

REVISTA DE ESTUDOS DA LINGUAGEM

Faculdade de Letras da UFMG

ISSN

Impresso: 0104-0588

On-line: 2237-2083

V.30 - N° 4



REVISTA DE ESTUDOS DA LINGUAGEM

Universidade Federal de Minas Gerais

REITORA: Sandra Regina Goulart Almeida

VICE-REITOR: Alessandro Fernandes Moreira

Faculdade de Letras

DIRETORA: Sueli Maria Coelho

VICE-DIRETOR: Georg Otte

Editores-chefes

Gustavo Ximenes Cunha (UFMG)

Janayna Maria da Rocha Carvalho (UFMG)

Revisão e Normalização

Gustavo Ximenes Cunha (UFMG)

Editoras-associadas:

Ana Larissa Adorno Macioto Oliveira (UFMG),

Carla Viana Coscarelli (UFMG),

Helcira Maria Rodrigues de Lima (UFMG).

Organizadores convidados

Aroldo Leal de Andrade (UFMG),

Ana Paula Quadros Gomes (UFRJ),

Carlos Felipe Pinto (UFBA).

Revisão de Língua Inglesa

Ana Larissa Adorno Macioto Oliveira (UFMG)

Junia de Carvalho Fidelis Braga (UFMG)

Mara Passos Guimarães (UFMG)

Marisa Mendonça Carneiro (UFMG)

Diagramação

Naila Catherine França Eleutério

REVISTA DE ESTUDOS DA LINGUAGEM, v.1 - 1992 - Belo Horizonte, MG,
Faculdade de Letras da UFMG

Histórico:

1992 ano 1, n.1 (jul/dez)

1993 ano 2, n.2 (jan/jun)

1994 Publicação interrompida

1995 ano 4, n.3 (jan/jun); ano 4, n.3, v.2 (jul/dez)

1996 ano 5, n.4, v.1 (jan/jun); ano 5, n.4, v.2; ano 5, n. esp.

1997 ano 6, n.5, v.1 (jan/jun)

Nova Numeração:

1997 v.6, n.2 (jul/dez)

1998 v.7, n.1 (jan/jun)

1998 v.7, n.2 (jul/dez)

1. Linguagem - Periódicos I. Faculdade de Letras da UFMG, Ed.

CDD: 401.05

ISSN: Impresso: 0104-0588

On-line: 2237-2083

REVISTA DE ESTUDOS DA LINGUAGEM

V. 30 - N. 4 - out.-dez. 2022

Indexadores

Diadorim [Brazil]

DOAJ (Directory of Open Access Journals) [Sweden]

DRJI (Directory of Research Journals Indexing) [India]

EBSCO [USA]

EuroPub [England]

JournalSeek [USA]

Latindex [Mexico]

Linguistics & Language Behavior Abstracts [USA]

MIAR (Matriu d'Informació per a l'Anàlisi de Revistes) [Spain]

MLA Bibliography [USA]

OAJI (Open Academic Journals Index) [Russian Federation]

Portal CAPES [Brazil]

REDIB (Red Iberoamericana de Innovación y Conocimiento Científico) [Spain]

SCOPUS [Amsterdam]

Sindex (Scientific Indexing Services) [USA]

Web of Science [USA]

WorldCat / OCLC (Online Computer Library Center) [USA]

ZDB (Elektronische Zeitschriftenbibliothek) [Germany]



REVISTA DE ESTUDOS DA LINGUAGEM

Editores-chefes

Gustavo Ximenes Cunha (UFMG, Belo Horizonte/MG, Brasil)

Janayna Maria da Rocha Carvalho (UFMG, Belo Horizonte/MG, Brasil)

Editoras-associadas

Ana Larissa Adorno Maciotto Oliveira (UFMG)

Carla Viana Coscarelli (UFMG)

Helcira Maria Rodrigues de Lima (UFMG)

Conselho Editorial

Alejandra Vitale (UBA, Ciudad Autónoma de Buenos Aires, Argentina)

Didier Demolin (Université de la Sorbonne Nouvelle Paris 3, Paris, França)

Ieda Maria Alves (USP, São Paulo/SP, Brasil)

Jairo Nunes (USP, São Paulo/SP, Brasil)

Scott Schwenter (OSU, Columbus, Ohio, Estados Unidos)

Shlomo Izre'el (TAU, Tel Aviv, Israel)

Stefan Gries (UCSB, Santa Barbara/CA, Estados Unidos)

Teresa Lino (NOVA, Lisboa, Portugal)

Tjerk Hagemeyer (ULisboa, Lisboa, Portugal)

Comissão Científica

Aderlande Pereira Ferraz (UFMG, Belo Horizonte/MG, Brasil)
Alessandro Panunzi (Unifi, Florença, Itália)
Alina M. S. M. Villalva (ULisboa, Lisboa, Portugal)
Aline Alves Ferreira (UCSB, Santa Barbara/CA, Estados Unidos)
Ana Lúcia de Paula Müller (USP, São Paulo/SP, Brasil)
Ana Maria Carvalho (UA, Tucson/AZ, Estados Unidos)
Ana Paula Scher (USP, São Paulo/SP, Brasil)
Anabela Rato (U of T, Toronto/ON, Canadá)
Aparecida de Araújo Oliveira (UFV, Viçosa/MG, Brasil)
Aquiles Tescari Neto (UNICAMP, Campinas/SP, Brasil)
Augusto Soares da Silva (UCP, Braga, Portugal)
Beth Brait (PUC-SP/USP, São Paulo/SP, Brasil)
Bruno Neves Rati de Melo Rocha (UFMG, Belo Horizonte/MG, Brasil)
Carmen Lucia Barreto Matzenauer (UCPEL, Pelotas/RS, Brasil)
Celso Ferrarezi (UNIFAL, Alfenas/MG, Brasil)
César Nardelli Cambraia (UFMG, Belo Horizonte/MG, Brasil)
Cristina Name (UFJF, Juiz de Fora/MG, Brasil)
Charlotte C. Galves (UNICAMP, Campinas/SP, Brasil)
Deise Prina Dutra (UFMG, Belo Horizonte/MG, Brasil)
Diana Luz Pessoa de Barros (USP/UPM, São Paulo/SP, Brasil)
Edwiges Morato (UNICAMP, Campinas/SP, Brasil)
Emília Mendes Lopes (UFMG, Belo Horizonte/MG, Brasil)
Esmeralda V. Negrão (USP, São Paulo/SP, Brasil)
Flávia Azeredo Cerqueira (JHU, Baltimore/MD, Estados Unidos)
Gabriel de Avila Othero (UFRGS, Porto Alegre/RS, Brasil)
Gerardo Augusto Lorenzino (TU, Filadélfia/PA, Estados Unidos)
Glaucia Muniz Proença de Lara (UFMG, Belo Horizonte/MG, Brasil)
Hanna Batoréo (UAb, Lisboa, Portugal)
Heliana Ribeiro de Mello (UFMG, Belo Horizonte/MG, Brasil)
Heronides Moura (UFSC, Florianópolis/SC, Brasil)
Hilario Bohn (UCPEL, Pelotas/RS, Brasil)
Hugo Mari (PUC-Minas, Belo Horizonte/MG, Brasil)
Ida Lucia Machado (UFMG, Belo Horizonte/MG, Brasil)
Ieda Maria Alves (USP, São Paulo/SP, Brasil)
Ivã Carlos Lopes (USP, São Paulo/SP, Brasil)
Jairo Nunes (USP, São Paulo/SP, Brasil)

Jairo Venício Carvalhais Oliveira (UFMG, Belo Horizonte/MG, Brasil)
Jean Cristtus Portela (UNESP-Araraquara, Araraquara/SP, Brasil)
João Antônio de Moraes (UFRJ, Rio de Janeiro/ RJ, Brasil)
João Miguel Marques da Costa (Universidade Nova da Lisboa, Lisboa, Portugal)
João Queiroz (UFJF, Juiz de Fora/MG, Brasil)
José Magalhaes (UFU, Uberlândia/MG, Brasil)
João Saramago (Universidade de Lisboa, Lisboa, Portugal)
José Borges Neto (UFPR, Curitiba/PR, Brasil)
Kanavillil Rajagopalan (UNICAMP, Campinas/SP, Brasil)
Laura Alvarez Lopez (Universidade de Estocolmo, Stockholm, Suécia)
Leo Wetzels (Free Univ. of Amsterdam, Amsterdã, Holanda)
Laurent Filliettaz (Université de Genève, Genebra, Suíça)
Leonel Figueiredo de Alencar (UFC, Fortaleza/CE, Brasil)
Livia Oushiro (UNICAMP, Campinas/SP, Brasil)
Lodenir Becker Karnopp (UFRGS, Porto Alegre/RS, Brasil)
Lorenzo Teixeira Vitral (UFMG, Belo Horizonte/MG, Brasil)
Luiz Amaral (UMass Amherst, Amherst/MA, Estados Unidos)
Luiz Carlos Cagliari (UNESP, São Paulo/SP, Brasil)
Luiz Carlos Travaglia (UFU, Uberlândia/MG, Brasil)
Marcelo Barra Ferreira (USP, São Paulo/SP, Brasil)
Marcia Cançado (UFMG, Belo Horizonte/MG, Brasil)
Márcio Leitão (UFPB, João Pessoa/PB, Brasil)
Marcus Maia (UFRJ, Rio de Janeiro/RJ, Brasil)
Maria Bernadete Marques Abaurre (UNICAMP, Campinas/SP, Brasil)
Maria Cecília Camargo Magalhães (PUC-SP, São Paulo/SP, Brasil)
Maria Cecília Magalhães Mollica (UFRJ, Rio de Janeiro/RJ, Brasil)
Maria Cândida Trindade Costa de Seabra (UFMG, Belo Horizonte/MG, Brasil)
Maria Cristina Figueiredo Silva (UFPR, Curitiba/PR, Brasil)
Maria Luíza Braga (PUC/RJ, Rio de Janeiro/RJ, Brasil)
Maria Marta P. Scherre (UNB, Brasília/DF, Brasil)
Micheline Mattedi Tomazi (UFES, Vitória/ES, Brasil)
Miguel Oliveira, Jr. (UFAL, Maceió, Alagoas, Brasil)
Monica Santos de Souza Melo (UFV, Viçosa/MG, Brasil)
Patricia Matos Amaral (UI, Bloomington/IN, Estados Unidos)
Paulo Roberto Gonçalves Segundo (USP, São Paulo/SP, Brasil)
Philippe Martin (Université Paris 7, Paris, França)
Rafael Nonato (Museu Nacional-UFRJ, Rio de Janeiro/RJ, Brasil)
Raquel Meister Ko. Freitag (UFS, Aracaju/SE, Brasil)

Renato Miguel Basso (UFSCar, São Carlos, SP, Brasil).
Roberto de Almeida (Concordia University, Montreal/QC, Canadá)
Ronice Müller de Quadros (UFSC, Florianópolis/SC, Brasil)
Ronald Beline (USP, São Paulo/SP, Brasil)
Rove Chishman (UNISINOS, São Leopoldo/RS, Brasil)
Sanderléia Longhin-Thomazi (UNESP, São Paulo/SP, Brasil)
Sergio de Moura Menuzzi (UFRGS, Porto Alegre/RS, Brasil)
Seung- Hwa Lee (UFMG, Belo Horizonte/MG, Brasil)
Sírío Possenti (UNICAMP, Campinas/SP, Brasil)
Suzi Lima (U of T / UFRJ, Toronto/ON - Rio de Janeiro/RJ, Brasil)
Thais Cristofaro Alves da Silva (UFMG, Belo Horizonte/MG, Brasil)
Tommaso Raso (UFMG, Belo Horizonte/MG-Brasil)
Tony Berber Sardinha (PUC-SP, São Paulo/SP, Brasil)
Ubiratã Kickhöfel Alves (UFRGS, Porto Alegre/RS, Brasil)
Vander Viana (University of Stirling, Stirling/Sld, Reino Unido)
Vanise Gomes de Medeiros (UFF, Niterói/RJ, Brasil)
Vera Lucia Lopes Cristovao (UEL, Londrina/PR, Brasil)
Vera Menezes (UFMG, Belo Horizonte/MG, Brasil)
Vilson José Leffa (UCPel, Pelotas/RS, Brasil)

Sumário / Contents

Apresentação

Sintaxe e semântica comparativas: abordagens atuais nos eixos sincrônico e diacrônico

Comparative Syntax and Semantics: Current Approaches in the Synchronic and Diachronic Axes

Aroldo Leal de Andrade

Ana Paula Quadros Gomes

Carlos Felipe da Conceição Pinto 1521

Formas pronominais

Uma proposta de análise para o objeto acusativo anafórico na variedade de espanhol de Madri e no português brasileiro de São Paulo

An Analysis Proposal for the Anaphoric Accusative Object in the Spanish Variety of Madrid and in the Brazilian Portuguese Variety of São Paulo

Adriana Martins Simões 1547

O pronome pleno de terceira pessoa no “pretuguês” oitocentista

The third Person Full Pronoun in the ‘Pretuguês’ Ninetheenth Century

Fernanda de Oliveira Cerqueira 1596

Resistência do dativo de primeira pessoa na batalha (quase) perdida dos clíticos pronominais do português brasileiro

The Survival of the First Person Dative Pronoun in the (Almost) Lost Battle of Pronominal Clitics of Brazilian Portuguese

Maria Aparecida Torres Morais

Heloisa Maria Moreira Salles 1621

Sintagmas Nominais

Estudo diacrônico do sintagma nominal descontínuo no português brasileiro

Diachronic Study of the Discontinuous Noun Phrase in Brazilian Portuguese

Nathalia Pereira de Souza-Martins

Sebastião Carlos Leite Gonçalves 1659

Posição dos adjetivos nos gêneros *cartas e narrativas* em Português Europeu do século XVII ao XIX

Adjective Position in Letters and Narratives in European Portuguese During the 17th-19th Centuries

Cristina de Souza Prim

Thais Luisa Deschamps Moreira 1688

Nominal Number in Sign Languages

Número nominal em línguas de sinais

Bruno Gonçalves Carneiro

Mônica Veloso Borges

Miroslava Cruz Aldrete..... 1719

Um estudo diacrônico sobre a polaridade negativa no sintagma nominal: o caso do indefinido “algum” na formação do IPN [N + algum]

A Diachronic Study of Negative Polarity in Nominal Phrases: the Case of the Indefinite “Some” in the Formation of IPN [N +algum (Some)]

Cristiane Namiuti

Fernanda Gusmão Silva 1752

Argumentos verbais

Locative Adverbs in the Subject Position in Brazilian Portuguese

Advérbios locativos na posição de sujeito no português brasileiro

Zenaide Dias Teixeira

Humberto Borges 1788

Null Subjects and Null Objects in Brazilian Portuguese: Correlations and Change

Sujeitos e objetos nulos em português brasileiro: correlações e mudança

Gabriel de Ávila Othero

Melissa Lazzari 1831

“Por uma Sociolinguística Românica ‘Paramétrica’” – relendo Tarallo (1987)
e virando a página

*“For a ‘Parametric’ Romance Sociolinguistics” – revisiting Tarallo (1987)
and turning the page*

Maria Eugenia Lammoglia Duarte

Eduardo Patrick Rezende dos Reis 1855

On Partial Null Subject Languages: Why Pro-Drop in Brazilian Portuguese
and Russian Became Similar But Not Identical

*Sobre as línguas de Sujeito Parcialmente Nulo: Por que o pro-drop em
português brasileiro e em russo se tornaram semelhantes, mas não idênticos*

Nerea Madariaga 1896

Semântica: Tempo e aspecto, léxico

A aquisição de Tempo e Aspecto: uma investigação sobre a emergência
de estruturas verbais e a influência do aspecto semântico sobre o aspecto
gramatical no processo de aquisição da linguagem

*The Tense and Aspect acquisition: an investigation on the emergence
of verbal structures and the influence of the semantic aspect on the
grammatical aspect in the language acquisition process*

Arabic Bezri Hermont

Kelly Cesário de Oliveira

Lucas Segantini Brito 1937

Entre accomplishments e atividades: mudanças direcionais e o caso dos
verbos de editoração

*Between accomplishments and activities: directed changes and the case of
publishing verbs*

Gabriela Vilela Souza Martins

Luana Lopes Amaral 1964

A gradação em português brasileiro e a variação translinguística na expressão da comparação

Degree constructions in Brazilian Portuguese and the translinguistic variation in the expression of comparative constructions

Luisandro Mendes de Souza

Roberta Pires de Oliveira

Lara Frutos

Kayron Beviláqua 2015

Mudança semântica e *word embeddings*: estudos de caso na diacronia do português

Semantic change and word embeddings: case studies on the diachrony of Portuguese

Lucas Fonseca Lage

Evandro Landulfo Teixeira Paradela Cunha 2043

Apresentação



Sintaxe e semântica comparativas: abordagens atuais nos eixos sincrônico e diacrônico

Comparative Syntax and Semantics: current approaches in the synchronic and diachronic axes

Aroldo Leal de Andrade

Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Minas Gerais / Brasil

aroldoleal@letras.ufmg.br

<http://orcid.org/0000-0002-1354-7916>

Ana Paula Quadros Gomes

Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Rio de Janeiro / Brasil

anaquadrosomes@letras.ufrj.br

<http://orcid.org/0000-0002-3476-0193>

Carlos Felipe da Conceição Pinto

Universidade Federal da Bahia (UFBA), Salvador, Bahia / Brasil

cfpinto@ufba.br

<http://orcid.org/0000-0002-4322-0199>

Resumo: O objetivo do presente texto é o de oferecer um panorama das direções atuais de pesquisa em sintaxe e semântica comparativas, considerando um breve histórico das duas áreas e as suas tendências internacionais atuais. Uma vez que a Linguística Contrastiva ainda apresenta desenvolvimento relativamente incipiente nos espaços lusófonos, as principais correntes de pesquisa em sintaxe e semântica diacrônica são apresentadas nas vertentes formalista, funcionalista e variacionista, assim como os grupos de pesquisa que os representam, sobretudo no Brasil e em Portugal. Finalmente, descrevem-se as contribuições dos artigos que compõem o dossiê que acompanha o presente texto, assim como uma reflexão sobre os rumos que eles apontam para estudos futuros.

Palavras-chave: sintaxe; semântica; comparação; mudança linguística.

Abstract: The aim of this text is to offer an overview of current research directions in comparative syntax and semantics, considering a brief history of the two areas and their current international trends. As Contrastive Linguistics still presents relatively incipient development in Portuguese-speaking areas, the main paths of research in diachronic syntax and semantics are presented in the formalist, functionalist and variationist strands, as well as the research groups that represent them, especially in Brazil and Portugal. Finally, the contributions of the articles that make up the special issue that accompanies this text are described, as well as a reflection on the directions that they point for future studies.

Keywords: syntax; semantics; comparison; linguistic change.

1 Introdução

O objetivo do presente texto é refletir sobre os rumos atuais da sintaxe e da semântica em perspectiva comparada e diacrônica. O evento que ensejou a presente publicação, o *V Congresso Internacional de Linguística Histórica* (V CILH: Constelações Diacrônicas), realizado online em julho de 2021, homenageou duas figuras que contribuíram para o desenvolvimento de estudos nessas áreas, e em especial da sintaxe diacrônica: Charlotte Galves e Mary Kato. Por essa razão, os organizadores do evento consideraram que seria bastante adequada a proposição de uma publicação especializada nessas áreas. Note-se, no entanto, que o presente volume esteve aberto a manuscritos que abordassem um desses temas, mesmo que indiretamente, e a trabalhos que não foram apresentados no referido evento.

O texto se organiza da seguinte forma: na seção 2, discorreremos sobre os desenvolvimentos atuais das áreas de sintaxe e semântica diacrônica e comparativas, dando especial relevância às novas vertentes teóricas. A seção 3 representa uma continuidade dessa discussão, ao apresentar os desenvolvimentos dessas áreas em países lusófonos, com especial relevância para Brasil e Portugal. A seção 4, por outro lado, apresenta de forma resumida os temas desenvolvidos nas contribuições que constam do dossiê que acompanha o presente texto, na *Revista de Estudos da Linguagem*. Na seção 5, tecem-se as considerações finais, reunindo as questões que se abrem a partir das reflexões feitas no presente volume.

2 Desenvolvimentos atuais em sintaxe e semântica comparativas

Nas áreas de sintaxe e semântica, nota-se uma progressiva aproximação temática dos campos formalista e funcionalista, ao mesmo tempo em que as duas perspectivas buscam firmar diferenças no campo epistemológico, tendo como fruto uma progressiva reflexão sobre o papel da mudança gramatical. Como se sabe, uma diferença crucial entre as duas visões é a possibilidade de interrelação entre os dois níveis de análise (semântico e sintático), bem aceita no funcionalismo, mas não no formalismo. Crucialmente, abordagens funcionalistas valorizam a dimensão diacrônica como peça-chave para a compreensão da gramática sincrônica, o que não é necessariamente verdade na visão formal. Talvez exatamente por isso os estudos da mudança, tanto sintática quanto semântica, tenham sido pautados desde sempre no funcionalismo – geralmente em correlação com mudanças em outras áreas do sistema linguístico –, enquanto no formalismo eles se iniciaram nos anos 1970 (sintaxe) e 2000 (semântica) – cf. Lightfoot (1979) e Eckardt (2006).

No entanto, novos desenvolvimentos da chamada biolinguística trouxeram uma discussão sobre a mudança linguística no campo formal, propiciada via compreensão da aquisição da linguagem.

2.1 Sintaxe

Na linguística formal há atualmente duas grandes propostas explicativas para a mudança sintática, que se diferenciam em torno do tratamento dado à interrelação entre mudança e tipologia linguística: de maneira previamente estruturada (Hierarquias de Parâmetros – vide ROBERTS, 2019) ou por meio de predições baseadas nos dados (Gatilhos nos dados de Aquisição – vide LIGHTFOOT, 2020). A abordagem de Roberts (2019) é na verdade uma revisão da Teoria de Princípios e Parâmetros em sua versão original (CHOMSKY, 1981), de tal forma a fazer face à granularidade da variação e aos pressupostos minimalistas que requerem a redução do aparato inato postulado em torno da Faculdade da Linguagem. Dela deriva uma série de postulados sobre caminhos para a mudança linguística – especialmente a sintática –, que levam a sério princípios de marcação (no sentido de *markedness*) e dialogam com reflexões minimalistas ligadas ao chamado Terceiro Fator na Aquisição da Linguagem (CHOMSKY, 2005, entre outros): a Economia de Traços, a Generalização do Input e o Princípio de Subconjunto. Assim, os percursos

de mudança sintática no nível macro dialogam com aqueles no nível micro – inclusive a gramaticalização –, indicando uma relação possível entre dados externos e tendências de mudança internas.

Por outro lado, a abordagem de Lightfoot (2020) considera a importância crucial dos gatilhos (*triggers*) presentes dos dados de aquisição para a mudança paramétrica. Os gatilhos, trechos de sentenças que expressam um determinado parâmetro, que também são denominados de pistas robustas, isto é, que não são ambíguas, geram assim possíveis casos de reanálise. Conforme Kroch (2021, p. 27-28), sua proposição pressupõe que a mudança gramatical depende da existência de consistentes desvios no uso durante longos períodos. Sem dúvida, nesta concepção, a relevância do componente externo à linguagem é muito maior: há uma sensibilidade à frequência dos dados, a qual pode ser desencadeada por interferências das interfaces (como a prosódia e a pragmática) e/ou por interferências advindas do contato linguístico. A grande vantagem desse modelo é sua maior plasticidade face aos dados de diferentes línguas, não apresentando expectativas prévias de mudanças que podem ou não ocorrer. Por outro lado, ele descarta a possibilidade de mudanças gramaticais motivadas internamente à gramática.

Um exemplo claro para distinguir as duas visões pode se orientar, como é frequente na literatura, em torno da discussão sobre o parâmetro do sujeito nulo. Na visão de Roberts (2019), esse parâmetro envolve na verdade uma série de marcações paramétricas relacionadas hierarquicamente em torno da expressão dos traços *phi* (marcas de gênero, número e pessoa), partindo do macroparâmetro, que separa línguas em que os traços *phi* não têm qualquer papel no sistema (como o japonês e outras línguas chamadas de sujeito nulo radical) daquelas em que esses traços têm um papel; passando pelo mesoparâmetro, que distingue línguas em que os traços *phi* estão especificados em todos os núcleos funcionais (como o basco e outras línguas de argumento pronominal) das que apresentam essa especificação em um subconjunto desses núcleos; e chegando ao microparâmetro, que especifica a presença de traços *phi* nos núcleos Tempo e Definitude (como o italiano e outras línguas de sujeito nulo consistente), distinguindo-as das que apresentam essa informação só na categoria Tempo (como o finlandês e outras línguas de sujeito nulo parcial), das que apresentam só na categoria Definitude (como o inglês e outras línguas de sujeito não nulo). Os chamados nanoparâmetros podem atuar junto a itens lexicais específicos, quebrando o comportamento

sintático uniforme esperado, como se observa em gramáticas com clíticos-sujeito – que podem ser vistas como um subcaso de línguas de sujeito nulo consistente, em que opera um requisito extra, de realização separada do núcleo Tempo na forma de um clítico-sujeito. Um exemplo é encontrado no dialeto vêneto de Cornuda (ROBERTS, 2019, p. 288), em que a realização do clítico-sujeito somente é obrigatória nos casos em que há um auxiliar:

- (1) a. Nisun l'è rivà.
 ninguém clS.3sg=é chegado
 'Ninguém chegou.'
- b. Nisun riva.
 ninguém chega
 'Ninguém chega.'

Já uma proposta como a de Lightfoot não prevê a existência de tipos definidos de línguas de sujeito nulo. Assim, a existência de línguas de sujeito nulo parcial geneticamente tão distantes como o português brasileiro, o finlandês e o hebraico é só o reflexo de mudanças independentes, o que prevê que haja diferenças importantes no sistema gramatical de cada uma delas que dificilmente poderão ser reduzidas a efeitos diretos de parâmetros.

No campo funcionalista, as explicações para a mudança (sintática) tendem a considerar tanto motivações internas quanto externas. As motivações internas não raro são guiadas por requisitos de eficiência comunicativa (e não por princípios de aprendibilidade). Assim, desde Martinet (1962), considera-se que as tendências de simplicidade e clareza podem atuar por meio de mecanismos compensatórios, de forma a manter a eficiência comunicativa quando uma mudança afeta uma área do sistema. De acordo com Mithun (2003), além desse desiderato geral de “equilíbrio de forças” – colocado em prática de diferentes formas, nos diferentes modelos funcionalistas – há que se ter em conta dois outros mecanismos gerais de mudança: a habilidade de rotinização em consequência da alta frequência de uso, o que facilita a reanálise de estruturas gramaticais; e a extensão de padrões de uso para novas funções comunicativas, com reflexos em novas áreas da gramática. Nos modelos de gramaticalização, esses dois últimos tipos de mudança têm sido apresentados em correlação, como etapas de um mesmo processo

mais amplo. Exemplos desses fenômenos no português brasileiro são encontrados abaixo:

- (2) a. Você/A gente não quer só comida.
b. Não se quer só comida.
- (3) a. Eu achei uma moeda na calçada.
b. Vai chover, acho.

O exemplo em (2) corresponde a uma situação típica de “equilíbrio de forças”, relacionada à expressão do sujeito – no caso, indeterminado – como consequência do enfraquecimento do paradigma de concordância no verbo (NEVES, 2000). Já em (3), tem-se um caso de gramaticalização de um verbo lexical como uma marca de evidencialidade, que se acompanha de maior frequência e de mobilidade posicional do elemento em causa (*acho*) (cf. CASSEB-GALVÃO, 1999).

Há ainda abordagens que, cada vez mais, indicam uma via que reforça metodologias próprias de compreensão da mudança sintática e/ou semântica, como aquelas apresentadas pela linguística variacionista e pela linguística contrastiva. A linguística variacionista apresenta uma metodologia sobre como relacionar variação e mudança. Dessa forma, apesar de historicamente se apresentar alinhada à perspectiva funcionalista, pode ser aplicada a diferentes teorias, desde que estas tenham sensibilidade às informações advindas da variação (DUARTE, 2016). Assim, alguns estudiosos têm desenvolvido estudos em sociolinguística paramétrica (aliando a metodologia variacionista à teoria gerativa, cf. TARALLO, 1985), sobretudo para estudos de mudança em tempo aparente, enquanto outros têm utilizado o conceito de competição de gramáticas (cf. KROCH, 1989) para trabalhar com estudos variacionistas em tempo real. Os estudos variacionistas podem apresentar novos detalhamentos sobre como a mudança pode se processar, a partir de uma visão empiricista e rigorosa.

No campo comparativo, cada uma das perspectivas teóricas – formalista e funcionalista – apresenta uma metodologia própria de avaliação do grau de distância entre línguas do ponto de vista tipológico, e propostas para explicar tais diferenças, como vimos em parte. No entanto, a linguística contrastiva tem oferecido novas ferramentas para comparação entre línguas, o que tem sido feito sobretudo no plano sincrônico. Como observa Ke (2019), essa área tem uma aplicação importante, ao se constituir como um campo de testes para o alcance de

diferentes teorias. No entanto, o mesmo pode ocorrer no plano diacrônico, ao se observarem sistematicidades em como os sistemas mudam com o tempo (HARBERT, 2012). Nesse intuito, a linguística contrastiva pode também fazer uso de corpora, sobretudo os anotados. Em consequência de esforços como esses, talvez seja possível, no futuro, elaborar uma lista (mesmo que não exaustiva) de mudanças sintáticas, assim como aquela que já está disponível para mudanças no plano lexical, envolvendo a gramaticalização (cf. HEINE; KUTEVA, 2002).

2.2 Semântica

O campo linguístico da semântica tem seu início situado no século XIX. A linguística histórica com ênfase na semântica se desenvolveu no século XX, em especial a partir de 1980, sobretudo com o clamor de cognitivistas de que a investigação diacrônica poderia afinar a nossa compreensão sobre processos semânticos fundamentais. É tudo bastante recente. Não obstante, há distintas respostas à pergunta sobre quais explicações dão conta da mudança de significação ao longo do tempo.

A linguística cognitiva, uma abordagem em que a semântica é central, reavivou o interesse pelos mecanismos e processos de mudança linguística. Os modelos de análise semântica baseados em protótipos promoveram a investigação de expressões individuais, muitas vezes sem considerar sentenças completas ou sua inserção em contextos gramaticais. O conhecimento sobre os mecanismos ligados a metáforas e metonímias lançou luz sobre a ligação entre sentidos e estruturas. Surgiram concepções pragmáticas e contextualizadas de léxico (GEERAERTS, 2019). Entre as propostas mais difundidas para explicar as regularidades da mudança semântica estão a gramaticalização, a subjetificação e a metaforização (TRAUGOTT; DASHER, 2001; LAKOFF, 2008).

Um hibridismo entre a abordagem funcionalista e a cognitiva gerou a abordagem construcional da gramática, que lança mão das ideias de Traugott (2008), Bybee (2010), Givón (1979, 1995), Croft (2001), Goldberg (1995, 2006) e Fillmore e Kiparsky (1968). O interesse da pesquisa funcionalista clássica pela gramaticalização incorpora a dimensão contextual, concebendo a estrutura linguística como derivada de processos cognitivos gerais. Já a visão construcional da gramática passa a ver a mudança gramatical como a teoria das relações entre pares de forma-sentido e tem como certa sua provável direcionalidade ao longo do

tempo, assumindo-se como uma teoria do uso (ROSÁRIO; OLIVEIRA, 2016). Um expoente dessa linha é Martelotta (2011).

A grande maioria dos estudos relacionando a evolução linguística à aquisição se concentra no estudo do léxico. Uma hipótese aventada é a de que os itens lexicais de aquisição mais tardia e mais difíceis de processar estejam mais sujeitos a mudanças.

Uma análise semântica lógica ou formal, à maneira de Montague (1974), era ainda vista como não estando equipada para estudar a mudança, como chegou a ser dito por várias estrelas da semântica cognitiva, tais como Geeraerts (1997) e Langacker (1990). Investigar a mudança ou reanálise semântica no paradigma da veridicondicionalidade é um empreendimento recente, mas nem por isso sem frutos. Nessa vertente, as palavras, os sintagmas e os morfemas que passam por mudança semântica ocupam diferentes posições sintáticas antes e depois de sua reanálise. A reanálise é moldada pela interface sintaxe-semântica, e respeita universais linguísticos. Nesse campo, as ferramentas da semântica formal são vistas como muito eficazes, pois a estrutura sintática conversa com a composicionalidade semântica, e a semântica de linha formal tem o poder de capturar o significado de qualquer expressão linguística, seja ela de que categoria lexical for (ECKARDT, 2006).

Enfim, a semântica diacrônica é um campo bem novo, mas bastante promissor em muitas frentes, pois fundamentos semânticos intrínsecos à linguagem tendem a vir à tona nas investigações da área.

3 Desenvolvimentos das áreas nos países lusófonos

No cenário lusófono, como em outros lugares do mundo, seguindo a tradição de estudos no Ocidente, pouca relevância tem sido dada às áreas de sintaxe e de semântica diacrônicas e/ou comparadas. Os estudos em Linguística Contrastiva, incluindo os que deram continuidade à Linguística Histórico-Comparativa, no entanto, estiveram bem menos representados no Brasil, assim como em outros países lusófonos. A exceção diz respeito aos estudos sobre a genética de línguas indígenas e aos estudos sobre Linguística Aplicada ao ensino de línguas estrangeiras. Evidentemente, há algumas exceções de trabalhos tipológicos de cunho funcionalista ou gerativista – mais recentemente, incluem-se preocupações sobre a Libras, face a outras línguas de sinais.

Os trabalhos em perspectiva diacrônica em Portugal têm se dedicado principalmente aos planos da fonologia, da morfologia, da sintaxe e da semântica lexical. Graça Maria Rio-Torto (2008) é um dos expoentes dos estudos de mudança lexical e morfológica. Os processos de gramaticalização têm vindo a ser observados e estudados em Portugal sobretudo no âmbito da Pragmática, com destaque para os estudos de Ana Cristina Macário Lopes, dedicados a vários operadores discursivos, mas também há estudos no âmbito da Linguística Cognitiva.

No nível de análise semântico, muitos dos estudos precursores da diacronia em língua portuguesa estão ligados à variação dialetal e à sociolinguística, incluindo a diferenciação entre o português europeu e o português brasileiro, e em relação ao português de países africanos, como o português angolano, o moçambicano etc. Um exemplo bem conhecido é o estudo da troca do verbo *haver*, que, no século XIII, era tanto um verbo existencial quanto um verbo de posse, pelo verbo *ter*, de posse, mudança iniciada ainda no período do português antigo. Em PB, essa mudança foi ainda mais longe que em PE, pois o verbo *ter* substituiu *haver* também no sentido existencial. Outra área de mudança na história linguística relacionada à semântica já bastante estudada no mundo lusófono é a da negação, incluindo palavras negativas e itens de polaridade (MARTINS, 2016). Em Portugal, estudos como os de Ana Maria Martins, na interface sintaxe-semântica, que traçam o percurso de itens lexicais do latim até o português contemporâneo, têm feito escola nos últimos anos.

Em Portugal há importantes grupos de pesquisa em sintaxe diacrônica, sobretudo em Lisboa, em que também há abordagens na interface sintaxe-semântica, o que reflete a maior influência dos estudos gerativistas naquela cidade, tanto na Universidade de Lisboa (UL) quanto na Universidade Nova de Lisboa (UNL). No entanto, em Coimbra também há estudos sobre semântica diacrônica. Os estudos ainda dão maior relevo ao português antigo, incluindo o período pré-literário.

Na Galícia, os estudos diacrônicos têm se desenvolvido sobretudo em conjunto com os estudos em filologia e dialetologia, na Universidade de Santiago de Compostela (USC), o que indica um grande campo aberto para a sintaxe e a semântica diacrônicas, que começam a ser feitas com o aparato da linguística contrastiva, com o português europeu e o asturiano.

No Brasil, podem ser considerados precursores dos estudos gramaticais comparativos Mattoso Câmara Jr., considerado o introdutor

da ciência linguística no Brasil, e Rosa Virgínia Mattos e Silva, fundadora do Programa para a História da Língua Portuguesa (PROHPOR), projeto que produziu muitos trabalhos em diacronia de natureza fônica, morfossintática e sintática. Os estudos em sintaxe diacrônica começaram a se desenvolver claramente nos anos 1990, ao mesmo tempo em várias perspectivas concomitantes: na visão formal, graças à colaboração inicialmente formulada entre Ian Roberts e Mary Kato (cf. ROBERTS; KATO, 1993); na visão funcionalista, com a tese de BITTENCOURT (1995), seguida dos trabalhos sobre gramaticalização envolvendo vários grupos de pesquisa (GONÇALVES; LIMA-HERNANDES; CASSEB-GALVÃO, 2007); e, na visão variacionista, em que logo de início se percebeu a relevância de não se limitar a trabalhar com a mudança em tempo aparente (cf. TARALLO, 1991). Os estudos dedicados ao fenômeno de gramaticalização desenvolvidos no Brasil adotavam ou a perspectiva da Gramática (Cognitivo-) Funcional ou a da Linguística Cognitiva (FERRARI, 1998), abordando verbos, preposições e, sobretudo, construções com funções discursivas.

Na semântica comparativa e diacrônica, os funcionalistas integram os achados diacrônicos à compreensão da semântica sincrônica, alimentando a formulação de modelos. Já a semântica formal aportou no Brasil por meio de nomes como Rodolfo Ilari, após cerca de duas décadas de sua aparição nos EUA. Ele foi o organizador, junto com Maria Helena Moura Neves, do vol. 2 da coleção *Gramática do Português Culto Falado no Brasil*, um trabalho sucedâneo do Projeto NURC, dedicado às classes de palavras e processos de construção (ILARI; NEVES, 2008), e, mais recentemente, de um volume específico da coleção *Para a História do Português Brasileiro*, dedicado à semântica diacrônica (cf. ILARI; BASSO, 2021), o que reflete a mais recente identificação dessa área como um campo autônomo. Essa linha tem se desdobrado em vertentes muito profícuas, mas com dedicação primordial a recortes sincrônicos. Poucos trabalhos sobre a diacronia nesta vertente são conhecidos, o que faz desse volume um marco inaugural para os estudos semânticos em diacronia no Brasil.

A linguística histórica nessas duas subáreas apresenta grande capilaridade no País em várias instituições, graças a esforços coletivos de grupos de pesquisa de relevância nacional e do apoio financeiro das agências de fomento a projetos de longa duração, que possibilitaram a criação de corpora históricos e o desenvolvimento de recursos

humanos especializados. As pesquisas, que num primeiro momento combinavam dois extremos – a descrição sincrônica do português brasileiro e a do português antigo – mais recentemente têm incluído as etapas intermediárias da mudança – o português clássico em Portugal e o português colonial no Brasil.

Nos países africanos – especialmente Angola e Moçambique, em que há centros universitários consolidados, bem como há em Macau –, os estudos diacrônicos têm se centrado sobre questões relativas ao contato linguístico, a fim de compreender seja a formação de novas variedades de português, seja de línguas crioulas. Muitas vezes, tais estudos têm tido apoio de pesquisadores no Brasil e em Portugal. O trabalho de Gonçalves (2010) é um bom exemplo de descrição de uma nova variedade (português moçambicano) ancorada numa visão de mudança linguística.

As várias edições do *Congresso Internacional em Linguística Histórica* (que vêm se sucedendo desde 2009) contribuíram para o desenvolvimento dessas áreas nos países lusófonos, pois esse é um evento tradicionalmente de fala galego-portuguesa, tendo já sido sediado no Brasil, em Portugal e na Galícia. O V CILH: Constelações Diacrônicas, que esteve organizado em workshops, sessões livres e sessões de pôsteres, contou com três workshops dedicados às áreas de sintaxe e semântica diacrônica ou comparativa, e mais de trinta trabalhos dedicados às áreas de sintaxe e semântica apresentados como comunicações em sessões livres e em pôsteres. Os workshops, propostos como espaços de debate aprofundado sobre temas específicos, tiveram maior participação de pesquisadores originários do exterior do país. Os três workshops foram “Comparative Syntax in the 21st century”, organizado por Sonia Cyrino (Universidade Estadual de Campinas); “Universals and Variation in Semantics and Pragmatics”, organizado por Ana Paula Quadros Gomes (UFRJ) e por Luciana Sanchez Mendes (UFF) e “Modelos para o estudo da mudança sintática: em que consistem, como são aplicados e quais as suas vantagens”, organizado por Maria Eugenia Duarte e Silvia Cavalcante (UFRJ). Nas sessões livres e de pôsteres, notou-se a grande variedade tanto de abordagens teóricas (formalistas e funcionalistas) quanto de métodos de pesquisa (trabalhos baseados em corpora ou em experimentos), ficando estes concentrados na área de Linguística Teórica. Não obstante, em outras áreas do evento houve também trabalhos que abordavam a sintaxe e semântica comparadas, já que o congresso tem abertura para qualquer estudo relacionado à chamada linguística

histórica *lato sensu* (MATTOS E SILVA, 1999, 2016), isto é, aquela que inclui trabalhos com descrições sincrônicas datadas e localizadas (historicamente contextualizadas).

4 As contribuições do presente dossiê

As contribuições do presente dossiê podem ser divididas em temas relacionados às áreas da estrutura da gramática com as quais mais se relacionam: formas pronominais; sintagmas nominais; a sintaxe dos argumentos verbais; flexão e traços lexicais. As abordagens sobre os três primeiros temas se concentram na sintaxe, enquanto a semântica está mais bem representada nos dois últimos temas.

4.1 Formas pronominais

Os artigos que tratam de formas pronominais apresentaram como principal ponto de interesse a variação entre formas foneticamente nulas e expressas. O artigo assinado por Adriana Simões, “Uma proposta de análise para o objeto acusativo anafórico na variedade de espanhol de Madri e no português brasileiro de São Paulo”, descreve a variação entre o uso de clíticos ou de objetos nulos no espanhol de Madri, ou entre pronomes fracos e objetos nulos no português falado em São Paulo, a partir de uma abordagem da Sociolinguística Paramétrica, considerando, para tanto, variáveis semânticas e sintáticas.

A contribuição de Fernanda Cerqueira (“O pronome pleno de terceira pessoa no ‘português’ oitocentista”) apresenta um estudo baseado em corpus de atas de afrodescendentes brasileiros sobre as variadas manifestações do pronome ‘ele’ e suas variantes, a partir de uma perspectiva de análise formal. Os resultados demonstraram que a reanálise do pronome de terceira pessoa de tal forma a ocupar novas funções gramaticais, típicas de um pronome fraco, ainda estava em estágio insipiente. O trabalho ilustra a importância de comparações com outros corpora, representativos de diferentes gêneros textuais, de forma a confirmar como se deu a reanálise relacionada ao estatuto desse pronome.

O artigo de Maria Aparecida Torres Morais e Heloisa Salles (“Resistência do dativo de primeira pessoa na batalha (quase) perdida dos clíticos pronominais do português brasileiro”) investiga a cisão pronominal na expressão dos dativos no português brasileiro (PB), a qual consiste na presença dos clíticos de 1ª e 2ª pessoas *me* e *te*, em oposição

aos pronominais tônicos *ele(s)*, *ela(s)*, introduzidos por *a* e *para*, em decorrência da perda dos clíticos dativos *lhe(s)*. As autoras propõem que as propriedades sintáticas e semânticas do clítico dativo *me* sejam unificadas como núcleo Aplicativo Alto, nos termos da tipologia proposta em Pylkkänen (2008).

4.2 Sintagmas Nominais

Os artigos que tratam de sintagmas nominais abordam a estrutura interna desse tipo de constituinte a partir de diferentes perspectivas teóricas. O artigo de Nathalia de Souza-Martins e Sebastião Gonçalves, intitulado “Estudo diacrônico do sintagma nominal descontínuo no português brasileiro”, estuda características de SNs descontínuos, comparando estratos de língua escrita e falada, a partir do quadro teórico da Gramática Discursivo-Funcional. Discute, para além dos princípios de peso por complexidade e de peso estrutural, enunciados em trabalho anterior citado, funções retóricas, de mitigação e de preservação do escopo. O texto demonstra que o uso de SNs descontínuos permaneceu em pequena frequência na evolução do português brasileiro, de acordo com o corpus estudado.

A contribuição de Cristina de Souza Prim e Thais Moreira (“Posição de adjetivos nos gêneros *cartas* e *narrativas* em português europeu do século XVII ao XIX”) descreve, a partir de pesquisa baseada em *corpus*, a posição de variados tipos de adjetivos em SNs com diferentes configurações sintáticas na história do português europeu. O artigo demonstra que há preferências que guiam a variação nos dois períodos (anteposição de adjetivos avaliativos e de nomes [+humanos]) e que a mudança teria se dado mais pelo aumento da frequência de adjetivos não avaliativos, nos quais não há variação entre as posições pré- e pós-nominal.

O artigo de Bruno Carneiro, Mônica Borges e Miroslava Aldrete (“Nominal number in sign languages”) descreve a manifestação da categoria Número em dez línguas de sinais a partir de uma perspectiva tipológica, partindo de observações feitas para as línguas orais. O texto observa que as marcas de número podem ser opcionais ou obrigatórias, havendo preferência por estratégias sintáticas (típicas de línguas isolantes). Já as estratégias morfológicas apresentaram uma hierarquia implicacional, pela qual o *mouthings* (produção de sílabas visuais com a

boca, concomitantes a um sinal manual) é mais marcado, sendo, portanto, a estratégia menos recorrente.

Finalmente, o artigo “Um estudo diacrônico sobre a polaridade negativa no sintagma nominal: o caso do indefinido ‘algum’ na formação do IPN [N + algum]”, de Cristiane Namiuti e Fernanda Gusmão Silva, debruça-se sobre a origem da posposição de *algum* com a função de item de polaridade negativa do PB [N + *algum*], analisando dados do português a partir do século XVI. A estrutura [N + *algum*] é examinada em textos portugueses do século XVII, XVIII e XIX, e em textos brasileiros do século XIX, em busca do seu processo de gramaticalização como um Item de Polaridade Negativa (IPN) na história da língua portuguesa. Como mostram as autoras, no português clássico dos séculos XVI e XVII a estrutura com inversão não tinha sempre valor negativo, e outros elementos podiam ocorrer no interior do sintagma nominal, inclusive com a possibilidade da flexão e não-adjacência entre o nome e o indefinido. Para elas, os fatos indicam que então ainda não havia ocorrido a gramaticalização da estrutura como um IPN. No português clássico, a negação era valorada sob o domínio sintático de um operador de negação sentencial, de uma preposição ou de uma conjunção negativa. A análise proposta para o item *algum* dessa época é a de que esse indefinido realiza o núcleo da categoria determinante (D) mesmo nas estruturas com inversão, sendo estas derivadas do movimento da parte lexical do sintagma à posição de especificador do Sintagma Determinante (D). A proposta é condizente com as propriedades de categoria D encontradas no item *algum* do português clássico.

4.3 Argumentos verbais

O grupo de artigos que trabalha com a sintaxe de argumentos verbais examina a variada realização de sujeitos e objetos, tendo o português brasileiro como língua central e sempre trabalhando numa perspectiva de sintaxe formal, às vezes com apoio de dados empíricos extraídos de corpora. No artigo “Locative adverbs in the subject position in Brazilian Portuguese”, Zenaide Dias Teixeira e Humberto Borges examinam sentenças do PB com um locativo (preposicionado ou não) em posição de sujeito, como *Aqui bate sol*, cotejando as análises sintáticas disponíveis na literatura com os dados, desde Pontes (1987). Com base nas propriedades nominais apontadas em Baker (2003), os

autores explicam o fato de advérbios e SNs locativos poderem ocupar a posição de sujeito em PB como uma consequência de possuírem índice referencial. Segundo essa análise, advérbios e SNs locativos apresentam parte das propriedades de SNs plenos, a saber, o traço de pessoa (embora não tenham o de número, o que resulta em concordância verbal *default*, na 3ª. pessoa do singular). A proposta é a de que esses advérbios em posição de sujeito sejam gerados como advérbio de SV ou argumento locativo. Não obstante a análise sintática proposta seja de cunho sincrônico, o fenômeno é situado num quadro de mudança do PB em direção a um maior preenchimento do sujeito.

O artigo “Null subjects and null objects in Brazilian Portuguese: correlations and change”, de Gabriel de Ávila Othero e Melissa Lazzari, desenvolve a proposta de Creus e Menezzi (2004), de que o gênero semântico (animado feminino, animado masculino, inanimado) determina a distribuição dos objetos diretos realizados por pronomes ou nulos, assumindo uma correlação entre sujeitos nulos e objetos diretos anafóricos na língua. Os autores constatam que o sujeito nulo de 1ª. e 2ª. pessoas do singular está estável, mas apontam que os sujeitos nulos e pronominais de 3ª. pessoa estão em mudança aparente. Após ter sido revisada a literatura sobre o fenômeno, bem como apresentada uma análise de corpus, um processo de mudança em curso foi desenhado: sujeitos nulos de terceira pessoa ainda não se estabilizaram no português brasileiro; ainda assim, a assimetria entre, de um lado, sujeitos de 1ª. e 2ª. pessoa e, de outro, sujeitos de 3ª. pessoa não tem mais sustentação empírica, segundo os achados dos autores.

A contribuição de Maria Eugénia Duarte e Eduardo dos Reis, intitulada “Por uma Sociolinguística Românica ‘Paramétrica’ – relendo Tarallo 1987 e virando a página”, apresenta uma discussão mais ampla sobre o legado de Fernando Tarallo, dentro da qual desenvolvem considerações em torno da realização da posição de sujeito. São apresentados os critérios de variação que condicionam a realização de sujeitos nulos no português europeu, uma língua de sujeito nulo consistente, e no português brasileiro. Com base na variação nas duas gramáticas em fatores semelhantes (com a devida diferença de aplicação da regra), os autores eliminam a possibilidade de classificar o português brasileiro como uma língua de sujeito nulo parcial, e em vez disso defendem que se trata de uma gramática em mudança para uma língua de sujeito não-nulo. Em conclusão, demonstram que a Teoria da Variação

e Mudança auxilia na identificação de parâmetros sintáticos, discussão que já deveria, a seu ver, ser considerada uma “página virada”.

O texto de Nerea Madariaga, “On Partial Null Subject Languages: Why pro-drop in Brazilian Portuguese and Russian became similar but not identical”, assume,, diferentemente daquele de Duarte e Reis, que o português brasileiro é uma língua de sujeito nulo parcial, assim como o russo. No entanto, o comportamento dessas duas línguas em relação a uma série de critérios sintáticos relacionados à manifestação do sujeito nulo é diferente. Adotando uma visão de diacronia baseada em Lightfoot (1999), Madariaga mostra que a suposta mudança de uma língua de sujeito nulo consistente a parcial não apresenta um correlato obrigatório com a perda de concordância, como os estudos sobre o português brasileiro fazem pensar. Em suma, a noção de mudança paramétrica se mantém, pois há uma alteração em comum, relacionada à perda do traço-D em T, por sua vez associada a uma série de propriedades, porém acompanhada de pequenas diferenças atribuídas a diferentes condições iniciais do estágio *pro-drop* para o russo e o PB, enquanto línguas de sujeito nulo parciais.

4.4 Semântica: Flexão e traços lexicais

O grupo de artigos que trata de questões semânticas trabalha fortemente com aparatos de orientação formal, mas também com metodologias bem consagradas da semântica lexical e da linguística computacional. Os três primeiros trabalhos dedicam-se a questões ligadas à flexão de tempo, modo e aspecto (incluindo aspecto lexical), ou à noção de gradação; e o último, a uma investigação sobre campos semânticos no léxico. O artigo “A aquisição de Tempo e Aspecto: uma investigação sobre a emergência de estruturas verbais e a influência do aspecto semântico sobre o aspecto gramatical no processo de aquisição da linguagem”, de Arabie Hermont, Kelly Cesário de Oliveira e Lucas Segantini Brito, pesquisa a aquisição de formas verbais relacionadas a tempo e a aspecto a partir da hipótese da Primazia do Aspecto (ANDERSEN; SHIRAI, 1996). Para tanto, em um estudo transversal, foram examinados os registros de produções de crianças monolíngues entre 2 anos de idade e 5 anos e 11 meses de idade. As duas hipóteses estabelecidas foram confirmadas pelos resultados: (i) na gramática infantil, a noção de aspecto surge, de forma mais recorrente, antes da noção de tempo e (ii) quando as formas verbais no presente e no pretérito perfeito surgem, o que está expresso é o aspecto. Delineou-se ainda um

quadro que enseja futuras investigações: (i) o aparecimento mais tardio do pretérito imperfeito, em relação ao presente e ao pretérito perfeito; (ii) o surgimento do pretérito imperfeito junto com as perífrases verbais, expressões analisadas como denotando tempo.

Já Gabriela Martins e Luana Amaral, no artigo “Entre *accomplishments* e atividades: mudanças direcionais e o caso dos verbos de editoração”, põem em xeque a operatividade dos critérios de distinção entre classes acionais vendlerianas, tomando como dados empíricos os verbos do campo semântico de editoração em língua inglesa e em português brasileiro. O conjunto de verbos estudado é especialmente permeável a oscilações entre as características de uma atividade e as características de um *accomplishment*. Adotando o modelo bidimensional de representação do aspecto lexical, numa abordagem cognitivo-funcional, as autoras apontam que verbos como ‘*to edit*’/ *editar* ou *editorar*, ‘*to copyedit*’/ *preparar*, ‘*to proofread*’/ *revisar*, ‘*to typeset*’/ *diagramar* e ‘*to format*’/ *formatar*, por especificarem um tipo de mudança direcional, têm potencial para serem (re)construídos como mudanças direcionais télicas ou atélicas. As autoras concluem que não é a natureza do objeto que distingue telicidade de atelicidade, nos dados examinados, mas é a incrementalidade da ação descrita pelo verbo que é responsável pelo predicado verbal se apresentar como uma atividade ou como um *accomplishment*.

O artigo “A gradação em português brasileiro e a variação translinguística na expressão da comparação”, de Luisandro Mendes de Souza, Roberta Pires de Oliveira, Lara Frutos e Kayron Beviláqua, se insere na vertente, bastante recente, que busca por universais e parâmetros semânticos. Uma vez constatada a coexistência de diversas formas para realizar a comparação, é preciso explicar teoricamente essa variação. Os autores assumem a proposta de três parâmetros para as comparativas nas línguas naturais, na linha de Beck, Oda e Sugisaki (2004) e de Beck et al. (2009). Os três parâmetros propostos são: (i) língua com predicados graduais (DSP – Degree Structure Parameter); (ii) língua com ligação de variáveis na sintaxe (DAP - Degree Abstraction Parameter); e (iii) língua com uma posição sintática de grau preenchida (DegPP – Degree Phrase Parameter). Os autores defendem que o português brasileiro apresenta características híbridas, recebendo marcação positiva nos dois primeiros parâmetros e negativa no terceiro. Embora, como previsto pela marcação negativa no DegPP, essa gramática não permita estruturas

semelhantes a *2 meters tall* (cf. **2 metros alto*), ele permite sintagmas de medida como *2 metros de altura*. Esse fato, bem como outras estruturas de difícil encaixe nas opções teóricas, devem ser mais bem discutidos no futuro, a fim de se verificar o alcance e a adequação empírica de tais parâmetros semânticos.

Por último, no artigo “Mudança semântica e *word embeddings*: estudos de caso na diacronia do português”, Lucas Fonseca Lage e Evandro Paradela Cunha utilizam técnicas computacionais como *word embeddings* para agrupar palavras pelos mesmos traços semânticos. A partir de estudo realizado sobre o Corpus do Português Histórico Tycho Brahe, foram analisadas ocorrências das formas *deus*, *homem*, *mulher*, *pai*, *mãe* e *terra* por tais técnicas, que permitem a representação de palavras como vetores em um espaço vetorial. Primeiramente, foi aplicado o algoritmo Skip-gram para gerar *word embeddings*, e, a seguir, foram produzidas visualizações para a rede de relações semânticas de cada palavra em três diferentes recortes temporais. Tais visualizações geram evidências da organização semântica do léxico. Como base teórica, os autores adotam a visão de Givón (2001), para quem o léxico é um repositório de conceitos relativamente estáveis no tempo, compartilhados socialmente e bem codificados, além de ser organizado em forma de rede, onde conceitos similares estão agrupados próximos uns aos outros, juntamente com a proposta de Matoré (1949), de que as palavras estabelecem relações associativas entre si. A viabilidade de tais métodos de análise de mudança semântica orientados pelos elementos em torno dos quais o léxico se organiza foi atestada pela aplicação dos conceitos de campos nocionais e de palavras-testemunho.

5 Considerações finais

Os artigos sobre temas sintáticos trabalham com riqueza de detalhes, seja numa perspectiva qualitativa ou quantitativa, permitindo compreender particularidades, e, em sua maior parte, tratam do português brasileiro. Os trabalhos de Simões, Duarte e Reis e Madariaga são de natureza comparativa, e permitem perceber com maior clareza as especificidades do português brasileiro face ora ao português europeu, ora a um conjunto de línguas românicas de sujeito nulo (espanhol, italiano e português europeu), ora ao russo. Os trabalhos de Cerqueira, Torres Morais e Salles, Souza-Martins e Gonçalves seguem numa linha

primordialmente diacrônica sobre a evolução do português brasileiro, porém, em alguns momentos, instaurando de maneira mais pontual sejam comparações com o português europeu, seja a atribuição de tais mudanças ao contato sobretudo com as línguas africanas. Já os trabalhos de Teixeira e Borges e Othero e Lazzari apresentam uma abordagem unicamente centrada na identificação da gramática do português brasileiro, se bem que a primeira seja de base formal e a segunda, de base semântica e variacional. Diversamente, os trabalhos de Prim e Moreira, de Namiuti e Silva e de Carneiro, Borges e Aldrete focalizam línguas e variedades diferentes do português brasileiro. Nos dois primeiros casos, trata-se da evolução diacrônica do português europeu, enquanto no último trata-se da Libras, comparando-a com outras línguas de sinais no mundo, no que diz respeito à categoria número.

Os trabalhos oferecem uma variedade de teorias e metodologias que, no seu conjunto, expandem nossa compreensão sobre a gramática do português brasileiro, de maneira específica, e sobre os parâmetros de variação intra- e interlinguísticos, além de indicarem os caminhos da mudança linguística. Esse é um ponto que nos leva a refletir sobre as atuais teorias de mudança linguística, aplicadas em vários trabalhos, os quais demonstram a renovada relevância da relação entre variação e mudança para a compreensão de sistema. Entre as questões que se colocam para verificação em estudos futuros temos: (i) os parâmetros de variação interlinguísticos são pré-configurados em termos universais ou dependem da estrutura de traços de cada língua (Madariaga)?; (ii) critérios de variação intralinguísticos refletem naturalmente os critérios extralinguísticos (Duarte e Reis)?; (iii) como diferenciar critérios de variação estáveis dos que apresentam mudança no tempo (Souza-Martins e Gonçalves)?

No nível de análise semântico, o volume conta com três artigos de abordagem sincrônica e um de abordagem diacrônica. Nota-se uma boa diversidade de linhas teóricas, o que faz deste volume uma rica amostra das perspectivas em voga nos estudos semânticos de hoje. A orientação gerativista (interface sintaxe-semântica) está bem marcada no artigo sobre aquisição de tempo e aspecto verbais por falantes nativos, por Hermont, Cesário de Oliveira e Segantini Brito. A orientação funcionalista se distingue no artigo sobre mudança semântica lexical, por Lage e Paradela Cunha. A semântica formal de ponta transparece tanto na abordagem sobre as comparativas do PB (por Mendes de

Souza, Pires de Oliveira, Frutos e Beviláqua) quanto no tratamento das classes acionais de verbos de editoração (por Martins e Amaral). O primeiro desses dois últimos artigos situa o PB em uma proposta de parâmetros semânticos para os modos de se produzir comparativas em línguas naturais, também aborda duas temáticas muito atuais: por um lado, é examinada a proposta de universais e parâmetros linguísticos que prevejam e expliquem a diversidade semântica das línguas naturais; por outro lado, é examinado o fenômeno da semântica de graus, que tem recebido muita atenção não só pela diversidade de construções relacionadas ao tema encontradas em línguas distintas, mas também pela controvérsia sobre a melhor teoria para tratar tal conjunto de fenômenos, se a delineadora (a dos neokleianos) ou se aquela que defende que há graus na ontologia (cf. KENNEDY, 1997). Considerada a insipiência dos estudos semânticos em diacronia no Brasil até há pouco, sendo esse um campo ainda praticamente em formação, contar com contribuições do naipe destas, e dessa complexidade, nos leva a considerar que a chamada para o volume obteve excelente resposta.

Esperamos que esses sejam apenas os primeiros passos num terreno cada vez mais fértil, e que a sintaxe e a semântica diacrônicas e comparativas se desenvolvam cada vez mais no Brasil. Este volume nos traz motivos para alimentar tais esperanças. Bom proveito aos leitores!

Agradecimentos

Os autores deste artigo agradecem a toda a equipe organizadora do V CILH: Constelações Diacrônicas, a todos os que submeteram manuscritos a este volume e aos editores permanentes da *Revista de Estudos da Linguagem* pela colaboração nesta publicação.

Referências

ANDERSEN, R. W.; SHIRAI, Y. Primacy of Aspect in First and Second Language Acquisition: The pidgin/creole connection. In: RITCHIE, W.; BHATIA, T. (eds.). *Handbook of second language acquisition*. San Diego: Academic Press, 1996. p. 527-570.

BAKER, M. *Lexical categories: verbs, nouns and adjectives*. Cambridge: Cambridge University Press, 2003.

BECK, S.; KRASIKOVA, S.; FLEISHER, D.; GERGEL, R.; HOFSTETTER, S.; SAVELSBERG, C.; VANDERELST, J.; VILLALTA, E. Crosslinguistic variation in comparison constructions. *Linguistic Variation Yearbook*, Amsterdam, v. 9, p. 1-66, 2009. DOI: <https://doi.org/10.1075/livy.9.01bec>

BECK, S.; ODA, T.; SUGISAKI, K. Parametric Variation in the Semantics of Comparison: Japanese vs. English. *Journal of East Asian Linguistics*, Berlin, v. 13, n. 4, p. 289-344, 2004. DOI: <https://doi.org/10.1007/s10831-004-1289-0>

BITTENCOURT, V. O. *Da expressão da causatividade no português do Brasil: uma viagem no túnel do tempo*. 1995. 341f. Tese (Doutorado em Linguística) - Pontifícia Universidade Católica, 1995.

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

CASSEB-GALVÃO, V. C. *O achar no português do Brasil: um caso de gramaticalização*. 1999. 167f. Dissertação (Mestrado em Linguística) - Universidade Estadual de Campinas, 1999.

CASTILHO, A.T. de. *A Gramaticalização*. *Estudos Lingüísticos e Literários*, Salvador, n. 19, p. 25-64, 1997.

CHOMSKY, N. Three Factors in Language Design. *Linguistic Inquiry*, Cambridge, v. 36, n. 1, p. 1-22, 2005. DOI: <https://doi.org/10.1162/0024389052993655>

CHOMSKY, N. *Lectures on Government and Binding*. Dordrecht: Foris, 1981.

CREUS, S; MENUZZI, S. O papel do gênero na alternância entre objeto nulo e pronome pleno em português brasileiro. *Revista da ABRALIN*, Florianópolis, v. 3, n. 1-2, p. 149-176, 2004. Acesso em: <https://revista.abralin.org/index.php/abralin/article/view/931>.

CROFT, W. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. New York: Oxford University Press, 2001.

DUARTE, M. E. L. Sociolinguística “Paramétrica”. In: MOLLICA, M. C.; FERRAREZI Jr., C. (orgs.). *Sociolinguística, sociolinguísticas: uma introdução*. São Paulo: Contexto, 2016. p. 33-44.

ECKARDT, R. *Meaning change in grammaticalization*. An enquiry into semantic reanalysis. Oxford: Oxford University Press, 2006.

FERRARI, L. A gramaticalização de formas não-finitas como evidência da motivação conceptual do léxico. *Veredas*, Juiz de Fora, v. 2, n. 1, p. 103-115, 1998.

FILLMORE, C. J.; KIPARSKY, P. *Universals in linguistic theory*. New York: Holt, Rinehart and Winston, 1968.

GEERAERTS, D. Cognitive approaches to diachronic semantics. In: MAIENBORN, C.; VON HEUSINGER, K; PORTNER, P. (eds.). *Semantics: Typology, Diachrony and Processing*. Berlin: De Gruyter Mouton, 2019. p. 147-176.

GEERAERTS, D. *Diachronic Prototype Semantics. A Contribution to Historical Lexicology*. Oxford: Clarendon Press, 1997.

GIVÓN, T. *Language typology in Africa: a critical review*. *Journal of African Languages and Linguistics*, Berlin, v. 1, n. 2, p. 199-224, 1979.

GIVÓN, T. *Functionalism and grammar*. John Benjamins Publishing, 1995.

GIVÓN, T. *Syntax: an introduction*. v. 1. Amsterdam/Philadelphia: John Benjamins, 2001.

GOLDBERG, A. E. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: Chicago University Press. 1995.

GOLDBERG, A. E. *Constructions at Work: the nature of generalization in language*. Oxford: Oxford University Press, 2006.

GONÇALVES, P. *A Génese do Português de Moçambique*. Lisboa: INCM, 2010.

GONÇALVES, S. C.; LIMA-HERNANDES M. C.; CASSEB-GALVÃO, V. C. (orgs.). *Introdução à Gramaticalização: em homenagem a Maria Luiza Braga*. São Paulo: Parábola, 2007.

HARBERT, W. Contrastive linguistics and language change. *Languages in Contrast*, Amsterdam, v. 12, n. 1, p. 27-46, 2012. DOI: <https://doi.org/10.1075/lic.12.1.03har>

- HEINE, B.; KUTEVA, T. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press, 2002.
- ILARI, R.; BASSO, R. (coord.) *História do Português Brasileiro: História Semântica do Português Brasileiro*. v. 8. São Paulo: Contexto, 2021.
- ILARI; R. NEVES, M. H. M. (orgs.). *Gramática do Português Culto Falado no Brasil: Classes de palavras e processos de construção*. v. 2. Campinas, SP: Editora da Unicamp, 2008.
- KE, P. *Contrastive Linguistics*. Peking/Singapore: Peking University Press/Springer, 2019.
- KENNEDY, C. *Projecting the adjective: The syntax and semantics of gradability and comparison*. 1997. 304f. Thesis (Ph.D., Philosophy in Linguistics) - University of California, 1997.
- KLEIN, E. A semantics for positive and comparative adjectives. *Linguistics & Philosophy*, Berlin, v. 4, n. 1, p. 1-45, 1980.
- KROCH, A. Mudança sintática. *Working Papers em Linguística, Florianópolis*, v. 22, n. 2, p. 23-61, 2021. DOI: <http://doi.org/10.5007/1984-8420.2021.e86882>
- KROCH, A. Reflexes of grammar in patterns of language change. *Language Variation and Change*, Cambridge, v. 1, p. 199-244, 1989.
- LAKOFF, G. *Don't think of an elephant: Know your values and frame the debate*. Chelsea, VT: Green Publishing, 2004.
- LAKOFF, G. *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago press, 2008.
- LANGACKER, R. *Concept, image, and symbol*. The cognitive basis of grammar. Berlin and New York: Mouton de Gruyter. 1990.
- LIGHTFOOT, D. *Born to Parse: how children select their languages*. Cambridge, MA: MIT Press, 2020. <https://doi.org/10.7551/mitpress/12799.001.0001>
- LIGHTFOOT, D. *Principles of Diachronic Syntax*. Cambridge: Cambridge University Press, 1979.

