Roberto Requião reproduz uma suposta fala de Roberto Jefferson diante de grande pressão política: “Ah, é assim? Então espera aí!”. A instanciação da construção *espera aí* no exemplo expressa o intuito de fazer com que o(s) interlocutor(es) se preparasse(m) para receber informações importantes naquele instante. Esse uso, portanto, gera expectativa sobre a enunciação iminente, o que pode ser confirmado pela narrativa “E entregou o esquemão” em sequência a *espera aí*.

- **Contexto isolado de grau 2**

No segundo grau do contexto isolado, encontramos, desde o século XIX, o emprego da microconstrução refrecedor-enunciativa (RE), originada



do contexto atípico de grau 2. O uso mais concreto, de refreamento de ações no âmbito biossocial, torna-se mais abstrato, circunscrito ao âmbito discursivo, passando a suspender momentaneamente o fluxo da enunciação. Detectamos os seguintes propósitos específicos no uso da RE: (i) recordar informações armazenadas na memória; (ii) refletir e formular resposta ou comentário; (iii) retificar o que foi dito anteriormente; (iv) redirecionar o tópico; (v) (re)tomar o turno. O dado (7) refere-se ao propósito (i):

(7) [...] Já não vive há muitos anos o cirurgião que a tratou; eu saí daqui há trinta e cinco e nunca mais o vi; se ele vivesse, poderia ajudar-me a recordar. **Espera lá...** Como a velhice nos varre tudo da memória! Ah! Uma circunstância... o aparecimento de uma criança no rio... - O quê? - Espera, António, não me quebres o fio das recordações. (*Corpus* do Português: Maria Moisés, de Camilo Castelo Branco, 1876).

Em (7), o uso de *espera lá* não é, ao menos diretamente, dirigido ao interlocutor. Seu emprego tampouco constitui propriamente um pedido ou ordem. A microconstrução *espera lá* marca a suspensão momentânea do discurso, que tem como objetivo, nesse exemplo específico, conseguir resgatar eventos armazenados na memória.

### • Contexto isolado de grau 3

No terceiro grau do contexto isolado, verificamos, a partir do século XX, a especialização do uso de microconstruções [*espera Afix*] em discursos argumentativos. Assim, *espera aí* e *espera lá* passam a ser empregados para refrear a proposição<sup>13</sup> do interlocutor, apoiando a alegação (justificação, ponderação ou refutação) do enunciador, que se dá por meios persuasivos.

---

<sup>13</sup> Segundo Charaudeau (1992), o mecanismo argumentativo se compõe de três elementos: proposição (tese), alegação (justificação, ponderação ou refutação) e persuasão.

(8) [...] - áh o excesso de liberdade no mundo tá provocando provocando a difusão de tudo que não presta. - **perai** isso não é verdade, isso é besteira, isso é besteira. - não, não sei de nada. Me perguntando tudo é válido sim, não é válido nada, sobretudo a violência. Violência é você fazer comunicar aquilo que não presta. “Olho para o mundo e por isso - isso aí é que é violência - cada vez me sinto mais solitário e aflito “. - não, eu cada vez acredito mais. - não sei não sei pra onde é que nós vamos. (*Corpus do Português: Linguagem Falada, Recife, sé. XX*).

No fragmento (8), o enunciador apresenta uma proposição (tese): “o excesso de liberdade no mundo está provocando a difusão de tudo o que não presta”. Seu interlocutor instancia a construção refreador-argumentativa “perai” para refrear a argumentação alheia ao mesmo tempo em que marca a sua posição diante da proposição apresentada: “isso não é verdade, isso é besteira, isso é besteira”. O expositor da proposição inicia, então, sua justificativa, valendo-se de alguns elementos persuasivos. Entendemos que, nesse último estágio da construcionalização de *espera aí* e *espera lá*, as microconstruções refreador-argumentativas perdem a composicionalidade original, de tal modo que ocorre redução morfofonêmica: *espera aí* > *perai* (erosão e aglutinação) e *espera lá* > *peralá* (erosão e justaposição).

- **Integração paradigmática**

A partir da construcionalização de *espera aí* e *espera lá*, ocorre estabilização dessas microconstruções no paradigma de macrofunção marcadora discursiva (MD), mais especificamente, particularizada pelo esquema formado por elementos indutores e afixoides [Indut Afix]<sub>MD</sub>. Além da refreador-enunciativa e da refreador-argumentativa, outras subfunções integram subesquemas associados a [Indut Afix]<sub>MD</sub> como, por exemplo, a exortativa (ex.: vamos lá!) e a de chamamento de atenção (ex.: olha aí!). Assim, as microconstruções *espera aí* e *espera lá* passam a ser confrontadas com esses membros em distinção/oposição do mesmo padrão, sendo gradualmente associadas com um significado mais abstrato, que serve como

um denominador comum para todo o paradigma: a marcação discursiva. Uma vez criado o esquema [Indut Afix]<sub>MD</sub> e, posteriormente, o subesquema [Indut<sub>R</sub> Afix<sub>Loc</sub>]<sub>RA</sub><sup>14</sup> no paradigma dos MD, ocorrem novas instanciações via analogização, como é o caso da microconstrução *pera aí*, verificada em função refrecedor-argumentativa no século XXI, sem vestígios de mudanças por micropassos ao longo dos séculos anteriores.

### Considerações finais

Neste capítulo, demonstramos como podemos lidar metodologicamente com a pesquisa diacrônica na LFCU, uma vez que a dinamicidade linguística é um dos pressupostos básicos do Funcionalismo. Enfatizamos que essa pesquisa requer viés qualitativo, com base no levantamento e na análise interpretativa dos contextos de uso e sua gradiência ao longo da trajetória da língua, e viés quantitativo, voltado para a produtividade desses padrões de uso contextual.

Devemos destacar, contudo, que a realização de pesquisa diacrônica pode trazer dificuldades, uma vez que, entre outras questões, podemos lidar com escassez de registros de fala sobretudo nos séculos mais longínquos e ainda com textos escritos de fontes duvidosas ou em mal estado de conservação, entre outros problemas. Outro ponto a ressaltar é que a aplicação da análise de nanopassos deve ocorrer apenas se de fato necessária, ou seja, caso estejam evidentes as nanomudanças intracontextuais, como verificado na trajetória de *espera Loc* (com exceção do contexto crítico). Em não havendo necessidade de tal esquadrinhamento, basta que o analista se concentre em um modelo prototípico de arranjo linguístico ou construção para cada contexto de mudança: fonte, atípico, crítico e isolado.

Por fim, vale mencionar que os fragmentos aqui exemplificados, que atestam a gradualidade dos usos linguísticos, em termos históricos, convivem na sincronia atual do português. Tal observação evidencia, conforme

---

<sup>14</sup> Indut<sub>R</sub>: indutor-refreedor; Afix<sub>Loc</sub>: afixoide de origem locativa; [ ]<sub>RA</sub>: refrecedor-argumentativo.

Bybee (2010), que a gradiência sincrônica é consequente da gradualidade diacrônica, o que enseja a pesquisa de ambas as perspectivas, no que se configura como uma abordagem pancrônica. No capítulo seguinte, que aborda a mudança linguística em perspectiva sincrônica, essa questão será mais detalhada e desenvolvida.

## Referências

- BOOIJ, G. Morphology in construction grammar. *In*: HOFFMANN, T.; TROUSDALE, G. (eds.). **The Oxford handbook of construction grammar**. Oxford: Oxford University Press, 2013. p. 255-273.
- BYBEE, J. **Language, usage and cognition**. New York: Cambridge University Press, 2010.
- BYBEE, J. **Language change**. Cambridge: Cambridge University Press, 2015.
- CHARAUDEAU, P. **Grammaire du sens et de l'expression**. Paris: Hachette Éducation, 1992.
- CROFT, W. **Radical construction grammar: syntactic theory in typological perspective**. Oxford: Oxford University Press, 2001.
- CRYSTAL, D. **A dictionary of linguistics and phonetics**. Oxford: Blackwell Publishing, 2008.
- DIEWALD, G. A model for relevant types of contexts in grammaticalization. *In*: WISHER, Ilse; DIEWALD, G. (eds). **New reflections on grammaticalization**. Amsterdam: John Benjamins, 2002. p.103-120.
- DIEWALD, G. Contexts types in grammaticalization as constructions. *In*: Special volume 1: **Constructions all over** – case studies and theoretical implications. Dusseldorf, 2006. Disponível em: <[www.constructions-online.de/009-4-6860](http://www.constructions-online.de/009-4-6860)>. Acesso em: 10 mar. 2019.
- DIEWALD, G; SMIRNOVA, E. “Paradigmatic integration”: the fourth stage in an expanded grammaticalization scenario. *In*: DAVIDSE, K. *et al.* (eds). **Grammaticalization and language change** – new reflections. Amsterdam/Philadelphia: John Benjamins, 2012. p.111-131.
- GOLDBERG, A. **Constructions: a construction grammar approach to argument structure**. Chicago: University of Chicago Press, 1995.
- HEINE, B. On the role of context in grammaticalization. *In*: WISCHER, I.; DIEWALD, G. (eds). **New reflections on grammaticalization**. Amsterdam: John Benjamins, 2002. p.83-101.

LACERDA, P. F. A. C. O papel do método misto na análise de processos de mudança em uma abordagem construcional: reflexões e propostas. **Revista Linguística/Revista do Programa de Pós-Graduação em Linguística da Universidade Federal do Rio de Janeiro**, vol. especial, p. 83-101, 2016.

MARTELOTTA, M. E. **Mudança linguística: uma abordagem baseada no uso**. São Paulo: Cortez, 2011.

OLIVEIRA, M. R. Contexto: definição e fatores de análise. In: OLIVEIRA, M. R.; ROSÁRIO, I. C. (orgs). **Linguística centrada no uso: teoria e método**. Rio de Janeiro: Lamparina/Faperj, 2015. p. 22-35.

OLIVEIRA, M. R.; ARENA, A. B. O viés funcional do pareamento simbólico *função < > forma* na abordagem construcional da gramática. **Soletras**, n. 37, p. 30-58, 2019.

PAULA, V. B. **A construcionalização de grau intensivo [[p(r)ara lá de] [X]] no português**. 131f. Tese (Doutorado em Estudos de Linguagem) Universidade Federal Fluminense, Instituto de Letras, Niterói: RJ, 2021.

ROSA, F. S. L. **A mesoconstrução marcadora discursiva refreador-argumentativa: uma análise cognitivo-funcional**. 216 f. Tese (Doutorado em Estudos de Linguagem) Universidade Federal Fluminense, Instituto de Letras, Niterói: RJ, 2019.

TEIXEIRA, A. C. M. **A construção verbal marcadora discursiva VLoc<sub>MD</sub>: uma análise funcional centrada no uso**. 297 f. Tese (Doutorado em Estudos de Linguagem) Universidade Federal Fluminense, Instituto de Letras, Niterói: RJ, 2015.

TRAUGOTT, E. A constructional perspective on the rise of metatextual discourse markers. **Cadernos de Linguística**. Abralín, v. 2, n. 1, p. 1-25, 2021a.

TRAUGOTT, E. C.; TROUSDALE, G. **Constructionalization and constructional changes**. New York: Oxford University Press, 2013.

## 4. Construcionalidade e mudança na sincronia

Ivo da Costa do Rosário (UFF/CNPq/FAPERJ)  
Monclar Guimarães Lopes (UFF)

### **Considerações iniciais**

Nos capítulos anteriores deste livro, apresentamos a metodologia de pesquisa funcionalista tanto em sua aplicação sincrônica (capítulo 2) quanto diacrônica (capítulo 3). Como foi apresentado, a opção pelo viés sincrônico ou diacrônico enseja diferentes escolhas a serem tomadas no processo de investigação dos fenômenos linguísticos em geral. Também é comum que ora alguns autores sejam mais selecionados que outros para dar sustentação teórica às investigações, já que alguns se debruçam mais sobre sincronia, ao passo que outros sobre a diacronia.

De certa forma, por um lado, este capítulo dá continuidade aos dois anteriores, tendo em vista que tratamos também da mudança linguística. Contudo, por outro lado, o viés é distinto, já que a mudança será tratada a partir de um ponto de vista sincrônico, o que foi denominado por Rosário e Lopes (2019) de *construcionalidade*.

Um princípio clássico funcionalista revela que, no plano da mudança, há uma relação indissociável entre sincronia e diacronia, de tal modo que o estado sincrônico das línguas revela traços de seu percurso histórico. Essa concepção já estava presente em Givón (1971, p. 413), para quem “a morfologia de hoje é a sintaxe de ontem” e na visão de Faarlund (1989, p. 71) ao defender que “a sintaxe de hoje pode ser o produto da pragmática discursiva de ontem”.

Em outras palavras, o exame da sincronia, com foco na gradiência dos elementos linguísticos, a partir de uma metodologia própria, pode nos ajudar a compreender a gênese e a trajetória histórica dos fenômenos

linguísticos, já que segundo Heine e Kuteva (2007, p. 210), “o estado sincrônico das línguas pode ser visto como o produto congelado de processos cognitivos e comunicativos ocorridos no passado”. Assim, a sincronia espelha processos desenvolvidos na diacronia.

Vale ressaltar que, de alguma forma, essa ideia já estava presente também nos estudos variacionistas labovianos, sob o rótulo de *mudança em tempo aparente*. Assim, nessa perspectiva teórica, a fala dos mais idosos atestada sincronicamente, por exemplo, serve de ponto de partida para o delineamento do sistema linguístico em fatias temporais pgressas.

Ainda na esteira dos estudos labovianos, vale destacar aqui, por fim, o princípio do uniformitarismo. Segundo esse princípio, “tendências hoje em curso devem ter atuado em estágios anteriores e possivelmente continuarão a atuar” (cf. CUNHA, 1999, p. 2). Logo, a literatura apresenta diversos estudos já realizados com o mote central de que a sincronia espelha a diacronia. Neste capítulo, a defesa do conceito de construcionalidade ancora-se justamente nessa premissa teórica, além de basear-se em uma série de outros trabalhos, com destaque para Lehmann ([1982]2015), Hopper (1991) e Himmelmann (2004), além de outras investigações de cunho teórico e empírico.

Este capítulo está organizado em cinco partes. Após estas considerações iniciais, na próxima seção, propomos algumas reflexões teóricas sobre a mudança linguística, tal como compreendida na visão funcional centrada no uso, e a associamos ao paradigma da construcionalidade, foco deste capítulo. Em seguida, apresentamos resultados de dois estudos em que é possível comprovar que o exame sincrônico das línguas recapitula os traços de sua trajetória histórica. Por fim, apresentamos as considerações finais e elencamos a bibliografia consultada.

## **Mudança linguística e construcionalidade**

A mudança linguística é um fenômeno natural de todas as línguas humanas. Em termos bem gerais, a mudança indica que, ao longo do tempo, os sistemas linguísticos vão sofrendo modificações de caráter formal e/ou funcional. A metodologia mais apropriada para captar e explicar esses movimentos históricos consiste na coleta e análise de dados extraídos de

outras sincronias pretéritas. Por outro lado, a literatura apresenta consideráveis apontamentos acerca dos reflexos da mudança histórica no quadro sincrônico das línguas, de modo que a análise de dados de nossa fatia temporal atual seria um recurso plausível para a postulação de trajetórias progressas dos fenômenos linguísticos, especialmente com foco nos distintos graus de gramaticalidade das construções.

Cunha (1999, p. 2) defende, de modo muito claro, uma dimensão pancrônica da mudança, com base na qual deveriam se fundar os estudos linguísticos de base funcionalista. De fato, uma pesquisa mais profunda e de caráter longitudinal implica uma metodologia que conjugue tanto o viés diacrônico quanto o sincrônico, o que requer tempo, expertise e experiência do pesquisador. Por outro lado, a própria autora admite que “as línguas têm passado, e o estado sincrônico é uma função desse desenvolvimento passado”.

Hilpert (2014, p. 194), por sua vez, assevera que “a mudança é o estado natural de uma língua viva”. Assim, forma e função estão continuamente se realinhando para dar conta de novas necessidades comunicativas. Essa permanente dinamicidade do sistema pode ser capturada a partir da noção de gramática emergente, cunhada por Hopper (1991), ou da metáfora das dunas de areia, nos termos de Bybee (2010). De diferentes maneiras, os teóricos sustentam que há permanente movimento nas línguas.

Na LFCU, tem sido cada vez mais frequente a incorporação dos conceitos de *gradiência* e *gradualidade* às pesquisas, sendo o primeiro associado à variabilidade sincrônica, e o segundo à variabilidade diacrônica. Trousdale (2013, p. 32) ressalva que não é muito acertado dizer que gradualidade (na diacronia) equivale exatamente à gradiência (na sincronia), pois assevera que “gradiência é uma consequência natural de uma sequência de mudanças”. Logo, comungamos com a visão do teórico na proposição de uma aproximação entre os recortes sincrônico e diacrônico, ainda que não se estabeleça uma relação de inexorável derivação entre ambos.

Há farta literatura comprobatória de que a mudança é um produto direto da variação. Esse postulado é mantido na abordagem construcional da gramática, já que para Traugott e Trousdale (2010, p. 4), “a variação é o resultado e a razão da mudança”. Na visão de Hoffmann e Trousdale (2011,



p. 12), essa ideia também se confirma: “mudanças são sempre manifestadas na variação sincrônica, e as mudanças do passado são comumente refletidas nas alternâncias sincrônicas”. De maneira ainda mais explícita, assim afirmam Trousdale e Traugott (2010, p. 26-27):

Várias hipóteses foram apresentadas sobre a probabilidade de a variação sincrônica refletir a história passada. [...] Heine e outros argumentaram que a história pode ser reconstruída com base na variação sincrônica [...]. É uma hipótese razoável que a gradiência é o resultado de mudanças em pequenos passos, associadas à gradualidade.

As reflexões realizadas até este ponto sustentam a defesa de que, na abordagem construcional da gramática, o estudo de análises sincrônicas permite reflexões acerca da gênese e das rotas de mudança das construções de uma língua. Aliás, como explicam Rosário e Lopes (2019), esse também foi o percurso teórico seguido pelo paradigma da gramaticalização, que nasce essencialmente voltado para a diacronia, mas, com o tempo, passa também a designar “um fenômeno sintático, discursivo-pragmático, a ser estudado do ponto de vista dos padrões fluidos de uso de língua ao longo do tempo ou em um momento sincrônico segmentado no tempo” (Traugott; Heine, 1991, p. 1). Essa perspectiva de análise foi chamada inicialmente de “gramaticalização sincrônica”. Após sucessivos refinamentos, chegou-se ao termo “gramaticalidade”, justamente para dar conta das análises sincrônicas preocupadas com os reflexos de desenvolvimentos históricos das construções no interior da gramática.

Uma vez que o estado sincrônico das línguas revela diferentes camadas históricas de mudança, a gramaticalidade passou a ser analisada em uma perspectiva de gradiência. Assim, atestar a gramaticalidade de uma construção implica(va) identificar os graus de sua gramaticalização sincrônica.

Uma grande contribuição para a defesa de que o estado sincrônico de uma língua pode nos trazer importantes reflexões sobre sua origem e trajetória está em Lehmann (2015[1982]). Segundo o autor, “meios estruturais distintos, presentes em uma língua em estado sincrônico, podem ser organizados em uma escala de gramaticalização” (Lehmann, 2015[1982],

p. 129). Ao investigar a gramaticalização sincrônica (ou gramaticalidade), o autor apresenta um conjunto de parâmetros e processos que passamos a brevemente apresentar. Vejamos:

**Quadro 1.** Correlação dos parâmetros de gramaticalização.

	Parâmetro	Gramaticalização fraca	Processo	Gramaticalização forte
E i x o  p a r a d i g m á t i c o	Integridade	Muitos traços semânticos, possivelmente polissilábicos	Atrição	Poucos traços semânticos, itens possivelmente monossegmentais
	Paradigmaticidade	Participação “frouxa” do item no campo semântico	Paradigmatização	Paradigma pequeno, fortemente integrado
	Variabilidade paradigmática	Escolha livre de elementos de acordo com as intenções comunicativas	Obrigatoriedade	Escolha sistematicamente restrita, com uso altamente obrigatório
E i x o  s i n t a g m á t i c o	Escopo estrutural	Relação do item com constituinte de complexidade arbitrária	Condensação	Item modifica palavra ou raiz
	Fixação	Justaposição independente do item	Coalescência	Item é afixo ou até mesmo um traço fonológico
	Variabilidade sintagmática	Movimentação livre do item	Fixação	Item ocupa posição fixa

Fonte: Lehmann (2015[1982], p. 174), adaptado pelos autores.

O autor defende que esses parâmetros e processos podem ser analisados de modo individual, mas a aplicação de todos eles ajuda, sem dúvida, a atestar um nível de gramaticalidade maior de um determinado fenômeno linguístico em relação a outras expressões linguísticas correlatas. Alguns desses parâmetros e processos se correlacionam positivamente, outros negativamente. Assim, por exemplo, na visão de Lehmann (2015[1982]), à medida que a gramaticalização aumenta, os parâmetros de coesão também aumentam, ao passo que os parâmetros de peso e variabilidade diminuem.

Vejamos uma breve explicação de cada parâmetro a seguir, com base no próprio autor e em Gonçalves e Carvalho (2007):

- a) *Integridade* - Refere-se ao tamanho substancial de um signo, em termos de sua matriz semântica e fonológica. O signo maior tem maior proeminência no contraste com outros signos.
- b) *Paradigmaticidade* - Diz respeito ao grau de coesão de um item com outros de um paradigma. Esse parâmetro é medido por meio da verificação do tamanho e da homogeneidade do paradigma, isto é, pelas similaridades entre os seus membros.
- c) *Variabilidade paradigmática* - Possibilidade de uso de um elemento em lugar de outro, sem grande prejuízo semântico.
- d) *Escopo estrutural* - Peso estrutural da construção que ele ajuda a formar. À medida que a gramaticalização avança, o escopo diminui.
- e) *Fixação* - Diz respeito ao grau de ligação de um item em relação aos demais. Se a gramaticalização é fraca, esse grau de conexão é menor.
- f) *Variabilidade sintagmática* - Possibilidade de mobilidade de um item na construção em que ocorre.

No processo de gramaticalização, esses parâmetros sofrem modificações sensíveis ou, nos termos de Lehmann ([1982]2015), passam por

processos que contribuirão para o que ele chama de gramaticalização forte. Vejamos:

- a) *Atrição* - Diz respeito à perda de integridade de um signo. A atrição pode se dar em termos fonológicos (erosão fonética) ou pode ocorrer em termos semânticos (dessemantização ou *bleaching*).
- b) *Paradigmatização* - Paradigmas altamente gramaticalizados tendem a ser menores que os menos gramaticalizados.
- c) *Obrigatoriedade* - Redução na liberdade de escolha de itens, tendo em vista uma maior restrição no paradigma.
- d) *Condensação* - Perda progressiva de arbitrariedade no escopo estrutural. Assim, o item passa de uma relação com constituintes de complexidade arbitrária para uma relação com palavras ou radicais.
- e) *Coalescência* - Grau maior de fusão entre as partes, levando, em casos extremos, a aglutinações e ao zero.
- f) *Fixação* - À medida que um item fica mais fixo na construção, mais gramaticalizado se torna.

Aos parâmetros de Lehmann ([1982]2015), acima brevemente explicitados, assomam-se os chamados princípios de Hopper (1991), que têm o mérito de ir além dos aspectos formais, mais frisados por Lehmann ([1982]2015). Além disso, Hopper (1991) tinha um foco maior na chamada gramaticalização incipiente, ou seja, nos estágios iniciais da mudança. Os princípios podem assim ser apresentados:

- a) *Estratificação* - Em um domínio funcional, novos usos vão surgindo. Esses novos usos passam a conviver com os antigos, em camadas. Assim, coexistem camadas novas e antigas na língua, em diferentes codificações.

- b) *Divergência* - Diferentes graus de gramaticalização de um mesmo item convivem em uma mesma fatia temporal, de modo que, em um contexto X, o item tem um significado ou uso distinto do aferido em um contexto Y.
- c) *Especialização* - Diz respeito a um progressivo afinamento nas escolhas de formas pertencentes a um mesmo domínio. Com o tempo, uma forma mais gramaticalizada ganha maior frequência de uso.
- d) *Persistência* - Manutenção de traços semânticos da forma-fonte na forma gramaticalizada, com consequentes restrições sintáticas para a forma nova.
- e) *Decategorização* - Perda parcial ou total de traços da categoria-fonte.

Todos os parâmetros e processos de Lehmann ([1982]2015), bem como os princípios de Hopper (1991), foram pensados para dar conta do processo de gramaticalização com foco maior na forma. Contudo, são conceitos forjados a partir de farto número de pesquisas empíricas envolvendo dezenas de línguas humanas, com grande rigor metodológico. Isso significa que, apesar de estarmos atualmente em uma outra fase dos estudos funcionalistas, agora, no campo da LFCU, defendemos a pertinência da aplicação adaptada dessas contribuições dos autores, especialmente na defesa realizada neste texto do conceito de construcionalidade. Isso será ilustrado mais à frente, ao tratarmos dos trabalhos de Acosta (2016, 2021) e de Paula (2021).

Além de Lehmann ([1982]2015) e de Hopper (1991), que se debruçaram sobre aspectos mais formais da gramaticalização, evocamos aqui também a proposta de Himmelmann (2004), mais associada ao plano semântico-pragmático. De acordo com o autor, a gramaticalização envolve basicamente um processo de expansão em três níveis, conforme pontuam Rosário e Oliveira (2016):

- a) *mudança da classe hospedeira*, com a ampliação paradigmática de membros de uma dada categoria, face à entrada de novo membro na classe;
- b) *mudança de contexto sintático*, envolvendo metonimização, com rearranjo na ordem dos constituintes internos e consequente formação de uma nova sintaxe regular de expressão;
- c) *mudança de contexto semântico-pragmático*, considerada por Himmelmann (2004) como a mais importante, uma vez que envolve desbotamento de sentido, com ressemantização e uso anafórico associativo.

Traugott e Trousdale (2013, p. 96-97) realizam um importante estudo sistemático das diferentes perspectivas que abrigam a gramaticalização. Os autores propõem que há autores e trabalhos mais afinados com uma visão de gramaticalização como redução e aumento de dependência, como a perspectiva de Lehmann ([1982]2015); de outro lado, há os autores que se alinham a uma visão mais aberta, na concepção de gramaticalização como expansão, que é a linha de Himmelmann (2004). Isso significa que a mudança linguística pode caminhar para diferentes direções, ressalvada a observação de que costuma haver, em termos gerais, redução no campo formal e expansão no campo funcional, o que seria uma forma de compatibilizar ambas as visões aqui apresentadas.

Na visão de Bybee (2010), as línguas humanas são sistemas adaptativos complexos. Dentre tantas implicações possíveis advindas dessa visão, podemos destacar que a mudança é sempre um produto complexo decorrente de múltiplos fatores, tanto linguísticos quanto extralinguísticos (Martelotta, 2009, p. 2). Logo, qualquer cálculo que intente rastrear o passado de uma língua a partir de seu estado sincrônico exige parcimônia. Assim, em hipótese alguma defendemos que o recorte sincrônico poderia nos proporcionar amplo e irrestrito conhecimento do passado das línguas. Essa também não era a defesa de Lehmann ([1982]2015) nem de Hopper (1991) nem de Himmelmann (2004). A defesa da construcionalidade apenas aponta para o fato de que a sincronia permite a apreensão de reflexos do passado a partir dos diferentes níveis de gramaticalidade e de usos de

uma dada construção, já que há um considerável conjunto de pesquisas que dá sustentação a essa tese.

Feitas todas essas observações preliminares, podemos agora recuperar o conceito de construcionalidade, tal como definido em Rosário e Lopes (2019, p. 92):

relação sincrônica estabelecida entre construções, de tal sorte que (i) duas construções A e B apresentam horizontalmente algum grau de parentesco, ou (ii) uma construção menos esquemática pode ser associada verticalmente a uma ou mais construções de natureza mais esquemática.

Neste capítulo, propomos um refinamento da proposta de 2019, de modo a acentuarmos mais fortemente o caráter gradiente das construções linguísticas. Portanto, considerando que os níveis de gramaticalidade sincrônica de uma dada construção linguística refletem aspectos da sua origem e trajetória histórica, redefinimos o conceito de construcionalidade e o estabelecemos como a

**relação sincrônica entre duas ou mais construções, de modo que uma construção pode ser apontada como base para outra(s), a partir de seus diferentes níveis de gradiência e gramaticalidade**

Assim, o estudo da construcionalidade permite uma reflexão acerca dos traços gradientes de uma dada construção gramatical em recorte sincrônico. Por sua vez, esse estudo permite o levantamento de informações sobre os seus caminhos diacrônicos, já que é possível apontar que uma dada construção X é base para (ou é anterior a) uma Y. Em outros termos, o conceito de construcionalidade, por um lado, prevê a remodelação da ideia de gramaticalidade agora aplicada à abordagem construcional; por outro lado, permite que um estudo essencialmente sincrônico aponte informações sobre a gênese e a trajetória de uma dada construção ao longo do tempo, visto que os processos de mudança deixam marcas no estado sincrônico das línguas.

Além disso, procuraremos demonstrar como os mesmos princípios e conceitos implicados no estudo da construcionalização (em viés diacrônico) podem igualmente servir a análises de construcionalidade (em viés sincrônico), com boa recuperação dos postulados de Lehmann ([1982]2015), de Hopper (1991) e de Himmelmann (2004).

Essencialmente, nossa compreensão acerca da construcionalidade permanece a mesma de Rosário e Lopes (2019), já que essa nova definição aqui apresentada também assegura a análise de relações horizontais e verticais entre construções distintas em diferentes níveis de abstração. Entretanto, agora reiteramos que queremos focalizar com mais clareza a questão dos níveis de gramaticalidade e de gradiência das construções, bem como a interface sincronia x diacronia.

Em entrevista concedida por Graeme Trousdale em 2019 à Revista *Soletras*, o pesquisador afirma que sua obra de 2013 (*Construcionalização e Mudanças Construcionais*), que tem servido de base para muitos trabalhos no país, “é um livro primariamente sobre mudança linguística [...] e, por esse motivo, seu foco é a diacronia, não a sincronia”. (Rosário, 2019, p. 10). Contudo, o mesmo teórico acrescenta que é possível “observar um grupo de padrões em qualquer ponto sincrônico da história de uma língua e tentar usar alguns dos princípios de trabalho da mudança construcional para explicar por que esses mesmos padrões têm o formato particular que os caracteriza” (ibidem).

O estudo da construcionalidade busca justamente desenvolver essa ideia indicada por Trousdale. Assim, também postulamos que é possível a aplicação da abordagem contextual de Diewald (2002, 2006) e de Diewald e Smirnova (2012), já apresentada no capítulo 3, por exemplo, para o estudo dos diferentes estágios de mudança linguística, ainda que de um ponto de vista sincrônico. Afinal, a sincronia reflete as camadas (cf. Hopper, 1991) pregressas da língua, em uma situação de convivência ou competição, sem que haja uma situação de necessária substituição de uma pela outra.



A partir dessas premissas, julgamos possível a aplicação dos fatores de construcionalização (esquematicidade<sup>1</sup>, produtividade<sup>2</sup> e composicionalidade<sup>3</sup>) ao estudo da construcionalidade. Isso significa que, além da sua aplicação ao exame histórico da língua, tais conceitos organizados por Traugott e Trousdale (2013) também passam a ser perfeitamente operacionalizáveis em um estudo que se debruce sobre o estado sincrônico de construções gramaticais marcadas por gradiência.

Aqui também evocamos os conceitos de *chunking*<sup>4</sup>, analogia<sup>5</sup> e neanálise<sup>6</sup>, igualmente relevantes para o estudo dos níveis de gradiência, de gramaticalidade e de mudança linguística em recorte sincrônico. É verdade que muitos desses conceitos já eram usados na pesquisa sincrônica *stricto sensu*, mas a proposição aqui defendida é que eles possam também dar conta da construcionalidade, nesse novo enquadre teórico proposto.

Uma das vantagens da construcionalidade é a possibilidade de reconstrução de traços do passado histórico das construções quando não há

---

<sup>1</sup> A *esquematicidade* diz respeito ao nível de abstração das construções, tendo em vista que esses pareamentos simbólicos podem ser bastante abstratos ou bastante específicos (além de diversos pontos intermediários). Essa organização esquemática das construções tem sido representada por meio de uma hierarquia comumente composta de três níveis: esquemas > subesquemas > microconstruções

<sup>2</sup> A *produtividade* diz respeito à extensibilidade da construção, associada a sanções e restrições. O levantamento de frequência *type* e *token* também é importante nesse aspecto.

<sup>3</sup> *Composicionalidade* diz respeito ao grau de transparência entre forma e significado das construções. Pode ser de natureza sintática ou semântica. Segundo Rosário e Oliveira (2016, p. 246), “*composicionalidade semântica* diz respeito à soma dos significados das partes. Assumimos que uma construção é mais composicional em termos semânticos quando o significado das partes ainda é recuperado no significado do todo. *Composicionalidade sintática*, por sua vez, diz respeito ao nível de integridade morfossintática das subpartes, no sentido de que quanto mais composicional, mais essas subpartes retêm as propriedades gramaticais de sua categoria fonte”.

<sup>4</sup> *Chunking* é processo que gera unidades complexas de organização a partir de partes singulares. O processo de *chunking* gera como produtos os *chunks*: sequências de unidades embaladas como unidades de sentido.

<sup>5</sup> *Analogia* (ou *analogização*) envolve a reconfiguração de traços ou dimensões internas de uma dada construção. Nessa perspectiva, a analogia é vista como *mecanismo* ou processo de mudança.

<sup>6</sup> *Neanálise* é um mecanismo que consiste em uma nova interpretação para uma forma já existente, o que normalmente ocorre por meio da negociação de sentidos. Segundo Traugott e Trousdale (2013, p. 36), o termo *neanálise* é preferível a *reanálise* por conta de não ser possível analisar novamente uma estrutura que acaba de nascer; logo, nem sempre há análise anterior.

fontes disponíveis para tal. Esse desafio já é conhecido no campo dos estudos linguísticos quando fazemos referência ao trabalho empreendido pelos comparativistas que, por meio de reconstruções e analogias, propunham as origens e trajetórias históricas das línguas advindas do proto-indoeuropeu. Esse tipo de metodologia passou por sofisticções e refinamentos ao longo do tempo, especialmente com o crescente conhecimento acerca da cognição humana e da própria estrutura das línguas. Logo, uma das razões para a aplicação da construcionalidade consiste na falta (ou na dificuldade de acesso) de registros históricos para o estudo do passado de determinadas construções. Bybee (2015, p. 8) é uma das defensoras desse método ao afirmar que “em línguas com ou sem registros escritos, as mudanças em curso criam variação, e o estudo dessa variação pode prover excelentes evidências sobre como a mudança ocorre”.

É sabido que, ao contrário disso, a língua portuguesa apresenta farta documentação histórica, em diferentes *corpora* de língua escrita. Entretanto, há outros fatores que podem ensejar a viabilidade de um trabalho de mudança a partir de dados sincrônicos, como os estudos em língua falada. A ausência de registros históricos nessa modalidade de uso, por exemplo, não permite uma reconstrução de diversos fenômenos linguísticos pautados em dados de língua real.

Outra razão importante para os estudos em construcionalidade está na própria decisão metodológica do pesquisador no momento de definir seus objetivos de pesquisa, o tempo disponível para o trabalho e outros fatores associados. Nem sempre há condições nem tempo de o pesquisador se enveredar na pesquisa histórica, o que acarreta inúmeros desafios e demanda conhecimentos mais aprimorados das estruturas morfossintáticas das línguas em suas fases pretéritas, além do necessário domínio das grafias antigas, dos contextos político-históricos etc.

Antes de concluir esta seção, vale destacar que a noção de construcionalidade foi originalmente proposta por Rosário e Lopes (2017) sob o rótulo de “construcionalização sincrônica”. Com base nas ideias lançadas pelos autores citados, Dall’Orto e Lacerda (2019, p. 185-186) compilaram cinco justificativas para a adoção desse constructo teórico e acrescentaram uma sexta:

1. “O conceito de mudança presente nos estudos sociolinguísticos de Labov (2008) já admitiria o trabalho com a sincronia”;
2. “O próprio conceito de língua de Traugott e Trousdale (2013, p. 44, tradução nossa), a partir de Bybee (2010, p. 1), pressupõe que a análise sincrônica é coerente com os estudos em mudança, conforme observamos a seguir: “(...) a língua é um fenômeno que exhibe estrutura aparente e regularidade de padrões ao mesmo tempo em que apresenta considerável variação em todos os níveis”.
3. “Traugott e Trousdale (2013) admitem a correlação gradualidade e gradiência quando assumem que as mudanças linguísticas apresentam uma pequena escala de gradiência, que se manifesta na variação”.
4. “Há um correlato entre construcionalização diacrônica e construcionalização sincrônica, assim como há a correlação gramaticalização e gramaticalidade”.
5. “Os próprios modelos de gramática de construções adotados no âmbito dos estudos em Linguística Funcional Centrada no Uso foram pensados originalmente para a representação do conhecimento linguístico na mente dos falantes, e não para a mudança linguística, como é o caso da Gramática de Construções Radical de Croft (2001)”.
6. “Uma análise sincrônica dos dados nos permite, também, observar a extensibilidade de padrões a partir da perspectiva da analogização, sem necessariamente negarmos a mudança linguística em micropassos – isto é, do ponto de vista de uma sucessão de neoanálises”.

Uma vez apresentadas as razões para a adoção do conceito de construcionalidade, bem como comprovada a sua plausibilidade teórica, vejamos dois exemplos de trabalhos já realizados que legitimam a adoção desse paradigma.

## Exemplos de aplicação

De antemão, é importante destacar que o percurso a ser adotado em uma pesquisa baseada no paradigma da construcionalidade pode adotar os passos/procedimentos indicados neste livro, no segundo capítulo, quais sejam: a) escolha do objeto de pesquisa; b) seleção ou constituição de *corpus* (ou *corpora*); c) revisão de literatura; d) análise piloto; e) levantamento das hipóteses e planejamento dos fatores de análise; f) análise sistemática, tabulação das ocorrências e apresentação dos dados.

Como vimos defendendo ao longo deste capítulo, existem fartas evidências de que a análise de dados sincrônicos permite compreender boa parte do processo de formação das construções, a partir de sua natureza gradiente. Para isso, vamos apresentar aqui alguns dados da pesquisa de Acosta (2017, 2021) e de Paula (2021).

Em Acosta (2017), uma pesquisa essencialmente sincrônica, a autora detectou o uso de pares correlativos disjuntivos aparentemente inovadores. Esses correlatores, denominados *não espelhados*, são os seguintes: *seja...ou, quer...ou, nem...ou*. Àquela altura, foi postulado que a gênese desses conectores poderia ser explicada a partir de possíveis mesclas entre pares espelhados (*seja...seja, quer...quer, nem...nem*) que tomaram *ou...ou* como base para seu processo de formação, já que este é o conector mais prototípico e transparente no campo da disjunção. Vejamos algumas ocorrências desses pares correlativos não espelhados:

- (01) O Brasil, é claro, vive segundo essa mesma regra. Mas a história, aqui, é muito mais quente, porque o osso em disputa é muito maior. Perder uma eleição lá fora é ruim – mas no fim é apenas isso, uma derrota. Aqui não. Se o PT perder a eleição presidencial de 2014, [seja com a presidente Dilma Rousseff] [ou com o ex-presidente Lula,] vai haver um terremoto na vida pessoal de dezenas de milhares de pessoas, possivelmente muito mais, a começar por seus bolsos. Fonte: Revista Veja on-line. Ed. 1609. Séc. XX, p.36)
- (02) O físico Niels Bohr tinha uma ferradura sobre o batente da porta da frente. Alguém perguntou a ele: “Niels por que a ferradura? Você

não pode acreditar nisso”. Ele respondeu: “É claro que não acredito. Mas isso funciona [**quer** você acredite] [**ou** não.]” Espero que Steven acredite que a oração funciona, [**quer** você acredite] [**ou** não.]. Fonte: Revista Veja *on line*. Ed. 1782. Séc. XXI, p.43

- (03) Ninguém faz nada quanto a isso; por alguma razão misteriosa, insondável, nada se corrige. Homens, mulheres, jogados aos magotes em celas que não admitiriam razoavelmente [**nem** seis] [**ou** dez.]. Fonte: Revista Veja *on line*. Ed. 1932. Séc. XXI, p.74.