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

MARTINET, A. *A Functional View of Language*. London: Oxford University Press, 1962.

MARTINS, A. M. Introdução: O português numa perspectiva diacrónica e comparativa. In: MARTINS, A. M.; CARRILHO, E. (org.). *Manual de linguística portuguesa*. Berlin: De Gruyter, 2016, p. 1-39.

MATORÉ, G. La lexicologie sociale. *L'Information Littéraire*, Paris, n. 2, mar.-abr. 1949.

MATTOS E SILVA, R. V. Orientações atuais da Linguística Histórica Brasileira. *DELTA, São Paulo*, v. 15, n. esp., p. 147-166, 1999. DOI: <https://doi.org/10.1590/S0102-44501999000300006>

MATTOSE SILVA, R. V. *A gramaticalização numa perspectiva diacrônica: contribuições baianas. Estudos Lingüísticos e Literários, Salvador*, n. 29/30, p. 135-147, 2003. DOI: <https://doi.org/10.9771/2176-4794ell.v0i23-24>

MITHUN, M. Functional Perspectives on Syntactic Change. In: JOSEPH, B. D.; JANDA, R. D. (eds.) *The Handbook of Historical Linguistics*. Malden, MA: Blackwell, 2003. p. 552-572.

MONTAGUE, R. *Formal Philosophy*. Selected Papers of Richard Montague. Ed. por Richmond H. Thomason. New Haven: Yale University Press, 1974.

NEVES, M. H. de M. *Gramática de usos do Português*. São Paulo: Editora UNESP, 2000.

PONTES, E. *O tópico no português do Brasil*. Campinas: Pontes, 1987.

PYLKKÄNEN, L. *Introducing arguments*. Cambridge: MIT Press, 2008.

RIO-TORTO, G. M. Mudança genolexical: teoria e realidade. *Linguística: Revista de Estudos Linguísticos da Universidade do Porto, Porto*, v. 3, p. 223-240, 2008. <https://ler.letras.up.pt/uploads/ficheiros/6883.pdf>

ROBERTS, I. *Parameter Hierarchies and Universal Grammar*. Oxford: Oxford University Press, 2019.

ROBERTS, I.; KATO, M. (orgs.). *Português Brasileiro: uma viagem diacrônica*. Campinas: Editora da Unicamp, 1993.

ROSÁRIO, I. da C.; OLIVEIRA, M. R. de. *Funcionalismo e Abordagem Construcional da Gramática*. Alfa, São Paulo, v. 60, n. 2, p. 233-259, 2016. DOI: <https://doi.org/10.1590/1981-5794-1608-1>

TARALLO, F. Reflexões sobre o conceito de mudança linguística. *Organon, Porto Alegre*, v. 5, n. 18, 1991. DOI: <https://doi.org/10.22456/2238-8915.39119>

TARALLO, F. Por uma Sociolinguística Românica Paramétrica: Fonologia e Sintaxe. *Cadernos de Linguística e Teoria da Literatura, Belo Horizonte*, v. 13, p. 51-83, 1985. DOI: <http://dx.doi.org/10.17851/0101-3548.7.13.53-85>

TRAUGOTT, E. C. The grammaticalization of NP of NP patterns. In: BERGS, A.; DIEWALD, G. (eds.) *Constructions and language change*. Berlin: De Gruyter, 2008, p. 23-46, 2008. DOI: <https://doi.org/10.1515/9783110211757.23>

TRAUGOTT, E. C.; DASHER, R. B. *Regularity in Semantic Change*. Cambridge University Press; Cambridge: 2001.

Formas pronominais



Uma proposta de análise para o objeto acusativo anafórico na variedade de espanhol de Madri e no português brasileiro de São Paulo

An Analysis Proposal for the Anaphoric Accusative Object in the Spanish Variety of Madrid and in the Brazilian Portuguese Variety of São Paulo

Adriana Martins Simões

Universidade Federal de Alfenas (Unifal), Alfenas, Minas Gerais / Brasil

adriana.simoes@unifal-mg.edu.br

<https://orcid.org/0000-0003-2911-8873>

Resumo: Neste trabalho, nosso objetivo é apresentar os resultados de nossas pesquisas sobre o objeto acusativo anafórico na variedade de espanhol de Madri e na variedade do português brasileiro de São Paulo, assim como sugerir uma análise teórica para essas construções a partir dos desenvolvimentos minimalistas (CHOMSKY, 1999, 2000, 2004). Analisamos entrevistas da variedade de espanhol de Madri (CESTERO MANCERA *et al.* 2014) e do português brasileiro (LIMA-HERNANDES; VICENTE, 2012). Como referencial teórico, aliamos as perspectivas biológica (CHOMSKY, 1981, 1986) e social de língua (LABOV, 2008; WEINREICH; LABOV; HERZOG, 2009). Observamos a atuação da referencialidade do antecedente, do traço de indefinidade do determinante, da ausência de determinante, da estrutura verbal e do aspecto lexical na possibilidade de elipses na variedade de espanhol de Madri. Em relação à variedade do português brasileiro de São Paulo, observamos objetos nulos quase categóricos e que os pronomes lexicais apresentam restrições na expressão de antecedentes menos referenciais. Esses resultados se assemelham aos de nossa pesquisa anterior (SIMÕES, 2015) e às variedades de espanhol do País Basco (LANDA, 1993) e de Quito (SUÑER; YÉPEZ, 1988).

Palavras-chave: espanhol; português brasileiro; objeto acusativo anafórico, variação linguística; concordância de objeto.

Abstract: In this paper, our aim is to present the results of our researches about the anaphoric accusative object in the Spanish variety of Madrid and in the Brazilian Portuguese variety of São Paulo and to propose a theoretical analysis to these constructions based on minimalist developments (CHOMSKY, 1999, 2000, 2004). Interviews of the Spanish variety of Madrid (CESTERO MANCERA *et al.* 2014) and Brazilian Portuguese variety of São Paulo (LIMA-HERNANDES; VICENTE, 2012) were analysed. Regarding the theoretical approach, the biological (CHOMSKY, 1981, 1986) and social (LABOV, 2008; WEINREICH; LABOV; HERZOG, 2009) perspectives on language were combined. We observed the effect of the antecedent referentiality, of the determinant indefinite features, of the bare nouns, of the structure and lexical aspect of verbs on possibility of elipsis in the Spanish variety of Madrid. Concerning the Brazilian Portuguese variety of São Paulo, the null objects were almost categorical and the lexical pronoun presented restrictions to less referential antecedents. These results resemble the tendencies of our previous research (SIMÕES, 2015) and the Spanish varieties of País Vasco (LANDA, 1993) and Quito (SUÑER; YÉPEZ, 1988).

Keywords: Spanish; Brazilian Portuguese; anaphoric accusative object; linguistic variation; object agreement.

Recebido em 02 de fevereiro de 2022

Aceito em 15 de março de 2022

1 Introdução

O objetivo deste trabalho é apresentar e discutir os resultados de nossas últimas pesquisas a respeito da expressão do objeto anafórico acusativo de 3ª pessoa na variedade de espanhol de Madri e na variedade do português brasileiro (doravante PB) de São Paulo, bem como sugerir uma análise para essas construções tendo em vista os desenvolvimentos minimalistas (CHOMSKY, 1999, 2000, 2004) e outros trabalhos gerativistas (EZEIZABARRENA, 1997; FUGINO; SANO, 2002; GALVES, 2001; GROPP, 1997; THRÁINSSON, 2008). Para a realização desses estudos, analisamos entrevistas orais da variedade de espanhol de Madri, pertencentes ao *Proyecto para el Estudio Sociolingüístico del Español de España y de América* (PRESEEA) (CESTERO MANCERA *et al.* 2014), cujos informantes apresentam escolaridade equivalente ao Ensino Médio, e entrevistas orais da variedade de PB de São Paulo,

pertencentes ao *Projeto História do Português Paulista (PHPP)* (LIMA-HERNANDES; VICENTE, 2012), cujos informantes apresentam Ensino Superior. Esses novos estudos dão continuidade à nossa pesquisa de doutorado¹ (SIMÕES, 2015), na qual investigamos a expressão do objeto acusativo nas variedades de espanhol de Madri e Montevidéu e realizamos uma comparação qualitativa com o PB. Essas novas pesquisas se justificam pela necessidade de avançarmos na descrição dessa área da gramática do espanhol e do PB, e detectar, no âmbito do espanhol, as possibilidades de elipse e, no âmbito do PB, os contextos em que o pronome é imprescindível para expressar o antecedente, bem como os contextos em que este não seria aceito. Essas pesquisas nos permitirão avançar na compreensão dos mecanismos envolvidos na expressão dos argumentos verbais por pronome ou elipse nas línguas naturais e serão nossa base para a análise do processo de aquisição/aprendizagem do espanhol como língua estrangeira.

Como perspectiva teórica, além da concepção biológica de língua (CHOMSKY, 1981, 1986), apoiamo-nos também em alguns aspectos da sociolinguística (LABOV, 2008; WEINREICH; LABOV; HERZOG, 2009). Nossas hipóteses são de que, na variedade de espanhol de Madri, as elipses de objeto seriam favorecidas por antecedentes com o traço semântico de indefinitude, os [-animados] e [-específicos], enquanto no PB a elipse seria favorecida por antecedentes com os traços semânticos de indefinitude, definitude, os [-animados] e [+/-específicos] e haveria restrições ou impossibilidade de realização do pronome lexical com antecedentes [-animados; -específicos] indefinidos ou quantificados. Veremos que, em ambas as línguas investigadas, as hipóteses foram parcialmente confirmadas. Os resultados revelam a atuação da referencialidade do antecedente, do traço semântico de indefinitude do determinante, da ausência de determinante, da estrutura verbal e do aspecto lexical na possibilidade de elipses de objeto na variedade de espanhol de Madri. Na esfera da variedade do PB de São Paulo, observamos que os objetos nulos foram quase categóricos e que os pronomes lexicais

¹ Naquele momento, a comparação com o PB se deu por meio da tradução a essa língua de alguns dados de objetos nulos encontrados nas variedades de espanhol de Madri e Montevidéu, configurando-se como uma análise qualitativa. Em nossas atuais pesquisas, temos a oportunidade de analisar entrevistas orais também do PB e realizar uma análise quantitativa de ambas as línguas.

apresentam restrições para retomar antecedentes menos referenciais. Os resultados tanto da variedade de espanhol de Madri quanto do PB se assemelham aos encontrados em nossa pesquisa anterior (SIMÕES, 2015), bem como se assemelham ao funcionamento do objeto nulo nas variedades de espanhol faladas no País Basco (LANDA, 1993, 1995) e em Quito (SUÑER; YÉPEZ, 1988).

Este artigo se estrutura da seguinte forma: na primeira seção, apresentamos os trabalhos sobre a expressão anafórica do objeto acusativo no espanhol e no PB. Na segunda seção, abordamos as teorias a partir das quais analisamos as entrevistas e interpretamos os resultados. A terceira seção será dedicada à metodologia. Na quarta seção, apresentamos e discutimos os resultados encontrados. Na quinta seção, apresentamos nossa proposta de análise teórica para a expressão do objeto acusativo. Concluímos o artigo com as considerações finais.

2 As gramáticas do espanhol e do português brasileiro na expressão do objeto anafórico acusativo

Conforme Campos (1986) e Fernández Soriano (1999), no espanhol, a elipse do objeto em função acusativa poderia ocorrer com antecedentes [-específicos; -definidos]. Por outro lado, quando o antecedente é [+específico], seria necessária a presença do clítico retomando o antecedente. Segundo Groppi (1997), essa tendência também se manifesta na variedade de espanhol de Montevidéu. Observem-se as sentenças a seguir:

(1a) — ¿Compraste flores_i?

— Sí, compré Ø_i.

— Sí, *las_i compré.

(1b) — ¿Compraste las flores_i?

— Sí, compré *Ø_i.

— Sí, las_i compré. [Exemplos extraídos e adaptados de Campos (1999, p. 1530)]

De acordo com Groppi (2009), no espanhol, apenas o clítico pode retomar um antecedente em função acusativa. Sendo assim, sua retomada pelo pronome tônico seria agramatical, como em (2a). Nessa língua, para que o pronome tônico possa aparecer em função acusativa, é necessário que seja correferente com um clítico, que haja a necessidade

discursiva de estabelecer contraste entre dois ou mais referentes, que o antecedente seja [+humano] e o pronome deve estar introduzido pela preposição *a*, como em (2b).

(2a) *Veo a ella.

(2b) La_i veo [a ella_i]. (GROPPI, 2009, p. 100)

Quando se trata de um sintagma nominal² (doravante SN) introduzido pelo artigo indefinido, conforme Leonetti (1999), esse tipo de SN pode ser retomado por um pronome definido, como em (3b). Nessa construção, o clítico retoma o SN *un caso de corrupción*. A tendência é que os SNs encabeçados pelo artigo indefinido recebam uma interpretação [-específica] devido ao traço semântico de indefinidade desse artigo.

(3a) Han denunciado un caso de corrupción_i en el juzgado nº 3.

(3b) Parece que lo_i ha descubierto un periodista. (LEONETTI, 1999, p. 838).

Quanto aos SNs quantificados, segundo Campos (1986), para sua expressão anafórica, seria necessário um quantificador, já que sua expressão pela elipse seria agramatical, como em (4).

(4) — ¿Compraste algunos regalos_i?

(4a) — *Sí, compré Ø_i.

(4b) — Sí, compré algunos. (CAMPOS, 1986, p. 354)

Em relação às construções nas quais o antecedente aparece como um tópico, conforme Groppi (2009), é preciso a presença de um clítico correferente no interior da oração. Na análise de Groppi (2009), o clítico corresponderia ao argumento do verbo. Em (5b), o antecedente é [+específico] e [+humano] e a ausência do clítico tornaria a construção agramatical.

² Embora trabalhos mais recentes optem por utilizar o termo sintagma determinante (*Determiner Phrase*), considerando o determinante como o núcleo e o nome como o complemento desse sintagma, em nosso trabalho, optamos pela designação clássica de sintagma nominal (*Noun Phrase*), no qual o nome seria o núcleo e o determinante ocuparia o especificador deste.

(5a) A Juan_i (,) lo_i vi en la playa ayer.

(5b)*A Juan vi ayer en la playa. (GROPPI, 2009, p. 110)

Com respeito aos SNs indefinidos que aparecem como tópico, segundo Leonetti (1999), sua retomada por um clítico é possível nos casos em que sua interpretação é genérica ou [+específica], como vemos em (6a) e (6b), respectivamente.

(6a) Un cumpleaños_i, es mejor celebrarlo_i fuera de casa.

(6b) A un amigo mío_i, este profesor le_i ha suspendido ya tres veces. (LEONETTI, 1999, p. 855)

Quanto aos SNs quantificados em posição de tópico, conforme Groppi (1997), poderia ou não aparecer um clítico correferente no interior da oração. Sua presença conduziria a uma leitura referencial, na qual o SN quantificado receberia uma interpretação partitiva, como em (7b).

(7a) Algunas tarjetas_i, yo también recibí Ø_i.

(7b) Algunas tarjetas_i, yo también las_i recibí. (GROPPI, 1997, p. 124)

Apesar de o espanhol ser uma língua na qual os objetos nulos estariam altamente restringidos, em algumas variedades é possível a ocorrência de omissão do objeto em contextos mais amplos, como as variedades de espanhol faladas no País Basco (LANDA, 1993, 1995) e em Quito (SUÑER; YÉPEZ, 1988).

Conforme Landa (1993, 1995), no espanhol falado no País Basco, a elipse ocorreria com antecedentes [-determinados] e [+determinados; +/-definidos], como se observa em (8). Nessa variedade, os objetos nulos aparecem em construções ditransitivas [exemplo (9)], com tópico [exemplo (10)], com predicação secundária [exemplo (11)]. Além disso, essa variedade não apresentaria restrição aspectual, de modo que a elipse ocorre em construções com aspecto [+/-perfectivo] [exemplo (12)].

(8a) — ¿Compraste el regalo_i?

— Sí, lo_i=compré e_i.

— Sí, Ø_i=compré e_i.

(8b) — ¿Compraste algunos regalos_i?

— Sí, Ø_i=compré e_i. (LANDA, 1993, p. 132-133)

(9) También tengo las fotos_i del bote de J., pero están muy desenfocadas, así que no os= \emptyset_i =mando e_i. Los padres de J. quieren que les= \emptyset_i =mandemos e_i, aunque estén desenfocadas, así que me imagino que J. les= \emptyset_i =mandará e_i.

(10) La boda_i me= \emptyset_i =pagó e_i éste de la Campa de Erandio.

(11) ¿Tienes bomba de bici?

Sí, una chiquita.

Pues cuando \emptyset_i =tenga e_i bajas_i te la pido y ya está. (i = ruedas)

(12) ¿Quién ha puesto ese disco?

Juan puso e_i. (LANDA, 1995, p. 101)

No que se refere à variedade de espanhol de Quito, segundo Suñer e Yépez (1988), os objetos nulos seriam possíveis com antecedente [+definido] em construções de tópico [exemplo (13)], nos casos em que o referente aparece na oração anterior [exemplo (14)], em construções com clítico dativo [exemplo (15)]. Nessa variedade, a elipse também ocorreria com antecedente [+humano], como em (16).

(13) Las elecciones_i yo nunca entendí \emptyset_i .

(14) A mi mamá se le quedó un poco mal cerrado el armario_i y logré abrir \emptyset_i .

(15) Bueno, yo te \emptyset_i saco.

(lo = el vestido_i)

(16a) Presentáme \emptyset .

(16b) Ya te \emptyset voy a presentar. (SUÑER; YÉPEZ, 1988, p. 512-514)

Em nossa pesquisa de doutorado (SIMÕES, 2015), na qual analisamos entrevistas orais das variedades de espanhol de Madri e Montevideú, encontramos 4,1% e 11,1% de objetos nulos nessas variedades, respectivamente. Observamos que os objetos nulos foram favorecidos pelos SNs sem determinante, os quantificados e os introduzidos pelo artigo indefinido, os [-animados] e, na variedade de Montevideú, também os SNs [-específicos].

Temos, nas construções a seguir, alguns casos de objetos nulos extraídos das amostras analisadas. Em (17), as elipses apresentam como antecedente SNs sem determinante. Nas construções em (18), os SNs antecedentes são quantificados e, em (19), introduzidos pelo artigo indefinido. Em todas as construções os antecedentes são [-animados] e em (17a-b), (18b), (20a-b) e (21a-b) são também [-específicos].

(17a) E: y los estudios ¿seguirías con la idea de hacer **oposiciones** o no?

I: yo creo que sí porque yo creo que en mi casa aburrida / o a lo mejor no haría Ø / yo realmente hago oposiciones para tener un trabajo seguro / (...) (Entrevista 5 – Madri)

(17b) E: y en el jardín ¿tenés **plantas**?

I: sí / en el fondo tenemos Ø sí / ahora / (...) (Entrevista 17 – Montevideú)

(18a) I: (...) bueno llevaba en el bolsillo **dos mi<alargamiento/>1 y algo** / hh si le llevo a dar Ø a mi hijo / pues el tío sale frustrado del todo ¿comprendes? (...) (Entrevista 16 – Madri)

(18b) E: ¿compraste **alguna rifa**?

I: no / mamá compra Ø en la de arquitectura <ruido = “ladrido”/> (Entrevista 20 – Montevideú)

(19a) I: y<alargamiento/> no sé / bueno / ee / yo hace tiempo lo pensaba y un **una mercería que hay un poquito más abajo<alargamiento/>** la traspasaron menos mal / porque iban cerrando Ø (...) (Entrevista 4 – Madri)

(19b) I: nunca he llegado al <risas = “todos”/> / este<alargamiento/> / cuando llegué a los / a **unos cubiertos** creo que tenía Ø Devoto / (...) (Entrevista 1 – Montevideú)

(20a) I: (...) en el arroz por ejemplo sí / en **el arroz** / pues primero rehogo Ø con aceite y cebolla (Entrevista 8 – Madri)

(20b) I: <entre_risas> me tenía que tomar **el ómnibus** </entre_risas>

E: claro

I: y tomaba Ø / salía de casa seis menos cuarto (...) (Entrevista 17 – Montevideú)

(21a) I: (...) aparte no me complico en **el cocido** la<alargamiento/>rgo Ø en plan de Lardi ni<alargamiento/> ni pollo ni gallina (...) (Entrevista 12 – Madri)

(21b) I: eran más definidas **las estaciones** ¿no? absolutamente / es más eh uno asociaba Ø a los juegos / la cometa por ejemplo / (...) (Entrevista 13 – Montevideú)

Observamos também que as variedades de espanhol de Madri e Montevideú apresentaram a elipse em construções com aspecto [-perfectivo], perífrase verbal, clítico dativo, predicação secundária,

tópico e verbos cognitivos. Embora essas construções não tenham sido selecionadas como relevantes para a ocorrência dos objetos nulos, constituem contextos que favorecem a elipse nas variedades de espanhol do País Basco (LANDA, 1993, 1995), no de Quito (SUÑER; YÉPEZ, 1988) conforme vimos, assim como no PB (DUARTE, 1986). Observem-se alguns dados:

(22a) I: los más modestos del barrio somos los que procedemos / hh de esa<alargamiento/> etapa / que nos adjudicaron / los **los pisos** que<alargamiento/> nos tocaron / no<alargamiento/> había opción de decir <cita> yo quiero Ø en esta calle en esta altura </cita> no // te te lo daba<alargamiento/>n por los hijos que tenías (Entrevista 18 – Madri)

(22b) I: (...) no sé de qué será **esa sal** // a mí me han regalado Ø tengo ahí un poco pero bueno / (...) (Entrevista 13 – Madri)

(22c) I: sí / una plazoleta chiquitita / estee / 21 de Setiembre / se engancha con Bulevar España por ahí / **a una de ellas** / violaron Ø / eran las seis de la mañana (Entrevista 9 – Montevideú)

(22d) I: (...) Belén ahora por ejemplo estudia en <vacilación/> en **unos libros** no donde si la tarea es sintetizar la información / eeh <vacilación/> nos ahogamos porque no hay lo que es sintetizar porque están previstos para que el niño <ênfasis/> ya </ênfasis/> tenga resumido Ø (Entrevista 11 – Montevideú)

(22e) I: (...) mi madre así <vacilación/> no se compraba **un helado** en la heladería <cita> porque Miguel no / porque / nos ayuda / y tú sabes que no / yo te hago Ø en casa de lo que tú quieras / de chocolate / de crema / pero **un helado de heladería** no te puedo comprar Ø </cita> (...) (Entrevista 14 – Montevideú)

Considerando-se os estudos de Campos (1986), Fernández Soriano (1999), Groppi (1997, 2009) e Leonetti (1999), observa-se que no espanhol é necessário que um SN [+determinado]³ em função acusativa se expresse pelo clítico, seja esse SN [+/-específico]. Entretanto, quando o antecedente constitui um SN sem determinante, a elipse seria possível. Esses trabalhos demonstram que os objetos nulos seriam altamente restringidos no espanhol.

Por outro lado, vemos que, nas variedades de espanhol do País Basco (LANDA, 1993, 1995) e de Quito (SUÑER; YÉPEZ, 1988), que estão em contato com línguas não indoeuropeias, é possível a ocorrência da elipse com antecedentes [+determinados] e em construções como as com clítico dativo e tópico.

³ Classificamos como [+determinado] um SN introduzido por algum determinante, seja este definido, indefinido ou quantificado.

Nosso trabalho (SIMÕES, 2015) revelou também algumas ocorrências de objetos nulos em contextos mais amplos nas variedades de espanhol de Madri e Montevidéu, que formariam parte das variedades nas quais as elipses apresentariam maiores restrições.

Em relação ao PB, conforme demonstraram diferentes pesquisas (CYRINO, 1993; DUARTE, 1986; GALVES, 2001; TARALLO, 1993), essa língua vem passando por variação e mudança no paradigma pronominal desde o século XIX. Esse processo de variação e mudança linguísticas se reflete na perda do clítico acusativo de 3ª pessoa (*o, a*), na ampliação das elipses de objeto e no emprego do pronome lexical (*ele/ela*) em função acusativa.

Segundo Galves (2001), a mudança no sistema pronominal teria ocorrido devido ao enfraquecimento da concordância verbal, desencadeado pela inserção do pronome *você* e pela perda do pronome *tu*, que mantinha a distinção semântica entre a 2ª e 3ª pessoa do discurso.

De acordo com os dados diacrônicos de Tarallo (1983, *apud* 1993), começou a ocorrer no PB uma diminuição na frequência da expressão pronominal do objeto acusativo a partir do século XIX. Observe-se a tabela 1:

Tabela 1 - Preenchimento do objeto direto pronominal

Data % de preenchimento do objeto direto	
1725	89,2%
1775	96,2%
1825	83,7%
1875	60,2%
1982	18,2%

Fonte: extraída e adaptada de Tarallo (1983, *apud* 1993, p. 84)

No PB contemporâneo, o clítico *o* já não se incorpora à gramática por meio de aquisição natural (MAGALHÃES, 2006), mas apenas pela aprendizagem formal durante o processo de escolarização, na medida em que se restringe à produção de falantes escolarizados (DUARTE, 1986). Nas propostas de Galves (2001) e Kato (2005), esse clítico integraria a periferia marcada. Portanto, no PB, apenas o objeto nulo e o pronome lexical corresponderiam à gramática adquirida naturalmente para expressar o objeto acusativo.

Em sua pesquisa, Cyrino (1990, *apud* CYRINO, 1993) analisou documentos de diferentes séculos e observou a aumento na frequência da elipse na 1ª metade do século XIX e o uso do pronome lexical em função acusativa na 2ª metade do século XIX, período no qual teve início uma diminuição na frequência do clítico.

Em um estudo posterior, Cyrino (1994) observa a possibilidade de elipses no PB com antecedentes oracionais e SNs [+/-específicos]. Seus resultados revelaram que as elipses sempre ocorreram no PB, mas, em um determinado momento, começaram a ter maior incidência. Houve um aumento das elipses com antecedentes [+específicos] ao longo do tempo, que se acentuou no século XX, período em que também se acentuou a frequência dos objetos nulos com antecedentes [-específicos], sendo esses os casos que apresentam mais incidência de apagamento.

Cyrino (1994) verificou que no PB do século XX são os antecedentes [-animados; -específicos] os que mais favoreceram a elipse, uma vez que a omissão foi de 93,1%. Em relação aos antecedentes [-animados; +específicos], estes tiveram também alta incidência, com 86,5% de objetos nulos. Entretanto, essa autora não observou nenhuma elipse com antecedentes [+animados; +específicos]. Quanto aos antecedentes [+animados; -específicos], a ocorrência de elipse foi de 57,1%.

Apesar de Cyrino (1994) não ter encontrado objetos nulos com antecedentes [+animados; +específicos], segundo Kato (2003), seria possível a elipse com esse tipo de antecedente quando o mesmo aparece como um tópico, como em (23).

(23)Esse ator, eu acho que (eu) não conheci Ø_i. (KATO, 2003, p. 139)

Duarte (1986) constatou 23,7% de objetos nulos com antecedente [+animado], como em (24a), e 76,3% com antecedente [-animado], como em (24b). Tanto o resultado de Duarte (1986) quanto o de Cyrino (1994) demonstram que as elipses no PB seriam favorecidas por entidades [-animadas].

(24a) A FEBEM é um dos elos dessa corrente que cria o menor infrator; não é ela o único responsável, o único elo que cria Ø_i, e como tal ela não consegue recuperar Ø_i. (entrevista) (46,47)

(24b) O Armando leu a peçai e aprovou Ø_i. (novela) (123) (DUARTE, 1986, p. 16)

O estudo de Casagrande (2012) revelou que a ocorrência do pronome lexical estaria relacionada com o fator especificidade do antecedente, já que apenas os antecedentes [+específicos] permitiram a retomada pelo pronome. Essa pesquisadora constatou, a partir de testes de julgamento de gramaticalidade aplicados a falantes adultos do PB, que os antecedentes [+/- animados; -específicos] não permitem a retomada pelo pronome lexical, seja o verbo perfectivo ou imperfeito, como se observa em (25). Sendo assim, os antecedentes com esses traços semânticos se expressariam pela elipse. Contudo, com antecedentes [-animados; +específicos] seria possível a expressão pela elipse ou pelo pronome lexical, como em (26), sendo que a perfectividade do verbo favorece o pronome. Com relação aos antecedentes [+animados; +específicos], estes permitem apenas a expressão pelo pronome lexical, com verbo perfectivo ou imperfeito, como se observa em (27).

(25a) Minha tia faz bolo_i e eu como Ø_i/*ele_i todo dia.

(25b) Minha tia tem filho_i e educa Ø_i/*ele_i com carinho. [Exemplos extraídos e adaptados de Casagrande (2010, p. 112-130)]

(26a) Eu tenho uma bicicleta_i e o meu pai sempre lava Ø_i/ela_i.

(26b) Eu fiz o suco de laranja_i e o meu pai tomou Ø_i/ele_i. [Exemplos extraídos e adaptados de Casagrande (2010, p. 120-137)]

(27a) Meu tio sempre chama o cachorro_i lá de casa e molha *Ø_i/ele_i.

(27b) Eu achei o meu cachorro_i e abracei *Ø_i/ele_i. [Exemplos extraídos e adaptados de Casagrande (2010, p. 124-139)]

No que se refere ao pronome lexical, de acordo com Kato (2002), no PB, esse pronome teria perdido a restrição a antecedente [+humano] e, portanto, poderia retomar antecedentes [-humanos]. Também seria considerado um pronome fraco homófono aos pronomes fortes ELE/ELA.

Galves (2001) apresenta uma ocorrência do pronome lexical extraída do projeto NURC, na qual o SN *o carro*, entidade [-humana], é retomado pelo pronome lexical, como se observa em (28). De acordo com Galves (2001), esse pronome no PB corresponderia aos clíticos nas outras línguas românicas. No entanto, enquanto é possível retomar uma oração pelo pronome clítico, o pronome lexical permite apenas a retomada de antecedentes referenciais, como se observa em (29).

(28) Se tiver muita pressa, eu largo ele num lugar proibido mesmo (SP). (GALVES, 2001, p. 163)

(29a) Je le sais.

(29b) Eu sei (*ele). (GALVES, 2001, p. 173)

Em Simões (2015) realizamos uma comparação qualitativa entre as variedades de espanhol de Madri e Montevideu e o PB a fim de verificar as possibilidades de expressão do objeto acusativo anafórico nesta última língua. Nossa estratégia consistiu em verificar, a partir dos dados de objetos nulos encontrados nas variedades de espanhol investigadas, se o PB permitia a variação entre a elipse e o pronome lexical.

Observamos que seria possível a elipse ou o pronome lexical no âmbito dos antecedentes [+animados; +específicos], com SN definido ou indefinido, como em (30). Vimos que essa possibilidade se mantém entre os antecedentes [+animados; -específicos], com SN definido ou quantificado, como se observa em (31).

(30a) (...) este_i espera que eu te mostro Ø_i/ele_i (...) (referente: o filho da informante) (cf. entrevista 16 – Madri)

(30b) uma amiga_i minha eu chamei Ø_i/ela_i quando ela estava vendo TV para dedicar um vídeo para ela (...) (cf. entrevista 5 – Madri)

(31a) porque a polícia_i nós não chamamos Ø_i/ela_i porque eles não tinham entrado... não tinham roubado nada (...) (cf. entrevista 18 – Montevideu)

(31b) A: e você considera que tem alguns amigos_i?

B: sim... E é uma

A: ah... sim

B: está bom então

A: por sorte eu encontrei Ø_i/eles_i/alguns_i... viu? (...) (cf. entrevista 2 – Montevideu)

Com respeito aos antecedentes [-animados; +específicos], a variação entre a elipse e o pronome lexical se mantém entre os SNs definidos e indefinidos. Por outro lado, com um SN quantificado, apenas a elipse seria possível. Observem-se os exemplos em (32):

(32a) (...) eu não sei do que é esse sa_i... me deram Ø_i/ele_i... eu tenho um pouco aí (...) (cf. entrevista 13 – Madri)

(32b) com uma barrinha de ferro que introduzia em uma tábua_i... eu tenho Ø_i/ela_i lá naquele quarto... depois eu te mostro Ø_i/ela_i (...) (cf. entrevista 12 – Montevidéu)

(32c) (...) eu levava no bolso dois mil e alguma coisa_i... se eu chego a dar Ø_i/*ele_i para o meu filho... o cara sai frustrado de tudo... entende? (...) (cf. entrevista 16 – Madri)

Em relação aos antecedentes [-animados; -específicos] e definidos, a variação entre a elipse e o pronome lexical é possível nas construções com verbos dinâmicos, como em (33a-b). No entanto, se a construção apresenta um verbo de comunicação, esta não seria bem formada com o pronome, como vemos em (33c). Entre as construções com verbos estativos, como *querer*, a presença do pronome conduz a uma interpretação [+específica] do antecedente, como se observa em (33d).

(33a) (...) no arroz... por exemplo... no arroz_i... primeiro eu refogo Ø_i/ele_i com óleo e cebola (...) (cf. entrevista 8 – Madri)

(33b) as estações_i eram mais definidas... né? absolutamente... as pessoas associavam Ø_i/elas, aos jogos (...) (cf. entrevista 13 – Montevidéu)

(33c) A: esses dados_i interessam pra vocês?

B: sim

A: porque talvez é como uma novela

B: nos interessam e a gente adora

A: pois então eu te conto Ø_i?/eles_i (cf. entrevista 18 – Madri)

(33d) (...) recebemos os apartamentos_i e não era possível dizer “eu quero Ø_i/ele_i nesta rua nesta altura” (...) (cf. entrevista 18 – Madri)

Na esfera dos antecedentes [-animados; -específicos] e indefinidos, apenas a elipse é aceita na maior parte das construções. Em (34a), construção que apresenta o verbo intensional *pedir*, o pronome lexical conduz a uma interpretação [+específica] do antecedente. Em relação a (34b), o pronome é aceito com o verbo dinâmico *fazer*, mas não com *comprar*, construção em que o antecedente desempenha a função discursiva de tópico. No âmbito do verbo estativo *ter*, o pronome também não é aceito, como se observa em (34c). Por fim, entre os SNs quantificados, apenas em (34e) foi possível o pronome.

- (34a) A: (...) como eu peço uma bolsa_i... um horário... documentos... e vou
B: e vai pra fora
A: estou ocupada com A
B: o único que você perde... mas... você perde a bolsa ou a ajuda de...
A: não... mas eu não posso pedir Ø_i/?ela_i... eu é que não me atrevo (...) (cf. entrevista 18 – Madri)
- (34b) (...) minha mãe não comprava um sorvete_i na sorveteria... “porque Miguel nos ajuda e você sabe que não... eu te faço Ø_i/ele_i em casa do que você quiser... de chocolate... de creme... mas um sorvete de sorveteria_i eu não posso te comprar Ø_i/?ele_i” (...) (cf. entrevista 14 – Montevidéu)
- (34c) A: e com ela você tem uma boa relação?
B: como eu tinha Ø_i/?ela_i com o meu irmão... como eu tenho Ø/?ela com a minha cunhada (...) (cf. entrevista 8 – Montevidéu)
- (34d) (...) depois que está tudo_i frito... eu misturo Ø_i/?ele_i... agrego Ø/?ele à massa (...) (cf. entrevista 3 – Madri)
- (34e) A: (...) pois eu tenho muitos exames_i muitos atestados de incapacidade_i que refazer é muito difícil
B: é verdade... é uma confusão... vai muito tempo
A: e só eu poderia fazer Ø_i/eles_i/ porque meus filhos não fizeram Ø/eles (cf. entrevista 18 – Madri)
- (34f) A: e você comprou alguma rifa?
B: não... mamãe compra Ø_i/?ela_i/alguma_i na de arquitetura (cf. entrevista 20 – Montevidéu)

As tendências observadas em nosso estudo (SIMÕES, 2015) nos permitiram concluir que, no PB, a possibilidade de expressão do objeto pelo pronome lexical se relaciona com os antecedentes [+animados] e [+específicos]. Contudo, na esfera dos antecedentes [-animados], apenas os [+específicos] definidos ou indefinidos parecem permitir o pronome em todas as construções. Ademais, ocorrem restrições que estão relacionadas à estrutura do SN antecedente e à semântica do predicado verbal.

Considerando-se as pesquisas sobre o PB, observamos que essa língua permite objetos nulos com antecedentes [+/-específicos], sobretudo com os [-animados]. Por outro lado, a possibilidade de expressão pelo pronome lexical se relaciona com os SNs [+específicos] e a perfectividade do verbo.

3 Referencial teórico

Em nossas pesquisas, aliamos a concepção biológica de língua (CHOMSKY, 1981, 1986) à concepção social (LABOV, 2008; WEINREICH; LABOV; HERZOG, 2009), de modo que, além de considerar gramática como um conhecimento internalizado na mente/cérebro, consideramos também alguns aspectos da sociolinguística.

Apesar de essas perspectivas teóricas apresentarem divergência quanto ao lugar das regras variáveis no sistema linguístico, se mostram complementares. A combinação dessas duas teorias foi denominada *Sociolinguística Paramétrica* e iniciou-se nos anos 80 por linguistas brasileiros. Kato e Tarallo (1989) realizaram um trabalho pioneiro que revela a importância de aliar as perspectivas intralinguística e interlinguística, na medida em que a primeira fornece subsídios para que se possa compreender e desenvolver melhor a segunda.

4 A perspectiva biológica de língua

De acordo com Chomsky (1981, 1986), os seres humanos teriam uma capacidade linguística inata, denominada *Faculdade da Linguagem*, que se encontra na mente/cérebro. Esse dispositivo seria o responsável pelo desenvolvimento do conhecimento linguístico. A aquisição da linguagem ocorre a partir da fixação de parâmetros. Esse processo ocorre mediante a interação entre os dados linguísticos do ambiente e a Gramática Universal, que abriga os princípios universais das línguas naturais.

O conhecimento linguístico compreende a *língua-I*, que se caracteriza por ser internalizada, intensional e individual e opor-se à *língua-E*, que se caracteriza por ser externa e extensional. Além disso, o conhecimento linguístico constitui a *competência* do falante, que se diferencia do *desempenho linguístico*, correspondente à língua em uso.

Após a fixação de parâmetros de uma língua, o falante teria uma *gramática nuclear*, originada do processo natural de aquisição da linguagem. Por outro lado, a gramática nuclear se expandiria ao longo da vida e surgiria a *periferia marcada*, que abriga os resíduos de mudança linguística e outros fenômenos.

Em nossa pesquisa, consideramos que, na variedade de Madri, tanto o clítico como o objeto nulo corresponderiam à língua-I. Sendo assim, ambas as formas de expressão do objeto integram o conhecimento linguístico internalizado e adquirido naturalmente pelos falantes nativos.

Em relação ao PB, o clítico já não constitui aquisição natural (GALVES, 2001; MAGALHÃES, 2006) e seria incorporado à gramática da língua apenas mediante o processo de escolarização (DUARTE, 1986) como vimos. Segundo Galves (2001) e Kato (2005), esse clítico integraria a periferia marcada. Portanto, no PB, apenas o objeto nulo e o pronome lexical integram a gramática adquirida naturalmente para a expressão do objeto acusativo.

Conforme Lightfoot (2006), a língua-E reflete o *output* das gramáticas das comunidades linguísticas e o uso da língua no discurso e na variação social. Em nosso trabalho, as entrevistas analisadas corresponderiam à língua-E. Entretanto, as ocorrências de clítico e objeto nulo no espanhol e de pronome lexical e objeto nulo no PB corresponderiam à língua-I enquanto conhecimento linguístico internalizado. Portanto, a pesar de as entrevistas representarem a língua em uso e constituírem tanto o conhecimento linguístico internalizado quanto o desempenho linguístico do falante, o objeto de análise de nossa pesquisa compreende a competência linguística.

No que se refere à arquitetura da linguagem, de acordo com os desenvolvimentos minimalistas (CHOMSKY, 2000), uma língua se compõe de um léxico e um sistema computacional e gera expressões linguísticas que são interpretadas nos sistemas de performance. O léxico é formado por um conjunto de traços fonéticos, semânticos e sintáticos, que podem ser interpretáveis ou não, e apenas os primeiros são legíveis nos sistemas externos. Em relação ao sistema computacional, nele atuam as operações de concatenação, concordância e movimento. A partir da primeira operação são formados os objetos sintáticos. Quanto às outras duas, estas seriam responsáveis pela eliminação dos traços [-interpretáveis].

Para formar uma construção com objeto acusativo, o objeto se concatena com o predicado verbal que o seleciona. O objeto, que constitui o argumento interno desse predicado, satisfaz suas propriedades semânticas e sintáticas e recebe seu papel temático. Os traços- ϕ , que constituem aos traços gramaticais de pessoa, número e gênero, são [+interpretáveis] no alvo (objeto) e [-interpretáveis] na sonda (predicado verbal). O objeto apresenta o traço [-interpretável] de Caso estrutural e *v* atribui Caso acusativo ao objeto nas construções transitivas. A operação de concordância seria suficiente para eliminar esses traços [-interpretáveis] para a derivação ser convergente. Contudo, se *v* seleciona o traço-EPP,

que é [-interpretável], é necessária a operação de movimento para que a derivação seja convergente.

Consideramos esse modelo de gramática para a compreensão das diferenças na expressão do objeto em função acusativa por um pronome ou elipse na variedade de espanhol e na variedade do PB investigadas, assim como em nossa pesquisa anterior (SIMÕES, 2015). Conforme veremos, propomos essa interpretação teórica para os resultados de nossos estudos recentes.

5 A sociolinguística

De acordo com Labov (2008), os sistemas linguísticos teriam um caráter heterogêneo. Isso significa que as línguas variam em decorrência do fato de estarem inseridas em comunidades de fala.

A variação se manifesta por meio de duas ou mais formas linguísticas equivalentes em significado ou valor funcional que são compartilhadas pelos membros de uma comunidade de fala, apesar de cada membro apresentar frequências diferentes em sua produção. O processo de variação linguística ocorreria porque não seria possível que uma forma inovadora substituísse a outra de maneira instantânea. Entretanto, há fenômenos de variação que se manifestam durante um longo período de tempo e que não resultam em mudança. Em casos assim, estaríamos diante do fenômeno de *variação estável*.

Conforme Weinreich, Labov e Herzog (2009), a variação e mudança linguísticas são regidas pelos seguintes princípios: (1) os fatores condicionantes, que constituem os contextos linguísticos e sociais que podem favorecer uma variante ou outra; (2) a transição, que revela a mudança linguística mediante o estudo de diferentes faixas etárias, estratégia que permite observar a coexistência de duas formas e a obsolescência de uma delas; (3) o encaixamento, que se manifesta nos âmbitos linguístico e social e compreende a alteração gradual das variantes no sistema linguístico e na estrutura social; (4) a avaliação, que compreende o juízo de valor, a significação social que o falante atribui a uma variante, sendo um aspecto que pode atuar na implementação ou retrocesso de uma mudança; (5) a implementação, que também se relaciona com a esfera linguística e à social. Na esfera linguística, considerando-se que a mudança estaria encaixada no sistema linguístico, a expansão de uma variante inovadora acontece de forma gradual.

Entretanto, como a mudança não constitui um processo instantâneo, a estrutura social poderia interferir no avanço da expansão de uma variante, já que a significação social da mesma atua tanto no sentido de sua implementação quanto no de seu retrocesso.

Em nossa pesquisa, nosso objetivo foi investigar os contextos linguísticos que poderiam favorecer o objeto nulo na variedade de espanhol de Madri e os contextos nos quais o pronome lexical seria imprescindível ou não seria aceito para a expressão do objeto no PB. Isso nos permitiria verificar o encaixamento e a implementação dessas variantes na estrutura linguística do espanhol de Madri e do PB de São Paulo.

6 Metodologia

Analisamos as 18 entrevistas da variedade de espanhol de Madri presentes na amostra *La lengua hablada en Madrid. Corpus PRESEEA — Madrid (Distrito de Salamanca)* (CESTERO MANCERA et al. 2014), pertencentes ao *Proyecto para el Estudio Sociolingüístico del Español de España y de América* (PRESEEA), cujos informantes apresentam escolaridade equivalente ao Ensino Médio e diferentes idades. Analisamos também as 11 entrevistas do PB presentes na amostra *A Língua Portuguesa falada em São Paulo: amostra da variedade culta do século XXI* (LIMA-HERNANDES; VICENTE, 2012), pertencente ao *Projeto História do Português Paulista* (PHPP), cujos informantes apresentam alta escolaridade e têm acima de 35 anos⁴.

Como variável dependente, estudamos o objeto anafórico acusativo de 3ª pessoa e, como variantes, o objeto nulo e o clítico, no âmbito da variedade de Madri, e o objeto nulo, o pronome lexical e o clítico, no âmbito da variedade do PB de São Paulo. Na variedade de Madri, a elipse foi a variante que escolhemos como o valor da regra variável. Quanto ao PB, não foi possível realizar a análise multivariada, devido ao fato de a ocorrência dos objetos nulos ter sido quase categórica, como veremos na seção a seguir. Sendo assim, nossa estratégia para verificar as possibilidades do pronome lexical foi observar para cada ocorrência de objeto anafórico da amostra analisada se a presença do

⁴ Contudo, nas entrevistas não consta a informação relativa à idade do informante.

pronome lexical seria possível. Essa verificação foi feita considerando a nossa intuição⁵ de falante paulistana do PB.

Investigamos os seguintes contextos linguísticos: (1) a estrutura do SN antecedente; (2) o traço semântico de animacidade; (3) o traço semântico de especificidade⁶; (4) a semântica verbal; (5) aspecto verbal I: construções [+/-perfectivas]; (6) aspecto verbal II: verbos dinâmicos e de estado; (7) as construções intensionais; (8) o objeto indireto; (9) a predicação secundária; (10) a perífrase verbal; (11) os contextos discursivos. Em relação aos contextos sociais, investigamos o fator social faixa etária⁷ apenas na variedade de espanhol de Madri, já que a amostra do PB não especifica a idade de cada informante. Os informantes foram divididos nas seguintes faixas etárias⁸: 20 a 34 anos (faixa etária 1); 35 a 54 anos (faixa etária 2); a partir dos 55 anos (faixa etária 3).

⁵ Esclarecemos que, até este momento da pesquisa, não realizamos testes de aceitabilidade e/ou de avaliação com falantes do PB sobre o emprego anafórico do pronome lexical. Temos consciência das limitações dos resultados de um estudo que considere nossa própria intuição sobre os dados e reconhecemos que seria relevante conduzir um estudo com essa metodologia em algum momento, a fim de complementar nossas conclusões.

⁶ Para a análise do traço semântico de especificidade, adotamos o critério pragmático conforme Leonetti (1999). Segundo esse autor (LEONETTI, 1999), um SN será considerado [+específico] se, ao empregá-lo, o falante referir-se a uma entidade determinada. Entretanto, não é necessário que o objeto seja conhecido pelo falante, basta a sua intenção de referir-se a uma entidade determinada. Sendo assim, em (1), o SN *un amigo tuyo* será [+específico] ainda que o interlocutor não identifique o referente. (1) *Un amigo tuyo te está esperando abajo.* (LEONETTI, 1999, p. 858)

⁷ Parece-nos relevante o estudo desse condicionador social na medida em que nos permitirá comparar os resultados com os de nosso trabalho anterior (SIMÕES, 2015), no qual analisamos entrevistas de informantes da variedade de espanhol de Madri distribuídos entre as mesmas faixas etárias, porém que apresentam Ensino Superior.

⁸ A princípio, pretendíamos manter a divisão em quatro faixas etárias utilizada em nossa pesquisa anterior (SIMÕES, 2015), que era de 19 a 29 anos (faixa etária 1), de 30 a 45 anos (faixa etária 2), de 46 a 59 anos (faixa etária 3) e de 60 a 89 anos (faixa etária 4). Contudo, mantendo essa divisão, haveria 5 informantes para as faixas etárias 2, 3 e 4 e apenas 3 para a faixa etária 1. Sendo assim, optamos por manter a divisão etária que foi a base para a seleção dos informantes para a amostra.

Os dados de ambas as línguas estudadas foram submetidos ao programa estatístico *Goldvarb X*, que nos permite realizar uma análise quantitativa e assimilar a sistematicidade subjacente à variação linguística.

7 Análise dos dados

Ao analisar as entrevistas da variedade de espanhol de Madri, encontramos 9,3% de objetos nulos nominais e 90,7% de clíticos para a expressão do objeto acusativo anafórico de 3ª pessoa. Observe-se a tabela 2:

Tabela 2 - Ocorrências de objetos nulos e clíticos na variedade de espanhol de Madri

	Variedade de espanhol de Madri					
	Objetos nulos		Clíticos		Total	
	n.	%	n.	%	n.	%
Objeto nominal	102	9,3%	990	90,7%	1092	100%

Conforme vimos na seção anterior, para realizar a análise multivariada, escolhemos a elipse como o valor de aplicação da regra variável no intuito de verificar os condicionadores que favoreceriam essa variante. Foram selecionados como significativos para a ocorrência da elipse os seguintes contextos linguísticos, que aparecem conforme a sua ordem de relevância: estrutura do SN antecedente; perífrase verbal; traço semântico de animacidade do antecedente; predicação secundária; verbos dinâmicos e de estado.

Em relação à estrutura do SN, nossa hipótese era que os objetos nulos seriam favorecidos pelos SNs introduzidos por determinantes com o traço semântico de indefinidade. Para a formulação dessa hipótese, consideramos que os quantificadores e o artigo indefinido teriam o traço semântico de indefinidade, que não lhes permitiria identificar o referente (LEONETTI, 1999). Além disso, sendo o clítico um pronome definido (DI TULLIO, 1997), haveria incompatibilidade em retomar um SN sem determinante, que não constitui uma expressão referencial em espanhol (LACA, 1999). Outra sustentação para essa hipótese são os resultados de nossa pesquisa anterior (SIMÕES, 2015), que mostraram que seriam esses os antecedentes que favoreceriam a elipse.

Podemos considerar que nossa hipótese foi confirmada, uma vez que encontramos 29,3% de objetos nulos e peso relativo de 0.82 com os SNs sem determinantes e 22% de objetos nulos e peso relativo de 0.72 com os SNs quantificados. Contudo, os SNs introduzidos pelo artigo indefinido não se mostraram como um contexto favorecedor da elipse, embora esse contexto tenha sido selecionado como relevante para as variedades de Madri e Montevideu investigadas em SIMÕES (2015). Observem-se os resultados na tabela 3:

Tabela 3 - Ocorrências de objetos nulos conforme a estrutura do SN antecedente

	Variedade de espanhol de Madri		
	n. de ocorrências/total	%	Peso relativo
Determinante definido	45/758	5,9%	0.43
Artigo indefinido	12/160	7,5%	0.49
Quantificador	18/82	22%	0.72
SN sem determinante	27/92	29,3%	0.82

Apresentamos algumas ocorrências de objetos nulos com SNs sem determinante em (35) e com SNs quantificados em (36).

(35a) I: calle un mes o un mes y medio y otra vez encuentra **trabajo** otros mm a lo mejor dos meses o otro aguanta un poco más y a la calle otra vez y así // ahora mismo está en un periodo de que no tiene / no lleva mucho sin trabajar / está buscando Ø (...) (Entrevista 22)

(35b) I: y depende / es que como me hacen un poco la pelota como les hago la comida <risas = "I"/> me hacen la pelota / no pero por ejemplo / me gusta mucho la / la comida sabrosa entonces busco muchas veces **recetas** y voy cambiando Ø / y el el guiso de ternera es uno de los (Entrevista 22)

(36a) E: sí // muy bien // bueno estamos terminando la / la entrevista // oye eeh // tengo que volver ahora // tengo que coger el autobús // eeh // ¿hay por aquí **alguna boca de metro** / cerca de?

I: sí // tienes Ø aquí andando // unos diez minutos / (...) (Entrevista 21)

(36b) I: (...) hay **muchas cosas que podría tener** y no tengo Ø (...) (Entrevista 24)

(36c) E: hm // ¿y has tenido alguna vez **algún problema?** / eeh de **algún accidente con motos** y /

I: hh sí he he tenido Ø (...) (Entrevista 26)

(36d) I: y a mí me tocan **doscientos millones** y desde luego los reparto en a entre aquella gente / yo no necesito Ø para vivir (...) Entrevista 36)

No que se refere à perífrase verbal, nossa hipótese era que essas construções favoreceriam a elipse. Essa hipótese se justifica pelos resultados dos trabalhos de Landa (1993, 1995) sobre o espanhol falado no País Basco. Os resultados revelam que nossa hipótese foi confirmada, já que houve 19% de objetos nulos em construções com perífrase verbal e peso relativo de 0.70. Observe-se a tabela 4:

Tabela 4 - Ocorrências de objetos nulos conforme a presença ou ausência de perífrase verbal na construção

	Variedade de espanhol de Madri		
	n. de ocorrências/total	%	Peso relativo
Presença de perífrase verbal	36/189	19%	0.70
Ausência de perífrase verbal	66/903	7,3%	0.70

Em (37), apresentamos alguns casos de elipses em construções com perífrase verbal. Em (37a), temos uma perífrase com gerúndio. Nas construções em (37b) e (37d), temos perífrases modais com o verbo deôntico *poder*. Em (37c), a perífrase é aspectual. Nas construções em (37e), as perífrases também são modais, porém com *tener que*, que expressa necessidade e, portanto, são deônticas também.

(37a) I: (...) te tienes que buscar **las habichuelas** // y cad / cada dos por tres estás pidiendo Ø fuera <silencio/> (Entrevista 19)

(37b) I: (...) es que de meterte en un chalecito con no sé qué / a meterte en **un piso con una habitación de hace cuarenta años** pues no / y claro es que aquí no puedes comprar Ø de otra manera / (Entrevista 20)

(37c) I: una vez hice unos es una vez hice **unos espaguetis** / y tuve que llamar por teléfono / a la chica esta // para que me guiara cómo tenía que ir haciendo los espaguetis / y desde ese día ya no

E: ¿y qué hiciste? /

I: no pues ese día cogí unos espaguetis con tomate que tampoco tiene mayor misterio // y la verdad es que no me gustaron porque le eché mucha sal / y no / no he vuelto a intentar hacer Ø (Entrevista 21)

(37d) I: yo desde luego desde que mm i desde que yo vivo aquí en Madrid desde que **tengo** así un poco / de conocimiento que iba al colegio y tal / pues yo veo a mis sobrinos que son poco más mayores porque a mi hija lógicamente no puedo comparar Ø (Entrevista 23)

(37e) I: claro / y si te dan **una exposición** te dan una exposición que

E: <ininteligible/> y unos medios /

I: que es de otro de otro sitio / claro / el que tenía que proponer Ø el que tenía que hacer Ø es el barrio / la gente del barrio (...) (Entrevista 25)

Com respeito ao traço semântico de animacidade do antecedente, nossa hipótese era que os SNs [-animados] favoreceriam a omissão do objeto. Nossa hipótese se sustenta em trabalhos sobre o espanhol (LANDA, 1993, 1995), em nossa pesquisa anterior (SIMÕES, 2015) e também no trabalho de Duarte (1986) sobre o PB, que revelam que as elipses são favorecidas pelos antecedentes [-animados]. Encontramos 11,9% de objetos nulos com SNs [-animados] e peso relativo de 0.58, de modo que confirmamos nossa hipótese. Observe-se a tabela 5:

Tabela 5 - Ocorrências de objetos nulos conforme o traço de animacidade do SN antecedente

	Variedade de espanhol de Madri		
	n. de ocorrências/total	%	peso relativo
[+animado]	12/338	3,6%	0.33
[-animado]	90/754	11,9%	0.58

Observemos algumas construções extraídas da amostra e que apresentam objetos nulos com antecedentes [-animados].

(38a) I: (...) y viene ya **el de la ONCE** o algo de eso y compran / **compras Ø** / pero por el hecho de joder / a ver si les va a tocar a estos que están comprando Ø al lado mío (...) (Entrevista 20)

(38b) I: un robo / un robo / fue / eeh / teníamos que hacer un biombo / para separar unas clases y tal / entonces yo diseñé **un biombo** / que tenía como unos triángulos que salían / y cuando cerrabas Ø (...) (Entrevista 22)

(38c) I: pero realmente en el barrio / pues no hay un espacio para ahora han hecho ese pero / no es **un espacio** para para jugar en la medida que conocemos Ø ¿no? / (...) (Entrevista 25)

Conforme a Hierarquia Referencial de Cardinaletti e Starke (1994), que é abordada por Cyrino, Duarte e Kato (2000), uma entidade [+animada] e/ou [+específica] estaria situada na extremidade mais referencial da escala, enquanto uma entidade [-animada] e/ou [-específica] ocuparia uma posição menos referencial. Tendo em vista essa gradação na referencialidade, a tendência seria que um argumento [+animado] e/ou [+específico] seja retomado por um pronome, ao contrário de um argumento [-animado] e/ou [-específico], que teria mais probabilidade de não se realizar foneticamente. Considerando-se essa proposta, vemos que, no que tange aos SNs [-animados], nossos resultados refletem essa hierarquia.

Quanto às construções com predicação secundária⁹, tendo em vista os trabalhos de Landa (1993, 1995) para o espanhol basco, nossa hipótese era que os objetos nulos seriam favorecidos nessas construções. No entanto, foram as construções sem predicação secundária que favoreceram a omissão. Encontramos 10,7% de elipse e peso relativo de 0.56 nas construções sem predicação secundária e 2,3% e peso relativo de 0.21 nas construções com predicação secundária, de maneira que nossa hipótese não foi confirmada. Por outro lado, esse resultado se assemelha ao encontrado por Duarte (1986) para o PB, já que esse contexto linguístico favorecia a realização pronominal do objeto. Observem-se nossos resultados na tabela 6:

⁹ De acordo com Demonte e Masullo (1999), a predicação secundária consiste em construções que apresentam um predicativo que modifica tanto o predicado verbal quanto um SN, que pode ser o sujeito ou o objeto acusativo. O papel desse predicativo é atribuir um estado ou propriedade a um SN. A dependência sintática do predicativo se evidencia na concordância de gênero e número com o SN quando aquele é um adjetivo. Observem-se as sentenças em (2), que apresentam predicativos que atribuem um estado ou propriedade aos objetos acusativos:

(2a) Juan guardó la camisa *sucia*.

(2b) Considero a tu hermana {*de muy buen carácter/inteligente*} (cf. #*Considero a tu hermana*). (DEMONTE; MASULLO, 1999, p. 2463)

Se em (2a) interpretarmos *sucia* como um predicativo, veremos que esse adjetivo denota o estado do SN *la camisa* no momento em que esta foi guardada. Esse estado da camisa está relacionado com o evento denotado pelo predicado verbal. Em (2b), os sintagmas *de muy buen carácter* e *inteligente* modificam o SN *a tu hermana*. A diferença entre essas duas construções é de que em (2a) o predicativo é um adjunto e, em (2b), o predicativo é um complemento, na medida em que é requerido pelo predicado verbal. Não teríamos em (2a) uma construção de predicação secundária se o adjetivo *sucia* integrasse o SN *la camisa sucia*.

Tabela 6 - Ocorrências de objetos nulos conforme a presença ou ausência de predicação secundária na construção

	Variedade de espanhol de Madri		
	n. de ocorrências/total	%	Peso relativo
Presença de predicação secundária	4/173	2,3%	0.21
Ausência de predicação secundária	98/919	10,7%	0.56

Em (39), apresentamos algumas ocorrências de objetos nulos em construções com predicação secundária encontradas na amostra. Em (40), apresentamos construções sem predicação secundária.

(39a) E: sí // muy bien // bueno estamos terminando la / la entrevista // oye eech // tengo que volver ahora // tengo que coger el autobús // eeh // ¿hay por aquí **alguna boca de metro** / cercade?

I: sí // tienes Ø aquí andando // unos diez minutos / un cuarto de hora andando en línea recta // llega a la boca de metro (Entrevista 21)

(39b) E: y vivías por el barrio / ¿ha cambiado mucho en cuanto al tiempo? / ¿el tiempo que hace?

I: ¿**el tiempo**

E: sí

I: **atmosférico?** pues yo casi recuerdo Ø antes más frío me parecía a mí / que hacía cuando yo era niña o / (...) (Entrevista 28)

(40a) I: lo comprenderías ¿no? o bueno / tampoco lo comprenderías ¿no? porque **esa violencia** / con la que ellos / ech actuaron pues tampoco // yo es que no lo entiendo o sea no entiendo a la gente que puede ir haciendo

E: es decir porque han mirado mal a / a mi / a mi novia ¿no?

I: no / sí / pero es que no / aun así es que yo no entiendo Ø (Entrevista 19)

(40b) I: realmente / mis preocupaciones son cuando alguien de mi familia está enfermo / ¿sabes? o sea que / no sé no no me gasto en pensar en en qué podría tener y no tengo qué podría ser y no soy / porque ¡hombre! hay **muchas cosas que podría tener** y no tengo Ø hay cosas que me gustaría ser y no soy ves (...) (Entrevista 24)

Por fim, com relação ao aspecto lexical dos verbos, partimos da hipótese de que a elipse seria favorecida nas construções com verbos de estado. Formulamos essa hipótese com base nos resultados do estudo de Landa (1993, 1995) para o espanhol basco. Observamos 11,5% de objetos nulos e peso relativo de 0.67 com verbos estativos, de modo que confirmamos nossa hipótese. Observe-se a tabela 7:

Tabela 7 - Ocorrências de objetos nulos conforme o aspecto lexical dos verbos

	Variedade de espanhol de Madri		
	n. de ocorrências/total	%	Peso relativo
Verbos dinâmicos	75/858	8,7%	0.46
Verbos estativos	27/234	11,5%	0.67

As construções em (41) apresentam verbos estativos, sendo eles *apreciar, conocer, tener y recordar*.

(41a) I: en lo que nosotros nos movemos es / pues eso // lo que son **las Scooter** por // porque es lo que nos gusta y porque sabes apreciar Ø realmente (Entrevista 19)

(41b) I: pero realmente en el barrio / pues no hay un espacio para ahora han hecho ese pero / no es **un espacio** para para jugar en la medida que conocemos Ø (...) (Entrevista 25)

(41c)E: sí digo a lo mejor en el cuartel había **una banda** o algo/

I: no / no no / no tenía Ø todavía (...) Entrevista 25)

(41d) I: a mí me ha compensado / porque mandarle a una guardería tan pequeño me hubiera horrorizado / tener a **una chica** no conocía Ø así de confianza (...) (Entrevista 28)

(41e) E: ¿y cómo cómo recuerdas el nacimiento de / de tu hija?

I: hombre pues **el nacimiento de mi hija** pues cómo voy a recordar Ø (...) (Entrevista 32)

A hipótese geral da pesquisa para a variedade de espanhol de Madri era que os objetos nulos seriam favorecidos por antecedentes com o traço semântico de indefinidade, os [-animados] e [-específicos]. Os resultados encontrados nos permitem confirmar parcialmente a nossa hipótese, já que os objetos nulos foram favorecidos pelos SNs sem determinante, os quantificados e os [-animados], mas também pelas construções com perífrase verbal, as construções em que não ocorre predicação secundária e as construções com verbos de estado.

Em relação aos demais contextos linguísticos investigados, observamos que os objetos nulos ocorreram em construções com verbos cognitivos, de comunicação, aspecto imperfeito, construções

intensionais¹⁰, objeto indireto e tópico, muitas das quais são contextos que favorecem os objetos nulos nas variedades de espanhol do País Basco (LANDA, 1993, 1995) e de Quito (SUÑER; YÉPEZ, 1988).

Na construção em (42a), temos objeto indireto e negação, que compreende contexto intensional. Em (42b), a primeira construção

¹⁰ Conforme Leonetti (1999), as construções intensionais seriam aquelas que apresentam elementos que levam os SNs a receberem uma interpretação [-específica] prototípica. Esses contextos seriam os elementos modais com âmbito oracional, a negação e os predicados intensionais. Em relação aos elementos modais, estes transformam a construção em um contexto intensional, supendendo as implicações existenciais dos SNs. Sendo assim, o evento descrito seria não factual. Esses elementos incluem as orações interrogativas, os tempos futuro e condicional, os modos imperativo e subjuntivo, as construções condicionais, os enunciados genéricos e caracterizadores, as orações com advérbios de possibilidade (*quizás*) e os predicados modais (*poder, deber, posible*). Observem-se os exemplos:

- (3a) ¿Quién me presta un bolígrafo?
- (3b) {Me compraré/Me compraría} una gabardina nueva.
- (3c) {Cuando/Aunque} te compres una gabardina nueva...
- (3d) Si encuentran un paquete sospechoso, avisen al portero.
- (3e) Cada vez que suena el teléfono tiene un sobresalto.
- (3f) Quizás me compre una gabardina.
- (3g) {Pueden/Deben} llegar unos agentes de refuerzo. (LEONETTI, 1999 p. 861-862)

No que se refere à negação, esta levaria um SN indefinido a ser interpretado como [-específico], já que anula a implicação de existência do referente. Para que essa leitura seja possível, é necessário que o SN esteja no âmbito da negação. Essa interpretação ocorre em (4a), mas não em (4b). Seriam elementos de negação o advérbio *no* e os quantificadores negativos; os predicados de impedimento, recusa, dúvida (*oponerse a, dudar de*); a preposição *sin*; e advérbios como *raramente*.

- (4a) El departamento no dispone de un especialista que supervise el experimento.
- (4b) Un especialista (*que supervise el experimento) no dispone de laboratorio. (LEONETTI, 1999, p. 862).