Em virtude da falta de evidências históricas que explicassem a formação desses pares correlativos no português (devido à natureza sincrônica da investigação à época), a autora aventou a hipótese de que essas mesclas teriam sido licenciadas por meio do mecanismo de analogização, em que o falante reconfigura e realinha traços de uma construção já existente para a formação de uma nova. A não prototipicidade desses correladores (com baixa frequência e ainda não retratados nos manuais normativistas), associada à presença constante de *ou* no segundo correlator serviram como fortes indícios para que fosse defendida a ideia de que havia ocorrido mudança formal no esquema dos correladores disjuntivos do português.

As evidências eram bastante fortes, contudo, a falta de comprovação histórica, especialmente por conta da ausência da análise de dados diacrônicos, não permitiu, àquela época, uma constatação cabal de que esses conectores eram mais recentes na língua. Afinal, o que garantia, de fato, que eles já não estivessem presentes no uso, em outros séculos, mesmo que de forma marginal?

Em tese de doutoramento, Acosta (2021) investigou essas mesmas construções em uma perspectiva pancrônica, dando continuidade ao seu trabalho de mestrado, realizado em 2016. Em uma agenda mais ampla de pesquisa, rastreou esses usos desde o século XIII até a sincronia atual. Ao final, a autora constatou que, de fato, os pares correlatos não espelhados haviam surgido apenas no século XX. Com isso, ela comprovou empiricamente que *seja...ou*, *quer...ou* e *nem...ou* são inovações na língua portuguesa, tendo em vista sua convencionalização no século XX, a partir de uma base que tomou *ou...ou* como modelo analógico.

O exame sincrônico realizado pela pesquisadora, ainda na etapa de mestrado, demonstrou algumas características da construcionalidade dos pares correlatos disjuntivos não espelhados (*seja...ou, quer...ou, nem...ou*). Em outras palavras, com base em um estudo de base sincrônica, foi possível levantar traços da origem da construção em tela, a partir dos níveis de gradiência e de gramaticalidade dos conectores investigados. A baixa frequência, a não prototipicidade, a mescla formal baseada em um par correlato canônico (*ou...ou*), a postulação plausível de uma relação analógica entre um esquema mais abstrato e as microconstruções citadas, a variabilidade formal e o percurso histórico de outros conectores seriam suficientes, enfim, para comprovar que *seja...ou, quer...ou, nem...ou* são pares correlatos inovadores na língua.

Com relação aos princípios de Hopper (1991), atestamos estratificação (tendo em vista que esses novos conectores são camadas em convivência com os anteriores) e decategorização (uma vez que os elementos *seja, quer* e *ou* são realinhados em novos usos). Nos termos de Himmelmann (2004), atestamos mudança de classe hospedeira (com a sensível ampliação do paradigma dos conectores disjuntivos) e mudança de contexto sintático (com rearranjo dos constituintes internos dos conectores).

Como vimos insistindo neste capítulo, somente a pesquisa pancrônica (com dados sincrônicos e diacrônicos) poderia comprovar cabalmente todos os traços principais do processo de construcionalização gramatical dos pares correlativos disjuntivos, com informações precisas e detalhadas acerca dos micropassos de mudança. Contudo, é inegável que o estado sincrônico gradiente dessas construções, por si só, é capaz de revelar muitas informações sobre sua gênese e trajetória histórica (como ficou comprovado no cotejo entre as pesquisas realizadas pela autora, respectivamente em 2016 e em 2021). Destacamos, ainda, que a construcionalidade se baseia em generalizações oriundas de inúmeras pesquisas funcionalistas já realizadas, que apontam claramente para uma direcionalidade da mudança, tipicamente marcada, por um lado, por diminuição de composicionalidade sintática e semântica e, por outro lado, por aumento de produtividade e de esquematicidade.

Outro trabalho que igualmente nos ajuda a pensar nas relações de construcionalidade é a tese de doutoramento de Paula (2021). A autora analisou a construcionalização de [[p(a)ra lá de] [X]] no português, como indicador de grau intensivo. Tomando como base um trabalho análogo desenvolvido por Venâncio (2015) na sincronia, a autora fez uma incursão na história do português em busca da gênese e do desenvolvimento dessa construção, que pode ser ilustrada por meio do seguinte dado:

(04) Em entrevista recente, Sharon Osbourne foi **para lá de crítica** com estrelas da cultura pop, alfinetando Justin Bieber, o congressista Anthony Weiner e o talk show The View. Mas ela não parou por aí. Fonte: CdP/ Web Dialectos, Brasil, <http://blog.clubnme.com.br/?p=8332>.

Em (04), observamos que [para lá de X] cumpre a função de intensificador na língua, de modo semelhante ao prototípico “muito”. Contudo, [para lá de X] é de uso mais informal, expressivo e marginal. O “peso” da construção (formada por três elementos, além do nome intensificado), o uso de um elemento originalmente dêitico espacial (“lá”), o recrutamento das preposições em novos arranjos estruturais, a falta de registro dessa estratégia de intensificação nos manuais de referência e outros fatores contribuem para a postulação de que esse é um uso inovador no português.

Mais uma vez destacamos que, em termos teóricos, somente um estudo diacrônico poderia comprovar a mudança em termos absolutos. Nesse sentido, alinhamo-nos a Martelotta (2009, p. 13), quando o autor diz que “há um conjunto de princípios atuando sobre a língua de modo relativamente estável, mas a sua efetivação real depende de uma série de fatores que tornam impossível fazer previsões acerca do destino exato de uma mudança linguística”. Contudo, a construcionalidade, em viés sincrônico, certamente nos credencia a indicar a existência de usos inovadores (em comparação com outros usos), como o intensificador [para lá de X], formado a partir da composição de elementos pré-existentes na língua, afetados por deslizamentos metafóricos.

Paula (2021, p. 114-115) elenca ocorrências distintas que demonstram a convivência de diferentes usos de [para lá de X] no português atual:

- (05) Como resposta, os soldados plebeus retiram-se para o Monte Sacro, **para lá do Aniene**, a cerca de 5 km de Roma (Pisão, citado por Lívio (II, 32), diz que se retiraram para o Aventino). Ali acampam e permanecem alguns dias. Iniciam-se conversações e os patrícios são obrigados a fazer algumas concessões. Fonte: CdP/Web Dialectos, Brasil, <http://www.azpmedia.com/espacohistoria/index.php/parte-iroma-antiga-ate-a-primeira-guerra-punica/-capitulo-viii>
- (06) A música tem simplesmente o poder de nos fazer tocar, dançar, cantar e rir... Os ecos da melodia de quem nos recita poemas e nos leva **para lá da imaginação**... Amar... Aqui... Ali... Além... Sem olhar a quem. [...] Fonte: CdP/Web Dialectos, Portugal, <http://algarve-saibamais.blogspot.com/2010/08/o-valor-de-um-abraco.html>
- (07) A promessa do Dr. Seara deixou-me até com aquele o ar daquele menino com a lágrima ao canto do olho que está naquele quadro muito conhecido e também lindíssimo que se vende em todas feiras. A história é simples e vai já **para lá dos quatro anos**, um médico psiquiatra violou comprovadamente uma doente a quem tratava devido a um quadro depressivo e estava grávida. Fonte: CdP/Web Dialectos, Portugal, <http://atentainquietude.blogspot.com/>
- (08) AH já me ia esquecendo de dizer que a minha decisão de ser FAT apenas teve uma condição, é que derivado a viver em um apartamento, apenas posso acolher gatos e não cães, com muita pena minha, pois sei que existem **pra lá de mil cães** a precisarem de um lar... Sei que estou a ajudar, mesmo só acolhendo gatos (gatas e gatinhos) Fonte: CdP/Web Dialectos, Portugal, <http://louzanimales.blogspot.com>
- (09) Olha, ainda bem que gostaste de a Dr. Organic. Tenho-te a dizer que aos poucos começo a comprar outros produtos da marca e estou **pra lá de satisfeita**. Comecei pelo shampô e amaciador, como



já te tinha contado, agora vou no sabonete para o rosto. Fonte: CdP/Web Dialectos, Portugal, <http://ideiasdebaixodotelhado.blogspot.com/2012/10/cabelo-tratado-e-perfumado.html>

Os dados apresentados por Paula (2021), todos coletados no século XXI, nas variedades brasileira e europeia do português, espelham os usos de [para lá de X] em diferentes funções, aqui respectivamente classificadas como circunstancial locativo concreto (05), circunstancial locativo virtual (06), circunstancial temporal (07), quantificador (08) e intensificador (09).

A hipótese localista (Traugott; Heine, 1991; Barotéo, 2000), bem como diversos outros estudos já empreendidos no âmbito do Funcionalismo, apontam que os usos mais concretos servem como base para usos mais abstratos, que se constituem na língua por meio de deslizamentos metafóricos. Assim, o exame sincrónico da construção em tela, em suas diversas manifestações gradientes na língua, permite a dedução de que os usos locativos são anteriores ao uso temporal que, por sua vez, dá origem às funções quantificadora e intensificadora. Em outros termos, esses usos espelham graus distintos de gramaticalidade na língua em uso nos dias de hoje.

Com relação aos parâmetros de Lehmann (1982[2015]), destacamos que a construção “pra lá de X” com valor intensificador passa por atrição, uma vez que seu sentido é concentrado na ideia de intensificação, o que reduz os muitos usos semânticos das preposições “p(a)ra” e “de” e do locativo “lá”, se tomados como partículas individuais. Também podemos aplicar o parâmetro da obrigatoriedade, tendo em vista que o slot X é comumente preenchido por adjetivos, o que reduz seu escopo de uso. Também está em cena o parâmetro da coalescência, já que as subpartes da construção são integradas, sem possibilidade de alteração em “p(a)ra lá de X”, com exceção da variação entre «para» e «pra».

Nos termos de Hopper (1991), atestamos estratificação, já que “para lá de X” passa a conviver com outros intensificadores como “muito”, “demais” etc. Também há divergência, tendo em vista os usos mais concretos originais e os usos mais abstratos da construção, como demonstrado por meio dos dados (05) a (09). Se pensarmos em *locus* de uso, há também especialização,

já que “para lá de X” é utilizado em contextos mais informais. Há persistência, pois a carga semântica de “lá”, de alguma forma, está metaforizada no uso intensificador (maior distância é concebida como maior intensidade), e também atestamos decategorização. Afinal, “para”, “lá” e “de” migram de suas classes para outra (preposições e locativos para intensificador).

Por fim, nos termos de Himmelmann (2004), a inclusão de “para lá de X” na classe dos intensificadores aponta para uma mudança da classe hospedeira (no caso, categoria dos intensificadores), que se expande atraindo esse novo elemento. Também se dá mudança de contexto sintático, haja vista o rearranjo interno dos constituintes da construção. Por fim, atestamos também mudança de contexto semântico-pragmático, com desbotamentos de sentido e ressemantizações dos constituintes de “para lá de X”.

O estudo diacrônico de Paula (2021), de fato, comprovou que a gradiência de [para lá de X], atestada por meio de diferentes dados na atual sincronia do português, espelha as várias camadas de seus usos do ponto de vista histórico. Assim, a gradiência (na sincronia) reflete a gradualidade (atestada na diacronia).

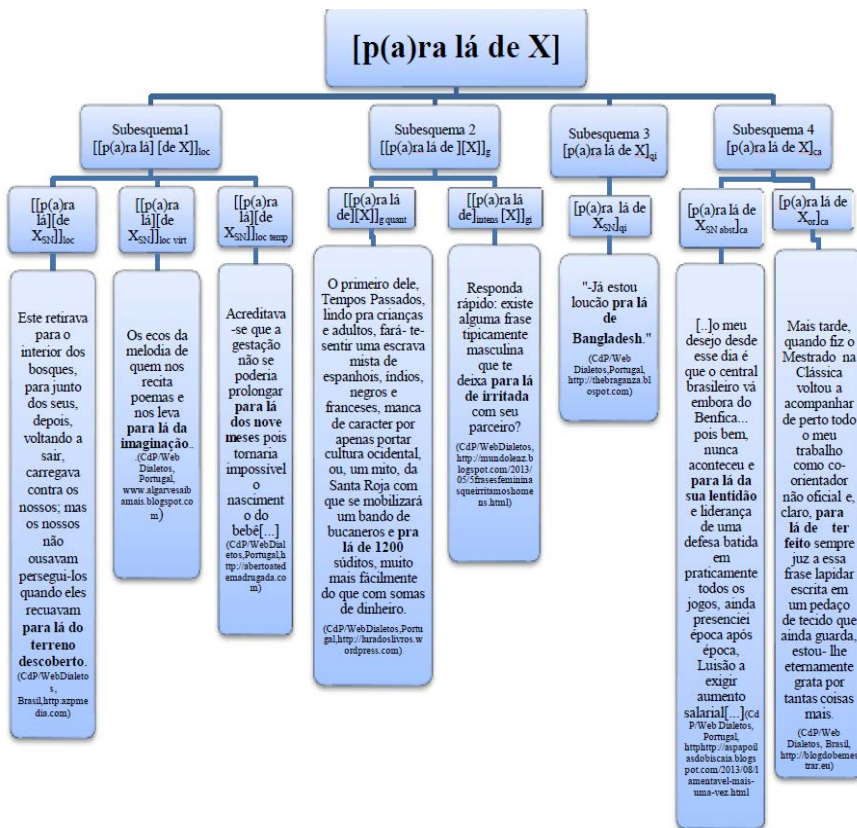
Apoiando-se na abordagem contextual de Diewald (2002, 2006) e de Diewald e Smirnova (2012), Paula (2021) comprovou também que o contexto fonte, ilustrado no dado (05), que expressa uso locativo concreto, é atestado desde o século XIV. O contexto atípico, que aponta para usos locativos abstratos e temporais, como está demonstrado nos dados (06) e (07), surge mais tarde, no século XVII. Os contextos crítico e isolado, por fim, representados pelos dados (08) e (09), respectivamente, são flagrados somente no século XX (sendo o contexto crítico de natureza muito idiossincrática).

Em síntese, os usos mais abstratos (relacionados com o [para lá de X] quantificador e intensificador) são derivados dos usos mais concretos (relacionados a [para lá de X] nos seus sentidos locativos e temporais). O exame da construcionalidade (calcado essencialmente em uma visão sincrônica) já seria capaz de comprovar esse fato. A pesquisa de Paula (2021), contudo, vai além e corrobora o postulado da construcionalidade (apesar de esse conceito não ter sido adotado no trabalho da pesquisadora) com

evidências históricas. Afinal, o estado sincrônico da construção gramatical [para lá de X], devido à sua gradiência, permite a apreensão de traços de sua origem e trajetória, calçada em usos mais concretos dos quais derivam os mais abstratos.

Os muitos usos da construção [para lá de X] são esquematizados por Paula (2021) por meio da seguinte rede:

**Esquema 1 - Rede construcional de [para lá de X] no português do século XXI**



Fonte: Paula (2021, p. 123).

A organização esquemática dos diversos usos da construção [para lá de X] traduz as suas diferentes funções em nossa atual sincronia. A rede serve para refletir claramente as relações de construcionalidade de [para lá

de X], tendo em vista que a disposição imagética do esquema espelha as derivações históricas da construção a partir de um retrato sincrônico de sua realidade. Assim, no subesquema 1, atestam-se os usos locativo concreto, locativo virtual (ou abstrato) e temporal, que são os mais antigos. No subesquema 2, estão os usos de [para lá de X] na função de quantificador e de intensificador. Por fim, a autora avança na investigação e apresenta usos ainda mais abstratos em que a construção cumpre o papel de qualificador intensivo e de conector argumentativo.

Assim, mesmo não adotando o conceito de construcionalidade, as pesquisas desenvolvidas por Acosta (2021) e Paula (2021) ajudam a comprovar a validade e a plausibilidade desse constructo teórico, que pode ser manejado metodologicamente nas pesquisas de língua em uso. Afinal, ambas as pesquisadoras, com fartas evidências, comprovam que o exame sincrônico das construções gramaticais, marcadas por gradiência e por diferentes níveis de gramaticalidade, revela camadas históricas de suas origens e trajetórias. Assim, a construcionalidade dos pares correlatos disjuntivos e da construção [para lá de X] pôde ser demonstrada empiricamente.

### **Considerações finais**

O estudo da mudança linguística é e sempre será um campo bastante vasto e desafiador. Por meio desse estudo, é possível depreender a natureza dos processos cognitivos que moldam as línguas e também compreender um pouco melhor a nossa própria constituição humana mediada pela linguagem.

Na visão funcionalista, a mudança é sempre atestada no uso e é dotada de um caráter altamente complexo. De fato, fatores comunicativos, sociais, cognitivos, estruturais, históricos e outros (cf. Martelotta, 2009) podem impactar as trajetórias históricas de mudança linguística, que nunca deverá ser concebida como um fenômeno puramente linear.

Por outro lado, as fartas evidências científicas na Linguística endossam a plausibilidade de uma abordagem teórica que permita reflexões sobre a gênese e trajetória de construções linguísticas a partir do seu exame

sincrônico. As pesquisas em gramaticalização sincrônica, o princípio do uniformitarismo, os estudos de mudança em tempo aparente, a postulação do conceito de gramaticalidade e as muitas evidências empíricas chancelam o paradigma da construcionalidade, que se apóia sobretudo em Lehmann (1982[2015]), Hopper (1991) e Himmelmann (2004).

Nesse sentido, este capítulo procurou defender o estatuto teórico da construcionalidade, que pode ser definida como a relação sincrônica entre duas ou mais construções, de modo que uma construção pode ser apontada como base para outra(s), a partir de seus diferentes níveis de gradiência e gramaticalidade. Com base nesse conceito, certamente a pesquisa sincrônica pode ganhar ainda mais relevância e abrangência, em termos tanto teóricos como metodológicos. Afinal, uma análise sincrônica que aponte para reflexos pretéritos das construções sempre será mais completa e instigante.

## Referências bibliográficas

ACOSTA, J. M. **Análise funcional das construções correlatas alternativas**. 97f. Dissertação (Mestrado em Estudos de Linguagem) - UFF, Instituto de Letras, Niterói, 2017.

\_\_\_\_\_. **Análise pancrônica das construções correlatas disjuntivas**. 190f. Tese (Doutorado em Estudos de Linguagem) - UFF, Instituto de Letras, Niterói, 2021.

BATORÉO, H. **Expressão do espaço no português europeu**: contributo psicolinguístico para o estudo da linguagem e cognição. Lisboa: Fundação Calouste Gulbenkian, 2000.

BYBEE, J. **Language, usage and cognition**. Cambridge: Cambridge University Press. 2010.

\_\_\_\_\_. **Language Change**. Cambridge: Cambridge University Press. 2015.

CROFT, W. **Radical construction grammar**: syntactic theory in typological perspective. Oxford: Oxford University Press, 2001.

CUNHA, M. A. F. da. A interação sincronia/diacronia no estudo da sintaxe. **Delta**, São Paulo, v. 15, n. 1, p. 1-13, 1999. Disponível em: <[https://www.scielo.br/scielo.php?pid=S0102-44501999000100004&script=sci\\_abstract&ctlng=pt](https://www.scielo.br/scielo.php?pid=S0102-44501999000100004&script=sci_abstract&ctlng=pt)>. Acesso em: 10 maio 2021.

DALL'ORTO, L. F. M.; LACERDA, Patrícia Fabiane Amaral da Cunha. Construcionalização gramatical sincrônica: evidências a partir da análise de construções avaliativas com “super” e “mega” na língua portuguesa. **Revista Soletras**, São Gonçalo - RJ, n. 37, p. 179-203, 2019. Disponível em: <<https://www.e-publicacoes.uerj.br/index.php/soletras/article/view/41946/29214>>. Acesso em: 15 maio 2021.

DIEWALD, G. A model of relevant types of contexts in grammaticalization. In: WISCHER, I; DIEWALD, G (eds). **New reflections on grammaticalization**. Amsterdam/Philadelphia: John Benjamins, 2002. p. 103-120.

\_\_\_\_\_. Context types in grammaticalization as constructions. **Constructions**, Düsseldorf, 2006. Disponível em: < www.constructions-online.de:0009- 4-6860>. Acesso em: 10 nov, 2017.

DIEWALD, G; SMIRNOVA, E. Paradigmatic integration: the fourth stage in an expanded grammaticalization scenario. In: DAVIDSE, K. et al (eds). **Grammaticalization and language change: new reflections**. Amsterdam/Philadelphia: John Benjamins, 2012. p. 111-131.

FAARLUND, J. T. Pragmatics and syntactic change. In: BREIVIK, L. E.; JAHR, E. H. (Orgs.). **Language Change: Contributions to the Study of its Causes**, . (Trends in Linguistics; Studies and Monographs, 43). Berlin/New York: Mouton de Gruyter. p.71-98

GIVÓN, T. Historical syntax and synchronic morphology: an archaeologist's field trip. In: **Regional Meeting of Chicago Linguistic Society**, 7, 1971, Chicago. Papers... Chicago: Chicago Linguistic Society, 1971. p. 394-415.

GONÇALVES, S. C. L. et al. (org.) **Introdução à gramaticalização: em homenagem a Maria Luíza Braga**. São Paulo: Parábola, 2007.

HEINE, B.; KUTEVA, T. **The genesis of grammar: a reconstruction**. Oxford: Oxford University Press, 2017.

HILPERT, M. **Construction Grammar and its Application to English**. Edinburgh: Edinburgh University Press, 2014.

HILMMELMANN, N. Lexicalization and grammaticalization: opposite or orthogonal? In: BISANG, W. et al. (Ed.). **What makes grammaticalization?** Berlin: Mouton de Gruyter, 2004. p.21-42.

HOFFMANN, T.; TROUSDALE, G. Variation, change and constructions in English. In: **Cognitive Linguistics** 22-1-23, 2011. p.1-23.

HOPPER, P.J. On some principles of grammaticization. In: TRAUGOTT, E.; HEINE, B. (Org) **Approaches to grammaticalization**. Vol I. Focus on theoretical and methodological issues, Amsterdam: John Benjamins, 1991. p. 17-35.

LEHMANN, C. Grammaticalization: synchronic variation and diachronic change. In: **Lingua e Stile** 20, n. 3, luglio-settembre, 1985. p. 303-318

LEHMANN, C. **Thoughts on grammaticalization**. 3<sup>rd</sup> edition. Erfurt: Universität Erfurt, 2015[1982].

MARTELOTTA, M. Funcionalismo e metodologia quantitativa. In: OLIVEIRA, M.; ROSÁRIO, I. (org). **Pesquisa em linguística funcional: convergências e divergências**. Rio de Janeiro: Léo Christiano Editorial, 2009. p. 1-20.

ROSÁRIO, I. C. Interview with Graeme Trousdale. **Revista Solettras**, São Gonçalo - RJ, n. 37, p. 10-19, 2019. Disponível em: < <https://www.e-publicacoes.uerj.br/index.php/soletras/article/view/36451>>. Acesso em: 15 maio 2021.

\_\_\_\_\_; LOPES, M. G. Construcionalização gramatical em perspectiva sincrônica. In: **Apresentação em XXII Seminário Nacional e IX Seminário Internacional do Grupo de Estudos Discurso & Gramática**. Niterói: UFF, 2017.

\_\_\_\_\_. Construcionalidade: uma proposta de aplicação sincrônica. **Revista Solettras**, São Gonçalo - RJ, n. 37, p. 83-102, 2019. Disponível em: <<https://www.e-publicacoes.uerj.br/index.php/soletras/article/view/36318>>. Acesso em: 15 maio 2021.

ROSÁRIO, I. C.; OLIVEIRA, M. R. Funcionalismo e abordagem construcional da gramática. **Revista Alfa**, São Paulo, v. 2, n. 60, p. 233-259, 2016. Disponível em: <<http://www.scielo.br/pdf/alfa/v60n2/1981-5794-alfa-60-2-0233.pdf>>. Acesso em: 15 maio 2021.

PAULA, V. B. **A construcionalização de grau intensivo [[p(a)ra lá de] [X]] no português**. 135f. Tese (Doutorado em Estudos de Linguagem) - UFF, Instituto de Letras, Niterói, 2021.

TRAUGOTT, E. C.; HEINE, B. **Approaches to Grammaticalization: Focus on Theoretical and Methodological Issues**. Amsterdam: John Benjamins, 1991.

TRAUGOTT, E.; TROUSDALE, G. **Constructionalization and constructional changes**. Oxford, Oxford University Press, 2013.

TROUSDALE, G. Gradualness in language change: a constructional perspective. In: RAMAT, Anna Giacalone; MAURI, Caterina; MOLINELLI, Piera. **Synchrony and Diachrony: a dynamic interface**. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2013. p. 27-42.

\_\_\_\_\_; TRAUGOTT, E. C. Preface. In: TRAUGOTT, E. C.; TROUSDALE, G. (Eds.). **Gradience, gradualness and grammaticalization**. Amsterdam/Philadelphia: John Benjamins, 2010. p. 1-18.

VENÂNCIO, E. **Instanciações da microconstrução intensificadora “para lá de X” no português contemporâneo**. 178f. Dissertação (Mestrado em Linguística) - UFF, Instituto de Letras, Niterói, 2015.

# 5. Análise Colostrucional

Karen Sampaio Braga Alonso (UFRJ)

Diego Leite de Oliveira (UFRJ)

## **Considerações iniciais**

Neste capítulo, apresentamos a análise colostrucional, uma família de métodos de medida de associação utilizados no campo de estudos linguísticos conhecidos como Modelos Baseados no Uso. Para tanto, o capítulo se organiza da seguinte forma: na primeira seção, tratamos da relação da análise colostrucional com o arcabouço teórico dos Modelos Baseados no Uso, bem como fornecemos algumas das bases operacionais do método; em seguida, aplicamos a análise colostrucional ao estudo de construções binominais quantificadoras do português, mostrando alguns resultados aos quais o método possibilita chegar; por fim, apresentamos uma breve discussão em torno da análise colostrucional, e algumas considerações gerais sobre o tema encerram o capítulo.

## **Modelos Baseados no Uso e análise colostrucional**

Abordagens linguísticas baseadas no uso defendem que a língua, como sistema mental de representação, emerge a partir da interação entre processos cognitivos de domínio geral e de dados da experiência do falante. Nesse sentido, a representação do sistema linguístico tende a ser concebida em termos de redes de padrões associativos convencionalizados, estabelecidos entre forma e sentido, também conhecidos como construções. Tais padrões se organizam hierarquicamente em níveis distintos de abstração, podendo exibir desde uma total especificação fonológica (substantivos) até uma estrutura não especificada fonologicamente ('esquemáticos'). Tendo isso em vista, tomamos, neste capítulo, as construções binominais



quantificadoras para ilustrar a aplicação da família de métodos da análise colostrucional. Vejamos alguns exemplos dessa construção, extraídos do jornal Folha de São Paulo.

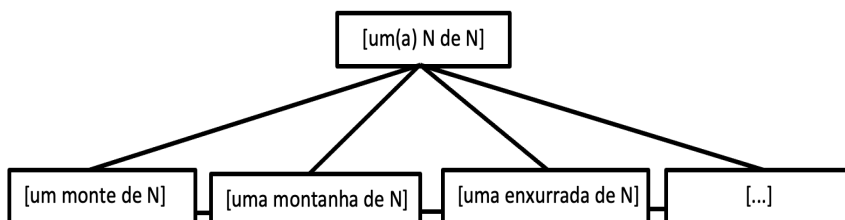
- (1) Minha curiosidade aumentou quando meus pais ganharam **um monte de livros** de um vizinho rico que havia morrido. (Folha de São Paulo 12/04/21)
- (2) Suas conclusões, consubstanciadas em relatório, poderão servir de base para ações de responsabilização, seja no âmbito civil, seja no criminal. Um elemento a mais em **uma montanha de indícios e provas** que só não é maior que a de vítimas. (Folha de São Paulo, 14/04/21)
- (3) Depois que Covas adiantou cinco feriados na capital, as cidades litorâneas decidiram por adotar medidas mais rígidas por receio de **uma enxurrada de turistas**. (Folha de São Paulo 26/03/21)

Nos exemplos de (1) a (3), é possível observar em negrito instâncias da construção esquemática [um(a) N de N] com função quantificadora. Nos casos em questão, a construção é utilizada para expressar grande quantidade de elementos pertencentes a uma dada categoria, sendo possível, em cada um desses casos, substituir a sequência “um(a) N de” pelo vocábulo “muitos/muitas” ou pela expressão “uma grande quantidade de” (muitos/uma grande quantidade de livros; muitos/uma grande quantidade de indícios e provas; muitos/uma grande quantidade de turistas).

Contudo, como falantes do português brasileiro, percebemos que os padrões específicos [um monte de N], [uma montanha de N] e [uma enxurrada de N] parecem expressar nuances distintas de significado atreladas, em maior ou menor grau, à semântica do lexema que ocupa a primeira posição (monte, montanha, enxurrada etc.). Em uma perspectiva construcionista baseada no uso, pode-se postular que tais nós consistem em padrões construcionais fonologicamente mais especificados e que exibem relação taxonômica com a construção esquemática mais geral [um(a) N de N] com sentido de grande quantidade de algo. Além disso, esses nós mantêm relações horizontais entre si, por conta de se encontrarem no mesmo

nível de abstração e por compartilharem aspectos de forma e significado, tal como indicado na figura 5.1, abaixo:

**Figura 1.** Representação da rede de construções [um(a) N de N].



Fonte: Elaboração própria.

Mas esses não são os únicos tipos de relações possíveis na rede apresentada. Ao identificarmos, por exemplo, padrões como [um monte de N], [uma enxurrada de N] ou [uma montanha de N], podemos indagar sobre os possíveis candidatos a ocupar a posição esquemática N nessas construções. Nesses casos, estamos considerando o tipo de relação que Diessel (2019) denomina relação de preenchimento de *slot* (*filler slot relation*), a qual especifica as associações entre lexemas individuais e *slots* específicos de esquemas construcionais.

Para calcular as associações entre lexemas e o *slot* (ou *slots*) específico(s) de uma construção em um dado *corpus*, existem diversos métodos estatísticos, rotulados como medidas de associação (cf. Levshina, 2015, p. 220). O tipo de medida de associação específico a ser apresentado neste capítulo compreende um conjunto de três métodos que, juntos, atendem pelo nome de análise colostrucional (do inglês *collostructional analysis*) e serão apresentados na próxima seção. O termo *collostructional* em inglês refere-se à combinação das palavras *collocation* e *construction*, para dar nome a um conjunto de medidas de associação elaborado por Stefan Gries e Anatol Stefanowitsch (cf. Stefanowitsch; Gries, 2003, 2005 e Gries; Stefanowitsch; 2004). Tais medidas buscam calcular a força de atração (ou repulsa, em alguns casos) dos lexemas em relação a uma ou mais posições esquemáticas de uma dada construção.

Desenvolvida mais especificamente no âmbito do modelo teórico da Gramática de Construções, a análise colostrucional geralmente considera testes estatísticos de independência – no caso, o Teste Exato de Fischer, o qual, além de revelar se a frequência de coocorrência é estatisticamente significativa ou não<sup>1</sup>, ou seja, se ela é ou não mero fruto do acaso, funciona muito bem inclusive com células pequenas, com baixo número de ocorrências. Para desenvolver esse tipo de análise, é aconselhável o uso do software de análises estatísticas R<sup>2</sup>.

O software R consiste em uma linguagem de programação para computação estatística e criação de gráficos, desenvolvida inicialmente por Ross Ihaka e Robert Gentleman, na Universidade de Auckland, Nova Zelândia, e atualmente incluindo um amplo grupo de colaboradores, dentre os quais John Chamber, criador da linguagem de programação S, na qual o R foi baseado. Trata-se de um software gratuito, que pode ser obtido em <https://cran.r-project.org/mirrors.html><sup>3</sup>.

Para que seja possível desenvolver a análise colostrucional, é necessário que se escolha um *corpus* que permita a obtenção das frequências necessárias (quanto à construção e aos lexemas investigados, assim como em relação ao número de palavras do *corpus*) e organizá-las em uma tabela cruzada 2X2, de acordo com o formato adequado para cada tipo específico de análise. Para ilustrar as análises a serem apresentadas na próxima seção, escolhemos o *corpus* CHAVE, que pode ser obtido integral e gratuitamente, mediante cadastro e solicitação no site <https://www.linguateca.pt>. O *corpus* compreende textos jornalísticos dos jornais Público e Folha de São Paulo dos anos de 1994 e 1995, perfazendo um total de 90.646.837 palavras. Aqui, selecionamos apenas os textos do jornal Folha de São Paulo, que consistem em 35.699.765 palavras. Na próxima seção,

---

<sup>1</sup> Geralmente a análise colostrucional adota um padrão, que considera o seguinte resultado em termos de relevância estatística: Se a força colostrucional é maior que 3  $\rightarrow p < 0.001$ ; se a força colostrucional é maior que 2  $\rightarrow p < 0.01$ ; se a força colostrucional é maior que 1.3  $\rightarrow p < 0.05$ .

<sup>2</sup> Mais informações podem ser obtidas em [www.r-project.org/contributors.html](http://www.r-project.org/contributors.html)

<sup>3</sup> Uma versão mais amigável para o usuário, a RStudio, pode ser obtida em [www.rstudio.com/products/rstudio/download/](http://www.rstudio.com/products/rstudio/download/). Para utilizá-la, é necessário já ter instalado o R. Para ter um panorama amplo sobre métodos estatísticos e formas de utilização do R para análises linguísticas, conferir Levshina (2015).

especificamos e discorremos sobre cada um dos métodos que compõem a análise colostrucional.

## **Análise colostrucional: o caso das construções binominais quantificadoras**

### *Análise colexêmica simples*

A análise colexêmica simples pode ser utilizada quando o pesquisador está interessado em observar as relações entre lexemas e o *slot* específico de apenas uma construção. Neste capítulo, vamos ilustrar a aplicação desse método, ao nos debruçarmos sobre o *slot* N da construção [um monte de N] em português.

Para começar a análise, coletamos todas as ocorrências da construção [um monte de N] do *corpus* selecionado (296, ao todo) e observamos todas as combinações possíveis entre lexemas e o *slot* N da construção. Na tabela 1, a seguir, apresentamos os três lexemas mais recorrentes, com sua frequência total no *corpus* e sua frequência na construção.

**Tabela 1:** Relação dos cinco lexemas mais frequentes na construção [um monte de N].

N	Lexema	Frequência no <i>corpus</i>	Frequência na construção
1	Gente	5323	46
2	Coisa	17475	38
3	Dinheiro	14177	7

Fonte: Adaptado de Alonso et al (2020).

É importante salientar que, em alguns casos específicos, será necessário executar uma filtragem dos dados. Diante da análise da construção [um monte de N], por exemplo, podemos nos deparar com sequências que não consistem em instâncias da construção, quando lidamos com os dados concretos. É preciso diferenciar instâncias como *um monte de gente interessante*, que de fato consiste em uma instância da construção [um monte de N] com valor quantitativo, de instâncias como *um monte de Telluride com mil metros de altura*, que nitidamente não representa uma instância da construção em pauta. Outro ponto importante refere-se à necessidade de

filtrar alguns lexemas específicos em relação à frequência total no *corpus*. Com relação ao estudo de [um monte de N], deparamo-nos com o lexema “gente”, cujo valor referencial pode estar associado a uma pessoa ou conjunto de pessoas, mas também pode se referir à forma de primeira pessoa do plural “a gente”. Nesse caso, foram consideradas apenas manifestações de “gente” relativas ao primeiro tipo. Na tabela, é possível observar que o lexema “gente”, portanto, ocorre 5323 vezes no *corpus*, sendo que, dessas ocorrências, 46 se dão na construção [um monte de N].

Após a coleta de todas as instâncias de manifestação da construção, de todos os lexemas que nela ocorrem e da frequência de cada um desses lexemas como um todo no *corpus*, é necessário gravar a tabela em um documento no formato .txt e processar os dados com auxílio do software R. Será necessário rodar os dados, por meio de um script disponibilizado por Stefan Gries, em <http://www.stgries.info/teaching/groningen/readme.txt><sup>4</sup>. O script oferece um passo a passo ao usuário, solicitando que insira respostas a perguntas sobre o tipo de análise a ser realizado (colexêmica, colexêmica distintiva ou colexêmica covariacional), o número de palavras no *corpus*, entre outras, permitindo que, ao final, seja anexada a tabela supracitada.

Tendo inserido o arquivo, o usuário deve decidir se deseja ter os dados da análise salvos no arquivo em .txt ou se deseja visualizar o resultado diretamente no terminal do programa R. Após atender aos requisitos do script, os cálculos são gerados automaticamente. A tabela abaixo, adaptada de Alonso et al. (2020), apresenta os três primeiros resultados com as maiores forças colostrucionais:

**Tabela 2:** Análise colexêmica da construção [um monte de N].

Lexema	Frequência no <i>Corpus</i>	Frequência na Construção	Frequência esperada	Força Colostrucional
Gente	5323	46	0,04	121,78
Coisa_coisas	17475	38	0,14	77,74
Dinheiro	14177	7	0,11	10,28

FONTE: Autores.

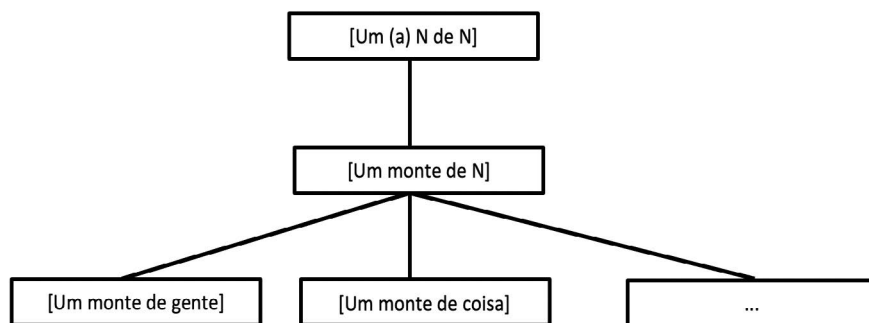
<sup>4</sup> Nesse link, além de informações sobre o script em si, há informações sobre como organizar a tabela em formato .txt para cada um dos tipos de análise sugeridos.

Na tabela, é possível observar a frequência de cada lexema coletado como um todo no *corpus* e sua frequência na construção. Considerando esses dados, o programa R calcula qual seria a frequência esperada para aquele lexema na construção, se não houvesse qualquer força de associação entre o lexema e a construção, ou seja, se a combinação entre o lexema e a construção fosse o mero fruto do acaso. Tomando-se a relação entre a frequência do lexema observada na construção e a frequência esperada para o lexema, o software calcula a força colostrucional (força de associação) entre o lexema e a construção. Quanto maior o número gerado, maior a força colostrucional. Considerando o lexema “gente”, podemos observar que sua frequência esperada na construção seria 0,04, correspondendo a bem menos do que uma ocorrência. No entanto, é possível observar que “gente” ocorreu 46 vezes na construção, o que equivale a mais de mil vezes o esperado. Isso gera uma força colostrucional de 121,78, a maior em nossa análise. Por outro lado, considerando a frequência do lexema “dinheiro” no *corpus* e sua frequência na construção, o programa calculou uma frequência esperada de 0,11 na construção, o que também equivale a bem menos de uma ocorrência, mas uma frequência ainda assim maior do que a esperada para o lexema “gente” - daí a diferença expressiva entre a força colostrucional de ambos os lexemas, o que nos informa que a força de atração entre [um monte de N] e “gente” é maior do que a de [um monte de N] e “dinheiro”.

Os resultados da análise colostrucional da construção [um monte de N] sozinhos não nos dizem muita coisa. Como afirma Hilpert (2014), é tarefa do pesquisador atribuir sentido aos dados. Assim, é preciso analisar qualitativamente os resultados. Observando o resultado completo, que não está sendo apresentado aqui por questões de espaço, percebemos que um total de 162 lexemas distintos ocorreram em combinação com a construção [um monte de N]. Isso permite inferir que a construção [um monte de N] exibe frequência *type* considerável, incluindo lexemas de natureza diversificada, o que pode estar indicando um grau elevado de abrangência da construção, corroborando a análise diacrônica de Fumaux (2018), que investigou o desenvolvimento da construção [um monte de N]. Considere-se também que os lexemas mais frequentes na construção, a saber “gente” e “coisa”, se destacam dos demais, exibindo força colostrucional substancialmente maior. Com isso, a análise

colexêmica simples nos permite postular a seguinte representação para a rede, no que diz respeito a [um monte de N]:

**Figura 2.** Representação de [um monte de gente] e [um monte de coisa(s)] na rede.



FONTE: Elaboração própria.