Quanto aos predicados intensionais, estes compreendem contextos nos quais não estaria assegurada a existência de um SN, já que esta depende dos desejos, esperanças, intenções, planos ou necessidades expressadas pelo predicado. Compreendem esses predicados verbos como *buscar, necesitar, pedir, querer, intentar, obligar a, esperar, planear, permitir*; adjetivos como *conveniente, necesario, suficiente, obligatorio, imprescindible*; e a preposição *para* em construções subordinadas finais. Também se incluem nesse grupo os verbos de atitude proposicional *creer* e *suponer*, pois seu comportamento se assemelha aos outros predicados. Observem-se os exemplos:

- (5a) Un refugiado está pidiendo [un cigarrillo] a un soldado.
- (5b) Sugerimos a un amigo [que hiciera [un viaje por el extranjero]].
- (5c) Es conveniente [dedicar [un día] al ocio]. (LEONETTI, 1999, p. 863)

apresenta os contextos intensionais de negação e tempo futuro e a segunda apresenta perífrase modal, que também compreende um contexto intensional. Em (42c), temos o verbo cognitivo *comparar*, a perífrase modal e a negação, ambos contextos intensionais conforme vimos. A construção em (42d) apresenta objeto indireto e verbo de comunicação. Em (42e), temos o verbo cognitivo *conocer* e negação.

(42a) I: este es **tu trabajo** / ¿se lo has dado? y yo ¡no no ni se lo he dado ni me han pagado Ø ni! ¡no podía creer nada! (...) (Entrevista 22)

(42b) E: ¿crees que está cambiando **el tiempo** o que / ahora con?

I: es cíclico / hay años que hay un invierno más frío / otros años que es un **invierno** más / más templado pero no /

E: pero no crees que sea por nada de esto del ozono que que dicen

I: no no / hombre / no sé si afectará Ø / si realmente puede que llegue a afectar Ø en un momento dado (...) (Entrevista 23)

(42c) I: yo desde luego desde que mm i desde que yo vivo aquí en Madrid desde que **tengo** así un poco / de conocimiento que iba al colegio y tal / pues yo veo a mis **sobrinos** que son poco más mayores porque a mi hija lógicamente no puedo comparar Ø (Entrevista 23)

(42d) I: exactamente sí / de más eeh / intimismo ¿no? / si por llamarlo de alguna forma ¿no? / y si quizá con él pues eeh ts hombre **las cosas que pudiera haber hecho allí** / me contaba Ø / (...) (Entrevista 25)

(42e) I: a mí me ha compensado / porque mandarle a una guardería tan pequeño me hubiera horrorizado / tener a **una chica no conocía Ø así de confianza** (...) (Entrevista 28)

Em relação ao fator social faixa etária, partimos da hipótese de que a variação linguística no âmbito da expressão do objeto acusativo corresponderia à variação estável na variedade de espanhol de Madri. De acordo com Labov (2008), a análise de diferentes gerações de falantes de diferentes faixas etárias constitui um estudo na dimensão do tempo aparente e permitiria observar se a variável em estudo constitui variação estável ou mudança em progresso.

Os resultados revelam que os informantes da faixa etária 1 apresentam 9,9% de elipse; os da faixa etária 2, 11,3%; e os da faixa etária 3, 7,2%. Entendemos, portanto, que se trata de um fenômeno de variação estável, de modo que confirmamos nossa hipótese. Em nossa pesquisa anterior (SIMÕES, 2015), os resultados também nos levaram a concluir que se tratava de um processo de variação estável na esfera da variedade de espanhol de Madri, embora na variedade de espanhol

de Montevideu os resultados tenham indicado mudança em progresso. Observe-se a tabela 8:

Tabela 8 - Ocorrências de objetos nulos conforme a faixa etária do informante

	Variedade de espanhol de Madri	
	n. de ocorrências/total	%
Faixa etária 1	36/365	9,9%
Faixa etária 2	38/336	11,3%
Faixa etária 3	28/391	7,2%

Em relação à análise das entrevistas do PB, observamos a ocorrência quase categórica de elipses para a expressão do objeto acusativo de 3ª pessoa, já que encontramos apenas cinco ocorrências do clítico e nenhuma do pronome lexical, como se observa na tabela 9.

Tabela 9 - Ocorrências de objetos nulos e clíticos na variedade do PB de São Paulo

	Variedade do PB de São Paulo					
	Objetos nulos		Clíticos		Total	
	n.	%	n.	%	n.	%
Objeto nominal	176	97,2%	5	2,8%	181	100%

Esse resultado revela o efeito da escolarização sobre os falantes cultos do PB, já que, por um lado, evitam o pronome lexical, que é a variante estigmatizada, e, por outro, produzem de maneira muito marginal o clítico, que é a variante prestigiada pela tradição escolar.

Tendo em vista esse cenário, não pudemos utilizar o programa estatístico *Goldvarb X* para realizar a análise multivariada e obter, assim, os pesos relativos, bem como verificar os contextos linguísticos que favoreceriam o pronome lexical. A ocorrência quase categórica de objetos nulos revela que essa variante se encontra em um estágio bastante avançado de implementação na variedade de São Paulo do PB contemporâneo.

Conforme vimos na seção anterior, para verificar as possibilidades de ocorrência do pronome lexical, nossa estratégia foi observar para cada ocorrência de objeto nulo e clítico se o pronome lexical seria possível na construção. Veremos essa verificação que fizemos ao longo de toda a análise.

Em relação à estrutura do SN antecedente, com base em nossa pesquisa anterior (SIMÕES, 2015), nossa hipótese era que os objetos nulos seriam mais frequentes com antecedentes que apresentassem traços semânticos de definitude ou de indefinitude. Os resultados da análise quantitativa revelam 96,8% de elipses com SNs introduzidos por determinantes definidos, 92% com SNs introduzidos pelo artigo indefinido e uso categórico das elipses com SNs quantificados ou que não apresentam nenhum determinante. Vemos, portanto, que nossa hipótese foi confirmada. Observe-se a tabela 10:

Tabela 10 - Ocorrências de objetos nulos e clíticos conforme a estrutura do SN antecedente

	Variedade do PB de São Paulo					
	Objetos nulos		Clíticos		Total	
	n.	%	n.	%	n.	%
Determinante definido	92	96,8%	3	3,2%	95	100%
Artigo indefinido	23	92%	2	7,4%	25	100%
Quantificador	18	100%	0	0%	18	100%
SN sem determinante	43	100%	0	0%	43	100%

No que se refere às possibilidades de pronome lexical, no âmbito dos SNs introduzidos por determinante definido, consideramos 77,9% das construções gramaticais e, entre os indefinidos, 80%. Em relação aos SNs antecedentes quantificados e sem determinante, consideramos agramaticais boa parte dessas construções, sendo as frequências 44,4% e 81,4%, respectivamente. Observe-se a tabela 11:

Tabela 11 - Possibilidades do pronome lexical conforme a estrutura do SN antecedente

	Variedade do PB de São Paulo									
	Gramatical		Agramatical		Inaceitável		Específico		Total	
	n.	%	n.	%	n.	%	n.	%	n.	%
Determinante definido	74	77,9%	7	7,4%	6	6,3%	8	8,4%	95	52,5%
Artigo indefinido	20	80%	2	8%	0	0%	3	12%	25	13,8%
Quantificador	8	44,4%	8	44,4%	2	11,1%	0	0%	18	9,9%
SN sem determinante	3	7%	35	81,4%	5	11,6%	0	0%	43	23,8%

Em (43), apresentamos dados da amostra analisada, nos quais os objetos nulos apresentam como antecedentes SNs introduzidos por artigo determinado, indefinido, quantificador ou não apresentavam determinante. Além disso, apresentamos as construções com pronome lexical que elaboramos tendo em vista as construções com elipse, cuja gramaticalidade julgamos, como já mencionado, considerando nossa intuição de falante nativa e paulistana do PB. Observamos que há incompatibilidade na retomada de um SN sem determinante pelo pronome lexical conforme constatou também Casagrande (2012).

(43a) [...] aí ela “a::i professo::ra que coisa bo::a que a senhora ve::io... eu vou lhe pedir... a senhora me trás **as apostilas de fisiologia do curso** enquanto... eu tiver aqui eu posso ficar lendo Ø aí quando eu sair eu já sei”... no dia seguinte fui lá... entreguei pra ela as apostilas da fisiologia [...] (Informante 1)

(43a’) ... a senhora me trás **as apostilas de fisiologia do curso** enquanto... eu tiver aqui eu posso ficar lendo **elas**...

(43b) L1 mas cês tão fazendo **um trabalho de final de curso**?

L2 é / não não é um trabalho de final...éh num é de final de curso é um trabalho pro semestre.

[
L1 fazendo ()Ø? cês tão

L2 é um trabalho para o semestre né?

L1 uhn (Informante 4)

(43b’) ... mas cês tão fazendo **um trabalho de final de curso**?

... é / não não é um trabalho de final...éh num é de final de curso é um trabalho pro semestre...

... cês tão fazendo ()ele?

(43c) L1 [...] eu não publiquei NEM um décimo dos resultados que eu tenho porque o proBLEma que eu também sou uma grande::... como é que a gente fala?... coeto **muitos dados** ((risos))

[
L2 ((risos))

L1 e depois eu nem dou conta de... de usar Ø sabe... [...] (Informante 4)

(43c’) ... como é que a gente fala?... coeto **muitos dados** ((risos)) e depois eu nem dou conta de... de usar **eles** sabe...

Com relação ao traço semântico de animacidade do antecedente, tendo em vista os trabalhos de Duarte (1986) e Cyrino (1994), nossa hipótese era que a elipse seria mais frequente com antecedentes

[-animados]. Observamos 75% de objetos nulos com SNs [+animados] e 98,8% com SNs [-animados]. Os resultados demonstram que as entidades [-animadas] são as que apresentam as frequências mais elevadas de omissão, apesar de os SNs [+animados] também compreenderem um contexto de implementação dessa variante. Sendo assim, confirmamos parcialmente a nossa hipótese. Observe-se a tabela 12:

Tabela 12 - Ocorrências de objetos nulos e clíticos conforme o traço de animacidade do SN antecedente

	Variedade do PB de São Paulo					
	Objetos nulos		Clíticos		Total	
	n.	%	n.	%	n.	%
[+animados]	9	75%	3	21,4%	12	100%
[-animados]	167	98,8%	2	1,1%	169	100%

Quanto às possibilidades de pronome lexical, consideramos gramaticais 83,3,4% das construções com SNs [+animados] e apenas 56,2% daquelas com SNs [-animados]. A tendência encontrada demonstra que o traço semântico de animacidade se relaciona com a possibilidade de expressão pronominal do objeto na variedade do PB investigada. Observe-se a tabela 13:

Tabela 13 - Possibilidades do pronome lexical conforme o traço de animacidade do SN antecedente

	Variedade do PB de São Paulo									
	Gramatical		Agramatical		Inaceitável		Específico		Total	
	n.	%	n.	%	n.	%	n.	%	n.	%
[+animados]	10	83,3%	1	8,3%	0	0%	1	8,3%	12	6,6%
[-animados]	95	56,2%	51	30,2%	13	7,7%	10	5,9%	169	93,4%

Na esfera do traço semântico de especificidade, com base nos resultados de Cyrino (1994), nossa hipótese era que a elipse seria mais frequente com antecedentes [+/-específicos]. Observamos 94,2% de objetos nulos com SNs [+específicos] e 98,4% com SNs [-específicos]. Conforme vemos nos resultados, confirmamos nossa hipótese. Observe-se a tabela 14:

Tabela 14 - Ocorrências de objetos nulos e clíticos conforme o traço de especificidade do SN antecedente

	Variedade do PB de São Paulo					
	Objetos nulos		Clíticos		Total	
	n.	%	n.	%	n.	%
[+específicos]	49	94,2%	3	5,6%	52	100%
[-específicos]	127	98,4%	2	1,5%	129	100%

Em relação à possibilidade de pronome lexical, consideramos gramaticais 94,2% das construções com SNs [+específicos] e apenas 43,4% das construções com os [-específicos]. Portanto, esse resultado nos permite observar que a possibilidade de pronome lexical se relaciona com os SNs [+específicos], assim como constatou Casagrande (2012). Observe-se a tabela 15:

Tabela 15 - Possibilidades do pronome lexical conforme o traço de animacidade do SN antecedente

	Variedade do PB de São Paulo									
	Gramatical		Agramatical		Inaceitável		Específico		Total	
	n.	%	n.	%	n.	%	n.	%	n.	%
[+específicos]	49	94,2%	2	3,8%	1	1,9%	0	0	52	28,7%
[-específicos]	56	43,3%	50	38,8%	12	9,3%	11	8,5%	129	71,3%

Cruzamos os fatores de animacidade e especificidade e observamos 98% de objetos nulos com SNs [-animados; +específicos], 67% com os [+animados; +específicos], 99% com os [-animados; -específicos] e 88% com os [+animados; -específicos]. A tendência encontrada revela que os objetos nulos na variedade de São Paulo do PB contemporâneo estão fortemente relacionados com o traço [-animado] do antecedente, assim como constataram Duarte (1986) e Cyrino (1994). Observe-se a tabela 16:

Tabela 16 - Ocorrências de objetos nulos e clíticos conforme o traço de animacidade e especificidade do SN antecedente

	Variedade do PB de São Paulo					
	Objetos nulos		Clíticos		Total	
	n.	%	n.	%	n.	%
[-animados; +específicos]	45	98%	1	2%	46	100%
[+animados; +específicos]	4	67%	2	33%	6	100%
[+animados; +específicos]	122	99%	1	1%	123	100%
[+animados; -específicos]	5	83%	1	17%	6	100%

Com respeito à possibilidade do pronome lexical, consideramos gramaticais 93% das construções com SN [-animado; +específico] e todas as construções com SN [+animado; +específico]. Entretanto, consideramos gramaticais apenas 42% das construções com SN [-animado; -específico] e 67% das com SN [+animado; -específico]. Essas tendências reforçam que a possibilidade de pronome lexical na variedade de São Paulo do PB contemporâneo estaria fortemente relacionada com entidades [+específicas]. Além disso, essa possibilidade se amplifica quando o antecedente é [+animado; +específico].

Tabela 17 - Possibilidades do pronome lexical conforme o traço de animacidade e especificidade do SN antecedente

	Variedade do PB de São Paulo									
	Gramatical		Agramatical		Inaceitável		Específico		Total	
	n.	%	n.	%	n.	%	n.	%	n.	%
[-animado; +específico]	43	93%	2	4%	1	2%	0	0%	46	100%
[+animado; +específico]	6	100%	0	0%	0	0%	0	0%	6	100%
[-animado; -específico]	52	43%	49	40%	12	10%	10	8%	123	100%
[+animado; -específico]	4	67%	1	17%	0	0%	1	17%	6	100%

Considerando-se esses resultados, vemos que os dados quantitativos confirmam parcialmente nossa hipótese geral da pesquisa no âmbito dos objetos nulos, uma vez que, por um lado, os antecedentes com os traços semânticos de indefinitude, definitude, [-animados] e [+/-específicos] manifestaram maior frequência da elipse e, por outro lado, os objetos nulos também apresentaram uma incidência considerável com os antecedentes [+animados].

No que se refere à hipótese relativa ao pronome lexical, esta também foi parcialmente confirmada, pois consideramos gramatical uma grande parcela das construções em que o pronome retoma um SN indefinido com antecedentes [-animados] e [-específicos] e as maiores restrições ocorreram entre os SNs quantificados.

As construções em (44) exibem SNs antecedentes com diferentes traços semânticos que se expressam pela elipse, mas aceitam o pronome lexical. Em (44a), o SN antecedente é definido [+animado; +específico], em (44b), é definido [-animado; +específico], em (44c), é definido [-animado; -específico], em (44d), é indefinido [-animado; -específico] e, em (44e), é quantificado [-animado; +específico].

(44a) L3 e o senhor conheceu **sua esposa** em toda essa jornada quando?

L1 eu conheci... bom... aí primeiro eu me casei muito jovem não é minha filha sabe não é... me casei com quarenta e cinco quase quarenta e seis anos não é... eu conheci Ø num curso que eu fui fazer no Rio de Janeiro [...] (Informante 7)

(44a') ... e o senhor conheceu **sua esposa** em toda essa jornada quando?

... eu conheci... bom... aí primeiro eu me casei muito jovem não é minha filha sabe não é... me casei com quarenta e cinco quase quarenta e seis anos não é... eu conheci ela num curso que eu fui fazer no Rio de Janeiro

(44b) L2 [...] você acaba não achando muito tempo mas eu gostaria muito tenho ainda **esse plano apesar de já tá já devia estar praticando Ø** tenho ainda esse plano (não teria tempo) (Informante 10)

(44b') ... mas eu gostaria muito tenho ainda **esse plano apesar de já tá já devia estar praticando ele** tenho ainda esse plano (não teria tempo)

(44c) L1 [...] eu eu encaro sempre como encomenda... mesmo **as canções que eu faço pros meus discos**... também são encomendas eu tenho mais prazer de fazer Ø [...] (Informante 6)

(44c') ... mesmo **as canções que eu faço pros meus discos**... também são encomendas eu tenho mais prazer de fazer elas

(44d) L1 [...] do que que cria a compatibilidade entre melodia e... e letra... aí é **um estudo que dá pra fazer em pós-graduação**... não posso introduzir Ø em graduação porque... não faz parte do programa ((sorriu))... [...] (Informante 6)

(44d') ... aí é **um estudo que dá pra fazer em pós-graduação**... não posso introduzir ele em graduação

(44e)((L1 voltando)) eu não encontro a nova fase já saíram **dois números da nova fase** mas eu fiquei tão decepcionado que eu acho que eu não sei se eu joguei Ø fora... ((risos)) [...] (Informante 5)

(44e') ... já saíram **dois números da nova fase** mas eu fiquei tão decepcionado que eu acho que eu não sei se eu joguei eles fora... ((risos))

Quanto aos demais contextos estudados, sua escolha se justifica pelo fato de que já haviam se mostrado relevantes na ocorrência de elipses no PB (DUARTE, 1986; CASAGRANDE, 2012) e em variedades do espanhol (LANDA, 1993, 1995; SUÑER; YÉPEZ, 1988).

Ao realizar a análise quantitativa dos dados, observamos que as elipses estão altamente implementadas em construções com verbos cognitivos, de comunicação e demais verbos, em construções com aspecto [+/-perfectivo], com verbos dinâmicos e estativos, em construções intensionais ou sem esses elementos, em construções com ou sem objeto indireto, perífrase verbal e predicação secundária e em construções com tópico e outros contextos discursivos. Observamos, ainda, objetos nulos categóricos nas construções com verbos de comunicação, aspecto perfectivo, objeto indireto, perífrase verbal e tópicos.

Em relação à possibilidade de pronome lexical, consideramos gramaticais 69,2% das construções com verbos cognitivos, enquanto consideramos agramaticais 60% das construções com verbos de comunicação. Consideramos o pronome lexical gramatical em 90% das construções com aspecto perfectivo e apenas em 54% das construções com aspecto imperfectivo. Com respeito ao aspecto lexical, consideramos gramaticais 62,2% das construções com verbos dinâmicos, mas agramaticais 57,6% das construções com verbos de estado. Na esfera das construções intensionais, consideramos gramaticais 56,6% das construções. Nas construções com objeto indireto, consideramos gramaticais 41,2% das construções. Quanto às construções com predicação secundária, consideramos gramaticais 75% dessas construções. No caso das construções com perífrase verbal, consideramos gramaticais 59,6% das construções. Por fim, com relação às construções com tópico, consideramos gramaticais 47,1% das construções, ao passo que nos demais contextos discursivos esse percentual foi de 59,1%.

De acordo com nossa intuição, o pronome lexical é melhor aceito nas construções com verbo de aspecto perfectivo, verbos dinâmicos e predicação secundária. Contudo, apresenta baixa aceitação nas construções com predicados estativos, de comunicação, objetos indiretos e tópico.

Em (45a), temos um SN indefinido [-animado; +específico] em construção perfectiva e o pronome lexical poderia retomar o antecedente. No entanto, em (45b) e (45c), as construções são imperfectivas e o pronome lexical seria aceito apenas na primeira, na qual o SN é indefinido [-animado; -específico]. Na segunda construção, que apresenta um SN quantificado [-animado; -específico] e também um objeto indireto, consideramos o pronome lexical agramatical.

(45a) L1 houve **uma inundação** no começo do ano

L4 e isso:: por exemplo talvez é::

L1 é que a inundação sempre ocorreu do lado de cá do lado dos professores... então ah:: todo ano a gente já estava até acostumado... o pessoal já tem uns rodos... grandes... bastante panos bastante baldes... o conjunto aqui do lado sempre inunda

L6 exatamente os vizinhos sempre passam (aperto) bem pior que a gente

L1 é porque o prédio tomba um pouco pro lado de lá então... chove mais do lado de lá... ah:: esse ano excepcionalmente estourou a canalização do lado da sala de aula e os alunos descobriram que o prédio inundava... mas ele inunda todo ano...

L6 (virou) um toboáguas... tem vídeo no Youtube... todo mundo viu

L1 todo mundo viu Ø no Youtube (Informante 9)

(45a') L1 houve **uma inundação** no começo do ano

(...)

L1 todo mundo viu **ela** no Youtube

(45b) L2 [...] mais fácil é ser advogado cê pega **um volume olha Ø lê Ø** e ... e diz ... (Informante 1)

(45b') mais fácil é ser advogado cê pega **um volume olha ele lê ele**

(45c) (L1) Leva isso aqui pra ele... pra vê se ele se interessa em faze alguma... **alguma melodia..**". eu recebo muitas letras na verdade... de muita gente e tal... acabo fazendo Ø pra pouca gente porque::... [...] (Informante 6)

(45c') pra vê se ele se interessa em faze alguma... **alguma melodia..**". eu recebo muitas letras na verdade... de muita gente e tal... acabo fazendo ***ela** pra pouca gente

Em (46a), o SN é definido, em [46b), indefinido, e, em (46c), quantificado. Em todos esses casos o SN é [-animado; -específico] e a construção apresenta verbo dinâmico. Observamos que o pronome lexical seria possível com o SN antecedente definido e indefinido, mas agramatical com o quantificado. Nas construções em (46d) a (46h), os verbos são de estado. Em (46d) o SN é definido [+animado; +específico], em (46e), definido [-animado; +específico], e, em (46f), definido [-animado; -específico]. Em todas essas construções, o pronome lexical seria aceito. Quanto às construções em (46g) e (46h), os SNs são quantificados e [-animados; -específicos] e o pronome lexical é agramatical. Destacamos que em (46g) o verbo de estado é *ter*. Tendo em vista essa análise, observamos que a possibilidade do pronome lexical se relaciona não apenas com os traços semânticos do antecedente, mas também com a estrutura do SN e o aspecto lexical do verbo da construção.

(46a) L1 [...] todo mundo que faz **o seu trabalho** quer mostra:: e eu entendo isso... então eu pego e levo... talvez um dia eu vá ouvi (e) numa/numas férias e vezes até faço isso... pego umas férias ouço um pouco mais e tal... mas seria é humanamente impossível::vel ouvir tudo aquilo... a pessoa que entrega **Ø** pra você... [...] (Entrevista 6)

(46a') ... todo mundo que faz **o seu trabalho** quer mostra:: e eu entendo isso... então eu pego e levo... talvez um dia eu vá ouvi (e) numa/numas férias e vezes até faço isso... pego umas férias ouço um pouco mais e tal... mas seria é humanamente impossível::vel ouvir tudo aquilo... a pessoa que entrega **ele** pra você...

(46b) L1 [...] às vezes faço **uma letra** eles não gostam muito eu refaço **Ø**... até dá certo né... (e é uma)... como como isso é atividade ... como qualquer outra atividade... (cê) tem que ficá ensaiando tentando faze até dá certo.... [...] (Informante 6)

(46b') ... às vezes faço **uma letra** eles não gostam muito eu refaço **ela**...

(46c) [...] Leva isso aqui pra ele... pra vê se ele se interessa em faze alguma... **alguma melodia**..". eu recebo muitas letras na verdade... de muita gente e tal... acabo fazendo **Ø** pra pouca gente [...] (Entrevista 4)

(46c') ... Leva isso aqui pra ele... pra vê se ele se interessa em faze alguma... **alguma melodia**..". eu recebo muitas letras na verdade... de muita gente e tal... acabo fazendo ***ela** pra pouca gente...

(46d) L3 e o senhor conheceu **sua esposa** em toda essa jornada quando?

L1 eu conheci... bom... aí primeiro eu me casei muito jovem não é minha filha sabe não é... me casei com quarenta e cinco quase quarenta e seis anos não é... eu conheci **Ø** num curso que eu fui fazer no Rio de Janeiro [...] (Informante 7)

(46d') ... e o senhor conheceu **sua esposa** em toda essa jornada quando?

... eu conheci... bom... aí primeiro eu me casei muito jovem não é minha filha sabe não é... me casei com quarenta e cinco quase quarenta e seis anos não é... eu conheci **ela** num curso que eu fui fazer no Rio de Janeiro ...

(46e) L1

[...] pelo número quarenta por aí acho que foi essa revista aqui (né) tinha **essa capa muito bonita** eu acho **Ø** muito bonita éh:: eu fiquei decepcionado agora com a diretoria penúltima diretoria (né) que modificaram a capa da revista quer dizer é o problema do brasileiro não éh de de não conservar a tradição não éh... (Entrevista 5)

(46e') ... pelo número quarenta por aí acho que foi essa revista aqui (né) tinha **essa capa muito bonita** eu acho **ela** muito bonita...

(46f) L2 [...] É um cartão de visita ... eu acho que ... **ESSE cartão de visita que a gente TEM PELA USP** ((batendo sobre a mesa)) ... a gente deve ... saber valorizar **Ø** ((batendo)) ... e MU::Ita geração jovem ... num tá mais valorizando **Ø** ... porque acha que só é uma universidade como outra qualquer ... não é ... eu DEFENDO a USP com unhas e dentes ... TEM defeitos ... todos eles sanáveis ... (Entrevista 1)

(46f') ... **ESSE cartão de visita que a gente TEM PELA USP** ((batendo sobre a mesa)) ... a gente deve ... saber valorizar **ele** ((batendo)) ... e MU::Ita geração jovem ... num tá mais valorizando **ele** ...

(46g) L1 [...] vocês têm **muito tempo eu não tenho Ø...** eu tenho: / eu... eu:: trabalho no Parque de Ciência e Tecnologia da Universidade de São Paulo [...] (Entrevista 9)

(46g') L1 [...] vocês têm **muito tempo eu não tenho *ele...** eu tenho: / eu... eu:: trabalho no Parque de Ciência e Tecnologia da Universidade de São Paulo [...] (Entrevista 9)

(46h) L1 você é brasileira mesmo né?

L2 sou

[

L1 mas gosta de **outras línguas**

[

L2 ai **adoro Ø amo Ø de paixão** acho que se uma área...se eu não fosse da matemática eu ia gostar de fazer letras

(46h') ... mas gosta de **outras línguas** ...

... ai **adoro *elas amo *elas de paixão** ...

Após a análise dessas construções, observamos que as tendências se assemelham às de nossa pesquisa anterior (SIMÕES, 2015), já que as restrições ou impossibilidades de expressão do objeto acusativo pelo pronome lexical ocorrem entre os antecedentes menos referenciais, sobretudo com os SNs quantificados, e nas construções com predicado de estado. No âmbito dos SNs quantificados, observamos também que o pronome lexical não seria possível em construção imperfectiva e receberia uma interpretação [+específica] em construção de predicação secundária.

Conforme vimos ao longo da análise, as tendências encontradas demonstram que os objetos nulos estariam altamente implementados na variedade de São Paulo do PB contemporâneo. Ao comparar com a frequência de elipses nos estudos de Duarte (1986) e Cyrino (1994), nosso resultado de objetos nulos quase categórico poderia sugerir uma ampliação na incidência dessa variante.

Contudo, no estudo de Othero e Spinelli (2019), no qual se investigou a elipse na variedade do PB de Porto Alegre, não se observou uma ampliação da elipse, uma vez que foi encontrada pouca diferença em sua frequência na amostra da década de 1990 e na coletada entre 2015 e 2018. Esse resultado levou os autores a concluírem que se trata de uma variação estável. Além disso, a ausência de pronomes lexicais nas entrevistas que analisamos poderia ser decorrente da elevada escolaridade dos informantes, pois, de acordo com Duarte (1986), essa variante seria estigmatizada e apresentaria uma baixa frequência na produção de falantes escolarizados.

5 Interpretação teórica

Nesta seção, proporemos uma análise para a derivação das orações de objeto anafórico acusativo na variedade de espanhol de Madri e na variedade do PB de São Paulo, tendo em vista os desenvolvimentos minimalistas (CHOSMKY, 1999, 2000, 2004) e outros trabalhos gerativistas. Essa análise é a que propomos em Simões (2015) para explicar o objeto acusativo anafórico nas variedades de espanhol de Madri e Montevidéu e no PB. Entendemos que a mesma análise pode ser proposta para os estudos atuais na medida em que as tendências encontradas são semelhantes.

Com respeito às construções nas quais o objeto é retomado por um pronome, apoiando-nos nas análises de Groppi (1997) e Galves (2001), propomos que seria desencadeada a operação de movimento, uma vez que o predicado verbal selecionaria o traço-EPP. Sendo assim, o clítico verificaria e apagaria esse traço adjuntando-se a *v*, enquanto o pronome lexical o faria no especificador de *Sv*.

No âmbito das construções nas quais o objeto se expressa pela elipse, em Simões (2015), encontramos evidências de que os objetos nulos na variedade de espanhol de Madri teriam uma natureza pronominal, assim como os objetos nulos no PB (GALVES, 2001). Portanto, essa elipse corresponderia a um *pro*. Estendemos essa análise à nova amostra que analisamos.

Em relação às elipses que têm como antecedente um SN sem determinante, tendo em vista a proposta de Chomsky (1999) de que a elevação visível do objeto seria favorecida pelos SNs definidos e específicos e os indícios das línguas escandinavas (THRÁINSSON, 2008), propomos que o elemento pronominal sem realização fonética permaneceria na posição na qual foi gerado, e, portanto, ocorreria apenas a operação de concordância.

Considerando-se essas propostas, o problema estaria em verificar se nas construções em que a elipse tem como antecedente um SN [+determinado] ocorreria ou não a operação de movimento. Para tanto, formulamos duas hipóteses. Na primeira, o *v* selecionaria o traço- EPP e seria desencadeada a operação de movimento. Na segunda, o *v* não selecionaria o traço- EPP, de modo que se manifestaria apenas a operação de concordância.

A primeira hipótese poderia ser sustentada com base na proposta de Galves (2001), já que essa autora defende que o objeto nulo pronominal também seria legitimado em adjunção ao vP. Outro argumento que sustentaria essa hipótese é o de Cardinaletti e Starke (1994) de que a categoria pronominal sem realização fonética corresponderia a um pronome deficiente que deve mover-se. O terceiro argumento seria a proposta de Chomsky (1999) de que a tendência seria os objetos definidos e [+específicos] se elevarem de maneira visível, que corresponderia à operação de movimento. Entretanto, como analisaríamos as elipses que têm como antecedente um SN [+determinado; -específico]?

No que se refere à segunda hipótese, nosso primeiro argumento para sustentá-la constitui a proposta de Chomsky (1999) de que os traços- ϕ se manifestariam de maneira visível nos casos em que também ocorre o movimento visível de um SN. Tendo em vista os desenvolvimentos mais recentes do minimalismo (CHOMSKY, 2000) e a ideia de que os clíticos seriam elementos que manifestam concordância, conforme sugere Chomsky (1999), se a concordância não se manifesta de maneira visível entre o verbo e o argumento, ocorreria apenas a operação de concordância. No entanto, se a concordância se manifesta de maneira visível, seria desencadeada também a operação de movimento, de modo que o predicado verbal selecionaria o traço-EPP.

Tendo em vista essa proposta, observamos que, na esfera do objeto, o pronome corresponderia ao elemento de concordância que se manifesta de maneira visível entre o verbo e o objeto. Por outro lado, nas construções em que o *pro* constitui o argumento do verbo, a concordância não se manifestaria de maneira visível, pois os traços- ϕ do argumento não teriam manifestação fonética. Portanto, ocorreria apenas a operação de concordância.

Em relação ao segundo argumento, de acordo com Chomsky (2004), a geração de um segundo especificador seria determinada no léxico. Sendo assim, se um núcleo seleciona o traço-EPP, seria gerado um segundo especificador e seria desencadeada a operação de movimento. Tendo em vista essa proposta de Chomsky (2004), o que as tendências de nossos estudos poderiam revelar?

Sendo o predicado verbal que se encarrega ou não da seleção do traço-EPP, em nossa pesquisa mais recente, na variedade de espanhol de Madri, a elipse foi favorecida nas construções com perífrase verbal e verbos de estado. Além disso, as elipses foram mais frequentes em

construções com aspecto [-perfectivo], que expressa eventos não delimitados, e em construções intensionais, que são interpretadas como contextos não factuais. Na variedade do PB de São Paulo, na qual os objetos nulos estão altamente implementados, verificamos que a elipse foi categórica nas construções com perífrase verbal, que tendem a ser contextos de aspecto [-perfectivo]. Há evidências, portanto, de que existiria uma relação entre os predicados que denotam eventos não acabados e/ou não factuais e a não seleção do traço-EPP.

Constatamos que há uma gradação na possibilidade de expressão do objeto pelo clítico ou pelo pronome lexical que está relacionada com a referencialidade do antecedente e com o predicado verbal. Sendo assim, enquanto na variedade de espanhol de Madri o clítico retoma SNs [-animados; -específicos], definidos, indefinidos ou quantificados, no PB, há restrições para a expressão de antecedentes menos referenciais pelo pronome. Nossa atual pesquisa revelou que essas restrições se intensificam entre os SNs quantificados e o pronome não seria possível se a construção apresenta um verbo de estado. Mais uma vez essa tendência nos revela a atuação do predicado verbal e do tipo de determinante na possibilidade de manifestação da concordância. Conforme Brucart (1999, p. 2805), no espanhol, não seria possível a expressão do objeto pelo clítico nos casos em que o antecedente constitui um SN sem determinante e a construção apresenta o verbo estativo *tener*, como em *¿Tienes cerillas? No tengo Ø*.

O terceiro argumento para sustentar nossa segunda hipótese constitui os resultados dos trabalhos de Fujino e Sano (2002) e Ezeizabarrena (1997) sobre a aquisição do espanhol como língua materna.

Fujino e Sano (2002) verificaram que, durante o processo de aquisição do espanhol, ocorre um período no qual as crianças omitem o objeto, embora a elipse diminua conforme o clítico se torna mais frequente.

Em sua pesquisa, Ezeizabarrena (1997) observou que as crianças começaram a apresentar os clíticos de maneira produtiva após os dois anos de idade. De acordo com essa autora, tanto os morfemas flexivos quanto os clíticos constituiriam morfemas de concordância. Sendo assim, ela propõe que os clíticos seriam a realização morfológica de Conc-O e conclui que a presença dos clíticos na produção das crianças sugere que a categoria funcional Conc já estaria disponível em sua gramática.

Nossa ideia é que esses resultados sobre a aquisição do espanhol como língua materna revelariam que haveria uma relação entre a realização

dos clíticos e a categoria funcional Conc-O, que corresponde a Sv nos desenvolvimentos minimalistas mais recentes (CHOMSKY, 2000).

Quanto ao último argumento, nossa proposta é que, nas construções em (47) e (48), a agramaticalidade se explicaria devido à ausência de um pronome para expressar um antecedente [+animado; +específico]. Sugerimos que isso ocorre porque o *v* teria selecionado o traço-EPP, mas, se, ao invés do pronome, o *pro* é selecionado, que não apresenta traços fonéticos e não é capaz de verificar e apagar o traço-EPP, as derivações não seriam convergentes.

(47) **María** llegó hace poco. Juan dijo que **la** / ***Ø** traje.

(48) **A Maria** chegou faz pouco tempo. O João disse que trouxe **ela** / ***Ø**. (Adaptado de Simões 2015: 378)

Tendo em vista os argumentos que apresentamos para defender a segunda hipótese, consideramos que esta seria a hipótese mais plausível. Portanto, nossa proposta de análise é que, nas construções que apresentam objetos nulos com antecedente que constitui um SN [+determinado], o predicado verbal não selecionaria o traço-EPP, de modo que não ocorreria a operação de movimento, mas apenas a de concordância. Os traços- ϕ de *pro* se encarregariam de identificar o antecedente do objeto.

As tendências encontradas em nossa pesquisa e a análise teórica que propomos sugerem que haveria uma coexistência de gramáticas (CHOMSKY, 1999; LIGHTFOOT, 1999) na variedade de espanhol de Madri e no PB, que se manifesta como variação linguística. Sendo assim, de acordo com o mecanismo proposto por Chomsky (1999), uma das gramáticas corresponderia à expressão do objeto por um pronome e a outra corresponderia à expressão pela elipse.

6 Considerações finais

As tendências encontradas revelam que, na variedade de espanhol de Madri, os objetos nulos foram favorecidos pelos SNs sem determinante, os quantificados, os [-animados] e pelas construções com perífrase verbal, as que não apresentam predicação secundária e as construções com verbos de estado. Esses resultados demonstram a atuação da referencialidade do antecedente, do traço semântico de indefinidade do determinante, da ausência de determinante e da estrutura e aspecto lexical do predicado verbal na possibilidade de objetos nulos.

Os objetos nulos ocorreram também em construções com verbos cognitivos, de comunicação, aspecto imperfectivo, construções intensionais, objeto indireto e tópico, que são contextos que favorecem as elipses no espanhol falado no País Basco (LANDA, 1993, 1995) e em Quito (SUÑER; YÉPEZ, 1988).

Quanto à amostra da variedade de São Paulo do PB contemporâneo, observamos que os objetos nulos ocorreram com antecedentes com os traços semânticos de indefinidade, definitude [-animados] e [+/-específicos], assim como também com os antecedentes [+animados].

No que se refere aos demais contextos linguísticos, observamos que os objetos nulos estão altamente implementados e foram categóricos nas construções com verbos de comunicação, aspecto perfectivo, objeto indireto, perífrase verbal e tópicos.

Na esfera do pronome lexical, observamos uma parcela considerável de construções gramaticais com SNs indefinidos [-animados] e [-específicos], sendo que as maiores restrições ocorreram entre os SNs quantificados.

Em relação aos demais contextos, o pronome foi melhor aceito nas construções perfectivas, com verbo dinâmico e predicação secundária e apresentou baixa aceitação nas construções com predicado estativo, de comunicação, com objetos indiretos e tópicos, sendo estas últimas construções os contextos que permitem as elipses nas variedades de espanhol do País Basco (LANDA, 1993, 1995) e de Quito (SUÑER; YEPEZ, 1988).

Considerando-se a análise de algumas construções, observamos que o pronome lexical não é aceito com SNs quantificados que apresentam predicados estativos, e há restrições com verbos dinâmicos em construções imperfectivas, tendências semelhantes às encontradas em nossa pesquisa anterior (SIMÕES, 2015).

Os resultados de nossa pesquisa para ambas as línguas analisadas refletem a Hierarquia Referencial de Cardinaletti e Starke (1994), uma vez que, no espanhol de Madri, as elipses são favorecidas pelos antecedentes menos referenciais, enquanto no PB estes constituem os antecedentes que menos aceitam a expressão pelo pronome lexical, já que, nessa língua, as elipses estão altamente implementadas, podendo realizar-se até mesmo com antecedentes mais referenciais.

As tendências encontradas nos permitiram propor que, nas construções com objetos nulos, seria desencadeada apenas a operação de concordância e seriam os traços- ϕ de *pro* que identificariam o antecedente do objeto.

Referências

- BRUCART, J. M. La elipsis. In: BOSQUE, I.; DEMONTE, V. (orgs.). *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa, 1999. p. 2787-2866.
- CAMPOS, H. Indefinite object drop. *Linguistic Inquiry*, Massachusetts, v. 17, n. 3, p. 354-359, 1986.
- CAMPOS, H. Transitividad e intransitividad. In: BOSQUE, I.; DEMONTE, V. (orgs.). *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa, 1999. p. 1519-1574.
- CARDINALETTI, A.; STARKE, M. The typology of structural deficiency: On three grammatical classes. *Working Paper in Linguistics*, Venice, v. 4, n. 2., p. 41-109, 1994.
- CASAGRANDE, S. Restrições de ocorrência do objeto direto anafórico no português brasileiro: gramática adulta e aquisição da linguagem. *ReVEL*, n. 6, p. 131-163, 2012.
- CASAGRANDE, S. *A correlação entre aspecto e objeto no PB: uma análise sintático-aquisicionista*. 2010. 301 f. Tese (Doutorado em Linguística) – IEL, UNICAMP, 2010. Disponível em: <https://bv.fapesp.br/pt/publicacao/76422/a-correlacao-entre-aspecto-e-objeto-no-pb-uma-analise-sint/>. Acesso em: 01 fev. 2022.
- CESTERO MANCERA, A. M.; MOLINA, I.; PAREDES, F. *La lengua hablada en Madrid. Corpus PRESEEA - Madrid (Distrito de Salamanca)*. V. I – Hablantes de instrucción superior. Alcalá de Henares: Universidad de Alcalá, 2012.
- CHOMSKY, N. *Lectures on Governing and Binding*. Dordrecht: Foris, 1981.
- CHOMSKY, N. *Knowledge of Language: its nature, origin and use*. New York: Praeger, 1986.
- CHOMSKY, N. *O Programa Minimalista*. Lisboa: Caminho, 1999.
- CHOMSKY, N. Minimalist inquires. In: MARTIN, R.; MICHAELS, D.; URIAGEREKA, J. (eds.). *Step by step*. Cambridge: MIT Press., 2000. p. 89-155.
- CHOMSKY, N. Beyond explanatory adequacy. In: BELLETTI, A. (ed.). *Structures and beyond*. Oxford: Oxford University, 2004. p. 104-131.

CYRINO, S. Observações sobre a mudança diacrônica no português do Brasil: objeto nulo e clíticos. In: ROBERTS, I.; KATO, M. (orgs.). *Português brasileiro: Uma viagem diacrônica*. Campinas: Editora da Unicamp, 1993. p. 163-184.

CYRINO, S. *O objeto nulo no português do Brasil: um estudo sintático diacrônico*. 1994. 217 f. Tese (Doutorado em Linguística) – IEL, Unicamp, 1994.

CYRINO, S.; DUARTE, M. E.; KATO, M. Visible subjects and invisible clitics in Brazilian Portuguese. In: NEGRÃO, E.; KATO, M. (orgs.). *Brazilian Portuguese and the null subject parameter*. Madrid, Frankfurt: Iberoamericana, Vervuert, 2000. p. 55-73.

DEMONTTE, V.; MASULLO, P. J. La predicación: los complementos predicativos. In: BOSQUE, I.; DEMONTTE, V. (orgs.). *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa, 1999. p. 1461-2524.

DI TULLIO, Á. *Manual de gramática del español*. Buenos Aires: Edicial, 1997.

DUARTE, M. E. *Variação e sintaxe: clítico acusativo, pronome lexical e categoria vazia no português do Brasil*. 1986. 73 f. Dissertação (Mestrado em Linguística Aplicada e Estudos da Linguagem) – Faculdade de Filosofia, Comunicação, Letras e Artes, PUC/São Paulo, 1986.

EZEIZABARRENA, M. J. Morfemas de concordancia con el sujeto y con los objetos en el castellano infantil. In: PÉREZ-LEROUX, A. T.; GLASS, W. (eds.). *Contemporary perspectives on the acquisition of Spanish*. Somerville: Casacadilla, 1997. p. 21-36.

FERNÁNDEZ SORIANO, O. El pronombre personal. Formas y distribuciones. Pronombres átonos y tónicos. In: BOSQUE, I.; DEMONTTE, V. (orgs.). *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa, 1999. p. 1209-1273.

FUJINO, H.; SANO, T. Aspects of the null object phenomenon in child Spanish. In: PÉREZ-LEROUX, A. T.; LICERAS, J. (eds.). *The acquisition of Spanish morphosyntax. The L1/L2 connection*. Dordrecht/Boston/London: Kluwer, 2002. p. 67-88.

GALVES, C. *Ensaio sobre as gramáticas do português*. Campinas: Ed. da Unicamp, 2001.

GROPPI, M. *Pronomes pessoais no português do Brasil e no espanhol do Uruguai*. 1997. 152 f. Tese (Doutorado em Filologia e Língua Portuguesa) – FFLCH, USP, 1997.

GROPPI, M. Estructuras con clíticos: revisión de terminología y datos del español. *Signo y Seña*, Buenos Aires, 20, p. 95-113, 2009. DOI: <https://doi.org/10.34096/sys.n20.5801>

KATO, M. Pronomes fortes e fracos na sintaxe do Português Brasileiro. *Revista Portuguesa de Filologia*, Coimbra, v. 20, p. 101-122, 2002.

KATO, M. Null objects and VP ellipsis in European and Brazilian Portuguese. In: QUER, J. et al. (eds.). *Romance Languages and Linguistic Theory*. Amsterdam: John Benjamins, 2003. p. 135-158.

KATO, M. A gramática do letrado: questões para a teoria gramatical. In: MARQUES, M. A., et alii (orgs.). *Ciências da Linguagem: trinta anos de investigação e ensino*. Braga: CEHUM (Universidade do Minho), 2005. p. 131-145.

KATO, M.; TARALLO, F. (orgs.). *Harmonia trans-sistêmica: variação intra- e interlinguística*. Campinas: R.G., 1989.

LABOV, W. *Padrões Sociolinguísticos*. São Paulo: Parábola, 2008.

LACA, B. Presencia y ausencia de determinante. In: BOSQUE, I.; DEMONTE, V. (orgs.). *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa, 1999. p. 891-928.

LANDA, A. *Los objetos nulos determinados del español del País Vasco*. *Linguística*, n. 5, p. 131-146, 1993.

LANDA, A. *Conditions on null objects in Basque Spanish and their relation to leísmo and clitic doubling*. 1995. 252 f. Tese (Doutorado em Linguística) - University of Southern California, 1995.

LEONETTI, M. El artículo. In: BOSQUE, I.; DEMONTE, V. (orgs.). *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa, 1999. p. 787-890.

LIGHTFOOT, D. *How new languages emerge*. New York: Cambridge, 2006.

LIMA-HERNANDES, M. C.; VICENTE, R. (orgs.). *A Língua Portuguesa falada em São Paulo: amostra da variedade culta do século XXI*. São Paulo: Humanitas, 2012.

MAGALHÃES, T. M. V. *O sistema pronominal sujeito e objeto na aquisição do Português Europeu e do Português Brasileiro*. 2006. 175 f. Tese (Doutorado em Linguística) – IEL, Unicamp, 2006. Disponível em: <<http://www.bibliotecadigital.unicamp.br/document/?code=vtls000385238>>. Acesso em: 21 jul. 2021.

OTHERO, G.; SPINELLI, A. C. Um tratamento unificado da omissão e da expressão de sujeitos e objetos diretos pronominais de 3ª pessoa em português brasileiro. *Cadernos de Estudos Linguísticos*, Campinas, v.61, p. 1-30, 2019. DOI: <https://doi.org/10.20396/cel.v61i1.8654211>

SIMÕES, A. M. *O objeto pronominal acusativo de 3ª pessoa nas variedades de espanhol de Madri e Montevideu comparado ao português brasileiro*: clíticos como manifestação visível e objetos nulos como manifestação não visível da concordância de objeto. 2015. 396 f. Tese (Doutorado em Letras) - DLM/FFLCH, USP, 2015. Disponível em: <<https://www.teses.usp.br/teses/disponiveis/8/8145/tde-09092015-175408/pt-br.php>>. Acesso em: 01 fev. 2022

SUÑER, M.; YÉPEZ, M. Null definite objects in Quiteño. *Linguistic Inquiry*, Massachusetts, v. 14, p. 561-565, 1988.

TARALLO, F. Diagnosticando uma gramática brasileira: o português d'aquém e d'além mar ao final do século XIX. In: ROBERTS, I.; KATO, M. A. (orgs.). *Português brasileiro: Uma viagem diacrônica*. Campinas: Editora da Unicamp, 1993. p. 69-105.

THRÁINSSON, H. Object shift and scrambling. In: BALTON, M.; COLLINS, C. (eds.). *The handbook of contemporary syntactic theory*. Oxford: Blackwell, 2008. p. 148-202.

WEINREICH, U.; LABOV, W.; HERZOG, M. *Fundamentos empíricos para uma teoria da mudança lingüística*. São Paulo: Parábola, 2009.



O pronome pleno de terceira pessoa no “pretuguês” oitocentista

The third Person Full Pronoun in the ‘Pretuguês’ Nineteenth Century

Fernanda de Oliveira Cerqueira

Universidade Federal da Bahia (UFBA), Salvador, Bahia / Brasil

f.cerqueira@hotmail.com / f.cerqueira@ufba.br

<https://orcid.org/0000-0002-2515-9371>

Resumo: O presente trabalho visa mapear o comportamento do pronome pleno de terceira pessoa, por meio de sua distribuição sintática e suas relações referenciais, através de suas leituras léxico-semânticas, partindo da Teoria de traços- ϕ (CARVALHO, 2008, 2017; CERQUEIRA, 2015, 2019, 2020b; HARBOUR; ADGER; BÉJAR, 2008), em atas produzidas, na Salvador do século XIX (OLIVEIRA, 2006), a fim de verificar quais características desse fenômeno já se manifestavam nessa modalidade do Português Brasileiro oitocentista. Diante disso, para compreensão de pronome como elemento decomponível, a pesquisa se pauta na perspectiva gerativista (CHOMSKY, 1995 e posteriores), cujas diretrizes norteiam as variáveis linguísticas selecionadas: posição sintática do pronome, leituras de definitude, de especificidade e de animacidade, além de gênero e de número. Quanto ao enquadramento sociohistórico, o método quantitativo possibilitou a verificação de caminhos para compreensão do retrato da variante plena de terceira pessoa, na comunidade linguística em questão (LABOV, 1972). Portanto, verificou-se que o pronome pleno de terceira pessoa no pretuguês oitocentista soteropolitano é favorecido em posição de sujeito (59,78%), embora já ocorresse em posição de objeto direto (2,11%) e de objeto preposicionado (8,98%). Tende a apresentar leitura definida (94,7%), específica (98,94%) e animada (75,13%). Também ocorre com maior frequência como masculino (81,48%) e singular (90,57%), aspectos fortemente atrelados ao gênero textual ata, assim como ao perfil da comunidade em questão.

Palavras-chave: Pronome pleno de terceira pessoa; pretuguês; Sociedade protetora dos desvalidos; distribuição sintática; referência léxico-sintática.

Abstract: The present work aims to map the behavior of the third person full pronoun, through its syntactic distribution and its referential relations, through its lexical-semantic readings, starting from the ϕ -feature theory (CARVALHO, 2008, 2017; CERQUEIRA, 2015, 2019, 2020b; HARBOUR; ADGER; BÉJAR, 2008), in minutes produced in 19th century Salvador (OLIVEIRA, 2006), in order to verify which characteristics of this phenomenon were already manifested in this modality of 19th century Brazilian Portuguese. Therefore, in order to understand the pronoun as a decomposable element, the research is guided by the generativist perspective (CHOMSKY, 1995 and later), whose guidelines guide the selected linguistic variables: syntactic position of the pronoun, readings of definiteness, specificity and animacy, in addition to gender and number. As for the socio-historical framework, the quantitative method enabled the verification of ways to understand the portrait of the full third-person variant, in the linguistic community in question (LABOV, 1972). Therefore, it was found that the third person full pronoun in nineteenth-century Soteropolitan Pretuguês is favored in subject position (59.78%), although it already occurred in the direct object position (2.11%) and in the prepositional object (8,98%). This pronoun tends to present a defined reading (94.7%), specific (98.94%) and animated (75.13%). It also occurs more frequently as masculine (81.48%) and singular (90.57%), aspects strongly linked to the textual genre minute, as well as to the profile of the community in question.

Keywords: Third person full pronoun; pretuguês; Society for the protection of the disabled; syntactic distribution; lexical-syntactic reference.

Recebido em 25 de fevereiro de 2022

Accepto em 11 de abril de 2022

1 Introdução

O presente trabalho¹ objetiva mapear o comportamento do pronome pleno de terceira pessoa, considerando sua distribuição sintática e suas relações referenciais, através de suas leituras léxico-semânticas,

¹ A primeira versão deste trabalho foi apresentada em comunicação, no V Congresso de Linguística Histórica, em homenagem às professoras Mary Kato e Charlotte Galves; a segunda versão foi apresentada em conferência, no Terceiro Encontro de Gramática Gerativa, em homenagem às professoras Sônia Cyrino e Maria Eugênia Duarte, ambas em 2021.

partindo da Teoria de traços- ϕ (CARVALHO, 2008, 2017; CERQUEIRA, 2015, 2019, 2020b; HARBOUR; ADGER; BÉJAR, 2008), em atas produzidas, na Salvador do século XIX (OLIVEIRA, 2006), a fim de verificar quais características desse fenômeno, ilustrado a seguir, já se manifestavam nessa modalidade do Português Brasileiro (PB) oitocentista, a qual tratamos aqui como “pretuguês”², como proposta por Gonzalez (1983, 1988).

- (1) aí Pre guntou o Senhor Prizidente ao dicto Socio se **elle** tinha vindo na Seicaõ passada respondeo o Socio que não (p. 581).
- (2) o escriturario tendo em seu puder quando lhe a prezentou foi por baixo do feixe da dita a cta escripto um artigo 6º. e algumas palavras que **elle** 2º. Secretario diz não ter dito (p. 658).
- (3) por que en outras suedade tem sifeito por Cotas do Socios então eu não axho bom sibulir bem si bolinerte bom bouilir <nelle> (p. 801).
- (4) uando sahia desua caza sahia só e acompanhado Com Estatutos I Rigimento para Fallar en Regra com u Artigo **delle** (p. 832).

Os dados de (1) a (4) apresentam ocorrências do pronome pleno de terceira pessoa em diferentes posições sintáticas e com diferentes estatutos referenciais, os quais serão melhor discutidos em suas seções correspondentes.

Assim, a escolha pelo fenômeno, como já anunciado, decorre do interesse em compreender o comportamento dessa variante, tradicionalmente acionada como inovadora no que concerne à função acusativa (DUARTE, 1986; OMENA, 1978), na história da língua portuguesa no Brasil. Ademais, também houve intuito de verificar o comportamento desse pronome, nessas sincronias passadas, por meio de dados produzidos pelos agentes de difusão do Português Popular

² A justificativa sociohistórica do presente conceito, bem como sua conceituação mais madura, será realizada na seção 2.

Brasileiro (PPB)³, a saber, a população negra afrodescendente (MATTOS E SILVA, 2004a). Se, por um lado, Galves, Paixão de Souza e Namiutti (2006) defendem que no século XVIII surgem no PB estruturas distintas daquelas verificadas em Português Europeu (PE), revelando um cenário de competição de gramáticas (GALVES, 2012). Por outro lado, mas na mesma direção, Mattos e Silva (2000, 2004a, 2004b) e Lucchesi (2006, 2009, 2015) defendem a importância sociohistórica de episódios nos oitocentos para pavimentação da polarização sociolinguística do PB. Logo, diálogos entre a estrutura linguística e a estrutura social mostram-se reveladores à intelecção de muitas singularidades de PB.

Com efeito, o referencial teórico metodológico dessa pesquisa considerou a Linguística Gerativa, mas não só, uma vez que, como o próprio título sugere, a intenção aqui é de apontar como a estrutura social do Brasil oitocentista contribui para fenômenos identificados na estrutura linguística, na direção do já sinalizado por Pinto e Andrade (2019).

Ao colocar a aquisição da linguagem como lugar central para mudança linguística, especialmente para mudança gramatical, o gerativismo mantém seu caráter cognitivista/mentalista dos estudos da linguagem. Contudo, abre uma porta de diálogo extremamente profícuo com as ciências sociais no sentido de compreender quais são os possíveis gatilhos para a mudança linguística e como ela acontece (PINTO; ANDRADE, 2019, p. 53).

Não por acaso, de acordo com Mufwene (1997, 2002, 2007), as línguas humanas nascem, crescem e morrem, na condição de espécies, tendo em vista que precisam estar hospedadas em seus falantes para garantia de sua vitalidade, ou seja, sua possibilidade de sobrevivência e/ou resistência frente a condições de dominação e/ou prestígio socioeconômico. Nesse sentido, toda mudança é motivada externamente⁴, ao passo que há internamente uma recombinação gramatical⁵, nos termos

³ Para maiores informações acerca da constituição do PPB, ver Mattos e Silva (2004a).

⁴ Embora saibamos que há outras perspectivas, tais como a deriva secular (SAPIR, 1949 [1921]) ou encadeamento de marcações paramétricas, quer por predições (LIGHTFOOT, 1999), quer por microparâmetros (ROBERTS, 2019), nesta etapa da pesquisa, não nos debruçaremos em suas particularidades.

⁵ A recombinação gramatical é fruto hipótese da “gramática híbrida” (ABOH, 2015), na qual se advoga em função de que mudanças paramétricas, na Gramática Universal (GU), são acionadas pela diversidade de *inputs* verificadas em situações de contato multilíngue.

de Aboh (2015). Portanto, de acordo com esses autores, a variação paramétrica pode ser motivada pela estrutura social, em contextos de contato linguístico, o que parece ser o caso do Brasil.

Por sua vez, De Graff (1994) e Roberts (1997) constataram que a ausência de clítico acusativo é uma característica presente em línguas advindas de contato linguístico, isto é, o *pool* linguístico formado, em contexto multilíngue, instaurado pela dominação e exploração territorial e humana, forneceu diferentes *inputs* para a formatação da Gramática Universal (GU) em uma nova Língua Interna (LI). Com efeito, dá-se a constituição de uma gramática híbrida, em que há diferentes valores para parâmetros e/ou microparâmetros (ROBERTS, 2019), a qual permite, por exemplo, relações de vinculação, de *Move*, de *Agree* e de *Concord* distintas daquelas verificadas nas línguas alvo ou língua de colonização.

Em confirmação, os trabalhos de Câmara Júnior (1972), Omena (1978), Tarallo (1983), Duarte (1986), Galves (1984), Nunes (1996) e Cyrino (1994) demonstram a proeminência de pronomes tradicionalmente nominativos comportarem-se como pronomes objeto e oblíquo, podendo concorrer com suas formas nulas⁶ em PB, ilustrado de (5) a (8).

- (5) a. Maria encontrou **ele**.
b. Maria encontrou **ec**.
- (6) a. Comprei esse livro para/prá **nós / a gente**.
b. Comprei esse livro **ec**.
- (7) a. Por isso que eu gosto de **você / tu**.
b. Por isso que eu gosto **ec**.
- (8)a. A vó **dele** é um doce.
b. A vó **ec** é um doce.

Em acordo com a literatura supracitada, tais resultados sugerem que a entrada dos pronomes “você” e “a gente” no paradigma pronominal

⁶ Nesta etapa do trabalho, não serão exploradas as categorias vazias, de modo que seu estatuto não será discutido. Além disso, cabe destacar que, nas sentenças de (5) a (8) tais categorias vazias são legíveis, quando seu referente é apresentado contextualmente, condição para seu licenciamento (CHOMSKY, 1995).

do PB, bem como a redução do paradigma de concordância verbal, acompanhado, conseqüentemente, pelo sincretismo⁷ das formas plenas sejam fruto da oferta de múltiplos *inputs*, na formação do PB, por meio do contato linguístico, no bojo da colonização.

Sendo assim, para compreensão de pronome como elemento decomponível, composto por traços formais de pessoa, de gênero e de número, a pesquisa se pauta na perspectiva gerativista (CHOMSKY, 1995 e posteriores), com base em Cardinaletti e Starke (1999), Déchaine e Wiltschko (2002), Harley e Ritter (2002), Cowper e Hall (2002), Béjar (2003), Carvalho (2008, 2017) e Marcotulio (2017), cujas diretrizes norteiam as variáveis linguísticas selecionadas para descrição do objeto de investigação, a saber: posição sintática do pronome, leituras de definitude, de especificidade e de animacidade, além de gênero e de número.

Já para o seu enquadramento sociohistórico do fenômeno, o trabalho valeu-se de alguns elementos da abordagem sociolinguística, ao passo que o método quantitativo possibilitou a verificação de caminhos para compreensão do retrato da variante plena de terceira pessoa, na comunidade linguística em questão (LABOV, 1972), ainda que o enfoque dessa etapa do trabalho não seja a variação propriamente dita, mas a descrição da variante plena, a fim de obter pistas de sua natureza léxico-sintática.

Portanto, para cumprimento do objetivo previsto, o trabalho se divide nas seguintes seções: i. O pretuguês e o *corpus*, em que serão apresentados o conceito de pretuguês, os marcadores sociohistóricos e a constituição do *corpus* selecionado; ii. O pronome pleno de terceira pessoa, cujas subseções são distribuição sintática e referência léxico-sintática, em que serão apresentados os dados; iii. Uma possível análise, em que se pretende testar a proposta de Cerqueira (2019) adotada para tratamento da terceira pessoa plena em PB contemporâneo; iv. Conclusões; e Referências.

2 O pretuguês e o *corpus*: breve sociohistória

Como efeito do colonialismo, bem como do seu desdobramento em sistemas de opressões de raça, de gênero, de classe e de sexualidade, estipulados por meio da hierarquização de território e de sujeitos

⁷ Emergência de pronomes tradicionalmente nominativos em outras posições casuais (GALVES, 1994, 2001; KATO, 1999; NUNES, 1996).

(KILOMBA, 2019), há, desde o contato entre o PE com as línguas africanas e indígenas, particularidades no que hoje concebemos como PB.

Tendo isso em vista, Gonzalez (1979, 1983, 1988) propõe que a modalidade linguística falada no Brasil, pelas camadas populares, seria, conforme proposto pela autora, uma evidência para a categoria político-cultural de amefricanidade. Assim, em se tratando mais precisamente do PPB, a autora o denomina como pretuguês, tomando por base “[...] as marcas de africanização do português falado no Brasil” (GONZALEZ, 1988, p. 54). Por isso, destaca que

[é] engraçado como eles gozam a gente quando a gente diz que é *Framengo*. Chamam a gente de ignorante dizendo que a gente fala errado. E de repente ignoram que a presença desse *r* no lugar do *l*, nada mais é que a marca linguística de um idioma africano, no qual o *l* inexistente. Afinal, quem é o ignorante? Ao mesmo tempo, acham o maior barato a fala dita brasileira, que corta os *erres* dos infinitivos verbais, que condensa *você* em *cê*, *está* em *tá* e por aí afora. Não sacam que tão falando pretuguês (GONZALEZ, 1983, p. 238).

Em outras palavras, pretuguês pode ser concebido como a modalidade linguística falada por sujeitos negros na formação do PB, bem como seus desdobramentos no PPB, sua modalidade contemporânea, a qual tende a acarretar estigmas (CERQUEIRA, 2020a, 2021a, 2021b).

Considerando o expressivo contingente de populações negras, de etnias diversas, submetidos ao regime de escravização que atuou como principal pilar econômico do Brasil, muito antes do capital, Gonzalez (1988) propõe a existência de uma categoria político-cultural com base na verificação intensa de referências culturais desses sujeitos na constituição de elementos diversos da cultura brasileira. Essa categoria política-cultural intitulada “Améfrica” atua como um sistema etnográfico de referência ancestral africana por meio de signos fenotípicos, musicais, estéticos, filosóficos, gastronômicos, religiosos e, sem dúvida, linguísticos. Logo, o pretuguês é acionado pela antropóloga como uma forte evidência empírica para fundamentação de tal categoria política-cultural.

Nesse sentido, fica cristalino que tal proposição se pauta na confluência entre estrutura linguística e estrutura social, considerando tanto a atuação das línguas africanas na sociohistória do PB – cuja

intervenção resultou uma nova língua/variedade⁸ de português; quanto à história da educação da população negra no pós-abolição, cujos acessos foram ora negados, ora comprometidos (ALBUQUERQUE, 2004; CERQUEIRA, 2020a; GOMES, 2012), resultando em sistemas de performance significativamente distintos dos produzidos pelos falantes da norma dita culta no Brasil oitocentista (LUCCHESI, 2015).

Com efeito, o *corpus* selecionado para esse trabalho, faz parte de *corpora* composto por atas de assembleias deliberativas da Sociedade Protetora dos Desvalidos (SPD) – Primeira irmandade negra do Brasil, fundada em 1832, instalada em 1851 e regulamentada pelo governo em 1861. Está, até os dias de hoje, situada no Centro histórico de Salvador, mais precisamente, no Pelourinho, espaço-território famoso àquela época por ser palanque para espetacularização do açoite contra sujeitos negros escravizados, muito embora atualmente seja ressignificado como importante ponto turístico da cidade.

Segundo Campos (2018), a SPD referendou-se como uma associação civil negra cujo objetivo foi a colaboração mútua de membros em casos como tratamento de doenças, auxílio funeral, suporte para invalidez, aprisionamento e velhice, além de assistência, através de pensões, para os familiares dos sócios, e educação de órfãos. Diante disso, o autor defende que a SPD apresentou papel preponderante diante dos cenários de efervescência dos movimentos abolicionistas e de crise do escravagismo no Brasil.

Os *corpora* contam com 290 atas produzidas por homens negros livres ou alforriados, dentre as quais, 6 autores eram africanos e 12 autores eram brasileiros. Contudo, além desses, presume-se que 1 autor provavelmente era africano e 11 provavelmente eram brasileiros. Os documentos são datados entre 1850 e 1870. Esses documentos obtiveram tratamento filológico em edição semidiplomática, realizada pelo Prof. Dr. Klebson Oliveira, publicizadas em sua tese de doutorado, defendida em 2006, pelo Programa de Pós-Graduação em Letras e Linguística da Universidade Federal da Bahia. Assim, o *corpus* selecionado para essa pesquisa conta com 122 atas produzidas exclusivamente por brasileiros, haja vista que há interesse em investigar pistas diacrônicas da natureza do pronome pleno de terceira pessoa nesses documentos.

⁸ Para algumas abordagens, PB e PE são línguas distintas (KATO, 2012). Para outras, são variedades de uma mesma língua (COSTA, 2012).

3 O pronome pleno de terceira pessoa

Tendo em vista a ampla gama de abordagens nos estudos formais da linguagem humana, entendemos como necessário destacar que a concepção de pronome feita aqui está para além da classe de palavra ou do item lexical cristalizado. Portanto, de acordo com os postulados do Programa Minimalista, conforme proposto por Chomsky (1995 e posteriores), pronomes são concebidos como unidades referenciais nas quais se encontram um conjunto de traços- ϕ , a saber, traços formais de pessoa, gênero e número.

No que tange à natureza estrutural do pronome, muitos trabalhos já argumentam, desde a década de 90, a possibilidade de subdivisão da categoria pronome:

- a. tanto em termos de decomposição, decorrente da ausência de camadas funcionais mais complexas em sua projeção, visitamos os trabalhos de Cardinaletti e Starke (1999), Kato (1999, 2002) e Déchaine e Wiltschko (2002); e
- b. quanto em termos de subespecificação, decorrente da carência de traços formais na composição do pronome em um léxico pré-sintático ou na primeira lista de um léxico fragmentado, visitamos os trabalhos de Harley e Ritter (2002), Cowper e Hall (2002), Béjar (2003), Carvalho (2008, 2017), Marcotulio (2017) e Nunes (2020).

Diante disso, assumimos que pronome seja uma categoria funcional composta por traços formais equivalentes aos de seu referente nominal, isto é, a forma do pronome pode ou não refletir sua sintaxe (CERQUEIRA, 2019, 2020b; CERQUEIRA; CARVALHO, 2020), uma vez que esse elemento não é visto como primitivo sintático, mas como produto da sua notação de traços formais.

Considerando que o pronome pleno de terceira pessoa comporta-se, em PB, como pronome fraco podendo ser mais subespecificado do que os de primeira e de segunda pessoa, é nosso interesse investigar seu comportamento nessas sincronias passadas, já que o sincretismo do

pronome pleno de terceira pessoa é um fenômeno inovador na formação do PB (GALVES, 2001; KATO, 1999), o que afeta as suas condições de referência/retomada (CYRINO, 1994; DUARTE, 1995; GALVES, 1984).

Ademais, também pressupomos encontrar um número considerável de ocorrências desse pronome no *corpus* selecionado, considerando que a retomada pronominal é uma estratégia de coesão produtiva no gênero textual ata, devido à sua capacidade de referenciar com maior objetividade autores, tópicos e eventos de fala, realizando a tessitura do texto (MARCUSCHI, 2008).

3.1 Distribuição Sintática

Quanto à distribuição sintática, foram encontradas ocorrências de pronome pleno de terceira pessoa em diferentes posições, conforme se pode ver nos dados de (9) a (16).

- (9) [...] O Soçio Caetano diz que **elle** não se importava com apartes pois não lhe per tu bava aoracaõ (p.649)
- (10) [...] o Senhor presidente do Conselho pelos Senhores Socios, na occasião em que lia o seu relatorio, dizendo que n'**elle** continham palavras offensivas aos mesmos (p. 719)
- (11) [...] por tanto de cá elle nos perdemos **elle** inos requeremos ugoverno (p. 802) [Referente: o dinheiro]
- (12) [...] pede a palavra o Senhor Socio Severiano Pedro, dizendo que elle como amigo da ordem, e da paz, que retirava a palavra escandalo, que tinha sido pronunciada por **elle** na sessão passada (p. 719)
- (13) [...] o 1º Secretario declara que o fisiando ao Senhor Visente Torquarto dos Santos anomiação **delle** pelo (p. 796)
- (14) [...] pedio palavra O Socio, Amaro Berinque disse eu não vou de em Contro ao trabalho da Sociedadade pois haxo munto bom mais o que eu quero e hum escraricimento este Socio Felecianno nao esta a pozentado e como e que **elle** <e> chamado para vim pagar (p. 628-629)

- (15) [...] depois da explicação foi posto o artigo a votos, sendo aprovado; entra o artigo 24º, usa da palavra sobre **elle**, o Senhor Socio Francisco Paraizo (p. 727)
- (16) [...] **elle** o Thezoureiro seriaõ Capaz para essa Comissaõ efoi lidas as respostas do Soçio Juvençio e Grigorio, e do Soçio Joaõ Theodorio da Solidade (p. 654)

Os dados de (9) a (16) representam algumas das ocorrências do pronome pleno de terceira pessoa em diferentes posições sintáticas. Em (9), está em posições de sujeito; em (10), de adjunto; em (11), de objeto direto; em (12), de objeto preposicionado; em (13), de complemento nominal; em (14), de clivagem⁹; em (15), com c-modificação¹⁰; e, em (16), como tópico¹¹. Assim, a emergência do pronome pleno de terceira pessoa em todas essas posições sintáticas reitera seu caráter estrutural híbrido, como proposto por Kato (1999), uma vez que é licenciado em posições restritas tanto a pronomes fortes, quanto a fracos.

A seguir, tabela representa a quantidade de ocorrências de pronome pleno de terceira pessoa, encontrados no *corpus* em questão.

⁹ Em que há leitura de foco contrastivo, confirmando o já previsto por Pinto e Ribeiro (2008), nas línguas românicas.

¹⁰ C-modificação é a modificação de todo o DP, decorrente de dominância, por meio de c-comando, cujo elemento modificador está projetado em SpecD. Tal teste, bem como sua nomenclatura, é proposto por Cardinaletti e Starke (1999), sendo uma das etapas para verificação de restrição de pronomes deficientes – fracos/clíticos. Nesse sentido, o que os autores assumem como c-modificação é a relação estrutural na projeção, não a classe projetada (A, P, N, V, Adv).

¹¹ De acordo com Cardinaletti e Starke (1999), ser licenciado em posição de tópico é comportamento esperado para pronomes fortes.

Tabela 1 – Distribuição sintática

Posição sintática	Ocorrências	Percentual
SUJ Matriz	51	26,98%
SUJ Encaixada	41	21,69%
SUJ Interpolado (inversão de SUJ)	2	1,05%
SUJ Posposto	12	6,34%
SUJ Inf Encaixada	6	3,17%
SUJ SC	1	0,52%
Adj. Adn.	10	5,29%
Adj. Adv.	8	4,23%
OD	4	2,11%
OI	5	2,64%
Rel. ou Circuns.	9	4,76%
Comp. Loc.	3	1,58%
Comp. Nom.	7	3,70%
TOP	6	3,17%
Clivagem	2	1,05%
C-modificação	19	10,05%
Reflexivo	1	0,52%
Total	189	100,00%

Conforme previsto, há, nessa sincronia passada, uma predileção pelas posições de sujeito, cujo percentual é de 59,75 %, quando os diferentes tipos de sujeito são vistos como unidade. Em acordo com Galves (2012), interpretamos que tal resultado corresponde ao fato de que, no século XIX, havia forte competição de gramáticas em PB, de modo que o sincretismo das formas plenas ainda era pouco emergente, se comparado, por exemplo, com o seu comportamento atual (CARVALHO, 2008, 2017). Logo, embora haja ocorrência de pronome pleno de terceira pessoa em posições de argumento¹², de adjunto e de periferia, suas ocorrências ainda são pouco expressivas se comparadas com as

¹² Destaca-se o caráter inovador do pronome pleno em posição de objeto direto, sobretudo, sob viés do contato linguístico. Sob esse prisma, a emergência do PB decorre da influência africana, especialmente das línguas de origem Bantu, cujos pronomes sujeito emergem em posição acusativa (AVELAR; GALVES, 2014).

modalidades contemporâneas de PB (CERQUEIRA, 2019). Todavia, destaca-se caráter inovador de seu sincretismo, com destaque para as posições de objeto direto (2,11%) e objetos preposicionados (8,98%).

Contudo, acreditamos que dois eventos possam ter colaborado com esse resultado, a saber: a. a formalidade do gênero textual ata, haja vista que o grau de monitoramento é acionado diante da modalidade escrita, bem como de sua função comunicativa escrita; e b. a ainda forte influência de Portugal sobre o Brasil no século XIX, com destaque para o âmbito linguístico. Tendo isso em vista, há intuito de, em momento oportuno, verificar o comportamento desse pronome em outros *corpora* da mesma natureza, a fim de cruzar tais resultados com os aqui apresentados.

Ademais, os resultados verificados nesse *corpus*, cujos produtos não se assemelham tanto aos de Duarte (1995), oriundos de peças teatrais sudestinas, sugerem que os segmentos racial e socioeconômico dos sujeitos produtores desses documentos informa a modalidade de PB oitocentista em uso, em confluência com a polarização sociolinguística (LUCCHESI, 2006, 2015; MATTOS E SILVA, 2004a, 2004b).

3.2 Referência Léxico - Sintática

Quanto à referência desse pronome, componente léxico-sintática responsável pela leitura semântica, assumimos, respectivamente, com Hertenberg (2015) e Partee (2004), que o traço de definitude aciona leitura identificável, enquanto o traço de especificidade resulta em leitura particular. Com efeito, também concebemos que a categoria pessoa é acarretada não só pelos traços de falante e ouvinte, mas por definitude e por especificidade ambos dominados por determinação (CERQUEIRA, 2019, 2020b; CERQUEIRA; CARVALHO, 2020; CERQUEIRA; MARIANO, 2020).

Tendo isso em vista, essa etapa da pesquisa teve por intuito descrever as condições de especificação ou subespecificação do pronome em questão, haja vista que, conforme a teoria de traços- ϕ , traços formais são aqueles cuja presença na notação da categoria afeta sua distribuição, por meio das operações de *merge* e *agree*, à medida que há identidade total ou parcial desses traços (BEJAR, 2008).

Assim, houve, como previsto, ocorrência de pronome pleno de terceira pessoa com leituras definida, em (17); específica, em (18); animada, em (19); inanimada, em (20); masculina, em (21); feminina, em (22); singular, em (23); e plural, em (24), as quais podem ser concomitantes.

- (17) [...] vortando o vizitador deu parte de todo o Corido e como Se tinha **elle** havido Com a dicta viu va (p. 554)
- (18) [...] aí Pre guntou o Senhor Prizidente² ao dicto Socio se **elle** tinha vindo na Seicaõ passada respondeo o Socio que não (p. 581)
- (19) [...] pede a palavra o Socio Terensio Dantas, dando uma explicação, terminando dizendo que, desde a ultima Assemblea, que **elle** pediu a palavra em nome da commissão (p. 740)
- (20) [...] o Prizidenti declara aberta açeção iprocedisi aleitura da Açta da ceccão anterior ina Prizidenti poz avotos não avendo quem uzasi dapalvra soubri **ella** deu po aprovada (p. 792)
- (21) [...] Senhor Sococio Thezoreiro deiopor conto do dinheiro que **elle** tomor inpretado de 5\$000 res. (p. 803)
- (22) [...] Ficou a diado para 1^o. sessão Thomou, possa o Senhor Manuel Leonardo e disse o Dito Senhor que o Prezidente deveria dar A pauta os Entencilio por **ella** Foi Feita a chamada do Dito Entencilho (p. 834)
- (23) [...] concluindo pede, que se mande o Senhor presidente do Conselho, redigir o seu Relatorio, retirando algumas palavras offensivas, que n'elle contenham (p. 720)
- (24) [...] Porque udinheiro da Lotaria uguverno pode lancal amaõ nelle diz elles não faz acaza de azilho por tanto de cá elle nos perdemos elle inos requeremos uguverno (p. 802) – [Referente: membros do governo]

A seguir, tabela representa a quantidade de ocorrências do pronome pleno de terceira pessoa, segundo sua referência no *corpus* em questão.

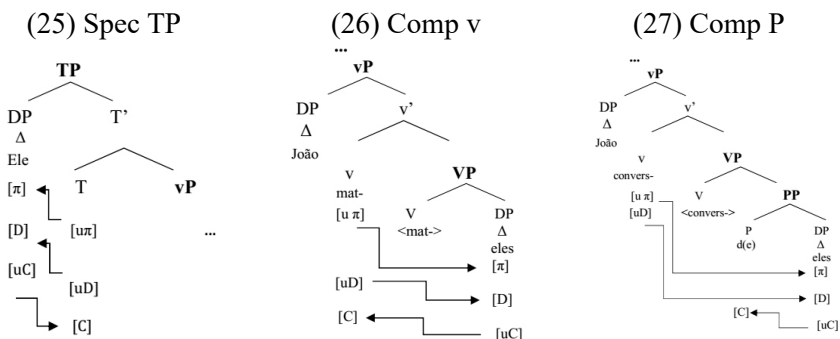
Tabela 2 – Referência

Relações referenciais		Ocorrências	Percentual
Pessoa	Definidos	179	94,70%
	Específicos	187	98,94%
Classe	Animados	142	75,13%
	Inanimados	47	24,87%
	Masculino	154	81,48%
	Feminino	35	18,52%
Número	Singular	171	90,57%
	Plural	18	9,43%

Como previsto, houve ocorrência de pronome pleno de terceira pessoa com leituras acarretadas por traços- ϕ dominados pelas categorias pessoa, classe e número, ao passo que houve taxas altas de definidos (94,7%), específico (98,94%), animados (75,13%), masculinos (81,48%) e singular (90,57%). Quanto à categoria pessoa, verificamos maior nível de referencialidade nessa modalidade histórica do PB, o que pode estar relacionado às condições de licenciamento do pronome pleno de terceira pessoa. Quanto às categorias classe e número, o perfil da comunidade, massivamente masculina, a situação comunicativa, assembleia, e o gênero textual, *ata*, parecem colaborativos para os resultados verificados.

4 Uma possível análise

Como foi possível observar, o pronome pleno de terceira pessoa tende a ocorrer em posições de Spec TP, Comp *v* e Comp P cujas representações arbóreas seguem, respectivamente, em (25), (26) e (27).



Assim, *merge* ocorre na condição de que haja identidade dos traços formais dos nós raiz da composição léxico-sintática do pronome, resultando na operação *agree*, no curso da derivação. Diante disso, as sondas buscam seus alvos a fim de satisfazer a valoração de traços, em seus pontos de fase correspondentes, como proposto em Chomsky (2005). No entanto, assumimos com Frapton e Gutmann (2000) e Costa (2012) que tais traços não sejam deletados, após a checagem, mas que entrem em estado de inércia para que possam ser lidos pelas interfaces articulatorio-perceptual e conceitual-intencional, após *spell-out*.

Esse caminho nos parece estratégico devido aos limites que autores com Béjar (2003, 2008), Cobert (2006) e Alexiadou (2009) estabelecem para as operações *agree*, estritamente sintática/computacional, e *concord*, de ordem morfossintática – portanto, afetada pelas interfaces, caso assumamos a existência de uma arquitetura da linguagem nos termos propostos por Hauser, Chomsky e Fitch (2002). Dito de outro modo, *match* é um requerimento sintático, cuja manifestação morfossintática pode configurar *mismatch*, como, segundo Aboh (2020), tende a ocorrer com línguas fruto de colonização.

Já o estabelecimento da referência pronominal é decorrente tanto da restrição estrutural, prevista no Princípio B da Ligação, quanto da identidade de traços- ϕ do pronome com seu antecedente expresso (quando ligado) ou com seu correferente (quando dêitico/logofórico).

Desse modo, a referência do pronome pleno de terceira pessoa é regulada pela cadeia- ϕ composta pela ligação entre o pronome e seu antecedente, em (28), ou entre o pronome e o operador léxico-semântico, em (29), que licencia sua correferência, respeitando o princípio da localidade.

- (28) [...] pediu palaver o Socio Bracete pedindo ao Senhor Presidente a entrega dos papeis pretencentes a [o Socio Martinianno]_a que **elle**_a queria ver e izaminar commo menbro da Comição de Contas (p. 634).
- (29) [OP_{Señor Francisco}]_b [...] quanto o homem le desse O dinheiro para **elle**_b entregar na Sociedade ninguen vio (p. 618).

Logo, a ligação entre o pronome pleno de terceira pessoa com seu antecedente expresso, em (28), ou com seu correferente, em (29), forma uma cadeia-φ para o estabelecimento de referência, já que atende às condições do Princípio da Homogeneidade (CARVALHO; BRITO, 2017; COLLINS; POSTAL, 2012) e de Minimalidade-φ (CARVALHO et al., 2018; NUNES, 2017; NUNES; MARTINS, 2017).

6 Conclusões

Considerando as discussões aqui realizadas, verificou-se que o pronome pleno de terceira pessoa no pretuguês oitocentista soteropolitano, como previsto, é favorecido em posição de sujeito (59,78%), apesar de já ocorrer em demais posições sintáticas, com destaque para as de objeto direto (2,11%) e objetos preposicionados (8,98%). Além disso, tende a apresentar leitura definida (94,7%), específica (98,94%) e animada (75,13%). Também ocorre com maior frequência como masculino (81,48%) e singular (90,57%), aspectos fortemente atrelados ao gênero textual ata, assim como ao perfil da comunidade em questão.

Logo, embora o sincretismo do pronome pleno de terceira pessoa já estivesse em curso nessa sincronia passada, sua incidência ainda é pouco expressiva se comparada com o comportamento atual desse pronome (CERQUEIRA 2019, 2020b; CERQUEIRA; CARVALHO, 2020), que tende a emergir em todas as posições casuais, mas também tende a apresentar uma geometria de traços mais especificada, nesses casos.

Ademais, quanto ao lócus referencial, os dados apontam que, nessa etapa de implementação do pronome pleno de terceira pessoa, quando há leitura animada, suas leituras também são, preferencialmente, definida e, categoricamente, específica, na direção da hierarquia referencial proposta por Cyrino, Duarte e Kato (2000), à medida que ora se comporta como pronome forte, em sua posição básica, ora comporta-se como pronome fraco, nas demais posições sintáticas.

Referências

ABOH, E. *The Emergence of Hybrid grammar: Language Contact, Change and Creation*. Cambridge: Cambridge University Press, 2015.

ABOH, E. Lessons From Neuro-(a)-Typical Brains: Universal Multilingualism, Code-Mixing, Recombination, and Executive Functions. *Front. Psychol*, Switzerland, v. 11, n. 488, p. 1-16, 2020. DOI: <https://doi.org/10.3389/fpsyg.2020.00488>

ABOH, E.; DEGRAFF, M. A. Null Theory of Creole Formation Based on Universal Grammar. In: ROBERTS, I. (ed.). *Oxford Handbooks: New York*, 2017. Disponível em: <<http://linguistics.mit.edu/wp-content/uploads/aboh-and-degraff-2017-null-theory-of-creole-formation.pdf>> Acesso em: 24 fev. 2022.

ALBUQUERQUE, W. *A exaltação das diferenças: racialização, cultura e cidadania brasileira (Bahia, 1880 - 1900)*. 2004. 250p. Tese (Doutorado em História) - Instituto de Filosofia e Ciências Humanas, Universidade do Estado de Campinas, 2004.

ALEXIADOU, A. On the role of syntactic locality in morphological processes: the case of (Greek) derived nominals. In: GIANNAKIDOU, A.; RATHER, M. (eds.). *Quantification, Definiteness and Nominalization*. Oxford University Press, 2009. p. 253-280.

AVELAR, J.; GALVES, C. M. C. O papel das línguas africanas na emergência da gramática do português brasileiro. *Linguística*, Rio de Janeiro, Linguística, v. 30, p. 241-288, 2014. Disponível em: < http://www.scielo.edu.uy/scielo.php?script=sci_arttext&pid=S2079-312X2014000200010&lng=es&tlng=pt > Acesso em: 24 fev. 2022

BÉJAR, S. *Phi-syntax: a theory of agreement*, 2003. 206p. Tese (Doutorado em Linguística) – University of Toronto, 2003.

BÉJAR, S. Conditions on phi-agree. In: HARBOUR, D.; ADGER, D.; BÉJAR, S. (eds.). *Phi Theory: Phi-Features across modules and interfaces*. New York: Oxford University Press, 2008. p. 130-154.

CÂMARA JR, J. M. Êle como um acusativo no português do Brasil. In: CÂMARA JR., J. M. *Dispersos*. Rio de Janeiro: Fundação Getúlio Vargas, 1972, p. 47 - 53.

CAMPOS, L. R. *Sociedade Protetora dos Desvalidos: mutualismo, política e identidade racial em Salvador (1861-1894)*. 2018. 149p. Dissertação (Mestrado em História) – Faculdade de Filosofia e Ciências Humanas, 2018.

CARDINALETTI, A.; STARKE, M. The typology of structural deficiency: a case of the three classes of pronouns. In: VAN RIEMSDIJK, H. (ed.). *Clitics in the language of Europe*. Berlin: Mouton et Gruyter, 1999. p. 41-109.

CARVALHO, D. S. *A Estrutura interna dos pronomes pessoais em português brasileiro*. 2008. 154p. Tese (Doutorado em Letras e Linguística) – Faculdade de Letras, UFAL, 2008.

CARVALHO, D. S. *The internal structure of personal pronouns*. London: Cambridge Scholars Publishing, 2017.

CARVALHO, D. S.; BRITO, D. B. S. Impostores, correferência e concordância em português brasileiro. *Revista Letras*, v. 96, p. 55-73, 2017. DOI: <http://dx.doi.org/10.5380/rel.v96i0.50421>

CARVALHO, D. S.; BRITO, BRITO, D.; SEDRINS, A. P. Referência de primeira pessoa e anáfora em português brasileiro. *ReVEL*, v. 6, n.30, p. 127-145, 2018. Disponível em: < https://www.researchgate.net/publication/323969485_REFERENCIA_DE_PRIMEIRA_PESSOA_E_ANAFORA_EM_PORTUGUES_BRASILEIRO> Acesso em 23 fev. 2022.

CERQUEIRA, F. O. Reflexos semânticos na sintaxe de terceira pessoa. *Letrônica* – Revista Digital do Programa de Pós-Graduação em Letras da PUCRS, Porto Alegre, v.8, n.2, p. 422-437, jul./dez., 2015. DOI: <https://doi.org/10.15448/1984-4301.2015.2.20384>

CERQUEIRA, F. O. *O pronome pleno de terceira pessoa: estrutura interna e relações referenciais*. 2019. 152p. Tese (Doutorado em Língua e Cultura) – Universidade Federal da Bahia, 2019.

CERQUEIRA, F. O. O pretuguês como comunidade de prática: concordância nominal e identidade racial. *Revista Traços de Linguagem*, Cáceres, v. 4, n. 1, p. 75-88, 2020a. DOI: <https://doi.org/10.30681/2594.9063.2020v4n1id4644>

CERQUEIRA, F. O. Pronomes pessoais: participante e determinação como componentes de referência. In: AZEVEDO, N. D. (org.). *Estudos interdisciplinares da linguagem* – v 01. Campina Grande: Realize Editora, 2020b. p. 95-114.

CERQUEIRA, F. O. Lélia Gonzalez e o Pretuguês: do racismo e sexismo ao epistemicídio. In: CARVALHO, D.; LIMA, P. E. (orgs.). *Linguagem, gênero e sexualidade*. Salvador: Edufba, no prelo.

CERQUEIRA, F. O.; CARVALHO, D. S. A configuração de referência nos pronomes de terceira pessoa em português. *Revista do GELNE*, Natal, v. 22, n2, p. 321-335, 2020. DOI: <https://doi.org/10.21680/1517-7874.2020v22n2ID23278>

CERQUEIRA, F. O.; MARIANO, V. C. Sobre a morfossintaxe da determinação nominal dos DPs em português. *Revista Estudos Linguísticos e Literários*. n. 68, p. 597-614, 2020. DOI: <https://doi.org/10.9771/ell.v0i68.38997>

CHOMSKY, N. *Programa Minimalista*. Tradução de Eduardo Raposo. Lisboa: Caminho, 1995.

CHOMSKY, N. *On phases*. Ms., 2005. Disponível em: <http://www.fossil.in/chomsky_phases.pdf> Acesso em: 24 fev. 2022.

COLLINS, C.; POSTAL, P. *Imposters: A study of pronominal agreement*. Cambridge, MA: MIT Press, 2012.

CORBETT, G. *Agreement*. Cambridge: Cambridge University Press, 2006.

COSTA, J. Variação PE-PB sem configuracionalidade discursiva: argumentos adicionais para a primazia da sintaxe. In: LOBO, T.; CARNEIRO, Z.; SOLEDADE, J.; ALMEIDA, A.; RIBEIRO, S. (orgs.). *Rosae: Linguística Histórica, História das línguas e outras histórias*. Salvador: Edufba, 2012, p. 109-121.

COWPER, E.; HALL, D. The syntactic manifestation of nominal feature geometry. In: *Proceedings of the 2002 Annual Conference of the Canadian Linguistic Association*. Montréal: Cahiers Linguistiques de l'UQAM, 2001. p. 55-66.

CYRINO, S. M. L. *O Objeto Nulo no Português do Brasil*. 1994. 217p. Tese (Doutorado em Ciências), Instituto de Estudos da Linguagem, UNICAMP, 1994.

DÉCHAINED, R. M; WITSCHKO, M. Decomposing Pronouns. *Linguistic Inquiry*, Cambridge v. 33, n.3. p. 409-442, 2002. Disponível em: <<https://faculty.georgetown.edu/rtk8/Dechaine%20and%20Wiltschko%202002%20decomposing%20pronouns.pdf>> Acesso em 20 fev. 2022.

DEGRAFF, M. *To move or not move?* Placement of verbs and Objects Pronouns in Haitian Creole and in French. In: BEALS, K.; DENTON, J.; KNIPPEN, R.; MELNAR, L; SUZUKI, H.; ZEILNFELD, E. (eds). *Papers from the 30th Meeting of the Linguistic Society University of Michigan*, 1994, p. 141 - 155.

DUARTE, M. E. L. *Variação e Sintaxe: clítico acusativo, pronome lexical e categoria vazia no português do Brasil*, 1986. Dissertação (Mestrado em Letras e Linguística). Universidade Católica de São Paulo, 1986.

DUARTE, M. E. L. *A perda do princípio “evite pronome” no português brasileiro*. 1995. 151p. Tese (Doutorado em Ciências), Instituto de Estudos da Linguagem, UNICAMP, 1995.

FRAMPTON, J.; GUTMANN, S. *Agreement is Feature Sharing*. Ms.: Northeastern University, 2000. Disponível em: <https://www.researchgate.net/publication/251639256_Agreement_is_Feature_Sharing> Acesso em 24 fev. 2022.

GALVES, C. M. C. Pronomes e Categorias Vazias em Português do Brasil. *Cadernos de estudos linguísticos*, Campinas: UNICAMP, v. 7, p. 107-136, 1984.

GALVES, C. M. *Ensaio sobre as gramáticas do português*. Campinas: UNICAMP, 2001.

GALVES, C. M. C. Periodização e competição de gramáticas: o caso do português médio. In: LOBO, T.; CARNEIRO, Z.; SOLEDADE, J.; ALMEIDA, A.; RIBEIRO, S. (orgs.). *Rosae: Linguística Histórica, História das línguas e outras histórias*. Salvador: Edufba, 2012. p. 65 - 74.

GALVES, C. M. C.; PAIXÃO DE SOUSA, M. C.; NAMIUTTI, C. T. Novas perspectivas para antigas questões: revisitando a periodização da língua portuguesa. In: ENDRUSCHAT, A.; KEMMLER, R.; SCHAFER-PRIEB, B. (orgs.). *Grammatische Strukturen des Europäischen Portugiesisch*. Tübingen: Calepinus Verlag, 2006. p. 45-74.

GOMES, N. L. Movimento Negro e Educação: ressignificando e politizando a raça. *Educ. Soc.*, Campinas, v. 33, n. 120, p. 717-744, 2012. Disponível em: <<https://www.scielo.br/j/es/a/wQQ8dbKRR3MNZDJKp5cfZ4M/abstract/?lang=pt>> Acesso em 20 fev. 2022.

GONZALEZ, L. Racism and its effects in Brazilian society. In: *Women's conference of human rights and mission*, Veneza, 24-30 jun. 1979. (Mimeo.).

GONZALEZ, L. Racismo e sexismo na cultura brasileira. *Revista Ciências Sociais Hoje*, ANPOCS, p. 223-244, 1983. Disponível em: <https://edisciplinas.usp.br/pluginfile.php/6969610/mod_resource/content/1/06%20-%20GONZALES%2C%20L%20A9lia%20-%20Racismo_e_Sexismo_na_Cultura_Brasileira%20%281%29.pdf> Acesso em 20 fev. 2022.

GONZALEZ, L. A categoria político-cultural de amefricanidade. *Tempo Brasileiro*, Rio de Janeiro, n. 92-93, p. 69-82, 1988. DOI: <https://doi.org/10.21057/10.21057/repamv15n1.2021.40454>

HARBOUR, D.; ADGER, D; BEJAR, S. *Conditions on phi-agree*. New York: Oxford University Press, 2008.

HARLEY, H.; RITTER, E. *Person and number in pronouns: a feature-geometric analysis*. *Language*, v. 78. p. 482-526, 2002. Disponível em: <<http://www.ai.mit.edu/projects/dm/harley-ritter-geometry.pdf>> Acesso em 24 fev. 2022.

HAUSER, M D; CHOMSKY, N.; FITCH, W. Tecumesh. The faculty of language: what is it, what has it, and how did it evolve?. *Science*, Washington, v. 298, p. 1569-1579, 2002. DOI: 10.1126/science.298.5598.1569

HERTEZENBERG, M. J. B. Third Person Reference in Late Latin: demonstratives, definite articles, and personal pronouns in the *Itinerarium Egeriae*. (*Library of Congress Cataloging-in-Publication Data*). De Gruyter Mouton: Berlin/Boston, 2015.

KATO, M. A. Strong and weak pronominals in the null subject parameter. *PROBUS – International Journal of Romance Linguistics*, v.11, p. 1-37, 1999. DOI: <https://doi.org/10.1515/prbs.1999.11.1.1>

KATO, M. O português são dois... ou três? In: LOBO, T.; CARNEIRO, Z.; SOLEDADE, J.; ALMEIDA, A.; RIBEIRO, S. (orgs.). *Rosae: Linguística Histórica, História das línguas e outras histórias*. Salvador: Edufba, 2012. p. 93 – 108.

KILOMBA, G. *Memórias de plantação: episódios de racismo cotidiano*. Rio de Janeiro: Conogó, 2019 [2008].

LABOV, W. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press, 1972.

LIGHTFOOT, D. *The development of language: acquisition, change and evolution*. Massachusetts, Oxford: Blackwell, 1999.

LUCCHESI, D. E. R. Parâmetros Sociolinguísticos do Português Brasileiro. *Revista da ABRALIN*, Cidade Aberta, v. 5, n. 1 - 2, p. 83-112, 2006. Disponível em: < <https://revista.abralin.org/index.php/abralin/article/view/941#:~:text=Com%20a%20formula%C3%A7%C3%A3o%20do%20conceito,%5B1%2D5%5D.>> Acesso em 20 fev. 2022.

LUCCHESI, D. E. R. A transmissão linguística irregular. In: LUCCHESI, D; BAXTER, A.; RIBEIRO, I. *O português afro-brasileiro*. Salvador: EDUFBA, 2009, p. 41-74.

LUCCHESI, D. E. R. *Língua e sociedades partidas: a polarização sociolinguística do Brasil*. São Paulo: Contexto, 2015.

MARCOTULIO, L. Sobre composicionalidade pronominal das formas de tratamento na história do português. In: CARVALHO, D (org.). *Traços-phi: contribuições para a compreensão da gramática do português*. EDUFBA: Salvador, 2017. p. 87-108.

MARCUSCHI, L. A. *Produção textual, análise de gêneros e compreensão*. São Paulo: Parábola, 2008

MATTOS E SILVA, R. V. Uma interpretação para a generalizada difusão da língua portuguesa no território brasileiro. *Gravatá*, Gravatá, n. 9, p. 11-27, 2000. Disponível em: < <https://periodicos.uff.br/gragoata/article/view/49033>> Acesso em 20 fev. 2022.

MATTOS E SILVA, R. V. *Ensaio para uma socio-história do português brasileiro*. São Paulo: Parábola Editorial, 2004a.

MATTOS E SILVA, R. V. O português no Brasil: sua formação na complexidade multilinguística do Brasil colonial e pós-colonial. *Leituras Contemporâneas*, Salvador, v. 1, n. 1, p. 95-105, 2004b. Disponível em: <<https://www.prohpor.org/anais>> Acesso em 22 fev. 2022.

MUFWENE, S. 'Jargons, pidgins, creoles, and koines: What are they?' In: SPEARS, A.; WINFOR, D. (eds.), *The structure and status of pidgins and creoles*, Amsterdam: Benjamins, 1997, p. 35–70.

MUFWENE, S. Pidgin and Creole Languages. In: SMERLSER, N. J.; BALTES, P. B. (Eds.) *International encyclopedia of the social and behavioral sciences*. Amsterdã: Elsevier, 2002.

MUFWENE, S. What does creoles and pidgins tell us about the evolution of language? In: LAKS, B.; CLEUZIQU, S.; DEMOULE, J. P.; ENCREVÉ, P. (eds.) *The origin and evolution of languages: approaches, models, paradigms*. London: Equinox, 2007.

NUNES, J. Direção de Cliticização, objeto nulo e pronome tônico na posição de objeto direto em português brasileiro. In: KATO, M.; ROBERTS, I (org.). *Português brasileiro: uma viagem diacrônica*. Campinas: UNICAMP, 1996. p. 207 - 222.

NUNES, J.; MARTINS, A. M. Subespecificação de Traços- ϕ em Infinitivos Flexionados e Variação Dialetal/Idioletal em Português. *Handout*. XI Ramania Nova. ALFAL XVIII. Bogotá, Colombia, 2017.

NUNES, J., MARTINS, A. M. Circumventing ϕ -minimality. On some unorthodox cases of A-movement in Brazilian Portuguese. In: LOPES, R.; AVELAR, J.; CYRINO, S. (eds.). *Romance Languages and Linguistic Theory 12*. Selected papers from the 45th Linguistics Symposium on Romance (LSRL), Campinas, Amsterdam/Philadelphia: John Benjamins, 2017. p. 159-184.

OLIVEIRA, K. *Negros e escrita no Brasil do século XIX: sócio-história, edição filológica de documentos e estudo linguístico*. 2006. 1198p. Tese (Doutorado em Letras) – Instituto de Letras, UFBA, 2006.

OMENA, N. P. *Pronome pessoal de terceira pessoa: suas formas variantes em função acusativa*. 1978. Dissertação (Mestrado em Linguística) - Faculdade de Filosofia, Comunicação, Letras e Artes, PUC, 1978.

PARTEE, B. Opacity, coreference, and pronouns. In: PARTEE, B (ed.). *Compositionally in Formal Semantics*. Malden: Blackwell, 2004. p. 26-49.

PINTO, C. F. C.; ANDRADE, A. L. de. Desmistificando a gramática gerativa como uma teoria associativa e a-histórica da mudança linguística. *Macabéa – Revista Eletrônica do Netlli*, v. 8, n.2, jul-dez, p. 36-66. 2019. Disponível em: < <https://core.ac.uk/download/pdf/230134115.pdf>> Acesso em 20 fev. 2022.

PINTO, C. F. C.; RIBEIRO, I. M. O. Um estudo sintático-discursivo comparativo da clivagem em línguas românicas. In: MOURA, D. (org.). *Os desafios da língua: estudos em língua falada e escrita*. Maceió: EDUFAL, 2008. p. 401-404.

ROBERTS, I. Creoles, Markedness and the Language Bioprogram Hypothesis. *Revista Estudos Linguísticos e Literários*, Salvador, p. 12 – 22, 1997.

ROBERTS, I. *Parameter hierarchies and Universal Grammar*. New York: Oxford University Press, 2019.

SAPIR, E. *Language – an introduction to the study of speech*. New York: Harcourt, Brace & World, Inc, (1949 [1921]).

TARALLO, F. *Relativization strategies in spoken Brazilian Portuguese*. 1983. Tese (Doutorado em Ciências) - University of Pennsylvania, 1983.



Resistência do dativo de primeira pessoa na batalha (quase) perdida dos clíticos pronominais do português brasileiro

The Survival of the First Person Dative Pronoun in the (Almost) Lost Battle of Pronominal Clitics of Brazilian Portuguese

Maria Aparecida Torres Morais

Universidade de São Paulo (USP), São Paulo, São Paulo / Brasil

torres.mariacida@gmail.com

<https://orcid.org/0000-0003-1012-3795>

Heloisa Maria Moreira Salles

Universidade de Brasília (UnB), Brasília, Distrito Federal / Brasil

heloisasalles@gmail.com

<https://orcid.org/0000-0002-3637-3453>

Resumo: Este estudo investiga a cisão pronominal na expressão dos dativos no português brasileiro (PB), a qual consiste na presença dos clíticos de 1ª e 2ª pessoas *me* e *te*, em oposição aos pronominais tônicos *ele(s)*, *ela(s)*, introduzidos por *a* e *para*, em decorrência da perda dos clíticos dativos *lhe(s)*. Em termos descritivos, focalizamos a resistência do clítico *me*, na sua ampla distribuição sintática em predicados ditransitivos e monotransitivos, com verbos inergativos e inacusativos, entre eles, os verbos psicológicos, e em estruturas com o dativo ético. Assumindo uma abordagem teórica da cisão com base na hipótese de um traço interpretável de pessoa nos núcleos funcionais aplicativos (cf. TORRES MORAIS; SALLES, 2010, 2019), propomos que opções microparamétrica nos valores do traço pessoa, de um sistema bivalente, para um sistema monovalente, causam que as propriedades sintáticas e semânticas do clítico dativo *me* sejam unificadas como núcleo Aplicativo Alto, nos termos da tipologia proposta em Pylkkänen (2002, 2008).

Palavras-chave: clítico dativo; cisão pronominal; aplicativo; oblíquo.

Abstract: The study investigates the pronominal split in the expression of datives in Brazilian Portuguese (BP), which consists in the use of first and second person clitics, as opposed to the third person strong pronoun in the complement position of the dative preposition (*para ele(s)/ ela(s)*), with the loss of the third person dative clitics (*lhe(s)*). The analysis argues in favor of the survival of the first person clitic *me* in view of its wide syntactic distribution, in ditransitive predicates, dynamic and stative, in monotransitive predicates, inergative and inaccusative, including psychological verbs, and in predicates expressing the ethical datives. Assuming an approach of the pronominal split in terms of an interpretable person feature in the applicative heads (cf. TORRES MORAIS; SALLES, 2010, 2019), we propose that a microparametric change in the values of the person feature, from a bivalent system to a monovalent one, causes that the syntactic and semantic properties of the dative clitic *me* be unified as a High Applicative head, in Pykkänen's (2002, 2008) typology.

Keywords: dative clitic; pronominal split; applicative; oblique.

Recebido em 01 de março de 2022

Aceito em 13 de abril de 2022

1 Considerações iniciais

Este estudo tem como objetivo central apresentar uma proposta de análise para o fenômeno da cisão que se estabeleceu no sistema pronominal do português brasileiro (PB) entre a realização clítica do objeto indireto (OI) de 1ª e 2ª pessoas, de um lado, e do objeto indireto (OI) de 3ª pessoa, de outro, exemplificada nas estruturas ditransitivas em (1a-c):

- (1) a. O João deu o livro *para ela*.
- b. O João *me* deu o livro/ João deu o livro *para mim*.
- c. O João *te* deu o livro/ João deu o livro *para você*.

Como se depreende dos dados (1a-c), o clítico *lhe* de 3ª pessoa está ausente, sendo realizado exclusivamente por um pronome pleno introduzido por preposição (preferencialmente *para*). No caso da 1ª e 2ª pessoas, porém, os clíticos coocorrem com as formas pronominais preposicionadas.

Vale ressaltar que a perda dos clíticos acusativos e dativos de 3ª pessoa tem sido amplamente documentada na literatura corrente, manifestando-se nos dados já a partir das primeiras décadas do século XX. Em particular, na perspectiva gerativista da aquisição da linguagem, a inclusão dos clíticos acusativos de 3ª pessoa no PB padrão é entendida como um desenvolvimento posterior à aquisição da gramática nuclear, creditando-se à escola o papel de resgatá-los, sob a pressão de restrições normativas (KATO et al., 2009; NUNES, 2019, entre outros). No entanto, a mesma afirmação não pode ser estendida para contemplar os clíticos de 1ª e 2ª pessoas, *me* e *te*. Partimos da hipótese de que a sua permanência no sistema pronominal não é um fenômeno residual, ou uma manifestação da gramática periférica, uma vez que os mesmos se mostram produtivos e espontâneos na língua falada e escrita.

A apresentação do texto está elaborada da seguinte forma. Na seção 2 trazemos resultados empíricos de estudos prévios, os quais revelam a perda dos clíticos dativos de 3ª pessoa, *lhe(s)*, bem como a reanálise da forma de tratamento *você(s)*, em variação com o clítico *te*, como pronome complemento no PB, em oposição ao PE. Fazendo um recorte que contempla o clítico *me*, a seção 3 descreve a sua ocorrência em estruturas ditransitivas, inergativas e inacusativas.¹ Incluímos, igualmente, a discussão dos contextos que licenciam o dativo ético, realizado quase exclusivamente como clítico de 1ª pessoa no PB. Por sua vez, a seção 4 traz uma proposta de análise da cisão pronominal, com base na teoria dos núcleos aplicativos altos e baixos, nos termos de Pylkkänen (2002, 2008). Propomos a atuação de opções microparamétricas nos valores do traço interpretável de pessoa do Aplicativo, os quais passam de um sistema bivalente para um sistema monovalente. Essa reanálise leva a que as propriedades sintáticas e semânticas do dativo *me*, no PB, estejam unificadas em todos os contextos verbais, pela ativação do núcleo Aplicativo Alto. Na seção 5, fazemos as nossas considerações finais.

¹ Nesta tipologia não incluímos os predicados incoativos e causativos, uma vez que manifestam propriedades de natureza sintática e semântica particulares, a serem tratadas em estudo futuro.

2 O sistema pronominal no PB: reanálise e uso de formas inovadoras

2.1 A perda dos clíticos de 3ª pessoa no PB

Estudos quantitativos sincrônicos e diacrônicos revelam que os clíticos acusativos de 3ª pessoa, *o(s)*, *a(s)*, e os clíticos dativos, *lhe(s)*, não mais fazem parte da gramática nuclear do PB contemporâneo. Essa perda ativa a emergência de várias outras estratégias pronominais na gramática inovadora do PB (cf. BERLINCK, 1996; CYRINO, 2018, 2019; GALVES, 2001; 2018; 2019; FREIRE, 2000, 2005; GOMES, 1996; 2003; RAMOS, 1992; TORRES MORAIS; BERLINCK, 2006, 2007, 2018; TORRES MORAIS; SALLES, 2010, 2019, entre outros).

Um corolário desse fenômeno, amplamente citado, é a perda da distinção entre o caso acusativo e dativo no sistema de clíticos pronominais. Em particular, verifica-se que a cisão no sistema pronominal se refere não só ao uso de clíticos na 1ª e 2ª pessoas, em oposição ao uso do pronome tônico/ pleno, na 3ª pessoa, mas também em relação à expressão de caso morfológico, restrita aos pronomes de 1ª e 2ª pessoas, enquanto, na 3ª pessoa, o caso oblíquo é expresso por meio da preposição, conforme ilustrado anteriormente. Essa cisão é discutida em Torres Morais e Berlinck (2018), pela retomada de vários estudos que investigam a distribuição do OI pronominal, em estruturas ditransitivas. As autoras apresentam a Tabela 1, extraída do estudo de Berlinck (1997), com dados de fala de estudantes universitários de Curitiba, os quais expressam o embate entre o uso do OI clítico e o uso do OI oblíquo (mediante a preposição), no contexto de verbos ditransitivos de transferência e movimento. A tabela evidencia, por um lado, o uso quase categórico do OI oblíquo (e a conseqüente perda do clítico dativo de 3ª pessoa), e, por outro, o uso dos clíticos de 1ª e 2ª pessoas (que varia com o OI oblíquo, num diferente estatuto percentual).

Tabela 1 - Expressão do OI pronominal na fala de universitários curitibanos

Tipos de OI/ Pessoa gramatical	OI clítico	OI oblíquo
1ª pessoa	84% (88/105)	16% (17/105)
2ª pessoa	77% (37/48)	23% (11/48)
3ª pessoa	2% (1/56)	98% (55/56)
Total	60% (126/209)	40% (83/209)

Fonte: adaptada de Berlinck (1997, apud. TORRES MORAIS; BERLINCK, 2018, p. 261).

Torres Morais e Berlinck (2018, p. 263) mencionam outros estudos com dados de falantes do PB de diferentes regiões, em que se evidenciam as mesmas tendências.² Também relevante para a presente discussão é o uso das preposições *a* e *para* introdutoras do OI oblíquo, conforme destacam as autoras:

Berlinck [1996] avaliou o uso de preposições no OI oblíquo na fala de jovens universitários curitibanos, observando o emprego quase categórico da preposição *para* (92%; 187 em 192 dados). Resultado semelhante é encontrado em Gomes (2003) em dados da fala de cariocas, de 1980 a 2000.

A discussão considera, em seguida, o percurso histórico da expressão pronominal do OI, em que se confirma o declínio gradativo do uso do clítico de 3ª pessoa, a culminar no cenário sincrônico acima descrito. Analisando cartas do século XIX e do século XX, estas últimas com textos produzidos em várias cidades do país (reunidas no acervo dos corpora do PHPB³), as autoras destacam a emergência, no século XX, de uma distinção importante:

² Referimos ainda os resultados de Pereira (2019) com dados coletados de corpus do português rural de Goiás e de Minas Gerais, que atestam a vitalidade do pronome *me*, bem como a do pronome *te*, ainda que com frequência menor do que a forma oblíqua *para você(s)*.

³ PHPB, referente a Projeto para a História do Português Brasileiro.

[a] 1ª e a 2ª pessoa tu são expressas por OI clítico (...); por outro lado, a 3ª pessoa apresenta o predomínio do OI oblíquo e a 2ª pessoa você aparece nas cartas pessoais e nas cartas/ anúncios das revistas como outro contexto comparativamente mais propício à forma preposicionada do OI (TORRES MORAIS; BERLINCK, 2018, p. 269).

Os dados em (2) e (3), a seguir, extraídos de cartas pessoais do século XX, ilustram a ocorrência do OI de 3ª pessoa na estrutura preposicionada, com os pronomes *você* e *ele*, destacando ainda o papel do tipo sintático-semântico do verbo – “(...) as construções com verbos de transferência material (como *dar*, *enviar*) emergem como o contexto que menos retém o clítico” (TORRES MORAIS; BERLINCK, 2018, p. 270).

(2)[20, 1 CP BA] Achei Ivo um “amoreco” mando *para ele* um beijo.

(3)[20, 1 CP RN] L., peço desculpa por eu ter demorado tanto a *lhe* escrever e depois de tudo dar essa resposta tão ruim *para você*.
(TORRES MORAIS; BERLINCK, 2018, p. 270)⁴

As autoras observam ainda que o clítico *lhe*, como exemplificado em (3), “não desapareceu completamente, mas foi realocado” (p. 299). Elas verificaram que, nos dados dos anos 60, “apenas 31% apresentam referência de 3ª pessoa”, enquanto “nos anos 1990 (...) os poucos casos de *lhe* de 3ª e 2ª pessoas parecem estar associados a uma interpretação de maior formalidade” (p. 300). Além da especialização como OI de 2ª pessoa mais formal, o pronome *lhe* também ocorre como OD de 2ª pessoa, em substituição aos clíticos acusativos *o(s)* e *a(s)* (14). As autoras concluem que “[e]ssa extensão de uso pode ser vista como uma parte do processo geral de perda de distinções morfológicas, em progresso, pelo menos, desde o século XX” (p. 300).

2.2 A alternância entre o clítico *te* e a forma forte *você*

Em contraponto ao cenário de cisão entre 1ª e 2ª pessoas, em oposição à 3ª pessoa, encontra-se a alternância entre o clítico *te* de 2ª pessoa e o uso do pronome forte *você* na estrutura oblíqua (introduzido pela preposição *para*). Os estudos sobre o chamado ‘voceamento’ no PB têm sido reveladores em relação a essa alternância e à formação do sistema pronominal, conforme demonstram os resultados de Lopes e Cavalcante (2011), e obras ali citadas. As autoras observam que “a

⁴ Com a numeração original alterada.

entrada de *você* não ocorreu na mesma velocidade em todo o quadro de pronomes (pessoais retos, oblíquos átonos e tônicos, possessivos)” (p. 11). E acrescentam: “[o] clítico dativo de 3ª pessoa *lhe* também perdeu espaço para sintagmas preposicionados introduzidos pelas preposições *a* ou *para* (esta mais produtiva do que aquela no português brasileiro falado, como demonstram vários estudos)” (p. 12). Referindo-se ao dialeto carioca e citando o estudo de Gomes (2003), mencionado anteriormente, as autoras destacam ainda o caráter variável do sistema na 2ª pessoa, decorrente da entrada do pronome *você*, no seguinte paradigma: “nominativo (*você/ tu*); acusativo (*te/lhe/você*); dativo (*te, lhe/a~para você*)” (p. 12).

De fato, o estudo de Lopes e Cavalcante (2011) aprofunda a questão do uso variável de *tu* e *você*, buscando estabelecer uma cronologia do ‘voceamento’. Na discussão, as autoras consideram as formas variáveis em contexto de complementação, que nos interessam em particular, observando que “as formas relacionadas ao pronome *tu* foram as mais produtivas na documentação relativa aos fins do século XIX” (p. 44), conforme ilustrado em (4).

- (4) “Estimando que tudo \emptyset *encontres* a teu gosto, peço-*te* que desculpes a demora.” (Carta de Alberto – família Cupertino)
(LOPES; CAVALCANTE, 2011, p. 44).⁵

Partindo de estudos prévios (RUMEU, 2008; LOPES, 2009 *et. al* 2009.), as autoras investigam cartas pessoais produzidas de 1870 a 1937, tendo em vista a hipótese de que a entrada do pronome *você* está vinculada a esse período. Em relação ao uso das formas *tu* e *você* na posição de sujeito, elas verificam que, nesse período, em termos globais, “o pronome de segunda pessoa *tu* é mais produtivo (60%) em relação a *você* (40%)” (p. 46), considerando-se a sua realização como pronome pleno ou na desinência verbal (embora a forma relacionada a *tu* predomine como sujeito nulo).⁶ No entanto, partir dos anos 1920, os dados evidenciam a superação do uso de *tu* pelo pronome *você* na posição de sujeito, com

⁵ Com a numeração original alterada.

⁶ Temos aqui um ponto importante a ser considerado na avaliação das inovações diacrônicas envolvendo as formas *tu* e *você*, qual seja: o alto percentual da forma *tu* no PB desse período histórico (1870-1937), na função de sujeito, não se refere realmente à sua realização lexical, mas à sua expressão na flexão verbal, como morfema de 2ª pessoa do singular: *compraste, escreveste, soubeste, etc.* Essa particularidade evidencia claramente o estatuto de língua de sujeito nulo do PB do século XIX e primeiras décadas

79% de ocorrências. As autoras destacam que fatores sociopragmáticos contribuíram para a entrada de *você*, como atenuadores de pedidos e ordens, e o uso em expressões fixas. Elas destacam também o aumento no uso do pronome pleno *tu*, alinhando-se ao uso de *você*, que já ocorria preferencialmente na forma plena.

Na posição de complemento, as autoras distinguem as formas pronominais em verbos monotransitivos e verbos ditransitivos, com papel semântico de alvo, fonte ou beneficiário e traço [+animado] de um lado, e pronominais preposicionados de outro. Dessa forma, manifesta-se a variação *te~a/para você/ti~zero~lhe*, enquanto o complemento oblíquo exclui o clítico. Em relação aos resultados globais, os índices de *te* ficam próximos dos 50%, seguindo-se o dativo nulo (20%) e o *lhe* (14%), enquanto os sintagmas preposicionados apresentam índices baixos: *para você* (4,3%), *a você* (3,4%) e *para ti* (2,4%). Os dados a seguir ilustram ocorrências desses complementos citados pelas autoras.

- (5) a. Agradeço-*te* desde já a fineza (...)
 b. O dicionário serviu bem e mamãe mandou pelo Tio agradecer *a você* (...)
 c. Já tive notícia de ter João recebido os 100\$, o que de novo *lhe* agradeço (...)
 d. Agradeço Ø muito o cuidado que *você* tem dito de mim.
 e. Um afetuoso abraço para *os teus* (...) e *para voce* minha santa que mandarei
 f. (...) este amor que só *a ti* dedico (...)
 (LOPES; CAVALCANTE, 2011, p. 51)⁷

Considerando a correlação entre as formas de tratamento na posição de sujeito e o uso das formas pronominais de complemento dativo (e também acusativo), as autoras concluem que existem três subsistemas: nas cartas com supremacia de *tu* como sujeito, o clítico *te* é mais frequente; nas cartas com a mistura dos pronomes *tu* e *você*, ocorre o predomínio da forma nula, que pode ser relacionada a ambas as formas; nas cartas em que prevalece *você*, a forma clítica *te* é a preferida.

do século XX, no sentido em que a flexão verbal continha os morfemas pronominais referentes às pessoas discursivas.

⁷ Com a numeração original alterada.

O aumento de *você* vincula-se ao declínio de *lhe*, que se manifesta conforme descrito a seguir.

Nas últimas décadas do século XIX (1870-1899), nota-se certo equilíbrio entre os clíticos *te* e *lhe* com índices percentuais próximos de 50%. A partir da primeira década do século XX, há uma acentuada diminuição de que se mantém com índices bem baixos nas décadas subsequentes. (LOPES; CAVALCANTE, 2011, p. 58)

Conclui-se que a relação entre a entrada do pronome *você* na posição de sujeito e o uso de *te* em contexto de complementação é o padrão que caracteriza o sistema pronominal do PB. Consequentemente, como as autoras observam, a relação entre *você* (sujeito) e *te* (complemento) não constitui um sistema misto, uma vez que a gramaticalização de *você* como pronome pessoal de 2ª pessoa permite a identificação com o pronome de 2ª pessoa *te*, o que indica uma identificação em termos dos traços semânticos.

Gostaríamos finalmente de salientar que os resultados obtidos em cartas pessoais têm relevância para o presente estudo, em dois pontos: (i) o uso do clítico *te*, na 2ª pessoa, confirma a cisão referida anteriormente, ou seja, o alinhamento com o clítico *me* de 1ª pessoa, em oposição à perda das formas clíticas na 3ª pessoa; (ii) o uso variável das formas *te* e *você*, este último introduzido pelas preposições *a* e *para* nos contextos dativos, corroboram a força do paradigma *me-mim-comigo* da 1ª pessoa, uma vez que o mesmo não se alterou ao longo da história do PB. Por fim, fica registrado que a forma *te* ainda se encontra em um cenário de variação, desta vez com o clítico *lhe*, reanalisado como pronome para se referir à 2ª pessoa pessoa do singular.

2.3 O sistema pronominal do PB em perspectiva comparada com o PE: o caso da gramaticalização da forma tratamental *você*

Neste ponto, seria interessante uma breve digressão comparativa com o sistema pronominal do português europeu padrão (PE) Conforme mencionado anteriormente, o pronome *você* passa a ocupar o campo do pronome *tu*, no PB, a partir de um uso gramaticalizado, em que deixa de integrar o sistema de honoríficos, marcado pelo traço de distanciamento, e pela vinculação à 3ª pessoa gramatical. Na perspectiva comparativa com o PE constatamos primeiramente que as formas *você(s)* estão incluídas em seu sistema pronominal. É o que se depreende do Quadro 1, a seguir,

adaptado de Gonçalves e Raposo (2013), no qual se expõe uma síntese do sistema pronominal do PE, nas suas formas átonas e tônicas, nas funções de sujeito e complementos acusativo, dativo e oblíquos preposicionados.

Quadro 1 - Pronomes pessoais no português europeu (PE)

	Formas tônicas		Formas átonas	
	Sujeito	Complemento Preposicionado (oblíquo)	Complemento Direto (acusativo)	Complemento Indireto (dativo)
1sg	eu	mim, comigo	me	me
1pl	nós	nós, conosco	nos, se	nos, se
	a gente	nós, conosco		
2sg.	tu	ti, contigo	te	te
	você	você, si, consigo	o, a, se	lhe
3sg	ele, ela	ele, ela, si, consigo	o, a, se	lhe
3pl.	eles, elas	eles, elas	eles, elas	lhes

O Quadro 1 deixa claro que, embora a forma *você* esteja presente no sistema pronominal em uso no PE padrão, a sua ocorrência está restrita à função de sujeito (caso nominativo) e à função oblíqua de complemento de preposição. Este é um ponto fundamental na distinção do sistema pronominal do PE e do PB, ou seja, a de que os complementos acusativos e dativos na variedade lusitana se realizam exclusivamente pelas formas clíticas, acusativas e dativas.

Considerando o Quadro 2, observa-se o cenário dos rearranjos pronominais no PB padrão, não só em função da perda dos clíticos pronominais de 3ª pessoa (citada anteriormente e amplamente referida na literatura), mas também sob o ponto de vista da gramaticalização da forma *você*, e as formas pronominais a ela relacionadas. As perdas estão indicadas pelas lacunas.

Quadro 2 - Pronomes pessoais no português brasileiro (PB)

	Sujeito	Formas tônicas		Formas átonas	
		Objeto Direto	Complemento Obliquo Preposicionado	Objeto Direto Acusativo	Objeto Indireto Dativo
1sg.	eu		mim, comigo	me	me
1pl.	nós		nós, conosco	nos	nos
	a gente	a gente	a gente		
2sg.	tu		ti, contigo	te	te
	você	você	você	se	--
2pl.	--	--	--	--	--
	vocês	vocês	vocês	se	--
3sg.	ele, ela	ele, ela	ele, ela	se	--
3pl.	eles, elas	eles, elas	eles, elas	se	--

O fato relevante sobre o sistema pronominal do PB contemporâneo, destacado no Quadro 2, é o que revela a resistência dos clíticos acusativos e dativos de 1ª e 2ª pessoas, *me* e *te*, ainda altamente frequentes, expressando a permanência de uma propriedade dos clíticos românicos, ainda que formas inovadoras estejam vinculadas. De fato, como dissemos, enquanto a 1ª pessoa conserva consistência no morfema pessoal, quer como clítico, quer como pronome tônico (*me*; *mim*; *comigo*), a 2ª pessoa apresenta uma variação que se estabeleceu em duas modalidades: (i) variação entre o clítico *te* e a forma tônica *você*, introduzido pelas preposições *a* e *para* nos contextos dativos; (ii) variação entre os clíticos *te* e *lhe*, este último reanalisado gramaticalmente como pronome de 2ª pessoa do singular. Por sua vez, a cisão entre 1ª e 2ª pessoas, por um lado, e a 3ª pessoa, por outro, se efetiva de forma radical com o desaparecimento dos clíticos acusativos *o(s)*, *a(s)* e o clítico dativo *lhe(s)*.⁸

Com isso, a 3ª pessoa no PB se expressa unicamente através de formas pronominais tônicas, *ele(s)*; *ela(s)*, o que constitui um traço fortemente inovador da gramática brasileira, em contraste com o PE,

⁸ O Quadro 2 não inclui o uso da forma forte *eu* e *tu* como complemento direto ou como complemento de preposição. Consideramos que essas ocorrências têm distribuição marcada dialetalmente.

no qual os complementos verbais, OD e OI, são unicamente realizados como clíticos (Quadro1).⁹

Importante notar ainda que a forma *nos* ocorre relacionada à forma forte de sujeito *nós* e à forma gramaticalizada *a gente*, o que indica uma situação de variação.

Por fim, vale ressaltar que o PB não é inovador unicamente em relação ao PE na cisão que expressa na sua gramática clítica, como ainda se identifica como uma “anomalia” quando considerado o universo das demais línguas românicas, nos termos de Roberts (2019).¹⁰

3 O clítico *me* dativo e a sua contraparte forte *mim* em diferentes contextos verbais

3.1 Os dativos *me/mim* com verbos ditransitivos de transferência e de movimento e verbos de criação

As construções ditransitivas canônicas se identificam sintaticamente pela realização de dois complementos verbais, denominados tradicionalmente como objeto direto (OD) e objeto indireto (OI). Em termos semânticos configuram um grupo coeso na expressão da transferência, seja ela, material, ou verbal e metafórica. Nesses contextos, o OD é a entidade transferida/movida e o OI clítico é tematicamente interpretado como recipiente/meta. Observe-se que, com alguns verbos, a sua realização como pronominal tônico *mim* requer categoricamente a preposição *para* (cf. 6a-c).¹¹

⁹ Nunes (2019, 2020) propõe um processo maciço de subespecificação morfológica no sistema pronominal do PB, quando então são eliminadas as especificações redundantes. Para uma exposição comparativa mais completa dos quadros pronominais do PE e PB, cf. Menuzzi e Lobo (2016).

¹⁰ Roberts (2019) afirma que certas línguas parecem ser morfossintaticamente “anômalas” por apresentarem gramáticas “marginais” e mudanças sintáticas “extremas”. Evidências para a anomalia provêm de propriedades do inglês medieval e do PB, apresentadas pelo autor, as quais incluem os rearranjos no nosso sistema pronominal, tanto nas formas nominativas, quanto acustivas e dativas.

¹¹ Essa restrição está condicionada ainda pela posição pós-verbal do OI. Veja que a estrutura de tópico contrastivo, a preposição *a* concorre com *para*:

A mim o João dedicou um belo poema. / Para mim o João dedicou um belo poema.

Observe-se ainda que certos verbos, entre eles, *extrair*, *furtar*, *roubar*, *tirar*, expressam semanticamente um direcionamento inverso, no qual o OI clítico é interpretado como a origem/fonte do movimento. Neste caso, o argumento oblíquo é introduzido pela preposição *de* (6d).

De fato, como ilustrado em (6d), infere-se que o relógio está na posse do referente no momento do roubo. Por sua vez, a contraparte preposicionada instancia a preposição *de* na expressão da origem do movimento.¹²

- (6) a. O João *me* deu/ofereceu o livro. / O João deu/ofereceu o livro *para mim*.
 b. O João *me* enviou/mandou um e-mail. /O João enviou/mandou um e-mail *para mim*.
 c. O João *me* dedicou um belo poema./O João dedicou um belo poema *para mim*.
 d. O ladrão *me* roubou/furtou o relógio./O ladrão roubou/furtou o relógio *de mim*.

Da mesma forma, com os verbos dinâmicos de criação, os quais não envolvem direcionalidade, entre eles, *preparar*, *fazer*, *bater (um bolo)*, *pintar*, *desenhar*, *escrever*, o OI *me* é interpretado como beneficiário, uma vez que a posse do objeto é igualmente inferida. Observe-se, novamente, que a variante oblíqua *mim* é introduzida categoricamente pela preposição *para* (7a -c).

- (7) a. O João *me* fez um delicioso bolo de chocolate .
 b. O João fez um delicioso bolo de chocolate *para mim*.
 c. * O João fez um delicioso bolo de chocolate *a mim*.

¹² Com base em Berlinck (1996), consideramos a seguinte tipologia dos verbos ditransitivos de transferência e movimento:

- (i) verbos que expressam transferência material, entre eles, *dar*, *entregar*, *levar*, *mandar*, *oferecer*;
 (ii) verbos que expressam transferência perceptual, verbal, entre eles, *confessar*, *falar*, *dizer*, *perguntar*, *prometer*, *responder*;
 (iii) verbos que expressam movimento físico, entre eles, *levar*, *trazer*;
 (iv) verbos que expressam movimento abstrato, *entre eles*, *dedicar*, *incorporar*.

3.2 O clítico *me* e a posse externa dativa

No contexto de verbos transitivos de atividade não direcional, entre eles, *cortar, examinar, lavar, operar* (8a-b), e verbos transitivos estativos, entre eles, *admirar, elogiar, interpretar, invejar*, incluindo os perceptivos *ver, ouvir*, entre outros (9a-b), configura-se uma estrutura de posse externa dativa, assim denominada na literatura por uma propriedade peculiar: o possuidor dativo se comporta sintaticamente como um argumento do verbo, mas é interpretado como possuidor do OD. No PB, a relação possessiva que se estabelece entre os dois argumentos, possuidor e possuído, pode ser de posse alienável ou inalienável, esta última em suas diferentes acepções de natureza inerente e cultural. A contraparte da posse dativa é a posse interna genitiva, envolvendo um pronome possessivo no interior do DP.

- (8) a. O vizinho *me* lavou/limpou o carro./ O vizinho lavou/limpou o *meu* carro.
 b. O médico *me* operou/examinou a garganta./ O médico operou/examinou a *minha* garganta.
- (9) a. O médico *me* admirou/elogiou a paciência com os enfermos./ O médico admirou/elogiou a *minha* paciência com os enfermos.
 b. O José *me* viu a tatuagem no braço. / O José viu a *minha* tatuagem no braço.

Vale ressaltar que, na expressão da posse inerente, obtém-se uma ambiguidade estrutural, na qual o clítico *me* é interpretado como beneficiário do evento expresso pelo verbo. Nesse caso, há uma contraparte oblíqua do pronome, categoricamente introduzido pela preposição *para* (10).

- (10) O João lavou/limpou o carro para mim.

Destaque-se ainda que, no contexto da posse envolvendo partes do corpo e verbos dinâmicos agentivos, é possível o apagamento do termo inalienável, quando o possuidor é realizado pelo clítico de 1ª pessoa, conforme ilustrado em (11a-b).

- (11) a. O médico *me* operou. (o joelho).
 b. O médico *me* examinou. (o joelho).

Acrescente-se que essa estrutura está restrita ao uso do clítico de 1ª pessoa, uma vez que a realização do objeto direto (OD) por pronome forte como *você* e *ele(s)/ ela(s)* exclui a realização do argumento possuído. Inversamente, a presença do argumento possuído requer o uso do sintagma nominal possessivo (12a-b).¹³

- (12) a. O João lavou/limpou o carro *para mim*.
 b. O médico operou *ele*. / (*o joelho)/ o joelho *dele*.

Outras expressões de posse inerente são muito produtivas no uso coloquial, expressando a interpretação de possuidor “afetado” do argumento clítico de 1ª pessoa (13a) e (13b), com a respectiva realização oblíqua, com a preposição *em* (13c).

- (13) a. *Me* dói o joelho. (dói em mim/ dói o meu joelho)
 b. *Me* ardeu o machucado. (ardeu em mim/ ardeu o meu machucado)
 c. Música: Dói *em mim* saber que a solidão existe e insiste no teu coração... (Angela Ro Ro) (= me dói saber...)

3.3 O dativo *me* introduzido nos predicados inacusativos

A vasta e intrigante literatura sobre os verbos inacusativos nas diferentes línguas, revela que eles não constituem uma classe homogênea, apresentando dificuldades para a sua identificação, meramente com base em testes de natureza sintática e semântica. Não pretendemos considerar as particularidades translinguísticas que os predicados inacusativos apresentam, levando em conta apenas três propriedades identificadoras de sua estrutura argumental (cf. MARANTZ, 2013), a saber: (i) ausência do argumento agente; (ii) ausência de argumento interno acusativo; (iii) propriedades sintáticas e semânticas do argumento interno.

Começamos com os predicados inacusativos construídos com verbos psicológicos do tipo de *agradar* (14a), com verbos existenciais

¹³ É interessante notar que as construções com a chamada posse externa têm uma contraparte que pode ser considerada um tipo de construção de tópico-sujeito (i) e (ii), em que o argumento na posição de sujeito não é o agente do evento expresso pelo verbo. (cf. CANÇADO; NEGRÃO, 2010; CIRÍACO, 2014).