A figura 2 nos informa que a construção [um monte de N] ganha em esquematicidade, assumindo posição intermediária na rede e licenciando construções mais específicas, como [um monte de gente] e [um monte de coisa(s)]. Para além disso, os resultados da análise colexêmica simples indicam uma alta previsibilidade de ocorrência de sequências sintagmáticas como “um monte de gente” e “um monte de coisa(s)” e permitem, portanto, levantar a hipótese de que tais sequências possam assumir certo grau de automatização na língua, a despeito de serem previsíveis pelo nó mais abstrato [um monte de N]<sup>5</sup>.

Análises que considerem um banco de dados mais diversificado em termos de gêneros textuais e modalidade oferecerão resultados mais robustos e confiáveis. Tal hipótese poderia ser, adicionalmente, objeto de estudo e de possível confirmação, com base em métodos complementares como experimentos psicolinguísticos – considerando tempo de processamento de sequências sintagmáticas mais e menos frequentes – ou análises acústicas com base em produção e percepção, métodos que não são objeto do presente capítulo.

<sup>5</sup> Conferir Goldberg (2019) e Bybee (2010) para uma discussão mais aprofundada sobre o impacto da frequência na representação.

### *Análise colexêmica distintiva.*

A análise colexêmica distintiva pode ser utilizada para comparar a força de atração de lexemas em relação a um *slot* de duas ou mais construções, geralmente exibindo algum grau de alternância. É comum observar sua aplicação à investigação de fenômenos como a possível alternância entre construções bitransitivas e de dativo com *to* em inglês, dentre outros (Gries e Stefanowisch, 2004). Formas alternantes podem ser compreendidas em algumas perspectivas de análise como sinônimas, mas em investigações baseadas no uso, diferenças de forma entre construções são entendidas como acarretando alguma diferença de significado e, nesse sentido, a análise contrastiva das preferências colocacionais de construções alternantes pode contribuir para a identificação das diferenças de sentido.

Nesta seção, aplicamos a análise colexêmica distintiva ao estudo das construções [um monte de N] e [uma enxurrada de N], com vistas a entender um pouco mais sobre as diferenças de significado entre essas construções. Diferentemente da análise colexêmica simples, a análise colexêmica distintiva lança luz sobre as relações horizontais entre essas construções. Uma análise mais detalhada pode ser conferida em Alonso et al. (2019).

Tendo coletado todas as ocorrências de ambas as construções, o que será necessário para calcular a força colostrucional na análise colexêmica distintiva é a frequência observada dos lexemas que se combinam com essas construções. A tabela 3 a seguir apresenta os três lexemas que ocorrem com mais frequência nas construções analisadas.

**Tabela 3.** Comparação da frequência de lexemas nas construções [um monte de N] e [uma enxurrada de N].

Lexema	Construção [um monte de N]	Construção [uma enxurrada de N]
Gente	46	1
Coisa_coisas	38	0
Dólares	1	21

Fonte: Elaboração própria.



Para realizar a análise colexêmica distintiva, é necessário utilizar o mesmo script da análise colexêmica simples, bastando seguir as instruções que são dadas diante da operacionalização do programa R (ou RStudio). Vejamos abaixo as tabelas 4 e 5, que apresentam o resultado da análise colexêmica distintiva para [um monte de N] e [uma enxurrada de N] respectivamente, considerando os três lexemas que apresentaram a maior força colostrucional para cada construção:

**Tabela 4.** Lexemas mais atraídos pela construção [um monte de N] em comparação a [uma enxurrada de N].

Lexema	Freq. observada enxurrada	Freq. observada monte	Freq. esperada enxurrada	Freq. esperada monte	Força colostrucional
Gente	1	46	14,94	32,05	6,93
Coisa_coisas	0	38	12,08	25,91	6,66
Amigos	0	5	1,58	3,41	0,83

Fonte: Elaboração própria.

**Tabela 5.** Lexemas mais atraídos pela construção [uma enxurrada de N] em comparação a [um monte de N].

Lexema	Freq. observada enxurrada	Freq. observada monte	Freq. esperada enxurrada	Freq. esperada monte	Força colostrucional
Dólares	21	1	6,99	15	9,72
Ações	7	0	2,22	4,77	3,52
Dinheiro_ estrangeiro	7	0	2,22	4,77	3,52

Fonte: Elaboração própria.

O que se pode observar, contrastando as tabelas apresentadas, é que as construções [um monte de N] e [uma enxurrada de N], no *corpus* analisado, exibem preferências colocacionais distintas. Diferentemente de [um monte de N] – que, além atrair um número maior de lexemas para ocupar o *slot* N, parece preferir referentes de sentido mais genérico (como ‘gente’

e ‘coisa’) –, a construção [uma enxurrada de N] parece atrair lexemas cujos referentes são compatíveis com um *frame* de deslocamento, por exemplo, indicando que essa construção apresenta uma maior restrição em relação aos elementos atraídos por ela, se comparada à construção [um monte de N], como será discutido na terceira seção deste capítulo.

### *Análise colexêmica covariacional*

A análise colexêmica covariacional pode ser utilizada por pesquisadores que desejem analisar o comportamento de dois *slots* em uma única construção. No caso específico apresentado aqui, a análise colexêmica covariacional pode ser utilizada quando voltamos nosso olhar para o nó mais abstrato [um(a) N de N] das construções quantificadoras de grande quantidade em português. Nesse caso, especificamente, podemos analisar quais são os elementos que ocupam a primeira posição de N em um dado *corpus* e como eles se combinam com os elementos que ocupam a segunda posição de N.

Como a construção em pauta parece ser bastante produtiva em termos de extensibilidade, ou seja, permite um amplo arcabouço de candidatos a ocuparem ambas as posições de N, restringimo-nos, por conta de espaço e tempo, a buscar, na primeira posição de N, apenas lexemas com semântica relacionada a elementos da natureza como um domínio conceptual relevante para as estratégias de quantificação (relevo, vegetação, fenômenos naturais etc.). Ao todo, no *corpus* analisado, foram encontrados 10 lexemas com maior ou menor expressividade na primeira posição de N (monte, montanha, mar, rio, oceano, nuvem, selva, chuva, enxurrada, avalanche), combinados com um amplo conjunto de lexemas na segunda posição de N. A busca no *corpus* retornou 559 ocorrências.

Após a coleta de dados, o procedimento para realização do cálculo da análise colexêmica covariacional seguiu os mesmos parâmetros das análises anteriores a partir do mesmo script já mencionado. Abaixo, na tabela 6, apresentamos as cinco combinações que apresentaram a maior força colostrucional.

**Tabela 6.** Principais colexemas covariacionais da construção [um(a) N de N] no *corpus* CHAVE.

	Frequência da combinação observada no <i>corpus</i>	Frequência da combinação esperada no <i>corpus</i>	Força Colostrucional
um monte de coisa	38	20,12	11,01
um monte de gente	46	25,42	10,99
uma enxurrada de dólares	21	6,56	8,78
uma enxurrada de ações judiciais	8	2,52	3,41
uma chuva de críticas	2	0,07	2,96

Fonte: Elaboração própria.

Os dados da tabela 6 permitem observar que as combinações “um monte de coisa” e “um monte de gente” foram as mais frequentes no *corpus* e exibem força colostrucional semelhante. Ainda que tenha exibido frequência maior, “um monte de gente” apresentou força colostrucional ligeiramente menor que a combinação “um monte de coisa”, pois no *corpus* o lexema “coisa” foi encontrado exclusivamente em combinação com “monte”, ao passo que o lexema “gente” também ocorreu em combinações como “uma enxurrada de gente” e “uma avalanche de gente”, impactando a distribuição da força colostrucional. Observamos também que a combinação “uma enxurrada de dólares” no *corpus* foi relativamente frequente, exibindo força colostrucional próxima de um “um monte de gente” e “um monte de coisa”. Os resultados da análise colexêmica covariacional, como um todo, permitem corroborar aspectos como a maior extensibilidade, generalidade e conseqüente produtividade da construção [um monte de N], em contraposição aos demais possíveis nós específicos investigados.

## Discussão

Como se viu, a análise colostrucional parte do princípio básico de que a forma como os elementos se distribuem na língua pode nos oferecer

informações valiosas sobre suas propriedades morfosintáticas, semânticas e pragmáticas (Gries, 2012). Nesse sentido, entende-se que, a partir do cálculo da força de atração entre lexemas e construções, é possível obter informações acerca do sentido mais central da construção, uma vez que conceitos que coocorrem sistematicamente tendem a estar próximos cognitivamente; sendo assim, lexemas com alta força colostrucional para ocorrer em um *slot* de uma dada construção (cf. análise colexêmica) contribuem para o entendimento do sentido geral da construção.

A análise colostrucional é vantajosa em relação à análise de frequência bruta, uma vez que uma frequência bruta alta de um lexema em um *slot* de uma dada construção pode se dar em função de fatores adversos que, por vezes, tangenciam os resultados. Por exemplo, pode ser que um determinado lexema seja extremamente frequente no *corpus* como um todo, de modo que sua recorrência no *slot* de uma dada construção pode ser tomada como um epifenômeno. Dessa forma, considerar a força de atração, como já demonstrado anteriormente, oferece um olhar privilegiado e mais acurado acerca das motivações que levam dois ou mais lexemas a coocorrerem dentro de uma mesma construção.

Críticos da análise colostrucional argumentam, por vezes, que os resultados advindos de uma análise estatística que não se apoia em fatores semânticos não podem levar a conclusões sobre o sentido das construções. De fato, a análise colostrucional oferece como resultado estatístico um ranqueamento dos lexemas que ocorrem em um ou mais *slots* de uma dada construção, de acordo com a força colostrucional que apresentam, sem que, para isso, tenham sido requeridas informações de ordem semântica. Por outro lado, uma vez que aceitamos o pressuposto de que construções e lexemas são mutuamente atraídos por motivações de ordem semântico-pragmática, é razoável admitir que propriedades de sentido associadas aos lexemas mais fortemente atraídos para uma construção podem ser tomadas como um conjunto de evidências acerca do sentido geral da própria construção.

Os resultados oferecidos pelo programa R fornecem diferentes informações para o analista. Ainda assim, é preciso tomar alguns cuidados, pois não se pode simplesmente descrever o ranqueamento dos lexemas com

maior força colostrucional e entender esse resultado como um indicador suficiente. O analista deve ter um olhar crítico e mais amplo sobre as informações geradas pelo algoritmo no qual a análise se baseou, com a devida atenção a diferentes pontos. Algumas vezes, lexemas apresentam uma alta força colostrucional, porque ocorreram apenas uma vez no *corpus* e essa única vez ocorreu na construção. Para o algoritmo, isso pode gerar uma força de atração entre lexema e construção relativamente alta. Entretanto, essa força de atração deve ser tomada com desconfiança pelo analista, uma vez que a relação entre lexema e construção não é estatisticamente consistente e pode ter sido gerada ao acaso. Isso significa que todas as frequências geradas pelo algoritmo devem ser levadas em conta na análise.

Certo é que, como cientistas, devemos sempre desconfiar de tudo que pareça antinatural, estranho e devemos sempre olhar os fenômenos e os resultados de forma mais ampliada. Observe, por exemplo, o caso da construção [uma enxurrada de N]. Em pesquisa anterior, Alonso et al (2019) demonstraram que o lexema *dólares* (e outros como *dinheiro*, *dinheiro estrangeiro*) apresentou uma alta força colostrucional para a construção [uma enxurrada de N]. Ao ler esses resultados, os pesquisadores não concluíram que [uma enxurrada de N] está relacionado a um sentido de quantificar dinheiro ou algo do tipo, mas atribuíram esses resultados mais altos ao *corpus* analisado, que, no caso, compunha-se de textos do jornal Folha de São Paulo. Os dados foram encontrados, sobretudo, nas seções de economia. Além disso, a comparação de *dólares* com outros lexemas que também se combinaram com [uma enxurrada de N] e apresentaram força colostrucional importante levaram os pesquisadores a concluir que o sentido de transferência, por exemplo, associado ao fenômeno natural da enxurrada, é mais determinante em relação ao grupo de lexemas que são atraídos pela construção do que uma associação mais direta à ideia de moeda, dinheiro. Ou seja, uma das formas como os falantes usariam a construção [uma enxurrada de N] é a quantificação de grande quantidade de referentes que podem ser transferidos. Poderíamos inferir, a partir daí, que, do conceito mais complexo de enxurrada, perspectiviza-se o sentido de deslocamento nos dados analisados.

Muitas vezes, observamos que os fatores que levam lexemas a serem fortemente atraídos para o *slot* de uma dada construção são motivados por fatores como, por exemplo, o *corpus* selecionado para análise. Também devemos levar em conta que a gramática se molda a partir de uma combinação de construções e que essa combinação não se dá por mero acaso. Por exemplo, [uma enxurrada de N] frequentemente recruta lexemas que são transferíveis, de um lado, e, de outro, ocorre frequentemente combinada com verbos de movimento, como *entrar* e *sair* – *entrou/saiu uma enxurrada de dólares do país*. Não dá para negar a importância de todo esse entorno numa perspectiva teórica que prevê uma gramática que se organiza em rede. Da mesma forma, não é de se esperar que o analista aposte todas as suas fichas no resultado estatístico, deixando de lado a análise crítica e qualitativa dos dados.

Considerando que análises colostrucionais lançam luz sobre a força de atração, aqueles elementos com maior força são frequentemente privilegiados na hora de os resultados serem expostos em tabelas, por exemplo – geralmente opta-se por uma ordenação baseada na força colostrucional, da maior para a menor. Embora o algoritmo gere resultados de força colostrucional para todos os dados, bem como as demais frequências, não é incomum que, ao utilizar o método em questão, o analista dê pouca atenção à análise de lexemas que apresentam força colostrucional baixa – já que dizem pouco sobre o sentido mais central da construção.

Tendo isso em vista, alguns autores mais críticos ao método apontam que análises colostrucionais, ao desvalorizarem, em certa medida, as combinações mais periféricas, prejudicam a percepção da gradiência da construção, uma vez que dados mais periféricos tendem a mostrar o alcance da construção, bem como as tendências de mudança. Isso porque, como frequentemente se observa nas análises funcionalistas, novas construções surgem de casos mais periféricos de construções pré-existentes, de forma que eles vão indicando, de um lado, o alcance da construção existente e, de outro, os limites entre ela e o surgimento de uma nova construção. Ressalta-se que, embora relações de atração sejam o foco da análise colostrucional, todos os resultados, incluindo os dados de repulsa,

podem ser extraídos das estatísticas geradas pelo experimento. Embora as informações estatísticas possam ser limitadoras em relação a uma análise qualitativa mais apurada, elas dão bons indícios do comportamento de uma construção num dado *corpus*. Ainda assim, entendemos que melhores resultados são obtidos combinando análises quantitativas com análises qualitativas.

Sobre a adoção de metodologias quantitativas, é importante se ter em vista que, quanto mais robusto for o *corpus*, em termos de volume, diversidade e equilíbrio, mais consistente será a descrição do fenômeno investigado. Análises colostrucionais usam dados obtidos geralmente via sistemas de busca em *corpora* anotados, construídos dentro das premissas da Linguística de *Corpus* - o que vem contribuindo bastante para os estudos linguísticos que se pautam em frequência, tendo em vista que é possível obter uma quantidade muito grande de dados a partir da automatização das buscas.

Se, de um lado, a tarefa hercúlea de procurar dados em *corpus* foi imensamente facilitada, ampliada e potencializada com a interface entre Linguística e Computação, algumas questões ainda permanecem. Mais acima no texto, por exemplo, falamos do caso do lexema “gente”, que aparece tanto em construções como “um monte de gente” como em construções como “a gente”. Para avaliar a força de atração entre “um monte de N” e “gente”, precisamos retirar dos dados as ocorrências de “a gente”. Nesse caso, especificamente, foi possível subtrair os casos pronominais dentre os milhares de usos de “gente” no *corpus*, porque havia um índice formal para que a ferramenta filtrasse os dados que interessavam à análise colostrucional.

Entretanto, nem sempre isso é possível. Muitas vezes lexemas homógrafos, como o caso de *ações* (que pode se referir a ações da bolsa de valores, a ações judiciais, a atos, atitudes etc.) são tratados indistintamente na análise colostrucional, porque não é possível olhar cada um dos milhares de dados desse lexema no *corpus* para distinguir quais se encaixam no sentido ‘a’, ‘b’ ou ‘c’. Então, essa diferença não é feita, mesmo que cada sentido tenda a se distribuir na língua de forma bem diferente - o que poderia interferir

no resultado da análise. Esse é um dos problemas que a análise enfrenta e ao qual o analista precisa estar atento, para que possa fazer ponderações ou aprofundamentos sempre que necessário.

### **Considerações finais**

Neste capítulo, buscamos tratar do conjunto de métodos conhecido em abordagens linguísticas baseadas no uso como Análise Colostrucional. O primeiro método apresentado foi a análise colexêmica simples, que observa a força de atração de lexemas em relação a um único *slot* em uma dada construção. Na rede de construções, esse método permite observações acerca do nível de abrangência da construção investigada, sua extensibilidade e esquematicidade. O segundo método exposto neste capítulo foi a análise colexêmica distintiva, que permite ao pesquisador contrastar construções em um mesmo nível de abstração, observando possíveis diferenças entre elas, no que diz respeito aos lexemas atraídos e/ou repelidos para um de seus *slots*. O método é interessante para se buscar diferenças em nuances de sentido entre construções que mantêm algum tipo de relação horizontal, por compartilharem similaridades formais e/ou semântico-pragmáticas. O terceiro e último método apresentado foi a análise colexêmica covariacional, que permite a investigação de dois *slots* dentro de uma mesma construção. No caso tratado neste capítulo, foi possível verificar que tal método permitiu a análise de uma construção ainda mais abstrata (com dois *slots* abertos) na rede de construções binominais quantificadoras, assegurando ao pesquisador, juntamente com os outros dois métodos mencionados, um retrato interessante sobre a organização e o funcionamento dessa rede de construções no português brasileiro. Esperamos que o presente capítulo tenha esclarecido os pontos básicos da análise colostrucional e que sirva de incentivo para que mais pesquisadores lancem mão do método, no caso de ele atender às demandas suscitadas pela pesquisa.



## Referências bibliográficas

ALONSO, K. S. B.; OLIVEIRA, D. L. de; FUMAUX, N. C. A. Construções binominais quantitativas em perspectiva distintiva: uma análise colostrucional. **Odisseia**, Natal, RN, v. 4, n. esp., p. 173-193, jul.-dez. 2019.

ALONSO, K. S. B. et al. Quantifying binominal constructions in portuguese and russian: the case of um monte de NP and kucha NPgen Work. **Pap. Linguíst.**, Florianópolis, v.21, n.1, p. 75-101, jan./jul., 2020.

BYBEE, J. **Language, usage and cognition**. Cambridge: Cambridge University Press, 2010.

DIESSEL, H. **The Grammar Network** - How Language structure is shaped by language use. New York: Cambridge University Press, 2019.

FUMAUX, N. C. A. **Construcionalização de 'um monte de SN'**: uma abordagem centrada no uso. 127f. Dissertação (Mestrado em Linguística). Universidade Federal do Rio de Janeiro, Programa de Pós-graduação em Linguística, 2018.

GOLDBERG, A. E. **Explain me this**: creativity, competition and the partial productivity of constructions. Princeton: Princeton University Press, 2019.

GRIES, S. Frequencies, probabilities, and association measures in usage-/ exemplar-based linguistics. **Studies in Language**, v.11, n.3, p.477-510, 2012.

GRIES, S.; STEFANOWITSCH, A. Extending colostrucional analysis. A corpus based perspective on alternations. **International Journal of Corpus Linguistics**, v.9, n.1, p.97-129, 2004.

HILPERT, M. Colostrucional analysis. Measuring associations between constructions and lexical elements. In: GLYNN, D.; ROBINSON, J. A. (ed.) **Corpus Methods for Semantics**. Quantitative studies in polysemy and synonymy. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2014. p. 391-404.

LEVSHINA, N. **How to do Linguistics with R**. Amsterdam: John Benjamins. 2015.

STEFANOWITSCH, A.; GRIES, S. Colostructions: Investigating the interaction between words and constructions. **International Journal of Corpus Linguistics**, v.8, n.2. p.209-243, 2003.

STEFANOWITSCH, A.; GRIES, S. Covarying collexemes. **Corpus Linguistics and Linguistic Theory**, v.1, n.1. p.1-43, 2005.

## 6. Uso do software *Antconc* na análise de dados do uso

Patrícia Fabiane Amaral da Cunha Lacerda (UFJF)

Lauriê Ferreira Martins Dall’Orto (UFJF)

### Considerações iniciais

O propósito deste capítulo é apresentar, de modo prático e pontual, de que maneira o programa *Antconc*<sup>1</sup>, utilizado no âmbito da Linguística de *Corpus*, pode contribuir, de modo substancial, para as pesquisas de natureza funcionalista – principalmente, para as que são realizadas no âmbito da Linguística Funcional Centrada no Uso – doravante, LFCU –, tal como concebida por Cunha *et al.* (2013), Cunha (2016), Rosário e Oliveira (2016) e Bispo e Silva (2016).

A preocupação com questões de natureza metodológica, nas pesquisas de cunho funcionalista, já está presente, por exemplo, em Martelotta (2009). Como destaca o autor, a busca pela aplicação da metodologia quantitativa, na análise funcional de fatos linguísticos, pode acabar levando muitos pesquisadores a empregar – implícita ou explicitamente – a metodologia quantitativa variacionista. Entretanto, como ressalta Martelotta (2009, p. 1), “como nem sempre esses fatos linguísticos constituem fenômenos variáveis, a metodologia corre o risco de resultar inadequada”.

Nesse sentido, levando em consideração a importância de uma metodologia adequada para o tratamento dos dados sob uma perspectiva funcionalista, este capítulo visa a demonstrar em que medida a Linguística de *Corpus* – ao operar a partir do tratamento da língua em uso e a partir do equacionamento entre a análise qualitativa e a análise quantitativa – pode contribuir com as pesquisas desenvolvidas no contexto da LFCU.

---

<sup>1</sup> O download do programa *Antconc* pode ser realizado gratuitamente no seguinte site: <https://www.laurenceanthony.net/software/antconc/>

Desse modo, a fim de cumprir o objetivo a que se propõe este capítulo, nós o organizamos da seguinte maneira: em um primeiro momento, discutimos por que, de fato, a Linguística de *Corpus* pode contribuir sistematicamente para o refinamento dos pressupostos metodológicos da LFCU; em um segundo momento, demonstramos, de modo prático, as ferramentas que constituem o programa *Antconc* e que, quando aplicadas no âmbito de pesquisas realizadas no contexto da LFCU, podem trazer importantes evidências empíricas de um ponto de vista tanto qualitativo como quantitativo; por fim, em um terceiro momento, encaminhamos algumas conclusões a partir das discussões e das evidências apresentadas nas duas seções anteriores.

### **Tema central do capítulo**

A Linguística de *Corpus* ocupa-se da coleta e da exploração de *corpora* que foram coletados criteriosamente com o propósito de servirem para a pesquisa de uma língua ou de uma variedade linguística. Como tal, dedica-se à exploração da língua através de evidências empíricas, extraídas por meio de computador (Berber Sardinha, 2000, 2004; Novodvorski; Finatto, 2014). O *corpus*, no contexto da Linguística de *Corpus*, é concebido como uma coletânea de textos em formato eletrônico, que é compilada segundo critérios específicos e considerada representativa de uma língua ou variedade linguística.

Como destaca Berber Sardinha (2000, 2004), quatro pré-requisitos são fundamentais para a formação de um *corpus* computadorizado, a saber: i) o *corpus* deve ser composto por textos autênticos, ou seja, os textos não podem ter sido produzidos com o propósito de serem alvo de pesquisa linguística; ii) o *corpus* deve ser constituído por textos produzidos por falantes nativos de determinada língua ou variedade linguística; iii) o *corpus* deve ser escolhido criteriosamente, e os textos que o compõem devem seguir, acima de tudo, as condições de naturalidade e autenticidade; e iv) o *corpus* deve, de fato, ser representativo da língua ou variedade linguística que representa.

Do ponto de vista metodológico, a Linguística de *Corpus*, além de se apoiar em *corpora* disponíveis on-line, conta com *softwares* cuja função

é processar um grande volume de textos a partir de algumas ferramentas específicas de processamento. Neste texto, apresentamos, na próxima seção, de forma sucinta, o programa *Antconc*, que constitui uma ferramenta com acesso gratuito e com versões para diferentes sistemas operacionais<sup>2</sup>.

Nesse sentido, como ilustraremos na próxima seção, a grande vantagem do programa *Antconc*, para as pesquisas desenvolvidas no âmbito do Funcionalismo – e, especialmente, no contexto da LFCU –, é justamente fornecer resultados que podem ser interpretados tanto qualitativamente como quantitativamente.

Nesse contexto, como destaca Lacerda (2016), caberia, principalmente, a uma análise qualitativa de dados, no âmbito da LFCU, as seguintes funções: i) caracterizar, por meio de ocorrências empiricamente atestadas, o pareamento entre forma e função no nível da microconstrução, do subesquema e do esquema; ii) descrever os contextos de uso em que emergem os construtos na língua. Já o levantamento da frequência de uso, que compreende uma análise de natureza quantitativa, se tornaria fundamental se, por exemplo, o objetivo for (Lacerda, 2016): iii) comprovar como os construtos, devido à sua alta frequência, passam a ser reconhecidos na língua como padrões microconstrucionais, que se estabelecem a partir de um pareamento simbólico entre forma e função; iv) compreender a extensibilidade dos níveis mais hierárquicos da rede, atestando que, quanto mais esquemático é um subesquema ou um esquema, maior será o número de padrões microconstrucionais sancionados; v) verificar que, quanto mais produtivo é determinado padrão microconstrucional, maior é a probabilidade de ele servir de exemplar, a partir do mecanismo da analogização – nos termos de Traugott e Trousdale (2013) –, para a emergência de novos padrões microconstrucionais na língua.

Sobre o equacionamento entre a metodologia qualitativa e a metodologia quantitativa, é fundamental mencionar aqui também o posicionamento de Traugott e Trousdale (2013). Apesar de os autores não apresentarem, de modo pontual, uma proposta de cunho metodológico para o tratamento

---

<sup>2</sup> O programa *Antconc* está disponível para os sistemas operacionais Windows, Macintosh e Linux.

construcional da mudança, eles trazem, em sua obra, algumas reflexões sobre o papel do método misto. Segundo os autores, embora sua obra se pautasse primordialmente em uma abordagem qualitativa para a construcionalização e para a ocorrência de mudanças construcionais, eles enfatizam que os dois métodos de análise são complementares. Como sinalizam Traugott e Trousdale (2013), o método quantitativo permitiria estabelecer, de fato, uma relação entre as propriedades da produtividade e da esquematicidade. Vejamos o que afirmam Traugott; Trousdale (2013, p. 238)

Neste livro, discutimos alguns dos caminhos a partir dos quais a esquematização parece se correlacionar ao aumento em produtividade, e em generalidade semântica, sem fornecer medidas concretas de tais mudanças. A abordagem quantitativa pode ser capaz de fornecer algumas pistas sobre a natureza do entrenchamento dos esquemas e sobre a formação prototípica no nível da microconstrução. [...] Uma vez que o chunking parece ser um importante fator no desenvolvimento de uma microconstrução, uma abordagem quantitativa baseada em corpus pode demonstrar como, ao longo do tempo, um chunk se torna entrenchado como uma microconstrução (tradução nossa)<sup>3</sup>.

Como se pode observar, os autores, de fato, apontam a importância do equacionamento entre a metodologia qualitativa e a metodologia quantitativa no âmbito da abordagem construcional da mudança. Como a LFCU assume, primordialmente, as proposições teóricas de Traugott e Trousdale (2013) para o tratamento da construcionalização e para a ocorrência de mudanças construcionais, se torna, portanto, extremamente importante a busca por um refinamento de natureza metodológica que vise, de fato, ao equacionamento entre as metodologias qualitativa e quantitativa, configu-

---

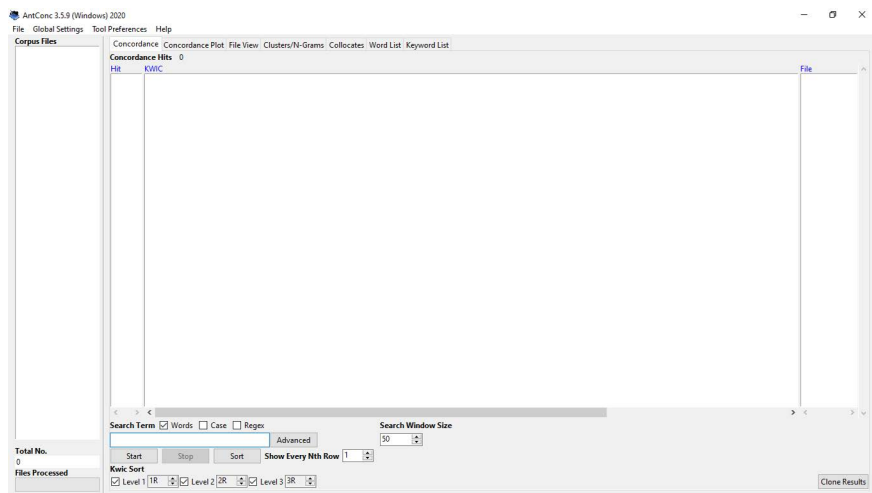
<sup>3</sup> Cf.: “In this book, we have discussed some of the ways in which schematization appears to correlate with increases in productivity, and semantic generality, without providing concrete measures of such changes. A quantitative approach may be able to provide some insights into the nature of entrenchment of schemas, and prototype formation at the level of the micro-construction. [...] since chunking appears to be an important factor in the development of a micro-construction, a quantitative corpus-based approach can demonstrate how, over time, a ‘chunk’ comes to be entrenched as a micro-construction (on which see Bybee 2010, who however does not use the term ‘micro-construction’) (Traugott; Trousdale, 2013, p. 238).

rando o que entendemos como método misto. Desse modo, na próxima seção, de modo bastante prático, buscamos justamente apresentar de que maneira o programa *Antconc* – desenvolvido no âmbito da Linguística de *Corpus* – pode nos levar a análises bastante refinadas nas pesquisas que assumem teoricamente os pressupostos da LFCU.

## Exemplo de aplicação

O programa *Antconc* foi desenvolvido pelo Professor Lawrence Anthony, da Universidade de Waseda, no Japão. Atualmente, a versão mais atual do programa é a 3.5.9. Vale ressaltar, nesse sentido, que o programa *Antconc* apresenta duas grandes vantagens, a saber: a) o tamanho pequeno do arquivo, que permite *download* rápido, não ocupando muito espaço em disco; e b) a utilização sem a necessidade de instalação e licenciamento. Além disso, a interface do programa *AntConc* é muito simples, já que, na mesma janela, é possível navegar por diferentes opções de análise, que permitem descobrir como o elemento pesquisado ocorre, o quanto ocorre e em que contextos ocorre. A seguir, apresentamos a tela inicial do programa *Antconc*:

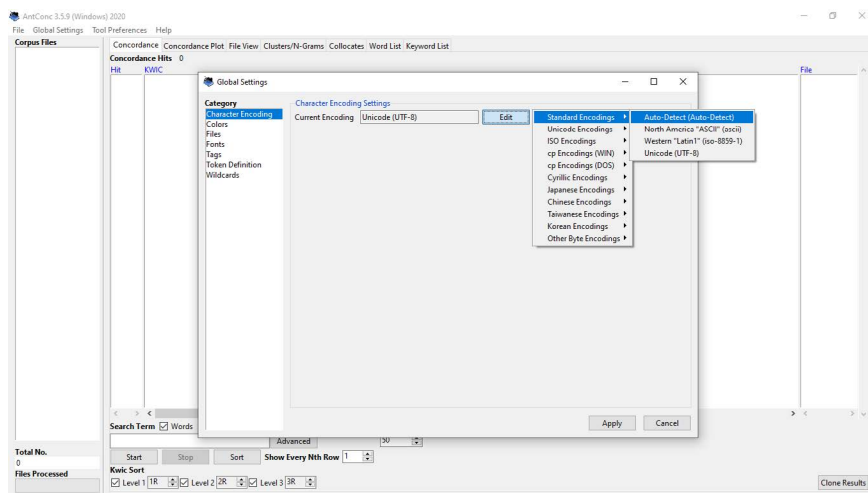
**Figura 1** – Tela inicial do programa Antconc



Fonte: imagem retirada do programa *Antconc*.

O primeiro passo, para utilizar o programa *Antconc*, consiste em converter o arquivo que compõe o *corpus* a ser analisado – ou os arquivos que o compõem, quando houver mais de um arquivo – para a extensão *.txt*. Após esse primeiro passo, devemos acessar a aba *File* e clicar em *Open File(s)* ou *Open Directory*. Quando clicamos em *Open File(s)*, selecionamos individualmente o(s) arquivo(s) que compõe(m) o *corpus*; por outro lado, se quisermos selecionar, de uma única vez, todos os arquivos que compõem o *corpus* e que estão em uma mesma pasta, devemos clicar em *Open Directory*. Vale destacar, ainda, que é extremamente importante realizar um ajuste específico na configuração do programa para que ele seja capaz de processar adequadamente os resultados a partir da acentuação gráfica da língua portuguesa. Caso esse ajuste na configuração não seja realizado, o programa, por exemplo, não conseguirá processar palavras que são acentuadas em português. Nesse caso, devemos acessar a aba *Global Settings* e clicar em *Character Encoding*. Na caixa intitulada *Standard Encodings*, que aparece em *Character Encoding*, devemos selecionar a opção *Auto-Detect* e, posteriormente, clicar em *Apply*. A tela a seguir ilustra esse ajuste necessário na configuração do programa.

**Figura 2** – Alteração na configuração do programa *Antconc* para o reconhecimento dos acentos gráficos.

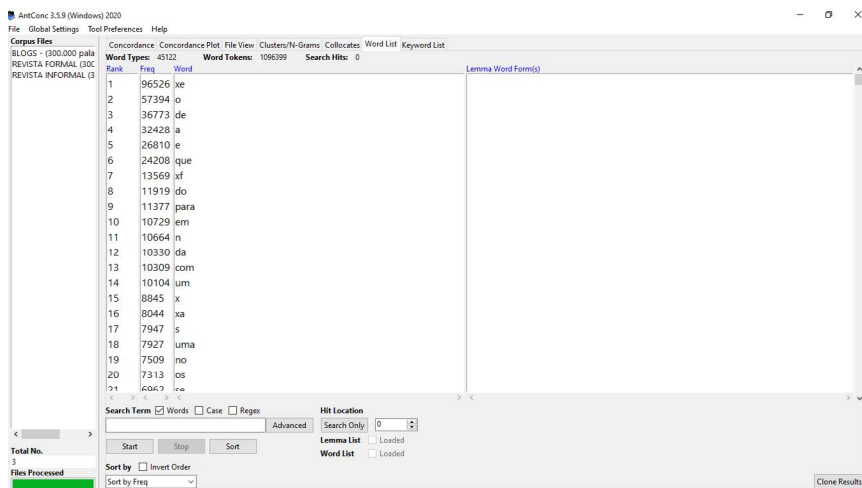


Fonte: imagem retirada do programa *Antconc*.

O programa *Antconc* disponibiliza as seguintes ferramentas para o processamento dos *corpora*: i) *WordList*; ii) *Concordance*; iii) *Concordance Plot*; iv) *File View*; v) *Clusters/N-Grams*; vi) *Collocates*; e vii) *Keywords*.

A ferramenta *WordList* gera uma lista, em ordem alfabética ou por frequência, de todas as palavras – no nosso caso, construções – que constam nos arquivos selecionados para a análise linguística. Assim, é possível encontrar rapidamente quais palavras – ou construções – são as mais frequentes e quais são as com menor ocorrência em um *corpus*. A figura, a seguir, mostra a tela da ferramenta *Wordlist*, a qual possui as seguintes colunas: i) *Rank*, que indica a classificação das palavras do *corpus* por ordem de frequência; ii) *Freq*, que mostra frequência com que cada palavra ocorre no *corpus*; iii) *Word*: que mostra cada palavra referente aos itens das duas colunas anteriores. Nesse caso, a lista é organizada conforme o critério escolhido dentre os três disponíveis: a) *Sort by Freq*, pautada na busca padrão do programa que gera uma lista de palavras por critério de frequência; b) *Sort by Word*, pautada na organização das palavras por ordem alfabética; c) *Sort by Word End*, pautada na ordenação segundo as terminações. A seguir, apresentamos três telas que são representativas dos três critérios de busca facultados pela ferramenta *Wordlist*.

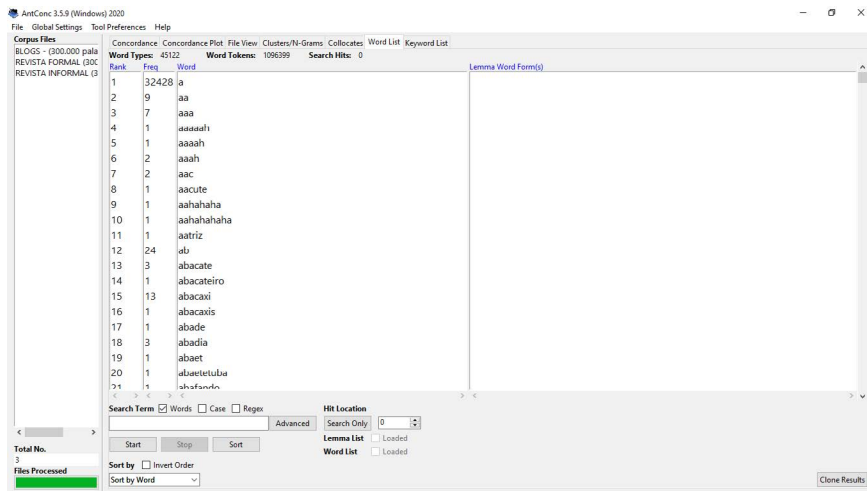
**Figura 3** – Tela representativa da ferramenta *Wordlist* distribuída pelo critério *Sort by Freq*.



Fonte: imagem retirada do programa *Antconc*.

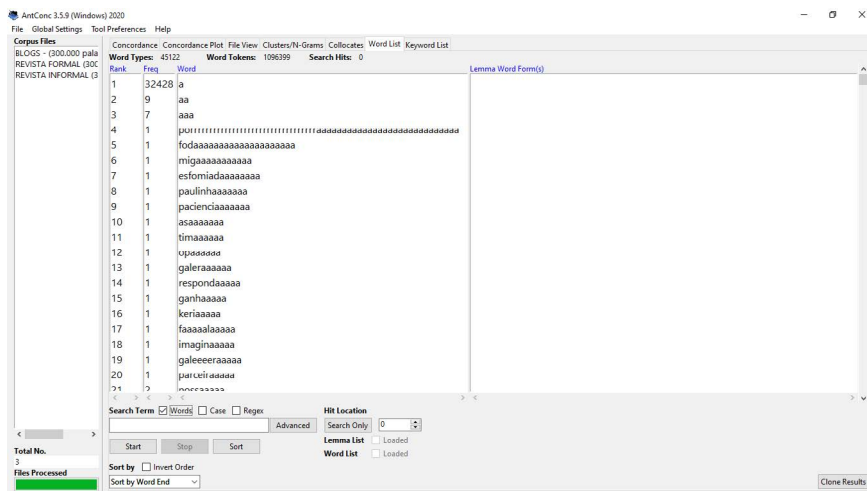


**Figura 4** – Tela representativa da ferramenta *Wordlist* distribuída pelo critério *Sort by Word*.



Fonte: imagem retirada do programa *Antcon*.

**Figura 5** – Tela representativa da ferramenta *Wordlist* distribuída pelo critério *Sort by Word End*.



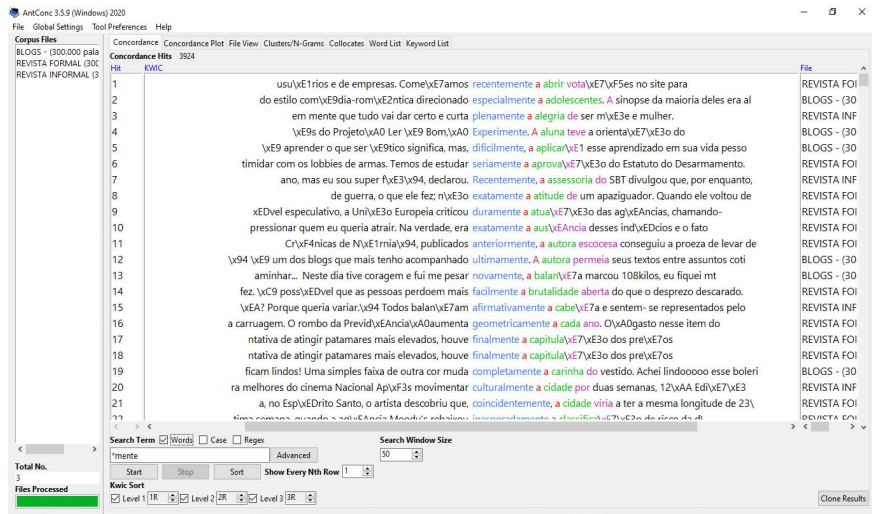
Fonte: imagem retirada do programa *Antcon*.

Por sua vez, a ferramenta *Concordance* mostra os resultados da pesquisa em um formato 'KWIC' (*KeyWord In Context*), o que permite ver

como as construções são comumente usadas em um *corpus* a partir de sua distribuição no contexto. O resultado, nesse caso, é apresentado em três colunas. Na primeira, quantificam-se as ocorrências; na segunda, apresentam-se as linhas de concordância com destaque para as *palavras-chave* no contexto; e na terceira, há a identificação do arquivo em que se encontra.

No caso específico das pesquisas no âmbito da LFCU, podemos pesquisar pontualmente os padrões construcionais que estamos investigando. E, nesse sentido, o programa nos fornece quantos *tokens* representativos de determinado padrão construcional foram verificados, além de apresentar também o contexto de que faz parte cada ocorrência empiricamente atestada. A seguir, apresentamos a tela da ferramenta *Concordance*:

**Figura 6** – Tela representativa da ferramenta *Concordance* a partir do levantamento do padrão construcional [X]mente.

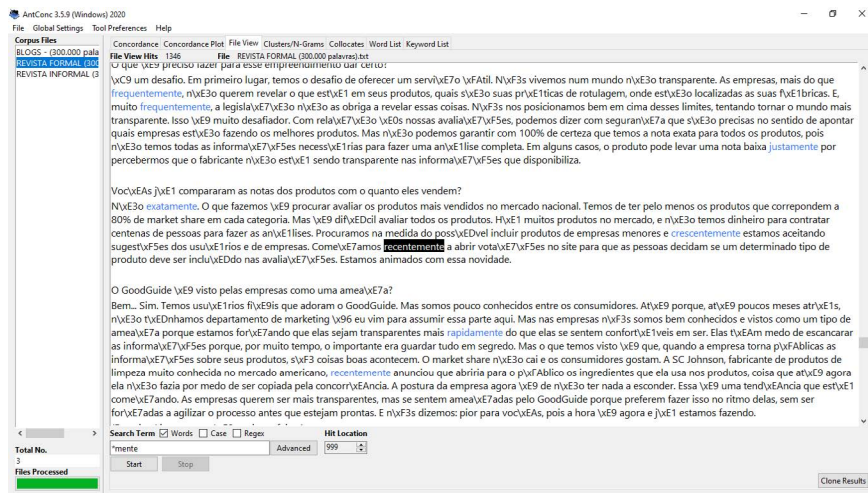


Fonte: imagem retirada do programa *Antcon*.

Como se pode observar na figura 6, foram atestadas 3.924 ocorrências para o padrão construcional [X]mente, em que [X] representa um *slot* com possibilidades diversas de preenchimento por um adjetivo. Nesse caso, ao utilizar um asterisco, podemos suprimir uma determinada parte formadora de um padrão construcional, considerando que esse asterisco atuaria

como um *slot*. Nesse sentido, a ferramenta *Concordance* pode contribuir substancialmente com as pesquisas no âmbito da LFCU, já que nos faculta obter resultados de natureza quantitativa – nos levando a observar a frequência de determinado padrão construcional nos *corpora* investigados – e de natureza qualitativa, tendo em vista que, ao clicarmos em cada ocorrência atestada, abre, por meio da ferramenta *File View*, o contexto de uso mais amplo em que ela figura. A função *File View* mostra o texto bruto de arquivos individuais. Isso permite investigar, com mais detalhes, os resultados gerados pelas ferramentas *Concordance* e *Concordance Plot*, pois as diferentes ocorrências ficam sinalizadas. Nesse caso, se um determinado padrão construcional tiver sido especificado, os resultados serão destacados em todo o texto. Também é permitido alterar a busca. Com o botão *Hit Location*, é possível saltar pelas ocorrências sem precisar percorrer todo o arquivo. A seguir, apresentamos a tela representativa da ferramenta *File View*:

**Figura 7** – Tela representativa da ferramenta *File View*.



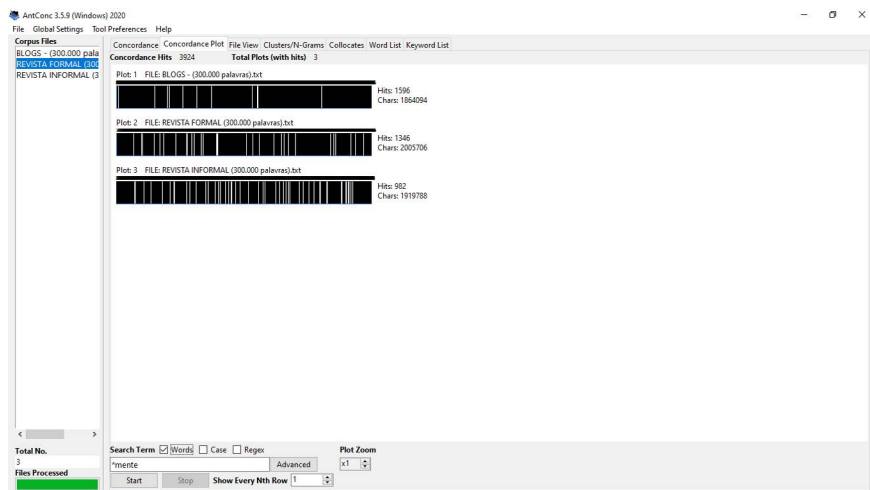
Fonte: imagem retirada do programa *Antconc*.

Outra ferramenta que também pode contribuir com as nossas pesquisas no âmbito da LFCU intitula-se *Concordance Plot*. Esse dispositivo constitui uma ferramenta de plotografia de concordância, que possibilita visualizar os resultados da pesquisa em forma de gráfico, semelhante a um

código de barras, mostrando a distribuição, no arquivo em questão, da palavra ou construção que se está analisando por *corpus*.

Com o intuito de representar a funcionalidade dessa ferramenta no contexto da LFCU, apresentamos abaixo uma tela em que se pode observar, por exemplo, a distribuição de construções [X]mente por *corpus* investigado. Conforme representado abaixo, das 3.924 ocorrências representativas do padrão construcional [X]mente, 1.596 se distribuem no *corpus* formado por textos de *blogs*, 1.346 pertencem ao *corpus* constituído por revistas consideradas mais formais e, por fim, 982 ocorrências foram identificadas no *corpus* constituído por revistas consideradas mais informais.

**Figura 8** – Tela representativa da ferramenta *Concordance Plot*.

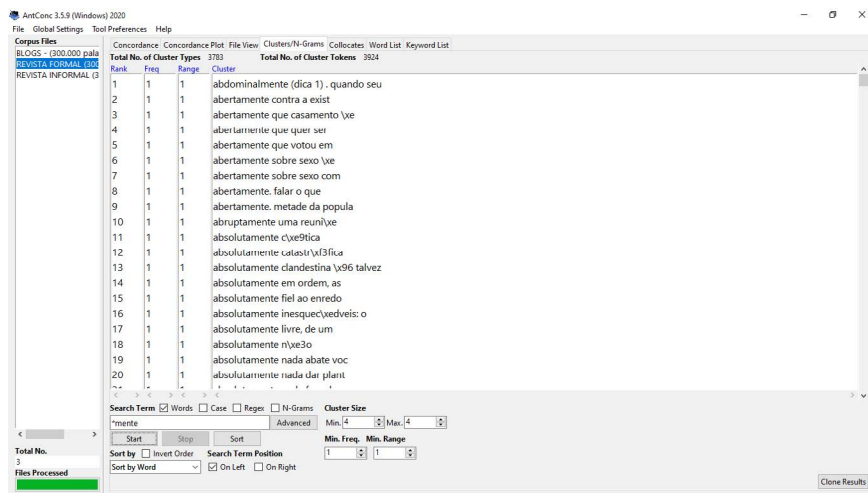


Fonte: imagem retirada do programa *Antcon*.

Por sua vez, a ferramenta *Clusters/N-Grams* gera uma lista com o resultado da pesquisa, conforme a ordem escolhida – alfabética, de frequência, de probabilidade ou de terminações. A ordem dessas listas também pode ser invertida, ativando a opção *Invert Order*. Além dessas possibilidades, a ferramenta *Clusters/N-Grams* agrupa os resultados gerados em *Concordance* ou *Concordance Plot* com as palavras mais próximas à esquerda (*On Left* – busca padrão do programa) ou à direita do termo de pesquisa (*On Right*), de acordo com a opção assinalada. Isso quer dizer que, ao utilizar esta

ferramenta, podemos, em nossas pesquisas no âmbito da LFCU, verificar quais elementos aparecem mais à direita ou mais à esquerda do padrão construcional investigado. Esse tipo de resultado pode, nesse sentido, refinar o olhar do analista em relação ao próprio padrão construcional, já que a verificação de que algum elemento ocorre com uma alta frequência à direita ou à esquerda pode talvez sinalizar que ele faça parte da própria representação formal do padrão construcional que está sob investigação. Nesse sentido, vale ainda destacar que é possível selecionar o comprimento mínimo e máximo – número de palavras – em cada *cluster* e a frequência mínima de *clusters* exibidos, o que possibilita localizar expressões comuns em todo o *corpus*. Esse procedimento é bem proveitoso quando se trabalha com o levantamento de lexias compostas e complexas. No caso específico de nossas pesquisas, a ferramenta *Clusters/N-Grams* pode nos facultar a identificação de *chunks*, nos termos de Bybee (2003, 2007, 2010). A seguir, apresentamos a tela representativa da ferramenta *Clusters/N-Grams*.

**Figura 9** – Tela representativa da ferramenta *Clusters/N-Grams*.



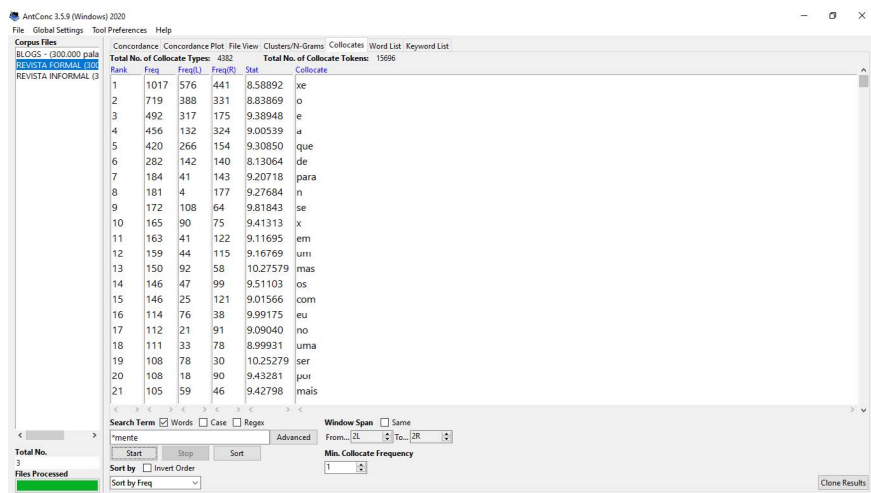
Fonte: imagem retirada do programa *Antconc*.

Na figura acima, observa-se que a busca do *Search Term* foi feita por palavras, o que é o padrão. No entanto, é possível fazer por *sequência de caracteres* desativando a opção *Words*. Ou, ainda, também pode ser feita por

meio de uma busca sensível a maiúsculas e minúsculas ao ativar a opção *Case* ou usando expressões regulares cheias (com caracteres), ativando a opção *Regex*. Essas opções de filtro para a busca estão disponíveis em todas as opções do *AntConc*.

A ferramenta *Collocates* também pode contribuir para o refinamento das nossas pesquisas na LFCU, já que ela permite investigar padrões não sequenciais na língua. Sua função principal é gerar listas ordenadas das palavras próximas ao termo pesquisado, chamadas de “colocados”. Nesse caso, os padrões são ordenados pela frequência total, pela frequência à esquerda e à direita do termo de pesquisa e pelo início e final da palavra. Por meio dessa ferramenta, também é apresentada uma média estatística, gerada automaticamente pelo programa, que mede o nível de relação entre o termo pesquisado e o colocado. Além disso, pode-se selecionar a extensão de palavras à esquerda e à direita do termo de pesquisa para localizar os colocados e a sua frequência mínima exibida. Apresentamos, a seguir, a tela representativa da ferramenta *Collocates*.

**Figura 10** – Tela representativa da ferramenta *Collocates*.

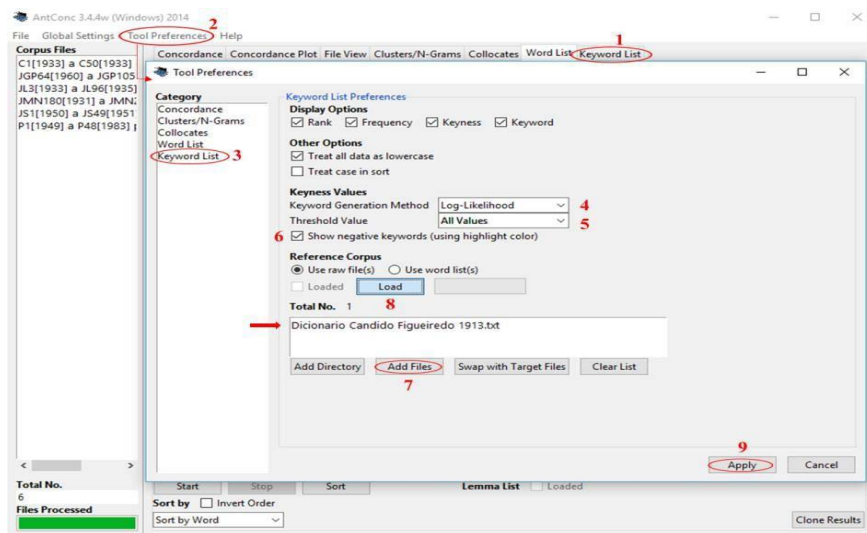


Fonte: imagem retirada do programa *Antcon*.

Por fim, a ferramenta *Keywords* gera uma lista de palavras-chave, comparando a frequência das palavras dos arquivos em análise com a

frequência das palavras de um outro *corpus* – que, nesse caso, atua como *corpus* de referência –, o que permite identificar palavras características no *corpus* de estudo como parte de um gênero ou de uma variedade linguística. Desse modo, para criar uma lista de palavras-chave, primeiro é preciso definir um *corpus* de referência que pertença a um gênero textual distinto do gênero que caracteriza o nosso *corpus* de estudo. A ferramenta *Keywords*, ao comparar, então, as ocorrências do *corpus* de estudo às ocorrências do *corpus* de referência, lista os padrões que realmente são mais idiossincráticos no *corpus* de estudo que está sendo pesquisado. A seguir, apresentamos a tela representativa da inicialização da ferramenta *Keywords*:

**Figura 11** – Tela representativa da inicialização da ferramenta *Keywords*.



Fonte: imagem retirada do programa *Antcon*.

Vemos, portanto, por meio de mais esta ferramenta, que a utilização do programa *Antcon*, de fato, pode permitir ao analista processar e analisar um *corpus* ou dois ou mais *corpora* de forma bastante rápida e precisa, o que possibilita a obtenção de uma grande quantidade de resultados em um curto intervalo de tempo.

## Considerações finais

A partir de uma breve apresentação das ferramentas que constituem o programa *Antconc* e de suas respectivas funcionalidades, reiteramos aqui a convergência entre os pressupostos fundamentais da Linguística de *Corpus* e os pressupostos a partir dos quais se estabelece a LFCU. Assim como a LFCU, a Linguística de *Corpus* também opera com a língua em uso, a qual é atestada em situações reais de produção e interação. Assim como a LFCU, a Linguística de *Corpus* também se baseia em uma perspectiva pautada no empirismo, a partir do levantamento de ocorrências em *corpora* que são representativos de uma língua ou de determinada variedade linguística.

Desse modo, findamos a proposta ensejada neste texto, reafirmando que as nossas pesquisas no âmbito da LFCU podem se beneficiar muito de uma aproximação com a Linguística de *Corpus*. A utilização da ferramenta *Antconc*, que foi brevemente exemplificada neste capítulo, é um exemplo disso. Operar com uma ferramenta como o *Antconc* pode, assim, refinar o nosso olhar para uma identificação mais apurada dos padrões construcionais com os quais trabalhamos, nos permitindo, de uma forma muito mais rápida e precisa, ter acesso a informações sobre frequência de uso e também acesso a todas as ocorrências empiricamente atestadas que são, de fato, representativas dos padrões construcionais investigados em determinado *corpus* ou em determinados *corpora*.

Nesse sentido, reafirmamos aqui a importância de operarmos, no âmbito da LFCU, com o equacionamento entre a metodologia qualitativa e a metodologia quantitativa. E, nesse caso, a assunção da Linguística de *Corpus* – e, de modo mais específico, a utilização do programa *Antconc* – pode, sim, como acreditamos e defendemos, contribuir substancialmente para que, do ponto de vista analítico, nosso olhar sobre os dados – a partir de uma perspectiva funcionalista – seja, cada vez mais, refinado.



## Referências

- BERBER SARDINHA, T. **Linguística de Corpus**. São Paulo: Editora Manole, 2004.
- \_\_\_\_\_. Linguística de Corpus: histórico e problemática. In: **DELTA [online]**, v. 16, n. 2, p. 323-367, 2000.
- BISPO, E. B.; SILVA, J. R. **Variação linguística, mudança linguística e construcionalização**. In: XXI SEMINÁRIO DO GRUPO DE ESTUDOS DISCURSO & GRAMÁTICA E VIII SEMINÁRIO INTERNACIONAL DO GRUPO DE ESTUDOS DISCURSO & GRAMÁTICA. Rio de Janeiro: UFRJ, 2016.
- BYBEE, J. L. **Language, usage and cognition**. Cambridge: Cambridge University Press, 2010.
- \_\_\_\_\_. **Frequency of Use and the Organization of Language**. Oxford: Oxford University Press, 2007.
- \_\_\_\_\_. Mechanisms of change in grammaticalization: the role of frequency. In: JOSEPH, B. D.; JANDA, R. D. (Eds.). **The handbook of Historical Linguistics**. Oxford: Blackwell, 2003. p. 602-623.
- CUNHA, M. A. F. da. Funcionalismo. In: MARTELOTTA, M. E. (Org.). **Manual de Linguística**. 2.ed. São Paulo: Contexto, 2016. P. 157-176.
- CUNHA, M. A. F. da; BISPO, E. B.; SILVA, J. R. Linguística funcional centrada no uso: conceitos básicos e categorias analíticas. In: CEZÁRIO, M. M.; CUNHA, M. A. F. da (orgs.). **Linguística centrada no uso: uma homenagem a Mário Martelotta**. Rio de Janeiro: Mauad, 2013. p. 13-44.
- LACERDA, P. F. A. da C. O papel do método misto na análise de processos de mudança em uma abordagem construcional: reflexões e propostas. **Revista Linguística**, volume especial, p. 83-101, 2016.
- MARTELOTTA, M. E. T. Funcionalismo e metodologia quantitativa. In: OLIVEIRA, M.; ROSÁRIO, I. (orgs.). **Pesquisa em linguística funcional: convergências e divergências**. Rio de Janeiro: Léo Christiano Editorial, 2009. p. 1-20.
- NOVODVORSKI, A.; FINATTO, M. J. (org.) **Linguística de Corpus: abordagem e metodologia em pesquisas linguísticas de base empírica**. **Letras & Letras**, Uberlândia: Universidade Federal de Uberlândia, 2014.
- TRAUGOTT, E. C.; TROUSDALE, G. **Constructionalization and constructional changes**. Oxford: Oxford University Press, 2013.
- ROSÁRIO, I. da C. do.; OLIVEIRA, M. R. de. Funcionalismo e abordagem construcional da gramática. **Revista Alfa**, São Paulo, v.60, n.2, p.233-259, 2016.

# 7. Metodologia para a análise prosódica no tratamento de dados do uso

Lauriê Ferreira Martins Dall'Orto (UFJF)

Patrícia Fabiane Amaral da Cunha Lacerda (UFJF)

## Considerações iniciais

No âmbito da Linguística Funcional Centrada no Uso – doravante, LFCU –, nos termos em que assumem Cunha *et al.* (2013), Cunha (2016), Rosário e Oliveira (2016) e Bispo e Silva (2016), as pesquisas de fenômenos linguísticos adotam uma postura cognitivo-funcional, coadunando pressupostos formulados no âmbito do funcionalismo clássico, ou de vertente norte-americana, e princípios da Gramática de Construções – mais especificamente, na esteira da Gramática Cognitiva (*Cognitive Grammar*), de Langacker (1987, 2008), da Gramática de Construções Cognitiva (*Cognitive Construction Grammar*), de Goldberg (1995, 2006), e da Gramática de Construções Radical (*Radical Construction Grammar*), de Croft (2001).

Três postulados básicos do Funcionalismo Clássico, adotados no âmbito da LFCU, são os seguintes: (i) (re)configuração da gramática da língua a partir do uso; (ii) investigação da língua sob o ponto de vista tanto da gramática quanto do discurso e (iii) correlação direta entre estruturas linguísticas e suas funções em situações reais de interação. No que concerne às contribuições da Gramática de Construções, destacam-se a noção de *rede*, proposta por Langacker (1987, 2008) e por Croft (2001), e o conceito de *construção*, tal como concebido por Goldberg (1995, 2006) e por Croft (2001).

Langacker (1987, 2008) destaca que a língua – assim como outros sistemas cognitivos – constitui uma rede de nós interligados por elos, que se estabelecem de maneira hierárquica, de modo que generalizações podem ser formuladas a partir da observação de características comuns que são

compartilhadas entre membros de diferentes níveis de especificidade. Também, para Croft (2001, p. 25, tradução nossa), cada “construção constitui um nó na rede taxonômica de construções”<sup>1</sup>.

Goldberg (1995, 2006) define *construção* como a unidade básica da língua, estabelecida a partir da convencionalização de um pareamento forma-significado, cujos traços comuns são captados a partir de usos individuais. Croft (2001) acrescenta, ainda, que o conceito de *construção* se aplica a qualquer estrutura gramatical, desde morfemas a padrões completamente esquemáticos.

A partir desses e de outros conceitos, Traugott e Trousdale (2013), na obra intitulada *Constructionalization and Constructional Changes*, propõem um modelo teórico para o tratamento da mudança linguística que ocorre tanto na gramática quanto no léxico – em uma perspectiva construcional e em rede –, o qual tem embasado diversas investigações linguísticas no contexto da LFCU.

Embora tanto Croft (2001) quanto Traugott e Trousdale (2013) apontem a propriedade fonológica como integrante do polo da forma de uma determinada construção, ainda são poucas as pesquisas que se dedicam, no âmbito da LFCU, à análise de aspectos fonológicos – como, por exemplo, tonicidade, intensidade, entoação, contorno, ritmo e pausa – na investigação de diferentes fenômenos linguísticos. É nesse sentido que o presente capítulo pretende demonstrar que: a) a análise da propriedade fonológica, a depender do objeto de estudo, pode se mostrar fundamental para a identificação e descrição de pareamentos forma-função<sup>2</sup>; b) o *software* de análise acústica PRAAT pode contribuir, do ponto de vista metodológico, para um refinamento na identificação de padrões construcionais na língua.

## **Tema central do capítulo**

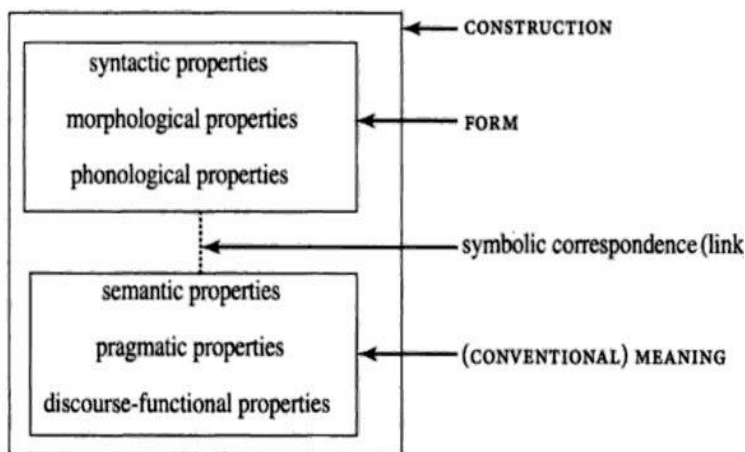
Croft (2001), na obra *Radical Construction Grammar*, propõe o seguinte modelo de representação simbólica para a construção, em termos de correspondência entre forma e significado:

---

<sup>1</sup> Cf.: “Each construction constitutes a NODE in the taxonomic network of constructions.” (Croft, 2001, p. 25).

<sup>2</sup> Assume-se, neste texto, a denominação “pareamento forma-função”, assim como Goldberg (2016), uma vez que “função” é um termo mais abrangente para designar a contraparte semântica, pragmática e discursiva da construção.

**Figura 1** – Representação da construção por Croft (2001).

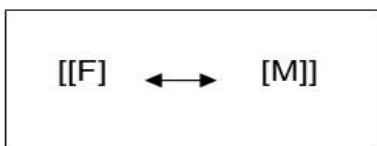


Fonte: Croft (2001, p. 18).

Tal como representada por Croft (2001), a construção constitui uma associação simbólica de propriedades relacionadas ao polo da forma – propriedades sintáticas, morfológicas e fonológicas – e ao polo do significado – propriedades semânticas, pragmáticas e discursivo-funcionais –, o que garante a interpretação da construção como um todo.

Traugott e Trousdale (2013), a partir da representação simbólica de construção, nos termos propostos por Croft (2001), propõem o seguinte:

**Figura 2** – Representação da construção por Traugott e Trousdale (2013).



Fonte: Traugott e Trousdale (2013, p. 8).

Os colchetes externos [ ] indicam que o pareamento forma-significado constitui uma unidade convencionalizada na língua, e a seta dupla  $\leftrightarrow$  sinaliza a correspondência entre forma e significado.

Tanto Croft (2001) quanto Traugott e Trousdale (2013) incluem a propriedade fonológica no polo da forma de uma construção. Desse modo,

visando a contemplar, ainda que de maneira não exaustiva, tal propriedade nas pesquisas em LFCU, a análise acústica, realizada a partir do programa PRAAT – *software* de análise acústica, desenvolvido por Paul Boersma e David Weenink, no *Institut of Phonetics Sciences*, da Universidade de Amsterdam –, é capaz de evidenciar resultados relevantes e fundamentais para a compreensão e a identificação da própria construção, no que diz respeito à sua contraparte formal. Neste capítulo, para ilustrar essa aplicação, utilizaremos dados de construções que atuam em contextos de modalização epistêmico-asseverativa (Barbosa, 2020).

### Exemplo de aplicação

Em contextos de modalização, a análise da prosódia tem se mostrado relevante e fundamental para uma análise mais completa das construções. Conforme destaca Neves (2013, p. 168), “seja qual for o meio segmental utilizado, os meios prosódicos sempre estão presentes na modalização em linguagem falada e, frequentemente, são os únicos responsáveis por ela”. Nesse sentido, isso é o que se percebe na análise de construções com “real” e “real oficial” em contextos de modalização epistêmico-asseverativa<sup>3</sup>, tal como apresentado na pesquisa de Barbosa (2020).

Inicialmente, a ideia da pesquisadora era trabalhar com um *corpus* constituído a partir de dados extraídos da rede social Twitter. Nesse momento da pesquisa, surgiram, entretanto, dois desafios em particular, a saber:

- (i) a limitação do *corpus*: na rede social Twitter, as pessoas compartilham informações e opiniões por meio de *tweets* de até 140 caracteres. Desse modo, os usuários da rede devem ser mais concisos ao comunicar suas mensagens.
- (ii) além de a pequena porção textual característica desse tipo de rede social dificultar a interpretação dos dados, a maneira como cada analista realizava a leitura de cada ocorrência determinava seu

---

<sup>3</sup> A modalização epistêmico-asseverativa diz respeito ao posicionamento do falante (que revela opiniões, crenças e julgamentos) acerca de uma proposição de maneira assertiva, de modo a se comprometer com o que é dito (para maiores detalhes, ver Neves, 2013).

padrão construcional. Em outras palavras, a entoação que cada analista atribuía a uma determinada ocorrência influenciava na identificação de seu pareamento forma-função. Por exemplo, em uma sentença como “eu vi um dragão real”, como saber, apenas a partir do dado escrito, se o falante está comunicando que realmente viu um dragão – indexando, assim, valor de verdade à proposição – ou se ele está caracterizando o dragão como sendo real, com o sentido de realidade?

Logo, identificou-se a necessidade de procurar as respostas para essas questões na análise prosódica das falas dos locutores. Sendo assim, para eliminar a ambiguidade das sentenças e para a compreensão da própria construção no que diz respeito à sua contraparte formal, passou-se à análise dos dados coletados de um *corpus* representativo da modalidade oral da língua – composto por um total de 24 horas e 31 minutos de vídeos extraídos da plataforma *online* YouTube, datados do período compreendido entre 2017 e 2020 – e recorreu-se à análise do *break* por meio do *software* PRAAT (Barbosa, 2020).

Para a identificação adequada do escopo da modalização epistêmico-asseverativa, um importante recurso prosódico investigado foi justamente o *break*, ou *Prosodic Break* (Bögels *et al.* 2011). Segundo Bögels *et al.* (2011, p. 424, tradução nossa), o *break*, entendido como limite prosódico, “consiste em um ou mais dos seguintes elementos: uma pausa em uma frase, um tom de limite anterior a essa pausa e o alongamento da palavra antes da pausa”<sup>4</sup>.

Conforme verificado em Barbosa (2020), em contextos de modalização epistêmico-asseverativa, a presença e a ausência de *break* podem determinar como os elementos de uma construção são organizados na mente do falante e, conseqüentemente, como são produzidos oralmente. Em outras palavras, a análise do recurso prosódico do *break*, mais especificamente do grau de integração entre “real” e “real oficial” e outros elementos que compõem a construção, permite a identificação do escopo

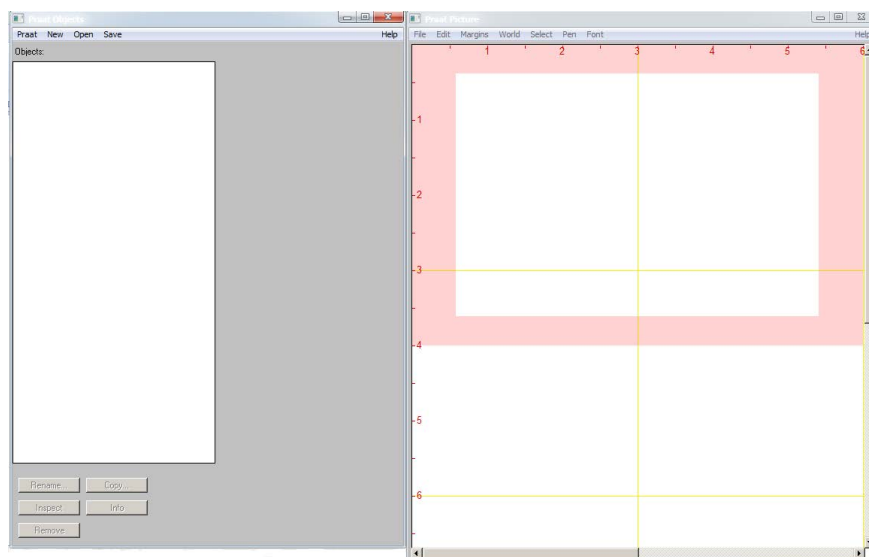
---

<sup>4</sup> Cf.: “A PB, also referred to as prosodic boundary or intonational phrase boundary, consists of one or more of the following elements: a pause in a sentence, a boundary tone preceding this pause and the lengthening of the word before the pause” (Bögels *et al.*, 2011, p. 424).

da modalização epistêmico-asseverativa realizada por tais elementos – se sintagma nominal, se sintagma verbal, se sintagma adjetival, se sentença.

A ferramenta de análise acústica PRAAT está disponível para *download* gratuito<sup>5</sup>, sendo possível optar pela versão que seja compatível com o sistema operacional utilizado. O programa, então, se inicia com duas janelas: a janela *Praat Objects* e a janela *Praat Picture*, como é possível observar na figura a seguir:

**Figura 3** – Tela representativa das janelas principais do programa PRAAT.



Fonte: imagem retirada do programa PRAAT.

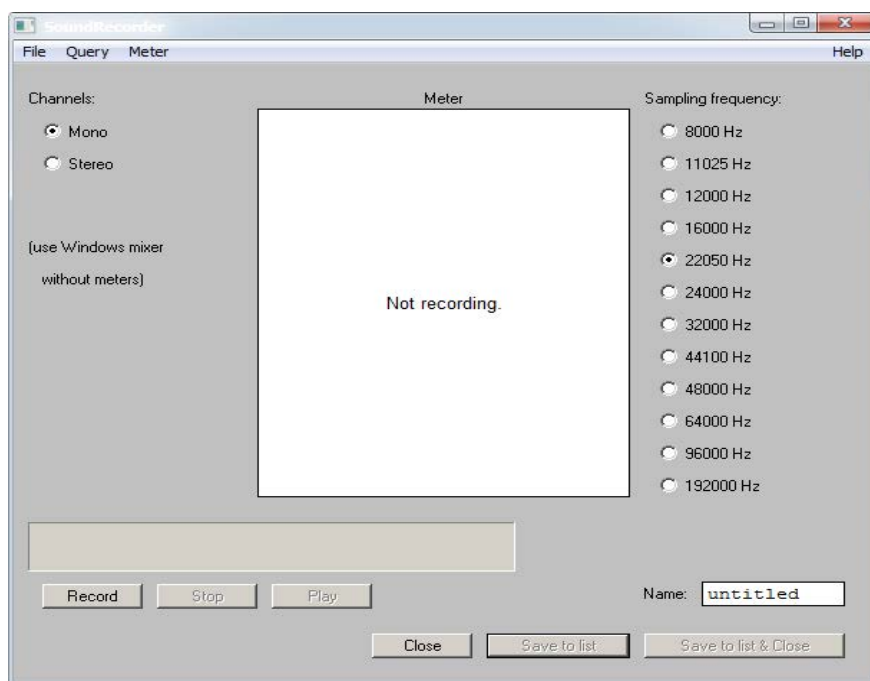
A *Praat Objects* constitui a janela principal do programa, na qual os dados são gravados e analisados; já a *Praat Picture* consiste em uma janela secundária, na qual podem ser editados diagramas e textos. Neste texto, especificamente, demonstramos algumas funcionalidades básicas da janela *Praat Objects* para a análise do *break*.

No caso da análise de dados disponíveis na internet, como em Barbosa (2020), deve-se abrir o áudio ou o vídeo no navegador. Por exemplo, para a análise de dados de fala presentes em vídeos da mídia social You-

<sup>5</sup> O programa PRAAT pode ser baixado no *site* <https://www.fon.hum.uva.nl/praat>.

Tube, deve-se abrir o vídeo escolhido no próprio YouTube e posicioná-lo no tempo em que se realizará sua gravação no programa PRAAT. Em seguida, para gravar a porção textual a ser analisada, devem ser seguidos os comandos na janela Praat Objects: *New* → *Record mono Sound*. Ao abrir a janela SoundRecord, deve-se selecionar uma faixa de frequência (Sampling frequency) adequada (normalmente, se o analista pretende verificar vários segmentos de fala, a partir de 22050 Hz já é suficiente), nomear a gravação e selecionar a opção Record. Em seguida, é necessário dar play no início do excerto de voz investigado e selecionar, no SoundRecord, a opção Stop quando a intenção for parar a gravação. Para ouvir a gravação, é necessário apertar Play; já para salvá-la, deve-se clicar em File e escolher o formato desejado. Por fim, para incluir tal gravação na lista de objects a serem analisados, deve-se clicar em Save to list & Close. A figura a seguir ilustra a janela SoundRecord:

**Figura 4** – Tela do *SoundRecord*.

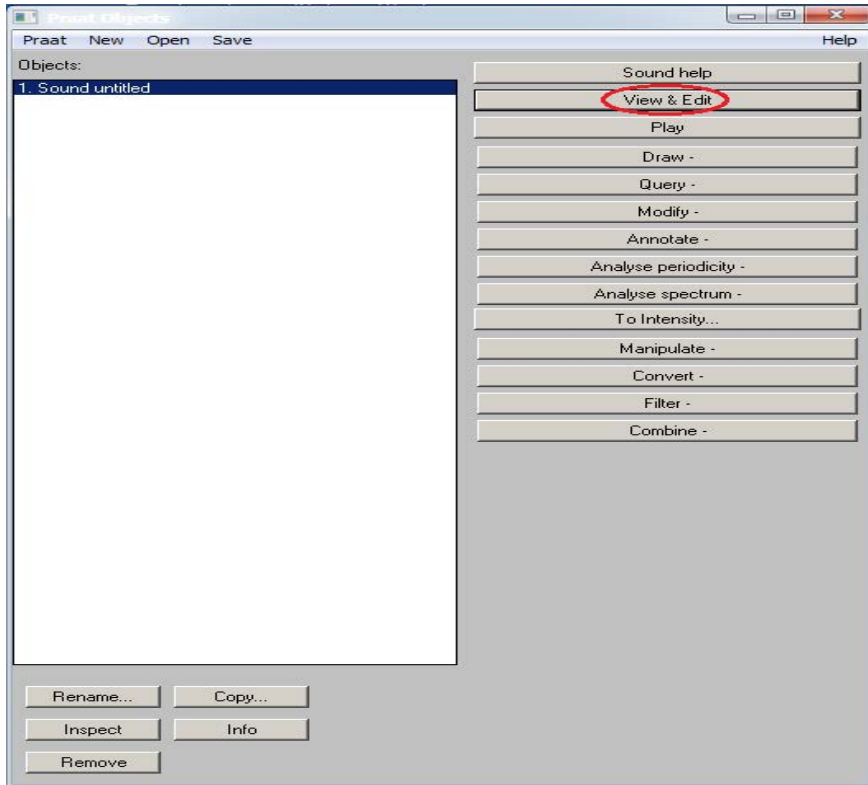


Fonte: imagem retirada do programa PRAAT.



Na janela *Praat Objects*, deve-se selecionar o arquivo gravado (*Sound Untitled*) e clicar em *View & Edit*, conforme ilustrado na figura a seguir:

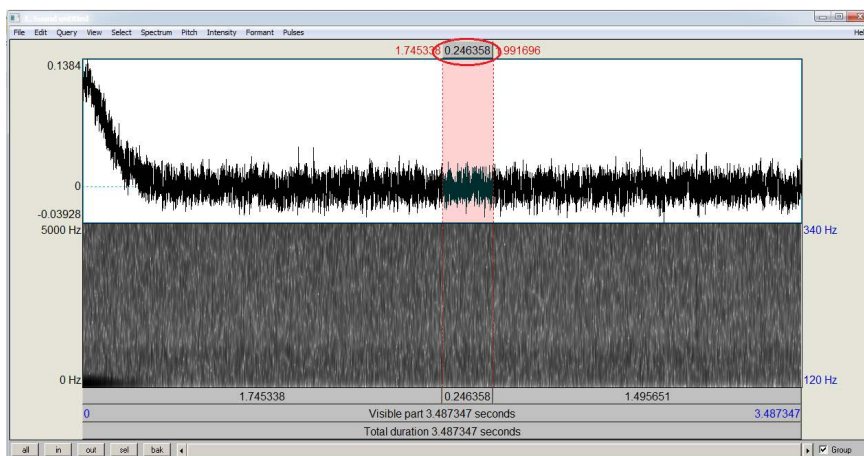
**Figura 5** – Tela do *Praat Objects*.