- (i) Eu me operei (*o joelho) com o Dr. Fulano/ Eu operei o joelho com o Dr. Fulano
 (ii) Eu cortei o cabelo com o Joãozinho.

como *faltar* (14b), e com verbos inacusativos de movimento, como *cair*, *chegar*, *crescer*, *vir*, *sair*, *suced*er (14c-e). Nessas estruturas, o clítico de 1ª pessoa é interpretado como experienciador (verbos psicológicos), ou como possuidor/locativo (verbos existenciais e de movimento, respectivamente).¹⁴

- (14) a. Os vinhos californianos *me* agradam muito
 b. *Me* falta uma boa babá.
 c. *Me* chegaram/vieram duas mensagens pelo celular.
 d. *Me* cresceu a vontade de beber vinho.
 e. *Me* saíram duas espinhas no rosto.
 f. *Me* sucedeu um fato interessante.

Importante ressaltar que o PB se alinha ao PE e ao espanhol na medida em que o verbo *dar* combina com nomes nus na formação de expressões de verbo leve como *dar apoio*, *dar raiva*, *dar medo*, *dar coragem*, *dar permissão* e tantas outras. Com base em Cuervo (2003), assumimos que tais expressões envolvem três argumentos: nominativo, dativo (realizado pelo clítico *me*) e nome nu como objeto direto do verbo (15a) ou como sujeito predicativo na formação de predicados psicológicos (15b-c).¹⁵

- (15) a. Meus pais sempre *me* deram apoio.
 b. Me deu calor esse casaco de lã.
 c. As baratas *me* dão medo / nojo.

Outros verbos do tipo de *ser*, *parecer*, *permanecer*, *interessar*, *importar*, formam predicados estativos associados a adjetivos/advérbios. Em todos eles, o clítico *me* é interpretado como experienciador (16 a-c).

- (16) a. A tese do aluno *me* parece muito bem escrita.
 b. *Me* importa muito essa notícia.
 c. Esse documento *me* interessa bastante.

¹⁴ Essa distinção se baseia em Cuervo (2003), que a desenvolve como parte da sua abordagem sobre o dativos do espanhol contemporâneo.

¹⁵ Como assinalado por um dos pareceristas, a restrição ao nome *nu* não poderia ser categórica, uma vez que os nomes podem ser modificados/determinados, como nos exemplos: *me deram muito apoio*; *me deram o maior apoio*. No entanto, apesar dessa possibilidade, continuam a se diferenciar dos DPs referenciais, definidos, uma vez que não se pode ter: *Meus pais me deram o apoio*; *As baratas me dão o medo, o nojo*.

Por sua vez, nos predicados monotransitivos (falsos ditransitivos), em que o verbo é um denominal (derivado de um nome que expressa o evento), o objeto direto (OD), realizado pelo clítico *me* é interpretado como experienciador. Tal interpretação não depende da raiz verbal, mas do evento expresso pelo verbo. Por hipótese, a raiz verbal seleciona um objeto cognato da categoria nominal (HALE; KEYSER 1993; MANZINI; FRANCO, 2016; PINEDA, 2013, 2020; TORREGO 2010). Observe-se a ausência de uma contraparte preposicionada com a forma oblíqua *mim* introduzida por *para* ou *a*, e a sua realização pela preposição *em* (17a-b). No caso de verbos como, *ajudar*, diferentemente, a preposição *a* é categórica (17c).

- (17) a. O João *me* abraçou / *me* deu um abraço. (*para/*a/ em mim)
 b. O João *me* beijou / *me* deu um beijo. (*para/ *a/ em mim)
 c. O João *me* ajudou / *me* deu uma ajuda. (a / *para mim)

3.4 O argumento *me* dativo com verbos monoargumentais

Verbos que expressam eventos dinâmicos, entre eles *telefonar*, *servir*, *agradecer*, *obedecer*, *pagar*, *perdoar*, podem tomar um OI interpretado como recipiente de algo físico ou metafórico. Esse OI pode ser realizado pelo clítico dativo de 1ª pessoa *me* (18a-c).

- (18) a. O João *me* telefonou .
 b. Meu amigo *me* agradeceu pelo convite.
 c. Este funcionário *me* serviu durante anos.

Esse grupo de verbos é definido pela gramática tradicional como verbos que apresentam uma variação intransitiva e transitiva, na qual se identifica a realização de um OD. Nesse caso, não se altera o estatuto semântico do OI, ou seja, a sua interpretação como recipiente afetado (19a-c).

- (19) a. Meu amigo *me* agradeceu o convite.
 b. Este funcionário *me* serviu o almoço durante anos.
 c. Os clientes *me* pagaram as dívidas.

3.5 O argumento *me* dativo com verbos inergativos

Como sabemos, verbos inergativos como *cantar*, *dançar*, *correr*, entre outros, expressam uma atividade que envolve produzir som ou movimento. A presença de um OD é que os torna verbos de atividade e, portanto, capazes de tomar um argumento dativo (20 a-d). Nesse caso, o clítico *me* pode ser licenciado e interpretado como beneficiário. A contraparte oblíqua exige a preposição *para* (21a-b).

- (20) a. A menina *me* dançou um samba.
 b. * A menina *me* dançou.
 c. A menina *me* cantou uma bela canção.
 d. * A menina *me* cantou.
- (21) a. A menina dançou (um belo samba) *para mim*.
 b. A menina cantou (uma bela canção) *para mim*.

3.6 O dativo ético

Por fim, consideremos as estruturas com o chamado dativo ‘ético’, identificado a partir das seguintes propriedades de natureza morfossintática e semântico/pragmática: (i) sua ocorrência se dá em estruturas transitivas e inacusativas que expressam eventos dinâmicos; (ii) está praticamente restrito ao clítico de 1ª pessoa *me* (cf. MENON, 2006) (iii) não apresenta alternância entre a realização clítica e DP lexical ou pronominal forte (a/para mim); (iv) não é um argumento do verbo, mas um elemento interpretado pragmaticamente como avaliador afetado do evento expresso pelo verbo; (v) tem a sua ocorrência condicionada por sentenças com forte força ilocucionária, tais como imperativas, exclamativas, e mesmo afirmativas. Vejamos a exemplificação em (22a-d).

- (22) a. Não *me* queime o arroz!
 b. O menino *me* atravessou a rua sem olhar para os lados.
 c. Como que a minha filha *me* engravida neste ano de pandemia?
 d. As crianças *me* comeram todos os brigadeiros antes da festa.

Ao final desta seção, buscamos evidenciar pelos dados apresentados a alta produtividade do clítico *me* nos variados contextos verbais, instanciando predicados ditransitivos de movimento/

transferência e criação, nos quais o clítico é interpretado como recipiente/meta benefactivo, em predicados inacusativos, nos quais é interpretado como experienciador, possuidor, locativo, em contextos de posse externa dativa, como possuidor, com verbos monoargumentais, e como dativo ético. Por sua vez, o uso das preposições *a* e *para* na contraparte oblíqua revela ainda um aporte semântico que se alinha aos traços da forma clítica e da forma oblíqua como argumentos afetados, envolvidos nos eventos expressos pelos verbos. Esta propriedade representa um aspecto importante para motivar a abordagem formal a ser proposta.

4 Uma proposta de análise para a resistência do clítico *me* no PB

4.1 O núcleo aplicativo baixo e os clíticos dativos de 3ª pessoa

Nesta seção, propomos uma análise para a cisão que atinge os clíticos dativos no PB, numa perspectiva sintática da estrutura de argumentos (cf. MARANTZ, 2013), a qual envolve a atuação de núcleos aplicativos como introdutores de argumentos. Pykkänen (2002, 2008) argumenta que as construções ditransitivas de objeto duplo (DOC) das línguas banto e do inglês expressam, respectivamente, dois tipos semanticamente distintos de núcleos aplicativos. O aplicativo alto (AppIP) é projetado acima do VP/*v*P e denota uma relação temática entre o argumento aplicado (objeto indireto) e o evento expresso pelo verbo (VP/*v*P) (23a). Por sua vez, o aplicativo baixo (AppIP) é projetado como complemento do VP/*v*P e denota uma relação possessiva entre duas entidades, o objeto indireto (objeto aplicado) e o objeto direto (23b):

(23) a. Aplicativo Alto

$$[_{AppIP} \text{OI} [_{AppI'} \text{AppI}_\emptyset [_{VP} \text{V OD tema}]]]$$

b. Aplicativo Baixo

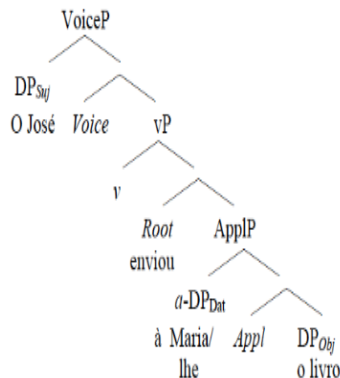
$$[_{VP} \text{V} [_{AppIP} \text{OI} [_{AppI'} \text{AppI}_\emptyset [_{DP} \text{OD tema}]]]]]$$

A proposta do aplicativo baixo para a DOC do inglês foi posteriormente adotada e refinada em Cuervo (2003, 2010, 2020) para as ditransitivas do espanhol; Fournier (2010) para o francês; Pineda (2013, 2020) para o espanhol e o catalão, para mencionar apenas alguns estudos mais recentes referentes às línguas românicas. Na mesma perspectiva,

Torres Morais e Salles (2010, 2016, 2019); Salles e Torres Morais (2020); Torres Morais e Berlinck (2018) tratam as construções ditransitivas do português europeu (PE) e do PB histórico em termos do aplicativo baixo (Fig.1), com verbos dinâmicos de transferência/ movimento e verbos de criação, nas quais o objeto indireto (OI) de 3ª pessoa é interpretado como recipiente, fonte, beneficiário (24a-c). Inclui-se, na mesma proposta, a construção de posse externa dativa, na qual o OI dativo é interpretado como possuidor (24d).

- (24) a. O José enviou um livro *à Maria/* enviou-*lhe* um livro.
 b. O ladrão roubou o relógio *ao Pedro/*roubou-*lhe* o relógio
 c. A Maria preparou uma festa *aos amigos/*preparou-*lhes* uma festa.
 d. O José lavou o carro *à vizinha/*lavou-*lhe* o carro.

Figura 1 - Aplicativa Baixo (Português)



Observe-se que, na estrutura argumental da construção aplicativo representada na Figura 1, a frase aplicativo (AppIP) se relaciona ao verbo como um constituinte e expressa configuracionalmente uma relação direta entre o OD e o OI aplicado. O núcleo aplicativo toma o OD como complemento e introduz o OI no SpecAppIP, assinalando a ele um papel temático associado ao caso dativo inerente. Seguindo Roberts (2007), Torres Morais e Salles (2010), e trabalhos subsequentes, assumem que o caso dativo inerente no argumento aplicado é expresso morfologicamente no clítico de 3ª pessoa ou na preposição *a*. Por sua

vez, o núcleo funcional *Voice*, gerado acima da projeção eventiva (vP), introduz o argumento externo, interpretado como agente. As autoras propõem ainda um traço interpretável de pessoa no núcleo aplicativo, em termos da distinção minimalista do traço [+/- interpretável] e da operação *Agree* (cf. CHOMSKY, 1995, 1999, 2001).

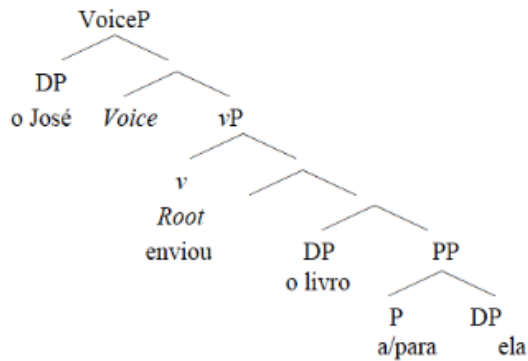
Um aspecto relevante dessa análise é a de que o traço interpretável de pessoa no núcleo aplicativo, associado ao caso dativo inerente, desativa o núcleo aplicativo baixo para a operação *Agree*. Ao contrário do núcleo aplicativo, porém, os traços- ϕ não interpretáveis em v ativam a operação *Agree* com os traços- ϕ interpretáveis do DP- argumento interno. A operação elimina os traços- ϕ não interpretáveis de v e valora o traço de caso do DP como acusativo. Obviamente, as posições primitivas do *Merge* não se manifestam na ordenação final dos constituintes, ou seja, a ordem não marcada OD-OI. As autoras argumentam que essa ordem resulta de um movimento do OD para uma posição mais alta, requerida pela operação de checagem do caso estrutural acusativo.

Entretanto, consideremos a hipótese de que o licenciamento do OI dativo por um núcleo aplicativo não parece ser plausível para o PB. Como vimos na seção 2, a documentação hoje disponível revela a perda dos clíticos dativos de 3ª pessoa e a sua substituição pelas formas pronominais fortes *ele(s)*; *ela(s)*, configurando uma estrutura ditransitiva preposicionada (Fig.2). Lembramos que essa possibilidade está descartada no PE, conforme demonstrado pelo contraste ilustrado em (25a-b).¹⁶

- (25) a. O José enviou-*lhe* o livro. (PE; *PB)
 b. O José enviou o livro *a/para ela*. (*PE; PB)

¹⁶ Para uma abordagem das ditransitivas preposicionadas no PB (cf. CALINDRO, 2015).

Figura 2 - Ditransitiva Preposicionada



Uma comparação entre a estrutura aplicativa (Fig. 1) e a ditransitiva preposicionada (Fig. 2) pode ser feita nos seguintes termos: enquanto o núcleo aplicativo está comprometido com uma marca de caso dativo expressa morfologicamente no objeto aplicado, a ditransitiva preposicionada do PB instancia o uso inovador dos pronomes tônicos/fortes introduzidos pelas preposições *a* e *para*.

Observe-se, porém, que a condição de animacidade do referente de 3ª pessoa se mantém nas formas oblíquas. Vale ressaltar que a noção de animacidade aqui considerada tem efeitos de natureza morfossintática e não se restringe a uma correlação conceptual com o mundo real. Portanto, estende-se a um uso metafórico e à relação parte-todo. Para exemplificar, tomemos a expressão locativa ‘Banco do Brasil’ na sentença “Enviei uma carta ao Banco do Brasil” e constatemos que a expressão locativa não poderá ser pronominalizada pelo clítico *lhe* ou pelo pronome tônico *ele*, a não ser que seja metaforicamente interpretada como “os funcionários do Banco do Brasil.” Neste caso, os pronomes ocorrerão no plural, resultando em sentenças gramaticais como “Enviei-lhes uma carta/ Enviei uma carta a/para eles.”

4.2 A perda dos clíticos dativos de 3ª pessoa no PB

Em Torres Morais e Salles (2010) e trabalhos subsequentes, assumimos a hipótese de que a perda do clítico *lhe(s)* em estruturas ditransitivas no sistema pronominal do PB evidencia a desativação do núcleo aplicativo baixo, ou seja, evidencia a perda de um traço

interpretável de pessoa no núcleo aplicativo. Essa análise, portanto, supõe uma dependência da DOC românica em relação à codificação morfológica do argumento aplicado, que não se verifica nem na DOC do inglês, nem na DOC dialetal do PB, ambas com diferentes propriedades codificadoras.

A hipótese da desativação do aplicativo baixo no PB, tomando como base a perda dos clíticos dativos de 3ª pessoa, está, porém, comprometida com o fato de que os clíticos de 1ª e 2ª pessoas não se perderam, muito pelo contrário, como vimos na seção 3 deste estudo, onde documentamos a alta frequência do clítico *me* em diferentes contextos verbais. Além disso, como também registramos, o clítico *te* resiste em variação com a forma gramaticalizada *você*.

Galves (2018), na discussão da análise de Torres Morais e Salles (2010), faz comentários sobre a mesma questão. A autora argumenta que a cisão não envolve a perda do núcleo aplicativo, mas decorre do fato de que “os pronomes clíticos que se mantiveram no paradigma têm morfologia dativa, mas o caso dativo, correspondendo efetivamente à função dativa, não existe mais na língua” (p. 93). Diante disso, propõe que a perda da distinção acusativo/ dativo na 3ª pessoa no PB resulta na reanálise do dativo, passando esses clíticos a serem licenciados localmente pelo verbo (V), exatamente como os pronomes fortes (*você*) e os DP plenos. Nosso entendimento é o de que as ponderações de Galves são pertinentes em relação à necessidade de se propor uma análise para os clíticos (dativos) sobreviventes de 1ª e 2ª pessoa *me* e *te* (em oposição à perda dos clíticos pronominais de 3ª pessoa) no PB, o que constitui nosso objetivo primordial neste estudo.¹⁷

¹⁷ A autora questiona ainda que estruturas de objeto duplo do PB dialetal sejam analisadas em Torres Morais e Salles (2010) em termos da projeção aplicativa (nos moldes da análise de Pylkkänen, (2002, 2008), para a construção de objeto duplo do inglês), alegando ser “um tanto paradoxal pensar que a construção com aplicativo se perdeu no PB padrão, mais próximo do seu ancestral português e se manteve no PB dialetal mais marcado pelo contato” (p. 95). Em relação ao contraste entre o PB dialetal e as demais variedades do PB, nosso entendimento é o de que as diferentes gramáticas viabilizam diferentes possibilidades de implementação, não havendo, presentemente, evidência para se postular que a desativação de uma categoria em uma variedade exclua sua manifestação em outra variedade (em termos de uma noção de *markedness* ou de um *cline* na manifestação de núcleos funcionais). O desafio que se coloca é explicar a existência de duplos na gramática, observada no PB dialetal (em que ocorrem DOC e oblíquo-PP), mas não no dito “PB padrão” (em que a DOC não ocorre) – uma questão também destacada no estudo de Galves (2018).

Antes de nos voltarmos à questão da cisão pronominal no PB, é importante considerar o fato de que numa mesma língua podemos encontrar diferentes tipos de configurações aplicativas. Cuervo (2003), por exemplo, afirma que, no espanhol, os diferentes tipos de aplicativos baixos introduzem o OI num conjunto de verbos ditransitivos. Da mesma forma, o aplicativo alto é identificado em certos predicados inacusativos, com verbos psicológicos. Por fim, o que a autora denomina como aplicativo afetado é projetado nas estruturas bieventivas, acusativas e incoativas numa distinta posição configuracional. Em todos esses casos, o dativo é um argumento aplicado, ou seja, introduzido pelo núcleo aplicativo. Nos mesmos moldes, Torres Morais (2007) reconhece a atuação dos diferentes núcleos aplicativos baixos e altos no PE.

Assim, no que se segue, buscaremos evidenciar que a distribuição do clítico *me* dativo no PB em certos contextos verbais motiva uma análise do aplicativo alto, lembrando que o mesmo relaciona semanticamente um indivíduo a eventos (*vP*), os quais podem ser subdivididos em evento estativo (*vPBE*), evento dinâmico não agentivo (*vPGO*), caracterizando um conjunto dos predicados inacusativos, e evento agentivo dinâmico (*vPDO*) Cf. Harley (1995).. Os significados do argumento aplicado estão, portanto, condicionados ao tipo de *vP* que o aplicativo alto toma como complemento e ao tipo de elemento do qual o núcleo aplicativo é o complemento (cf. CUERVO, 2003, 2020).

4.3 Aplicativo alto no PB¹⁸

4.3.1 O dativo *me* com predicados inacusativos

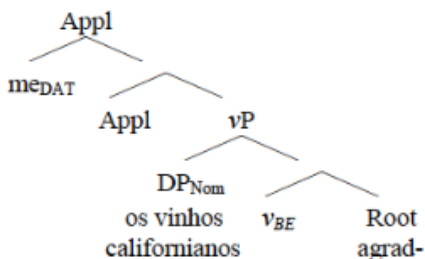
Como vimos anteriormente (cf. seção 3.3), o clítico dativo *me* é licenciado com predicados inacusativos construído com verbos psicológicos do tipo de *agradar*, *parecer*, com verbos existenciais como *faltar*, e com verbos inacusativos de movimento, como *cair*, *chegar*, *crescer*, *vir*, *sair*, *suced* e outros. Nessas estruturas, o clítico de 1ª pessoa é interpretado como experienciador (verbos psicológicos), ou como possuidor/locativo (verbos existenciais e de movimento, respectivamente).

¹⁸ As seções que se seguem contemplando os aplicativos altos, segue Cuervo (2003, 2020) para as estruturas com verbos psicológicos e inacusativos.

Propomos uma análise generalizada do aplicativo alto para os predicados inacusativos do PB. No caso dos verbos psicológicos *agradar* e *parecer*, o aplicativo alto toma um vP estativo como complemento e introduz o argumento aplicado na posição mais alta, ou seja, externa ao evento (Fig. 3) e (Fig. 4), relativas às sentenças (26a-b):

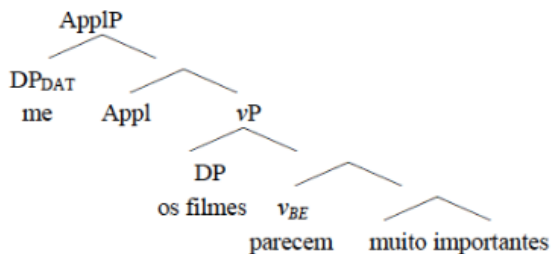
(26) a. Os vinhos californianos *me* agradam.

Figura 3 – Verbos Psicológicos



b. Esses filmes *me* parecem muito importantes.

Figura 4 – Predicados inacusativos

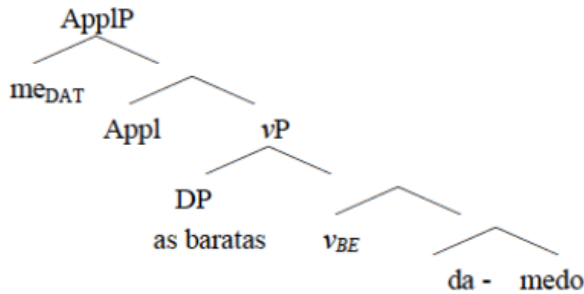


As estruturas acima expressam ainda que o DP tema é o sujeito dos predicados estativos *agradar* e *parecer*, e não o seu objeto. Com Cuervo (2003), argumentamos que essa propriedade do argumento interno, associada à projeção externa do argumento dativo expressam de forma natural duas das propriedades dos predicados psicológicos. A primeira é a de que o argumento nominativo não pode ser um nome *nu*, o que caracteriza uma propriedade identificadora dos sujeitos predicativos. A segunda propriedade relevante é a de que o argumento dativo não participa da relação de predicação entre o verbo e o DP nominativo.

Na mesma linha de análise, será identificado o uso do verbo *dar* (verbo leve) na expressão de estados psicológicos (27).

(27) As baratas *me* dão medo.

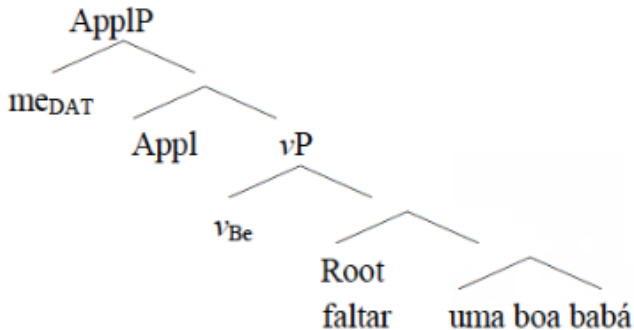
Figura 5 – Verbos leves



Propomos ainda que a análise do aplicativo alto contempla os predicados existenciais. A única diferença entre os verbos existenciais e os psicológicos é a de que o DP do evento estativo é realizado na posição de argumento interno (28).

(28) *Me* falta uma boa babá.

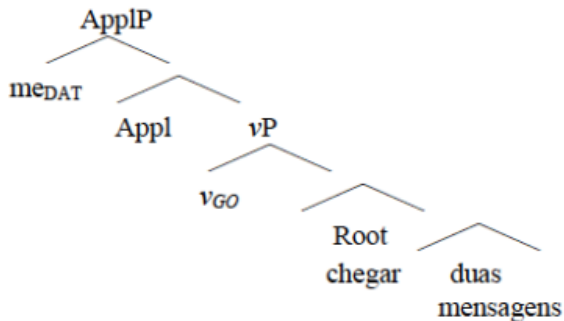
Figura 6 – Verbos existenciais



Da mesma forma, com verbos de movimento como *chegar*, o DP nominativo é gerado na posição de argumento interno da raiz verbal. Porém, neste caso, temos um evento dinâmico não agentivo (vGO) selecionado pelo núcleo aplicativo, como expresso na Fig. 7, relativa à sentença (29).

(29) Me chegaram duas mensagens.

Figura 7 – Verbos inacusativos dinâmicos

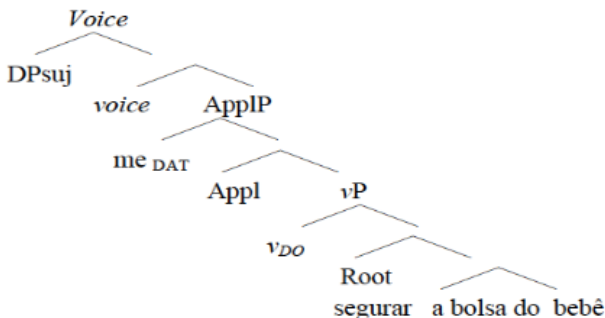


4.3.2 O dativo *me* nas construções ditransitivas

Uma vasta literatura tem se dedicado a identificar algumas construções que são exemplos claros de interpretação, na qual o evento é direcionado para uma pessoa em seu detrimento ou benefício. São os denominados dativo *commodi/incommodi*. Vamos considerar que esses casos são expressos tanto sintaticamente como semanticamente pelos aplicativos altos. No PB temos exemplos claros da interpretação de um objeto aplicado como indivíduo afetado, de modo positivo ou negativo, pelo evento expresso pelos verbos, como ilustrado em (30) e Figura 8.

(30) A babá me segurou a bolsa (do bebê)

Figura 8 – Verbos ditransitivos estativos



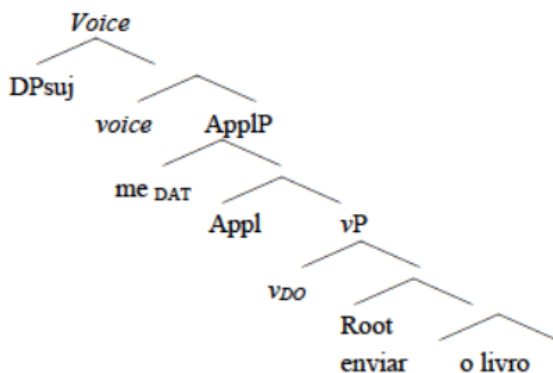
O clítico dativo introduzido na estrutura argumental eventiva por um aplicativo alto nos predicados inacusativos e agentivos representa um fator relevante na nossa abordagem da cisão. Procuraremos defender a hipótese de que a permanência do clítico *me* evidencia dois movimentos inovadores no PB: o primeiro decorre de uma mudança microparamétrica no sistema dos núcleos sintáticos aplicativos, nos seguintes termos: diferentemente de outras línguas românicas, o PB aciona unicamente o núcleo Aplicativo Alto, o qual introduz os clíticos de 1ª (e 2ª) pessoas na posição de especificador e o relaciona tematicamente a diferentes tipos de eventos. Isso sugere que os aplicativos altos podem ter o mesmo significado básico, algo como “o evento é orientado para o dativo”.

A interpretação de experienciador, beneficiário e malefativo, portanto, não decorre do conteúdo semântico proporcionado pelo núcleo licenciador. Um traço interpretável de pessoa no núcleo aplicativo alto codifica gramaticalmente a pessoa e caso dativo inerente do objeto clítico aplicado (cf. TORRES MORAIS; SALLES, 2010).

Como era de se esperar, portanto, com a perda dos aplicativos baixos no PB, a nossa proposta engloba dentro do cenário dos aplicativos altos, os verbos ditransitivos de movimento e transferência, verbos de criação e posse externa, ilustrado pela estrutura ditransitiva em (31).

(31) O José *me* enviou o livro.

Figura 9 – Verbos ditransitivos dinâmicos



Dessa forma, a relação possessiva entre o objeto aplicado e o OD nas ditransitivas canônicas é recuperada na configuração do aplicativo alto. O aplicativo alto relaciona semanticamente o argumento aplicado

ao evento de transferência de posse, possibilitando que o OI possa ser interpretado como o possuidor, recipiente/meta.

O mesmo ocorre com verbos de criação em sentenças ditransitivas como *O José me preparou um jantar, me bateu um bolo* e outros. Nesse contexto, o clítico argumental é interpretado como beneficiário do evento de mudança de estado do OD.¹⁹

Quanto ao segundo movimento inovador para a causa da cisão, propomos que este se refere aos rearranjos no sistema pronominal do PB, em particular, os que afetaram a realização clítica dos complementos acusativos e dativos de 3ª pessoa.

4.3.3 A perda dos clíticos de 3ª pessoa no PB e efeitos da restrição de pessoa

A hipótese de uma generalização do aplicativo alto no PB está associada intimamente com a reanálise referente à perda dos clíticos dativos e acusativos de 3ª pessoa, criando uma cisão entre clíticos pronominais puros e clíticos determinantes (RAPOSO, 1999).

Recentemente, Torres Morais e Salles (2016, 2019) retomam estudos sobre os efeitos da *Restrição Pessoa Caso*²⁰ na realização clítica dos argumentos internos das construções ditransitivas. Dados provenientes de documentação histórica evidenciam que a formação dos grupos de clíticos no português se enquadra na variação denominada *Strong PCC*, descrito por Bonet (1991:182), nos seguintes termos: “Se Dativo, então acusativo de 3ª pessoa”, como ilustrado nas sentenças (32a-d).

- (32) a. O João não *me/ te/ lhe* apresentou o Pedro.
 b. O João não *mo/ to/ lho* apresentou. 1a /2a /3/a P – dat; 3a P acus.
 c. * O João não *te me* apresentou. *2a P – dat; 1a P - acus.
 d. * O João não *me te* apresentou. * 1a P – dat; 2a P - acus

Seguindo Pancheva e Zubizarreta (2018), assumimos que o PCC é um fenômeno da interface entre a sintaxe e a semântica, na

¹⁹ A realização do beneficiário na posição de aplicativo alto está contemplada na análise do dativo ético no PB (cf. ROCHA, 2017).

²⁰ *Person Case Constraint* (PCC).

medida em que um traço de pessoa interpretável no núcleo aplicativo é responsável pela marcação gramatical do objeto indireto OI, introduzido no SpecApplP, como centro do ponto de vista. Dependendo do valor do traço interpretável de pessoa nas gramáticas individuais, o OI pode estar restrito à 1ª pessoa ou à 2ª pessoa, ou a ambas, ou incluir as 3as pessoas, contanto que estas sejam igualmente apropriadas para centro da perspectiva, ou sejam tenham o traço de animacidade, no sentido de estarem associadas aos papéis temáticos do OI, a saber, recipiente/meta, experienciador, possuidor, afetado.

Segundo as autoras, a codificação sintática da noção semântica de ponto de vista é responsável pelos efeitos de sensibilidade de pessoa nas configurações de PCC. Adotando as especificações de traços relacionados à pessoa, proposta em Nevins (2007), as autoras acrescentam a elas o traço [aproximativo]. Assim, como elas esclarecem, a assimetria entre 1ª e 2ª pessoas e argumentos de 3ª pessoa é capturada pela especificação positivo (+) *versus* negativo (-) do traço participante. Os dois argumentos participantes são ainda distinguidos através do traço [+/- autor]. Por sua vez, a noção aproximativo, como a noção de 1ª e 2ª pessoas, está relacionada com a situação da fala, entendida em termos de perspectiva. Nas palavras das autoras “*1P and 2P arguments are inherently proximate, being part of the speech event. 3P arguments may or may not be proximate, depending on context.*” (PANCHEVA; ZUBIZARRETA, 2018: 13).

Ainda segundo esta proposta, o OI introduzido no SpecAppl ativa uma relação de concordância com o traço interpretável de pessoa do Appl expressando uma concordância com o seu próprio traço interpretável de pessoa. Portanto, dentro da fase aplicativo, manifestam-se os diferentes tipos de PCC, ativados por uma restrição de pessoa, com cláusulas que atuam de forma variável nas diferentes línguas. Para nós, o importante é reconhecer que o português histórico e o PE contemporâneo ativam em suas gramáticas o *Strong PCC* com três restrições relevantes: (i) restrição imposta ao OD de 3ª pessoa (ii) proibição das combinações entre 1ª e 2ª pessoas; (iii) permissão para as combinações de duas 3as pessoas. Com essas propriedades, conclui-se que o traço de pessoa interpretável no Appl é valorado [+ aproximativo], como caso *default*.

Com a perda dos clíticos de 3ª pessoa no PB, não se formam mais grupos de clíticos nos moldes das restrições impostas pelo PCC forte, o que nos leva a reconhecer uma reanálise radical na caracterização

dos traços de Pessoa no PB, de modo que valores binários contrastivos tais como [+participante] *versus* [-participante] no evento de fala, ou [+ aproximativo] *versus* [-aproximativo] deixam de ser relevantes. Consequentemente, a 1ª e 2ª pessoas são identificados no traço monovalente [participante] do núcleo funcional aplicativo.

Entende-se por que, embora sem instanciar o PCC forte com a perda dos clíticos acusativos e dativos de 3ª pessoa, o PB não permite a formação de dupla pronominalização com os clíticos *me*, *te/te*, *me* no domínio da fase aplicativo (33).

- (33) a. *O João *me te* apresentará. [*1-DAT, 2-ACUS]
 a'. O João te apresentará *a mim*.
 b. *A Maria *te me* apresentará. [*2-DAT, 1-ACUS]
 b'. A Maria me apresentará *a ti/a você*.

Vamos propor, portanto, que a mudança de natureza microapamétrica efetivada na cisão no sistema pronominal dos complementos clíticos no PB se reflete no traço de Pessoa interpretável do núcleo aplicativo.

5 Considerações finais

Neste texto investigamos o fenômeno da cisão pronominal na expressão do dativo no português brasileiro (PB), em que se verifica o uso dos clíticos de 1ª e 2ª pessoas no singular, *me* e *te*, em oposição ao uso do pronome forte de 3ª pessoa como complemento da preposição dativa (*para ele(s)* e *ela(s)*), diante da perda do clítico dativo *lhe(s)*.

Argumentamos a favor da hipótese de que a cisão dos clíticos dativos no quadro pronominal, em particular a permanência do clítico *me*, evidencia dois movimentos inovadores no PB, decorrentes de opções microparamétricas altamente específicas. O primeiro movimento se efetiva no sistema dos núcleos sintáticos aplicativos, nos seguintes termos: diferentemente de outras línguas românicas, o PB aciona unicamente o núcleo Aplicativo Alto, o qual introduz os clíticos de 1ª (e 2ª) pessoas na posição de especificador e os relaciona tematicamente a diferentes tipos de eventos. Um traço interpretável de pessoa no núcleo aplicativo alto codifica gramaticalmente a pessoa e o caso dativo inerente do objeto clítico aplicado. A segunda inovação decorre diretamente da perda dos

clíticos dativos e acusativos de 3ª pessoa no PB, criando uma cisão entre clíticos pronominais e clíticos determinantes. Em particular, mostramos que a ausência dos efeitos da Restrição Pessoa Caso na gramática é uma evidência adicional de reanálise na caracterização dos traços de pessoa de modo que valores binários deixam de ser relevantes. Consequentemente, a 1ª e 2ª pessoas são identificados no traço [participante] do núcleo funcional aplicativo.

Declaração de autoria

Este artigo foi elaborado pelas autoras dentro de uma estratégia de discussões que envolveram desde a escolha do tema, os fatos descritivos e abordagem teórica de orientação gerativista. Além disso, foi decidido em comum acordo a estrutura do texto, com divisão em 4 seções, acrescidas de introdução, conclusão e referências bibliográficas. Partes do texto final foram apresentadas em dois eventos, dos quais ambas as autoras participaram, a saber, o VCILH e o 3EGG.

Agradecimentos

Versões anteriores deste artigo foram apresentadas no V Congresso Internacional de Linguística Histórica (VCILH), em homenagem à Mary Kato e Charlotte Galves, e no III Encontro de Gramática Gerativa (IIIIEGG), em homenagem à Sonia Cyrino e Eugênia Duarte. As autoras agradecem às audiências pelas importantes contribuições. Expressamos ainda os nossos agradecimentos à leitura cuidadosa e importantes sugestões enviadas pelos dois pareceristas anônimos deste texto.

Referências

BERLINCK, R. Datives. In: VAN BELLE, W.; VAN LANGENDOMCK (eds.) *The Dative, Descriptive Studies*. Amsterdam: John Benjamins, 1996. p. 119-151.

CALINDRO, A. R. *Introduzindo argumentos: Uma proposta para as sentenças ditransitivas do português brasileiro*. 2015. 206p. Tese (Doutorado) Universidade de São Paulo, 2015.

- CANÇADO, M. Verbal alternations in Brazilian Portuguese: a lexical semantic approach. *Studies in Hispanic and Lusophone Linguistics*, De Gruyter Mouton, v. 3, n. 1, p. 77-111, 2010. DOI:10.1515/shll-2010-1066
- CANÇADO, M.; NEGRÃO, E. *Two possessor raising constructions in Brazilian Portuguese*. Comunicação apresentada ao VIII Workshop on Formal Linguistics. Universidade de São Paulo, São Paulo, 2010.
- CHOMSKY, N. *The Minimalist Program*. Cambridge MA: MIT Press, 1995.
- CHOMSKY, N. Minimalist inquiries: The framework. In: MARTIN, R. et al. (eds.). *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*. Cambridge: MIT Press, 2000. p. 89-155.
- CHOMSKY, N. Derivation by phase. In: KENSTOWICZ, M. (ed.). *K. Hale: A Life in Language*. Current Studies in Linguistics 36. Cambridge: MIT Press, 2001. p. 1-52.
- CIRÍACO, L. S. A construção transitiva de sujeito agente-beneficiário no português brasileiro. *Caligrama*, Belo Horizonte, v. 19, n. 2, p. 83-98, 2014.. doi:<http://dx.doi.org/10.17851/2238-3824.19.2.83-98>.
- CUERVO, M. C. *Datives at Large*. 2003. 212p. PhD thesis. Department of Linguistics, Massachusetts Institute of Technology, 2003.
- CUERVO, M. C. Against ditransitivity. *Probus*, De Gruyter Mouton, v. 22, p. 151-180, 2010. <https://doi.org/10.1515/prbs.2010.006>
- CUERVO, M. C. Datives as applicatives. In: PINEDA, A.; MATEU, J. (eds.). *Dative constructions in Romance and beyond*. Berlin: Language Science Press, 2020. p. 1- 42.
- CYRINO, S.M.L. Observações sobre a mudança diacrônica no português do Brasil: objeto nulo e clíticos. In: ROBERTS, I.; KATO, M. (orgs.). *Português brasileiro: uma viagem diacrônica*. Campinas, Editora Contexto, 2018. p. 129-142.
- CYRINO, S.M.L. O objeto nulo do português brasileiro: sincronia e diacronia. In: GALVES, C.; KATO, M.; ROBERTS, I. (orgs.). *Português brasileiro: uma segunda viagem diacrônica*. Campinas, Editora da Unicamp, 2019. p. 173- 200.
- DUARTE, M.E.L. O papel da sociolinguística na descrição da gramática da escrita contemporânea. In: MARTINS; M. A.; TAVARES, M.A. (orgs.).

Contribuições da Sociolinguística e da Linguística Histórica para o ensino de língua portuguesa. 1ed. Natal: EDUFERN, 2013. p. 113-142.

DUARTE, M.E.L.; FREIRE, G. C. Como a escrita padrão recupera formas em extinção e implementa formas inovadoras. In: PAIVA, M. C.; GOMES, C. A. (orgs.) *Dinâmica da Variação e da Mudança na Fala e na Escrita*. 1a. ed. Rio de Janeiro: Contra Capa, 2014. p. 121-142.

FOURNIER, D. *La structure du prédicat verbal: une étude de la construction à double objet en français*. 2010. 274p. PhD, University of Toronto, 2010.

FREIRE, G. C. *Os clíticos de terceira pessoa e as estratégias para sua substituição na fala culta brasileira e lusitana*, 2000.118p. Dissertação (Mestrado) - Universidade Federal do Rio de Janeiro, 2000.

FREIRE, G. C. *A Realização do Acusativo e do Dativo Anafórico de 3ª Pessoa na Escrita Brasileira e Lusitana*. 2005. 215p. Tese (Doutorado) Universidade Federal do Rio de Janeiro, 2005.

GALVES, C. *Ensaio sobre as gramáticas do português*. Campinas, Editora da UNICAMP, 2001.

GALVES, C. Competition, stability and change in the emergence of Brazilian Portuguese. In: BREITBARTH, A.; BOUZOUITA, M.; DANCKAERT, L.; FARASYN, M. (eds.) *The determinants of diachronic stability*. Amsterdam, John Benjamins Publishing Company, 2019. p. 191-2014.

GALVES, C. Ainda sobre os pronomes do português brasileiro. Sintaxe, morfologia e variação. In: NEVINS, A.; BOECHAT, A. (orgs.). *O apelo das árvores*. Campinas: Pontes Editores, 2018. p. 79-100.

GOMES, C. A. *Aquisição e perda da preposição no português do Brasil*. 1996. Tese (Doutorado) Universidade Federal do Rio Janeiro, 1996.

GOMES, C. A. Variação e Mudança na Expressão do dativo no português brasileiro. In: PAIVA, M. da C.; DUARTE, M. E. L. (orgs.) *Mudança linguística em tempo real*. Rio de Janeiro, FAPERJ/Contracapa, 2003. p. 81-96.

GONÇALVES, A.; RAPOSO, E. Verbo e sintagma verbal. In: RAPOSO, E. et al. (coord.) *Gramática do português*. Volume 2. Cap. 28, Lisboa: Fundação Calouste Gulbenkiano, 2013. p. 1153-1218.

HALE, K.; KEYSER, S. J. On argument structure and the lexical expression of syntactic relations. In: HALE, K.; KEYSER, S. J. (eds.). *The View from Building 20*. Cambridge: MIT Press, 1993. p. 53-109.

- HARLEY, H. *Subjects, events and licensing*. 236 p. PhD Dissertation. Massachusetts Institute Technology, 1995. Dept. of Linguistics and Philosophy, 1995.
- KATO, M. A gramática do letrado. In: MARQUES, M. A.; TEIXEIRA, J.; LEMOS, A. S. (orgs.) *Ciências da linguagem: trinta anos de investigação e ensino*. Braga: CEHUM/ Universidade do Minho, 2005. p. 131-145.
- KATO, M.; CYRINO, S.; CORRÊA, V. R. Brazilian Portuguese and the recovery of lost clitics through schooling. In : PIRES, A.; J. ROTHMAN. (eds.). *Minimalist inquiries into Child and Adult Language Acquisition*. New York: Mouton de Gruyter, 2009. p. 245-272.
- LOPES, C. R. *et al.* Sobre norma e tratamento em cartas a Rui Barbosa. In: AGUILERA, V. (org.) *Para a História do Português Brasileiro*. Vol. VII. Londrina: Vozes/Veredas/Voragens, 2009. p. 45-92.
- LOPES, C. R. S.; CAVALCANTE, S. R. O. A cronologia do voçamento no português brasileiro: expansão de você-sujeito e retenção do clítico-te. *Linguística*, v. 25, p. 30-65, 2011.
- MARANTZ, A. Verbal argument structure: Events and participants. *Lingua*, v. 130, p. 152-168, 2013. <https://doi.org/10.1016/j.lingua.2012.10.012>
- MANZINI, R.; FRANCO, L. Goal and DOM datives. *Natural Languages and Linguistic Theory*, v. 34, p. 197-240, 2016. ISSN-1573-0859 (web)
- MENON, O. P. S. E não me fique grávida! Ou o caso do dativo ético. In: GORSKI, E. M.; COELHO, I. L. (orgs.) *Sociolinguística e ensino: contribuições para a formação do professor de língua*. v. 1. Florianópolis: Editora da UFSC, 2006. p. 155-171.
- MENUZZI, S.; LOBO, M. Binding and Pronominal Forms in Portuguese. In WETZELS, L.; COSTA, J.; MENUZZI, S. (eds.) *The Handbook of Portuguese Linguistics*. Wiley -Blackwell, 2016. p. 338-355.
- NEVINS, A. The representation of Third-person and its consequences for person-case effects. *Natural Language and Linguistic Theory*, v. 25, p. 273-313, 2007.
- NUNES, J. Clíticos acusativos de terceira pessoa em português brasileiro como concordância de objeto. In: GALVES, C.; KATO, M.; ROBERTS, I. (orgs.) *Português brasileiro*. Uma segunda viagem diacrônica. Campinas, Editora da UNICAMP, 2019. p. 151-172.
- NUNES, J. Especificação morfológica de pronomes nominativos, concordância verbal e sujeitos nulos em português brasileiro. *Forum*

Linguístico, Número Especial. Florianópolis, Universidade Federal de Santa Catarina. v. 17, p. 4658-4672, 2020. <https://doi.org/10.5007/1984-8412.2020v17nespp4658>

PANCHEVA, R.; ZUBIZARRETA, M. L. The Person Case Constraint: The syntactic encoding of perspective. *Natural Language and Linguistic Theory*. v.36. 2018. p.1291-1337.

PEREIRA, M. B. *Predicados bitransitivos do português dialetal do Brasil Central (PBC)*. Construções de objeto duplo e de redobro do clítico. 2019. 260p. Tese Doutorado Universidade de Brasília, 2019.

PINEDA, A. Double object constructions and dative/accusative alternations in Spanish and Catalan: A unified account. *Borealis: An International Journal of Hispanic Linguistics*, v. 2, p. 57–115, 2013. <https://doi.org/10.7557/1.2.1.2524>

PINEDA, A. From dative to accusative. An ongoing syntactic change in Romance. *Probus*, v. 32, n. 1, p. 129-173, 2020. <https://doi.org/10.1515/probus-2019-0001>

PYLKKÄNEN, L. *Introducing arguments*. 2002, 137p. Ph.D thesis. MIT, Cambridge 2002.

PYLKKÄNEN, L. *Introducing arguments*. Cambridge, Massachusetts/London: The MIT Press, 2008.

RAMOS, J. *Marcação sintática e mudança sintática no português*. 1992. 354 p. Tese (Doutorado), Universidade Estadual de Campinas, 1992.

RAPOSO, E. Some observations on the pronominal system of Portuguese. *Catalan Working Papers in Linguistics*, Barcelona, v. 6, p. 59-93, 1999. ISSN 1132-256X

ROBERTS, I. *Diachronic Syntax*. Cambridge: Oxford University Press, 2007.

ROBERTS, I. Gramáticas ‘marginais’ e mudanças sintáticas ‘extremas’: o inglês e o português brasileiro. In: GALVES, C.; KATO, M.; ROBERTS, I. (orgs.) *Português brasileiro*. Uma segunda viagem diacrônica. Campinas, Editora da UNICAMP, 2019. p. 23-56.

ROSA, B. G. *Applicatives in dialectal Brazilian Portuguese*. 2017. 157p. Dissertação (Mestrado) Universidade Federal de Minas Gerais, 2017.

RUMEU, M. C. de B. A implementação de *você* no português brasileiro oitocentista e novecentista: um estudo de painel. 2008. 276p. Tese (Doutorado) - Universidade Federal do Rio de Janeiro, 2008.

SALLES, H.M.L.; TORRES MORAIS, M.A. Estrutura argumental no português brasileiro: perdas e alinhamentos inovadores nos sistemas pronominal e preposicional. Cuadernos de la ALFAL. n.12, (2) p. 467-490, 2020. ISSN 2218- 0761

TORREGO, E. *The Dependencies of Objects*. Cambridge: MIT Press, 2010.

TORRES MORAIS, M. A.; BERLINCK, R. de A. A caracterização do objeto indireto no português: aspectos sincrônicos e diacrônicos In: MATTOS e SILVA, R. V.; LOBO, T. (orgs.). *Novos Dados, Novas Análises*. v. VI, Tomo I, Salvador, EDUFBA, 2006. p. 73-106.

TORRES MORAIS, M.A. *Os dativos*. 2007. 242p. Tese de Livre Docência. Universidade de São Paulo, 2007.

TORRES MORAIS, M. A.; BERLINCK, R. de A. ‘Eu disse pra ele’ ou ‘Disse-lhe a ele’: a expressão do dativo nas variedades brasileira e europeia do português. In: CASTILHO, A. T. de; TORRES MORAIS, M. A.; LOPES, R. E. V.; CYRINO, S. M. L. (orgs.). *Descrição, História e Aquisição do Português Brasileiro*. São Paulo, Fapesp; Campinas, Editora Pontes, 2007. p. 61-83.

TORRES MORAIS, M. A.; BERLINCK, R. de A. O objeto indireto. Argumentos aplicados e preposicionados. In: CYRINO, S.; TORRES MORAIS, M. A. (orgs.). *Mudança sintática no português brasileiro: perspectiva gerativista*. São Paulo, Editora Contexto, 2018. p. 252-307.

TORRES MORAIS, M. A. ; SALLES. H. M. L. Parametric change in the grammatical encoding of indirect objects in Brazilian Portuguese. *Probus*, n. 22, p. 181- 209, 2010. <https://doi.org/10.1515/probus-2019-0001>

TORRES MORAIS. M. A.; SALLES, H.M.L. The external possessor construction in European Portuguese and Brazilian Portuguese. In: KATO, M.; ORDÓNEZ, F. (eds.). *The Morphosyntax of Portuguese and Spanish in Latin America*. Oxford: Oxford University Press, 2016. p. 204-235.

Sintagmas Nominais



Estudo diacrônico do sintagma nominal descontínuo no português brasileiro

Diachronic Study of the Discontinuous Noun Phrase in Brazilian Portuguese

Nathalia Pereira de Souza-Martins

Universidade Estadual Paulista (UNESP), São José do Rio Preto, São Paulo / Brasil

nathaliapsouza12@gmail.com

<https://orcid.org/0000-0003-0365-9591>

Sebastião Carlos Leite Gonçalves

Universidade Estadual Paulista (UNESP), São José do Rio Preto, São Paulo / Brasil

sebastiao.goncalves@unesp.br

<https://orcid.org/0000-0002-1798-729X>

Resumo: Este trabalho tem como objetivo principal examinar, sob a perspectiva da Gramática Discursivo-funcional (HENGEVELD; MACKENZIE, 2008), quais são as motivações pragmáticas, semânticas e morfossintáticas subjacentes à ordenação dos constituintes do SN descontínuo em diferentes sincronias do português brasileiro, com o intuito de verificar se a caracterização sincrônica do fenômeno tem fundamentação diacrônica e se há diferenças na produtividade do fenômeno na fala e na escrita. A análise empírica do fenômeno toma por base amostras provenientes de diferentes fontes históricas: (i) cartas dos séculos XVIII, XIX e XX do Projeto Para a História do Português Brasileiro (PHPB) e do *Corpus* Histórico do Português *Tycho Brahe*; (ii) cartas de leitores do século XXI dos jornais *Folha de São Paulo* e *Diário do Norte*; e (iii) registros de língua falada do século XX, retirados do *corpus* mínimo do Projeto NURC/Brasil, e do século XXI, extraídos do banco de dados Iboruna (Amostra Censo). A metodologia inclui parâmetros de análise de ordem interpessoal (funções retórica e pragmática dos constituintes do SN), representacional (tipo de entidade semântica designada pelo núcleo do SN e a relação semântica entre núcleo e constituinte deslocado) e morfossintática (constituição morfossintática dos constituintes do SN e

peso estrutural do constituinte deslocado). Com base nas análises, a conclusão é a de que a descontinuidade na estruturação do SN tem um comportamento estável ao longo do tempo e é motivada por fatores de ordem mais pragmática do que morfosintática.

Palavras-chave: sintagma nominal; descontinuidade; diacronia; gramática discursivo-funcional.

Abstract: This work has as its main objective to examine, from the perspective of Functional Discourse Grammar (HENGEVELD; MACKENZIE, 2008), which are the pragmatic, semantic and morphosyntactic motivations underlying the ordering of the constituents of the discontinuous Np in different synchronies of Brazilian Portuguese, in order to verify whether the synchronic characterization of the phenomenon has a diachronic basis and whether there are differences in the productivity of the phenomenon in speech and writing. The empirical analysis of the phenomenon is based on samples from different historical sources: (i) letters from the 18th, 19th and 20th centuries from the *Projeto Para a História do Português Brasileiro* (PHPB) and from the *Corpus Histórico do Português Tycho Brahe*; (ii) readers' letters from the 21st century from the newspapers *Folha de São Paulo* and *Diário do Norte*; and (iii) spoken language records from the 20th century, extracted from the minimum *corpus* of the NURC/Brazil, and from the 21st century, taken from the *Iboruna* database (*Amostra Censo*). The methodology includes analysis parameters of interpersonal (rhetorical and pragmatic functions of the Np constituents), representational (type of semantic entity designated by the head noun and the semantic relationship between the head and the displaced constituent) and morphosyntactic nature (morphosyntactic configuration of the Np constituents and structural weight of the displaced constituent). Based on the analyses, the conclusion is that the discontinuity in the structuring of the Np has a stable behavior over time and it is motivated by factors of a more pragmatic order rather than morphosyntactic.

Keywords: noun phrase; discontinuity; diachrony; functional discourse grammar.

Recebido em 02 de dezembro de 2021

Aceito em 02 de fevereiro de 2022

1 Introdução

Este artigo busca descrever, com base na Gramática Discursivo-funcional (doravante, GDF), proposta por Hengeveld e Mackenzie (2008), o comportamento de sintagmas nominais (doravante, SN) que apresentam ordenação não canônica de seus constituintes, denominados

por Keizer (2007, p. 264) de *SN descontinuos*.¹ Com base em amostras de escrita, dos séculos XVIII ao XXI, e em amostras de fala, dos séculos XX e XXI, o objetivo é investigar as motivações que possam estar na base da ativação da descontinuidade, bem como verificar possíveis mudanças diacrônicas que expliquem a configuração desse tipo de SN na gramática do português brasileiro (doravante, PB) contemporâneo, de modo a constatar se o fenômeno é resultante de processo de mudança e, portanto, inovador, ou se é estável na gramática do PB. Essa perspectiva diacrônica se justifica com base na ideia defendida por Bybee (2016, p. 167) de que a mudança linguística não é fenômeno periférico para teorias linguísticas, e, portanto, sincronia e diacronia devem ser consideradas dimensões integradas, para uma melhor compreensão das representações cognitivas de padrões linguísticos sincrônicos. Além disso, entender como as estruturas surgem na gramática nos aponta possibilidades de explicações nem sempre disponíveis em descrições puramente sincrônicas. Segundo a autora, padrões morfossintáticos podem ser sincronicamente arbitrários, porque resultam de longas trajetórias de mudanças; assim, a única fonte segura de explicações de suas propriedades é a diacrônica. Caracteriza-se como descontinuo o SN que comporta interferência de elementos na adjacência de seus constituintes, como em (1), ou que tem seus próprios constituintes repositados de forma peculiar, como em (2).²

- (1) e até agora depois da Ordem de Sua Excelencia tudo está a meu cargo em mandar fazer **exercícios todos os Domingos, do manejo das armas** (Século XVIII – PHPB/ SP: Cartas da administração colonial em circulação pública, Documento 48, linha 29)
- (2) vai fazê(r) um:: banheiro no quarto de::le... e vai/ ta acaban(d)o um banhe(i)ro... vai começá::(r)... só tem assim o buraco [Doc.: pra fazê(r) o banhe(i)ro] pra pa fazê(r) o banhe(i)ro né?... da minha mã::e... vai quebrá(r) **aquela parede que eu te falei do meu quarto** vai pintá(r) ele de::/ agora eu num decidi muito bem a cor (Século XXI – Banco de dados Iboruna: AC-006, DE, linha 301)

Como se observa, nas ocorrências anteriores, enquanto em (1) há a interveniência de um elemento pertencente ao nível da oração e, portanto, externo ao domínio do SN (a saber, *todos os domingos*), em (2) há a ordenação não prototípica de dois modificadores pós-nucleares (*do*

¹ No original: “discontinuous noun phrases”.

² Ao longo deste artigo, ao final de cada ocorrência exemplificativa das análises, são indicados o século a que ela pertence e a fonte de onde ela foi extraída (cf. seção de metodologia).

meu quarto e que eu te falei) do núcleo *parede*. Neste trabalho, ambos os tipos de configuração do SN são considerados descontínuos.

Com base em Keizer (2007) e para explicar o fenômeno de SN descontínuos, é necessário recorrer à atuação de princípios que determinam a ordenação de seus elementos. Neste trabalho, consideram-se dois princípios: o de peso comunicativo e o de peso estrutural. O primeiro, evocado por Dik (1997, p. 403) como *Princípio de Saliência Pragmática*, determina que estruturas que veiculam informação saliente ou nova no discurso sejam posicionadas mais ao final da expressão (no caso deste trabalho, ao final do SN); já o segundo, evocado pelo mesmo autor como *Princípio de Complexidade Crescente*, determina que as estruturas sejam ordenadas da menos para a mais complexa. Esses princípios podem favorecer uma mesma ordem, isto é, o elemento mais complexo é também focal e/ou saliente, o que garante que ele ocorra ao fim do sintagma, ou podem entrar em competição, cada um favorecendo uma ordem diferente. Nesse caso, o usuário da língua seleciona a ordem de elementos que melhor cumpre seu objetivo comunicativo.

Diante dessa caracterização inicial, e tomando a descontinuidade dos constituintes internos do SN como fenômeno a ser investigado diacronicamente, são triplos os objetivos deste trabalho: (i) verificar como os princípios de Saliência Pragmática e de Complexidade Crescente ativam a descontinuidade; (ii) examinar, nas diferentes sincronias do PB, parâmetros pragmáticos, semânticos e morfossintáticos que, correlacionados à descontinuidade, destacam a configuração desse tipo de SN, bem como observar se esses parâmetros apresentam comportamento distinto em cada uma das sincronias, o que apontaria para uma mudança diacrônica na caracterização do fenômeno; (iii) comparar como a descontinuidade de SN se manifesta na fala e na escrita do PB contemporâneo, considerando que, na modalidade falada, podem surgir estratégias típicas da interação verbal face a face não verificáveis na modalidade escrita.

Este artigo se organiza em quatro seções. Na primeira seção, apresenta-se o suporte teórico para a discussão da descontinuidade e da ordem de elementos no SN. Na segunda seção, caracterizam-se o *corpus* de análise e os procedimentos metodológicos. A terceira seção traz a descrição e a análise dos resultados, examinando-se as motivações da descontinuidade nas diferentes sincronias do PB. A quarta seção é destinada às considerações finais.

2 Descontinuidade na ordenação de constituintes

García Velasco (2010, p. 412) reconhece, de modo geral, que o fenômeno da descontinuidade na ordenação de constituintes diz respeito à presença de material linguístico interveniente na morfossintaxe linear dos constituintes de uma expressão linguística. O reconhecimento da descontinuidade, no entanto, depende da concepção de *constituente* que cada teoria linguística assume. De acordo com Huck e Ojeda (1987 apud GARCÍA VELASCO, 2010, p. 415), *interpretação semântica constante, dependência sintática e unidade semântica* constituem três critérios distintos de que teorias linguísticas lançam mão para a abordagem da descontinuidade. Sob o primeiro, *constituente* é definido como “uma sequência fonética [que] mantém a mesma contribuição semântica quando seus membros aparecem tanto contíguos como não [...]” (HUCK; OJEDA, 1987, p. 5 apud GARCÍA VELASCO, 2010, p. 415).³ Sob o segundo, um *constituente* se define com base em “relações de dominância em uma configuração arbórea à qual transformações por movimento possam subsequentemente ser aplicadas, produzindo assim vários níveis de análise sintática”.⁴ O terceiro critério estabelece que “elementos sintáticos não contíguos não podem formar um constituinte, embora possam ser mapeados em uma representação semântica na qual suas interpretações formem uma unidade”.⁵ Diante dessas três concepções teóricas de *constituente*, Garcia Velasco (2010) defende que, para a abordagem do fenômeno da descontinuidade no âmbito da GDF (HENGEVELD; MACKENZIE, 2008), a terceira noção de *constituente* é a que se mostra mais adequada, tendo em vista que esse quadro teórico adota, para a caracterização do componente gramatical das línguas, uma organização em níveis hierarquicamente estruturados: Interpessoal (NI) > Representacional (NR) > Morfossintático (NM) > Fonológico (NF), cada um deles comportando camadas de organização própria do nível. Para o tratamento do fenômeno da descontinuidade, essa arquitetura

³ No original: “a phonetic sequence [that] maintains the same semantic contribution both when its members appear contiguously and when they do not.”

⁴ No original: “relations in a tree configuration to which movement transformations can subsequently apply, thus producing various levels of syntactic analysis.”

⁵ No original: “Non contiguous syntactic elements cannot form a constituent, although they may be mapped onto a semantic representation in which their translations form a unit.”

da GDF permite que constituintes representados separadamente no Nível Morfossintático (NM) possam ser interpretados adequadamente no Nível Representacional (NR), conforme se observa na formalização simplificada do exemplo dado em (3).

- (3) Aquelas polpas que cê compra de maracujá (AC-090, RP, linha 389) (SOUZA-MARTINS, 2020, p. 91).

NI: (A_i: [(R_i: [(T_i) (T_j)_{FOC} ...] (T_j)] (A_i))

NR: (p_i: [(x_i) aquelas polpas de maracujá (x_i) (e_i) cê compra (e_i) (p_i)

NM: (Np_i: aquelas polpas (Np_i)) (^{dep}Cl_i: que cê compra (^{dep}Cl_i)) (Pp_i: de maracujá (Pp_i))

Em casos como esse, a relação entre o NR, no qual se codificam representações semânticas, e o NM não é transparente, ou seja, não é de um-para-um, já que uma única entidade semântica do NR (x_i: *aquelas polpas de maracujá*) é codificada no NM em duas posições diferentes ((Np_i: *aquelas polpas* (Np_i) e (Pp_i: *de maracujá* (Pp_i)). Esse desalinhamento entre os níveis viola os princípios de ordenação que regem as relações entre o NM e os dois níveis mais altos, a saber: o princípio de Iconicidade (relação não arbitrária entre forma e conteúdo), o de Integridade de Domínio (preferência universal pela justaposição, no NM, das unidades que, juntas, pertencem ao NI e ao NR) e o de Estabilidade Funcional (preferência por colocação de constituintes com mesma especificação pragmática ou semântica em mesma posição fixa relativamente a outras categorias) (HENGEVELD; MACKENZIE, 2008). A atuação conjunta desses princípios contribui para a maximização do paralelismo entre as estruturas, porque reforça a transparência da estrutura linguística e facilita sua interpretação, mas qualquer um deles pode ser violado, produzindo enunciados descontínuos, como em (3).

Considerando o modelo de representação gramatical da GDF e os propósitos deste artigo, o SN descontínuo em (3) (*aquelas polpas de maracujá*) é descrito em três níveis da GDF, como segue: (i) no NI, é um Subato Referencial que compõe o Ato Discursivo (A_i) (*aquelas polpas que cê compra de maracujá*); no NR, é uma entidade semântica do tipo indivíduo (x_i) e, no NM, é codificado por um sintagma nominal (Np_i) descontínuo, com o sintagma preposicionado (Pp_i) deslocado para depois do sintagma verbal (Cl_i). Assim, é um caso típico de descontinuidade, que emerge justamente “quando a ordem linear não reflete relações semânticas, no sentido de que duas unidades que são semanticamente

relacionadas aparecem separadas uma da outra” (GARCÍA VELASCO, 2010, p. 416).⁶ Em um contexto sem descontinuidade, as estruturas internas do SN dado em (3) apareceriam contíguas umas às outras no NM, como se representa em (4).

(4) Aquelas polpas de maracujá que cê compra.

NI: (A_i: [(R_i: [(T_i) (T_j)_{FOC}...] (T_j)] (A_i))

NR: (p_i: [(x_i) aquelas polpas de maracujá (x_i) (e_i) cê compra (e_i)] (p_i))

NM: (Np_i: aquelas polpas (Np_i)) (Pp_i: de maracujá (Pp_i)) (^{dep}Cl_i: que cê compra (^{dep}Cl_i))

Buscando compreender as motivações da produção de enunciados que apresentam descontinuidade entre seus constituintes, Keizer (2007) considera a atuação dos princípios de peso estrutural e de peso comunicativo como circunstâncias favorecedoras do deslocamento de constituintes que justificariam a decisão do falante de optar por uma ou por outra ordem de palavras.

O princípio de peso estrutural prevê que a ativação da ordem de palavras ocorre por meio de um processamento cognitivo de otimização das estruturas da língua que explica a tendência de estruturas menos complexas ocorrerem em posição inicial, e as mais complexas, em posição final da expressão linguística (HAWKINS, 1983), tendência denominada por Dik (1997, p. 404) de *Princípio de Complexidade Crescente* (“*Principle of Increasing Complexity*”). Segundo esse princípio, prevalece, nas línguas, uma preferência por ordenar os constituintes internos de uma expressão linguística, observando-se, entre eles, a complexidade estrutural crescente, o que significa que, na estruturação da expressão linguística, quanto maior a complexidade de um constituinte, maior é a possibilidade de ele ocorrer em posição posposta ao de menor complexidade.

Equiparado ao princípio de peso estrutural, o *Princípio de Saliência Pragmática* (“*Principle of Pragmatic Highlighting*”), como Dik (1997, p. 403) nomeia, estabelece que constituintes de uma expressão linguística podem comportar funções pragmáticas especiais e, por isso, ocorrem também em “posições especiais”. É a atuação desse princípio que explicaria o deslocamento de constituintes para a posição tanto inicial quanto final da oração (como em (5) e em (6), respectivamente), em razão da intenção do falante de realçar uma informação em seu discurso.

⁶ No original: “when linear order does not reflect semantic relations, in the sense that two units which are semantically related appear separated from each other.”

- (5) *Doutores* sempre houve muito poucos. (PEZATTI, 2014, p. 98)
- (6) Um anúncio foi feito *de que ele estava indo para o Departamento de Educação e Ciência*. (KEIZER, 2007, p. 288, exemplo adaptado e traduzido livremente).

Enquanto a sentença em (5) é exemplar de casos em que constituintes portadores de informação conhecida são deslocados de seu domínio próprio para a posição inicial da oração, típica da função pragmática de tópico no português brasileiro, a sentença (6) exemplifica o chamado peso comunicativo (KEIZER, 2007, p. 267): o SP com função focal (*de que ele estava indo para o Departamento de Educação e Ciência*) é deslocado para a posição final da oração, geralmente reservada a constituintes portadores de informação nova no discurso. Cumpre observar, na estruturação da sentença em (6), a convergência dos dois princípios, porque o constituinte deslocado, ao mesmo tempo que é portador de informação focal, é estruturalmente mais complexo do que os demais que com ele compõem a sentença. Assim, os dois princípios atuam a favor de uma mesma ordenação de constituintes e, juntos, explicam a descontinuidade no contexto dado.