Fonte: imagem retirada do programa PRAAT.

Na janela que abre em seguida, intitulada *Sound Untitled*, segurando o botão esquerdo do *mouse* e arrastando o cursor para a direita ou para a esquerda, é possível selecionar uma parte do excerto de fala a ser analisado. No caso da análise do *break*, ao selecioná-lo, a sua duração é indicada em segundos na *Play Bar*, acima do retângulo na cor rosa.

**Figura 6** – Tela ilustrativa do *Sound Untitled*.



Fonte: imagem retirada do programa PRAAT.

É com o auxílio dessa ferramenta que Barbosa (2020) identifica os seguintes padrões para as construções com “real”<sup>6</sup> em contexto de modalização epistêmico-asseverativa:

**Quadro 1** – Representação de padrões construcionais com “real”.

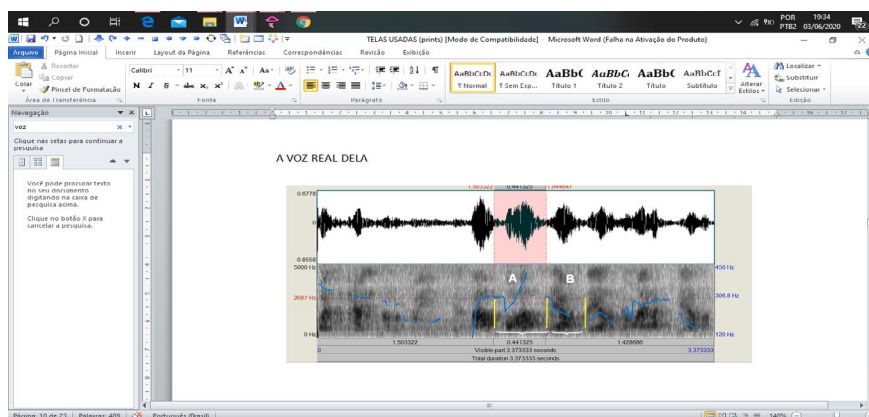
Padrões construcionais	
<b>Padrão construcional 1</b>	<b>Forma:</b> {[SN + real <sup>prosódia</sup> ]} *Prosódia: sem <i>break</i> entre o SN e “real” <b>Função:</b> Atribuir veracidade
<b>Padrão construcional 2</b>	<b>Forma:</b> {[SV + real <sup>prosódia</sup> ]} *Prosódia: sem <i>break</i> entre o SV e “real” <b>Função:</b> Asseverar uma situação ou um aspecto da realidade
<b>Padrão construcional 3</b>	<b>Forma:</b> {[SADJ + real <sup>prosódia</sup> ]} *Prosódia: sem <i>break</i> entre o SADJ e “real” <b>Função:</b> Asseverar uma qualidade
<b>Padrão construcional 4</b>	<b>Forma:</b> {[SS]+[real <sup>prosódia</sup> ]} *Prosódia: com <i>break</i> entre o SS e “real” <b>Função:</b> Asseverar o conteúdo de toda a proposição

Fonte: Adaptado de Barbosa (2020, p. 74).

<sup>6</sup> Para as construções com “real oficial”, ver Barbosa (2020).

As telas a seguir apresentam quatro análises – representativas de quatro ocorrências – de *break* pelo *software* PRAAT. Cada uma delas é representativa de cada um dos padrões construcionais identificados por Barbosa (2020):

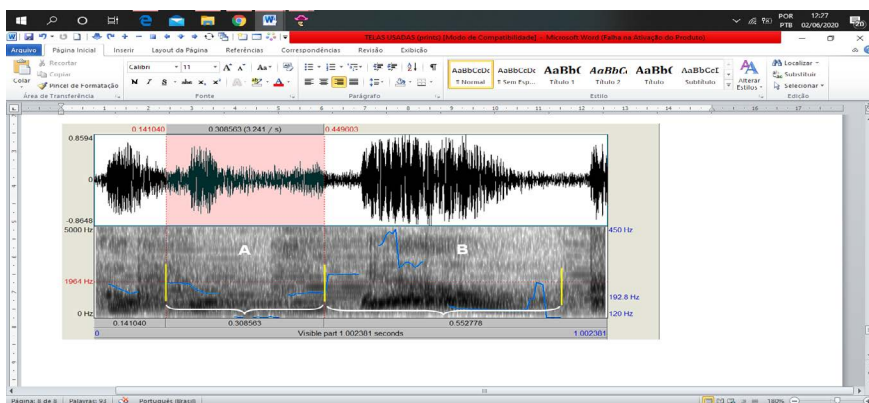
**Figura 7** – Análise acústica representativa do padrão construcional 1.



Fonte: retirado de Barbosa (2020, p. 82).

A Figura 7 apresenta uma análise do *break* na sentença “(...) a voz real dela não era aquela” (Corpus YouTube, 2020). No contexto, o adjetivo “real” indexa o sentido prototípico de realidade com escopo no sintagma nominal “a voz”. A porção selecionada e destacada de rosa (letra “A”) corresponde à produção do sintagma nominal “a voz” – e, conseqüentemente, ao escopo da modalização epistêmico-asseverativa neste caso; na sequência (letra “B”), há a produção de “real”. Conforme se observa na figura, o sintagma nominal e o adjetivo “real” são produzidos de maneira contínua e integrada, sem *break*. E, no caso da ocorrência analisada, as três linhas delimitadoras em amarelo indicam o início e o término do padrão construcional (Barbosa, 2020).

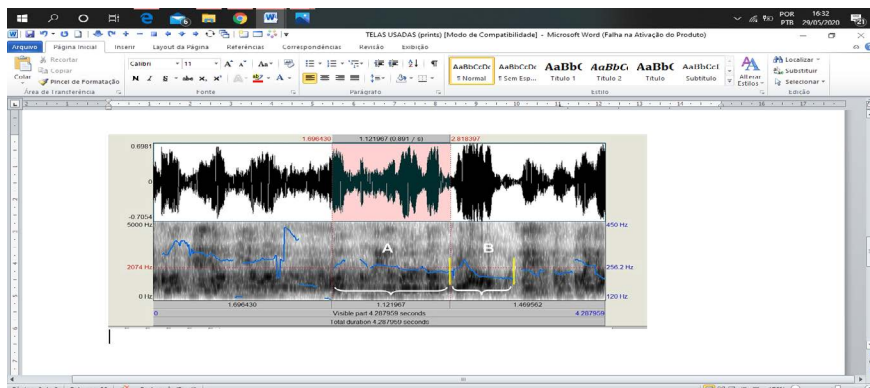
Figura 8 – Análise acústica representativa do padrão construcional



Fonte: retirado de Barbosa (2020, p. 96).

Na Figura 8, a análise do *break* ocorre em relação à ocorrência “(...) tô pagando o preço porque eu comi **real**, eu comi tudo”– e, conseqüentemente, ao escopo da modalização epistêmico-asseverativa neste caso” (*Corpus YouTube*, 2020). Nesse contexto, o advérbio “real” assevera a atitude do falante, “de comer exageradamente”, indexada pelo sintagma verbal “comi”. A porção selecionada e destacada de rosa (letra “A”) corresponde à produção do sintagma verbal, que constitui o escopo da modalização epistêmico-asseverativa neste caso; na sequência (letra “B”), há a produção de “real”. Conforme se observa na figura, o sintagma verbal e o advérbio “real” são produzidos de maneira contínua e integrada, sem *break* (Barbosa, 2020).

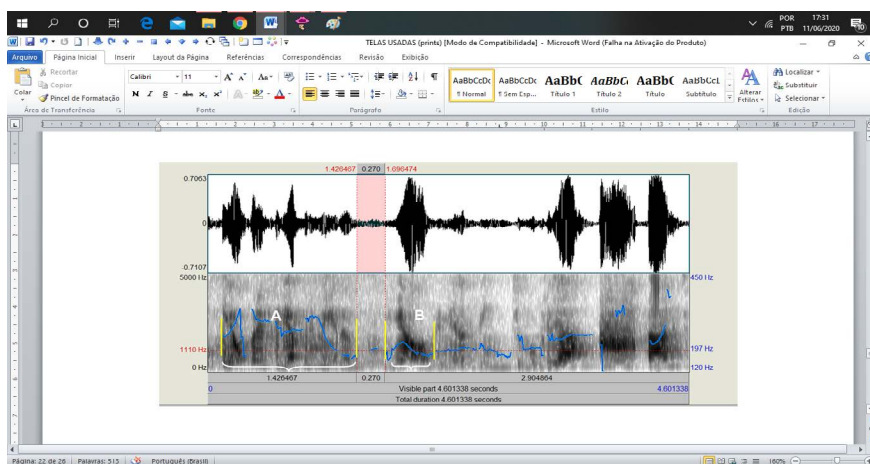
Figura 9 – Análise acústica representativa do padrão construcional 3.



Fonte: retirado de Barbosa (2020, p. 108).

Já a Figura 9 apresenta uma análise do *break* na ocorrência “(...) quantas meninas maravilhosas real!” (*Corpus YouTube*, 2020). Nesse caso, o advérbio “real” é utilizado para asseverar uma qualidade, expressa pelo sintagma adjetival “maravilhosas”. A porção selecionada e destacada de rosa (letra “A”) corresponde à produção do sintagma adjetival, que representa o escopo da modalização epistêmico-asseverativa; na sequência (letra “B”), há a produção de “real”. Conforme se observa na figura, o sintagma adjetival e o advérbio “real” também são produzidos de maneira contínua e integrada, sem *break* (Barbosa, 2020).

Figura 10 – Análise acústica representativa do padrão construcional 4.



Fonte: retirado de Barbosa (2020, p. 113).

Por fim, a Figura 10 apresenta uma análise do *break* na ocorrência “(...) eu preferi a primeira real” (Corpus YouTube, 2020). Nesse contexto, o advérbio “real” é utilizado para asseverar o conteúdo de toda a proposição. A porção selecionada e destacada de rosa, nesse caso, corresponde ao *break*, ou intervalo de tempo, entre a sentença e “real”. Desse modo, conforme se verifica na figura, o advérbio “real” é produzido de maneira não integrada ao escopo da modalização, que representa toda a sentença. No caso da ocorrência analisada, ocorre um intervalo de tempo de 0,27 segundos entre a sentença e “real” (Barbosa, 2020).

### Considerações finais

A partir dos resultados revelados pelo trabalho de Barbosa (2020), consideramos que a análise prosódica pode ser determinante para a identificação de pareamentos forma-função, como ocorre, por exemplo, em contextos de modalização epistêmico-asseverativa, em dados representativos da modalidade oral da língua. Vale ressaltar, aqui, que a própria definição de construção assumida por Croft (2001) e Traugott e Trousdale (2013) sustenta a importância da análise acústica na descrição de padrões construcionais, uma vez que, segundo os autores, aspectos da forma de uma construção – relacionados à fonologia, à morfologia e à sintaxe – estão diretamente associados aos aspectos de seu significado – relacionados à semântica, à pragmática e ao discurso.

Conforme se observa por meio da análise das telas do programa PRAAT, em três padrões construcionais, o escopo da modalização epistêmico-asseverativa – SN, SV e SAdj – estrutura-se de maneira mais integrada em relação a “real”, não apresentando *break* entre os elementos constituintes da construção do ponto de vista formal. Já no padrão construcional em que o escopo da modalização é todo o conteúdo proposicional, a sentença e “real” apresentam-se de maneira menos integrada, o que é sinalizado pela presença do *break*.

Dessa forma, esperamos que a reflexão e as evidências que ensejamos aqui possam, de algum modo, contribuir para um refinamento da descrição dos padrões construcionais do ponto de vista formal, contemplando, de fato, a dimensão fonológica que é preconizada por Croft (2001) quando o autor trata das propriedades que constituem o modelo de representação simbólica para a construção.

## Referências

- BARBOSA, L. S. **Construções modalizadoras epistêmicas asseverativas com “real” e “real oficial”**: uma proposta de rede construcional a partir da Linguística Funcional Centrada no Uso. 2020. 129 f. Dissertação (Mestrado em Linguística) – Universidade Federal de Juiz de Fora, Programa de Pós-Graduação em Linguística, Juiz de Fora, 2020.
- BISPO, E. B.; SILVA, J. R. Variação linguística, mudança linguística e construcionalização. In: XXI SEMINÁRIO DO GRUPO DE ESTUDOS DISCURSO & GRAMÁTICA E VIII SEMINÁRIO INTERNACIONAL DO GRUPO DE ESTUDOS DISCURSO & GRAMÁTICA. Rio de Janeiro: UFRJ, 2016.
- BÖGELS, S. *et al.* Prosodic Breaks in Sentence Processing Investigated by Event-Related Potentials. **Language and Linguistics Compass**, v.5, p.424-440, 2011.
- CROFT, W. **Radical construction grammar**: syntactic theory in typological perspective. New York: Oxford University Press, 2001.
- CUNHA, M.A.F. da; BISPO, E. B.; SILVA, J. R. Linguística funcional centrada no uso: conceitos básicos e categorias analíticas. In CEZARIO, M. M.; CUNHA, M. A. F. da (Orgs.). **Linguística centrada no uso**: uma homenagem a Mário Martelotta. Rio de Janeiro: Mauad X/FAPERJ, 2013. P. 13-39.
- CUNHA, M. A. F. da. Funcionalismo. In: MARTELOTTA, M. E. (Org.). **Manual de Linguística**. 2.ed. São Paulo: Contexto, 2016. P. 157-176.
- GOLDBERG, A. E. A constructionist approach to language. In: **Workshop em XXI Seminário do Grupo de Estudos Discurso & Gramática e VIII Seminário Internacional do Grupo de Estudos Discurso & Gramática**, 2016.
- \_\_\_\_\_. **Constructions at work**: the nature of generalization in language. Oxford: Oxford University Press, 2006.
- \_\_\_\_\_. **Constructions**: a construction grammar approach to argument structure. Chicago: University of Chicago Press, 1995.
- LANGACKER R. W. **Cognitive Grammar**. New York: Oxford University Press, 2008.
- \_\_\_\_\_. **Foundations of cognitive grammar theoretical prerequisites**. Stanford: Stanford University Press, 1987.
- NEVES, M. H. M. **Texto e Gramática**. 2. ed. São Paulo: Contexto, 2013.
- ROSÁRIO, I. da C. do.; OLIVEIRA, M. R. de. Funcionalismo e abordagem construcional da gramática. **Revista Alfa**, São Paulo, v.60, n.2, p.233-259, 2016.
- TRAUGOTT, E. C.; TROUSDALE, G. **Constructionalization and Constructional Changes**. Oxford: Oxford University Press, 2013.

# 8. Metodologia no tratamento da variação construcional

Marcia dos Santos Machado Vieira (UFRJ/CNPq/Faperj)

Marcos Luiz Wiedemer (UERJ/Faperj)

Pâmela Fagundes Travassos (UFRJ, Université de Lille)

## Considerações iniciais

Convidamos o leitor a trilhar conosco caminhos para a descrição de variantes linguísticas e a explicação do fenômeno de variação construcional, inerente à linguagem humana, sistematizado e probabilisticamente predizível. Assim, neste capítulo, queremos que nos acompanhe no seguinte: a) como se dá o tratamento da variação na abordagem construcional da gramática? b) quais e como são os passos de uma investigação em um *design* de pesquisa empírica multimetodológica e multidimensional?

Para nossa viagem, vamos “*dar uma olhadinha/dar uma olhada/olhar*” (n)os dados de pesquisas sobre predicadores verbais e verbo-nominais do Projeto Predicar (<https://projeto-predicar.wixsite.com/predicar>). O objetivo é mostrar passos de um percurso multimetodológico que articula análise experimental, análise multivariada de regressão logística e análise colostrucional do fenômeno de variação, bem como tratamento qualitativo e quantitativo, com base na Sociolinguística Variacionista e na Gramática de Construções (GC). Aqui, focalizaremos, em maior medida, o passo de análise multivariada.

### **1. Encaminhamento teórico-metodológico para o tratamento de variação construcional centrada em amostra do comportamento linguístico observável em *corpora***

Para que o leitor nos acompanhe, começamos por introduzir alguns conteúdos basilares: **(i)** o que é variante, variável, variedade? e **(ii)** o que é



variação construcional, metaconstrução e aloconstrução? Traremos à tona nosso entendimento de outros pontos correlacionados: a) a relação entre variação e mudança; b) a heterogeneidade ordenada, predizível e inerente ao nosso conhecimento multilinguístico e multidialetal, perspectivada diversamente (entre linguistas) e mensurada multimetodologicamente.

### **O que é variante, variável, variedade?**

Variante linguística é o rótulo mais usado em estudos da área de Letras/Linguística para designar uma unidade linguística que, por ter relação de similaridade com outra(s) unidade(s), é conceptualizada e convencionalizada como uma unidade alternante, a conviver com essa(s) outra(s) ao atingir fim semelhante ao desta(s) ou a competir com ela(s) ao pender a substituí-la(s). Está relacionada a uma área de sobreposição/semelhança de valores de atributo(s) ou função. E essa alternância é condicionada probabilisticamente por parâmetros ((extra)linguísticos) também variáveis e, portanto, é estatisticamente predizível.

Uma variável, diferentemente de uma constante, pode exibir pelo menos dois valores, estados ou níveis diferentes, desempenha um papel (variável dependente ou variável independente) na hipótese científica. É realizada diferentemente em contextos distintos; pode ser discreta (a variável diatópica - carioca e catarinense) ou contínua (inclusive, no sentido de assumir graus de significância social), bastando, neste caso, apreender a sua influência em graus/níveis (por exemplo, maior ou menor conexão discursiva na configuração da referencialidade, grau de definitude do referente sujeito). A perspetivação de variantes e variáveis pode tomar como referência uma língua, mais de uma língua em comparação, variedades de uma língua ou de mais de uma língua em comparação.

Variedade é o termo usado para designar conjuntos de regularidades linguísticas diferentes e relativas a um referencial linguístico. A noção de variedade pode envolver delimitação de uma configuração político-territorial de nação (por países em que uma língua é idioma nacional), delimitação por centros normatizadores, por regiões de contatos linguísticos, por nor-

mas e até por perfis tipológicos (variedades de línguas românicas em comparação com variedades de línguas anglo-germânicas).

A perspectiva diassistêmica de GC parte de uma projeção em que os limites do que chamamos de língua são difusos e também considera variedades dialetais (concepção multidialetal de língua). Essa perspectiva diassistêmica não é nova, pois está nas bases da Dialetologia, da Sociolinguística, da Sociolinguística de Contato, conforme já apresentado em Weinreich (1954, p. 390).

É importante esclarecer que o que apreendemos por variedade de língua depende dos recortes a que o pesquisador dá atenção: há quem se refira à língua portuguesa no Brasil considerando dados do eixo Rio de Janeiro-São Paulo e à língua portuguesa em Portugal considerando dados do eixo Lisboa-Porto. De todo modo, a apreensão de variedades/comunidades linguísticas já põe em evidência um projeto de associações e diassociações que depende da compreensão do pesquisador sobre o fenômeno linguístico.

### **O que é variação construcional, metaconstrução e aloconstrução?**

Em seu caminho de pesquisa, você encontrará vários modelos de GC, que nos podem servir de base como aparato teórico-metodológico, que abrigam a concepção de que as línguas são moldadas pela complexa interação de princípios cognitivos e funcionais. Você já deve estar ciente de que a língua é conhecimento e atividade dinâmicos, ou seja, um objeto de observação em constante mudança. Também deve estar ciente de que, para haver mudança, em lugar de constância há variação. Além disso, saberá que mudança linguística, variação, aquisição de língua e aprendizagem de línguas envolvem adaptabilidade.

Porém, não podemos perder de vista que diferentes pesquisadores, e não é diferente entre funcionalistas-construcionistas, podem divergir quanto à apreensão de padrões de variação, pondo diferentes ênfases em termos do que captam, em dados linguísticos, como variantes ou invariantes no *design* do objeto de uma pesquisa. Não é raro encontrar pesquisadores que olhem para o mesmo conjunto de dados e reconheçam neles o fenômeno de variação, os fenômenos de variação e mudança, ou só o de mudança.

Um questionamento sempre feito sobre o tratamento da variação na GC é a relação entre variação e o princípio da não sinonímia<sup>1</sup> (Goldberg, 1995, p. 67). Sobre isso, conforme Machado Vieira; Wiedemer, 2020, p. 274:

Goldberg refere, por um lado, a potencialidade de sinonímia ou semântica (corolário A) ou pragmática (corolário B). Por outro, indica a impossibilidade de identidade entre as propriedades de dois ou mais padrões construcionais. Tendo em vista os multiníveis das faces forma e função de uma construção, está descartada a possibilidade de sinonímia entre as propriedades de todos os atributos que caracterizam um pareamento X e os que caracterizam um pareamento Y, no caso de emergir uma relação de similaridade entre X e Y ou de o falante traçar, mediante processo cognitivo analógico, uma relação entre eles (e essa relação se convencionalizar). Neste sentido, as variantes construcionais de uma área de sobreposição semântica/funcional, de uma alternância/variação, podem ser caracterizadas, segundo Cappelle (2006), como aloconstruções.

Assim, a noção de construção não se pauta na especificação de um sentido fixo, por isso podemos considerá-la polissêmica por natureza. Ademais, de acordo com Machado Vieira e Wiedemer (2020, p. 16), “a construção está relacionada a um enquadre semântico subjacente à cena representada pelos elementos linguísticos que a constituem. Esses enquadres integram o conhecimento partilhado pelos falantes, que podem codificar a cena (enunciação) de diferentes formas”.

Leino & Östman (2005) discutem a ideia de a GC incorporar também generalizações sobre as metaconstruções. Referem-se a um tipo de variabilidade em que certas expressões precisam de referência a analogia: casos de construções (independentes) que são usadas como modelos de uma maneira tal que nenhuma delas por si licencia determinadas expressões resultantes. Para os autores, a metaconstrução não deve ser vista como um

---

<sup>1</sup> “Se duas construções são sintaticamente distintas, devem, portanto, ser semanticamente ou pragmaticamente distintas: Corolário A – se duas construções são sintaticamente distintas e semanticamente sinônimas, então elas têm de ser pragmaticamente distintas. Corolário B – se duas construções são sintaticamente distintas e pragmaticamente sinônimas, então elas têm de ser semanticamente distintas”.

“nível” mais abstrato, geral ou esquemático de construções e nem derivam de generalizações que captam apenas as semelhanças de um dado grupo de construções. Em vez disso, capturam semelhanças e diferenças sistemáticas que ocorrem entre construções. Um exemplo dessa aplicação, denominada de variação por analogia, está em Machado Vieira e Wiedemer (2020). Além disso, os autores associam a noção de metaconstrução à noção de paradigma discursivo, que é o correlato cognitivo do tipo de texto definido linguisticamente e do gênero definido socioculturalmente. Em relação aos valores/atributos, os autores introduzem a noção de *value pool*. A ideia básica é a de que um conjunto de valores pode ser atribuído/especificado a um atributo, semelhante aos fatores de variáveis (socio)linguísticas.

Cappelle (2006) defende que formas alternantes podem ser tratadas pelo arcabouço da GC, e cita as construções de partículas verbais do inglês: a) Verbo-Partícula-Objeto, por exemplo, *pick up the book* – pegar o livro; ou b) Verbo-Objeto-Partícula, por exemplo, *pick the book up* – pegar o livro. Para explicar tais relações entre as construções, sugere que ambas devem ser vistas como variantes construcionais de uma área de sobreposição semântica/funcional, de uma alternância/variação, e caracterizadas como aloconstruções, já que elas correspondem a diferentes possibilidades esquemáticas de um significado/propriedade. Cappelle (2006, p. 18) define aloconstruções como possibilidades alternativas de uma unidade linguística parcialmente especificada.

Perek (2015) retoma a proposta de Cappelle (2006) e também argumenta, com base em pesquisa experimental, que a abordagem construcional da gramática não é incompatível com a proposta de integrar as “alternâncias”. Cada construção envolve propriedades multidimensionais próprias, embora apresente algumas similares/sinônimas quando correlacionada a outra construção. Para Perek (2015), podemos representar construções como generalizações de um significado comum (ou de uma área de funcionalidade semântica, discursiva, pragmática, social ou cognitiva partilhada) entre dois ou mais padrões construcionais, associadas a uma forma não específica (subespecificada) denominada de metaconstrução/*constructeme*. A metaconstrução captura o nível no qual as construções

são equivalentes, e as aloconstruções especificam exatamente em que essas construções diferem.

Uma proposta promissora para a modelagem da variação no âmbito da GC e realizada no cenário brasileiro é a abordagem “Socioconstrucionista”, desenvolvida por Machado Vieira e Wiedemer (2020) a partir de duas premissas teóricas: (i) variação/alinhamento de construções independentes; e (ii) variação por compatibilização de (co)lexemas a *slots* de construção. Para dar conta desses dois encaminhamentos, os autores apresentam as seguintes possibilidades teórico-metodológicas para o tratamento da variação de construções: (i) variação por aloconstruções e metaconstrução; (ii) variação por semelhança simbólica; e (iii) variação por paradigma discursivo (cf. Machado Vieira; Wiedemer, 2020). Destes, focalizamos o primeiro.

A vantagem desta proposição é que ela une as propostas de Leino & Östman (2005), Cappelle (2006) e Perek (2015). Baseia-se na correlação entre atributos/valores associados às aloconstruções como parâmetros de usos de cada construção. Estes podem ser calculados em termos de atração/coerção representativa/esquemática de cada construção:

**Esquema 01** - Metaconstrução/alocnstruções associadas aos atributos/valores.

<p><b>alocnstrução</b> [a &lt;v1;v2;v3&gt;]</p> <p><i>dar uma olhada</i> <i>dinha</i></p> <p>[a&lt;Brasil; modal; papel temático – não identificável/ agente&gt;]</p>	<p><b>alocnstrução</b> [a&lt;v1;v2;v3&gt;]</p> <p><i>dar uma olhadela</i></p> <p>[a&lt;Portugal; aspecto não-durativo; papel temático – não identificável/ agente &gt;]</p>	<p><b>alocnstrução</b> [a&lt;v1;v2;v3&gt;]</p> <p><i>dar uma olhada</i></p> <p>[a&lt;modal; predador complexo; papel temático – não identificável/ agente&gt;]</p>	<p><b>alocnstrução</b> [a&lt;v1;v2;v3&gt;]</p> <p><i>lançar um olhar</i></p> <p>[a&lt;aspecto não-durativo; predador complexo; papel temático – não identificável/ agente&gt;]</p>
<b>metaconstrução</b>			

Fonte: adaptado de Machado Vieira e Wiedemer (2020).

No esquema acima, o retângulo no espaço em cinza representa a metaconstrução (cf. Perek, 2015). Além disso, temos as aloconstruções

(cf. Cappelle, 2006), e os valores são dados entre colchetes. Por exemplo, em [a < v1; v2; v3 >], “a” significa “atributo” e “v” indica “valor” (cf. Leino & Östman, 2005).

## **2. Metodologia para tratamento de variação construcional centrada em amostra do comportamento linguístico observável em *corpora***

Nesta segunda etapa, você já trilhou metade do caminho, e agora chega o momento de colocarmos em prática o que vimos até aqui. Assim, vamos introduzir os principais passos para uma pesquisa sobre variação a partir destes caminhos: **(i)** quais os passos da caracterização do objeto linguístico a ser analisado (o que é variável e o que não é?); **(ii)** quais são as possibilidades (multi)metodológicas? **(iii)** como é, em linhas gerais, a análise da variação a partir do modelo da regressão logística e da aplicação de valores/atributos?

O movimento que dá início à trajetória de pesquisa é o de projetar, inclusive com base no **método introspectivo** (reflexão do pesquisador a partir de sua expertise), uma primeira caracterização do fenômeno variável em que sejam apreendidas pelo menos duas variantes relacionadas por similaridade e cujas variantes sejam estatisticamente acionadas devido a condicionamentos linguísticos e extralinguísticos (atributos formais e funcionais).

A título de ilustração, podemos citar o funcionamento de duas variantes da construção com verbo-suporte: *DAR + (uma) X-(z)inh[o,a]* e *DAR + (uma) X-adela*, cujo acionamento difere a depender, por exemplo, do valor em foco, da natureza do “X” do elemento não verbal e do tipo de construção (Travassos, 2019). Se, por um lado, a variante *DAR + (uma) X-(z)inh[o,a]* é usada, preferencialmente, para indicar valor modal, tem predominantemente o *slot* “X” de natureza denominal e se compatibiliza mais a um tipo de construção como V P1 (com apenas participante 1), por outro lado, a variante *DAR + (uma) X-adela* é utilizada, na maior parte das vezes, para expressar valor aspectual não durativo, tem o “X” com natureza verbal e se encontra mais em estruturas sintáticas do tipo V P1 P2 (com participante 1 e participante 2), como vemos nos seguintes exemplos:

- (1) “Poxa... É normal eu dar um risinho, mas essa manchete foi especialmente inspirada. Entra até o clichê: Seria cômico se... (então pq eu estou rindo?)” [<https://twitter.com/acidhub/status/1410009299934494721>, acesso em 06 de julho de 2021]
- (2) “Vou de vizinhos – breve resenha no TeleSéries- e NCIS hoje. Se der vou dar uma espiadela na Debra Messing” [<https://www.tele-series.com.br/destaques-na-tv-terca-1905/>, acesso em 01 de julho de 2021]

É importante também nos perguntarmos quais foram as facetas já trabalhadas sobre o assunto, via outras pesquisas já realizadas acerca da temática. Na verdade, antes mesmo de iniciarmos nossa própria pesquisa, é importante que seja feita uma **revisão da literatura**, ainda que panorâmica no início e aprofundada ao longo do processo. Desse modo, sabemos quais lacunas ainda precisam ser preenchidas e quais delimitações do nosso objeto de estudo são necessárias. Após o recorte da temática e das primeiras questões da pesquisa, é necessário pensar na metodologia que nos leve a chegar às respostas. Um estudo com perfil multimetodológico (cf. Klavan, 2012) tem maior probabilidade de apreender a realidade linguística, já que, em cada método usado, nossa atenção se volta a um aspecto do fenômeno.

Essa primeira série de movimentos já auxiliará o pesquisador a garimpar as primeiras generalizações indutivas com base nas observações preliminares que faz. Os padrões encontrados alimentarão o raciocínio dedutivo do pesquisador, que o levará a estipular o envelope de variantes construcionais em relação de similaridade (também chamado de variável dependente) que tomará como objeto (variável) de estudo e, então, a configurar hipóteses com base nas quais operacionalizará a investigação desse objeto. O desenho desse envelope de unidades construcionais guiará a detecção de dados do uso com configuração propensa à semelhança conceptual (em termos de um ou mais atributo(s)), de dados que são mais ou menos relacionados a essas unidades construcionais (gradiência) e, assim, que se ligam como membros exemplares prototípicos ou não prototípicos às construções em variação, além de guiar o discernimento de dados não exemplares das unidades em variação.

Uma hipótese científica pode ter formato textual ou estatístico. Em formato textual, é uma declaração potencialmente falsificável, ou seja, passível de ser contradita, com estrutura explícita ou implícita de uma oração condicional (*se... então...*). É uma afirmação sobre a relação entre duas ou mais variáveis ou, ainda, sobre construções ou níveis de uma variável, sobre algum tipo de efeito ou sobre alguma diferença. Pode apresentar relação entre duas variáveis, uma variável independente (constitutiva da parte *se...*) e a variável dependente (constitutiva da parte *então...*), como a seguir:

“se o grau de intersubjetividade é maior, então a construção [dar (uma) X-adinha] será acionada com mais frequência do que quando o grau de intersubjetividade for menor”;

“se o grau de intersubjetividade é maior, então a construção [dar (uma) X-adinha] será acionada diferentemente do que a construção [dar (uma) X-ada]”.

A primeira formulação de texto no exemplo prediz que há algum tipo de efeito, diferença ou relação e, ainda, prevê a direção do efeito com base na média (uma das fórmulas matemáticas de cálculo/maior, menor ou igual à média do tamanho da amostragem – “com mais frequência”). Já a segunda prediz que há algum tipo de efeito, diferença ou relação sem especificar a direção do efeito. Pode, ainda, ter apenas uma variável dependente, como:

“as duas/três construções (“dar uma olhada”, “dar uma olhadinha” e “dar uma olhadela”) não são igualmente frequentes (frequência type) no corpus”.

Esse segundo tipo de hipótese científica constitui uma afirmação sobre uma variável em uma amostragem, que, se supõe, também ser verdadeira em contextos semelhantes e/ou para objetos semelhantes na população. Além desse tipo de hipótese, conhecida como hipótese alternativa H1, formulamos a hipótese nula H0. Esta é o oposto lógico da H1:



“se o grau de intersubjetividade é maior, então a construção [dar (uma) X-adinha] NÃO será acionada com mais frequência do que quando o grau de intersubjetividade for menor”.

Numa hipótese científica em formato estatístico, o pesquisador prevê resultados numéricos (distribuições, frequências, correlações, médias, dispersões) que tem a expectativa de encontrar em relação às suas hipóteses em formato textual. Ao esperar encontrar com mais ou menos frequência uma variante numa amostragem, ele trabalha com o resultado da média.

Para **formular hipóteses**, também contamos com introspecção sobre uso(s) de um dado recurso linguístico, tanto com nossa intuição como usuários da língua, como com nosso olhar de pesquisadores inclinados a explicar as ocorrências da língua em um referencial teórico-metodológico. Nossas intuições sobre contextos específicos em que determinadas construções são usadas, nossas impressões sobre as preferências colocacionais, nossas expectativas sobre prováveis combinações com outros elementos linguísticos ou apreensão de um perfil sociolinguístico do usuário da expressão vão contar nessa caminhada. Uma pesquisa linguística não se limita à introspecção, vale-se dela para empreender uma jornada no sentido de obter evidências empíricas para as impressões, de modo a confirmá-las ou não.

Estabelecida essa etapa, o pesquisador poderá passar a trilhar caminhos que viabilizem a verificação das hipóteses que lhe servem de ponto de partida:

- (i) definir as fontes em que procederá à observação de textos para coleta de dados do uso (centrando-se em pesquisa de usos observáveis em textos, orais e/ou escritos);
- (ii) antes do procedimento de coleta, testar a relação de similaridade entre as unidades que percebe como aloconstruções/variantes mediante pesquisa experimental (combinando pesquisa experimental à pesquisa de usos), inclusive se tencionar retomar esse tipo de tratamento em etapas futuras da investigação;
- (iii) a partir da análise preliminar dos dados e das generalizações indutivas obtidas inicialmente, recorrer à metodologia de pesquisa

experimental para a verificação de relações de similaridade e outras relações (iniciando o percurso pela pesquisa de usos e, num segundo momento, recorrendo à pesquisa experimental, e não só para a testagem de relações de similaridade).

Assim, uma **sistematização dos dados observados** pode ser realizada a partir de coleta e análise de dados de uso e coleta e análise de dados de percepção mediada por avaliação subjetiva (pesquisa experimental) e, então, a partir deles, podemos alcançar generalizações. Quando dizemos “usos da língua”, estamos fazendo referência a usos retirados de situações comunicativas que, de fato, ocorreram naturalmente (*corpus*), não sendo, pois, fruto de projeção ou invenção do pesquisador, com base na experiência que ele tem de língua.

A centralidade da análise empírica envolve desafios. Um deles está no volume de dados com que podemos ter de lidar; afinal, estudos socioconstrucionistas tendem a lidar com grandes volumes de dados. Também contemplam leitura e anotação qualitativa e manual de dados examinados em (dimensões de) contextualidade discursiva, o que demanda tempo e requer o balanceamento/equilíbrio do pêndulo que nos guia entre o tratamento qualitativo e quantitativo de *corpora*. Sabemos que, quanto mais rico é o conjunto das anotações que fazemos dos dados, melhor é entender a gradualidade dos resultados estatísticos e cogitar/garimpar implicações potencialmente envolvidas (que, por sua vez, podem levar a verificar a coatuação de outras influências discursivas ou pragmáticas). Outro desafio está no processo de **coleta ou extração de dados**, que deve cercar-se de cuidados de modo a garantir a obtenção de uma amostra representativa (que represente todos os segmentos da população investigada) e balanceada (que represente proporcionalmente as partes da população investigada na comparação feita entre variedades), bem como promover a aleatorização no levantamento dos membros da população amostral, dos dados textuais das fontes.

É preciso fazer escolhas quanto à materialidade dos dados a coletar (em produções orais, escritas, ambas), quanto a suporte dessas produções (inquéritos ou documentos de acervos de dados de projetos de pesquisa,

cartas, textos em jornais e revistas acadêmicos ou não, roteiros de cinema, tweets, por exemplo) e técnicas para obtenção de dados (ler texto a texto à procura de dados ou buscas mediadas por sistema/aplicativos de concordância). É importante destacar que as fontes e textos que servirão de base da busca dos dados são pré-estabelecidos em função da organização da proposta de operacionalização da pesquisa.

Ainda sobre a **seleção dos dados**, poderíamos levar em conta, por exemplo, dados em variação em uma sincronia específica ou podemos ter um enfoque diacrônico, com direcionamento para possíveis mudanças linguísticas ao longo do tempo. Outra possibilidade de análise acerca da mesma temática seria a verificação de como se dá a variação entre construções específicas entre variedades diferentes de uma língua. Poderíamos observar, por exemplo, diferenças de usos preferenciais entre construções com verbo-suporte no português, levantando a hipótese de que, no português do Brasil, essa construção tem um uso mais produtivo do que no português europeu. Se estamos analisando a variação entre construções com verbo-suporte, podemos levar em conta a alternância entre predicadores complexos somente ou entre estes e predicadores simples (*dar uma bobeadada* e *bobear*, por exemplo), verificando a possibilidade de “equivalência” formal-funcional entre esses elementos e, assim, mapeando indícios de variação por similaridade. Por outro lado, podemos verificar extensões de sentido que uma mesma construção possa ter, como nos exemplos que seguem, em que a construção *dar uma entrada* apresenta, respectivamente, os sentidos de *entrar*, *empurrar*, *protocolar*, *calçar* e *atracar*:

- (3) “Estacando em frente à sede do governo, **deram entrada** no seu recinto”.<sup>2</sup>

[PB, Jornal online, [www.oglobo.globo.com](http://www.oglobo.globo.com)]

- (4) “Felipe Melo disse que Robinho **deu uma entrada** em Kaká”.

[PB, Jornal online, [www.oglobo.globo.com](http://www.oglobo.globo.com)]

- (5) “Ninguém poderá **dar entrada** em processos em número superior a cinco”.

[PB, Jornal online, [www.oglobo.globo.com](http://www.oglobo.globo.com)]

---

<sup>2</sup> Todos os dados utilizados neste capítulo são extraídos de Travassos (2019).

- (6) “Arranjaram um par 42, que teve os bicos cortados para **dar entrada** aos pés do arqueiro juvenil”.  
[PB, Jornal online, [www.oglobo.globo.com](http://www.oglobo.globo.com)]
- (7) “O pacote chinês “Yue-Ying-Wa” (...) não **deu entrada** no porto”  
[PB, Jornal online, [www.oglobo.globo.com](http://www.oglobo.globo.com)]

Portanto, a natureza dos dados coletados importa para a análise, para a busca de respostas às questões que mobilizam o processo de pesquisa. E a amostra de dados de pesquisa deve ser capaz de gerar generalizações mais amplas a partir dela, deve ser representativa e deve nos permitir reconhecer padrões.

Extraídos os dados das fontes escolhidas para a pesquisa, é preciso proceder à **triagem dos dados**, tendo em conta que há exemplares mais ligados às variantes e outros menos ligados, e até alguns cujo uso não é associado a exemplar de variante construcional em jogo na análise. Para ilustrar essas possibilidades, citamos os exemplos que seguem a respeito de predicadores complexos usados na conceptualização de percepção visual.

- (8) “Queria que você me levasse até a sacada do meu quarto, para eu **dar uma olhada** - respondi.” (Livro “As Bruxas”, de Roald Dahl, tradução em versão brasileira, p. 145)
- (9) “Fui dar uma volta ontem para **dar uma analisada** na região.” (<https://www.noticiasagricolas.com.br/fala-produtor/mensagem-122924/> - acesso em 02 de julho de 2021)
- (10) “Muito bacana. Onde encontrou esse livro? Queria **dar uma lida**.” (<https://medium.com/@iagokv/muito-bacana-onde-encontrou-esse-livro-queria-dar-uma-lida-e517385fa90c> - acesso em 02 de julho de 2021)

Se, por um lado, no exemplo (8), há o uso de uma variante construcional (*dar uma olhada*) que, prototipicamente, expressa percepção visual (inclusive, possui “olh-” na base do elemento não verbal), por outro lado, os exemplos (9) e (10) se afastam desse significado, revelando-se usos mais periféricos da construção sob análise, já que envolvem outras nuances

semânticas e cognitivas para além da noção de visão. Cabe ao pesquisador tomar decisões relativas ao encaminhamento que será dado a exemplares como esses. Nesse caso, é necessário decidir se os dados serão considerados ou não e, se considerados, qual será o tratamento diferencial dado a eles. É importante que o investigador do fenômeno faça considerações sobre a triagem feita, assim como sobre os valores envolvidos nessas construções mais periféricas e como elas se ligam com a construção prototípica sob análise.

Definidas a natureza e as fontes de coleta dos dados, é preciso pensar na forma de **reunir e armazenar os dados**. Em se tratando de pesquisa construcionista, vale lembrar que a contextualidade é uma propriedade importante. Logo, é fundamental reunir, numa pasta, os documentos (textos orais ou escritos, inquéritos, vídeos, transcrições) e, num documento de editor de texto, os dados do fenômeno variável em estudo. Para o tratamento quantitativo, é importante anotar os dados por meio do recurso a aplicativos de planilhas eletrônicas (por exemplo, *Excel*) e da organização dos dados por caso ou fator da variável (dependente e independente) do estudo. Logo, é necessário preparar uma cópia dos dados reunidos num documento de editor de texto que permita a codificação/anotação parametrizada de cada dado, tendo em conta as variantes/aloconstruções previstas na variável dependente e os valores/fatores de cada variável independente/atributo levado em conta.

A **parametrização dos dados** depende de quais são nossas variáveis dependentes e independentes. No português do Brasil, por exemplo, há diferença formal entre construções com verbo-suporte acionadas para a atenuação discursiva a depender da natureza do elemento não verbal ligado ao verbo-suporte DAR (tipo de X-afixo: dar X-ada, dar X-adinha, dar X-idinha, dar X-ida, dar X-adela, dar X-(z)inh[o,a]). Essa pode ser nossa variável dependente (eneária) ou pode redundar noutra configuração de variável dependente (binária), neste caso por meio, por exemplo, da análise probabilística da aplicação/do acionamento ou não de “dar X-adela” em relação às demais variantes em contextos de predicação atenuadora. Nessa variedade, encontramos, com frequência, por exemplo, expressões como

*dar uma espiada* e *dar uma espiadinha*. De modo a verificar os fatores que influenciam mais o uso de uma ou de outra estrutura, realizamos anotações relativas a possíveis variáveis (lexicais, semânticas, discursivas, pragmáticas, cognitivas, sociais, morfológicas, morfossintáticas e sintáticas) influenciadoras do comportamento linguístico das construções em alternância, tendo em vista que a variação pode não ser livre, havendo condicionamentos. Exemplos de variáveis dependentes e independentes na análise das perífrases verbo-nominais em variação são ilustradas a seguir:

**Quadro 02** – Parâmetros de análise: variável dependente e variáveis independentes.

<b>Variável dependente</b>	<b>1 – Tipo de predicador verbo-nominal em função de X-afixo</b> [dar X-ada]predicador complexo [dar X-ida]predicador complexo [dar X-adinha]predicador complexo [dar X-idinha]predicador complexo [dar X-adela]predicador complexo [dar X-(z)inh[o/a]]predicador complexo		
<b>Variáveis independentes</b>			
	<b>(i) Com relação à configuração da construção de predicação verbal como um todo ou ao predicador complexo</b>	<b>(ii) Com relação à configuração do sujeito do predicador complexo (“[dar X-afixo]predicador complexo”)</b>	<b>(iii) Com relação à configuração do elemento não-verbal do predicador verbal complexo</b>
	2- Valor em foco no estado de coisas 3- Perspectiva na conceptualização do estado de coisas 4- “Equivalência” de predicador complexo verbo-nominal com predicador verbal simples 5- Tipo de configuração morfossintática da predicação verbal 6- Tempo e modo do verbo DAR 7- Tipo de construção de estrutura de participantes 8- Traços da tipologia dos estados de coisas da construção 9- Marcas de oralidade	10- Número, grau de definitude e tipo de sujeito 11- Pessoa gramatical do participante 1 12- Animacidade e traço humano do participante 1 13- Concretude do participante 1 14- Papel temático do participante 1	15- Presença ou ausência de determinante e número 16- Presença ou ausência de quantificador 17- Presença ou ausência de modificador/qualificador 18- Presença ou ausência de intensificador 19- Natureza do “X 2. Extralinguísticas 20- Gênero textual 21- Recorte temporal

Fonte: Travassos (2019, p. 260).

Uma vez estipuladas as variáveis, planejamos um sistema de **codificação** dos dados:

**Quadro 03** - Sistema de codificação de dados.

Afixo	Valor	Perspectiva
a - X-adinha b - X-idinha c - X-ada d - X-ida e - X-adela f - X-(z)inh[o,a]	1 - Modal 2 - Aspectual – não-durativo 3 - Aspectual – reiterativo 4 - Intensidade	s- [+ subjetividade] o - [- subjetividade] i - [+ intersubjetividade]
Equivalência com verbo simples	Tipo de configuração morfossintática da predicação verbal	Tipo de construção de estrutura de participantes
9 – Sim 8 – Não	d - Estrutura finita/desenvolvida i - Estrutura não-finita (reduzida de infinitivo) g - Estrutura não-finita (reduzida de gerúndio) p - Estrutura não-finita (reduzida de participio)	1 - Vpredicador complexo 2 - - Vpredicador complexo SNArg/Participante1 3 - - Vpredicador complexo SNArg/Participante1 SPArgParticipante2 4 - - Vpredicador complexo SPArgParticipante 2

Fonte: adaptado de Travassos (2019).<sup>3</sup>

Estabelecemos previamente para cada fator/valor de atributo um código a pôr na planilha. Ao estruturarmos uma planilha para codificação, organizamos as anotações de modo que cada variável ocupe uma coluna, a começar pela coluna da variável dependente prevista logo após a coluna do dado categorizado, conforme ilustramos.

Após a reunião dos dados e a codificação de cada dado por variável, passa-se à etapa de **análise quantitativa**. Com a análise quantitativa, por vezes, é possível observar, por meio da distribuição dos dados, indícios de direcionamentos para usos preferenciais de uma determinada construção em detrimento de outra. Dessa forma, podemos fazer inferências estatísticas, que guiam a abordagem probabilística: observamos distribuições e tendências de uso, vemos como a língua está se moldando e também configuramos uma análise qualitativa buscando relacionar dados numéricos a referencial teórico-explicativo. E podemos detectar, com base nos números de frequência, ou indícios de variação (em que, pelos diversos atributos formais e funcionais tomados como variáveis independentes, se encontram as variantes em estudo, mesmo que, em relação a alguns, elas revelem inclinações percentuais de acionamento diferentes) ou de distribuição complementar

<sup>3</sup> Exemplo de dado codificado: (a1i9i3: “Qualquer ação envolvendo dinheiro público entre agentes públicos para **dar uma arrumadinha**, um ajeitamento em algo que pode estar fora de ordem é condenável, é mais que condenável, é execrável, falou ele” [PB, Jornal online, www.oglobo.globo.com]. Basta buscar o significado de cada item no quadro.

(em que, a depender do contexto, uma variante é acionada em 100% ou 0% dos casos), conforme exemplificamos a seguir:

Na indicação da perspectiva em foco no estado de coisas, a variante construcional *dar (uma) X-ada* é acionada em 100% dos casos em que há sobreposição da perspectiva [- subjetiva] e [+ intersubjetiva] (Travassos, 2019), como vemos no exemplo que segue:

- (11) “Você tem que ser criativo na hora de **dar uma cantada**, mas isso depende muito do “teor alcoólico” para ter cara de pau”.

[PB, Jornal *online*, [www.oglobo.globo.com](http://www.oglobo.globo.com)]

A construção *dar cantada* tem o significado de *paquerar* e não apresenta determinante no elemento não verbal. Além disso, parece estar convencionalizada com esse sentido, o que nos faz classificar esse dado como [- subjetivo], embora haja um diálogo direto com o interlocutor [+ intersubjetivo]. As outras variantes da construção (*dar (uma) X-adinha*, *dar (uma) X-idinha*, *dar (uma) X-ida*, *dar (uma) X-adela* e *dar (uma) X-(z)inh[o,a]*) não apresentaram situações em que houvesse sobreposição dessas duas perspectivas.

Além disso, podemos entrever pressuposições relacionadas ao funcionamento cognitivo do uso das estruturas linguísticas sob análise e seu grau de rotinização. A frequência de distribuição de dados e exemplares das variantes pode revelar indícios da etapa em que se encontra o processo de convencionalização social de cada construção/aloconstrução. Segundo Bybee (2002), duas medidas de frequência merecem especial atenção: a frequência *type* e a frequência *token*. A primeira faz referência à frequência de tipo construcional, de cada microconstrução. Já a frequência *token* diz respeito à frequência de ocorrência. Ilustramos, a seguir, a distribuição dos dados em função da frequência:

**Tabela 01:** Distribuição de dados em função das frequências *type* e *token*.

DAR UMA X-IDINHA						
<i>Type</i> (6 tipos)	Dar uma batidinha	Dar uma corridinha	Dar uma dormidinha	Dar uma fugidinha	Dar uma lidinha	Dar uma saidinha
<i>Token</i> (20 ocorrências)	2	2	3	7	2	4

Fonte: Travassos (2019, p. 70).



Por sua vez, com a **análise qualitativa**, podemos observar os contextos em que as expressões em variação foram usadas, suas características, verificar se há extensões de uso e de sentido, averiguar quais são as motivações para o uso de cada construção em cada contexto específico, como em Travassos (2021), que fez um estudo sobre a variação por paradigma discursivo entre construções que indicam aspecto inceptivo: *dar saída, dar arrancada, dar largada, dar partida, dar encetada, dar começada, dar iniciada, dar entrada, dar principiada e dar inaugurada*.

Nessa investigação, a autora detectou diferenças de usos preferenciais de determinadas construções em função de certos aspectos não só semântico-discursivo-pragmáticos, mas também formais (temática, gênero textual, tipo de ato de fala, tipo e número de determinante no elemento não verbal, concretude do complemento das perífrases verbo-nominais e possibilidade de “equivalência” com predicador simples). Observou, por exemplo, quanto à temática, que “dar (a) partida” é mais atraída para o tema de Automóvel, “dar (a/uma) saída”, para o de Informática, “dar (a) largada” e “dar (a/uma/sua) arrancada”, para o de Esporte, “dar (uma) começada”, para o de Entretenimento, “dar uma iniciada”, para o de Educação e “dar (a/uma) entrada”, para o de Sociedade. Do mesmo modo, podemos detectar construções em variação por convivência (por exemplo, *dar uma freada* e *dar uma desacelerada*, em um mesmo contexto, cf. Travassos, 2019) ou por competição (por exemplo, *dar uma cochiladinha* e *dar uma dormidinha* ou *dar uma lida* e *dar uma lidinha*, cf. Travassos, 2019). Dessa forma, realizamos um trabalho descritivo e interpretativo.

Para a **análise quantitativa** podemos utilizar aplicativos de acesso/uso livre como: o programa computacional *Rstudio* (<https://cran.rproject.org/>) ou *Goldvarb X* (programa estatístico desenvolvido por Sankoff, Tagliamonte e Smith (2005), <http://individual.utoronto.ca/tagliamonte/goldvarb.html> - ferramenta muito explorada em estudos sociolinguísticos), *AntConc* (<http://www.laurenceanthony.net/software/antconc/>, um concordanciador, que permite listar as ocorrências de uma determinada unidade em uma quantidade definida de contextos) e também recursos de quantificação do *Excel*.

Por meio destes e por outros recursos do *Rstudio* e do *Rbrul* (<http://www.danielezrajohnson.com/rbrul.html>) ou do *Goldvarb X*, primeiramente observamos a distribuição dos dados (valores numéricos absolutos) e a frequência percentual das variantes por valor/fator de cada variável independente (valores relativos), conforme anteriormente visto. Em segundo lugar, examinamos a frequência percentual de *tokens* em função de co-atuação de fatores mensurada via tabulação cruzada de variáveis, procedimento de análise viabilizado pelo cruzamento de duas variáveis independentes no *Rstudio/Rbrul* ou no *Goldvarb X* e que permite ratificar a hipótese de coatuação de variáveis independentes ou negá-la.

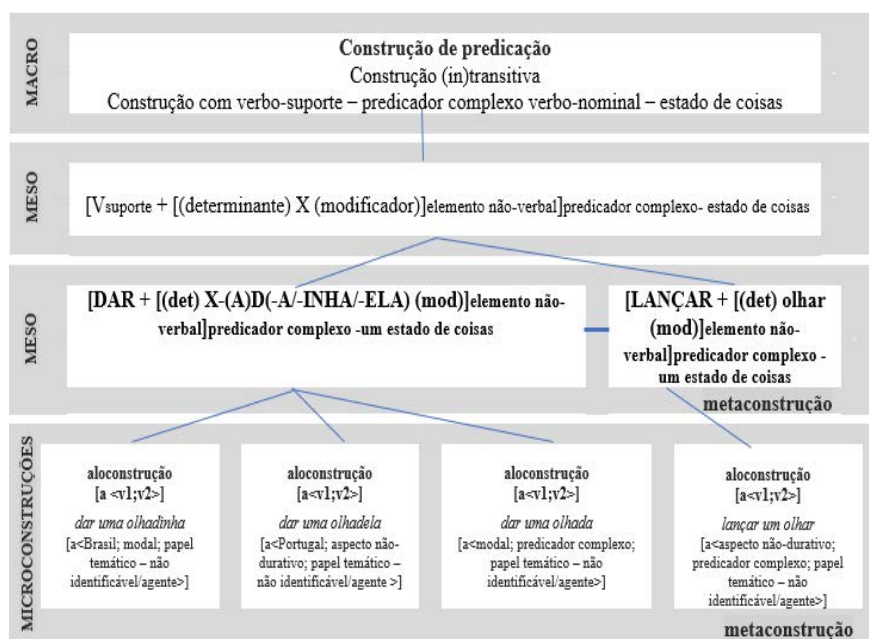
Se queremos verificar, por exemplo, se há relação entre valor em foco no estado de coisas e recorte temporal em usos de construções com verbo-suporte, podemos cruzar essas variáveis. Analisar esses dois grupos de fatores em co-atuação pode ser interessante para verificar, por exemplo, se há indícios de mudança construcional, ou seja, se há alteração, nesse caso, no polo funcional da construção ao longo do tempo. No estudo, Travassos (2019) tinha a hipótese de que o valor modal seria acionado com mais frequência nos decênios mais atuais, suposição que se confirmou com a análise.

Por fim, procedemos à **análise multivariada/multidimensional** de regressão logística linear, para verificar que valores numéricos (pesos relativos/probabilidades) são medidos estatisticamente à relação entre variantes da variável dependente e as múltiplas variáveis estipuladas (efeitos destas correlacionados), bem como quais são as indicadas pelo programa como mais significativas e quais são as eliminadas por ele ao final da rodada estatística. Medimos, com os resultados fornecidos por essas três etapas de quantificação, a influência de fatores linguísticos e extralinguísticos no acionamento de uma ou outra aloconstrução, bem como a propensão a uma ou outra variante no *corpus* ou a convivência entre elas. Se a análise tiver envolvido uma variável que perspective, em alguma medida, diferença geracional, a variável faixa etária ou a variável tempo cronológico, é possível observar a relação entre variação e mudança: em faixas etárias diferentes e/ou em sincronias diferentes. Naquele caso, procedemos a uma investigação em tempo aparente. No segundo caso, procedemos a uma investigação de

tendências em tempo real (de curta ou de longa duração). Um exemplo disso é a pesquisa de Travassos (2019, p. 94), que mostra a distribuição (crescente) de dados de perífrases verbo-nominais, coletados em texto jornalístico, ao longo de 90 anos.

Cabe, agora, ao pesquisador, compreender e **interpretar os resultados** fornecidos pelo processo de tratamento quantitativo (que propicia valores numéricos relativos a *tokens* e *types*) em função das hipóteses e da teorização explicativa em que baseia a investigação. Dessa análise, advêm as generalizações sobre o que está socialmente convencionado e o que é possível projetar em termos de representação mental da variação construcional. Vale destacar que o tratamento qualitativo está presente desde o início da trajetória e também no momento do tratamento quantitativo. Em função desse tratamento qualitativo e quantitativo, cabe-lhe proceder à projeção do fenômeno variável na rede de construções, representando formal e funcionalmente os valores detectados em termos de atributos das aloconstruções que se relacionam por similaridade numa metaconstrução, como em Travassos (2019):

**Esquema 2** - Mapeamento da rede de construções com verbo-suporte.



Fonte: Baseada em Travassos (2019, p. 248).

O esquema 2 consiste em uma representação da construção com verbo-suporte DAR. Partindo do pareamento forma-função constitutivo de cada construção, percebemos que essa construção se compatibiliza a uma estrutura de predicação, por meio de predicador complexo verbo-nominal, cuja função é a conceptualização de percepção visual. A seguir, vemos diferentes graus de esquematicidade, com uma especificação gradiente do verbo-suporte e do tipo de X-afixo. Por fim, na rede, há um espaço para marcar a ideia de que construções como “dar uma olhada” e “dar uma olhadinha” são variantes construcionais (aloconstruções), apresentando diferenças e semelhanças, estas últimas captadas na metaconstrução.

E o pesquisador pode recorrer também à análise colostrucional (método que é objeto de outro capítulo deste livro), para mensurar estatisticamente as variantes que preenchem um mesmo *slot* construcional de predicador (como, no caso de Travassos; Cappelle; Machado Vieira, 2021) no sistema de dados textuais que serviu de fonte de coleta. Outro método a somar evidências diz respeito à **análise de percepção** mediada por avaliação subjetiva dos falantes, ao qual podemos nos referir como **pesquisa experimental**. Observamos a atitude dos usuários da língua em relação a um aspecto do objeto de estudo, bem como percepção ou avaliação subjetiva de aspectos que nem sempre conseguimos apreender por uma análise de *corpora*: aspectos discursivos (associações relativas a graus de formalidade, intersubjetividade, ponto de vista) ou afetivo-emocionais (preferências estilísticas por certos exemplares, certas associações individuais), envolvidos no acionamento de uma ou outra variante.

Há uma série de fatores que podem influenciar nas respostas a uma pesquisa experimental. Por isso, é importante que esse método se articule a outros de modo a alcançar resultados mais confiáveis do ponto de vista da validade científica. O pluralismo metodológico propicia cercar um fenômeno linguístico da forma mais completa possível, levando em consideração suas diversas facetas. Algumas requerem, além de observação de *corpus*, técnicas de comparação de *corpora* aos acervos que lhes servem de fonte ou técnicas de exploração de dados manipulados ou não em laboratório e testados em relação a certas hipóteses metalinguisticamente (*offline*)

ou no processamento online, por meio do recurso a participantes de pesquisa (olhares e percepções não-linguistas).

### **Considerações/Destaques finais**

Qualquer modelo de linguagem ou gramática precisa tratar dos fenômenos de estabilização, variação e/ou mudança da língua. Isso exige do pesquisador uma investigação a partir de um perfil (multi)metodológico que considere representação social e psicológica/mental da variação construcional observada em *corpus/corpora* e/ou em registros obtidos por pesquisa experimental, ou seja, um passo a passo básico, que envolve etapas e movimentos.

É particularmente crucial para a abordagem (funcionalista-)construcionista incorporar as demandas estabelecidas pela variação em sua composição básica; porém, conforme já alertado por Leino & Östman (2005), o compromisso da GC de não se afastar da “periferia” representa um fardo extra para o trabalho. Assim, uma diversidade de movimentos, ações e práticas metodológicas está em jogo, assim como ocorre na língua viva, dinâmica e heterogênea. E, para aprofundamento do que aqui tratamos, basta *dar uma pesquisada/olhada* nas referências citadas.

### **Referências**

- BYBEE, J. Mechanisms of change in grammaticization: the role of frequency. In: JOSEPH, B. D.; JANDA, R. D. (eds.) **Handbook of Historical Linguistics**. Oxford: Blackwell. 2002. p.602-623.
- CAPPELLE, B. Particle placement and the case for ‘allostructions’. In: SCHÖNEFELD, D. (ed.) **Constructions All Over: Case Studies and Theoretical Implications, Special volume of Constructions SV1- 7/2006**, 2006. p. 01-28.
- GOLDBERG, A. **Constructions: A Construction Grammar Approach to Argument Structure**. Chicago: The University of Chicago Press, 1995.
- KLAVAN, J. **Evidence in linguistics: corpus-linguistic and experimental methods for studying grammatical synonymy**. University of Tartu: Institute of Estonian and General Linguistics, 2012.

LEINO, J.; ÖSTMAN, J-O. Constructions and variability. In: FRIED, M.; BOAS, H. (eds.) **Grammatical Constructions: Back to the roots**. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2005. p. 191-213.

MACHADO VIEIRA, M. dos S.; ESTEVES, G. A. T. Metodologia de avaliação subjetiva de usos linguísticos em variação. In: LOPES, C.; REICH, 112 **Uma história de investigações sobre a Língua Portuguesa Uli**. Romania. Variação linguística em megalópoles latino-americanas, 39, 2009. p. 237-266.

MACHADO VIEIRA, M. dos S.; WIEDEMER, M. L. Sociolinguística Variacionista e Gramática de Construções: os desafios e as perspectivas de compatibilização. In: MACHADO VIEIRA, M. dos S.; WIEDEMER, M. L. **Dimensões e experiências em Sociolinguística**. Blucher, 2019. p. 85-120.

MACHADO VIEIRA, M. dos S.; WIEDEMER, M. L. A variação no modelo construcionista da Linguística Funcional-Cognitiva. In: BRESCANCINI, C. R.; MONARETTO, V. N. O. (orgs.). **Sociolinguística no Brasil**: textos selecionados. Porto Alegre: Editora da PUCRS, 2020. p. 265-304

PEREK, F. **Argument Structure in Usage-Based Construction Grammar**: experimental and corpus-based perspectives. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2015.

SANKOFF, D.; TAGLIAMONTE, S.; SMITH, E. GoldVarb X – a multivariate analysis application. Toronto: Department of Linguistics; Ottawa: Department of Mathematics. 2005.

TRAUGOTT, E. C., TROUSDALE, G. **Constructionalization and Constructional changes**. Oxford: OUP, 2013.

TRAVASSOS, F. P. **Variação e mudança construcional: um olhar funcional-cognitivo sobre usos de construções com verbo-suporte DAR no PB**. 260f. Dissertação (Mestrado em Língua Portuguesa) - Universidade Federal do Rio de Janeiro, Faculdade de Letras, Rio de Janeiro. 2019.

TRAVASSOS, P. F.; MACHADO VIEIRA, M. dos S. Uma análise construcionista da variação entre construções com verbo-suporte DAR no PB. **Revista Soletras**, v. 1, p. 272-298, 2019.

TRAVASSOS, P. F.; MACHADO VIEIRA, M. dos S. O que brasileiros dizem acerca de usos de construções com verbo-suporte? **(Con)textos linguísticos**, v. 14, p. 198-217, 2020.

TRAVASSOS, F. P. A perífrase verbo-nominal no português brasileiro: um estudo da variação por padrão discursivo entre construções com o verbo-suporte DAR. **Revista E-escrita: Revista do Curso de Letras da UNIABEU**, v. 12, p. 41-58, 2021.

TRAVASSOS, F. P.; CAPPELLE, B.; MACHADO VIEIRA, M. dos S. Alternance de prédicats complexes: comparaison de variétés de langue portugaise et des langues romanes (portugais et français). **I Colóquio Internacional VariaR**. 2021.

WEINREICH, U. Is a structural dialectology possible? *WORD, Journal of the linguistic circle of New York*, v. 10, p.388-400, 1954.

WEINREICH, W.; LABOV, W.; HERZOG, M. (1968). "Empirical Foundations for Theory of Language Change". In: LEHMANN, Paul; MALKIEL, Yakov. (eds.) **Directions for Historical Linguistics**. Austin: University of Texas Press: 95-188. [*Fundamentos empíricos para uma teoria da mudança linguística*]. Trad.: Marcos Bagno; revisão técnica: Carlos Alberto Faraco. São Paulo: Parábola, 2006.

WIEDEMER, M. L.; MACHADO VIEIRA, M. dos S. Lexemas e construção: atração, coerção e variação. **Caderno Seminal Digital Especial**, v. 30, n. 30, p. 81- 132. 2018a.

WIEDEMER, M. L.; MACHADO VIEIRA, M. dos S. Sociolinguística e Gramática de Construções: o envelope da variação. In: FRANCESCHINI, L.T.; LOREGIAN-PENKAL, L. (orgs.) **Sociolinguística**: estudos de variação, mudança e atitudes linguísticas. Guarapuava: Editora da Unicentro, 2018b. p. 41-77.

# 9. Contribuições da abordagem experimental para o tratamento de dados do uso

Eduardo Kenedy (UFF / CNPq)

Ana Cláudia Machado dos Santos (UFF)

## Considerações iniciais

A pesquisa experimental foi introduzida de forma sistemática nas práticas dos pesquisadores no início da Idade Moderna e rapidamente se popularizou como a abordagem metodológica por excelência das ciências empíricas, sobretudo entre as ciências da natureza. Ao final do século XIX, a experimentação se consolidava como método relevante em algumas ciências sociais, principalmente na psicologia cognitiva, com o primeiro laboratório de psicologia experimental inaugurado por Wilhelm Wundt em 1879, na Alemanha. Ao longo do século XX, no entanto, a difusão do método experimental entre as chamadas “ciências humanas” foi dificultada em parte pela hegemonia do paradigma comportamentalista estadunidense para o estudo da mente (cf. Skinner, 1957) e, em certa medida, pelo bem-sucedido projeto epistemológico pós-moderno proposto para as humanidades por diversos filósofos, sociólogos e cientistas políticos (cf. Feyerabend, 1994; Rorty, 1990). No século XXI, são ainda poucas as pesquisas sobre o comportamento humano que adotam a experimentação como a sua principal ferramenta metodológica e, na linguística, isso não é exceção. Portanto, são ainda majoritárias as pesquisas em ciências sociais que empregam a intuição do analista, a observação naturalista (análise de *corpus*) e a hermenêutica interpretativista como principal ou única abordagem metodológica.



Com efeito, ainda que minoritária, ao longo dos anos a experimentação sobreviveu em distintas “ciências humanas”, especialmente na psicologia cognitiva. Na linguística, a metodologia experimental vem ganhando força nas últimas décadas, com a publicação de diversas obras e pesquisas, nos EUA, na Europa e no Brasil, que recorrem à experimentação como metodologia empírica central (cf. Bybee, 2010; Abraçado; Kenedy, 2014; França, Ferrari & Maia, 2016). O presente capítulo neste livro sobre metodologias de pesquisa funcionalista é mais uma evidência da consolidação do método experimental no seio da linguística contemporânea.

Conforme será detalhado ao longo do capítulo, a pesquisa experimental em linguística é uma abordagem metodológica relevante porque permite a verificação empírica de previsões comportamentais derivadas de algum modelo teórico específico em qualquer área da gramática ou do discurso. Essa testagem se torna factível quando um modelo teórico em linguística define com clareza as variáveis de pesquisa que deseja manipular e, assim, consegue elaborar previsões comportamentais precisas que, em tese, decorrem de tais variáveis, como certos índices de acerto/erro numa tarefa, dado tempo de resposta a um estímulo específico, certo padrão de escaneamento ocular durante a leitura de um texto ou a observação de uma imagem, dado percentual de aceitação ou rejeição de uma estrutura gramatical específica etc.

Mas não apenas isso, pois a maneira pela qual a experimentação conduz a testagem de hipóteses e a validação (ou a rejeição) de previsões comportamentais se dá de uma maneira altamente formalizada, isto é, padronizada e explícita, de modo que outros pesquisadores podem compreender e reproduzir uma pesquisa experimental. É essa possibilidade de replicação de protocolos experimentais que confere à experimentação o seu poder de generalização empírica e de autocorreção, consagrados nas ciências da natureza, uma vez que pesquisas independentes que convergem para um mesmo resultado ou para uma mesma conclusão provavelmente identificaram de maneira correta as variáveis responsáveis por determinado fenômeno em estudo, assim como os erros de uma pesquisa experimental particular podem ser corrigidos por outro estudioso atento, que depurará

tais falhas e seguirá com a exploração experimental de um assunto a partir do ponto em que outro pesquisador se estagnou.

Neste momento, talvez uma questão passe pela imaginação do leitor: por que ou em que circunstâncias estudiosos funcionalistas considerariam a adoção de uma abordagem experimental em suas pesquisas? Afinal, que tipo de validação empírica a experimentação consegue fazer, e outras opções metodológicas não fazem? A resposta para essa questão tem a ver com o tipo de problema que o linguista de orientação funcionalista deseja investigar. A metodologia experimental cria circunstâncias artificiais de uso da língua, como as aduzidas nas tarefas experimentais, de modo que os dados ali obtidos são qualitativa e quantitativamente muito distintos de dados colhidos em situações naturais de uso efetivo da língua. Por outro lado, o uso “real” de uma língua nunca será tão extraordinariamente simplificado de modo que apenas um pequeno número de variáveis linguísticas ou não linguísticas esteja condicionando um dado fenômeno da gramática ou do discurso – o natural é que um gigantesco número de variáveis de diversas naturezas atue simultaneamente e interaja com alta intensidade em qualquer uso de uma língua.

Sendo assim, um pesquisador precisará fazer uma escolha: ou ele enfrenta toda a complexidade do mundo linguístico natural e desiste de investigar a atuação de uma variável específica (ou de um pequeno número de variáveis), ou ele abre mão da complexidade do mundo real e abraça a artificialidade dos estímulos linguísticos experimentais que permitem o controle de variáveis específicas. Tipicamente, um pesquisador funcionalista assume a primeira opção. A razão para que ele ultrapasse essa tradição e recorra à experimentação será, portanto, apenas uma: a necessidade de medir o efeito de uma variável específica (ou de algumas poucas variáveis) que não pode ser observada em atuação isoladamente no mundo naturalista do uso da língua. Assim, para que sua opção pela metodologia experimental seja cientificamente motivada, o tipo de problema que o linguista de orientação funcionalista deseja investigar deverá ser suficientemente restrito aos efeitos de uma variável (ou de algumas poucas variáveis) particular, sobre a qual o mundo linguístico natural pouco ou nada pode dizer. Na linguística,

a experimentação é, antes de tudo, uma metodologia capaz de aferir o efeito comportamental de uma variável manipulada pelo pesquisador.

O presente capítulo apresentará indicações práticas de como pesquisadores funcionalistas podem delinear projetos experimentais em suas pesquisas. Nesse sentido, em vez de conduzir uma discussão abstrata e mais densa sobre o fazer experimental em linguística – para o que leituras como Butler (2006), Johnson (2008) e Oehlert (2010) já se encontram disponíveis no mercado das ideias –, este texto assumirá um tom deliberadamente didático e recuperará informações já publicadas em Kenedy (2015 e 2019) e exploradas em Teixeira (2015). Nas seções que se seguem, serão apresentadas as informações fundamentais que configuram um verdadeiro passo a passo para o delineamento, a implementação e a análise de um projeto experimental.

Na maior parte do tempo, os conceitos experimentais serão ilustrados com uma pesquisa hipotética acerca da concordância verbal e da ambiguidade estrutural em português. Porém, numa seção específica, serão apresentados e discutidos os aspectos experimentais de uma pesquisa funcionalista real conduzida por Teixeira (2015).

## **Projeto experimental**

O *modus operandi* da experimentação se assenta no modelo proposto por Popper (1959), conhecido como hipotético-dedutivo. Para esse filósofo da ciência, tal modelo seria poderoso o suficiente para delimitar claramente as fronteiras entre ciência empírica e hermenêutica interpretativista, uma vez que apenas as abordagens científicas podem derivar previsões testáveis experimentalmente que, assim, correm risco de falseamento. Uma pesquisa experimental se inicia com a delimitação de um conjunto de problemas que interessam ao pesquisador. Tais problemas são interpretados por uma dada teoria ou certa hipótese específica de algum arcabouço teórico. É dessa teoria/hipótese que pelo menos uma previsão comportamental deve ser derivada, a qual é então projetada num desenho experimental específico. Finalmente, esse desenho se apresenta aos participantes numa determinada tarefa experimental, da qual serão colhidos e interpretados os resultados relevantes para refutar ou confirmar as previsões do experimento. Esses

dados experimentais realimentarão de alguma maneira os problemas que deram origem à pesquisa e, dessa maneira, farão a ciência sobre o assunto avançar. A lógica da experimentação é ilustrada na sequência abaixo.

**Figura 1:** o método hipotético-dedutivo de Popper (1959).

Problema > Hipótese > Previsão > Experimentação > Resultados > Problema

Fonte: Autoria própria.

A experimentação em linguística está ao alcance de qualquer pesquisa que possa se submeter ao ciclo acima ilustrado. Como será visto nas seções a seguir, os resultados obtidos num experimento linguístico qualquer são dados de natureza comportamental que, por hipótese, decorrem das variáveis controladas nos estímulos apresentados durante uma tarefa experimental.

Qualquer pesquisa experimental deve possuir um protocolo metodológico que explicita, com a máxima clareza, a maneira pela qual a pesquisa é concebida e conduzida. Tal protocolo garante, por um lado, a objetividade do processo científico e, por outro, enseja a replicação de um estudo em particular, com adaptações ou não, por outros pesquisadores interessados no tema. O trabalho experimental tem início com a especificação dos problemas da pesquisa e, a partir desse ponto, o pesquisador formulará, com base nas hipóteses derivadas de sua teoria, alguma previsão comportamental passível de observação empírica. Há um número incalculável de previsões dessa natureza, em todos os níveis linguísticos. Uma vez estabelecida sua previsão, o pesquisador deverá dar continuidade a seu projeto experimental selecionando uma técnica de pesquisa compatível com as suas previsões. Logo depois, uma série de especificações metodológicas devem ser cumpridas: o pesquisador definirá a tarefa experimental, delineará as variáveis independentes e as variáveis dependentes relevantes para a pesquisa, estabelecerá as condições do experimento, formulará os estímulos experimentais, os estímulos de controle e os estímulos distrativos, selecionará e distribuirá os participantes da tarefa experimental de alguma maneira organizada, aplicará o experimento e, finalmente, organizará e interpretará os seus resultados.

A experimentação é uma metodologia muito diferente da análise de *corpus* porque, nesta, um sem-número de variáveis não controladas concorrem para a realidade ali registrada, de modo que diferentes *corpora* ou diferentes análises de um mesmo *corpus* podem engendrar conclusões muito diferentes acerca de um mesmo fenômeno, ao passo que, naquela, as variáveis em questão são previamente controladas e as previsões comportamentais são enunciadas antes de os dados empíricos da pesquisa serem produzidos. As previsões de um mesmo experimento devem ser confirmadas (ou refutadas) por pesquisadores independentes, desde que esses sigam os mesmos protocolos experimentais. Isso não acontece com as pesquisas observacionais, já que nelas um pesquisador pode selecionar o corpus de seu interesse e, dentro dele, pode selecionar os dados que lhe convêm (e ignorar deliberadamente ou não os dados que não lhe convêm). A experimentação é também muito diferente da hermenêutica interpretativa porque essa última é uma metodologia que pode se abster completamente de sustentação empírica ou pode selecionar alguns dados empíricos a dedo, conforme a conveniência do pesquisador.

### **Modelos de experimento**

Existem dois tipos de medidas comportamentais que definem dois grandes grupos de técnicas experimentais: medidas on-line e medidas off-line. As medidas comportamentais on-line são aferidas durante o processamento cognitivo que uma pessoa realiza inconscientemente enquanto recebe um estímulo linguístico oral ou escrito. Medidas on-line envolvem o registro de tempos muito rápidos (anotados em milésimos de segundo) e/ou a mensuração de algum comportamento automático, como a movimentação dos olhos durante a leitura. Experimentos que recolham medidas on-line precisam ser necessariamente programados e aplicados em equipamentos especializados, como softwares desenvolvidos para a realização de tarefas comportamentais e hardwares projetados para monitorar os movimentos oculares humanos. Equipamentos dessa natureza são capazes de registrar dados comportamentais precisos e finos. Os softwares mais utilizados em pesquisas experimentais em linguística são Psychopy, Paradigm e Pyscope – que podem ser baixados pela internet com diferentes custos.

As técnicas experimentais on-line mais produtivas na pesquisa em linguística no Brasil são o monitoramento ocular, a leitura segmentada autocadenciada (também chamada leitura automonitorada) e a audição segmentada autocadenciada. Dados obtidos por meio de técnicas on-line refletem mais diretamente o funcionamento natural de uma língua, já que envolvem a medição de comportamentos reflexos independentes da inspeção consciente dos participantes de um experimento.

Por sua vez, medidas comportamentais off-line são aferidas após a conclusão das reações reflexas disparadas por uma informação linguística e, por conseguinte, envolvem reflexões mais conscientes e deliberadas por parte daqueles que participam de uma tarefa experimental. Esse tipo de experimentação não impõe a necessidade de softwares ou hardwares especializados, embora preferencialmente seja realizado em equipamentos simples, como programas gratuitos e computadores com caixas de respostas (joysticks), que permitem, inclusive, a medição do tempo consumido durante a realização de uma tarefa. Alguns experimentos linguísticos off-line podem ser realizados até mesmo por meio de um formulário impresso, a ser respondido com uma caneta, como o julgamento de aceitabilidade de frases ou a produção linguística induzida (num preenchimento de questionário), ou por meio de gravações de áudio ou vídeo com equipamentos amadores comuns.

As técnicas experimentais off-line mais utilizadas em pesquisas empíricas são julgamento de aceitabilidade, produção induzida de fala ou escrita, reconhecimento de palavras, respostas a perguntas interpretativas e *priming*. Em função de sua natureza mais consciente, dados comportamentais off-line apenas indiretamente refletem a realidade natural de uma língua.

Os paradigmas experimentais on-line e off-line citados são apenas os mais comuns na pesquisa linguística. A maior parte deles reflete interesses específicos de estudiosos preocupados com a noção de competência linguística (no sentido de Chomsky, 1965) e, assim, podem não despertar o interesse de pesquisadores funcionalistas. Não obstante, as técnicas experimentais de quaisquer tipos são infinitamente ampliáveis, de modo que um pesquisador imaginativo pode criar um novo paradigma ou adaptar livremente paradigmas existentes de modo a tornar viável a experimentação linguística sob seus interesses funcionalistas específicos.

Ao definir a técnica experimental de sua pesquisa, um pesquisador funcionalista deverá considerar qual tipo de medida comportamental é a mais adequada para caracterizar, no desempenho linguísticos dos falantes, o efeito da variável sob investigação: uma medida on-line – mais automática, involuntária e inconsciente – ou uma medida off-line – mais deliberada e consciente. Medidas on-line geralmente são exploradas em tarefas de compreensão linguística, tais como leitura ou audição de estímulos, ao passo que medidas off-line podem ser exploradas tanto em tarefas de produção linguística, como julgamento de aceitabilidade, fala ou escrita induzida, quanto em tarefas de compreensão como a leitura.

### **Tarefa dos participantes**

Um experimento linguístico prototípico utiliza participantes que não sejam especialistas em linguística ou estudiosos de gramática. Por essa razão, a tarefa experimental de uma pesquisa deve ser a mais clara, simples e objetiva possível – despojada de metalinguagem. Se um pesquisador desejar, por exemplo, verificar se os falantes de uma língua percebem a transitividade de um verbo varia de acordo com a afetação que ele promove sobre seu respectivo objeto, tal estudioso deve elaborar alguma tarefa comportamental qualquer que possa refletir essa interpretação – como apontar para uma figura, parafrasear o dito – e nunca explicitamente perguntar algo como “o objeto foi afetado?”.