Com base em Dik (1997), Keizer (2007) elabora as seguintes hipóteses explicativas sobre a ação mútua dos dois princípios: (i) os princípios de peso estrutural e de peso comunicativo atuam de modo independente na preferência do falante por deslocar material do SN para fora dele; (ii) a atuação dos dois princípios, na maior parte dos casos, leva à mesma ordem; (iii) se os dois princípios favorecem a competição entre ordens diferentes, pelas circunstâncias discursivas, o falante decide qual deles suplanta o outro na eficiência comunicativa; (iv) em caso de empate, outros fatores independentes podem ser o fiel da balança, mesmo que não tenham por si força suficiente para determinar a ordenação dos constituintes; (v) mesmo violando um dos princípios de ordenação, o falante julga que, na atual circunstância discursiva, a ordem por ele escolhida é a mais eficaz, dentre as disponíveis. A análise da autora, para dados do inglês, comprova que os dois princípios, de fato, são relevantes para a explicação da descontinuidade de constituintes do SN.

No presente trabalho, em que se objetiva descrever como a descontinuidade na estruturação do SN do PB se comporta desde o século XVIII até o XXI, à luz dos princípios anteriormente descritos, será também considerado o emprego de outras estratégias linguísticas que, possivelmente, possam exercer algum papel na determinação da ordem dos constituintes, como funções retóricas, mecanismos de mitigação e de modalização, preservação das relações de escopo, dentre outras.

3 Metodologia

A análise empírica de SN descontínuos toma por base ocorrências recolhidas das seguintes fontes:

(i) cartas (pessoais, oficiais, de leitores e de redatores) datadas dos séculos XVIII, XIX e XX, que compõem os *corpora* do Projeto Para a História do Português Brasileiro (PHPB) (disponível em: <https://sites.google.com/site/corporaphpb/home>) e *Corpus* Histórico do Português *Tycho Brahe* (disponível em: <http://www.tycho.iel.unicamp.br/corpus/>);

(ii) cartas de leitores do século XXI, coletadas dos jornais *Folha de São Paulo* (disponível em: <https://www1.folha.uol.com.br/folha/sinapse/cartas.shtml> e <https://www1.folha.uol.com.br/paineldoleitor/>) e *Diário do Norte* (disponível em: <https://www.jornaldiariodonorte.com.br/>);

(iii) amostras de língua falada do século XX, que compõem o *corpus* mínimo do Projeto NURC/Brasil (Norma Urbana Linguística Culta) (disponível em: <http://www3.iel.unicamp.br/cedae/>);⁷

(iv) amostras de língua falada do século XXI do Projeto Amostra Linguística do Interior Paulista (ALIP) (disponível em: <http://www.alip.ibilce.unesp.br>), que registra, no Banco de dados Iboruna, a variedade do PB falado no noroeste paulista.⁸

O recorte temporal escolhido, que parte do século XVIII, leva em conta as postulações de Mattos e Silva (2004), que considera que a gramática do PB se define em meados do século XVIII para o século XIX, sendo esse século o mais propício a uma “expansão” da língua portuguesa no Brasil, uma vez que o século XVIII serviu de ponto de partida a esse processo.

⁷ As amostras de fala do Projeto NURC/Brasil compõem-se de inquéritos coletados nas cidades de Porto Alegre (POA), Recife (REC), Rio de Janeiro (RJ), Salvador (SSA) e São Paulo (SP), em três diferentes estilos: DID (Diálogo entre documentador e informante), D2 (Diálogo entre dois informantes) e EF (Elocução formal). Do *corpus* mínimo, fazem parte: de D2, os inquéritos REC-05; SSA-98, RJ-355, SP-360, POA-291; de DID, os inquéritos REC-131, SSA-231, RJ-328, SP-234, POA-45; de EF, os inquéritos REC-337, SSA-49, RJ-379, SP-405, POA-278.

⁸ O banco de dados Iboruna compõe-se de dois tipos de amostras: Amostra Censo, com 151 entrevistas sociolinguísticas, e Amostra de Interação, com 11 interações dialógicas coletadas secretamente, com consentimento posterior dos participantes. Para o presente trabalho, considerou-se apenas a Amostra Censo.

Considerando a língua falada como lugar privilegiado para o aparecimento de descontinuidades (RIJHKOFF, 2002; SOUZA-MARTINS, 2020), elegeu-se, inicialmente, o gênero *carta pessoal* como fonte de dados históricos, com base na pressuposição de que a escrita de cartas particulares, sendo mais informal (se comparada a documentos oficiais), seria a que mais se aproximaria da modalidade falada.⁹ Posteriormente, em vista do escasso número de ocorrências encontradas nesse tipo específico de carta, expandiu-se a coleta de dados para qualquer tipo de carta, a saber, cartas oficiais, de leitores, de redatores, entre outras.

Quanto aos séculos mais recentes, XX e XXI, em que há a possibilidade de coletar dados a partir de transcrições de língua falada, incluíram-se também o *corpus* mínimo do Projeto NURC/Brasil e as entrevistas sociolinguísticas da Amostra Censo do banco de dados Iboruna. Com relação a esse último tipo de amostra, reanalisam-se aqui os mesmos dados já analisados por Souza-Martins (2020). O objetivo, então, é fazer uma comparação entre as ocorrências de língua falada dos séculos XX e XXI, em busca de diferenças entre essas duas sincronias.

Elencam-se, no quadro 1, características das diferentes amostras utilizadas na composição de nosso *corpus* de análise, com relação aos tipos de texto, e o número aproximado de palavras de cada uma. O quadro 2, por sua vez, sintetiza a quantidade de palavras das amostras de cada século, como tentativa aproximada de balanceamento do *corpus*, e a quantidade de dados encontrados em cada sincronia analisada.

⁹ Desnecessário dizer que não adotamos aqui uma concepção dicotômica de escrita e de fala, mas compartilhamos da ideia de Mattos e Silva (2008, p. 21) de que, “sendo a documentação escrita que permanece e sendo ela uma representação convencional da fala, teremos o reflexo da fala, que permite tirar conclusões, até certo ponto, seguras, no nível morfofônico, já que, não havendo uma normatização ortográfica, a análise da variação da escrita oferece indícios para alguma percepção da voz, ou seja, da língua no seu uso primeiro, em qualquer dos níveis em que se pode estruturá-la: fônico, mórfico, sintático, discursivo.”

Quadro 1 – Sistematização das amostras representadas no *corpus* (tipos de texto e número de palavras)

Século	Fonte das amostras	Tipo de texto	Palavras
Século XVIII	PHPB – Pernambuco	Cartas Oficiais	13.974
		Cartas de comércio e cartas comuns	14.745
	Cartas pessoais e documentos particulares		
	PHPB – Bahia	Cartas oficiais	41.712
	PHPB – São Paulo	Cartas de aldeamentos de índios	8.736
		Cartas oficiais	11.493
Cartas da administração colonial (circulação pública)		25.174	
Século XIX	<i>Tycho Brahe</i>	Cartas Brasileiras: cultos	42.082
		Cartas para vários destinatários	4.565
	PHPB – Rio de Janeiro	Cartas escritas no RJ (documentos particulares)	5.939
		Cartas de leitores e redatores	29.776
		Documentos oficiais	5.681
	PBPB – São Paulo	Cartas de aldeamentos de índios	1.246
		Cartas oficiais	7.573
	PHPB – Pernambuco	Cartas pessoais	1.275
		Cartas de leitores	10.841
	PHPB – Bahia	Cartas de administração privada	15.328
Século XX	PHPB – Pernambuco	Cartas pessoais	38.281
		Cartas de leitores	19.963
	<i>Tycho Brahe</i>	Cartas Brasileiras: cultos	21.001
	PHPB – Bahia	Cartas particulares	45.640
	NURC/Brasil	Textos falados	74.358

Século XXI	Folha de São Paulo	Cartas de leitores	105.482
	Diário do Norte	Cartas de leitores	7.732
	ALIP/ Iboruna	Textos falados da Amostra Censo ¹⁰	1.000.000 ¹¹

Fonte: Autoria própria

Quadro 2 – Resumo da quantidade de palavras das amostras e de dados por sincronia

Língua escrita	Século	Quantidade de palavras	Quantidade de SN descontínuos
Língua escrita	Século XVIII	115.834	15
	Século XIX	124.486	17
	Século XX	124.885	13
	Século XXI	113.214	6
	Total	478.419	51
	Língua falada	Século XX	74.358
Século XXI		1.000.000	77
Total		1.074.358	124

Fonte: Autoria própria

Para que se obtivesse uma amostragem homogênea da língua escrita, buscou-se controlar a quantidade de palavras lidas, mantendo-se um intervalo de 110 a 125 mil palavras por século. Quanto à amostra de língua falada, leu-se todo o material do *corpus* mínimo do NURC/Brasil (referente ao século XX) para que se pudesse comparar com os dados já analisados por Souza-Martins (2020) (referentes ao século XXI). Ainda que não haja uma equivalência entre a quantidade de palavras de cada uma dessas amostras de fala em relação às amostras de escrita, isso já nos parece suficiente para fornecer dados relevantes para o objetivo pretendido.

Com relação à quantidade de dados encontrados, o balanceamento do *corpus* por número de palavras reflete a distribuição também equilibrada dos

¹⁰ As entrevistas da Amostra Censo registram 5 tipos de textos orais: NE (narrativa de experiência), NR (narrativa recontada), DE (descrição de local), RP (relato de procedimento) e RO (relato de opinião).

¹¹ Conforme Gonçalves (2019).

dados entre as duas modalidades, pois, embora o número de dados de língua falada seja maior do que os de língua escrita, ele é proporcional à quantidade de palavras lidas, que também é maior nas amostras faladas. Em outras palavras, há mais dados de língua falada porque mais palavras foram lidas nas amostras dessa modalidade. Mesmo se cercado de cuidados com o equilíbrio do *corpus*, percebe-se a baixa frequência de SN descontínuos em ambas as modalidades, o que conduz à opção pelo desenvolvimento de análises mais qualitativas do que quantitativas, a fim de se evitarem generalizações e afirmações categóricas. Com base nessa ponderação, muito provavelmente a ampliação do *corpus* de análise com aumento do número de palavras das amostras em todas as sincronias não produziria resultados frequenciais muito diferentes dos aqui apresentados, dada a produtividade do próprio fenômeno na língua em uso.

Como procedimento metodológico fundamental, realizou-se uma leitura exaustiva de todos os textos, uma vez que SN descontínuos podem se constituir de qualquer forma lexical, o que inviabiliza o uso de recursos de buscas automáticas por constituintes específicos. Somente por meio de busca manual, não eletrônica, torna-se possível identificar todos os SN potenciais alvos de ordenação descontínua, considerando-se sua constituição interna, a saber, núcleo, operadores, argumentos e modificadores pré e pós-nucleares, para, assim, detectar se há descontinuidade no que diz respeito a interrupções de elementos externos à estrutura do SN ou à ordenação não prototípica de seus próprios constituintes.

Finalizada essa etapa, estipulou-se um grupo de parâmetros que abrangem propriedades de ordem interpessoal (pragmáticas), representacional (semânticas) e morfossintática do SN, visando à descrição da natureza do SN descontínuo e à investigação das circunstâncias sob as quais ele se instancia na língua. Dentre os parâmetros utilizados para verificar a motivação da descontinuidade, buscou-se identificar, por um lado, se algum constituinte do SN (núcleo, material interveniente, modificadores ou argumentos separados do núcleo) exerce alguma função interpessoal (pragmática ou retórica) ou é enfático; por outro lado, também se avaliou a complexidade morfossintática das estruturas deslocadas para fora do domínio do SN, se são mais ou menos complexas que o material interveniente. Esses dois critérios levam em conta os dois princípios evocados por Keizer (2007) como motivadores da descontinuidade no nível do SN.

Além desses parâmetros, consideraram-se, também, outros que poderiam influenciar na organização dos elementos do SN, como estratégias linguísticas de mitigação, modalização e preservação das relações de escopo entre os constituintes, especialmente nos casos em que o material interveniente tem relação apenas com um elemento da construção. Como exemplo dessa possibilidade, em (7), a expressão “vamos dizer” tem escopo sobre o modificador “especial” como tentativa de mitigar a incerteza do falante com relação ao uso do termo.

- (7) mas não tive ainda **um... motivo vamos dizer especial** mesmo a não ser quando o meu marido às vezes tem que... conversar alguma coisa. (Século XX – NURC, DID POA 45)

Como parâmetros representacionais, verificou-se se haveria uma relação entre o tipo de entidade semântica (indivíduo, estado-de-coisas, proposição)¹² denotada pelo núcleo do SN e a motivação da descontinuidade, permitindo, também, apontar qual entidade semântica o SN descontínuo prototipicamente designa. Além disso, analisou-se a relação que se estabelece entre núcleo do SN e o constituinte deslocado, se este é argumento ou modificador do nome, buscando-se investigar qual tipo de elemento tende a permitir interrupções e deslocamentos para longe do núcleo.

Descreveu-se morfossintaticamente a constituição formal dos constituintes, considerando-se a presença de operadores, modificadores pré e pós-nucleares e argumentos (se são adjetivos simples, sintagmas preposicionais ou orações), com o objetivo de definir qual tipo de material morfossintático tende a intervir no SN e qual tende a ser deslocado para fora dele.

Aplicados os parâmetros, por fim, foi possível detectar quais fatores predominam na motivação da descontinuidade dos constituintes do SN e qual a configuração do SN descontínuo prototípico. Além disso, comparou-se tanto o comportamento desse tipo de sintagma em cada

¹² Conforme distinção ontológica de entidades possíveis de serem representadas por expressões linguísticas, Lyons (1977) reconhece os seguintes tipos de entidade: (i) de 1ª. ordem, *indivíduo*, entidade que ocupa lugar no espaço e, por isso, é avaliada pela sua existência; (ii) de 2ª. ordem, *estado-de-coisas*, entidade que reporta evento que, localizável no tempo e no espaço, ocorre, dura, e só pode ser avaliada em termos de realização; (iii) de 3ª. ordem, *proposição*, entidade que refere construtos mentais, sem lugar no tempo e no espaço e, por isso, avaliada apenas em termos de verdade/falsidade.

sincronia do PB, em busca de diferenças que apontassem ou não para uma mudança ao longo do tempo, quanto a caracterização do fenômeno com relação às modalidades escrita e falada da língua, com a intenção de captar especificidades típicas das ocorrências em cada uma delas.

4 Análise diacrônica do SN descontínuo

Entendendo a descontinuidade como uma ordem de constituintes escolhida, ainda que inconscientemente, pelo usuário da língua para estrategicamente organizar informações de seu enunciado, presumiu-se que sua presença seria detectada nos séculos mais remotos do PB. De fato, já no século XVIII, é possível verificar manifestações de SN descontínuos, o que prova que esse fenômeno não é um modo novo de estruturação da língua, como ilustram os exemplos em (8) e (9).

- (8) bem| pudera tratar aqui da **conveniencia** *que terà oinimigo* **em vir bombear a villa dapar-|te de fora**, ficando tam impossibiLitado para aganhar, como antes de abombear [...] (Século XVIII – PHPB/PE, Cartas Oficiais, Carta 9, linha 50)¹³
- (9) e por elles me-requererem pelo **in comó** | **doque padecem em mandar vir as cartas deUanças para as no-|vas Justiças**, procedi a Pillouros nas ditas villas [...]” (Século XVIII – PHPB/BA, Cartas oficiais, Carta 78, linha 13)

Em ambas as ocorrências, tem-se, de modo similar, uma ordenação não prototípica dos elementos pós-nucleares do SN: o complemento do núcleo (“em vir bombear a villa dapar-|te de fora”, em (8), e “em mandar vir as cartas deUanças para as no-|vas Justiças”, em (9)) é posicionado após o modificador (“que terà oinimigo”, em (8), e “que padecem”, em (9)). Conforme os princípios de peso comunicativo e de peso estrutural (KEIZER, 2007, p. 267), esse distanciamento entre núcleo e complemento aciona, ao mesmo tempo, a natureza focal do complemento, por introduzir informação nova e/ou saliente no discurso, e sua maior complexidade estrutural como mecanismos estratégicos de otimização da organização das estruturas da língua. Como prova disso, as paráfrases em (8’) e (9’), dadas a seguir, demonstram que, embora a ordem complemento-modificador seja esperada, ela não é a mais eficiente nesse contexto, uma vez que

¹³ Nas ocorrências exemplificativas, decidiu-se manter a transcrição original do corpus, o que justifica a falta de espaçamento entre algumas palavras e a ortografia distinta da norma-padrão do PB atual.

produz ambiguidade e estranheza, exigindo do leitor um esforço maior para interpretar o sentido pretendido pelo escritor.

- (8') bem| pudera tratar aqui da **conveniencia em vir bombear a villa dapar-|te de fora**, que terá oinimigo ficando tam impossibiliLitado para aganhar, como antes de abombear [...]
- (9') e por elles me-requererem pelo **in comó | do em mandar vir as cartas deUzaças para as no-|vas Justiças que padecem**, procedi a Pillouros nas ditas villas [...]

Ainda com relação aos dois princípios atuantes na descontinuidade do SN, é necessário considerar que eles nem sempre operam na mesma direção, isto é, favorecem a mesma ordem de constituintes. Em alguns casos, pode haver competição entre eles, em razão de a ordem final dos elementos se caracterizar como uma decisão do usuário da língua que leva em conta o que é mais relevante para a construção em questão. Essa situação de competição foi observada já no século XVIII, conforme mostra o exemplo em (10).

- (10) Quando fui para hesidade beiar os pes de vossa excellenca não tive Lugar dedar **Contas pelamuita oCopasam emque vossa excellenca SeaChava naocaziam sobre algreia da Aldea denosasenhora daesCada**, aqual tenho Reetificado aminha CuSta por servir anosasenhora casuamagestade [...] (Século XVIII – PHPB/SP, Aldeamento de índios, Carta 12)

Nesse caso, o material interveniente (“pelamuita oCopasam emque vossa excellenca SeaChava naocaziam”) é um pouco mais complexo do que o complemento (“sobre algreia da Aldea denosasenhora daesCada”) e, portanto, deveria estar posicionado mais ao final da expressão, para que, assim, houvesse o atendimento ao princípio de complexidade crescente. O que acontece, na verdade, é a violação desse princípio, motivada pelo atendimento ao princípio de peso comunicativo, que faz com que a informação mais saliente seja alocada na posição final. Em outras palavras, entre obedecer ao princípio de peso estrutural e o de peso comunicativo, numa competição de motivações, o escritor decide pelo último. Quanto a essa decisão, uma confirmação de que a posição do complemento “sobre aIgreia da Aldea denosasenhora daesCada” ao final da oração se deve ao seu caráter focal é o fato de que ele inicia o tópico desenvolvido na sequência, como evidencia, no contexto mais amplo da ocorrência, reproduzido em (10'), o conteúdo temático da carta: relatar a respeito da igreja da aldeia.

- (10') Quando fui para hesasidade beiar os pes de vossa excellenca não tive Lugar dedar **Contas pelamuita oCopasam emque vossa excellenca SeaChava naocaziam sobre aIgreia da Aldea denosasenhora daesCada**, aqual tenho Reetificado aminha CuSta por servir anosasenhora casuamagestade *que deos guarde para que* asim **SeaumentaSeasua Aldea que estava iaqua ge para cair noxam por Ser feita deparede deman em aLgû tempo eantes deterRetificado a dita Igreja pasando o muito reverendo padre e vigitador odoutor aLeixandre marques dovaLe ovigitou e naõChan do os ornamentos que opadre bispo ordenava e estar so mente empoder dos indios oquis fexar e deRubar epelo muito Rogo que lhe fis odeixou ficar mas deixando hû edital para que dentro de hû anno sefizesse todos os ornamentos nesarios (Século XVIII – PHPB/SP, Aldeamento de índios, Carta 12)**

Por outro lado, um fator que também parece motivar a ordenação das estruturas nesse contexto, em especial do material interveniente, é a intenção do escritor de fazer uma ressalva ao que é dito anteriormente: tenta-se amenizar o fato de dizer “não tive lugar de dar contas” com a inserção da justificativa de que “vossa excelência se achava muito ocupada na ocasião”. Essa explicação poderia ter sido introduzida na posição inicial (“pelamuita oCopasam emque vossa excellenca SeaChava naocaziam, não tive Lugar dedar Contas sobre aIgreia da Aldea denosasenhora daesCada”), mas, por algum motivo, o escritor decide inseri-la posteriormente. Em contrapartida, posicionar essa retificação ao fim do enunciado não parece ser a melhor escolha, já que poderia causar ambiguidade, além de interferir no tópico já iniciado e mais relevante (a igreja da aldeia). Desse modo, a expressão construída cumpre o propósito comunicativo para o qual é empregada, ainda que haja descontinuidade.

Comportamento semelhante ao dos exemplos do século XVIII, analisados anteriormente, se verifica também em ocorrências dos séculos XIX, XX e XXI, em amostras de escrita, como se pode notar de (11) a (13).

- (11) porque lhe restão muitas outras [p]osseções, como Cabo verde [etc] para onde precisa transportar Gente para novos estabelecimentos, e em 2 o quando seja impossivel conservar-se aquella convenção, não annuir a **infame condi-ção, que admetio o Brasil de serem considerados os Contrabandis-tas, como Piratas,** [...] (Século XIX – PHPB/RJ, Cartas escritas no Rio de Janeiro, Carta 10, linha 16)

- (12) O Rui mandou-me um cartão de visita, quando cheguei e pelo Car los mandou-me pedir que eu fosse jantar com ele domingo (5 de outubro,) para conversar mos sobre as bases do requerimento de informações e do **protesto, que ele enunciará da tribuna do senado contra esse pretendido direito da União a 18.000:000\$ de garantia de juros à retra [...] inglesa** (Século XX – Tycho Brahe, Cartas brasileiras, Carta 289, linha 3013)
- (13) É estarrecedor tomar conhecimento por meio da Folha de que a ANS (Agência Nacional de Saúde Suplementar) simplesmente deixou de cobrar das seguradoras de saúde particulares os milhões de reais pela **utilização que elas fizeram de recursos públicos** para atender a seus pacientes durante 2008 e 2009 [...] (Século XXI – Jornal Folha de São Paulo, Cartas de leitores, <https://www1.folha.uol.com.br/paineldoleitor/secaodecartas/935732-parada-gay-planos-de-saude-bonde-de-santa-teresa-corinthians.shtml>)

A tendência que se observa é a de elementos que veiculam conteúdos focais e que são mais complexos serem distanciados do núcleo do SN, como ocorre em (11) e (12). Em alguns casos, como em (13), não há maior complexidade morfossintática, mas apenas a veiculação de informação nova e/ou saliente, que geralmente é desenvolvida na sequência. Sendo assim, o princípio de peso comunicativo parece ter maior influência na determinação da ordem dos elementos, visto que nem sempre o elemento deslocado é complexo, mas sempre é focal e/ou saliente. Essa constatação indica que, na grande maioria dos dados analisados das diferentes sincronias, esse princípio atua em conjunto com o de peso estrutural, mas também pode operar sozinho, o que não acontece com o de peso estrutural, que, nos dados analisados, apenas atua em conjunto com o de peso comunicativo.

Quanto ao material interveniente na estruturação do SN, ele tende a veicular informação compartilhada/dada, menos relevante para a sequência discursiva. Morfossintaticamente, tende a assumir, nos dados analisados, a forma de uma oração relativa, como nas ocorrências dadas de (11) a (13): “que admetio o Brasil”, “que ele enunciará da tribuna do senado” e “que elas fizeram”, respectivamente. Já o elemento distanciado do núcleo do SN assume a forma de um sintagma preposicional (SP): “**de** serem considerados os Contrabandistas, como Piratas”, em (11), “**contra** esse pretendido direito da União a 18.000:000\$ de garantia de juros à retra inglesa”, em (12), e “**de** recursos públicos”, em (13). Isso se deve ao fato de que, mesmo distanciados linearmente, esses SP são facilmente identificados como parte integrante do SN por conta do seu

vínculo semântico ainda preservado. Essa possibilidade é prevista num modelo de gramática como a GDF, pois, uma vez que a relação entre núcleo e modificador/complemento é definida no Nível Representacional, o leitor/ouvinte pode, então, estabelecer as relações semânticas necessárias na decodificação da mensagem, apesar da descontinuidade e do distanciamento dos constituintes tal como codificados no Nível Morfossintático (GARCÍA VELASCO, 2010; SOUZA-MARTINS, 2020; VAN DE VELDE, 2012). Em (14), a representação formalizada de (13) permite visualizar esse desalinhamento entre os Níveis Representacional e Morfossintático.

(14) A utilização que elas fizeram de recursos públicos.

NR: (e_i: [f_i: (f_i: utilização (f_i)) (x_i: recursos públicos (x_i))_U (f_i)] (e_i) : (e_j: elas fizeram (e_j)) (e_i))

NM: (Np_i: a utilização (Np_i)) (^{dep}Cl_i: que elas fizeram (^{dep}Cl_i)) (Pp_i: de recursos públicos (Pp_i))

Pode-se perceber que, embora haja separação entre núcleo (“utilização”) e complemento (“de recursos públicos”) no NM, eles permanecem semanticamente unidos no NR, compondo uma única propriedade configuracional,¹⁴ em que “utilização” é o predicado e “recursos públicos”, o seu argumento.¹⁵ É por conta desse vínculo semântico estabelecido no NR que a interpretação esperada do enunciado, mesmo sendo ele morfossintaticamente descontínuo, não é prejudicada.

Ainda em termos semânticos, a entidade designada pelo núcleo é quase sempre um conteúdo proposicional, como “condição”, em (11), ou um estado-de-coisas, como “protesto”, em (12), e “utilização”, em (13), enquanto o elemento deslocado tende a ser complemento do núcleo. Essas observações de caráter semântico, no entanto, não são vistas como motivações para a descontinuidade no SN, mas servem para definir como se manifesta, prototipicamente, o SN descontínuo.

Com base nessas considerações, constata-se que a descontinuidade no SN exhibe um padrão prototípico, apresentando uma estabilidade

¹⁴ A Propriedade Configuracional (f^c) (uma propriedade que consiste em mais de uma unidade) é a camada que funciona como núcleo do Estado-de-Coisas (e) e, é, por sua vez, nucleada por uma propriedade (f) e um ou mais Indivíduos (x) (KEIZER, 2015).

¹⁵ Na representação, o símbolo U subscrito indica que “recursos públicos” é o argumento com função semântica Inativo (*Undergoer*), isto é, o participante exercendo um papel passivo no Estado-de-Coisas (KEIZER, 2015).

sintático-semântica-pragmática ao longo do tempo, como é possível confirmar pela análise de ocorrências de diferentes sincronias do PB, no que tange à modalidade escrita. Quanto à modalidade falada, é necessário apontar algumas especificidades.¹⁶

Em interações verbais face a face, o planejamento e a execução linguísticos ocorrem concomitantemente, diferentemente do que acontece na escrita, em que é possível realizar revisões e alterações no que se pretende enunciar. Frente a essa simultaneidade entre planejamento e execução, o usuário da língua lança mão de diferentes estratégias linguísticas, como correções, reformulações, entre outras, para minimizar possíveis mal-entendidos na interpretação de seu enunciado. Alguns mecanismos dessa natureza foram observados apenas nos dados de língua falada, o que reforça a ideia de que são estratégias mais frequentemente utilizadas em interações verbais face a face, embora não seja impossível que ocorram em certos contextos de produção escrita. Como exemplo dessas estratégias, detectou-se o uso da função retórica Esclarecimento (ou Correção) em dados de língua falada tanto do século XX como do século XXI. Essa função, para Hengeveld e Mackenzie (2008) e Keizer (2015), serve para esclarecer ou corrigir alguma informação mencionada anteriormente que o falante julga incompleta ou inadequada do ponto de vista comunicativo. As ocorrências em (15) e (16) ilustram o emprego dessa função.

(15) L2 porque:...chega um ponto

L1 que o **acúmulo** é muito grande né? **de... processos** (Século XX – NURC: D2 SP 360)

(16) Doc.: tem bastante computador lá?

Inf.: não... computador lá num tem... porque tem o **laboratório bem do lado tam(b) ém né?... de computação** (Século XXI – Banco de dados Iboruna: AC-053, DE, linha 211)

Percebe-se que o falante, inicialmente, não julga necessário expressar o complemento “de processos” do núcleo “acúmulo”, em (15), e o modificador “de computação” do núcleo “laboratório”, em (16), por se tratar de informação contextualmente acessível ou inferível pelo

¹⁶ Limitamo-nos a destacar os pontos particulares dos dados de língua falada que se distinguem dos de língua escrita. Isso significa que, quanto ao restante dos parâmetros de análise, não houve diferenças relevantes a serem abordadas aqui.

ouvinte. Contudo, ele decide acrescentá-la posteriormente num ato de esclarecimento,¹⁷ revelando sua preocupação em antecipar informações que poderiam ser solicitadas por seu interlocutor (“acúmulo de quê?”, em (15), e “laboratório de quê?”, em (16)). Como se vê, a descontinuidade, nesses casos, tem motivação interpessoal, especificamente retórica, pois é caracterizada pela organização estratégica das estruturas em prol do objetivo comunicativo do usuário da língua, que, por sua vez, busca garantir que o ouvinte interprete adequadamente seu enunciado.

De modo similar, o falante pode recorrer a outras estratégias linguísticas, como mitigação e modalização, a fim de salvar sua face e de evitar possíveis mal-entendidos na interação verbal. Muitas vezes, o emprego de tais estratégias produz enunciados descontínuos, como se pode notar em (17) a (20).

- (17) o indivíduo... nao pode figurar... como o senhor todo poderoso... ele tem que... repartir... **aquele poder... digamos assim... de exclusivis:mo..** com os seus... assessores imediatos... (Século XX – NURC: DID REC 131)
- (18) poderíamos inclusive estabelecer... como uma das diferenças... **a questao por exemplo acredito eu que... da assistência MEDica hospitalar...** que eu acredito que as cooperativas nao... prestam... aos seus associados (Século XX – NURC: DID REC 131)
- (19) às vezes a pessoa num BEbe... às vezes a pessoa num num é desequilibRAda... mas ela num tem **uma condição...** êh:: eu diria que:: **psicológica suficiente** pra educá(r) uma criança... (Século XXI – Banco de dados Iboruna: AC-099, RO, linha 466)
- (20) Há cinco departamentos há vários departamentos mas... o prédio principal... é o E um o prédio principal... havia antes pra inauguração o projeto pra tê(r) **dois prédios de oito andares se eu num me engano... de vidro...** (Século XXI – Banco de dados Iboruna: AC-081, DE, linha 120)

Nessas ocorrências, o uso de expressões como “digamos assim”, “acredito eu que”, “eu diria que” e “se eu num me engano” indica a intenção do falante de atenuar certos enunciados ou de mostrar

¹⁷ Morfossintaticamente, tem-se uma articulação dentro do domínio da Expressão Linguística via extraoracionalidade, o que significa dizer que um ato de esclarecimento está fora dos limites do sintagma e da oração. O que se propõe aqui como descontinuidade, na verdade, é uma relação semântica ainda preservada de elementos que, morfossintaticamente, se encontram distanciados, embora não seja possível propor uma relação de constituição morfossintática entre eles.

incerteza ou descomprometimento com relação ao conteúdo deles. Nota-se que a inserção de tais expressões no domínio do SN, embora cause descontinuidade, se justifica pela preservação das relações de escopo, pois, geralmente, elas têm escopo em apenas uma parte do SN: de (17) a (19), as expressões atenuam o conteúdo do que é enunciado em seguida, enquanto, em (20), a incerteza recai sobre o conteúdo do enunciado anterior (a quantidade de andares nos prédios). A ordenação interna das estruturas, nessas circunstâncias, é, portanto, a mais eficiente em termos de propósitos comunicativos, por levar em conta a intenção do falante, por um lado, e as relações de escopo, por outro.

Por fim, ainda vale mencionar interferências na linearidade do SN motivadas por uma estratégia de manutenção da interação face a face, como mostram as ocorrências de (21) a (24).

- (21) ele está se referindo exatamente a essa essência tradicional da economia japonesa tá? quer dizer uma **uma situação** (,) *eu vou repetir*(,) **muito diferente do início da economia americana**, tá dando pra situar a diferença? (Século XX – NURC: EF RJ 379)
- (22) as incursoes (ruído -ou aquilo que) eu estou rotulando de incursoes foram **quaisquer tipos de quê? de relações**, em função de aumento de ampliação de território (Século XX – NURC: EF RJ 379)
- (23) dependendo o o que fô(r) feito no cabelo o corte de cabelo que você qué(r) fazê(r)... faz **o pezi::nho entendeu? do cabelo**... em volta da ore::lha... (Século XXI – Banco de dados Iboruna: AC-072, RP, linha 394)
- (24) Inf.: olha éh... se/ lá teria que sê(r) **um quarto diferente como disse agora pouco d'uma pessoa normal**... apesar da idade dela... mas ainda tá um pouquinho bagunçado acho que ainda num num lembrô(u) que ela já passô(u) daquela idade ainda né? (Século XXI – Banco de dados Iboruna: AC-103, DE, linha 311)

Novamente, tem-se uma inserção de expressões no domínio do SN, mas, agora, com papel de manutenção da interação verbal, sinalizando o desejo do falante de assegurar que seu interlocutor está acompanhando seu raciocínio e interpretando adequadamente seu enunciado. Em (21) e (22), por se tratar de um contexto de sala de aula, o professor (que é o locutor) preocupa-se em cativar a atenção de seu auditório e, para isso, faz uso de expressões como “eu vou repetir” e “o quê?”. Já em (23) e (24), numa conversação mais informal, o falante emprega o marcador discursivo “entendeu?”, para confirmar o entendimento do interlocutor, e a expressão, “como disse agora pouco”,

para resgatar informação previamente introduzida no discurso. De modo geral, tais exemplos demonstram que o falante faz escolhas a favor de seus propósitos comunicativos, mesmo que elas sejam em detrimento da adjacência morfossintática das estruturas linguísticas, produzindo, conseqüentemente, a descontinuidade.

Com relação à distinção entre modalidade escrita e modalidade falada, é nítido que, em termos de interação verbal face a face, existem estratégias linguísticas contribuindo para o cumprimento dos objetivos comunicativos do falante, o que não ocorre, pelo menos não da mesma forma, na modalidade escrita, que dispõe de manifestações mais regulares do fenômeno da descontinuidade, isto é, apresenta um padrão de ocorrência muito similar nas diferentes sincronias do PB.

Quanto à mudança diacrônica, por outro lado, os dados encontrados e analisados não autorizam afirmar que essas estratégias linguísticas detectadas na modalidade falada são inovadoras (por terem aparecido nos dados dos séculos XX e XXI), uma vez que não é possível investigar dados de língua falada dos séculos mais remotos e verificar o ponto de origem do seu uso. Também não se pode afirmar que houve mudança entre os séculos XX e XXI, no que tange à modalidade falada, pois as ocorrências apresentam padrões muito semelhantes, o que, mais uma vez, reforça a afirmação da estabilidade do fenômeno em análise.

Pode-se, no entanto, postular uma diferença entre a natureza do fenômeno da descontinuidade em contextos falados e em contextos escritos, confirmando o que já havia proposto Souza-Martins (2020). Essa distinção, como já dito, deve ser feita em termos da situação de uso da língua, dada a especificidade de certas estratégias linguísticas em interações verbais face a face.

Além dessas estratégias presentes exclusivamente nos dados de língua falada, é importante destacar uma última diferença entre as modalidades concernente à atuação dos princípios de peso comunicativo e o de peso estrutural. Na modalidade falada, percebe-se uma presença menos significativa do princípio de peso estrutural, de modo que o princípio de peso comunicativo atua sozinho com mais frequência do que conjuntamente ao de peso estrutural, diferentemente do que ocorre na escrita, em que predomina a atuação conjunta dos dois princípios. Essa constatação pode encontrar respaldo na hipótese de que, na escrita, em face da possibilidade que tem o leitor de resgatar informações por meio de uma eventual releitura, o emprego de estruturas mais complexas

seria mais frequente,¹⁸ enquanto na fala o usuário da língua tende a evitar construções com maior complexidade, cujo processamento exigiria maior esforço cognitivo por parte de seu interlocutor. Em tais contextos, é esperado que o princípio de peso comunicativo desempenhe um papel mais relevante na definição da ordem dos elementos, havendo primazia de fatores pragmáticos em relação a fatores estruturais.

5 Considerações finais

A análise aqui empreendida teve por objetivo investigar o comportamento de SN descontínuos nas sincronias do PB, desde o século XVIII até o século XXI, buscando desvelar possíveis mudanças na manifestação do fenômeno ao longo do tempo. Além disso, propôs-se uma comparação entre as modalidades escrita e falada, a fim de pontuar diferenças ou semelhanças entre as ocorrências nas duas modalidades de uso do PB contemporâneo.

Foi possível constatar que as primeiras ocorrências de SN descontínuos datam do século XVIII, não sendo, portanto, um modo novo de estruturação da língua. Por outro lado, não foi possível verificar nenhum tipo de mudança nos dados dos diferentes séculos, o que nos permite afirmar que o fenômeno apresenta uma estabilidade sintático-semântica-pragmática na história do PB. Isso talvez se deva à própria natureza da descontinuidade, uma vez que ela, apesar de produzir enunciados que subvertem a ordem esperada da língua, atua em prol da otimização da organização das estruturas linguísticas, seja em termos interpessoais, como focalização de informações, seja em termos de processamento cognitivo, como a disposição das estruturas conforme sua complexidade crescente.

Essa estabilidade também se estende à frequência do fenômeno, que se mantém baixa em todas as sincronias e em ambas as modalidades analisadas. Esse resultado era esperado, em vista dos dados do PB contemporâneo analisados por Souza-Martins (2020), e pode ser explicado pela caracterização do fenômeno da descontinuidade, que se define como uma ordenação não prototípica e, portanto, menos esperada da língua.

¹⁸ É evidente que essas considerações sobre a escrita levam em conta o *corpus* analisado, já que, em outros tipos de textos escritos, como mensagem *online*, por exemplo, o uso de construções complexas parece ser evitado.

Em termos de motivações da descontinuidade, verificou-se, como esperado, a atuação dos dois princípios de peso comunicativo e de peso estrutural propostos por Keizer (2007, p. 267). Em todas as sincronias analisadas, eles operam a favor de uma mesma ordem de constituintes, o que implica dizer que o material distanciado do núcleo do SN é tanto focal/saliente pragmaticamente quanto complexo morfossintaticamente, sendo raros os contextos de competição de motivações. Em alguns contextos, o princípio de peso comunicativo pode atuar sozinho (o que não ocorre com o de peso estrutural), indicando que ele tem uma maior influência na motivação do fenômeno da descontinuidade.

Com relação à descrição das características do SN descontínuo, constatou-se que a interrupção de elementos no domínio do SN ou a ordenação não prototípica de seus próprios elementos ocorrem em sua porção pós-nuclear, distanciando núcleo e complemento/modificador. Esse resultado vai ao encontro do que Van de Velde (2012, p. 14) chama de “relação mais frouxa” entre núcleo e elemento pós-nuclear (seja na função de complemento, seja na de modificador), que permite, mais facilmente, interferências de fora do domínio do SN e reorganizações de seus próprios elementos. Em vista disso, a arquitetura da GDF viabiliza uma interpretação que permite uma separação entre os constituintes no Nível Morfossintático, ao passo que preserva sua conexão semântica no Nível Representacional.

Semanticamente, a entidade denotada pelo núcleo do SN é um conteúdo proposicional ou um estado-de-coisas, assumindo a forma, quase sempre, de nominalizações, e o elemento deslocado apresenta uma relação argumental com o núcleo. Morfossintaticamente, o material interveniente tende a assumir a forma de uma oração relativa, enquanto o elemento deslocado para o final da expressão tende a ser um sintagma preposicional.

Quanto às especificidades da descontinuidade em contextos falados, o princípio de peso comunicativo mostrou-se mais relevante, atuando sozinho mais frequentemente do que em conjunto com o de peso estrutural. Além da atuação desses princípios, existe uma gama de estratégias linguísticas disponíveis ao falante que contribuem para o cumprimento de seus propósitos comunicativos, cujo emprego, por outro lado, produz enunciados descontínuos. Nesse sentido, há certa primazia do interpessoal sobre o estrutural: em busca de construir enunciados mais eficientes do ponto de vista comunicativo e, também, cognitivo, o falante rompe com a ordem esperada para as estruturas da

língua. Esses resultados corroboram a adequação do modelo teórico da GDF, principalmente no que diz respeito às motivações provenientes dos níveis mais altos de formulação das expressões linguísticas (o interpessoal e o representacional) para os mais baixos de sua codificação (o morfossintático e o fonológico).

Declaração de autoria

Este artigo foi realizado de forma colaborativa pelos dois autores. A primeira autora, Nathalia Pereira de Souza-Martins, foi responsável pela elaboração e escrita da parte teórica, proposição de parâmetros de análise, organização do *corpus* e levantamento e análise de dados. O segundo autor, Sebastião Carlos Leite Gonçalves, por sua vez, foi responsável pela elaboração e sistematização da metodologia, discussão dos resultados e edição e revisão do texto do artigo. Por fim, coube aos dois autores a elaboração conjunta do resumo, do *abstract*, da introdução e das considerações finais.

Referências

ATAÍDE, C.; GOMES, V. S. *Cartas Particulares – Pernambuco*. Recife: Projeto PHPB/PE, 2010, CD-rom. Cartas Particulares. Disponível em <<https://www.ledoc.com.br>>. Acesso em: 24 jun. 2021.

ATAÍDE, C.; LIMA, T. J. S. *Corpus do Sertão do Pajeú do Laboratório de Documentação Linguística de Pernambuco*, 2016. Disponível em <<https://www.ledoc.com.br>> Acesso em: 24 jun. 2021.

BARBOSA, A.; LOPES, C. R. dos S. *Corpus diacrônico do Rio de Janeiro: cartas pessoais – séculos XVIII-XIX*. Rio de Janeiro, UFRJ/ PIBIC-CNPq/Labor-Histórico, 2003 (versão eletrônica). Disponível em: <<https://www.phpbrj.letras.ufrj.br>>. Acesso em: 16 ago. 2021.

BYBEE, J. *Língua, uso e cognição*. Tradução de Maria Angélica Furtado da Cunha. São Paulo: Cortez, 2016.

CARNEIRO, Z. de O. N. (org.). *Cartas brasileiras (1809-1999): coletânea de fontes para o estudo do português*. Vol. 2 - CD-ROM 1. Feira de Santana: UEFS Editora, 2011.

CARNEIRO, Z. de O. N. *Cartas Brasileiras (1808-1904): um estudo lingüístico-filológico*. 2005. 2360p. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, 2005. 2º volume, 3ª parte: cartas para Cícero Dantas Martins, barão de Jeremoabo.

DIK, S. *The Theory of Functional Grammar – Part I: the structure of the clause*. Berlim/New York: Mouton de Gruyter, 1997.

FERREIRA, P. S.; OLIVEIRA, K.; LOBO, T.; GONÇALVES, U. S. *Cartas Baianas Setecentistas*. 1. ed. São Paulo: Humanitas, 2001.

FOLHA DE SÃO PAULO (online). *Folha Sinapse*. São Paulo. Disponível em <<https://www1.folha.uol.com.br/folha/sinapse/cartas.shtml>> Acesso em: 24 jun. 2021.

FOLHA DE SÃO PAULO (online). *Painel do leitor*. São Paulo. Disponível em <<https://www1.folha.uol.com.br/paineldoleitor/secaodecartas/>>. Acesso em: 24 jun. 2021.

GALVES, C.; ANDRADE, A. L. de; FARIA, P. *Tycho Brahe Parsed Corpus of Historical Portuguese*. Disponível em: <<http://www.tycho.iel.unicamp.br/~tycho/corpus/texts/psd.zip>>. Acesso em: 24 jun. 2021.

GANDRA, A. S. *Cartas de amor na Bahia do século XX: normas linguísticas, práticas de letramento e tradições do discurso epistolar*. 2010. 525f. Dissertação (Mestrado em Letras e Linguística) – Instituto de Letras, Universidade Federal da Bahia, 2010. 2v. + 1 CD-ROM.

GARCÍA VELASCO, D. Discontinuity and Displacement in a Functional Theory of Grammar. 34th INTERNATIONALAEDEAN CONFERENCE ALMERIA. *Proceedings...* Oviedo: University of Oviedo, 2010. p. 412-420.

GONÇALVES, S. C. L. *Banco de dados Iboruna: amostras eletrônicas do português falado no interior paulista*. 2007. Disponível em <<http://www.alip.ibilce.unesp.br>>. Acesso em: 24 jun. 2021.

GONÇALVES, S. C. L. Projeto ALIP (Amostra Linguística do Interior Paulista) e banco de dados Iboruna: 10 anos de contribuição com a descrição do português brasileiro. *Revista Estudos Linguísticos*, São Paulo, v. 48, n. 1, p. 276-297, 2019. DOI: <https://doi.org/10.21165/el.v48i1.2430>

HAWKINS, J. A. *Word order universals*. New York: Academic Press, 1983.

HENGEVELD, K.; MACKENZIE, J. L. *Functional Discourse Grammar: a typologically-based theory of language structure*. Oxford: Oxford University Press, 2008.

HUCK, G.; OJEDA, A. E. (eds.). *Discontinuous constituency*. Londres: Academic Press, 1987.

IAPECHINO, M. N. K. *Cartas Oficiais – Pernambuco*. Recife: Projeto PHPB/PE, 2010, CD-rom. Cartas Oficiais. Disponível em <<https://www.ledoc.com.br>> Acesso em: 24 jun. 2021.

JORNAL DIÁRIO DO NORTE. *Cartas do leitor*. Disponível em: <<https://www.jornaldiariodonorte.com.br/cartas-do-leitor>>. Acesso em: 24 jun. 2021.

KEIZER, E. *A Functional Discourse Grammar for English*. United Kingdom: Oxford University Press, 2015.

KEIZER, E. *The English Noun Phrase*. Cambridge: Cambridge University Press, 2007.

LIMA, T. J. S.; ATAÍDE, C. *Banco Informatizado de Textos: cartas particulares*. Serra Talhada: Universidade Federal Rural de Pernambuco (UFRP), 2017. Disponível em <<https://www.ledoc.com.br>> Acesso em: 24 jun. 2021.

LYONS, J. *Semantics*. Cambridge: Cambridge University Press, 1977.

MATTOS E SILVA, R. V. *Caminhos da linguística histórica – ouvir o inaudível*. São Paulo: Parábola Editorial, 2008.

MATTOS E SILVA, R. V. *Ensaio para uma Sócio-História do Português*. São Paulo: Parábola Editorial, 2004.

PEZATTI, E. G. *A ordem das palavras no português*. São Paulo: Parábola Editorial, 2014.

RIJKHOFF, J. *The Noun Phrase*. Oxford: Oxford University Press, 2002.

SOUZA-MARTINS, N. P. de. *Motivações funcionais da descontinuidade sintagmática*. 2020. 96 f. Dissertação (Mestrado em Estudos Linguísticos)

- Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista “Júlio de Mesquita Filho”, 2020.

SIMÕES, J. da S.; KEWITZ, V. *Edição das cartas da capitania de São Paulo – Aldeamento de índios – Século XVIII e XIX*. São Paulo. 2005-2006. Disponível em: <<http://phpp.fflch.usp.br/corpus>>. Acesso em: 24 jun. 2021.

SIMÕES, J. da S. *Sintaticização, discursivização e semanticização das orações de gerúndio no português brasileiro*. 2007. 475f. Tese (Doutorado em Filologia e Língua Portuguesa) – Faculdade de Filosofia, Letras e Ciências Humanas (FFLCH), Universidade de São Paulo, 2007.

VAN DE VELDE, F. PP extraction and extraposition in Functional Discourse Grammar, *Language Sciences*, Leuven, v. 34, p. 433-454, 2012. Disponível em: <https://www.academia.edu/7100339/Vande_Velde_2012final_PP_extraposition_extraction> Acesso em: 03 set. 2019.



Posição dos adjetivos nos gêneros *cartas* e *narrativas* em Português Europeu do século XVII ao XIX

Adjective position in letters and narratives in European Portuguese during the 17th-19th centuries

Cristina de Souza Prim

Universidade Tecnológica Federal do Paraná (UTFPR), Curitiba, Paraná / Brasil

cristinaprim@utfpr.edu.br

<https://orcid.org/0000-0002-7199-313X>

Thais Luisa Deschamps Moreira

Universidade Federal do Paraná (UFPR), Curitiba, Paraná / Brasil

thais.deschamps@ufpr.br

<https://orcid.org/0000-0001-9066-9792>

Resumo: O trabalho analisa diacronicamente a sintaxe de adjetivos adnominais no português europeu dos séculos XVII e XIX. Para isso, utilizamos dados anotados sintaticamente e morfológicamente, presentes no *Corpus* Histórico do Português Europeu, do Projeto Tycho Brahe. Ao todo, foram analisadas cinco cartas e quatro narrativas escritas por autores nascidos entre o século XVII e XIX, o que resultou, após verificação manual dos dados, em 7.996 dados válidos. Nossa hipótese foi de que a mudança da posição preferencial de adjetivos avaliativos (da anteposição para a posposição) poderia ser observada primeiro em gêneros textuais mais próximos da oralidade e da informalidade, como a carta, e que tal mudança não ocorreria da mesma forma para sintagmas definidos e indefinidos. Os resultados confirmam a hipótese de que há, de fato, diferença no comportamento dos adjetivos quando olhamos para gêneros textuais distintos e para determinantes diferentes. Por fim, trazemos uma análise qualitativa de duas cartas, uma do século XVII e outra do XIX, a fim de avaliar a ocorrência de tipos de determinantes, tipos de adjetivos e tipos de nomes em conjunto. Os resultados mostram uma preferência pela anteposição dos adjetivos avaliativos em contextos de determinante definido e nomes contendo traço [+humano]

tanto no século XVII quanto no XIX, mas preferência pela posposição nos contextos de determinante indefinido e nome [+humano] para o mesmo recorte temporal. Ainda, vemos também um aumento no uso de adjetivos não avaliativos, que acaba por favorecer desproporcionalmente a posposição.

Palavras-chave: sintaxe dos adjetivos; diacronia; adjetivos avaliativos.

Abstract: This case-study analyzes the syntax of adnominal adjectives in 17th-19th century European Portuguese diachronically. We utilize syntactically and morphologically annotated data from the Tycho Brahe Project's Parsed *Corpus* of Historical Portuguese. A total of five letters and four narratives written by authors born between the 17th and 19th centuries were analyzed. They yielded 7,996 valid DPs after manual verification of the data. Our hypothesis was that the shift in the preferential position of evaluative adjectives (from anteposition to postposition) would be observed first in text genres closer to orality, such as letters, and that such shift has not taken place in the same way for definite and indefinite DPs. The results confirm that there is a difference in the behavior of adjectives when we look at different text genres and determiners. We also provide a qualitative analysis of two letters, one from the 17th century and another from the 19th century, in order to investigate the concomitant effect of determiner types, adjective types, and noun types. The results show a preference for the anteposition of evaluative adjectives with definite determiners and with names [+human], but a preference for postposition with indefinite determiners and names [+human], in both the 17th and 19th centuries. In addition, the use of non-evaluative adjectives increased from the 17th to the 19th century, disproportionately favoring postposition.

Keywords: syntax of adjectives; diachrony; evaluative adjectives.

Submetido em 23 de fevereiro de 2022

Aceito em 20 de abril de 2022

1 Introdução

Desde o fortalecimento da Linguística Histórica no Brasil na década de 1980, diversos aspectos da estrutura do português (em especial, brasileiro e europeu) têm sido objeto de análise sob uma perspectiva diacrônica. São relativamente numerosos os estudos sobre pronomes (e.g., mudanças no quadro pronominal, clíticos) e construções a nível frasal,

como, por exemplo, ordem dos elementos da frase, orações relativas e interrogativas (como visto, por exemplo, em ROBERTS; KATO, 2018). Já o Sintagma Nominal (SN) foi, em comparação, alvo de menos pesquisas; e é a este que nos dedicamos neste trabalho.

Em relação a adjetivos adnominais ou atributivos (isto é, adjetivos que modificam o nome-núcleo internamente ao SN), a literatura diacrônica comumente assume que sua posição preferencial tenha se alterado ao longo dos séculos – da pré-nominal para a pós-nominal (COHEN, 1990; BOFF, 1991; SALES, 2006). A aferição de tal mudança depende, por um lado, da compreensão das variáveis que condicionam o posicionamento adjetival no sintagma nominal. Nesse sentido, a sintaxe dos adjetivos no português é tema de diversos trabalhos sincrônicos, que apontam uma série de diferenças no comportamento de adjetivos pré- e pós-nominais (BORGES NETO, 1991; DESCHAMPS, 2015; MENUZZI, 1992; MÜLLER; NEGRÃO; NUNES-PEMBERTON, 2002; PRIM, 2015; SUDRÉ, 2020).

A análise do desenvolvimento diacrônico desse comportamento, por outro lado, também pode nos fornecer evidências para a compreensão das motivações da ante- e da posposição. Alguns trabalhos se propuseram a examinar o percurso diacrônico do posicionamento adjetival (BOFF, 1991; COHEN, 1990; PRIM, 2018; SALES, 2006; SERRA, 2005). Além da relativa escassez bibliográfica sobre o tema, diferenças metodológicas e de foco de cada estudo, como mostraremos, também deixam algumas lacunas em nosso entendimento desse processo. Ainda que haja fortes indícios de que houve queda quantitativa na proporção de anteposição do adjetivo ao longo dos séculos, a motivação para essa mudança não é clara. Assim, neste artigo procuramos conciliar algumas das variáveis aventadas em estudos prévios, a fim de examinar qual a contribuição de cada uma para essa mudança (se ela de fato ocorreu) na posição preferencial de adjetivos adnominais, com foco inicial no português europeu.

Abaixo, trazemos dois exemplos de dados do século XVIII e XIX para ilustrar o fenômeno observado:

- (1)[...] sem que a natureza pudesse cousa alguma contra os bisonhos costumes das nossas Madres.
(A-004,22.236, texto do século XVIII)

- (2) [...] eu considero o seu estado permanente de irritação e de azedume como a doença pior que pode atacar o espírito e o carácter de um homem.
(O-001,78.356, texto do século XIX)

O trabalho está organizado da seguinte forma: na próxima seção, apresentamos a literatura já existente sobre o tema, indicando suas contribuições e as lacunas que apontam para exploração futura. Nas seções seguintes, detalhamos nossa metodologia de pesquisa e descrevemos os dados encontrados em estudo quanti e qualitativo. Por fim, fechamos com a análise proposta e indicamos tanto nossas conclusões quanto os limites de nosso estudo.

2 A literatura especializada

Como já mencionamos, a literatura sobre o tema é relativamente escassa e conta com diferentes metodologias. Apresentaremos nesta seção tais metodologias e os resultados encontrados de Cohen (1988), Boff (1991), Serra (2005), Martinez (2009), Prim (2018) e Deschamps (2019).

Cohen (1988) faz uma análise quantitativa de 2.100 DPs extraídos de textos portugueses escritos (gêneros diversificados) entre o século XIV e o século XX. A autora selecionou dois textos representativos de cada século e buscou 150 DPs por texto, contabilizando 300 DPs por século. O quadro abaixo sumariza parte de seus resultados.

Quadro 1 - Frequência de ocorrência de anteposição e posposição do adjetivo em relação ao nome em termos percentuais

		XIV	XV	XVI	XVII	XVIII	XIX	XX
1o texto	Anteposição	63%	82%	31%	64%	72%	37%	23%
	Posposição	37%	18%	69%	36%	28%	63%	77%
2o texto	Anteposição	89%	82%	65%	78%	30%	31%	18%
	Posposição	11%	18%	35%	22%	70%	69%	82%

Fonte: COHEN (1988, p. 60).

Observando os dados acima, a autora aponta que a mudança no posicionamento dos adjetivos ocorreu no século XVIII, em que os

dados contraditórios dos textos 1 e 2 indicariam a mudança em curso. Interessantemente, a autora não considera que a mudança tenha ocorrido no século XVI, apesar de também ser possível observar dados do texto 1 e do texto 2 indo em direções contrárias. A razão para isso está nos dados do século XVII, que não reforçam o comportamento observado nos adjetivos do século XVI.

É preciso chamar a atenção para a quantidade de dados observados pela autora. Trata-se de uma limitação do trabalho, mas também da época em que o trabalho foi realizado. Hoje, contamos com ferramentas de busca automática de dados, o que nos permite observar um número muito maior de dados e chegar a conclusões mais precisas.

Ainda na década de 1990, Boff (1991), olhando para cartas oficiais e pessoais, diários, documentos e manifestos dos séculos XVII ao XX, todos escritos por autores falantes de português brasileiro, também conclui que houve uma mudança no percentual de adjetivos avaliativos¹, que teria se iniciado no século XVIII e se tornado mais evidente nos séculos XIX e XX. A hipótese da autora é que teria ocorrido uma perda no movimento opcional da posição pós-nominal para a pré-nominal dos adjetivos avaliativos. É preciso dizer que também neste caso o banco de dados utilizado pela autora é bastante restrito, formado por 446 dados no total (aproximadamente 50 por metade de século), e os textos investigados são de autores brasileiros, o que difere dos estudos de Cohen (1988).

Serra (2005), por sua vez, foca em aspectos semântico-discursivos e estilísticos para a compreensão da ordem dos adjetivos nos SNs do português europeu e brasileiro. A autora, analisando documentos notariais da primeira metade do século XVII, cartas de comércio do final do século XVIII, anúncios e cartas de redatores/editoriais dos séculos XIX e XX, conclui que

¹ Boff explica que adjetivos avaliativos são aqueles que podem ser subcategorizados por verbos de julgamento, como *julgar* ou *considerar*, por exemplo. Uma outra explicação possível é a de que adjetivos avaliativos, como *feio* e *bom*, são aqueles que podem ocorrer em ambiente de desacordo sem erro (DE CONTO, 2018). Por exemplo: se o falante A considera um certo ator bonito, mas o falante B o considera feio, não há contradição, tampouco resolução objetiva para o desacordo, já que se tratam apenas de pontos de vista distintos.

[...] levando em conta todos os corpora, do século XVII ao XX, a anteposição esteve sempre relacionada:

(1) aos adjetivos avaliativos, quer quando o núcleo é imaterial, quer quando é material,

(2) aos adjetivos de menor peso fônico com relação ao substantivo e

(3) aos adjetivos de base nominal [...]

(SERRA, 2005, p.107)

Segundo Serra, os adjetivos de base nominal (como o adjetivo público em *escrivão público*), ainda que sejam preferencialmente pós-nominais, ocorrem antepostos com mais frequência se comparados aos de base participial.

Outro aspecto que teria influência sobre o posicionamento de adjetivos é apresentado por Martínez (2009); apesar de ela não tratar do português, a direção metodológica observada pela autora é relevante para este trabalho. Martínez realizou sua pesquisa em nove textos em prosa em espanhol, correspondentes à segunda metade dos séculos XIII, XVI, XVII e XIX. De acordo com esses dados, a partir do século XVII fica evidente em espanhol a preferência pela posposição do adjetivo em relação ao nome. A autora aponta que a indeterminação do SN e a extensão do adjetivo (maior que o nome) favorecem a posposição deste - este último aspecto já corroborado para o português em Serra (2005). Quanto a fatores semânticos, a autora conclui que a anteposição é privilegiada com substantivos [+humanos], e a posposição, com substantivos concretos.

Prim (2018), seguindo a intuição de Martínez (2009), considera a influência do tipo de determinante no posicionamento dos adjetivos em um conjunto de textos de gêneros textuais diversos do século XVI ao XIX. Com base nesses dados, a autora observa que, com determinantes indefinidos, há sempre um percentual maior de posposição. Prim também conclui que a partir do séc. XVII não há mais influência da posição sintática do SN sobre o posicionamento do adjetivo. Dentre suas conclusões, destacamos que a informação contida no SN é mais decisiva para o posicionamento dos adjetivos do que o peso fônico, ao menos até o século XVII.

Por fim, Deschamps (2019) busca problematizar algumas questões metodológicas de estudos anteriores sobre o percurso diacrônico do posicionamento de adjetivos. A autora faz um levantamento de fatores que sincronicamente interferem na possibilidade de anteposição dos adjetivos avaliativos, tais como:

- (a) coordenações;
- (b) modificação dos adjetivos por advérbios (como muito, muy ou tão);
- (c) presença de comparativos/superlativos (como maior, mais belo ou outros terminados em -íssimo);
- (d) itens classificados como adjetivos pelo anotador morfológico utilizado em estudos prévios, cujo estatuto como tal é questionável (e.g., tal, tanto, numerais ordinais, etc.).

Como o efeito da presença de modificadores ou de morfologia comparativa/superlativa sobre a posição de adjetivos adnominais ainda não é bem compreendido, assim como o comportamento de coordenações de adjetivos no interior do SN, Deschamps defende que estudos sobre o posicionamento de adjetivos em português devem controlar tais variáveis para garantir a comensurabilidade de seus resultados. Ainda, a autora salienta a importância do gênero textual no posicionamento dos adjetivos, como já ressaltado por Sales (2006) e Serra (2005) — fator que deve ser considerado também do ponto de vista diacrônico.

Apesar da quantidade de pesquisas já realizadas tanto sobre a diacronia quanto sobre a sincronia de adjetivos adnominais em português, a natureza da distinção entre ante- e posposição ainda não é bem compreendida. Como a revisão da literatura apresentada demonstra, isso possivelmente se deve ao fato de esse ser um fenômeno multifacetado, de modo que o controle das variáveis em ação se faz essencial para que possamos examinar o peso de cada uma sobre a anteposição. Assim, é preciso levar em conta o gênero dos textos observados; a variação diatópica; o determinante que acompanha os adjetivos e nomes; a existência de modificação do adjetivo por advérbios e quantificadores; a presença de morfologia comparativa e superlativa; além de aspectos semânticos dos substantivos que os acompanham. Neste artigo, procuraremos pôr essa metodologia em prática, com o objetivo de verificar se o controle de tais variáveis corrobora ou se distancia de resultados de trabalhos anteriores.

3 Metodologia e Resultados

Para este estudo, utilizamos dados presentes no Corpus Histórico do Português Europeu, do Projeto Tycho Brahe. A escolha por este corpus foi motivada por dois fatores: primeiramente, por ele dispor de textos de mais de um gênero textual ao longo de vários séculos; em segundo lugar, pela anotação morfológica (i.e., por classes de palavras, flexão etc.) e sintática estar disponível para diversos textos do catálogo, o que garante maior agilidade à busca, uma vez que as buscas são feitas utilizando ferramentas automáticas disponíveis².

Delimitamos nossa pesquisa a textos de autores nascidos entre os séculos XVII e XIX, por tratar-se de um momento crucial para mudanças diversas em português – e também para os adjetivos, pelo já observado na literatura existente sobre o assunto. Para evitar variação decorrente de localidade, limitamos nossa seleção a autores portugueses.

A fim de avaliar o impacto do gênero textual no uso de adjetivos, optamos por comparar dois gêneros textuais disponíveis no Corpus: cartas e narrativas. A escolha pelas cartas foi motivada por esse gênero textual permitir uso de um registro relativamente menos formal, mais espontâneo e subjetivo, em contraposição às narrativas, em que o autor comumente não se coloca em foco e a linguagem tende a ser mais monitorada.

Assim, nosso corpus consistiu de duas narrativas e duas cartas de autores nascidos no século XVII; duas cartas de autores nascidos no século XVIII; e duas narrativas e uma carta de autores nascidos no século XIX. O pareamento da quantidade de textos por gênero textual e por período não foi possível pela ausência de textos narrativos do século XVIII já anotados morfológicamente e sintaticamente no Corpus, o que será levado em consideração na análise.

Apresentamos abaixo os textos que formam o corpus selecionado para esta pesquisa:

² A ferramenta *Corpus Search* permite a busca *online* nos textos anotados morfológica e sintaticamente que compõem o catálogo do *corpus* Tycho Brahe. A ferramenta, assim como o manual de anotação e de buscas, pode ser acessada em: <<http://www.tycho.iel.unicamp.br/corpus/texts/csquery/csquery.html>>.

Quadro 2 - Corpus da pesquisa

Código	Autor	Local de Nascimento e Falecimento	Ano de nascimento e morte	Título do texto observado
CARTAS				
m_003	Francisco Manuel de Melo	Lisboa/ Portugal – Lisboa/Portugal	1608 - 1666	Cartas Familiares, F.M. de Melo
b_008	José da Cunha Brochado	Cascais/Portugal – Cascais/Portugal	1651 - 1733	Cartas, J.C. Brochado
c_001	Cavaleiro de Oliveira (Fco Xavier)	Lisboa/Portugal – Londres/Reino Unido	1702 - 1783	Cartas, Cavaleiro de Oliveira
a_004	Marquesa de Alorna	Lisboa/ Portugal – Lisboa/Portugal	1750 - 1839	Cartas, Marquesa de Alorna
o_001	Ramalho Ortigão	Porto/Portugal – Lisboa/Portugal	1836 - 1915	Cartas a Emília, Ramalho Ortigão
NARRATIVAS				
b003	Manuel Bernardes	Lisboa/Portugal - Lisboa / Portugal	1644 - 1710	Nova Floresta
c002	Maria do Céu	Lisboa/ Portugal - Lisboa / Portugal	1658 - 1753	Vida e Morte de Madre Helena da Cruz
a003	Marquês de Fronteira e d'Alorna	Lisboa/Portugal - Lisboa / Portugal	1802 - 1881	Memórias do Marquês da Fronteira e d'Alorna
b005	Camilo Castelo Branco	Lisboa/ Portugal - Vila Nova de Famalicão/ Portugal	1825 - 1890	Maria Moisés

Após a escolha dos textos, conduzimos as buscas utilizando a ferramenta Corpus Search. Essa ferramenta permite levar em consideração as etiquetas de cada item, assim como relações de precedência³. A título

³ Optamos pela busca de dados utilizando relações de precedência e não de c-comando, por exemplo, para controlarmos melhor o ambiente de ocorrência dos adjetivos. Os casos em que não havia relações de constituintes foram manualmente eliminados para a análise final.

de exemplificação, o quadro 3 traz as buscas que foram realizadas por adjetivos antepostos com nomes comuns. As mesmas condições de busca foram também realizadas, feitas as modificações adequadas, com nomes próprios e com adjetivos pospostos.