Em experimentos on-line, as tarefas experimentais mais comuns são (1) ler ou ouvir uma frase apresentada em segmentos – que podem ser sintagmas ou palavras – numa tela de computador, os quais são disparados pelo próprio participante, conforme sua velocidade natural de leitura/audição, enquanto o computador registra o tempo que é consumido em cada segmento da frase, (2) ler com naturalidade palavras ou frases numa tela de computador, enquanto um equipamento monitora o comportamento ocular inconscientemente produzido ao longo da leitura. No caso de (1), os tempos de reação mais lentos ou mais rápidos em cada segmento da frase servirão para indicar medidas comportamentais (*reaction time*, no caso) que, a princípio, decorrem de alguma previsão da pesquisa – por exemplo, objetos menos

afetados pelo verbo serão lidos mais rapidamente do que objetos mais afetados. A mesma lógica se aplica a (2), já que esses dados comportamentais (no caso, *total fixation*, *look back*, *first pass* etc.) podem indicar que uma dada variável – por exemplo, a afetação do objeto – influenciou a percepção cognitiva da frase, algo refletivo no movimento ocular.

Já em experimentos off-line, as tarefas mais típicas são (1) julgar frases binariamente (declarando-as aceitáveis *versus* inaceitáveis), (2) julgar frases por meio de escalas de aceitabilidade (atribuindo-lhes uma nota, por exemplo de 0 a 5), (3) preencher formulários dando continuidade a uma frase oral ou escrita, (4) declarar o reconhecimento ou a familiaridade com uma palavra ou uma expressão apresentada por escrito ou oralmente, (5) responder a questionários variados, (6) declarar o reconhecimento ou a familiaridade com um determinado estímulo após a apresentação de outro (*priming*). Todas essas medidas comportamentais off-line devem refletir algum tipo de fator controlado pelo pesquisador durante a pesquisa – por exemplo, estruturas mais gramaticalizadas tenderão a receber notas mais altas em tarefas de julgamentos e certos tipos de complemento verbal podem ser mais previsíveis em certas circunstâncias ou certas construções.

Numa etapa preliminar fundamental no delineamento de um projeto experimental em linguística funcional, o pesquisador deverá definir que tipo de tarefa poderá produzir dados comportamentais tais que permitam verificar o eventual efeito da variável selecionada para a pesquisa. Com efeito, seja por meio de técnicas on-line ou off-line, uma tarefa experimental produzirá um conjunto de respostas que deverão decorrer, por hipótese, das variáveis selecionadas para a pesquisa. Não é por outra razão que a delimitação precisa das variáveis em investigação em seu projeto experimental constitui a base de todas as pesquisas que adotam essa metodologia.

### **Variável independente e variável dependente**

Definir as variáveis de pesquisa é uma das etapas mais importantes no delineamento de um projeto experimental. Por um lado, o linguista deverá delimitar os fenômenos que, de acordo com sua hipótese de trabalho,



são capazes de provocar certo comportamento durante a execução de uma tarefa experimental. Por outro lado, ele também deve estabelecer qual é o tipo de resposta que registrará esse comportamento. Os fenômenos selecionados como possíveis causadores do comportamento são denominados variáveis independentes, ao passo que as medidas comportamentais aferidas numa tarefa se denominam variáveis dependentes.

**Figura 2:** as variáveis de um experimento

**Estímulo (variável independente) > Comportamento (variável dependente)**

Fonte: Autoria própria

Variáveis independentes são também denominadas variáveis controladas, enquanto variáveis dependentes podem ser chamadas de variáveis de resposta ou medidas dependentes. Por exemplo, numa pesquisa hipotética sobre a concordância verbal em português, um pesquisador poderia definir a posição do sujeito relativamente ao verbo como uma variável independente capaz de desencadear maior ou menor índice de estabelecimento da concordância numa tarefa de produção induzida. Nesse caso, os índices reais de concordância encontrados no experimento caracterizariam a variável dependente em tal estudo.

Num experimento, pode haver mais de uma variável independente e mais de uma variável dependente. O importante é que todas essas variáveis sejam definidas com a maior clareza e, mais do que isso, é fundamental que haja o máximo de controle sobre outras variáveis que podem igualmente provocar ou influenciar determinado comportamento nos participantes do experimento. Continuando com o exemplo de uma pesquisa sobre concordância verbal, existem muitos fatores, para além da posição do sujeito em relação ao verbo, que podem influenciar o estabelecimento ou não da concordância, tais como natureza do verbo, o número e o tipo de itens intervinientes entre sujeito e verbo, a distinção morfofonológica entre formas verbais do singular e do plural, a frequência e a familiaridade de verbos específicos, o grau de instrução e literacia dos participantes da tarefa etc. Sendo assim, numa pesquisa sobre concordância, o estudioso não poderia deixar de

controlar tais variáveis, do contrário as chamadas variáveis de confusão, isto é, aquelas que não foram controladas pelo pesquisador, podem influenciar o comportamento registrado como variável dependente e, assim, enfraquecer ou anular o poder explicativo atribuído à variável independente.

### **Níveis de variáveis e condições experimentais**

Nos estímulos a serem apresentados aos participantes de uma tarefa experimental, as variáveis independentes irão se concretizar em formas linguísticas específicas que realizam as condições experimentais da pesquisa. Por exemplo, no hipotético estudo sobre concordância verbal, se a seleção da posição relativa entre verbo e sujeito for eleita como uma variável independente, então essa variável será concretizada em dois níveis que, no caso de um experimento com uma única variável independente, serão também as duas condições experimentais: (1) estímulos em que o sujeito antecede o verbo e (2) estímulos em que o sujeito sucede o verbo.

Nos experimentos que contam com mais de uma variável independente, as condições são concretizadas a partir da multiplicação entre os níveis de cada variável selecionada. Imagine-se, por exemplo, que, ainda num experimento sobre concordância verbal, o traço de animacidade do sujeito também fosse selecionado como variável independente, ao lado da posição do sujeito. Nesse caso, o respectivo experimento teria quatro condições experimentais, resultantes da multiplicação entre os dois níveis de cada uma das duas variáveis independentes escolhidas: (1) sujeito animado + anteposição ao verbo, (2) sujeito inanimado + anteposição ao verbo, (3) sujeito animado + posposição ao verbo e (4) sujeito inanimado + posposição ao verbo. Nesse exemplo, ter-se-ia um experimento com o desenho fatorial  $2 \times 2$ , no qual há duas variáveis independentes, cada qual com dois níveis que, combinados, geram quatro condições experimentais. Se, nesse experimento, fosse incluída uma terceira variável independente, suponha-se, a conjugação verbal, com três níveis (1ª, 2ª e 3ª conjugações), então o desenho do experimento passaria a ser  $2 \times 2 \times 3$ , o que daria à luz doze condições experimentais.

**Quadro 1:** ilustração dos fatores de um hipotético projeto experimental sobre concordância verbal.

Técnica off-line, tarefa experimental: numa pequena frase, preencher uma lacuna com verbo flexionado.
Variáveis independentes: (1) posição do verbo em relação ao sujeito, (2) animacidade do sujeito.
Níveis das variáveis independentes: variável (1) anteposição ou posposição do verbo ao sujeito, variável (2) sujeito animado ou sujeito inanimado
Desenho fatorial: 2x2.
Condições experimentais: (1) anteposição ao verbo + sujeito animado, (2) anteposição ao verbo + sujeito inanimado, (3) posposição ao verbo + sujeito animado e (4) posposição ao verbo + sujeito inanimado.
Variável dependente: índice de concordância entre sujeito e verbo.

Fonte: Autoria própria.

No quadro 1, apresenta-se um hipotético projeto experimental funcionalista que selecionou a posição do verbo relativamente ao sujeito e o traço semântico de animacidade no sujeito como as variáveis independentes relevantes para provocar a medida comportamental de produzir mais ou menos concordância verbal (variável dependente). Nesse projeto, a técnica off-line de preencher lacunas em frase foi selecionada como a mais adequada para registrar o comportamento linguístico derivado da hipótese de pesquisa como previsão – por exemplo, se suponha que sujeitos antepostos e animados terão maiores taxas de concordância com o verbo.

### **Estímulos linguísticos experimentais**

Na prática de uma tarefa de um experimento qualquer, as condições experimentais assumem alguma forma linguística concreta e específica, seja um morfema, uma palavra, um sintagma, uma oração etc. Essas formas concretas são denominadas estímulos.

A tradição das pesquisas experimentais estabelece que cada condição experimental deve ser apresentada aos participantes de uma tarefa pelo menos quatro vezes, na forma de quatro estímulos verbais distintos,

de modo que um padrão de reação a tal condição, se houver, possa ser detectado. De fato, se um estímulo fosse apresentado a um participante uma única vez, não seria possível saber se o respectivo comportamento provocado por esse estímulo decorreu de uma variável independente ou se deveu-se a um evento único e aleatório. Para que o efeito de uma variável possa ser identificado e separado de um evento aleatório, quatro exposições de estímulos de uma mesma condição experimental a cada participante específico costuma ser o suficiente.

Com efeito, na elaboração de estímulos, o pesquisador deve aplicar o máximo de esmero. Ele deve controlar, dentre outros fatores, a extensão dos estímulos de cada condição, em número de sílabas ou palavras, bem como a frequência e a familiaridade dos itens lexicais utilizados. Esse controle visa à tentativa de evitar que fatores outros (variáveis de confusão), diferentes da variável independente, possam afetar o comportamento dos participantes da tarefa – por exemplo, o tempo de reação a um determinado estímulo em, suponha-se, duas condições experimentais deve variar em função da variável independente selecionada na pesquisa, e não porque os estímulos de uma condição possuem muito mais palavras do que o da outra condição e, assim, obviamente, demandam mais tempo de reação; ou, ainda, a sensação de estranhamento a um estímulo deve decorrer da variável controlada pelo pesquisador e não da presença de uma palavra rara ou ambígua que interferiu no julgamento.

É esse controle esmerado que torna os dados obtidos via experimentação tão distintos daqueles coletados em pesquisas de *corpus*. Num ambiente natural, como o registado em um *corpus*, um sem-número de variáveis linguísticas e não linguísticas se encontram totalmente embaraalhadas, de modo que o comportamento ali registrado pode ter sido causado por inúmeros fatores – e não apenas por aqueles que o pesquisador acredita ser especialmente determinante para o fenômeno investigado.

### **Estímulos linguísticos distrativos**

Além do cuidado necessário na equivalência entre os estímulos de cada condição, o pesquisador também deve utilizar, num experimento, estímulos

distrativos. Tais estímulos não devem possuir nenhuma relação com as variáveis independentes da pesquisa e cumprem apenas a função de evitar que o participante reconheça (explicitamente ou não) o tipo de fenômeno linguístico que está sendo apresentado nos estímulos das condições experimentais.

Por convenção, os estímulos distrativos de um experimento devem perfazer pelo menos dois terços do número total de estímulos da tarefa. Sendo assim, num experimento com apenas uma variável independente e duas condições experimentais, cada participante será exposto a pelo menos oito estímulos experimentais (quatro estímulos de cada condição) e dezesseis estímulos distrativos.

### **Estímulos linguísticos de controle**

Um experimento pode conter também estímulos de controle. A função desses é permitir o cotejo direto entre uma condição em que determinado fenômeno está presente com outra condição (a condição de controle) em que ele é ausente. Por exemplo, numa pesquisa sobre ambiguidade sintática, o desempenho dos participantes diante de estímulos estruturalmente ambíguos (ex. “O policial viu o suspeito com um binóculo”) deve ser comparado com o que ocorre com estímulos de controle não ambíguos (ex. “O policial viu o suspeito com sotaque do sul”), o que permitirá a identificação de eventuais reações específicas na condição com ambiguidade. Nesses exemplos, caso a frase não ambígua provoque o mesmo comportamento registrado diante da frase ambígua – por exemplo, tempo de reação durante a leitura –, isso indicará que o participante não percebeu imediatamente a ambiguidade sintática manipulada na frase. Esse resultado poderá confirmar (ou refutar) alguma previsão formulada pelo estudioso dos efeitos do contexto sobre a sintaxe, sobre a percepção dos sintagmas preposicionados etc.

### **Randomização**

Os estímulos experimentais e distrativos (bem como os de controle, se houver) devem ser apresentados ao participante de maneira randomi-

zada, isto é, aleatória, sem qualquer padrão de sequência linear. Softwares especializados em experimentos já produzem randomização de estímulos de maneira automática. No caso de pesquisas mais simples, com formulário de papel a ser preenchido a caneta, é o próprio pesquisador que deve sortear aleatoriamente a ordem de apresentação dos estímulos, embaralhando os dois terços de distrativos ao terço final de experimentais. A randomização é necessária para evitar que o participante perceba a existência de padrões na apresentação dos estímulos e, então, passe a se comportar de acordo com esses padrões e não em função das variáveis da pesquisa.

No quadro a seguir, ilustra-se a randomização entre estímulos experimentais e distrativos relativos ao hipotético desenho experimental citado mais acima no quadro 1. Note-se que, no caso, apenas o segundo e o último estímulos são experimentais, e todos os demais, distrativos. No exemplo, os estímulos experimentais pertencem, respectivamente, às condições experimentais “anteposição ao verbo + sujeito animado” e “posposição ao verbo + sujeito animado”. Trata-se, portanto, de apenas um fragmento do total de estímulos que esse estudo hipotético deveria conter. Para que tal projeto fosse considerado completo, faltariam ainda a inserção de estímulos das demais condições experimentais, a criação de mais frases, de modo a configurar a necessária exposição mínima de quatro estímulos de condição experimental, e a elaboração de estímulos distrativos na proporção “dois para um” em relação às frases experimentais.

**Quadro 2:** ilustração de duas frases experimentais e quatro distrativas num hipotético estudo sobre concordância.

A CPI da _____ investigará omissões governo durante a pandemia.
_____ as férias na semana passada em todo RJ.
As disputas futebolísticas em todo o mundo foram _____ em 2020.
Não existem _____ de que aprender uma nova língua faz bem às pessoas.
_____ que maneira as pessoas idosas devem se exercitar sem riscos de lesões?
As obras _____ no mês de março, mas ainda estão em curso.

Fonte: Autoria própria

## Agrupamento de participantes

Antes de mais nada, é preciso considerar que os participantes de um determinado experimento podem configurar, por eles mesmos, uma variável independente. Isto é, se o pesquisador assumir que o comportamento a ser registrado numa tarefa pode variar de acordo com o tipo de participante (por exemplo, bilíngues *versus* monolíngues, com patologia *versus* sem patologia, estudantes de L2 fluentes *versus* não fluentes, pessoas com nível superior *versus* pessoas analfabetas etc.), então o experimento possuirá uma variável grupal (também chamada de fator grupal). Se não for esse o caso, o linguista deverá apenas determinar o perfil sociocultural das pessoas que podem participar da tarefa – fatores como idade, sexo, escolaridade, região e outros que se mostrem relevantes – bem como deverá estabelecer como se dá a distribuição das condições experimentais pelos participantes.

Na distribuição dos participantes pela tarefa experimental, existem duas possibilidades a serem adotadas. Na primeira delas, todos os participantes são expostos a todas as condições experimentais. Essa distribuição se denomina “dentre participantes” (*within-subjects*, em inglês) ou intraparticipantes. Na outra, cada participante é exposto a uma e somente uma condição experimental. Nesse caso, haveria um grupo de participantes separado para cada condição do experimento, razão pela qual tal distribuição se denomina “entre participantes” (*between subjects*, em inglês) ou interparticipantes.

A distribuição dentre participantes tem a vantagem de exigir menos indivíduos desempenhando as tarefas do experimento, já que todos são utilizados em todas as condições. Porém, essa opção tem a desvantagem de facilitar o efeito de familiaridade (isto é, o aprendizado da tarefa durante o experimento) e a identificação explícita ou tácita de padrões nos estímulos, se for levado em consideração que uma mesma pessoa é estimulada por todas as condições experimentais pelo menos quatro vezes em cada condição. Com a distribuição entre participantes, as chances de ocorrer o efeito de familiaridade são menores, mas para isso é demandado um número maior de participantes, dado que eles devem se distribuir em igual número por cada uma das condições do experimento.

Quando o pesquisador opta pela distribuição dentre participantes, os estímulos experimentais devem receber um tratamento adicional: o controle num quadrado latino. Esse recurso permite o balanceamento dos estímulos presentes em cada condição experimental, evitando-se que o mesmo participante seja exposto a estímulos muito parecidos, de diferentes condições, distintos apenas em função da variável independente do experimento. Com o quadrado latino, se estabelece que um participante numa distribuição *within-subjects* será exposto a todas as condições (*type*) do experimento, mas em cada condição serão usados itens lexicais específicos (*tokens*), de modo que a relação entre essas condições não se torne evidente durante a realização da tarefa. Por exemplo, na citada pesquisa sobre concordância verbal em português, um participante que veja um estímulo da condição “verbo + sujeito” como “Chegaram as encomendas” seria exposto, na condição “sujeito + verbo”, a um estímulo como “As reclamações cessaram” (e não “As encomendas chegaram”). O quadro a seguir ilustra a distribuição balanceada de estímulos no quadrado latino.

**Quadro 3:** Controle da distribuição dos estímulos experimentais num quadrado latino. Note-se que “B” é uma versão idêntica do estímulo “A” com o mesmo número, exceto pela condição experimental manipulada em que se insere.

	Grupo 1	Grupo 2
Condição 1	Estímulos 1a, 2a, 3a, 4a	Estímulos 5a, 6a, 7a, 8a
Condição 2	Estímulos 5b, 6b, 7b, 8b	Estímulos 1b, 2b, 3b, 4b

Fonte: Autoria própria

## Rodando o experimento

Após o longo percurso de elaboração de um experimento, o pesquisador deve manter a vigilância mesmo durante a aplicação das tarefas experimentais com os participantes. Eles devem receber todas as instruções e demonstrações necessárias para a perfeita realização da tarefa e devem ser submetidos a um breve aquecimento, na presença do experimentador, por



meio de um pré-teste constituído somente de estímulos distrativos, cujo objetivo é evitar que o desempenho durante o experimento propriamente dito possa ser prejudicado devido a questões mecânicas ou a fatores decorrentes da incompreensão da tarefa. Quando os participantes demonstram ter compreendido perfeitamente o que devem fazer durante o experimento, o aquecimento pode ser finalizado e o experimento, iniciado.

Dando início ao experimento, os participantes devem se encontrar sozinhos, numa sala isolada, sem elementos que possam distrair a sua atenção e interferir na realização da tarefa. Deve-se registrar o tempo médio despendido na tarefa, bem como, após a realização do experimento, é conveniente receber um feedback dos participantes, a fim de verificar se eles reportam alguma anomalia ou mesmo se confessam ter identificado o padrão subjacente à tarefa ou ao fenômeno linguístico em análise.

### **Conselho de ética e Termo de consentimento livre e esclarecido**

Em função de lidar com seres humanos, um projeto experimental em linguística deve ser aprovado pelo Conselho (ou Comitê) de Ética em Pesquisa da área de concentração respectiva em que o projeto se insere em sua Universidade. Esse projeto deve indicar claramente que as tarefas experimentais não oferecem quaisquer riscos de constrangimento físico ou moral aos participantes e são completamente não invasivos e inofensivos. A aprovação pelo Conselho de Ética, com o respectivo número processual atribuído ao projeto, pode ser um fator condicionante para a publicação de artigos derivados dos achados do experimento, sobretudo em revistas científicas prestigiadas na Europa e nos EUA.

Uma das consequências da aprovação do projeto pelo Conselho será a necessidade de obter de cada participante a sua anuência formal em se submeter à pesquisa experimental em questão, por meio de sua assinatura num termo de consentimento livre e esclarecido. Esse termo deve conter um cabeçalho com a identificação da Universidade e do grupo de pesquisa responsável pelo experimento, além de indicar com clareza e brevidade que o participante se dispõe de maneira livre e esclarecida a se voluntariar no projeto experimental, que não lhe oferece nenhum risco de qualquer natureza.

## Análise estatística

Com o experimento concluído, é possível passar à análise dos resultados para verificar se os dados comportamentais coletados junto aos participantes se encaminham ou não em favor das previsões da pesquisa. Nesse momento, o pesquisador precisará ou contratar os serviços de um profissional de estatística ou utilizará, ele mesmo, softwares de pacotes estatísticos para organizar e interpretar os resultados numéricos do experimento. A depender da variável dependente em questão, do tipo de distribuição dos participantes e da normalidade distributiva dos dados comportamentais coletados, diferentes tipos de análise estatística podem ser aplicados. Os mais comuns são análise da variância, teste T, qui-quadrado e regressão. Há uma grande gama de softwares estatísticos no mercado. O mais recomendado é o R, que possuiu uma adaptação para a interface do Microsoft Excel no aplicativo brasileiro ActionStat. Outro software largamente empregado no Brasil é o SPSS (Statistical Package for the Social Sciences<sup>1</sup>).

O objetivo de um teste estatístico é, por um lado, descrever a distribuição dos dados obtidos no experimento e, por outro, verificar se o comportamento típico encontrado nesses dados pode ser interpretado como provavelmente decorrente das variáveis independentes selecionadas na pesquisa – ou se, alternativamente, são grandes as chances de tal comportamento ter sido provocado por fatores aleatórios. O famoso p-valor utilizado nas análises estatísticas é, justamente, o resultado de cálculos matemáticos complexos que medem a atuação das variáveis da pesquisa no cotejo com o acaso. Simplificadamente, estudos experimentais em linguística assumem um nível de significância de no mínimo 95%, o que significa dizer que um p-valor igual ou inferior a 0.05 indica baixa probabilidade (igual ou inferior a 5%) de os resultados da pesquisa terem sido gerados

---

<sup>1</sup> R é um programa de código aberto disponível em <https://www.r-project.org/>. Sua utilização requer conhecimento da linguagem de programação utilizada pelo aplicativo, com os respectivos comandos exigidos para a formatação de dados e as análises estatísticas específicas. ActionStat é uma versão paga de um suplemento do Microsoft Excel, disponível em <http://loja.portalaction.com.br/>, que utiliza em segundo plano os fundamentos do próprio R e, dessa forma, dispensa conhecimentos de programação. O SPSS é um software corporativo da IBM, disponível em <https://www.ibm.com/br-pt/analytics/spss-statistics-software>, para cujo aprendizado existem diversos tutoriais em vídeo e livros na internet.

aleatoriamente. Com resultados dentro dessa margem, os dados reunidos por um projeto experimental em linguística podem ser interpretados como indicadores da atuação de uma ou mais variáveis independentes e/ou da interação entre duas ou mais delas.

### **Um exemplo com pesquisa funcionalista**

Esta seção destina-se à apresentação dos experimentos utilizados por Teixeira (2015), conduzidos a fim de verificar efeitos comportamentais resultantes da interpretação de certas expressões como marcadores discursivos, que funcionam como uma forma verbal e um clítico, representando um único signo linguístico, um pareamento de forma e sentido, e não mais como parte de um predicado transitivo circunstancial, constituído por dois elementos: um verbo e um pronome adverbial locativo. Serão apresentados três experimentos com as suas informações fundamentais.

Em Teixeira (2015), adotou-se um protocolo experimental off-line. Seu objetivo foi o de fornecer evidências para a confirmação da hipótese de que os marcadores sob estudo seriam formalmente configurados por um verbo e um locativo, mas que seriam acessados pelos falantes holisticamente, isto é, como um único signo linguístico.

Teixeira (2015) utilizou três técnicas experimentais off-line: escolha de paráfrase, escala de aceitabilidade e seleção de segmentação. Apesar de serem técnicas diferentes, todas elas estiveram a serviço do mesmo objetivo. Na primeira técnica, a tarefa do participante foi escolher uma frase entre quatro alternativas cujo sentido ele julgasse mais parecido com uma frase alvo de referência. Na segunda, a tarefa foi medir o grau de aceitabilidade do participante em relação à combinação “verbo + locativo”, que atuava como marcador discursivo. Na terceira técnica, a tarefa foi escolher, entre um conjunto de opções, a alternativa que melhor apresentasse a segmentação entre palavras ou grupo de palavras de acordo com a percepção de unidade de sentido manifestada pelo participante.

No total, 73 alunos voluntários participaram dos experimentos. Todos eram falantes nativos do português brasileiro e estudantes universitários. Desse número, 26 alunos do 1º semestre participaram da técnica de escolha

de paráfrase; 23, do 1º semestre, da técnica de julgamento de aceitabilidade; e 24, do 6º semestre, da de segmentação. Para atender aos objetivos da pesquisa, foi elaborado um questionário para cada técnica, cada qual com 11 sentenças estímulo, sendo que o número de alternativas de resposta para cada questionário variou de acordo com a técnica utilizada. A escolha do número de alternativas de respostas possíveis se deu em virtude da complexidade do comando contido na tarefa-experimento. Os detalhes de cada experimento serão descritos nas subseções a seguir.

Além dos resultados em termos percentuais, foi empregado o teste Qui-quadrado de Pearson,  $(X^2)^2$ , que tomou os resultados estatísticos da variável independente única “construção marcadora discursiva” adotada nos três experimentos, cujas condições, coincidentes com os níveis da variável, por tratar-se de monovariável, foram: ‘escuta aqui’, ‘(es)per(a) aí’, ‘(es)pera lá’, ‘(es)tá aí’, ‘olha aí’, ‘olha aqui’, ‘olha lá’, ‘vá lá’, ‘vamo(s) lá’, ‘vê lá’ e ‘vem cá’. Na apresentação dos resultados a seguir, a diferença estatisticamente significativa das respostas-alvo em relação às opções de controle será apresentada considerando o p-valor menor do que 0,05. A resposta alvo representa a escolha derivada da previsão da hipótese de pesquisa de Teixeira (2015); portanto, o comportamento dos participantes na escolha dessa resposta iria ao encontro das hipóteses do trabalho, mas qualquer outra opção, não.

### **Experimento 1: escolha de paráfrase**

O objetivo particular deste experimento consistiu em o participante identificar a função específica dos marcadores citados. Para tanto,

---

<sup>2</sup> Um teste  $X^2$  é uma análise estatística que estima se a distribuição de frequência de um determinado fenômeno provavelmente decorreu do acaso e, assim, possui baixa probabilidade de se repetir no futuro ou, alternativamente, se tal distribuição possui pouca chance de ter sido gerada por fatores aleatórios e, por isso, provavelmente decorre de alguma variável independente. Se são baixas as chances de uma distribuição ter sido gerada por acaso (por convenção, abaixo de 5%), o  $X^2$  indicará a ação de alguma variável e a consequente chance de tal distribuição voltar a ser encontrada em futuras análises com as mesmas condições experimentais. Esse tipo de inferência estatística decorre de cálculos probabilísticos complexos que em muito se distinguem de uma descrição estatística simples baseada apenas em percentuais de ocorrência. O nome desse tipo de teste de estatística inferencial é uma homenagem ao matemático britânico Karl Person, que muito contribuiu para o estabelecimento da estatística como área do conhecimento, no início do século XX.

elaborou-se uma “sentença aceitável”, a resposta alvo, isto é, uma paráfrase que procurava manter o sentido mais similar possível ao do da “sentença teste”, ou seja, o estímulo de referência. Nesse primeiro experimento, a tarefa do participante foi a de escolher uma entre quatro alternativas cujo sentido ele, por si mesmo, julgasse mais semelhante em comparação à sentença teste apresentada no enunciado.

Dessa maneira, além de se verificar se a combinação foi lida como uma unidade, testou-se também se a unidade de sentido foi percebida como a marca explícita que introduz a opinião do autor, ou seja, se ela foi aceita como um marcador discursivo. Nesse sentido, se o participante escolhesse a sentença aceitável, isto é, a paráfrase alvo, ele demonstraria (indiretamente, por meio dessa variável comportamental) ter compreendido que a combinação verbo e locativo estava atuando como um marcador discursivo. Consequentemente, assumiria a unidade de sentido e forma como aceitável. Já se escolhesse uma das outras paráfrases (estímulo de controle), essa previsão não seria confirmada.

### ***Variáveis e condições***

Como dito, a variável independente, em todos os experimentos de Teixeira (2015), é a construção verbal marcadora discursiva, cujas condições, isto é, o conjunto das ocorrências destacadas mais acima, estão inseridas em 11 trechos de textos distintos. Dessa forma, o total dos estímulos apresentados aos participantes perfaz 11 ocorrências (*trials*), uma para cada marcador. A sentença aceitável (isto é, a paráfrase alvo) e as três sentenças de controle (isto é, paráfrases com sentidos não previstos pela hipótese da pesquisa) compõem as quatro alternativas de resposta dispostas randomicamente em cada um dos 11 *trials*.

### ***Materiais***

Como material (sentenças teste), foram selecionados trechos de textos curtos do *corpus* utilizados em outras seções de Teixeira (2015), com algumas adaptações necessárias, mas com informações suficientes para que

o participante pudesse compreender o contexto específico da atuação do marcador. A sentença teste, destacada no quadro que se segue, está inserida no trecho aludido. Dessa maneira, tornou-se possível verificar como a função particular de cada marcador, como ocorrência da variável independente, foi percebida pelo participante. A distribuição da sentença estímulo teste, da paráfrase alvo e das três paráfrases de controle estão descritas a seguir:

**Quadro 4** – Descrição dos materiais do experimento 1.

Sentença teste (estímulo)		Paráfrase alvo	Paráfrase de controle
Descrição	Função Mar- cador		
Olha aqui, Adriano, isto é um aviso.	Repreensão Asseveração	4) – Entenda, Adriano, isto foi um aviso.	1) – De certo feitio, Adriano, isto foi um aviso. 2) – Melhor, Adriano, isto foi um aviso. 3) – Rapaz, Adriano, isto foi um aviso.
Eu nem sei não - perai dizem que - jumento é...	Interrupção Retificação	4) - eu nem sei não - retificando - dizem que - jumento é...	1) - eu nem sei não - entenda... - dizem que - jumento é... 2) - eu nem sei não - verdade - dizem que - jumento é... 3) - eu nem sei não - certamente - dizem que - jumento é...
Escuta aqui, Arão.	Repreensão Intimidação	2) – Atenção, Arão.	1) – Por favor seu Arão. 3) – Companheiro Arão. 4) - Senhor Arão.
Espera lá, marciano,	Temporização Contra-argumentação	3) – Alto lá, marciano.	1) – Ok, marciano. 2) – Será, marciano. 4) – Mentira, marciano.
Como não! Está aí!	Constatação	3) Como não! Constate!	1) Como não! Mesmo! 2) Como não! Já era! 4) Como não! Conseguiu!
Olha aí (risos), estou do lado do ministro Pertence.	Repreensão Provocação	1) Perceba (risos), estou do lado do ministro Pertence.	2) Por favor (risos), estou ao lado do ministro Pertence. 3) Tecnicamente (risos), estou ao lado do ministro Pertence. 4) Sinceramente (risos), estou ao lado do ministro Pertence.

Olha lá, hein! Aquilo é segredo	Repreensão Advertência	1) – Previno-te, hein! Aquilo é segredo.	2) – Verdade, hein! Aquilo é segredo. 3) – Sabe, hein! Aquilo é segredo. 4) – Senhor, hein! Aquilo é segredo.
Se ao menos oferecesse cursos, biblioteca ou terapias para essas pessoas, vá lá	Concessão	3) Se ao menos oferecesse cursos, biblioteca ou terapias para essas pessoas, concordaria.	1) Se ao menos oferecesse cursos, biblioteca ou terapias para essas pessoas, você sabe. 2) Se ao menos oferecesse cursos, biblioteca ou terapias para essas pessoas, seria verdade. 4) Se ao menos oferecesse cursos, biblioteca ou terapias para essas pessoas, temo por elas.
Um café forte para dar uma sacudida, e vamos lá!	Exortação	3) Um café forte para dar uma sacudida, e recomêço!	1) Um café forte para dar uma sacudida, e tudo certo. 2) Um café forte para dar uma sacudida, e que desanimo. 4) Um café forte para dar uma sacudida, e preciso fazer.
Na hora de sair de casa, falo com Nossa Senhora Aparecida: Vê lá, hein?	Repreensão Prevenção	2) "Na hora de sair de casa, falo com Nossa Senhora Aparecida: cuidado, hein!"	1) "Na hora de sair de casa, falo com Nossa Senhora Aparecida: não sei, hein?" 3) "Na hora de sair de casa, falo com Nossa Senhora Aparecida: tristeza, hein?" 4) "Na hora de sair de casa, falo com Nossa Senhora Aparecida: alegria, hein?"
Mas, vem cá, fulano, sou teu amigo, que diabo!	Indagação Reflexão	1) - Mas, se atente, fulano, sou teu amigo, que diabo!	2) – Mas, que triste, fulano, sou teu amigo, que diabo! 3) – Mas, compreenda, fulano, sou teu amigo, que diabo! 4) – Mas, alegria, fulano, sou teu amigo, que diabo!

Fonte: Adaptado de Teixeira (2015, p. 267).

### ***Procedimentos***

O experimento foi elaborado em formulário impresso intitulado “Questionário”, composto de três folhas, preenchido pelo participante no mesmo momento de sua apresentação na frente do pesquisador e respondido individualmente. O tempo médio de sua realização foi de 25 minutos.

A aplicação foi realizada em sala de aula com todos os participantes presentes, distribuídos estrategicamente pelas carteiras da sala a fim de evitar o contato entre eles.

Inicialmente, os 26 participantes passaram por uma fase de aquecimento, composta por três exercícios com o mesmo formato do experimento em si, porém desenvolvidos com frases genéricas sem a variável manipulada no experimento. Nessa fase, informou-se aos participantes a dinâmica da tarefa e esclareceram-se eventuais dúvidas que surgiram, a fim de evitá-las na resolução da tarefa experimental. Na primeira parte do experimento, foram lidas as orientações contidas no questionário, em conjunto, esclarecendo que a intenção era a de perceber como cada participante compreendia a frase destacada. Portanto, apesar de a escolha ser de apenas uma entre as quatro alternativas, esclarece-se que não existia uma única opção certa ou errada. O comando da tarefa consistia em selecionar a alternativa que manteria o mesmo sentido da frase de estímulo, fazendo isso da forma mais espontânea possível, de modo que as respostas refletissem as intuições naturais dos participantes sobre o uso da língua.

### ***Resultados***

Como cada marcador é a expressão da variável independente, a expectativa foi a de que todos seriam lidos como unidade de sentido, destituídos das marcas separadas de verbo e de complemento com que atuavam no predicado transitivo circunstancial. O percentual exposto na tabela 1 representa a escolha da sentença alvo, ou seja, a alternativa que continha a resposta considerada como a previsão da pesquisa. Dessa forma, nesse experimento, para *olha aqui*, por exemplo, 24 dos 26 participantes escolheram a alternativa que representa a sentença alvo. Nesse sentido, 92.31% é o percentual de respostas que ocorreu de acordo com as previsões do trabalho.



**Tabela 1** – Resultado da sentença alvo no experimento 1.

MARCADOR	SENTENÇA ALVO (Previsão)		Qui-quadrado de Pearson	
	Relação escolha da previsão/total participantes	Percentual de resposta da previsão de escolha	p-valor	X <sup>2</sup>
ESCUTA AQUI	20/26	76,92%	p = 0,006039559	X <sup>2</sup> = 7,538461538
ESPERA AÍ	21/26	80.77%	p = 0,001701872	X <sup>2</sup> = 9,846153846
ESPERA LÁ	24/26	92.31%	p = 0,0000159	X <sup>2</sup> = 18,61538462
ESTÁ AÍ	14/26	53.85%	p = 0,694886602	X <sup>2</sup> = 0,153846154
OLHA AQUI	24/26	92.31%	p = 0,0000159	X <sup>2</sup> = 18,61538462
OLHA AÍ	19/26	73.08%	p = 0,01860293	X <sup>2</sup> = 5,538461538
OLHA LÁ	17/26	65.38%	p = 0,116664465	X <sup>2</sup> = 2,461538462
VÁ LÁ	19/26	73.08%	p = 0,01860293	X <sup>2</sup> = 5,538461538
VAMOS LÁ	21/26	80.77%	p = 0,001701872	X <sup>2</sup> = 9,846153846
VÊ LÁ	19/26	73.08%	p = 0,01860293	X <sup>2</sup> = 5,538461538
VEM CÁ	17/26	65.38%	p = 0,116664465	X <sup>2</sup> = 2,461538462

Fonte: Adaptado de Teixeira (2015, p. 269).

Como se vê, 8 entre as 11 condições do experimento registraram dados comportamentais convergentes com a previsão da pesquisa. Isso significa que tais escolhas foram provavelmente motivadas pela atuação da variável selecionada, e não por algum fator aleatório não controlado. Considera-se, portanto, que os participantes perceberam a mudança linguística ocorrida nas combinações “verbo + locativo” e compreenderam-nas intuitivamente como marcadores discursivos, fato que teria induzido tacitamente a escolha da paráfrase alvo, formulada como a melhor redescrição da frase alvo dentre as quatro opções apresentadas.

### **Experimento 2: julgamento de aceitabilidade**

O objetivo particular desse experimento de aceitabilidade em escala foi o de identificar o grau de aceitação das combinações verbo e locativo como marcadores. Foi utilizada a escala mais comum em pesquisas desse

tipo designada “escala Likert” que, segundo Derwing e Almeida (2005, p. 412), “envolve cinco alternativas, com um rótulo atribuído a cada uma delas”. Considerou-se que o modelo seria adequado para o teste da hipótese da pesquisa, já que provê um “número suficientemente grande de alternativas de modo a permitir que mesmo que diferenças nos escores apareçam, não sejam em número tão grande que possam vir a sobrecarregar a memória e o poder de discriminação do participante” (Derwing; Almeida, 2005, p. 412). A tarefa experimental consistia na apresentação ao participante de uma sentença alvo, a qual continha uma expressão da variável independente única selecionada para a pesquisa, isto é, continha um dos marcadores discursivos já descritos no experimento 1. Após esse estímulo, apresentava-se ao participante um comando que descrevia a função comunicativa dessa frase alvo. Por fim, o participante deveria emitir um julgamento, entre as cinco opções disponíveis, acerca da descrição da função comunicativa atribuída à sentença alvo. As alternativas de julgamento foram “sem dúvida alguma”, “provavelmente sim”, “não posso decidir”, “provavelmente não” e “de jeito nenhum”.

O experimento foi desenhado com duas previsões de resposta derivadas da hipótese da pesquisa. Na primeira, o participante deveria escolher a alternativa “sem dúvida alguma”, uma vez que, nesse caso, o comando solicitava que se escolhesse a expressão que poderia ser entendida como uma maneira de o falante chamar a atenção do interlocutor para algo que se estabeleceria na sequência da frase, o que indicava a percepção do composto “verbo + locativo” como uma unidade de sentido. Na segunda, a opção a ser escolhida seria, “de jeito nenhum”, tendo em vista que, na sentença respectiva, a combinação atuaria como marcador; porém, a pergunta contida no comando questionaria a atuação do predicado verbal, desfazendo a sua percepção como marcador.

### *Variáveis e condições*

Conforme descrito acerca do primeiro experimento, as condições experimentais coincidem com os níveis da variável independente única.

O experimento seguiu o mesmo padrão do anterior, ou seja, cada um dos 11 trechos contém uma sentença estímulo teste composta por um dos 11 marcadores.

### ***Materiais***

Com relação ao material, mais uma vez foram utilizados trechos de textos do *corpus* usado por Teixeira (2015), com algumas adaptações necessárias, tais como: suprimir trecho de diálogo, adequar o formato do texto, colocar os diálogos em linhas separadas. Nenhuma dessas adaptações prejudicou a ideia do trecho original, além de conter informação suficiente para que o participante pudesse compreender o contexto específico da atuação do marcador.

O quadro 5 assinala a distribuição da sentença estímulo teste, da alternativa prevista e as quatro alternativas fora da previsão.

**Quadro 5** – Descrição das sentenças do experimento 2.

Sentença teste			Previsão	Fora da previsão
Descrição	Função Marcador	Comando		
Está aí, talvez, em esses 25 anos de experiência com a velocidade, a explicação para o que se veria depois	Constata- ção	Função da expressão: <i>chamar a atenção</i>  Questionamento do comando acerca da função de <i>chamar a atenção</i>	(1) provavelmente sim	(5) de jeito nenhum (4) sem dúvida alguma; (3) não posso decidir; (2) provavelmente não;
- Escuta aqui. (Olhando em torno) Chega! Já vem você com essa enrolação para tirar-me dinheiro.	Repreensão Intimidação	Função da expressão: <i>chamar a atenção</i>  Questionamento do comando acerca da função de <i>chamar a atenção</i>	(4) sem dúvida alguma	(5) de jeito nenhum; (3) provavelmente não; (2) não posso decidir; (1) provavelmente sim

<p>- Enfim - concluiu a velha condescendendo - vá lá.. Eu tomo, mas no meu estado, só um milagre.</p>	<p>Concessão</p>	<p>Função da expressão: <i>chamar a atenção</i></p> <p>Questionamento do comando acerca da atuação como <i>um pedido do falante para que o interlocutor se desloque em direção a um lugar distante de ambos</i></p>	<p>(5) de jeito nenhum</p>	<p>(4) não posso decidir (3) sem dúvida alguma (2) provavelmente sim (1) provavelmente não</p>
<p>O PASQUIM - Use o magicismo. RAUL - Perai. Eu vou falar uma coisa aqui. Eu vou falar sobre os cabeludos.</p>	<p>Interrupção Reformulação</p>	<p>Função da expressão: <i>chamar a atenção</i></p> <p>Questionamento do comando acerca da função de <i>chamar a atenção</i></p>	<p>(3) sem dúvida alguma</p>	<p>(5) provavelmente sim (4) provavelmente não (2) não posso decidir (1) de jeito nenhum</p>
<p>ANGÉLICA - Espera lá! Ou falas tu ou eu falo!</p>	<p>Temporização Objecção</p>	<p>Função da expressão: <i>chamar a atenção</i></p> <p>Questionamento do comando acerca da função de <i>chamar a atenção</i></p>	<p>(3) sem dúvida alguma</p>	<p>(5) provavelmente não (4) não posso decidir (2) de jeito nenhum (1) provavelmente sim</p>
<p>Gabriela - Vamos lá! A intenção era boa.. Só debes olhar para a intenção..</p>	<p>Exortação</p>	<p>Função da expressão: <i>chamar a atenção</i></p> <p>Questionamento do comando acerca da atuação como <i>um pedido do falante para que eles se desloquem em direção a um lugar distante de ambos</i></p>	<p>(3) de jeito nenhum</p>	<p>(5) não posso decidir (4) sem dúvida alguma (2) provavelmente sim (1) provavelmente não</p>

- Olha aí. Quarenta e oito horas depois do chafurdo, Osvaldão já é herói na Albânia - disse Gil. - E o Clementino não consegue descobrir nada com essa parafernália eletrônica dele.	Repreensão Provocação	Função da expressão: <i>chamar a atenção</i>  Questionamento do comando acerca da função de <i>chamar a atenção</i>	(5) sem dúvida alguma	(4) não posso decidir (3) provavelmente não (2) provavelmente sim (1) de jeito nenhum
- Olha aqui, XXXXXX, responda logo.	Repreensão Asseveração	Função da expressão: <i>chamar a atenção</i>  Questionamento do comando acerca da função de <i>chamar a atenção</i>	(5) sem dúvida alguma	(4) não posso decidir (3) provavelmente sim (2) de jeito nenhum (1) provavelmente não
Eis aí te fez Cristo digno de levares a sua cruz e segui-lo; vê lá não faças pé atrás e acudas mais pela tua carne que pela tua alma e te furte o diabo o reino da glória	Repreensão Prevenção	Função da expressão: <i>chamar a atenção</i>  Questionamento sobre a atuação como <i>um conselho do falante para que o interlocutor veja algo em algum lugar distante de ambos</i>	(3) de jeito nenhum	(5) não posso decidir (4) sem dúvida alguma (2) provavelmente sim (1) provavelmente não
O pai dizia-lhe: - Olha lá, minha jóia! não vá isso fazer-te mal..	Repreensão Advertência	Função da expressão: <i>chamar a atenção</i>  Questionamento do comando acerca da função de <i>chamar a atenção</i>	(1) de jeito nenhum	(5) provavelmente sim (4) provavelmente não (3) sem dúvida alguma (2) não posso decidir
Vem cá, você está achando que está certa?	Indagação	Função da expressão: <i>chamar a atenção</i>  Questionamento do comando acerca da função de <i>chamar a atenção</i>	(5) sem dúvida alguma	(4) provavelmente não (3) de jeito nenhum (2) não posso decidir (1) provavelmente sim

Fonte: Adaptado de Teixeira (2015, p. 273).

## ***Procedimentos***

Os procedimentos de elaboração e aplicação desse experimento foram semelhantes aos dos experimentos 1, diferindo apenas em dois pontos: i) o tempo médio de sua realização foi de 35 minutos e ii) o comando consistia em selecionar uma das alternativas que melhor respondesse ao questionamento (sobre a função comunicativa da frase alvo). Foi informado aos participantes que, apesar de uma única resposta dever ser selecionada, não havia uma resposta certa e, sim, uma resposta que eles considerassem livremente mais adequada. A intenção foi incentivá-los a escolher o mais intuitivamente possível.

## ***Resultados***

Foram selecionados cinco graus de aceitabilidade em relação ao estímulo apresentado. Assim, o valor mínimo da escala, “de jeito nenhum”, corresponderia à rejeição total, e o valor máximo, “sem dúvida alguma”, à aceitação total. O valor intermediário, “não posso decidir”, demonstraria uma posição neutra em relação ao comando. Já os outros valores situados entre as extremidades e o ponto médio seriam julgamentos de rejeição parcial (“provavelmente não”) ou aceitação parcial (“provavelmente sim”). Nesse experimento, optou-se por alternar a questão contida no comando, ora perguntando acerca do sentido vinculado à função de “chamar a atenção do leitor”, ora acerca da de “deslocamento espacial” ou “percepção visual no espaço”, conforme o quadro 5. O percentual exposto na tabela 2 representa a escolha da sentença alvo, ou seja, a alternativa que continha a resposta considerada esperada pela previsão da pesquisa. Dessa forma, nesse experimento, para *escuta aqui*, por exemplo, 15 dos 23 participantes escolheram a alternativa que representa a sentença esperada (alvo da hipótese da pesquisa). Nesse sentido, 65.22% é o percentual de respostas que ocorreram de acordo com as previsões do experimento, algo que não se mostrou estatisticamente significativo.

**Tabela 2** – Resultado da sentença alvo no experimento 2.

MARCADOR	SENTENÇA ACEITÁVEL (Previsão)		Qui-quadrado de Pearson	
	Relação escolha da previsão/total participantes	Percentual de resposta da previsão de escolha	p-valor	X <sup>2</sup>
ESCUA AQUI	15/23	65,22%	p = 0,144399792	X <sup>2</sup> = 2,130434783
ESPERA AÍ	09/23	39.13%	p = 0,29714653	X <sup>2</sup> = 1,086956522
ESPERA LÁ	08/23	34.78%	p= 0,144399792	X <sup>2</sup> = 2,130434783
ESTÁ AÍ	07/23	30.43%	p = 0,06056886	X <sup>2</sup> = 3,52173913
OLHA AQUI	14/23	60.87%	p = 0,29714653	X <sup>2</sup> = 1,086956522
OLHA AÍ	14/23	60.87%	p = 0,29714653	X <sup>2</sup> = 1,086956522
OLHA LÁ	18/23	78.26%	p = 0,00671439	X <sup>2</sup> = 7,347826087
VÁ LÁ	15/23	65.22%	p = 0,144399792	X <sup>2</sup> = 2,130434783
VAMOS LÁ	21/23	91.30%	p = 0,00000743	X <sup>2</sup> = 15,69565217
VÊ LÁ	17/23	73.91%	p = 0,021810119	X <sup>2</sup> = 5,260869565
VEM CÁ	14/23	60.87%	p = 0,29714653	X <sup>2</sup> = 1,086956522

Fonte: Adaptado de Teixeira (2015, p. 275).

Tais resultados indicam que a variável da pesquisa provavelmente não atuou sobre as respostas e, nesse sentido, não se obteve argumento em favor da hipótese da pesquisa.

### **Experimento 3: seleção de segmentação**

O objetivo particular deste experimento consistiu em verificar se o participante identificaria um marcador discursivo como unidade de sentido e de forma, a partir da leitura indivisível de suas subpartes. Dessa maneira, teve a intenção de verificar se o entrincheiramento<sup>3</sup> dos elementos do mar-

<sup>3</sup> Conforme Traugott e Trousdale (2013), entrincheiramento é uma propriedade relacionada ao grau de analisabilidade das partes componentes de uma construção. Quanto maior for a possibilidade de analisar as partes de uma construção complexa, menor é o nível de entrincheiramento dessas partes. Quando não houver possibilidade de analisar as partes de uma construção complexa, diz-se que a construção está entrincheirada. Por sua vez, ainda conforme ambos os autores, analisabilidade é a propriedade de uma construção complexa ser percebida pelos falantes como constituída de partes analisáveis.

gador foi percebido e aceito como tal pelos leitores. A tarefa do participante foi a de escolher uma entre cinco alternativas de separação entre palavras cuja segmentação indicasse ou não unidade de sentido. A sentença teste, portanto, foi reproduzida nas cinco opções de resposta, cada uma com uma segmentação diferente. A sentença com a segmentação alvo, isto é, aquela de acordo com as hipóteses da pesquisa, continha verbo + locativo (o marcador discursivo) juntos entre as barras que indicavam a segmentação, retratando assim o acesso integral do marcador. Dessa maneira, além de verificar se a combinação foi lida como uma unidade, o experimento testou se a unidade de sentido foi compreendida como a marca explícita que introduz a opinião do autor. Nesse sentido, se o participante escolhesse a alternativa em que cada um dos 11 marcadores figurasse juntos entre as barras de segmentação, o reconhecimento das expressões indicadoras dessa marca explícita teria sido identificado. Caso contrário, ou seja, se os participantes optassem por segmentações que separassem o verbo do locativo, então a previsão derivada da hipótese de pesquisa seria desconfirmada.

### ***Variáveis e condições***

Como as condições do experimento são compostas pelas 11 ocorrências da variável independente única, foram então dispostos 11 trechos de texto distintos, cada qual contendo uma sentença estímulo com diferentes tipos de segmentação. A sentença pré-determinada como alvo, isto é, com a segmentação prevista pela hipótese de pesquisa, e as quatro sentenças controle, com segmentações que separam verbo e locativo, compõem as cinco alternativas de resposta, dispostas randomicamente em cada *trial*.

### ***Materiais***

Com relação ao material, da mesma forma que nos experimentos anteriores, foram selecionados trechos de textos curtos do *corpus* utilizado por Teixeira (2015), com informação suficiente para que o participante pudesse compreender o contexto específico da atuação do marcador. A sentença teste destacada, na tabela 8, está inserida no trecho aludido. Dessa



maneira, intencionou-se verificar a função marcadora a partir da compreensão e interpretação como unidade de sentido de cada uma das ocorrências da variável. A distribuição da sentença estímulo, da segmentação alvo e das três segmentações de controle estão descritas na tabela seguinte.

**Tabela 3** – Descrição das sentenças do experimento 3.

Sentença teste	Segmentação alvo	Segmentação de controle
Mas pera aí o que você está pensando?	(3) Mas / pera / aí / o / que / você / está pensando?	(1) Mas / pera / aí / o / que / você / está / pensando? (2) Mas pera / aí o que você / está pensando? (4) Mas /pera / aí o que / você está pensando? (5) Mas / pera / aí o /que você está pensando?
Escuta aqui, Gê, o que você pretende com meu namorado?	1) Escuta aqui / Gê/ o / que / você / pretende / com / meu / namorado?	(2) Escuta / aqui Gê / o / que você / pretende / com / meu / namorado? (3) Escuta aqui Gê / o / que / você / pretende com / meu / namorado? (4) Escuta / aqui Gê / o que você / pretende / com / meu / namorado? (5) Escuta / aqui / Gê / o / que você / pretende / com / meu namorado?
Ah! espera lá! Já dou um jeito nisso.	(4) Ah / espera lá / Já / dou / um / jeito / nisso.	(1) Ah espera / lá! / Já / dou / um jeito nisso. (2) Ah /espera / lá / Já / dou / um jeito / nisso. (3) Ah espera / lá / Já / dou um jeito nisso. (5) Ah espera / lá / Já dou um jeito nisso.
Está aí! Era isso o que eu receava!	(1) Está aí / Era /isso / o / que / eu / receava!	(2) Está / aí Era /isso / o / que eu /receava! (3) Está / aí Era /isso / o que / eu /receava! (4) Está / aí Era isso o que eu /receava! (5) Está aí Era isso / o / que / eu receava!
Olha aí, quem diria, legal...	(4) Olha aí / quem / diria / legal...	(1) Olha / aí / quem / diria / legal... (2) Olha aí / quem diria / legal... (3) Olha / aí / quem / diria legal... (5) Olha / aí / quem diria legal...
Estas mesmas – vá lá – mulheres do PT ...	(5) Estas / mesmas / – vá lá – / mulhe- res / do / PT	(1) Estas /mesmas – vá / lá – mulheres / do / PT (2) Estas mesmas – vá / lá – mulheres / do PT (3) Estas /mesmas – vá lá – mulheres do PT (4) Estas mesmas – vá / lá – mulheres do / PT

Olha aqui, Paulo Cesar, quero te falar uma coisa importante	(5) - Olha aqui /Paulo Cesar /quero / te /falar / uma / coisa / importante	1) - Olha aqui Paulo / Cesar quero te falar / uma / coisa importante 2) - Olha /aqui Paulo / Cesar quero /te falar / uma / coisa /importante 3) - Olha aqui /Paulo / Cesar quero te falar / uma / coisa importante 4) - Olha /aqui Paulo / Cesar /quero te /falar / uma / coisa /importante
Vamos lá, se alguém engolir sopa de números.	(1) Vamos lá/ se /alguém / engolir /sopa / de / números	(2) Vamos / lá/ se alguém /engolir /sopa de / números (3) Vamos / lá se /alguém engolir /sopa /de números (4) Vamos lá se /alguém engolir /sopa /de números (5) Vamos / lá/ se /alguém /engolir sopa /de / números
Olha lá.. Estamos combinados, filha!	(3) Olha lá / Estamos combinados /filha	(1) Olha / lá / Estamos combinados filha (2) Olha / lá / Estamos /combinados /filha (4) Olha lá / Estamos /combinados filha (5) Olha / lá Estamos combinados / filha
Vem cá, Silvia, então você estava em São Lourenço, né	(5) Vem cá / Silvia /então / você /estava / em /São Lourenço /né?	(1) Vem cá /Silvia então /você estava em São /Lourenço /né? (2) Vem /cá /Silvia então /você estava em São Lourenço /né? (3) Vem / cá /Silvia / então /você estava /em São / Lourenço /né? (4) Vem /cá Silvia /então /você /estava em /São / Lourenço /né?
Vê lá se vou me assustar com atores que não sabem fazer conta	(2) Vê lá / se /vou /me / assustar /com /atores /que / não sabem / fazer / conta.	(1) Vê lá /se /vou me assustar com /atores /que não / sabem /fazer /conta. (3) Vê / lá se vou me /assustar com /atores que /não sabem fazer conta. (4) Vê /lá se vou me /assustar com atores que /não sabem /fazer conta. (5) Vê lá se vou me /assustar /com atores que não / sabem fazer /conta.

Fonte: Adaptado de Teixeira (2015, p. 280).

### ***Procedimentos***

Esse experimento também foi elaborado em formulário impresso intitulado “Questionário”, composto de três folhas, preenchido no mesmo momento de sua apresentação na frente do pesquisador e respondido individualmente pelos participantes. O tempo médio de sua realização foi de 30

minutos. A aplicação foi realizada em sala de aula com todos os participantes presentes, distribuídos estrategicamente a fim de evitar o contato entre eles.

### Resultados

Esse experimento obteve o maior nível de confirmação das previsões da pesquisa, ou seja, a maioria das escolhas feitas pelos participantes atestou a hipótese do trabalho. Esse fato é importante, a julgar pelo tipo de teste – segmentar as palavras de acordo com a unidade de sentido – e o objetivo da pesquisa – identificar a interpretação do marcador discursivo como um todo de forma e sentido.

As ocorrências da variável e a construção marcadora discursiva foram as mesmas, uma vez que se buscou identificar os 11 marcadores discursivos estudados. Porém, os fragmentos de textos foram distintos dos demais testes, retirados do *corpus* da pesquisa. A pergunta contida no comando referiu-se, em primeiro plano, à identificação da unidade de sentido e, em segundo, à testagem da aceitação do marcador resultante do processo de mudança linguística.

O percentual exposto na tabela 3 representa a escolha da sentença alvo, ou seja, a alternativa que continha a resposta considerada conforme a previsão (alvo). Dessa forma, neste experimento, para *espera lá*, por exemplo, 23 dos 24 participantes escolheram a alternativa que representa a sentença alvo. Nesse sentido, 95.83% é o percentual de respostas que ocorreram de acordo com as previsões do experimento.

**Tabela 3** – Resultado da sentença alvo no experimento 3.

MARCADOR	SENTENÇA ALVO (Previsão)		Qui-quadrado de Pearson	
	Relação escolha da previsão/ total partici- pantes	Percentual de resposta da previsão de escolha	p-valor	X <sup>2</sup>
ESCUTA AQUI	19/24	79,17%	p = 0,004266725	X <sup>2</sup> = 8,166666667
ESPERA AÍ	20/24	83.33%	p = 0,001090835	X <sup>2</sup> = 10,66666667

ESPERA LÁ	23/24	95.83%	p = 0,00000709	X <sup>2</sup> = 20,16666667
ESTÁ AÍ	19/24	79.17%	p = 0,004266725	X <sup>2</sup> = 8,166666667
OLHA AQUI	17/24	70.83%	p = 0,041226833	X <sup>2</sup> = 4,166666667
OLHA AÍ	22/24	91.67%	p = 0,00003105	X <sup>2</sup> = 18,33326665
OLHA LÁ	20/24	83.33%	p = 0,001090835	X <sup>2</sup> = 10,66666667
VÁ LÁ	20/24	83.33%	p = 0,001090835	X <sup>2</sup> = 10,66666667
VAMOS LÁ	22/24	91.67%	p = 0,00003105	X <sup>2</sup> = 18,33326665
VÊ LÁ	17/24	70.83%	p = 0,041226833	X <sup>2</sup> = 4,166666667
VEM CÁ	22/24	91.67%	p = 0,00003105	X <sup>2</sup> = 18,33326665

Fonte: Adaptado de Teixeira (2015, p. 282).

Nesse experimento, todas as escolhas favoreceram a previsão da pesquisa, obtendo-se diferenças significativas em termos estatísticos no X<sup>2</sup>.

### **Discussão dos resultados dos três experimentos em conjunto**

A pesquisa identificou que alguns marcadores são menos produtivos como tais do que outros. Esse fato pode ter impactado o resultado dos três experimentos. Dois motivos são relevantes para esse quadro: i) a maior analisabilidade do sentido das subpartes em relação ao sentido da unidade, ou seja, a menor abstração do sentido do verbo e do locativo sugere a incipiência dos usos - nesse caso, a indicação seria a de que o participante não percebeu ou aceitou a mudança linguística de itens como *(es)t(á) aí*, *(es) per(a) aí* e *(es)pera lá*; ii) a obsolescência<sup>4</sup> dos marcadores, ou seja, algumas unidades não são muito utilizadas atualmente - nesse caso, o participante pode ter optado pela alternativa que se aproximou mais de seu conhecimento linguístico ou ter marcado aleatoriamente.

Os experimentos 1 e 2 promoveram escolhas, seja por paráfrase, seja em julgamento escalar, que demandam esforço cognitivo maior, uma vez que, além da compreensão como unidade de sentido, o participante deveria: i) perceber a função atribuída ao marcador-teste, ii) checá-la com o

<sup>4</sup> De acordo com Traugott e Trousdale (2013), obsolescência é a propriedade de uma construção de ter decréscimo em seu uso ou cair em desuso.

comando presente no enunciado e iii) escolher uma das opções que melhor se assemelhasse à função. Outra questão importante é o êxito do experimento 3, que objetivou testar a segmentação dos itens da sentença teste em unidades de sentido. Nesse caso, independente do marcador em questão, a grande maioria dos participantes confirmou a hipótese principal de que as ocorrências testadas são compreendidas holisticamente.

Em todo caso, a pesquisa aqui rapidamente resumida é uma ilustração de como a linguística funcionalista centrada no uso pode lançar mão da experimentação como uma abordagem metodológica útil na condução de pesquisas de teste de previsões comportamentais.

## **Conclusões**

Por tudo o que se disse neste capítulo, é possível compreender que uma pesquisa experimental deve ser conduzida por meio de diferentes etapas interdependentes. Delinear um projeto experimental demanda esforços de diferentes naturezas, desde a seleção de um problema de análise passível de teste de previsão comportamental, até a análise estatística de resultados, passando pela adoção de uma técnica experimental adequada, o estabelecimento de uma tarefa clara e simples, o delineamento de variáveis e de condições experimentais, bem como pelo tratamento adequado dos estímulos, da distribuição dos participantes e da aplicação do experimento.

Quando esse longo processo é conduzido com zelo e os resultados experimentais vão ao encontro das hipóteses do pesquisador, dados robustos podem ser utilizados em favor de uma hipótese de pesquisa em linguística, seja a respeito de um fenômeno específico, seja a respeito de uma hipótese teórica mais abrangente. É por essa razão que a pesquisa experimental em linguística vem crescendo como uma abordagem metodológica extremamente produtiva, no Brasil e no exterior.

## **Referências**

ABRAÇADO, J.; KENEDY, E. **Transitividade traço a traço**. Niterói: EDUFF, 2014.

BUTLER, C. **Statistics in linguistics**. Oxford: Basil Blackwell, 2006.

BYBEE, J. **Language, usage and cognition**. Cambridge: Cambridge University Press, 2010.

CHOMSKY, N. **Aspects of the theory of syntax**. MASS: MIT Press, 1965.

DERWING, B. L.; ALMEIDA, R. G. Métodos experimentais em linguística. In: MAIA, M.; FINGER, I. (Eds.). **Processamento da linguagem**. Pelotas: Educat, 2005. Disponível em: <<http://alcor.concordia.ca/~almeida/Derwing-deAlmeida-BP-7d.pdf>>. Acesso em: 10 jun. 2014.

FEYERABEND, P. **Farewell to Reason**. London: Verso, 1994.

FRANÇA, A.; FERRARI, L.; MAIA, M. (orgs.) *A linguística no século XXI: convergência e divergências no estudo da linguagem*. SP: Contexto, 2016.

JOHNSON, K. **Quantitative methods in linguistics**. Oxford: Blackwell Pub., 2008.

OEHLERT, G. **A first course in design and analysis of experiments**. NY: Freeman, 2010.

KENEDY, E. Psicolinguística na descrição gramatical. IN.: MAIA, M. (Org.). **Psicolinguística, psicolinguísticas: uma introdução**. SP: Contexto, 2015. p. 143-156.

KENEDY, E. Uma breve introdução aos estudos experimentais em linguística. In.: WIEDEMER, M. (Org.) **Estudos linguísticos contemporâneos: questões e tendências**. RJ: Autografia, 2019. p. 159-194.

POPPER, K. **The logic of scientific discovery**. London and New York, 1959.

RORTY, R. **Philosophy and the mirror of nature**. Oxford: Blackwell, 1990.

SAMPAIO, T. A escolha de *software e hardware* na psicolinguística: revisão e opinião. **Revista de Estudos da Linguagem**, v. 25, p.971-1010, 2017.

SKINNER, B. F. **Verbal behavior**. MA: Copley Publishing Group, 1957.

TEIXEIRA, A. **A construção verbal marcadora discursiva VLoc: uma análise funcional centrada no uso**. 297f. Tese (Doutorado em Estudos de Linguagem) – Universidade Federal Fluminense, Instituto de Letras, Niterói, 2015.

TRAUGOTT, E. C.; TROUSDALE, G. **Constructionalization and Constructional changes**. Oxford, 2013.

# INFORMAÇÕES SOBRE OS AUTORES

## ORGANIZADOR

### **Ivo da Costa do Rosário**

É graduado em Letras (Português, Inglês e respectivas literaturas) pela UERJ e graduado em Pedagogia pela UNIRIO. É mestre e doutor em Letras Vernáculas pela UFRJ e é mestre e doutor em Letras pela UFF. Atualmente é professor associado de Língua Portuguesa e docente permanente do Programa de Pós-graduação em Estudos de Linguagem da UFF. É líder do CCO (Grupo de Pesquisa Conectivos e Conexão de Orações - cco.sites.uff.br) e membro do grupo D & G (Grupo de Estudos Discurso e Gramática - deg.uff.br), ambos sediados na UFF. É membro do GT Descrição do Português da ANPOLL. É Jovem Cientista do Nosso Estado, pela FAPERJ. É membro da comissão científica da área de Sintaxe da ABRALIN. É bolsista de produtividade em pesquisa pelo CNPq (PQ-2). É avaliador do SINAES (Sistema Nacional de Avaliação da Educação Superior). Atua principalmente nas seguintes áreas: funcionalismo, construcionalização, mudanças construcionais, morfossintaxe, conexão de orações e conectivos.

## AUTORES

### **Ana Claudia Machado dos Santos**

Doutora em Estudos de Linguagem, ênfase em Linguística, pela Universidade Federal Fluminense (UFF - 2015). Mestre em Língua Portuguesa pela UFF (2010) e Especialista em Língua Portuguesa pela Universidade Veiga de Almeida (UVA-2007). Professora Adjunta de Língua Portuguesa na UFF. É pesquisadora, atuando no grupo “Conectivos e Conexão de Oração” (CCO), voltado para estudos na Linguística Funcional Centrada no

Uso, considerando o Domínio Funcional da Conexão e do Grupo “Linguagem em uso, cognição e gramática” que visa à cooperação internacional Brasil-Portugal. Coordenadora do projeto de monitoria “Prática de análise linguística: a sintaxe no contexto e o ensino de Língua Portuguesa” no GLC-UFF. É coordenadora do projeto de extensão PROEX-UFF “Refletindo acerca do domínio da conexão em textos de opinião de cunho político na mídia digital: colóquios sobre ensino de língua materna e cidadania”. Tenho experiência na área de Letras, atuando principalmente nos seguintes temas: Língua portuguesa, Morfossintaxe, Leitura, Produção Textual, Funcionalismo centrado no uso, Gramaticalização e gramaticalização de construções, Mudança construcional e construcionalização.

### **Diego Leite de Oliveira**

É graduado em Letras Português-Russo, mestre e doutor em Linguística pela Universidade Federal do Rio de Janeiro (UFRJ), onde atua no Departamento de Letras Orientais e Eslavas e no Programa de Pós-Graduação em Linguística. Lidera o SLAV (Núcleo de Estudos em Eslavística) e é membro do PEUL (Programa de Estudos sobre o Uso da Língua), ambos sediados na UFRJ. Desenvolve pesquisas em Gramática de Construções Baseada no Uso, atuando principalmente com os seguintes temas: variação e mudança, morfossintaxe, estrutura da informação, estudos comparados entre russo e português, ensino de língua estrangeira e tradução.

### **Eduardo Kenedy**

Possui doutorado e mestrado em Linguística pela Universidade Federal do Rio de Janeiro (UFRJ). Durante o seu doutoramento, realizou estágio na Universidade de Lisboa. É licenciado em Letras pela Universidade Federal Fluminense (UFF). Foi professor da Universidade do Estado do Rio de Janeiro (UERJ, 2000 a 2008) e da Universidade Federal da Paraíba (UFPB, 2008 a 2009). Também atuou como docente na Educação Básica, tanto no Ensino Fundamental, nas Secretarias Municipais de Educação de Duque de Caxias (1999 a 2004) e de Niterói (2003 a 2004),



quanto no Ensino Médio, na Secretaria Estadual de Educação do Rio de Janeiro (1999 a 2004). Desde 2009, é professor de Linguística da UFF, atuando na Graduação em Letras presencial e à distância, esta última vinculada à Universidade Aberta do Brasil (UAB). É membro do Programa de Pós-Graduação Stricto-Sensu em Estudos da Linguagem da UFF. Orienta pesquisas de iniciação científica, de mestrado e de doutorado, bem como supervisiona estágios de pós-doutorado. É líder do Grupo de Estudos e Pesquisas em Linguística Teórica e Experimental (GEPEX), por ele fundado em 2009. Atua como pesquisador nas áreas da Psicolinguística Experimental e da Sintaxe Gerativa das Línguas Naturais. É bolsista de produtividade em pesquisa do CNPq e membro da Rede Nacional de Ciência para a Educação. Seus interesses de pesquisas atualmente são: (i) Psicolinguística translacional para a Educação; (ii) Processamento cognitivo da linguagem e (iii) Arquitetura do conhecimento linguístico em línguas naturais.

### **Edvaldo Balduino Bispo**

Doutor em Estudos da Linguagem pela UFRN. Professor Associado do Departamento de Letras dessa instituição e docente permanente do Programa de Pós-Graduação em Estudos da Linguagem, do qual foi Coordenador entre 2015 e 2019. Coordenou o GT Descrição do Português da ANPOLL (2016-2018). Pesquisador do grupo *Discurso & Gramática*, seus temas de interesse voltam-se à morfossintaxe do português em perspectiva funcional e/ou construcionista e ao ensino de língua portuguesa. É coorganizador de livros e autor de vários capítulos e artigos científicos relacionados a esses temas. É bolsista de produtividade em pesquisa do CNPq (PQ-2).

### **Flávia Saboya da Luz Rosa**

Doutora e Mestre em Estudos de Linguagem, Licenciada em Português-Francês e Bacharel em Língua e Literatura Italiana pela Universidade Federal Fluminense. Tem experiência profissional no ensino de Italiano,

Francês, Espanhol, Inglês e Português, este último como bolsista da CAPES na graduação de Letras da UFF, nas áreas de morfologia e sintaxe. Pesquisadora do Grupo de Estudos Discurso & Gramática - D&G UFF. Desenvolve investigações na área da Linguística Cognitivo-Funcional com foco nos seguintes temas: construcionalização, mudança construcional, propriedades da forma construcional (sintática, morfológica e fonológica) e do conteúdo construcional (semântico, pragmático e discursivo), paradigmaticização de construções.

### **Ivo da Costa do Rosário**

É graduado em Letras (Português, Inglês e respectivas literaturas) pela UERJ e graduado em Pedagogia pela UNIRIO. É mestre e doutor em Letras Vernáculas pela UFRJ e é mestre e doutor em Letras pela UFF. Atualmente é professor associado de Língua Portuguesa e docente permanente do Programa de Pós-graduação em Estudos de Linguagem da UFF. É líder do CCO (Grupo de Pesquisa Conectivos e Conexão de Orações - cco.sites.uff.br) e membro do grupo D & G (Grupo de Estudos Discurso e Gramática - deg.uff.br), ambos sediados na UFF. É membro do GT Descrição do Português da ANPOLL. É Jovem Cientista do Nosso Estado, pela FAPERJ. É membro da comissão científica da área de Sintaxe da ABRALIN. É bolsista de produtividade em pesquisa pelo CNPq (PQ-2). É avaliador do SINAES (Sistema Nacional de Avaliação da Educação Superior). Atua principalmente nas seguintes áreas: funcionalismo, construcionalização, mudanças construcionais, morfossintaxe, conexão de orações e conectivos.

### **Karen Sampaio Braga Alonso**

Professora Adjunta 40h DE do Departamento de Linguística e Filologia da UFRJ. Substituta Eventual do Diretor de Cultura e Extensão da Faculdade de Letras da UFRJ (2018). Coordenadora do Grupo Discurso & Gramática - UFRJ. Bolsista de Produtividade em Pesquisa do CNPq (PQ-2). Possui graduação em Português/Literaturas pela Universidade

Federal do Rio de Janeiro (2002), Mestrado em Linguística pela Universidade Federal do Rio de Janeiro (2005), Doutorado em Linguística pela UFRJ (2010) e Pós-doutorado (*visiting scholar*) na Universidade da Califórnia (Berkeley). Docente do Programa de Pós-graduação em Linguística da UFRJ. Docente do Mestrado Profissional em Letras da UFRJ. Tem experiência na área de Linguística, com foco em gramática de construções; mudança linguística; Linguística Baseada no Uso.

### **Lauriê Ferreira Martins Dall’Orto**

Doutora em Linguística pela Universidade Federal de Juiz de Fora (2018), mestre em Linguística pela Universidade Federal de Juiz de Fora (2013), especialista em Ensino de Língua Portuguesa pela Universidade Federal de Juiz de Fora (2012) e graduada em Letras-Língua Portuguesa pela Universidade Federal de Juiz de Fora (2010) e em Letras-Língua Italiana também pela Universidade Federal de Juiz de Fora (2015). Atualmente, realiza um pós-doutorado em Linguística na Universidade Federal de Juiz de Fora (PNPD CAPES/UFJF). Pesquisadora do Grupo de Estudos Discurso & Gramática e vice-coordenadora do NUPACT/UFJF - Núcleo de Pesquisa em Abordagem Construcional e Tradução. Desenvolve pesquisas a partir da Linguística Funcional Centrada no Uso, atuando, mais especificamente, sob a perspectiva da construcionalização gramatical.

### **Marcia dos Santos Machado Vieira**

Doutora (2001) e Mestre (1995) em Língua Portuguesa e Bacharel e Licenciada em Português-Inglês (1992) pela Universidade Federal do Rio de Janeiro (UFRJ). Professora Associada IV de Língua Portuguesa do Departamento de Letras Vernáculas e do Programa de Pós-Graduação em Letras Vernáculas, coordenadora dos Projeto Predicar (<https://projeto-predicar.wixsite.com/predicar>), desenvolvido na UFRJ, e VariaR (<https://variara.wixsite.com/variara>), projeto franco-brasileiro desenvolvido na Université Paul-Valéry/Montpellier e Universidade Federal do Rio de Janeiro. Foi professora da Secretaria de Educação do Município do Rio de

Janeiro. Membro (pesquisador) do grupo de pesquisa *Discurso & Gramática* (UFF). Integra, na gestão de 2014-2020, a coordenação do Eixo 1 do GT de Sociolinguística da ANPOLL. Coordena esse GT no período de 2018-2021. Coordena a Comissão de Sociolinguística da ABRALIN. É membro do Comitê de Ética em Pesquisa do IESC/UFRJ e de NDE de cursos da Faculdade de Letras/UFRJ. É editora-chefe da Revista Diadorim: estudos linguísticos e literários do PPGLEV/UFRJ. É bolsista de produtividade em pesquisa do CNPq (PQ-2).

### **Marcos Luiz Wiedemer**

Professor adjunto (Linguística) da Faculdade de Formação de Professores da UERJ, atuando no curso de Letras (Português/Inglês) e no Programa de Pós-Graduação em Letras e Linguística. Procientista (UERJ/Faperj). Doutor em Estudos Linguísticos (UNESP), com bolsa sanduíche na Erfurt Universität, Alemanha. Mestre em Linguística pela UFSC e licenciado em Letras (Português/Inglês) pela FURB. Foi Coordenador Geral do Programa de Pós-Graduação em Letras e Linguística (PPLIN), gestão 2016-2018/2018-2020. Membro (pesquisador) dos grupos de pesquisa *Discurso & Gramática* (UFF) e *Estudos Sociofuncionalistas* (UFMS). Líder do grupo de pesquisa *Interfaces linguísticas* (UERJ). Membro do Conselho da ANPOLL (Estudos Linguísticos) e Coordena a Comissão – Linguística e Cognição da Abralín. Desde 2018 coordena, com Márcia dos Santos Machado Vieira, o GT de *Sociolinguística* da ANPOLL. Membro do Corpo de Embaixadores da Olimpíada de Linguística. Coordenador do Laboratório de Formação Permanente em Letras: ações coletivas, docência e ensino (LABLETRAS/UERJ).

### **Maria Angélica Furtado da Cunha**

Professora titular de Linguística da Universidade Federal do Rio Grande do Norte. Doutorado em Linguística pela Universidade Federal do Rio de Janeiro, mestrado em Linguística pela Universidade de Brasília. Dois estágios de pós-doutoramento na University of California, Santa

Barbara, e um na Universidade Federal Fluminense. Coordenadora do Grupo de Estudos Discurso & Gramática na UFRN. Coorganizadora de livros e autora de vários capítulos e artigos sobre transitividade, estrutura argumental, gramática de construções e ensino de gramática sob o enfoque da Linguística Funcional. Pesquisadora do CNPq.

### **Mariangela Rios de Oliveira**

Professora titular de Língua Portuguesa da UFF, com pós-doutorado na Universidade Aberta - Lisboa. Docente permanente do Programa de Pós-Graduação em Estudos de Linguagem da UFF. Professora convidada da UERJ/FFP, com atuação no Programa de Pós-Graduação em Letras e Linguística. Pesquisadora 1C do CNPq, Cientista do Nosso Estado pela Faperj e investigadora colaboradora da FCT - Portugal. Líder do Grupo de Estudos *Discurso & Gramática*. Ex-presidente da Abralin e atual conselheira desta entidade. Sua produção intelectual, na área da Linguística Funcional Centrada no Uso, contempla estudos sobre a morfossintaxe do português (variabilidade e mudança), na perspectiva da gramática de construções.

### **Monclar Guimarães Lopes**

É graduado em Letras (Português, Inglês e respectivas literaturas) e Especialista em Língua e Literatura Brasileira pela Ferlagos. É Mestre em Língua Portuguesa e Doutor em Estudos da Linguagem pela UFF. Atualmente, é professor adjunto de Língua Portuguesa do Departamento de Letras Clássicas e Vernáculas e professor permanente do Programa de Pós-Graduação em Estudos da Linguagem da UFF. É vice-líder do grupo D&G (Grupo de Estudos Discurso e Gramática), pesquisador do CCO (Grupo de Pesquisa Conectivos e Conexão de Orações) e membro do GT Descrição do Português da ANPOLL.

## **Pâmela Fagundes Travassos**

Doutoranda do curso de Língua Portuguesa do Programa de Letras Vernáculas da Universidade Federal do Rio de Janeiro (UFRJ). Integra o projeto PREDICAR (Formação e expressão de predicados e predicções: estabilidade, variação e mudança construcional), coordenado pela Dr<sup>a</sup>. Marcia dos Santos Machado Vieira (UFRJ). De março a agosto de 2021, está em estágio de Doutorado na Université de Lille, no laboratório “*Savoirs, Textes, Langage*”, sob a supervisão do Dr. Bert Cappelle. Integra, como estudante colaboradora, o Projeto VariaR - Variação em Línguas Românicas. É membro do GT de Sociolinguística da ANPOLL (Associação Nacional de Pós-Graduação e Pesquisa em Letras e Linguística). Concentra-se em estudos de Morfossintaxe à luz das perspectivas Sociofuncionalista, Funcional-Cognitiva e da abordagem da Gramática de Construções. cursou Mestrado em Língua Portuguesa (Letras Vernáculas) na UFRJ. É especialista em Língua Portuguesa pela Universidade Federal Fluminense (UFF). Possui licenciatura em Letras (UFRJ), habilitação: Português-Literaturas.

## **Patrícia Fabiane Amaral da Cunha Lacerda**

Pós-doutora em Linguística pela Universidade Federal de Minas Gerais (2009), doutora em Linguística pela Universidade Federal de Minas Gerais (2007), mestre em Linguística pela Universidade Federal de Juiz de Fora (2003) e graduada em Letras pela Universidade Federal de Juiz de Fora (2002). Professora Associada da Faculdade de Letras da Universidade Federal de Juiz de Fora, atuando no Departamento de Letras Estrangeiras Modernas e no Programa de Pós-graduação em Linguística. Pesquisadora do Grupo de Estudos Discurso & Gramática e coordenadora do NUPACT/UFJF - Núcleo de Pesquisa em Abordagem Construcional e Tradução. Tem experiência na área de Linguística, desenvolvendo pesquisas e orientando trabalhos nos seguintes temas: Linguística Funcional Centrada no Uso, abordagem construcional da mudança, construcionalização gramatical e construcionalização lexical.

*Metodologia da Pesquisa Funcionalista* é uma obra destinada a suprir uma grande lacuna no cenário acadêmico nacional. De fato, há carência de um manual destinado a divulgar organicamente, de modo claro e didático, os métodos disponíveis à condução de investigações baseadas na língua em uso, especialmente na vertente da LFCU (Linguística Funcional Centrada no Uso).

Ao longo de nove capítulos, o leitor é conduzido a diferentes perspectivas analíticas. Partindo da caracterização teórico-metodológica da LFCU, demonstra-se o passo a passo da metodologia sincrônica e diacrônica, bem como da investigação na moderna linha da construcionalidade. Por fim, apresentam-se métodos originados da interface da Linguística Funcional com a Computação, a Linguística de Corpus, a Prosódia e a Linguística Experimental.