Quadro 3 - Buscas por adjetivos na anteposição para nomes comuns

SN Indefinido - Singular	((D-UM D-UM-F iPrecedes ADJ*) AND (ADJ* iPrecedes N))
	Exemplo: “um íntimo amigo de meu pai”
SN Indefinido - Plural	((D-UM-P D-UM-F-P iPrecedes ADJ*) AND (ADJ* iPrecedes N-P))
	Exemplo: “umas estranhas humanidades”
SN Definido - Singular	((D D-F P+D P+D-F iPrecedes ADJ*) AND (ADJ* iPrecedes N))
	Exemplo: “aquela cruel matança”
SN Definido - Plural	((D-P D-F-P P+D-P P+D-F-P iPrecedes ADJ*) AND (ADJ* iPrecedes N-P))
	Exemplo: “os maus homens”
SN Nu após complementizador	((C iPrecedes [1]ADJ ADJ-*) AND ([1]ADJ ADJ-* iPrecedes N N-P))
	Exemplo: “que semelhantes comércios”
SN Nu após WHs exclamativos/interrogativos	((WD* iPrecedes ADJ*) AND (ADJ* iPrecedes N N-P))
	Exemplo: “Que galante tradução será a do nosso italiano!”
SN Nu após Interjeição	((INTJ iPrecedes ADJ) AND (ADJ iPrecedes N N-P))
	Exemplo: “Oh palpáveis trevas”
SN Nu após Pronome relativo	((WPRO* iPrecedes ADJ*) AND (ADJ* iPrecedes N N-P))
	Exemplo: “a quem soberana mão”
SN Nu após Verbos	((SR* HV* ET* TR* VB* iPrecedes [1]ADJ ADJ-*) AND ([1]ADJ ADJ-* iPrecedes N N-P))
	Exemplo: “se requer grande entendimento”
SN Nu após Preposições	((P iPrecedes [1]ADJ ADJ-*) AND([1]ADJ ADJ-* iPrecedes N N-P))
	Exemplo: “de louvável vida”

SNs com pronome possessivo - Singular	((PRO\$ PRO\$-F iPrecedes ADJ*) AND (ADJ* iPrecedes N))
	Exemplo: “seu antigo lustro”
SNs com pronome possessivo - Plural	((PRO\$-P PRO\$-F-P iPrecedes ADJ*) AND (ADJ* iPrecedes N-P))
	Exemplo: “seus diversos timbres”
Adjetivos antecedidos por numeral	((NUM NUM-F iPrecedes ADJ*) AND (ADJ* iPrecedes N N-P))
	Exemplo: “dois valentes pontapés”
SN com dois adjetivos	((ADJ ADJ-* iPrecedes [1]ADJ ADJ-*) AND ([1]ADJ ADJ-* iPrecedes N N-P))
	Exemplo: “pobre obscuro diabo”
CONDIÇÕES DE EXCLUSÃO	
Adjetivos comparativos e exclamativos	(ADJ-R* iPrecedes N N-P)
Adjetivos superlativos	(ADJ-S* iPrecedes N N-P)
Adjetivos precedidos por advérbio	((ADV* iPrecedes ADJ*) AND (ADJ* iPrecedes N N-P))
Adjetivos precedidos por quantificadores	((Q* iPrecedes ADJ*) AND (ADJ* iPrecedes N N-P))
Adjetivos em coordenações	((ADJ ADJ-* VB-AN iPrecedes [1]CONJ CONJ*) AND ([1]CONJ CONJ* iPrecedes [2]ADJ ADJ-*) AND ([2]ADJ ADJ-* iPrecedes N N-P))

Por interferirem no posicionamento de adjetivos de formas ainda não bem compreendidas, como apontado na seção anterior, foram consideradas condições de exclusão, a fim de permitir um melhor controle das variáveis:

- (1) Adjetivos coordenados, e.g. “a perseguição atroz e despótica”;
- (2) Adjetivos modificados por quantificador ou advérbio, e.g. “um jantar bastante numeroso”;

- (3) Adjetivos com complemento, e.g. “uma ovação digna do Duque”;
- (4) Adjetivos comparativos (*maior, menor, mais bela etc.*);
- (5) Adjetivos superlativos (*terminados em -íssimo*).

Ainda que fosse possível embutir algumas das condições de exclusão nas buscas diretas, optamos por realizá-las separadamente para averiguarmos seu peso em relação ao total de ocorrências de adjetivos adnominais nos textos selecionados.

Por último, todas as buscas foram examinadas manualmente a fim de filtrar:

- (a) Expressões fixas, e.g. “Cabo da Boa Esperança”, “Quartel-general”;
- (b) Erros de segmentação ou categorização, e.g. “não era preciso licença”;
- (c) Erros do anotador, e.g. “estou doente cá por dentro” (cá anotado como Nome);
- (d) Adjetivos numerais (primeiro, segundo etc.);
- (e) Adjetivos comparativos/superlativos categorizados como adjetivos comuns.

O quadro 4 apresenta os totais finais válidos (já eliminadas as ocorrências acima) em relação aos absolutos para cada tipo de SN, somando-se nomes comuns e próprios. Adjetivos com pronomes possessivos, com e sem artigo, foram mantidos numa coluna à parte.

Quadro 4 - Total final das buscas por tipo do SN (dados válidos/número absoluto de ocorrências)

SN	Indefinido		Definido		Com Possessivo		Nu	
	Pré	Pós	Pré	Pós	Pré	Pós	Pré	Pós
m003	53/65	32/44	279/404	109/146	46/58	5/12	264/472	70/114
b008	42/50	28/41	291/393	119/136	33/38	7/13	161/271	44/100
c001	59/76	63/105	335/398	189/215	43/46	5/6	219/334	84/143
a004	57/66	112/157	190/274	274/308	63/68	30/320	133/193	97/153
o001	55/61	63/89	170/230	158/179	22/26	16/18	130/181	106/139

b003	54/66	36/46	457/525	269/333	73/85	19/23	208/330	177/243
c002	33/37	31/52	291/341	82/101	26/30	1/7	107/194	36/79
a003	177/184	92/111	414/559	249/304	100/118	34/46	275/362	92/131
b005	21/22	24/31	112/140	120/130	15/16	8/8	53/67	54/66

Assim, foi observado o comportamento do adjetivo em 10.920 SNs, mas apenas 7.996 foram consideradas ocorrências válidas para as análises subsequentes.

A fim de facilitar a visualização dos dados, o quadro 5 traz a proporção de anteposição em relação ao total de anteposição+posposição do adjetivo, organizados por tipo de SN para cada texto:

Quadro 5 - Percentagem de anteposição por tipo do SN (anteposição/total de dados válidos)

Texto	Indefinido		Definido		Com Possessivo		Nu	
	Pré/ Total	%	Pré/ Total	%	Pré/ Total	%	Pré/ Total	%
m003	53/85	62,35%	279/388	71,91%	46/51	89,80%	264/334	79,15%
b008	42/70	60,00%	291/410	70,47%	33/40	82,50%	161/205	78,92%
c001	59/121	48,76%	335/524	62,62%	43/48	89,58%	219/303	72,09%
a004	57/169	33,73%	190/464	40,43%	63/93	67,74%	133/230	57,83%
o001	55/118	46,61%	170/328	51,83%	22/38	56,76%	130/236	55,32%
b003	54/90	60,67%	457/726	59,82%	73/92	80,00%	208/385	54,33%
c002	33/64	53,23%	291/373	77,16%	26/27	96,30%	107/143	74,65%
a003	177/269	65,80%	414/663	60,54%	100/134	74,05%	275/367	75,14%
b005	21/45	46,67%	112/232	47,37%	15/23	65,22%	53/107	49,53%

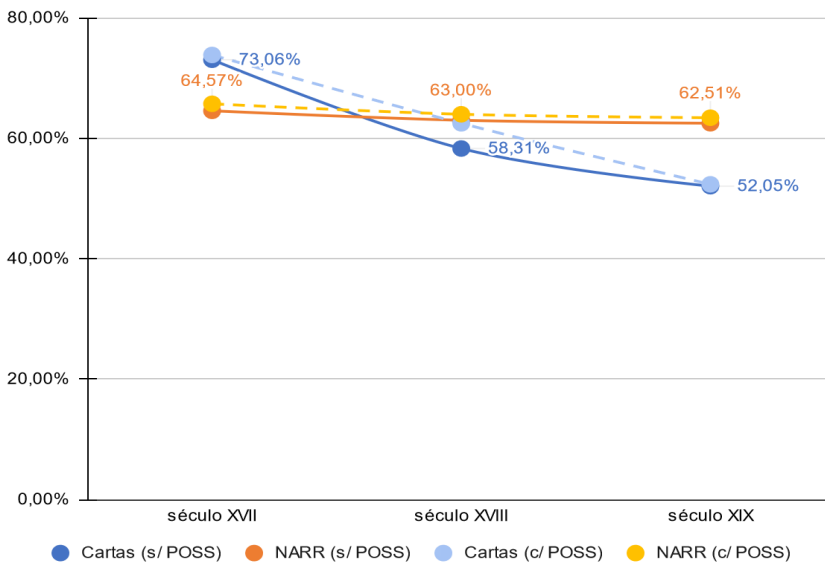
Em resumo, a partir da coleta de dados de adjetivos atributivos em cartas e narrativas dos séculos XVII ao XIX no Corpus Histórico do Português Europeu Tycho Brahe, discutiremos a distribuição pré- ou pós-nominal de 7.996 ocorrências de adjetivos atributivos, com especial atenção aos adjetivos avaliativos, em relação ao tipo de SN (indefinido, definido ou nu) e ao gênero textual (cartas ou narrativas) em que ocorrem.

4 Discussão

Considerando os resultados apresentados pela literatura especializada, nossa hipótese inicial foi de que teria havido uma mudança no comportamento dos determinantes definidos e de sintagmas nominais nus, com conseqüente mudança da posição preferencial de adjetivos avaliativos (da anteposição para a posposição). Sendo a carta um gênero textual mais próximo da oralidade pela sua hipotética maior informalidade, esperávamos que houvesse diferença entre os gêneros textuais observados: nossa expectativa é que a mudança de posição preferencial de adjetivos fosse primeiramente observada nas cartas e, mais tardiamente, nas narrativas.

Apresentamos o gráfico 1 abaixo, que reúne as percentagens de anteposição por século. As linhas contínuas apresentam os valores de anteposição e totais de adjetivos adnominais excluindo-se os SNs com pronomes possessivos, enquanto as linhas tracejadas incluem esses dados.

Gráfico 1 - Percentagem de anteposição por século e gênero textual



Ainda que não disponhamos de narrativas do século XVIII (o valor indicado no gráfico 1 para as narrativas do século XVIII é apenas a média entre os dois pontos), a proximidade entre os valores do século XVII e XIX é forte indício de que o século XVIII dificilmente fugiria ao esperado. Além disso, a diferença entre os gêneros é perceptível. Os dados acima nos mostram que há, de fato, diferença no comportamento dos adjetivos quando olhamos para gêneros textuais distintos. Vemos que, com cartas, a anteposição dos adjetivos já sofre um decréscimo no século XVIII, entrando na faixa dos 50%, índice que não é alcançado em narrativas mesmo olhando até o século XIX.

Quanto aos SNs com pronome possessivo, sua inclusão favorece sutilmente a anteposição, mas não altera os dados de forma significativa. Dada a variação que esses SNs apresentam (com e sem determinante), por ora deixaremos esses dados de lado, mas retomaremos a discussão sobre a presença dos possessivos mais adiante.

A fim de examinar os textos analisados individualmente, os gráficos 2 e 3 ilustram a proporção de anteposição por tipo de SN para cada um deles.

Gráfico 2 - Percentagem de anteposição no gênero Carta

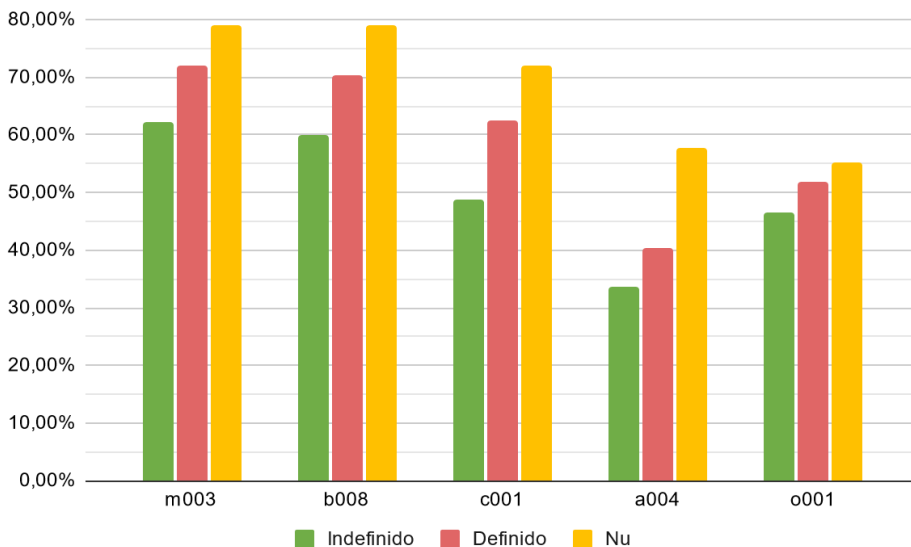
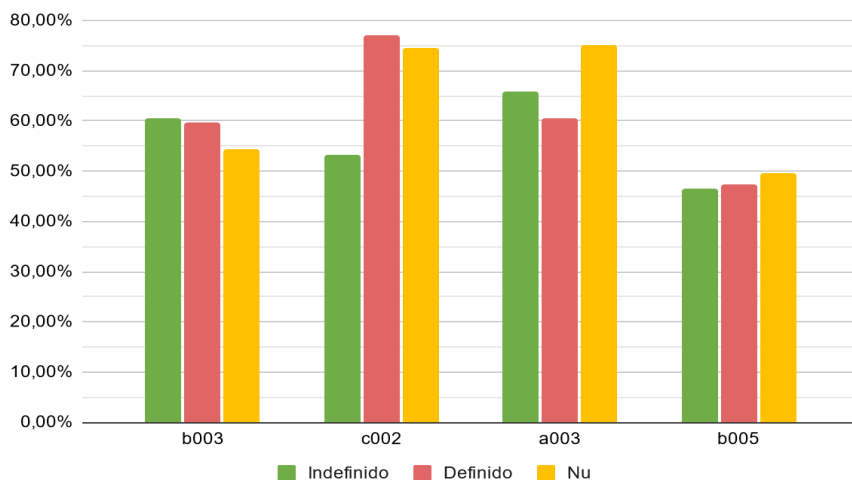


Gráfico 3 - Percentagem de anteposição no gênero Narrativa



Conforme já havíamos visto no gráfico 1, há uma queda no uso de adjetivos antepostos ao longo dos séculos. Essa queda observada é mais estável e consistente nas cartas (apesar do texto o001 representar um ponto de inflexão) que nas narrativas, em que uma redução substancial na proporção da anteposição só é atingida em um dos textos do século XIX (texto b005), como mostra o gráfico 3.

Os dados acima também nos mostram que, quando olhamos para as cartas, com determinantes indefinidos há sistematicamente uma tendência maior à posposição do adjetivo se comparado a contextos com determinantes definidos ou SNs nus. Nas narrativas, não vemos o mesmo padrão de maneira tão consistente; ainda assim, as percentagens de anteposição da última narrativa observada são bastante próximas da última carta, de modo que é possível que haja um fenômeno de maior escopo em jogo.

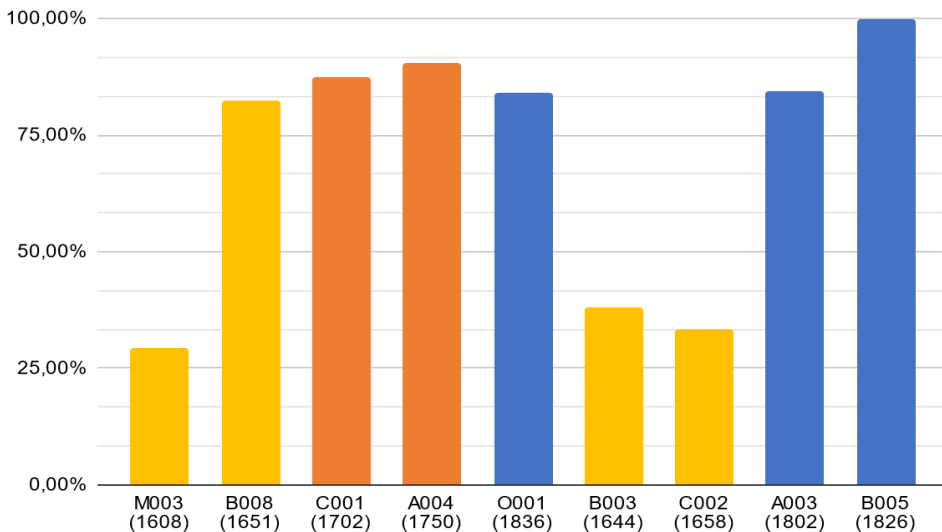
Considerando esses dados, seguiremos duas vertentes de análise que podem nos ajudar a compreender que fatores influenciaram em nossos resultados. Primeiramente, para entendermos o que há por trás da relação entre determinantes e posição dos adjetivos, exploraremos trabalhos que tratam diacronicamente dos determinantes na história do português europeu e brasileiro. Na sequência, faremos uma análise qualitativa da primeira e última cartas observadas (m003 e o001), uma vez que esse foi o gênero que apresentou maior variação.

4.1 Os determinantes na diacronia

Galves (2009) aponta que os usos dos determinantes eram mais restritos no português clássico do que no português moderno - essa transição de periodização tradicionalmente é relacionada ao século XVIII. Uma das diferenças do português clássico para o moderno é que o uso dos determinantes se mostra variável com o quantificador todo no plural (cf. GALVES, 2009) e em sintagmas possessivos, segundo Floripi (2008). Esta autora nos mostra que o percentual de presença de determinantes em SNs possessivos, em textos de autores nascidos no século XVI, variou entre 37 e 62%, a depender do contexto sintático em que tais SNs estavam inseridos; já no século XIX, o percentual chega próximo aos 100%.

Ao considerarmos a proporção de adjetivos em sintagmas possessivos com e sem artigos, vemos que os números que encontramos em nossos corpora (reproduzidos abaixo no Gráfico 4) estão de acordo com aqueles apontados por Floripi. A maioria dos textos do século XVII apresenta em torno de 30% a 40% de sintagmas possessivos com artigo definido - a exceção é b008, que já apresenta artigo em aproximadamente 83% dos sintagmas possessivos. Como b008 se trata de uma carta, podemos hipotetizar que a presença do artigo definido antecedendo um possessivo tenha ocorrido primeiramente nas cartas e mais tardiamente nas narrativas, assim como temos visto para a mudança com os adjetivos. Nos séculos XVIII e XIX, porém, não há exceções: os sintagmas possessivos apresentam artigo em mais de 84% das ocorrências em todos os textos. No gráfico 4, vemos com mais detalhes estes dados em cartas (as cinco primeiras colunas) e em narrativas (as quatro últimas colunas).

Gráfico 4 - Percentagem de sintagmas possessivos com adjetivos (pré- ou pós-nominais) e com artigo definido



Costa (2016) investigou diacronicamente os artigos definidos em português clássico, europeu e brasileiro. Ela observa uma ampliação gradual na ocorrência de artigo com nomes próprios e com nomes inalienáveis entre o português clássico e o português europeu moderno, cuja distribuição sofreria influência do tipo de nome e do gênero textual, assim como uma ampliação do uso dos artigos definidos, que está diretamente relacionada a uma mudança no comportamento dos SNs nus no português europeu.

Por hipótese, essa mudança no comportamento dos artigos definidos com nomes próprios e inalienáveis poderia estar relacionada à mudança no posicionamento preferencial de adjetivos. Dissemos anteriormente que há influência do tipo de determinante no posicionamento dos adjetivos, e essa mudança ocorrida nos determinantes definidos e nus poderia repercutir nesse posicionamento. No entanto, se seguirmos a proposta de Floripi (2008), esta não seria uma mudança de definitude. A autora argumenta que nomes nus com pronomes possessivos já possuem uma projeção de definitude mais alta, para a qual os possessivos se movem a fim de checar seus traços de definitude. O licenciamento ou

não do artigo ocorre em decorrência da relação estabelecida entre o núcleo nominal e o tipo de possessivo empregado, ou seja, é a relação de definitude/referência do núcleo nominal (nomes relacionais, partes do corpo, nomes próprios etc.) que é responsável pelo licenciamento do tipo de possessivo, se [+definido] e [+possessivo] ou se apenas [+possessivo]. Assim, apenas uma busca nos dados poderia nos fazer compreender se se trata de uma influência da definitude ou do tipo de determinante de fato no posicionamento dos adjetivos.

Utilizamos novamente a ferramenta de buscas automáticas para selecionar os contextos com determinante definido e nu junto de adjetivo e nomes relacionais ou nomes que indicam partes do corpo. Contudo, mesmo contando com um total de 7.996 dados analisados, foram encontrados poucos casos para se analisar essa hipótese, como mostra a tabela a seguir.

Quadro 6 - Cruzamento de dados que contêm determinante definido (Def) ou nu (sem det) e/ou pronome possessivo (poss) com adjetivos (Pré e Pós) e nomes relacionais (Nrel) e de partes do corpo (Ncorpo)

texto	Definido				Definido com possessivo				Nu sem possessivo	
	Ncorpo		Nrel		Ncorpo		Nrel		Ncorpo ou Nrel	
	Pré	Pós	Pré	Pós	Pré	Pós	Pré	Pós	Pré	Pós
b003	2	5	6	2	0	3	0	0	10	8
c002	0	2	1	0	0	0	1	0	2	2
a003	0	2	16	5	0	1	35	4	23	16
b005	2	11	1	5	0	0	1	0	2	6
m003	2	2	4	2	0	0	0	0	3	2
b008	1	1	0	0	0	0	0	0	3	1
c001	2	5	6	1	3	1	0	0	6	8
a004	0	4	0	0	0	2	2	1	1	2
o001	1	2	0	2	0	1	1	1	4	5

Os dados, como vemos, são insuficientes para conclusões mais assertivas. No entanto, podemos indicar uma tendência desses nomes, em especial os relacionais, ocorrerem com adjetivos pré-nominais em

contextos de DP definido, sendo que o mesmo comportamento não parece ocorrer com SNs nus.

Fizemos ainda a pesquisa com os nomes próprios. Novamente não foi encontrada uma quantidade significativa de dados, o que nos impede de estabelecer relações mais claras. Sem dúvida, seria necessário ampliar o corpus de pesquisa para que possamos estabelecer uma relação mais confiável entre nomes relacionais, inalienáveis e nomes próprios com o comportamento dos adjetivos. Deixamos essa sugestão para pesquisas futuras. De toda forma, ainda que nossos dados não nos permitam validar propostas anteriores, a concomitância da mudança no comportamento de artigos, mencionada pelos trabalhos acima descritos, e a mudança de adjetivos, abordada na seção 2, indica que parece haver fundamento para a investigação da relação entre determinantes e adjetivos adnominais.

As discrepâncias observadas no comportamento dos adjetivos entre os textos, especialmente a partir do século XVIII, ressaltam a importância de olharmos para determinantes qualitativamente. Costa (2016, p.149), por exemplo, observando o mesmo corpus Tycho Brahe com dados do português europeu, ao tratar do uso do artigo antes de nomes próprios, aponta que “nos dados do século XIX, enquanto o autor Camilo Castelo Branco apaga o artigo definido em 86% dos casos, Ramalho Ortigão só o faz em 29% das ocorrências.” Tendo este e outros resultados já mencionados como motivação, faremos a análise qualitativa de SNs de duas cartas, considerando o tipo de determinante, de nome e de adjetivo utilizado.

4.2 Análise qualitativa de duas cartas

Dado o grande volume de dados contemplados neste estudo, não foi possível realizar a análise qualitativa de todos os textos observados. Uma vez que as cartas apresentaram diferenças mais significativas que as narrativas em termos do decréscimo nas taxas de anteposição, optamos por conduzir uma análise qualitativa por amostragem de duas cartas.

Para seleção da amostra, adotamos uma baliza temporal: examinamos o conjunto de cartas mais antigo (m003, do século XVII) e o mais recente (o001, do século XIX). Limitamos nossa análise aos sintagmas definidos e indefinidos a fim de reduzir as variáveis envolvidas,

já que sintagmas nus representam uma outra ordem de fatores do ponto de vista teórico⁴.

A partir da observação feita em pesquisas anteriores e apresentada no começo deste artigo de que o decréscimo nas taxas de anteposição de adjetivos adnominais teria sido causado pela mudança na posição preferencial de adjetivos avaliativos, classificamos os adjetivos pré- e pós-nominais dos textos m003 e o001, separando-os em sintagmas indefinidos e definidos e considerando a possibilidade de mudança de posição de adjetivo, isto é, se seria possível antepor ou pospor o adjetivo em um mesmo SN sem perda de sentido. Por exemplo:

(1) [...] o grande peso, que sôbre elas tem vindo de adversidades.
(M-003,62.862)

(2) E partindo comigo dessas grandes cousas, que por lá soam, se vem e observam.
(M-003,110.1541)

(3) e o tenho por do mesmo autor.
(M-003,57.774)

(4) que as leis civis são filhas e dependentes do direito canónico;
(M-003,222.3022)

Em (3), o adjetivo anteposto poderia ser posposto sem que isso acarretasse em mudança de significado. O mesmo não se pode dizer dos demais casos. Em (4), o sentido de grande enquanto avaliação (i.e., como sinônimo de admiráveis) não está disponível na posposição; em (5), ainda que mesmo possa ser posposto, assim como em (4), haveria mudança de sentido; e em (6), por último, vemos um adjetivo classificativo, que não admite anteposição. Assim, na prática, apenas os adjetivos avaliativos têm posição variável. Nos dados, apenas o exemplo (3) foi marcado com possibilidade de mudança de posição.

⁴ A estrutura sintática que subjaz um sintagma nominal nu não é clara; por hipótese, é possível que isso tenha influência sobre o posicionamento de adjetivos. É provável que haja diferença, por exemplo, entre sintagmas nominais nus no singular e no plural (e.g., ausência de uma projeção de Número), assim como entre nus referenciais, genéricos e aqueles que denotam *kinds*.

Ainda, contamos as instâncias únicas de cada adjetivo por posição, a fim de verificar se algum em particular era responsável por um grande número de dados e poderia ter enviesado os resultados (como apontado, por exemplo, por Cohen (1989)).

Quadro 7 - Adjetivos de posição variável em comparação a todos os adjetivos em SNs indefinidos e definidos (m003), com indicação de percentual de anteposição dos adjetivos

Condições	Indefinido			Definido		
	Pré	Pós	% Ant	Pré	Pós	% Ant
Posição variável do adjetivo	44	22	66,7%	199	72	73,4%
TOTAL (variável e não variável)	53	31	63,1%	273	116	70,2%
Posição variável sem repetições	29	19	60,4%	72	54	57,1%
TOTAL (variável e não variável) sem repetições	31	28	52,5%	79	80	49,7%

Quadro 8 - Adjetivos de posição variável em comparação a todos os adjetivos em SNs indefinidos e definidos (o001), com indicação de percentual de anteposição dos adjetivos

Condições	Indefinido			Definido		
	Pré	Pós	% Ant	Pré	Pós	% Ant
Posição variável	53	31	63,1%	89	33	73%
TOTAL (variável e não variável)	56	63	47,1%	170	157	52%
Posição variável sem repetições	27	27	50%	54	27	66,7%
TOTAL (variável e não variável) sem repetições	29	54	34,9%	67	108	38,3%

Em primeiro lugar, é interessante observar como certos adjetivos são bastante frequentes na anteposição em SNs definidos: ao removermos adjetivos repetidos (isto é, contando cada item lexical apenas uma vez), essa categoria foi desproporcionalmente afetada em ambos os textos analisados. Os adjetivos grande e mesmo foram responsáveis pela maioria das ocorrências em ambos os textos: grande teve 30 ocorrências em m003 e 37 em o001, enquanto mesmo teve 44 em m003 e 21 em o001.

Ao removermos os adjetivos repetidos da contagem, vemos que a proporção de anteposição em SNs definidos em m003, que a princípio era

de aproximadamente 70,2%, mostra-se menor, 49,7% - mais próxima da proporção dos SNs indefinidos (52,5%). Já no caso do texto o001, 52% em SNs definidos cai para 38,3%, contra 34,9% em SNs indefinidos. Por outro lado, quando consideramos adjetivos de posição variável, m003 e o001 apresentam tendências contrárias: enquanto, em m003, temos uma proporção bastante próxima de anteposição em SNs indefinidos tanto com adjetivos de posição variável (60,4% em SNs indefinidos contra 57,1% em SNs definidos) quanto no geral (52,5% contra 49,7%), em o001 a taxa de anteposição com adjetivos de posição variável é muito maior em SNs definidos que em indefinidos (66,7% e 50%, respectivamente), ainda que a proporção tenha se mostrado próxima quando comparados todos os tipos de adjetivos, como já mencionado (38,3% contra 34,9%, respectivamente).

Comparando esses dois textos, ainda que tenha havido uma queda na anteposição geral tanto em SNs definidos quanto indefinidos, houve, para os definidos, aumento no uso pré-nominal de adjetivos de posição variável, de 57,1% em m003 para 66,7% em o001, e diminuição no percentual de anteposição dos adjetivos avaliativos para os indefinidos (de 60,4% para 50%), o que demonstra que, ao menos no contexto definido, a anteposição continuou a ser a estratégia preferida para adjetivos avaliativos.

Assim, esses dados apontam em duas direções. Por um lado, vemos uma queda da anteposição em SNs indefinidos, que não ocorre igualmente em definidos, o que pode indicar influência do determinante no posicionamento do adjetivo, como já apontado. Por outro, a divergência entre a queda na anteposição geral, mas aumento na anteposição de adjetivos de posição variável em SNs definidos quando comparamos m003 a o001 sugere que tais mudanças de distribuição envolvam outros fatores que não o posicionamento de adjetivos de posição variável.

Também categorizamos os SNs de acordo com os traços [+/- humano] e [+/- concreto] do nome-núcleo, sob a hipótese de que tais traços poderiam ter impacto sobre a posição do adjetivo, como já argumentado para o espanhol (MARTINEZ, 2009) e para o catalão (LIMA, 2004).

Quadro 9 - Percentagem de anteposição de adjetivos de posição variável com nomes [+humano] e [+concreto] por tipo de SN

Texto	Tipo do SN	[+HUMANO]		[+CONCRETO]	
		Total	S/ Repetição	Total	S/ Repetição
M003 (séc. XVII)	Indefinido	7/15 (46,7%)	5/11 (45,5%)	15/27 (55,6%)	9/19 (47,4%)
	Definido	32/39 (82%)	17/24 (70,8%)	52/69 (75,4%)	30/45 (66,7%)
o001 (séc. XIX)	Indefinido	5/11 (45,5%)	4/10 (40%)	29/44 (65,9%)	15/29 (51,7%)
	Definido	14/18 (77,8%)	12/16 (75%)	39/51 (76,5%)	27/38 (71,1%)

No cruzamento dos dados encontrados com adjetivos que aceitam ambas as posições com as outras categorias analisadas, verificou-se que o traço [+humano] dos nomes se mostrou mais relevante para o anteposicionamento dos adjetivos com determinantes definidos: os dados apresentados no quadro 9 indicam um aumento de anteposição dos adjetivos do século XVII para o XIX quando olhamos para nomes com traço [+humano]. No caso dos determinantes indefinidos, contudo, a presença de nomes [+humano] não parece afetar o posicionamento dos adjetivos; de fato, para os indefinidos, o traço [+concreto] se mostrou, por comparação, levemente mais relevante. A conclusão de Martinez (2009), então, parece ser parcialmente válida para o português. Também não verificamos uma tendência acentuada que indique preferência pela anteposição ou pela posposição condicionada pelo traço [+concreto] do nome em SNs encabeçados por determinantes indefinidos, mas com determinantes definidos sim.

Voltando aos quadros 7 e 8, chamamos a atenção para adjetivos em posição pós-nominal. Em m003, das 147 ocorrências totais (adjetivos de posição variável e não variável) nessa posição, 94 (63,9%) eram adjetivos avaliativos ou de posição variável. Por outro lado, em o001, foram apenas 64 (29,1%) de um total de 220. Os demais se tratam de adjetivos exclusivamente pós-nominais, que consistem de:

Adjetivos gentílicos, e.g. “colônia brasileira”;

Adjetivos descritivos, e.g. “luvas brancas”;

Adjetivos classificativos, e.g. “certeza científica”.

Nenhuma dessas categorias mudou de posição preferencial nos textos analisados. Como já reportado na literatura, tais adjetivos são pós-nominais desde pelo menos o século XIV (como apontado, por exemplo, em COHEN, 1989)⁵. Assim, o decréscimo da proporção de adjetivos avaliativos na posposição, ou aumento de uso de adjetivos exclusivamente pós-nominais, não indica uma mudança gramatical, e sim de alguma outra natureza.

Antes de finalizarmos, gostaríamos de levantar a hipótese de que as diferenças observadas se tratem de uma mudança estilística nos adjetivos avaliativos proporcionada pelos gêneros textuais observados. De fato, é necessário primeiramente analisarmos com mais cuidado as características dos textos em observação. O texto m003, chamado “Cartas Familiares”, de autoria de F.M. de Melo, consiste numa coletânea de cartas escritas pelo escritor e político durante o período em que esteve preso. As várias cartas são destinadas a uma série de interlocutores distintos: amigos, parentes, outros nobres, membros do clero, juízes, entre outros. Já o texto o001, “Cartas a Emília”, de autoria do escritor e jornalista Ramalho Ortigão, é composto apenas de cartas do escritor à sua esposa.

Ainda que haja muitos trabalhos que utilizam cartas como meio para conduzir pesquisas de cunho diacrônico, há poucos que se debruçam especificamente sobre suas características discursivas ou sobre o uso de adjetivos. Leite (2008) abordou estratégias discursivas em cartas oficiais e não oficiais, mas não abordou adjetivos em particular; já Serra (2005) estudou o percurso diacrônico de adjetivos a partir de um conjunto variado de gêneros (incluindo um subgênero das cartas), mas com foco na prosódia.

Em relação ao gênero textual, Leite (2008) salienta alguns aspectos relevantes no estudo de cartas, dentre os quais destacamos os seguintes: 1) proximidade entre os interlocutores; 2) uso de formas de tratamento; 3) tipo de carta; e 4) época histórica. Ela observa, ainda, que entre os séculos XVI a XVIII a assimetria discursiva (i.e., a distância entre os interlocutores em termos da hierarquia social) era fortemente

⁵ Ainda que Serra (2005) aponte um decréscimo do uso de adjetivos “descritivos” na anteposição entre os séculos XVII e XX, parece-nos, a partir dos exemplos citados pela autora, que os adjetivos que ela considerou como *descritivos* na anteposição são usados com valor subjetivo e seriam classificados por nós de outra maneira. Por exemplo, ela traz o dado “**eternos** inimigos” (2005, p.73) como sendo de um adjetivo descritivo, mas nós o consideraríamos avaliativo.

marcada em cartas, mesmo em relações familiares, com reflexo no uso de marcas de formalidade - assimetria esta que se torna menos díspar com o passar dos séculos.

Serra (2005), por sua vez, analisa o percurso diacrônico da anteposição ao longo dos séculos XIX e XX em anúncios e cartas ao redator/editoriais. Assim como neste artigo, a autora encontrou diferenças marcantes entre os gêneros textuais: enquanto a proporção da anteposição diminuiu de 47% para 30% nos editoriais, ela aumentou nos anúncios, passando de 36% no século XIX para 42% no século XX. Ainda que defenda que houve, sim, uma diminuição geral na anteposição em português, Serra aponta que a disparidade entre os gêneros textuais “reafirma que sua ocorrência [da anteposição] está relacionada a determinados tipos de textos e suas intenções comunicativas” (SERRA, 2005, p. 59). Ela hipotetiza que o aumento da anteposição nos anúncios esteja relacionado ao desenvolvimento de um discurso mais subjetivo, propagandístico, que almeja o convencimento do leitor - característica não presente nos anúncios do começo do século XIX.

Já em relação aos editoriais, podemos recuperar o trabalho Leite (2003), que analisou as estratégias argumentativas desse gênero também entre os séculos XIX e XX. A autora aponta que os editoriais da primeira metade do século XIX faziam mais uso de “argumentação pura” e menos uso de trechos narrativos e descritivos; em contraste, com o passar do tempo o discurso dos editoriais passa a se constituir de um maior imbricamento da argumentação com a narração e descrição. Leite argumenta que isso evidencia que passa a ser importante, para garantir o convencimento do leitor, fundamentar a argumentação na realidade, a partir do relato e descrição de fatos objetivos.

Devido ao escopo deste artigo, não caberá aqui fazer uma análise esmiuçada de nossos corpora. No entanto, podemos elencar alguns possíveis caminhos de análise que consideramos interessantes de serem explorados em trabalhos futuros.

Considerando os fatores relevantes no estudo de cartas apresentado por Leite (2008), uma primeira distinção entre m003 e o001 que se destaca é o destinatário: enquanto as várias cartas incluídas em m003 são destinadas a um conjunto heterogêneo de interlocutores, muitos deles com maior distância social do autor (como magistrados, padres, outros nobres etc.), o001 é inteiramente composto de cartas de Ramalho Ortigão à sua esposa. Nisso temos uma diferença não apenas

de interlocutor, como também de intenção comunicativa: enquanto, em suas cartas, Ramalho Ortigão (o001) em geral relata à esposa fatos de seu cotidiano, algumas das cartas de Francisco Manuel de Melo (m003) são de cunho mais dissertativo.

Se tomamos as características do discurso jornalístico ao longo dos séculos relatada em Leite (2003) como não limitadas ao estilo jornalístico em si, mas também sinalizando tendências em termos das estratégias argumentativas mais presentes na sociedade letrada, podemos especular que a combinação de interlocutor, intenção comunicativa e estratégias de argumentação de cada época faria com que m003, como um conjunto de textos do século XVII, fizesse menos uso de trechos de narração e descrição (isto é, tivesse uma linguagem menos objetiva) que o001. Essa tendência crescente ao uso de narração e descrição, combinada a uma relação familiar próxima entre os interlocutores, poderia, então, explicar o aumento significativo no uso de adjetivos pós-nominais não variáveis (classificativos, étnicos e descritivos em geral) em o001, assim como o aumento de anteposição dos avaliativos em definidos. Esses resultados, afinal, não são excludentes.

Por fim, nesse sentido, caberia um olhar mais acurado sobre as cartas, considerando o destinatário, e não apenas o gênero textual, para termos essa variável também controlada. Nosso corpus novamente nos apresenta essa limitação, e por isso deixaremos este desenvolvimento para pesquisas futuras.

5 Conclusão

Nossa hipótese de que teria havido mudança na posição preferencial de adjetivos avaliativos, causada por uma alteração no comportamento em especial de determinantes definidos, que teria se manifestado primeiramente em cartas por estas serem um gênero textual mais próximo da oralidade, foi parcialmente confirmada.

Dado o que vimos até este ponto, de fato parece haver ocorrido uma queda no uso de adjetivos antepostos ao nome ao longo dos séculos, mas que não se mantém estável quando olhamos para gêneros textuais diferentes: em cartas, há uma queda que se acentua entre os séculos XVII e XVIII, mas em narrativas os dados se mostram mais estáveis entre os séculos XVII e XIX. Esse resultado é inconclusivo na medida em que seria necessário observar dados de narrativas posteriores, a fim

de confirmar se essa foi uma mudança que se manifestou primeiramente nas cartas, mas que depois também ocorreu nas narrativas, ou se os dois gêneros se mantêm em trajetórias distintas até hoje.

Quando olhamos para os determinantes, por sua vez, também obtivemos uma confirmação parcial de nossa hipótese: vimos que, no gênero carta, os indefinidos sempre favorecem a posposição quando comparados aos definidos e nus; mas essa discrepância é menos acentuada nas narrativas.

Ainda, sobre a influência do tipo de nome, dados com nomes relacionais e nomes que indicam partes do corpo ocorreram em número inexpressivo, mas parece haver uma tendência de nomes relacionais com SNs definidos favorecerem a anteposição dos adjetivos. Um olhar mais acurado se mostra necessário para que se verifique essa hipótese.

Nosso olhar qualitativo, focado em duas cartas, uma do século XVII e outra do século XIX, buscou classificar os sintagmas definidos e indefinidos considerando a possibilidade de mudança de posição do adjetivo, as repetições dos adjetivos e os traços semânticos [+/- humano] e [+/- concreto] dos nomes. Os resultados mostraram que, desconsiderando as repetições, os dados não mostram discrepância no comportamento de adjetivos de posição variável quando olhando para sintagmas definidos e indefinidos no século XVII, mas no século XIX sim: com sintagmas definidos, encontramos 66% de anteposição do adjetivo avaliativo, contra 50% nos sintagmas indefinidos. Interessantemente, quando olhamos para todos os tipos de adjetivos, essa discrepância nos dados do século XIX se torna bastante discretas, uma vez que o aumento no uso de adjetivos de posição não variável aproximou os números de ante- e posposição. Nesse sentido, sugerimos como possível percurso de análise a ser investigado que o destinatário da carta, a intenção comunicativa e o estilo de escrita (se mais argumentativo, mais narrativo, mais descritivo etc.) sejam fatores que condicionem a proporção de anteposição em cartas.

Ainda foi observado que a queda na anteposição geral dos adjetivos, tanto em sintagmas definidos e indefinidos, não é acompanhada proporcionalmente pelos adjetivos de posição variável. Assim, observamos aumento no uso de adjetivos avaliativos antepostos ao nome nos contextos de sintagmas definidos, mas diminuição para os indefinidos. Essa divergência aponta que a queda na anteposição dos adjetivos avaliativos, defendida pela literatura especializada, não é generalizada para todos os contextos.

Sobre os traços dos nomes, verificou-se que o traço [+humano] dos nomes se mostrou mais relevante para o anteposicionamento dos adjetivos com determinantes definidos; no caso dos determinantes indefinidos e nomes [+humano], a preferência pela posposição do adjetivo já ocorria desde o século XVII. Por fim, não verificamos uma tendência acentuada que indique preferência pela anteposição ou pela posposição condicionada pelo traço [+concreto] do nome em SNs encabeçados por determinantes indefinidos.

Por fim, apesar dos vários resultados inconclusivos encontrados em nossa análise, nosso estudo nos permite afirmar com razoável grau de confiabilidade que apenas a comparação entre a proporção de anteposição e posposição de adjetivos em textos ao longo dos séculos não é suficiente para a investigação do percurso diacrônico de adjetivos atributivos. Além de termos observado diferenças sistemáticas no posicionamento de adjetivos de acordo com a (in)definitude do SN e do gênero textual, nossa pesquisa também aponta para a necessidade de se levar em conta o tipo de adjetivo, uma vez que a queda observada na anteposição não parece ser resultado somente de uma mudança na posição preferencial de adjetivos avaliativos (como já hipotetizado na literatura), como também de um aumento no uso de adjetivos descritivos, que pode estar relacionado à estilística de cada século ou do gênero textual observado.

Declaração de autoria

As duas autoras realizaram colaborativamente a coleta e a análise dos dados, a revisão bibliográfica e a redação do artigo.

Referências

BOFF, A. *A Posição dos Adjetivos no Interior no Sintagma Nominal: perspectivas sincrônica e diacrônica*. 1991. 110 p. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, 1991.

BORGES NETO, J. *Adjetivos: predicados extensionais e predicados intensionais*. Campinas: Editora da UNICAMP, 1991.

COHEN, M. A. A. M. O posicionamento do adjetivo no sintagma nominal português: um estudo diacrônico. *Boletim do Centro de Estudos Portugueses*, Belo Horizonte, v. 9/10, n. 12, p. 58-62, 1988.

COHEN, M. A. A. M. *Syntactic Changes in Portuguese*. 1990. 257 p. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, 1990.

COSTA, T. M. *Determinantes definidos: um estudo sobre a estrutura dos DPs na história do Português*. 2016. 294 p. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, 2016.

DE CONTO, L. *Tese é complicado: a leitura de situação em sentenças copulares com concordância não marcada*. 2018. 110 p. Tese (Doutorado em Letras) - Universidade Federal do Paraná, 2018.

DESCHAMPS, T. *A sintaxe dos adjetivos atributivos*. 2015. 214 p. Dissertação (Mestrado em Letras) - Universidade Federal do Paraná, 2015.

DESCHAMPS, T. Adjetivos na diacronia: fatores gramaticais e discursivos. *Revista Virtual de Estudos da Linguagem - ReVEL*, v. 17, n. 32, p. 113-136, 2019.

FLORUPI, S. *O determinante em sintagmas nominais possessivos na história do português*. 2008. 254 p. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, 2008.

GALVES, C. A Sintaxe da Grammatica. In: ABAURRE, M. B.; PFEIFFER, C.; AVELAR, J. (orgs). *Fernão de Oliveira: Um Gramático na História*. Campinas: Pontes Editores, 2009. p. 183-204.

GALVES, C.; FARIA, P. *Corpus Histórico do Português Europeu Tycho Brahe*. Campinas: Universidade Estadual de Campinas. Disponível em: <<http://www.tycho.iel.unicamp.br/~tycho/corpus>>. Acesso em: 25 nov. 2020.

LIMA, B. F. Z. de. A mudança da ordem dos constituintes adjetivo e substantivo na língua catalã: uma análise diacrônica. *Revista de Estudos da Linguagem*, Belo Horizonte, v. 12, n. 01, p. 159-174, 2004.

MARTÍNEZ, A. La frase adjetiva. El orden del sustantivo y del adjetivo. In: *SINTAXIS histórica de la lengua española*. Segunda parte: La frase nominal. Volumen 2. Coyoacán: FCE, UNAM, 2009. p. 1225-1320.

MENUZZI, S. *Sobre a Modificação Adjetival do Português: uma teoria da projeção dos adjetivos*. 1992. 194 p. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, 1992.

MÜLLER, A. L.; NEGRÃO, E. V.; NUNES-PEMBERTON, G. Adjetivos no Português do Brasil: Predicados, Argumentos ou Quantificadores?. In: ABAURRE, M. B. M.; RODRIGUES, A. C. S. (orgs.). *Gramática do Português Falado*. Vol. VIII. Campinas: Editora da UNICAMP, 2002. p. 317-344.

PRIM, C. *A Sintaxe dos adjetivos em português brasileiro*. 2015. 158 p. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, 2015.

PRIM, C. O Percurso Diacrônico dos Adjetivos Adnominais do Português Europeu: Séculos XVI ao XIX. *Fórum Linguístico*, Florianópolis, v. 15, n. 1, p. 2828-2842, 2018. DOI: <https://doi.org/10.5007/1984-8412.2018v15n1p2828>

ROBERTS, I.; KATO, M. A. *Português brasileiro: uma viagem diacrônica*. Homenagem a Fernando Tarallo. São Paulo: Contexto, 2018.

SALES, S. *A Ordem dos Adjetivos no Discurso Midiático: séculos XIX e XX*. 2006. 202 p. Dissertação (Mestrado em Língua Portuguesa) - Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2006.

SERRA, C. R. *A Ordem dos Adjetivos no Percurso Histórico: Variação e Prosódia*. 2005. 153 p. Dissertação (Mestrado em Língua Portuguesa) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2005.

SUDRÉ, T. G. *Efeitos gramaticais da classe semântica do modificador: o licenciamento e a interpretação dos adjetivos de grau relativos*. 2020. 99 p. Dissertação (Mestrado em Letras Vernáculas - Língua Portuguesa) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2020.



Nominal Number in Sign Languages

Número nominal em línguas de sinais

Bruno Gonçalves Carneiro

Universidade Federal do Tocantins (UFT), Porto Nacional, Tocantins / Brasil

brunocarneiro@uft.edu.br

<https://orcid.org/0000-0002-7417-2548>

Mônica Veloso Borges

Universidade Federal de Goiás (UFG), Goiânia, Goiás / Brasil

mvborges@ufg.br

<https://orcid.org/0000-0002-7863-7736>

Miroslava Cruz Aldrete

Universidad Autónoma del Estado de Morelos (UAEM), Cuernavaca, Morelos / México

miroslsm@gmail.com

<https://orcid.org/0000-0001-8110-4300>

Abstract: This study describes the manifestation of the category of number in sign languages, aiming to identify the values, the main forms and strategies available and some intramodal and intermodal manifestation patterns, considering the noun phrase. To this end, we worked with a sample of 10 sign languages, from different regions and historical groups. Based on secondary data, we identified that some sign languages have optional number marking, which presupposes the presence of the general number, and that other sign languages have obligatory number marking. In sign languages where number marking is optional, the general form is similar to the singular form, expressed by zero. The values of the category of number are singular, plural and dual, expressed by syntactic and morphological strategies, but preference is given for the former. This preference suggests that sign languages are isolating languages with regard to the category of number, differing typologically from spoken languages in this respect. Furthermore, the number system in sign languages appears to be phonologically driven.

The supposed trial and quadral values can be expressed by using the iconic plural, in a process of repetition of the singular form in the sign space, with a distinctive and localized pause. Therefore, they were considered to be an instance of direct counting and would be outside the number category. This strategy could create values that go beyond the amount of expected values for the category, in a typological perspective. Also, when mentioned, these values can be substituted for the plural. From the descriptions, morphological strategies suggest an implicational hierarchy, with mouthing being a rarely used strategy and reduplication with displacement being the most prevalent strategy. Another intra-modal feature is the spatial arrangement of the referent, expressed in the category of number.

Keywords: sign language typology; universals; number; plurality.

Resumo: Este estudo descreve a manifestação da categoria de número em línguas de sinais, com o objetivo de identificar os valores, as principais formas e estratégias disponíveis e alguns padrões de manifestação intramodal e intermodal, considerando o sintagma nominal. Para tanto, trabalhamos com uma amostra de 10 línguas de sinais de diferentes regiões e grupos históricos. Com base em dados secundários, identificamos que algumas línguas de sinais possuem marcação de número opcional, o que pressupõe a presença do número geral, e que outras línguas de sinais possuem número obrigatório. Nas línguas de sinais em que a marcação é opcional, a forma geral é semelhante à forma singular, expressa por zero. Os valores da categoria de número são singular, plural e dual, expressos por estratégias sintáticas e morfológicas, com predileção das primeiras sobre as segundas. Essa preferência sugere que as línguas de sinais estão isolantes no que diz respeito à categoria de número, diferindo tipologicamente das línguas orais. Além disso, o sistema de número nas línguas de sinais parece ser orientado fonologicamente. Os supostos valores trial e quadral podem ser expressos a partir do plural icônico, em um processo de repetição da forma singular no espaço de sinalização, com uma pausa distinta e pontual. Por isso, foram considerados como uma instância da contagem direta e estaria fora da categoria número. Essa estratégia poderia criar valores que extrapolam a quantidade de valores esperada para a categoria, em uma perspectiva tipológica. Além disso, quando mencionados, estes valores podem ser substituídos pelo plural. A partir das descrições, as estratégias morfológicas sugerem uma hierarquia implicacional, com o uso do *mouthing* manifestando-se como uma estratégia rara e a reduplicação com o deslocamento sendo a estratégia mais prevalente. Outra característica intramodal é o arranjo espacial do referente, expresso na categoria de número.

Palavras-chave: tipologia de línguas de sinais; universais; número; pluralidade.

Recebido em 20 de janeiro de 2022

Aceito em 15 de maio de 2022

1 Introduction

This article examines the category of number in sign languages from a typological perspective. Sign Languages Typology has been consolidated as a disciplinary field with the growing number of descriptive studies on these languages. One of its objectives is to survey similarities and differences between sign languages and between sign languages and spoken languages, thus identifying manifestation patterns both within and across language modalities.

In general, our conception of the category of number refers to our ability to group and quantify, in terms of an intuitive distinction between a single entity and a number of entities. In this process, we conceptually conceive sets which sometimes are made up of a single referent and sometimes of more than one referent.

In languages, the manifestation of this category varies evidencing the existence of a system, whose values cannot be subsumed under a simple opposition between singular and plural. According to Steinbach (2012), a comprehensive typological study on the subject in sign languages is still lacking, although there are descriptions of pluralization in individual sign languages. In this article, we present some similarities and differences in the manifestation of the category in terms of referents and, therefore, our focus is on the manifestation of nominal number.

Our objectives were to identify the values for the category of number in sign languages, considering the noun phrase, and to describe the strategies for manifesting these values, considering preference, restriction and distribution. This allowed us to identify trends in sign languages and compare them with other findings for (spoken) languages around the world. To this end, we surveyed the topic in ten sign languages using secondary data. We probably had access to prototypical constructions in these languages corresponding to the most reported phenomena.

2 Nominal number in (spoken) languages

According to Corbett (2000), some languages have complex systems that include many different number values, while in other languages number is optional, irregular and inferred by the context. Furthermore, there are languages in which number is manifested in combination with other properties.

In languages where number is optional, nouns can be expressed in a neutral way that would be outside the number system. Corbett (2000) calls this possibility “general form” or “general number,” and it depends on the speakers’ perception of the relevance of marking (or not) some number value. This and other phenomena may be restricted to part of the noun repertoire. Typologically, the general form is not widespread in languages.

Bayso is a language that allows nouns to be expressed in a general way, in situations where the manifestation of the category of number is irrelevant.

(1) Bayso language (Dick Hayward’s personal communication to Corbett, 2000, p. 11)

lúban	foofe
lion.general	watched
‘I watched lion (one or more lions).’	

lúban-titi	foofe
lion-sg	watched
‘I watched a lion.’	

lúban-jaa	foofe
lion-paucal	watched
‘I watched (a few) lions.’	

lúban-jool	foofe
lion-pl	watched
‘I watched (a lot of) lions.’	

In languages where number marking is obligatory, the meaning of the general form can be inferred by context and is shared with the singular form. A system in which the general form shares meaning with the plural form is considered non-existent, as explained by the principle of markedness. On this topic, Greenberg (1963) indicates that non-singular values are considered marked in relation to the singular value and, if a language has a neutral value for number, this must be considered the unmarked form. If singular and plural forms can also be used to express a neutral value, it is expected that the singular form is used. Turkish and Tagalog are languages that use singular/general and plural forms.

(2) Turkish language (CORBETT, 2000, p. 14)

ev	ev-ler
house	house-pl
‘house/houses’	‘houses’

(3) Tagalog Language (David Gil’s personal communication to Corbett 2000, p. 16)

aso	mga	aso
dog/dogs	pl	dog/dogs
‘dog/dogs’	‘dogs’	

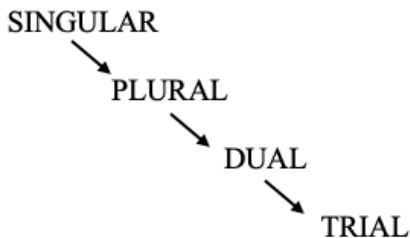
In Turkish, the form *ev* can mean *house* or *houses*, while the plural form *evler* must means *houses*. In Tagalog, the form *mga* indicates plurality, but its absence in the noun phrase leaves the possibility of singular or general meaning. Therefore, in (3) the form *aso* can indicate *dog* or *dogs*. In this type of language, the distinction between singular and general can be made, for example, by using articles and numerals.

In languages where number expression is obligatory, nouns are expressed within the values available for the category. According to Corbett (2000), languages can have the following number systems: (i) singular and plural; (ii) singular, dual and plural; (iii) singular, dual, trial and plural; (iv) singular, paucal and plural; (v) singular, dual, paucal and plural; and (vi) singular, dual, trial, paucal and plural. Thus, the category of number is expected to have maximum of five number values.

It is interesting to consider that in languages that have a *singular/plural* system, the plural form can mean both “more than one” and “two or more.” The existence of other values also implies the plural meaning. In languages that have a *singular/dual/plural* system, for example, the plural form has the meaning of “three or more real world entities”. Also according to the author, both the trial and the (rare) quadral values would be better understood as paucal, as they end up being used in other contexts.

Greemberg (1963) presents an implicational hierarchy regarding the values of the category of number. According to the author, no language has a trial number unless it has a dual, and no language has a dual number unless it has a plural. This implicational hierarchy is illustrated in Image 1.

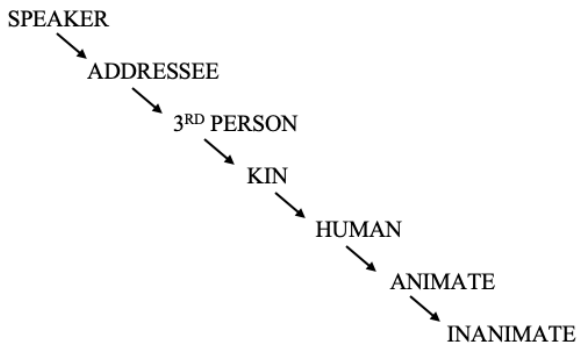
Image 1 - Implicational hierarchy of number values



Source: prepared by the authors based on Greenberg (1963, p. 74).

In languages in which number is expressed in part of the noun repertoire, animacy is an important feature for determining number marking (CORBETT, 2000; DRYER, 2013; HASPELMATH, 2013). According to Corbett (2000), personal pronouns are given preference over nouns and the first person pronoun even more so. Considering animate referents, kin-related terms are given preference over human-related terms, which, in turn, are given preference over other animate referents. Finally, the least marked class is that of inanimate referents. Image 2 illustrates this hierarchy, also considering the pronominal system.

Image 2 – Implicational hierarchy for number expression in nouns



Source: Corbett (2000, p. 57) – adapted by the authors.

Also with regard to preferences in the category of number, Haspelmath (2013) presents six possibilities of manifestation of plural

markers in the languages of the world¹, which varies along the animacy and obligatoriness dimensions. These possibilities are the following: (1) no nominal plural², (2) plural only in human nouns, optional; (3) plural only in humans, obligatory; (4) plural in all nouns, always optional; (5) plural in all nouns, optional in inanimates, and (6) plural in all nouns, always obligatory.

In the animacy dimension, the most important contrast is between animate (mainly human) and inanimate nouns. Human nouns are more likely to have plural marking than non-human (especially inanimate) nouns, which is evidenced by the fact that other logically possible values were not attested: (7) plural only in inanimate nouns, obligatory; (8) plural only in inanimate nouns, optional, and (9) plural in all nouns, optional in human nouns.

Also regarding the manifestation of the plural, according to Dryer (2013), (spoken) languages vary with some using primarily morphological strategies, which involve the use of affixes, stem change, tone change and reduplication, and other primarily syntactic strategies, which involve the use of free morphemes, including both plural words and clitics. There are also languages that use more than one strategy, none of which are considered primary. According to the author, (in spoken languages) the use of morphological strategies prevails over syntactic strategies, with the use of suffixes being the most prevalent and widespread strategy throughout the world. The following data illustrate these strategies, respectively.

¹ The author only considers full nouns, that is, excluding personal pronouns.

² Also according to the author, the non-occurrence of plural marking in languages does not mean that there is no manifestation of plural meaning. A language of type (i) having no nominal plural does not mean that only a singular meaning can be expressed, but that a non-number-marked noun form is used for both a single referent and a group of referents (general number). One example is Shigatse (Tibet), in which speakers are vague about the number of entities. The noun *ri* can mean either *mountain* or *mountains*. If they wish to be explicit, they can use numerals or quantity words (HASPELMATH, 2013). According to Dryer (2013), with regard to languages without a nominal plural, plurality can be coded on the verb.

(4) Portuguese language – affixing (suffixing) strategy (CUNHA; CINTRA, 2007, p. 195)

mesa	mesa-s
book	book -pl
‘book’	‘books’

(5) Maricopa Language – noun stem change strategy (GORDON 1986, p. 29 in DRYER, 2013, n.p)

humar	humaar	nchen	nchiin
‘child’	‘children’	‘older sibling’	‘older siblings’
hat haat		mhay	mhaa
‘dog’	dogs’	‘boy’	‘boys’

(6) Ngiti Language – tone strategy (KUTSCH LOJENGA, 1994, p. 135 in DRYER, 2013, n.p)

kamà	kámá	màlàyikà	màlàyíká
‘chief’	‘chiefs’	‘angel’	‘angels’
màlimò	màlímó	ad`òdu	ad`òdu
‘teacher’	‘teachers’	‘my brother’	‘my brothers’

(7) Indonesian Language – reduplication strategy (SNEDDON, 1996, p. 16-17 in DRYER, 2013, n.p)

rumah	rumah ~ rumah	perubahan	perubahan~perubahan
house	house ~ pl	change	change ~ pl
‘house’	‘houses’	change	‘changes’

(8) Tagalog Language – free morpheme strategy (David Gil’s personal communication to Corbett (2000, p. 134)

mga	bahay	mga	tubig
pl	house	pl	water
‘houses’		‘cups of water’	

- (9) Cayuvava Language – clitic strategy (KEY, 1967, p. 50 in DRYER, 2013, n.p)

me=rišɔ	raɓ iri
pl new	paddle
'new paddles'	

Number values can also be expressed by zero and thus inferred by context. The use of the numeral also seems to dispense with number markers in some languages, even where those markers are obligatory. Number manifestation can also be determined by certain pragmatic situations: topic versus non-topic, first mention versus subsequent mention, referential versus non-referential use, human versus non-human, definite versus indefinite (CORBETT, 2000; ELSON; PICKETT, 1978; HASPELMATH, 2013).

Still on zero expression, Greenberg (1963, p. 74) states that “there is no language in which the plural does not have some nonzero allomorphs, whereas there are languages in which the singular is expressed only by zero. The dual and trial are almost never expressed only by zero.”

In this section, we present the values and strategies of manifestation of the number category in the (spoken) languages of the world, in a typological perspective. Next, we present the methodological procedures and the languages that made up the sample.

3 Methodology

Research in linguistic typology requires a broad definition of grammatical categories in terms external to the system, thus enabling a comprehensive, reliable identification of specific linguistic phenomena, as well as the comparability between languages. In this sense, it is essential to adopt semantic and functional criteria during the investigation.

The study sample consisted of ten sign languages from diverse historical groups and areas. Grouping sign languages according to historical relationship relied on data obtained in the literature consulted, as well as on the evolution of sign languages over the 19th and 20th centuries (POWER; GRIMM; LIST, 2019). Table 1 illustrates the groups of historically related sign languages that make up the sample and indicates the data sources consulted.

Table 1 – Groups of historically related sign languages and sources consulted

Historical Relationship	Sign Language	Data Source
French SL Group	Libras (Brazil)	Ferreira (2000; 2010); Finau (2014); Lara (2017); Sanchez-Mendes; Xavier (2016); Sanchez-Mendes; Segala; Xavier (2017); Quadros; Karnopp (2004); Xavier; Barbosa (2015).
	LSM (México)	Cruz-Aldrete (2008); Smith-Stark; Cruz-Aldrete (2008)
	ISL (Ireland)	Leeson; Saeed (2012)
	NGT (Netherlands)	Zwitsersloot and Nijhof (1999),
Austrian SL Group	GDS (Germany)	Steinbach (2012), Pfau; Steinbach (2005b; 2006); Herbert (2015)
	ISL (Israel)	Meir and Sandler (2008); Stavans (1996)
British SL Group	Auslan (Australia)	Johnston and Schembri (2006).
	IPSL (India, Pakistan, Nepal)	Zeshan (2003)
Russian SL Group	ESL (Estonia)	Miljan (2003)
Isolated SL	IUR (Canada)	Schuit (2013)

Source: prepared by the authors

With respect to signing communities, nine of the sign languages are national languages that, in general, are legally recognized and have maintenance and dispersal policies and tend to be used by bilingual deaf people. Only IUR (Canada) is a village sign language.

Data for analytical and comparison purposes were collected using secondary sources. We surveyed the literature for information about noun phrases, sign formation processes, general morphological processes, word class, pronouns and classifiers. We were also attentive to the data collected. Languages for which there were many gaps in the descriptions obtained were excluded from the sample.

At first, we sought data on (1) optionality of the category and the presence of general number, (2) obligatoriness, (3) values, (4) forms and strategies available for value manifestation, (5) articulatory and semantic

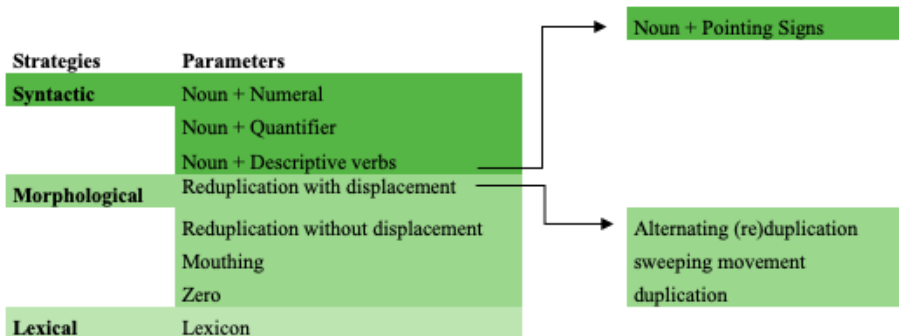
constraints, and (6) the presence of other properties in the manifestation of the category.

We were attentive to diverse possibilities of manifestation. In certain cases, it was necessary to adapt and standardize the terminologies used by the various sign language researchers to enable comparison within modalities. It was also necessary to (re)adapt the terminologies for performing inter-modal comparisons.

Forms and strategies were identified gradually over the process of collecting the descriptions of the category of number for the various languages. The list of forms and strategies was thus created simultaneously with data collection, that is, the parameters were listed as the analysis took place, and then we reexamined the available material using a cyclical approach (PALFREYMAN; SAGARA; ZESHAN, 2015).

Syntactic strategies include (1) juxtaposition with numeral signs, (2) juxtaposition with quantifiers and (3) juxtaposition with entity classifiers, which also included juxtaposition with pointing signs. Morphological strategies include (1) reduplication with displacement, which includes alternating (re)duplication, sweeping movement and doubling, (2) reduplication without displacement, (3) mouthing and (4) zero marking. The lexical strategy corresponds to the existence of signs expressing collective referents. Table 2 illustrates the list of parameters.

Table 2 – List of parameters for data collection on sign languages.



4 Nominal number in sign languages

Values and strategies in nouns

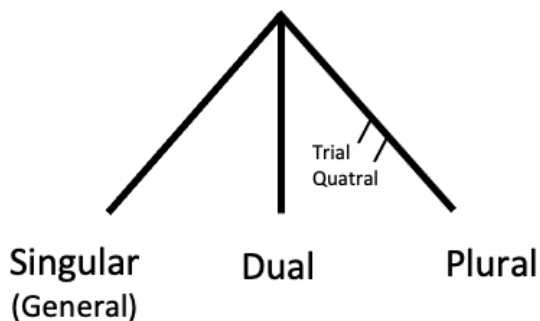
IPSL (India, Pakistan and Nepal) is a general number language, i.e. nouns can be expressed outside the category of number, while ESL (Estonia) is a language with obligatory number marking, i.e. nouns can only be expressed within the category. The descriptions we had access for the other sample languages do not provide information regarding general number.

Data on optionality and the presence of zero marking suggests the existence of general number in IUR (Canada), LSM (Mexico), Libras (Brazil), NGT (Netherlands), GDS (Germany) and Auslan (Australia). In these languages, as well as in IPSL (India, Pakistan and Nepal), the general number form coincides with the singular form. The descriptions of ISL (Ireland) and ISL (Israel) do not provide information on zero marking and optionality. Therefore, we identified sign languages where number marking is optional and others where it is obligatory.

The values we found for nominal numbers include the singular, the plural and the dual. The trial and the quadral were also mentioned, but deserve to be considered separately and will be discussed later. They are instances of the iconic plural and can be replaced with the plural. According to Zeshan (2003), iconic plural refers to a specific number above one and can represent either the number of referents or the spatial arrangement of referents or both. For example: three pointing signs (indexes) in a horizontal line representing three siblings; only the number of referents is indicated, not their spatial arrangement; three pointing signs in a vertical line representing the levels of a three-story house; both number and arrangement are indicated. About IPSL (India, Pakistan, Nepal) “in principle there is no limit to how many referents can be iconically represented, but in practice most occurrences concern numbers up to about five” (p. 185).

The singular form corresponds to the form zero and, in sign languages with the general number, the singular form also covers it. Image 3 illustrates the values and the optionality/obligatory nature of the category of number in sign languages.

Image 3 – Values for the category of number in sign languages



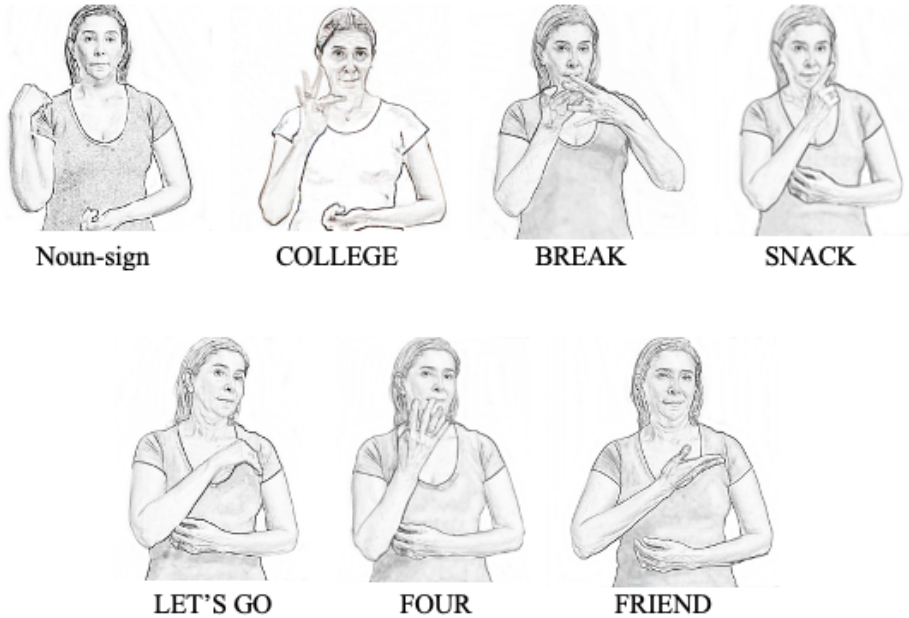
Source: prepared by the authors.

The plural is present in all sign languages and is expressed through syntactic and morphological strategies. The syntactic strategies we identified for marking the plural were the juxtaposition of nouns and (1) numerals, (2) quantifiers, and (3) descriptive verbs, and they are present in all sign languages. There are no articulatory constraints on the use of these strategies for expressing plurals. Juxtaposition of nouns and descriptive verbs³ involves the reduplication of the (verbal) form with displacement in the sign space. We also included in this strategy the juxtaposition with pointing signs as it presents the same reduplication pattern.

³ Different authors used different terminology for the strategy described above. We paid attention to the data presented in the descriptions, to ensure that they were the same phenomenon, and we chose to standardize the terminology. We adopt Liddell's terminology (2003): "depicting verbs, like verbs in general, encode meanings related to actions and states. What distinguishes depicting verbs from other verbs is that, in addition to their encoded meanings, these verbs also depict certain aspects of their meanings" (p. 261). Also according to the author, "depicting verbs can be divided at least three broad categories. The first consists of verbs signifying the presence of an entity at a place. Verbs in the second category signify the shape and extent of a surface or the extent of a linear arrangement of individual entities. Verbs in the third category signify movements of actions" (p. 262). Here, we refer to the first category. As it is a verbal form, according to Liddell (2003), it can refer the reader to the idea of pluractionality, which refers to the notion of plural events/actions, event/action performed by several agents, or even, event/action performed in several patients, or even a combination of these notions (CORBETT, 2000). Despite the use of the terminology "descriptive verbs", we consider this strategy as an instance of the nominal number.

In (10), the sign FRIEND in Libras (Brazil) is juxtaposed with the numeral four, which yields the plural reading.

(10) Libras (Brazil) - Juxtaposition with numeral sign (plural)



‘Maria, during a break between classes in college, while having a snack, called four friends.’

Source: Miranda (2020), personal collection – adapted

In (11), the DRINK and FOOD signs in Libras (Brazil) have a plural reading and are juxtaposed with the VARIOUS quantifier sign. It illustrates a syntactic strategy for expressing the plural.

(11) Libras (Brazil) – Juxtaposition with quantifier (plural)

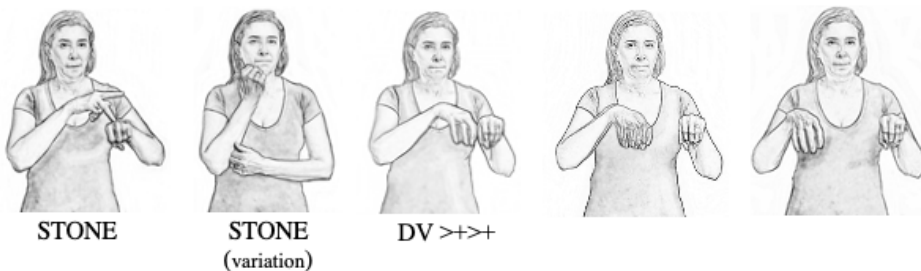


‘(...) selling food and drinks.’

Source: Entrance Examination for the Libras Undergraduate Program – UFSC Ead (2008 Edition). Available at <<http://antiga.coperve.ufsc.br/ead2008/libras/provasegabaritos.html>>. Access on: Jan. 15, 2020. – adapted

In (12) and (13), the plural is expressed using the juxtaposition of noun and descriptive verbs and the juxtaposition of noun and pointing signs, respectively.

(12) Libras (Brazil) – Juxtaposition with descriptive verbs (plural)



‘Stones’

Source: Entrance Examination for the Libras Undergraduate Program – UFSC Ead (2008 Edition). Available at <<http://antiga.coperve.ufsc.br/ead2008/libras/provasegabaritos.html>>. Access on: Jan. 15, 2020. – adapted

(13) NGT (Netherlands) - Juxtaposition with pointing signs (plural)



APLE

IX >+>+>+

'Aples'

Source: Zwitserlood and Nijhof (1999, p. 70) – adapted

The morphological strategies for marking the plural are (1) reduplication with displacement, which includes alternating (re) duplication, sweeping movement and doubling, (2) reduplication without displacement, (3) mouthing and zero marking. Pointing verbs⁴ also mark the plural, through a sweeping movement and reduplication with displacement. The following items illustrate reduplication with displacement and without displacement, respectively.

(14) HOUSE in ISL (Ireland) - reduplication with displacement (plural)



HOUSE >+>+

'houses'

Source: Leeson and Saeed (2012, p. 96) – adapted

⁴ We adopt Liddell's terminology (LIDDELL, 2003), which seems to name the agreement verbs of indication verbs

(15) NGT (Netherlands) – Reduplication without displacement



‘Problems.’

Source: Zwitserlood and Nijhof (1999, p. 61) – adapted

Regarding the number of repetitions, there seems to be a triplication, as mentioned by Steinbach (2012) and Zwitserlood (1999), although this number is irrelevant as a morphological manifestation⁵.

The use of zero marking for expressing the plural seems to be prevalent in sign languages and suggests the presence of general number, as mentioned earlier. The plural expressed by mouthing is reported only in ESL (Estonia) (MILJAN, 2003). This strategy corresponds to two forms, originating in the nominative plural form (-d), as well as in the genitive plural form (-de) of the Estonian language, which suggests a linguistic loan of grammatical forms between modalities. This strategy is widespread in ESL (Estonia)⁶.

Lexical forms for plural marking were not reported in the sample data. In ESL (Estonia) there is a sign for GROUP, a collectivity marker that also yields this meaning when juxtaposed with other signs. Table 4 illustrates the strategies for expressing plurals in the sample’s sign languages.

⁵ The authors do not detail in which contexts or subcategories of names, the number of repetitions would or wouldn’t be relevant.

⁶ Miljan (2003) reports mouthing as a strategy used for number marking, but does not present data.

Table 4 – Strategies for expressing plurals

Strategies	Sign Languages										
	IUR Canada	LSM Mexico	Libras Brazil	ISL Ireland	NGT Netherlands	GDS Germany	ESL Estonia	ISL Israel	IPSL India/Pakistan Nepal	Aulan Australia	
Noun + Numeral	+	+	+	+	+	+	+	+	+	+	+
Noun + Quantifier	+	+	+	+	+	+	+	+	+	+	+
Noun + Descriptive Verbs	+	+	+	+	+	+	+	+	+	+	+
Reduplication with Displacement	+	+	+	+	+	+	+	+	+	+	+
Reduplication without Displacement	+	-	-	-	+	+	+	+	+	+	-
Mouthing	-	-	-	-	-	-	+	-	-	-	-
Zero	+	+	+	+	+	+	-	-	+	+	+
Lexical	-	-	-	-	-	-	-	-	-	-	-

Source: prepared by the authors.

All syntactic strategies were found in all sign languages in the sample, but not all morphological strategies. According to Haspelmath (2013), languages that use only by syntactic strategies for number marking, for example by using numerals and quantifiers, or even by inference, are considered languages with no nominal plural. This is not the case with sign languages, as all sign languages use at least one morphological strategy.

Reduplication with displacement is a strategy found in all languages. But reduplication, both with displacement and without displacement, has articulatory and syntactic constraints and seems not to be widespread as a plurality strategy, which reflects in a restricted distribution.

In general, articulatory characteristics such as body-anchored signs and repetition of signs, as well as the presence of numerals and quantifier in the noun phrase, block reduplication (STEINBACH, 2012; ZESHAN, 2013). Syntactic blocking seems to obey some principle of economy in morphological marking for plural (NEIDLE; NASH, 2012; PFAU; STEINBACH, 2005; 2012). But ESL (Estonia) has no syntactic constraints that block reduplication. In this language, reduplication can happen in the presence of numerals and quantifiers (MILJAN, 2003).

In IUR (Canada), reduplication is blocked in balanced bimanual non-body anchored signs, in addition to being blocked by the articulatory characteristics described above. However, at least one body anchored sign can be reduplicated (with displacement). This is the CHAR signal, a type of fish. Example (16) shows this sign.

(16) IUR (Canada) – body anchored sign with reduplication with displacement.



CHAR >+>+

‘Chars’ (kind of fish)

Source: Schuit (2013, p. 39) – adapted

In languages with no articulatory constraints, the data discussed by the authors suggest at least one prototypical characteristic. Therefore, even the reduplication with displacement that is present in all sign languages occurs only in part of the lexicon.

Zeshan (2003) presents an interesting overview of reduplication in IPSL (India, Pakistan, Nepal). Reduplication is a strategy belonging to the category of number and occurs in part of the lexicon, due to articulatory constraints. Semantic effects arising from reduplication will depend on the syntactic function of the reduplicated sign. According to the author, signs that make up the open classes are multifunctional, sometimes functioning as nouns, sometimes as verbs, even though they show a preference for one domain or another. Reduplication without displacement (iterative form) in verbs implies repeated action, and in nouns it implies several referents. Reduplication with displacement (distributive form) in verbs implies action repeated and distributed by places and, in nouns, implies several referents in various places⁷.

⁷ According to Corbett (2000), the notion of iterativity is usually treated in the aspect category, which refers to how an event is executed, while the tense category refers to when an event is executed. But, the notion of iterativity can be considered by the number category as an instance of pluractionality.

In sign languages, sweeping motion yields a collective plural, while reduplication with displacement yields a distributive plural. The use of space also activates a property related to the spatial relations of the pluralized referent. Quantifiers can also be reduplicated with displacement in the sign space.

Reported strategies suggest that sign languages have an analytical preference rather than synthetic and, therefore, appear to isolate the category of number. Considering the constraints discussed by the authors, we also suggest that sign languages are guided by articulatory characteristics.

The dual value is mentioned for all sign languages and is expressed through the doubling of hands, either through the doubling of the sign itself as a morphological strategy, or through the juxtaposition of duplicated descriptive verbs as a syntactic strategy. In the ESL (Estonia), for example, dual refers to two members of the class identified by the noun and the distinction for dual can be shown by two hands, “which have a classifier handshape and stand for the number ‘two’” (MILJAN, 2003, p. 206).

The presence of two manual articulators favors dual number marking. The juxtaposition of descriptive verbs, in the doubled form, marks the dual and has no constraints, whereas the doubling of lexical signs is restricted to monomanual signs, that is, it only occurs in a part of the lexicon.

Doubling in noun signs can indicate both the dual and the plural (BÖRSTELL, 2011; MILJAN, 2003; XAVIER; BARBOSA, 2015; ZESHAN, 2003). In the case of verb signs, doubling is related to actions performed by dual/plural participants, as well as to reciprocity, which also indicates dual participants (JONSHTON; SCHEMBRI, 2007; PFAU; STEINBACH, 2005a; QUADROS; KARNOPP, 2004; SANCHEZ-MENDES; XAVIER, 2017; SANCHEZ-MENDES; SEGALA; XAVIER, 2017; WILBUR, 2005; KLIMA; BELLUGI, 1979; ZESHAN, 2003).

According to Miljan (2003), if two entities are mentioned in ESL (Estonia), the dual is normally used to mark them and the dual is almost never replaced by the plural. The trial and quadral values, on the contrary, are facultative and are often replaced by the plural.

The dual can also be manifested by using a strategy called the iconic plural, which is also employed for the trial and the quadral. Although some authors call this process reduplication, this strategy deserves a brief examination.

This strategy constitutes a repetition of the singular form in the sign space, or even a localized reduplication, different from the reduplication mentioned so far, which is articulated in a non-localized way, although it also yields a spatial arrangement of the pluralized referent. The repetition of a singular form in a localized way is an instance of this iconic plural in sign languages (SCHLENKER; LAMBERTON, 2019, ZESHAN, 2003; ZWITSERLOOK; NIJHOF, 1999).

The iconic plural refers to the manifestation of non-singular values through the spatial arrangement of forms individually repeated. The number of repetitions and the distinctive pause between them can indicate a value for the category of number. In the ESL (Estonia), according to Miljan (2003), in addition to the spatial arrangement and the distinctive pause between the repetitions, there is also the eye gaze of the signer directed towards this construction.

This strategy can be used to describe the spatial arrangement of a large number of referents, which would generate a large number of values for the category. The iconic plural is therefore perhaps best understood as a direct counting. In principle, there is no limit to how many referents can be iconically represented, but according to Zeshan (2003), in IPSL (India, Pakistan and Nepal) most occurrences refer to numbers up to five.

Numeral incorporation has a very restricted distribution in the lexicon of these languages. This strategy allows for a direct quantification of the referent involved. Auslan (Australia), for example, allows the incorporation from 2 up to 12, which would be a number of values far beyond what is typologically observed for the category of number. Therefore, this strategy is also more like a direct counting than an indirect one. In IPSL (India, Pakistan and Nepal), there is considerable individual and dialectal variation across different signs as to the extent of numeral incorporation. Figure (17) illustrates the incorporation of the numeral seven in YEAR in ISL (Israel) to illustrate the large number of possible values.

(17) ISL (Israel) – numeral incorporation



YEAR

'one year'



SEVEN – YEAR

'seven years'

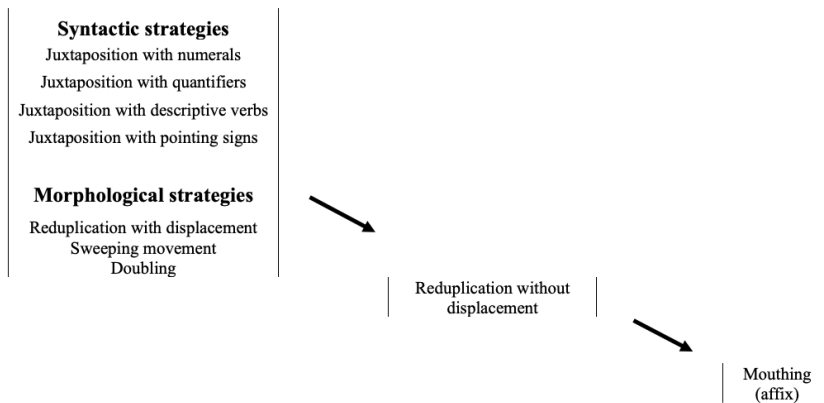
Source: Meir and Sandler (2008, p. 104) – adapted

As mentioned above, sign languages show a preference for syntactic strategies over morphological strategies for the manifestation of the category of number. Syntactic strategies are present in all sign languages, have a wide distribution and have no constraints. Morphological strategies, despite being present in all languages, are subject to distribution constraints, both in the sample languages and within the lexical repertoire of individual languages.

Reduplication with displacement, although used in all the sample's sign languages, presents articulatory constraints and a limited distribution in the lexicon of sign languages. Reduplication without displacement is not present in all languages and is also subject to articulatory constraints, as well as having a limited distribution. Furthermore, reduplication tends to be blocked by the presence of numerals and quantifiers in the noun phrase. Mouthing is considered an affixal strategy and is present in only one sign language, showing a wide distribution in ESL (Estonia).

Image 4 illustrates an implicational hierarchy of number manifestation strategies in sign languages. The diagram does not include zero marking.

Image 4 – Implicational hierarchy of number manifestation strategies in sign languages.



Source: prepared by the authors.

The implicational hierarchy indicates that syntactic strategies are considered primary and, regarding morphological strategies, mouthing (affix) is seldom used and presupposes the use of reduplication without displacement, which, in turn, presupposes reduplication with displacement.

A comparison between modalities shows that sign languages and spoken languages are typologically distinct. Sign languages favor syntactic strategies, while spoken languages favor morphological strategies.

With regard to morphological strategies, sign languages favor reduplication while spoken languages favor affixation. This may be an intrinsic effect of each modality, since in sign languages a non-concatenative morphology prevails. Another modality-related effect is the use of spatial strategies. Sign languages can yield the referent’s spatial arrangement in strategies based on the use of sign space.

Spatial strategies prevail over non-spatial strategies, which also suggests an implicational hierarchy. In sign languages, the use of non-spatial strategies presupposes the use of spatial strategies for the manifestation of number.

Sample data does not provide information on the manifestation of number in mass nouns. In brief, mass nouns in LSM (Mexico) and ESL (Estonia) are coded based on their physical dimension, using descriptive verbs that express the signers’ conception of the real world referent’s shape and size (CRUZ-ALDRETE, 2008; MILJAN, 2003). In ISL (Israel), quantifiers tend to denote intensity when juxtaposed to mass nouns (STAVANS, 1996).

All examples of strategies, as reported in the descriptions we had access to, were based on countable referents. Therefore, in sign languages, prototypicality seems to be present at least in countable referents for the expression of number. Although the manifestation of number in mass nouns was not reported in our survey, this category of nouns can become countable and behave in such a way as to allow the use of the same strategies (WACHOMICZ, 1997; WINTER; SCHA, 2015).

From a typological perspective, considering only nouns, animacy is an important feature in determining the manifestation of number. In the case of our sample's sign languages, there is no mention of any semantic preference/constraint for the manifestation of number in nouns. Based on the examples presented by the authors, we suggest that sign languages are not subject to semantic constraints. According to the data collected, the manifestation of number occurs in discrete referents, both animate and inanimate.

Values and strategies in pronouns

According to Corbett (2000), personal pronouns are given preference over nouns in the manifestation of number. The first person pronoun would be given an even stronger preference.

The personal pronoun system of Lihir, a language spoken on Lihir Island in Papua New Guinea's New Ireland Province, presents the maximum number of values expected for the category of number. Table 5 illustrates Lihir's pronominal system

(18) Lihir Language

Table 5 - Independent Pronouns in Lihir

	singular	dual	trial	paucal	plural
1 exclusive	<i>yo</i>	<i>gel</i>	<i>getol</i>	<i>gehel</i>	<i>ge</i>
1 inclusive	-	<i>kito</i>	<i>kitol</i>	<i>kitahel</i>	<i>giet</i>
2	<i>wa</i>	<i>gol</i>	<i>gotol</i>	<i>gohet</i>	<i>go</i>
3	<i>e</i>	<i>dul</i>	<i>dietol</i>	<i>diehet</i>	<i>die</i>

Source: Ross (1988, p. 258 in COBERTT, 2010, p. 25).

The strategies for the manifestation of number in personal pronouns, in general, differ from those for the manifestation in nouns.

The first, second and third person pronouns can also have different manifestations (COBERTT, 2000; DANIEL, 2013; HASPELMATH, 2013). According to Haspelmath (2013), the manifestation of number in pronouns in (spoken) languages usually involves stem changes, differing from the manifestation in nouns, which generally involves affixation. Stem changes in nouns for the manifestation of number are rare.

The manifestation of number in sign language pronominal systems involves numeral incorporation. Changing the handshape to incorporate numerals in pronouns allows for the expression of singular, dual, trial, quadral and plural values. According to Zeshan (2003), this strategy seems to be an instance of the iconic plural. The difference would reside in its paradigmatic organization.

If we consider numeral incorporation as an instance of stem change (MASSONE; JOHNSON, 1991; MASSONE; MACHADO, 1994), the pattern of manifestation of number in pronouns in sign languages is similar to that of spoken languages (HASPELMATH, 2013). Sign languages and spoken languages can thus be considered typologically similar. Stem change can be considered an inter-modal pattern of manifestation of number in pronouns.

In ISL (Israel), according to Meir and Sandler (2008), the pronominal system shows a strong tendency to incorporate numerals, limited only by the number of fingers, so that it allows denoting from 2 to 10 referents. This would illustrate a device outside the category of number.

The dual has an articulatory characteristic distinct from other incorporated forms and the plural generally involves a sweeping movement (CORMIER, 2012; CRUZ-ALDRETE, 2008; FELIPE, 2007; STEINBACH, 2012; ZESHAN, 2003). LSM (Mexico) has a plural form, based on a handshape with five extended fingers and palm facing down in a circular movement (CRUZ-ALDRETE, 2008).

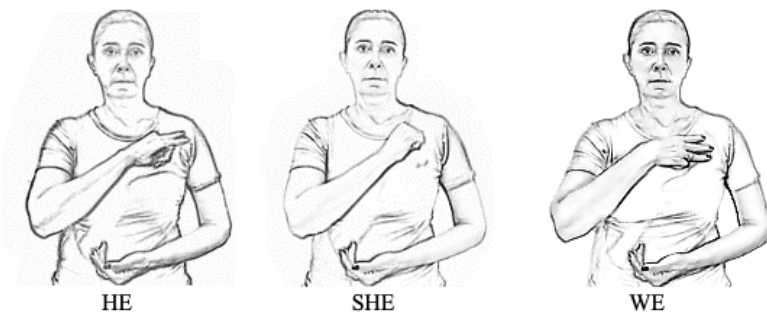
In IPSL (India, Pakistan and Nepal), both the first person pronoun and non-first person pronouns seem to be not marked for number, that is, IPSL has the general number, since the singular pointing can represent a group of referents (ZESHAN, 2003).

LSM (Mexico) and ISL (Ireland) have initialized pronominal forms, which are infrequently used. In the case of LSM (Mexico), there is a way to indicate first person with a Y-handshape touching the signer's chest, and a way to indicate second person (*usted*), with a U-handshape. The first person plural form can be performed with an "N" handshape,

in which it performs a sweeping arc movement juxtaposed with the MUCHO sign (CRUZ-ALDRETE, 2008).

In ISL (Ireland), according to Leeson and Saeed (2012), there are non-pointing initialized forms not for HE, SHE, THEY and WE, which are also infrequently used. Figure (19) illustrates these signs. The authors did not provide data for the THEY sign.

(19) ISL (Irlanda) – lexicalized forms for HE, SHE and WE.



Source: Leeson and Saeed (2012, p. 156-157) – adapted

5 Final considerations

Linguistic typology is an approach to the empirical study of human language interested in describing the structure and functioning of languages, taking into account the similarities and differences between them. One of its goals is to identify the types of systems that make up languages by comparing the various systems, surveying what varies between them and explaining the observed phenomena. For linguistic typology, sign languages present a much greater diversity than what has been established up to now, and only a systematic investigation can identify all the existing variation.

According to Palfreyman, Sagara and Zeshan (2015) and Zeshan and Palfreyman (2017), typological studies that include hundreds of sign languages are not yet possible, as is the case with spoken languages. The number of sign languages in the world is small and that of documented

sign languages is even smaller⁸. The large-scale typological studies of sign languages carried out so far have involved from 30 to 40 languages. Most studies on the topic are small scale, as in the case of this research.

In this article, we described the manifestation of the category of number in sign languages, with the aim of identifying possible organizational similarities and differences between them. We identified the values, forms and strategies available for use in the noun phrase and some intra-modal and inter-modal patterns were found.

- Some sign languages have optional number marking and others have obligatory number marking.
- In sign languages where number marking is optional, the general form coincides with the singular form, expressed by zero.
- Values of the category of number, in sign languages, are singular, plural and dual.
- Trial and quadral values can be expressed using the iconic plural, which consists of a process of repetition of the singular form in the sign space, with a distinctive and localized pause.
- Values of the category of number in nouns are expressed by syntactic and morphological strategies, with preference given to the former.
- Sign languages seem to be isolating languages in relation to the category of number, thus differing typologically from spoken languages.
- Morphological strategies in sign languages suggest an implicational hierarchy, with mouthing being a rarely used strategy and reduplication with displacement a more prevalent strategy.
- The manifestation of the category of number in sign languages seems to be phonologically driven.

⁸ According to the ethnologue (<https://www.ethnologue.com/subgroups/sign-language>, access on: Apr. 19, 2022), the number of known sign languages would be 157. According to Glottolog (<https://glottolog.org/resource/languoid/id/sign1238>, access on: Apr. 19, 2022) this amount would be 210.

- Personal pronouns in sign languages have singular, dual, trial, quadral and plural values.
- Numeral incorporation (stem change) seems to be the more prevalent strategy for the manifestation of the category of number in pronouns, with sign languages and spoken languages being typologically similar in this respect.

These generalizations were based on the data to which we had access. More descriptions regarding this topic could certainly be found, as well as forms not yet reported. This is a permanent topic of the linguistic typology agenda, as this is a discipline driven by empirical investigation. In this sense, the continuing search for new data and the review of previously identified patterns are within the linguistic typology's scope. Thus, new findings lead to new generalizations.

In collecting the data, we were careful to reconsider some information obtained from secondary sources, in order to standardize the terminology on the basis of our study parameters. The term reduplication without displacement, for example, was named by some authors as recycling, simple reduplication, or no distinction was made between reduplication with and without displacement. Doubling was also referred to as repetition, chameleon hand, or even reduplication. For this reason, we paid close attention to the reported examples, which were useful to us to establish some parameters.

Terminological adaptation was also necessary for enabling comparisons between modalities, as in the case of mouthing, which was considered an affixal strategy, and articulatory constraints in sign languages, considered phonological constraints in spoken languages.

This study presents the forms and strategies reported in the literature examined regarding the category of number, but does not provide information about the contexts of use, due to the nature of the investigation. New descriptions prepared using language corpora can provide new data, mainly on structures not reported in the sample's languages.

Another research challenge was inaccuracy in the information collected. Once again, we had to rely on reported data to decide whether a certain manifestation strategy was present or not, based on the established parameters. The sources we had access to do not necessarily describe the language from a typological perspective. In some cases, strategies were mentioned without supporting data being presented.

Some generalizations presented here can serve as a basis for describing how the category of number is manifested in other individual sign languages. This study needs to be complemented by further research with data from other sign languages, thus increasing the number of languages in the sample and collecting more data on number expression in nouns, pronouns and entity classifiers.

Finally, sign languages, as they are natural languages, are fundamental for the survey and (re)consideration of structural and functional aspects of languages in general. Sign Language Typology can open new paths for our understanding of human language.

Authorship statement

Bruno Gonçalves Carneiro was responsible for drafting and writing the article proposal, data analysis and textual review. Mônica Veloso Borges was responsible for the textual outline on the number in spoken languages section, data analysis and textual review. Mirosvala Cruz Aldrete was responsible for the textual outline on the number in sign languages section, data analysis and textual review.

Acknowledgements

The authors are grateful to Evandro Silva Carvalho and Roselba Gomes de Miranda by adapting the illustrations about sign language data.

References

- BÖRSTELL, C.; LEPIC, R.; BELSITZMAN, G. Articulatory plurality is a property of lexical plurals in sign language, *Lingvistica e Investigationes*, Amsterdã, v. 39, n. 2, p. 391-407, 2016, DOI: <https://doi.org/10.1075/li.39.2.10bor>
- BÖRSTELL, C. Revisiting Reduplication. *Toward a description of reduplication in predicative signs in Swedish Sign Language*. 2011. 86f. Master's Thesis. (General Linguistics). Stockholms Universitet, 2011.
- CORBETT, G. G. *Number*. New York: Cambridge University Press, 2000.
- CORMIER, K. Pronouns. In: PFAU, R.; STEINBACH, M.; WOLL, B. (eds.). *Sign Language. An International Handbook*. Berlin: Mouton de Gruyter, 2012. p. 227-244.

CRUZ-ALDRETE, M. *Gramática de la Lengua de Señas Mexicana*. 2008. 1.122f. Tesis (Doutorado en Linguística) – Centro de Estudios Lingüísticos y Literarios, El Colegio de México, 2008.

CUNHA, C.; CINTRA, L. *Nova gramática do português contemporâneo*. 4. ed. Rio de Janeiro: Lexikon, 2007.

DANIEL, M. Plurality in Independent Personal Pronous. In: DRYER, M.; HASPELMATH, M. (eds.). *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013. n.p. Available at: <http://wals.info/chapter/131>. Access on: Oct. 31, 2019.

DRYER, M. Coding of Nominal Plurality. In: DRYER, M.; HASPELMATH, M. (eds.). *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013. n.p. Available at: <http://wals.info/chapter/33>. Access on: Oct. 29, 2019.

FELIPE, T. A. *Libras em Contexto*. 8. ed. Brasília: WalPrint Gráfica e Editora, 2007.

FENLON, J; WILKINSON, E. Sign languages in the world. In: SCHEMBRI, A.; LUCAS, C. (eds.). *Sociolinguistics and deaf communities*. Cambridge: Cambridge University Press, 2015. p. 5-28.

FERREIRA, L. *Por uma gramática de línguas de sinais*. Rio de Janeiro: Editora Tempo Brasileiro, 2010.

FERREIRA, L. Repetição e reduplicação em Língua Brasileira de Sinais. *Papia*, São Paulo, v. 11, n. 1, p. 6-17, 2001. Recuperado de <http://revistas.flch.usp.br/papia/article/view/1849>.

FINAU, R. Uma análise do Sistema Quantificacional da Libras. In: STUMPF, M.; QUADROS, R. M.; LEITE, T. A. (orgs.). *Estudos da Língua Brasileira de Sinais II*. Florianópolis: Insular, 2014. p. 119-144.

GREENBERG, J. H. Some universals of grammar with particular reference to the order of meaningful elements. In: GREENBERG, J. H. (ed.). *Universals of Language*. Standford: Standford University Press, 1963. p. 58-90.

HASPELMATH, M. Occurrence of Nominal Plurality. In: DRYER, M.; HASPELMATH, M. (eds.). *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013. n.p. Available at: <http://wals.info/chapter/34>. Access on: Oct. 31, 2019.

- HERBERT, M. *Pluralization in German Sign Language*, 2015. Available at: <https://www.swarthmore.edu/sites/default/files/assets/documents/linguistics/2012_Herbert.pdf>. Access on: Jan. 10, 2019.
- KLIMA, E.; BELLUGI, U. *The signs of language*. Cambridge: Harvard University Press, 1979.
- LARA, M. C. P. *A Pluralidade em Libras*. 2017. 82 fls. Dissertação (Mestrado em Letras) – Universidade Federal do Paraná, 2017.
- LEESON, L.; SAEED, J. I. *Irish Sign Language*. A cognitive Linguistic Account. Edinburgh: EDINBURGH University Press, 2012.
- MASSONE, M. I.; JOHNSON, R. E. *Numbers and numeral classifier suffixes in Argentine Sign Language*. XI World Congress of the World Federation of the Deaf Tokyo Japón, 1991, p. 742-762.
- MASSONE M. I.; MACHADO, E. *Lengua de Señas Argentina*. Análisis y Vocabulario bilingüe. Buenos Aires: Edical, 1994.
- MEIR, I.; SANDLER, W. *A language in Space*. The Story of Israeli Sign Language. New York: Lawrence Erlbaum Associates, 2008.
- MILJAN, M. Number in Estonian Sign Language. *Trames*, Tallinn, n. 7, v. 3, p. 203-223, 2003. Disponível em: https://viipekeel.ee/wp-content/uploads/2021/01/number_ESL_Miljan_Trames.pdf.
- NEIDLE, C.; NASH, J. The noun phrase. In: PFAU, R.; STEINBACH, M.; WOLL, B. (eds.). *Sign Language*. An International Handbook. Berlin: Mouton de Gruyter, 2012. p. 265-292.
- PALFREYMAN, N.; SAGARA, K.; ZESHAN, U. Methods in Carrying out Language Typological Research. In: ORFANIDOU, E.; WOLL, B.; MORGAN, G. (eds.). *Research Methods in Sign Language Studies*. A practical Guide. Chichester: Wiley Blackwell, 2015. p. 173-192.
- PFAU, R.; STEINBACH, M. Pluralization in sign and in speech: a cross-modal typological study, *Linguistic Typology*, Berlin, v. 10, p. 135-182, 2006, DOI: <https://doi.org/10.1515/LINGTY.2006.006>
- PFAU, R.; STEINBACH, M. Backward and sidward reduplication in German Sign Language. In: HURCH, B. (ed.). *Studies on reduplication*. New York: Mouton de Gruyter, 2005a. p. 569-594.
- PFAU, R.; STEINBACH, M. Plural Formation in German Sign Language: Constraints and Strategies. In: LEUNINGER, H.; HAPP, D. (eds.). *Gebärdensprachen: Struktur, Erwerb, Verwendung*. Linguistische Berichte Sonderheft 13. Hamburg: Helmut Buske Verlag, 2005b. p. 111-144.

POWER, JM; GRIMM, GW; LIST, JM. Evolutionary dynamics in the dispersal of sign languages. *R. Soc. Open Sci*, London, v. 7, p. 1-15, 2019. DOI: <https://doi.org/10.1098/rsos.191100>

QUADROS, RM; KARNOPP, LB. *Lingua de Sinais Brasileira: Estudos Linguísticos*. Porto Alegre: Editora Artmed, 2004.

QUER, J.; CECCHETTO, C.; DONATI, C.; GERACI, C.; KELEPIR, M.; PFAU, R.; STEINBACH, M. (eds.). *Sign Gram Blueprint. A Guide to Sign Language Grammar Writing*. Berlin/ Boston: Walter de Gruyter GmbH, 2017.

SANCHEZ-MENDES, L.; XAVIER, A. N. A expressão de pluracionalidade em libras. *Estudos Linguísticos*, São Paulo, v. 45, n. 1, p. 292-304, 2016. DOI: <https://doi.org/10.21165/el.v45i1.633>

SANCHEZ-MENDES, L.; SEGALA, R. R.; XAVIER, A. N. O papel da (re)duplicação na expressão de pluracionalidade em libras. *Revista Letras*, Curitiba, n.96, p. 487-508, 2017. DOI: <http://dx.doi.org/10.5380/rel.v96i0.51047>

SCHUIT, J. M. *Signs of the Arctic*. Typological aspects of Inuit Sign Language. Amsterdam: Universiteit van Amsterdam, 2013.

SCHLENKER, P.; LAMBERTON, J. Iconic plurality. *Linguistic and Philosophy*, Berlin, n. 42, p. 45-108, 2019. DOI: <https://doi.org/10.1007/s10988-018-9236-0>

STARK, T. C. S.; CRUZ-ALDRETE, M. *La morfología en la lengua de señas mexicana*. II Congreso Internacional de Logogenia México 2006, del 20 al 22 de septiembre de 2006. Museo Nacional de Antropología, Bosque de Chapultepec, México, D.F.

STAVANS, A. One, Two, or More: The Expression of Number in Israeli Sign Language. In: EDMONDSON, W. H.; WILBUR, R. B. (eds.). *Internacional Review of Sign Linguistics*. Volume 1. New Jersey: Lawrence Erlbaum Associates, 1996. p. 95-114.

STEINBACH, M. Plurality. In: PFAU, R.; STEINBACH, M.; WOLL, B. (eds.). *Sign Language. An International Handbook*. Berlin: Mouton de Gruyter, 2012. p. 112-136.

WACHOWICZ, T. C. *Uma Semântica de Reticulados para os Plurais e os Termos de Massa*. 2017. 120 f. Dissertação (Mestrado em Linguística), UFPR, 1997.

WINTER, Y.; SCHA, R. Plurals. In: LAPPIN, S.; FOX, C. (eds.). *The Handbook of Contemporary Semantic Theory*. 2 ed. Chennai: SPi Global, 2015. p. 77-113.

WILBUR, R. A reanalysis of reduplication in American Sign Language. In: HURCH, B. (ed.). *Studies on reduplication*. Berlin/ New York: Mouton de Gruyter, 2005. p. 595-623.

XAVIER, A. N.; BARBOSA, P. A. A duplicação do número de mãos de sinais da libras e seus efeitos semânticos. *Fórum linguístico*, Florianópolis, v.12, n. 1, p. 505-514, 2015. DOI: <https://doi.org/10.5007/1984-8412.2015v12n1p505>

ZESHAN, U. Indo-Pakistani sign language grammar: a typological outline. *Sign Language Studies*, Washington, v. 3, n. 2, p. 157-212, 2003. DOI: [doi:10.1353/sls.2003.0005](https://doi.org/10.1353/sls.2003.0005)

ZESHAN, U.; PALFREYMAN, N. Typology of sign languages. In: AIKHENVALD, A. Y.; DIXON, R. M. (eds.). *The Cambridge Handbook of Linguistic Typology*. Cambridge: Cambridge University Press, 2017. p. 178-216.



Um estudo diacrônico sobre a polaridade negativa no sintagma nominal: o caso do indefinido “algum” na formação do IPN [N + algum]

A Diachronic Study of Negative Polarity in Nominal Phrases: the Case of the Indefinite “Some” in the Formation of IPN [N +algum (Some)]

Cristiane Namiuti

Universidade Estadual do Sudoeste da Bahia (UESB), Vitória da Conquista, Bahia / Brasil
cristianenamiuti@gmail.com / cristianenamiuti@uesb.edu.br
<https://orcid.org/0000-0002-1451-8391>

Fernanda Gusmão Silva

Universidade Estadual do Sudoeste da Bahia (UESB), Vitória da Conquista, Bahia / Brasil
fgsilva031@gmail.com
<https://orcid.org/0000-0002-0940-3629>

Resumo: O presente artigo trata da posposição do indefinido “algum” em relação ao nome substantivo para instanciar a polaridade negativa no sintagma nominal na diacronia da língua portuguesa, buscando compreender: (i) a natureza da estrutura [N+algum] em textos portugueses do século XVI, XVII, XVIII e XIX, e em textos brasileiros do século XIX; e (ii) o processo de gramaticalização da estrutura como um Item de polaridade Negativa (IPN) na história da língua. Verificou-se que a estrutura [N+algum], com inversão, nos séculos XVI e XVII, Português Clássico (PCL), podia ser usada com valor não negativo e que outros elementos podiam ocorrer no interior do sintagma nominal, inclusive com a possibilidade da flexão e não adjacência entre o nome e o indefinido. Tais fatos corroboraram a hipótese de que no PCL a estrutura [N+algum] não estava gramaticalizada como um IPN e o valor negativo era valorado pela presença de um operador de negação sentencial, preposição ou conjunção negativa em domínio sintático superior. Conclui-se que o indefinido no PCL realiza o núcleo da categoria determinante (D) mesmo nas estruturas com inversão, sendo estas derivadas

do movimento da parte lexical do sintagma à posição de especificador do sintagma determinante (DP, do inglês *Determiner Phrase*) fato que contempla as propriedades de categoria D de “algum” no PCL.

Palavras-chave: sintagma nominal; indefinido “algum”; polaridade negativa.

Abstract: The present paper deals with the posposition of the indefinite “algum” (some) in relation to the noun to instantiate the negative polarity in the nominal phrase in the Portuguese language diachrony, seeking to understand: (i) the nature of the structure [N+algum] in sixteenth-, seventeenth-, eighteenth- and nineteenth-century Portuguese texts, and in nineteenth-century Brazilian texts; and (ii) the process of grammaticalization of the structure as a Negative Polarity Item (IPN, in Portuguese *Item de Polaridade Negativa*) in the history of the language. It has been found that the structure [N+algum], with inversion, in the 16th and 17th centuries, Classical Portuguese (PCL, in Portuguese *Português Clássico*), could be used with non-negative value and that other elements could occur in a noun phrase (NP), including the possibility of inflection and non-adjacency between the noun and the indefinite. Such facts supported the hypothesis that in PCL the structure [N+algum] was not grammaticalized as an IPN and the negative value was valued by the presence of a sentential negation operator, preposition or negative conjunction in a higher syntactic domain. It has been concluded that the indefinite in PCL is a determiner (D) even in structures with inversion, these being derived from the movement of the lexical part of the phrase to the specifier position of Determiner Phrase (DP) a fact that contemplates the D properties of “algum” in PCL.

Keywords: noun phrase; indefinite “algum”(“some”); negative polarity.

Recebido em 25 de fevereiro de 2022

Aceito em 11 de abril de 2022

1 Introdução

A instanciação da polaridade negativa dos indefinidos na história da língua portuguesa foi objeto de investigação de Martins (1997, 2000, 2015), Pinto (2015) entre outros trabalhos. Nesta empreita, a posposição do indefinido “algum” em relação ao nome substantivo para instanciar a polaridade negativa no sintagma nominal na diacronia da língua portuguesa tem sido contemplada e neste trabalho buscamos discutir especificamente este caso, partindo das análises dos trabalhos mencionados anteriormente.

Para tanto, realizamos um estudo diacrônico comparado, no âmbito do Projeto Temático “*Do português pré-clássico às variantes modernas: contribuições para o estudo da Sintaxe e suas interfaces*” (FAPESB, APP 0007/2016), coordenado por Cristiane Namiuti (PPGLin/UESB)¹ buscando compreender: (i) a natureza da estrutura [N + algum] nos textos portugueses do século XVI, XVII, XVIII e XIX, e em textos brasileiros do século XIX; e (ii) o processo de gramaticalização² da estrutura como um Item de Polaridade Negativa (IPN) na história da língua.

O corpus para este estudo contempla textos do Português Clássico (PCL), do Português Europeu dos séculos XVIII e XIX (PE) e do Português Brasileiro do século XIX (PB) pertencentes ao *Corpus* Anotado do Português Histórico Tycho Brahe (CTB) (GALVES; ANDRADE; FARIA, s/d) e ao *Corpus* de Documentos Oitocentistas de Vitória da Conquista e Região (DOViC) (SANTOS; NAMIUTI, 2014). Em consonância com os trabalhos aqui revisitados, que abordam o fenômeno em perspectiva diacrônica, a exemplo de Martins (2015), o quadro teórico utilizado como guia para a descrição e análise dos dados foi o da Gramática Gerativa (CHOMSKY, 1995) por fornecer um instrumental descritivo que nos serve para explicar formalmente o fenômeno e ainda garantir a possibilidade de comparação e avanço em relação aos trabalhos e hipóteses de nosso ponto de partida.

¹ O recorte utilizado para esta publicação foi objeto da dissertação de mestrado de Fernanda Gusmão Silva (PPGLin/UESB), defendida em 2021, sob a orientação de Cristiane Namiuti, no âmbito do projeto temático.

² O processo diacrônico de mudança linguística que transforma uma forma lexical em uma forma gramatical foi denominado por Meillet (1948) de gramaticalização, nesse processo os conteúdos lexicais vão sendo esvaziados de sentido lexical e a forma/estrutura passa a desempenhar funções gramaticais. Segundo Neves (2001), o processo de gramaticalização ocorre quando uma determinada categoria migra para uma outra condição na língua, é o que acontece na transformação diacrônica de um verbo pleno a um verbo auxiliar, de uma palavra lexical ou pronome a um afixo ou clítico. Nesse sentido, a transformação diacrônica da estrutura sintagmática [N + algum], em que [N] e [algum] são núcleos sintáticos independentes, para um IPN, em que [N] e [algum] realizam um único núcleo sintático de polaridade negativa, funcionando como uma única palavra negativa (*n-word*), ou seja expressões como “coisa alguma” deixa de ser uma estrutura sintática com palavras independentes e valor negativo, passando a ter no nome o indefinido incorporado e funcionando como uma *n-word* (a exemplo de “nada”), fato evidenciado, entre outras coisas, pela perda da possibilidade de flexionar o substantivo nas variantes atuais do português, esse é o tipo de fenômeno linguístico que ficou conhecido por “gramaticalização”.

Este artigo está organizado em 5 seções, incluindo esta introdução a que se segue a apresentação do estado da arte, na seção 2, com a revisão do referencial que nos serve de ponto de partida e sustentação. Os resultados da pesquisa estão descritos na seção 3 do artigo e a ela segue-se a proposta de análise na seção 4. Por fim, na seção 5, conclui-se o artigo com as considerações finais para este trabalho.

2 A especificação da polaridade dos Indefinidos na diacronia da língua portuguesa

A negação está presente em todas as línguas e é registrada em cada uma delas por operadores diferentes e em estruturas sintáticas diversas. Sua variedade de usos e formas tem sido objeto de diversos estudos, os quais têm apresentado análises relevantes para o conhecimento desse fenômeno linguístico.

No Português Brasileiro, Português Europeu e em fases antigas da língua, Português Antigo e Clássico (doravante PB, PE, PA e PCL, respectivamente), a negação sentencial é realizada pelo operador de negação “não” (MARTINS, 1997; MIOTO, 1992; NAMIUTI 2008, entre outros trabalhos). A polaridade negativa, no entanto, pode figurar em outros níveis além do sentencial e outros marcadores também são utilizados: pronomes indefinidos/quantificadores (ninguém, nenhum), conjunção (nem), preposição (sem), dentre outros.

Em relação aos pronomes indefinidos como “nenhum” e “ninguém”, seu valor negativo, segundo Martins (2000, 2015), precisava ser valorado pelo operador de negação sentencial da sentença no PA, como podemos constatar no exemplo 02 nesta seção, não sendo, portanto, intrinsecamente negativo no léxico, como no PE atual.

De acordo com Martins (2000, 2015), no PE, os indefinidos “ninguém” e “nenhum” são intrinsecamente negativos e o indefinido “algum” é positivo no léxico. Todavia, a inversão de “algum”, segundo Martins (2015), em relação ao núcleo substantivo terá sempre um valor negativo no PE (1.a); manterá sempre adjacência absoluta entre o substantivo e o indefinido, ou seja, não terá a possibilidade de complementos ou adjuntos nominais (1.b), não terá a possibilidade de se flexionar (1.c) e ainda permitirá a graduação (1.d).

(1)

- a. **Lugar algum** se parece com este.
- b. ***Animal do deserto algum** vive aqui.
- c. ***Animais alguns** vivem aqui.
- d. Ainda não o vi fazer **coisíssima nenhuma/alguma**

(MARTINS, 2015, p. 1, 4)

Na ordem não marcada “algum lugar” a interpretação é positiva, nesta ordenação a flexão é permitida “alguns animais”, também é permitida a complementação/modificação “alguns animais do deserto” e vedada a graduação “*alguma coisíssima”, marcando uma diferença morfosintática importante entre a estrutura [algum + N] e [N + algum].

Há uma interpretação negativa sempre que a inversão do indefinido algum é atestada, como em (1.a) “Lugar algum”, que equivale na interpretação a “nenhum lugar” ou “lugar nenhum” indicando que a inversão do indefinido “algum” no sintagma nominal (NP) é a operação que realiza essa polaridade negativa. Segundo Martins (2015), a inversão que gera a polaridade negativa do sintagma no PE é derivada de um processo de incorporação de núcleos na formação de um item de polaridade negativa (IPN), como veremos mais adiante, a agramaticalidade de inversão sem contiguidade entre o substantivo e o indefinido, como ilustrada em (1.b), a agramaticalidade da flexão, como ilustrada em (1.c), e a possibilidade de graduação, como ilustrada em (1.d), corroboram, segundo Martins (2015), a hipótese do IPN derivado de um processo de incorporação do substantivo ao indefinido no PE. Não obstante, ao analisar os indefinidos negativos, ou *n-words*, no romance antigo, Martins (2000) descreve que eles coocorriam com o marcador negativo, não anulando a negação sentencial, e em alguns casos, eram usados em contextos não negativos, os quais poderiam ser não-assertivos (não afirmativos) (2.a) ou modais (2.b) (imperativo, condicional, dentre outros).

(02)

- a. que **nenh~uu nã** scapou (Crônica Geral de Espanha de 1344)
(cf. CINTRA, 1954, p. 107, apud MARTINS, 2000, p. 216)
- b. E por decreto publico foi defeso que **ninguém** navegasse.
(cf. ALI, 1931, p. 199, apud MARTINS 2000, p. 196)
[significando “e por um decreto público foi proibido que alguém navegasse”]

De acordo com Martins (2000), no século XV, o marcador de negação sentencial “não” passa a ser opcional no português (03 a-b) para valorar a polaridade negativa do indefinido “nenhum” em posição pré-verbal.

(03)

- a. “**Nenh~uu nom** mostrava que era famiinto”
(Fernão Lopes, Crônica de D. João I. apud Martins 2000:194)
- b. “**Nenh~uu** poderá seer emlegido a semelhante honra”
(Fernão Lopes, Crônica de D. João I. apud Martins 2000:194)

No português contemporâneo, estruturas com o indefinido negativo, em posição pré-verbal, coocorrendo com o marcador de negação sentencial, tornam a sentença agramatical (04).

(04) ***Ninguém não** vive aqui.

(MARTINS, 2015, p. 3)

Ou seja, no PE atual, a concordância negativa entre o indefinido pré-verbal e o marcador de negação sentencial, obrigatória no PA, opcional no período que abrange o PCL, tornou-se agramatical em um longo processo de mudança, passou de obrigatória no PA (séculos XIII, XIV, XV), passou a opcional no PCL (séculos XV, XVI, XVII) e se tornou agramatical apenas no PE moderno (séculos XIX, XX, XXI).

Com relação ao indefinido “algum” em posição pós-nominal em um NP, Martins (2015) ao descrever o comportamento da estrutura/ordem [N + algum] na diacronia do português, constata que nos séculos XVII e XVIII, o valor negativo de:

(i) [N+algum] era legitimado no escopo da negação, ocorrendo normalmente em posição pós-verbal; (ii) era possível a legitimação da inversão nominal negativa em contextos modais (também chamados “contextos negativos fracos”); (iii) a adjacência entre o nome e algum não era obrigatória e (iv) coisíssima alguma não ocorria. (MARTINS, 2015, p. 12)

Assim, Martins (2015), ao descrever as ocorrências de [N + algum] nos textos dos séculos XVII e XVIII e constatar que a inversão do indefinido, diferentemente do que acontece no PE atual, para ter interpretação negativa, precisava estar sob o escopo da negação sentencial; que o NP com o indefinido pós-nominal poderia não ter interpretação negativa em contextos modais; e ainda, a adjacência entre o substantivo e o indefinido invertido não era obrigatória, conclui que nesse período a estrutura [N + algum] não estava gramaticalizada como um IPN (Item de Polaridade Negativa), como no PE atual. Nos textos dos séculos XVII e XVIII a estrutura [N + algum] se comportava como um sintagma e não como uma estrutura de incorporação de núcleos em um núcleo funcional Neg no interior do NP (estrutura do IPN, segundo Martins, 2015), uma vez que a adjacência entre o nome (N) e o indefinido (algum) não era obrigatória e o valor negativo não dependia exclusivamente desta inversão. Além disso, segundo a autora,

No Corpus do Português não foi possível encontrar nenhum exemplo de [N+algum] na posição canônica de sujeito ou em qualquer outra posição fora do escopo da negação ao longo do século 17. Raros exemplos aparecem no século 18. É necessário esperar pelo século 19 para se encontrarem facilmente atestações da inovação (MARTINS, 2015, p. 12)

Segundo Martins (2015), a possibilidade de [N + algum] ocorrer fora do escopo da negação com valor negativo em posição pré-verbal, como em (05), e ainda a possibilidade de gradação pela morfologia, como em (06), constituem evidências da inovação que transformou [N + algum] em um IPN no PE.

(05) Em **época alguma** tinham os criados conhecido Maurício tão caseiro.

(Século XIX: *Corpus* do Português, MARTINS, 2015, p. 12)

(06) Nunca recebi favor do Sr. D. Pedro II, nem ele me deve **coisíssima alguma**.

(Século XIX: *Corpus* do Português, MARTINS, 2015, p. 12)

No português antigo, como vimos, os indefinidos apresentavam-se como itens de polaridade fracos (MARTINS, 1997, 2000; NAMIUTI, 2008, PINTO, 2015), seu valor negativo, precisava ser valorado pelo operador de negação sentencial no PA, não sendo, portanto, intrinsecamente negativo no léxico, como no PE atual, devendo por isso seu valor negativo ser legitimado por um marcador de negação regular para serem interpretados como negativos em uma configuração de concordância negativa (ZANUTTINI, 1997), podendo ocorrer em contextos negativos ou modais. Pinto (2015), baseando-se em Martins (1997, 2000), classifica os itens, observando que cada um dos traços pode ser catalogado como especificados [+] ou subespecificados, estes últimos podem ser não-variáveis [0] ou variáveis [α]. A leitura dos traços e seus valores permitem distinguir os itens de polaridade fortes dos que apresentam polaridade fraca, pois enquanto o item de polaridade forte apresenta um dos traços com valor especificado e, não possui traços com valores de subespecificação variável, o item de polaridade fraca exhibe, no mínimo, um valor subespecificado variável e nenhum traço com valor especificado [+].

O quadro 1, adaptado de Pinto (2015), apresenta os traços do indefinido algum no PA (cf. exemplos 02).

Quadro 1 – Análise de polaridade do Indefinido “algum”

TRAÇOS	ESPECIFICADO	SUBESPECIFICADO	RESULTADO
Afirmativo		0	Item de polaridade fraca
Negativo		α	
Modal		α	

Fonte: Adaptado de PINTO (2015)

Ao tratar dos quantificadores indefinidos “nenhum” e “algum”, Martins (2016) atesta que a coocorrência do marcador de negação sentencial “não” e indefinidos negativos torna-se opcional em fases mais recentes da língua (século XIX), pois estes passam a itens de polaridade negativa capazes de expressar sozinhos sua negação, ou seja, passam a ser

um item de polaridade forte, especificado para a negação. Porém, a autora descreve que, no português médio, houve uma separação em relação aos contextos nos quais passaram a ocorrer, resultando em uma distribuição complementar, em que “algum” é registrado em contextos afirmativos (assertivos) e modais (contextos negativos fracos) e “nenhum” está nos contextos estritamente negativos, diferentemente do PA em que ambos eram subespecificados para o traço de negação/afirmação.

Tal evolução também é descrita por Pinto (2015), segundo a autora, por meio da especificação de traços (cf. MARTINS, 1997, 2000), os indefinidos nada, nenhum e ninguém passaram a exibir o traço negativo [+ neg], não mais o traço subespecificado com valor variável [α], deixando assim de ocorrer em contextos modais. O mesmo aconteceu com a estrutura [N+algum], segundo Martins (2015), a partir do século XVIII, “algum” em posição pós-nominal apresenta o mesmo tipo de contraste interpretativo e gramatical que o item “ninguém” em contextos pós-verbais (cf. 07 a-b), e em contextos pré-verbais (cf. 07 a-d).

(07)

- a. Não vive aqui ninguém.
- b. Não vive aqui animal algum.
- c. Ninguém vive aqui.
- d. Animal algum vive aqui.

(MARTINS, 2015, p. 3)

O desenvolvimento dos itens de polaridade negativa no português parece acompanhar o de outras línguas românicas como o espanhol. A inversão nominal negativa do português e espanhol com algum/alguno, segundo Martins (2015), não tem as mesmas propriedades nas duas línguas, pois a sequência invertida [N+alguno], no espanhol, ainda precisa ocorrer sob o escopo da negação para ser legitimada (10 a), não pode ocorrer antes do verbo exigindo sempre a presença do marcador de negação predicativa (08 b-c) ou de um outro operador que crie o contexto negativo apropriado (08 d), contrastando com o português atual, que permite casos sem o escopo da negação (09 a-b)

(08)

- a. No he visto película alguna esta semana.
- b. *Ayuda alguna fue necesaria.
- c. No fue necesaria ayuda alguna.
- d. A los ricos los dejó sin cosa alguna.

(MARTINS, 2015, p. 9)

(09)

- a. Ajuda alguma foi necessária.
- b. Aos ricos, coisa alguma desejo.

(MARTINS, 2015, p. 9)

Os contrastes observados são descritos por Martins (2015) como resultado da distribuição sintática dos nomes nus, os quais não possuem determinante visível, pois, como o [N + alguno], eles são registrados em posição pós-verbal nas línguas românicas. Longobardi (1994 apud MARTINS, 2015) sugere que assim como os nomes nus a distribuição sintática mais limitada do espanhol [N + alguno], em contraste com o português [N + algum], pode ocorrer devido a necessidade da legitimação do núcleo D (Determinante) nulo na estrutura. E como o português não se submete às restrições de posicionamento oracional como o espanhol, pode ser esse o indicativo que no português há movimento sintático de Neg para D, ocasionando a incorporação da unidade [N +alguno] no núcleo D.

Martins (2015) destaca o bloqueamento da flexão de plural na inversão nominal negativa em ambas as línguas (10 a-b) como característica importante para entender a natureza da estrutura.

(10)

- a. *No hay soluciones algunas para ese dilema.
- b. *Não há soluções algumas para esse problema.

(MARTINS, 2015, p. 9)

Para a autora, o bloqueio para o plural destas construções e a necessidade de ocorrerem amalgamadas sugere que tais construções são

derivadas por processo de incorporação, sendo [N + algum] a realização de uma unidade complexa que se comporta como uma única palavra negativa.

Martins (2015) argumenta que no português europeu contemporâneo (PE), o indefinido “algum” apresenta polaridade positiva fraca e ao ser realizado em estrutura pós-nominal [N + algum], realiza, juntamente com o núcleo nominal, um item de polaridade negativa (IPN) com polaridade negativa forte, assim como [N + nenhum] (11 a-b).

(11)

- a. Animal algum vive aqui.
- b. Animal nenhum vive aqui.

Como demonstrou a autora, o IPN com o indefinido algum em português também não admite a pluralização do sintagma nominal na inversão nominal negativa (12 a), embora estruturas com o indefinido pré-nominal (com valor positivo) a admitam (12 b).

(12)

- a. *Animais alguns vivem aqui.
- b. Alguns animais vivem aqui.

(MARTINS, 2015, p. 4)

A hipótese defendida por Martins (2015) consiste no fato da flexão de plural ser bloqueada quando há formação do IPN [N + algum], pois segundo a autora, quando o NegP faz parte do DP, o Pl(ural)P não é projetado.

Em síntese, as características do IPN [N + algum/a] observadas por Martins (2015) foram: (i) na posição pré-verbal, a sequência invertida [N + algum] assegura, só por si, a interpretação negativa da frase (13 a) e não pode coocorrer com o marcador de negação sentencial “não”, quando em posição pré-verbal (13 b), e, (ii) há obrigatoriedade de adjacência entre o quantificador indefinido e o nome (13 c), pois quando ocorre um termo entre os itens, a sentença se torna agramatical (13 d).

(13)

- a. Homem algum vive aqui.
- b. *Homem algum não vive aqui.
- c. Animal algum vive aqui
- d. *Animal do deserto algum vive aqui.

Com relação aos indefinidos negativos pós-verbais, Martins (1996, 2000) descreve que tanto o indefinido negativo “ninguém” como a sequência invertida [N+algum], podem coocorrer com o marcador de negação predicativa mesmo sendo inerentemente negativos. Essa coocorrência era obrigatória no PA (Martins, 2000) e deixa de ser no século XIX com a gramaticalização da estrutura [N + algum] como um IPN, explicado por um processo de incorporação do Nome e do Indefinido em um núcleo negativo nulo na estrutura funcional do DP.

O paralelismo entre o [N + algum] e o [N + nenhum] também foi objeto de estudo para Martins (2015) que, analisando o percurso das construções do português clássico ao contemporâneo, constatou a existência de uma concorrência entre o indefinido “nenhum” e o quantificador indefinido “algum” em posição pós-nominal, entre os séculos XVIII e XIX, com prevalência do “algum” em posição pós-nominal, mas com um aumento no século XIX das estruturas [N + nenhum], fato que, para a autora, é indício de que também a estrutura [N + nenhum] se gramaticalizou em IPN seguindo os mecanismos que levaram à gramaticalização de [N + algum].

De acordo com Martins (2015, p. 13-14), os dados parecem indicar que na etapa final do processo de gramaticalização, a inversão nominal negativa se alargou de “algum” a “nenhum”. Para a autora, numa “possível interpretação dos dados empíricos, o processo de inversão nominal negativa será uniforme, sendo a diferença entre [N+algum] e [N+nenhum] superficial.” A diferença reside no fato de que no primeiro caso o núcleo Neg não possui expressão fonológica, enquanto que no segundo caso os traços do núcleo Neg se realizariam sob a forma de um morfema negativo ligado ao quantificador indefinido.

3 O indefinido “algum” no corpus Tycho Brahe e no corpus DOViC

Considerando a particular história dos indefinidos na língua portuguesa, em Silva (2001), ampliamos a investigação sobre o tema da valoração negativa do indefinido “algum” em posição pós-nominal [Nome + algum] olhando para o fenômeno em outros corpora. Utilizamos, como corpus da pesquisa, textos de autores portugueses nascidos entre os séculos XVI e XIX, pertencentes ao CTB, e documentos notariais brasileiros, pertencentes ao Corpus DOViC. Assim, mediante a investigação revisitada sobre os indefinidos negativos, a inversão negativa

[N + algum] e o levantamento dos dados realizados por Silva (2021), indagamos o estatuto/natureza da inversão [N + algum] nos corpora representativos do PCL, PE e PB, questionando:

- (i) Qual é o *status* da inversão [N + algum] no PCL século XVI e XVII? A estrutura [N+algum] atestada nos textos do *corpus Tycho Brahe* pode ser considerada um IPN como no PE ou o valor negativo se dá via presença de um operador de negação que precede o NP e desencadeia a valorização do traço negativo do NP, como no PA para as *n-words*?
- (ii) Qual é o *status* da inversão [N + algum] no PB do século XIX? A estrutura [N+algum] atestada nos textos do *corpus DOViC* pode ser considerada um IPN como no PE?
- (iii) É possível identificar o período em que ocorre a mudança do Item de Polaridade Positiva, (doravante IPP), [algum] para um IPN [Nome + algum] no tempo?

Para responder a essas questões, delineamos a metodologia de seleção e análise dos dados, incluindo os fatores de descrição/classificação. Verificamos como ocorre a valoração negativa dos indefinidos que não eram intrinsecamente negativos ou positivos em textos do português clássico e se havia uma concorrência entre o indefinido negativo “nenhum”, e o quantificador indefinido “algum/a” nos textos do *corpus* da pesquisa, controlando a interpretação (se negativa ou não negativa).

Nesta empreita, atestamos dados que não são compatíveis com uma estrutura gramaticalizada de IPN, nos termos definidos por Martins (2015) para o PE, mas que confirmam as conclusões relacionadas à diacronia da língua, como a necessidade de concordância negativa (presença de operador de negação na sentença em que a inversão nominal em relação ao indefinido é atestada com valor negativo).

- (14)
 - a. e que pagaria todos os anos de pareas dous mil fardos de arroz, assim como se obrigara ao Viso-Rei Dom Francisco de Almeida. “E que **não recolheria** em seus portos **navios alguns** de cossairos.

“E que daria lugar pera feitoria pera estarem os Officiaes d’ElRei feitorizando suas cousas.

(COUTO, Diogo do, Décadas, séc. XV, Português europeu)

- b. As asas dobradas se podem pôr aos anjos, e assi mesmo nos pés por mostrarem sua presteza, mas também podem ser pintados **sem terem asas algumas**, e com tal extremidade e tão angélica, que pareça serem anjos, como já os alguém pintou.

(HOLANDA, Francisco de, Da pintura antiga, séc. XVI, Português europeu)

- c. onde seus vassalos o agazalharam, e recolhêram, mandando dali espias a saber de Dom Christovão, e da Rainha, de que **não tinham novas algumas**. Assim os deixaremos todos em sua tristeza, até tornarmos a êles.

(COUTO, Diogo do, Décadas, séc. XV, Português europeu)

- d. Discrição, a quem até a idade de vinte anos se **não tinham feito civilidades algumas**, viu-se obrigada a aperfeiçoar os talentos naturais para suprir ao defeito em que se achava de tantas prendas, quais eram as que observava em sua irmã.

(XAVIER, Francisco. Cartas, Cavaleiro de Oliveira, séc. XVIII, Português europeu)

- e. porém devia permitir-se-lhe dizer, que **não se tinham poupado diligências algumas** para fazer que a Nação portuguesa tivesse um perfeito conhecimento da sua situação;

(Vários, Gazeta de Lisboa, séc. XVIII, Português europeu)

- f. Como se há de ver, lendo esta história, o doente, **não pode prestar**, no estado em que se acha, **informações algumas** sobre os diferentes turnos por que tem passado a moléstia; seu pai porém fez-me o favor de fornecer as informações de que eu precisava.

(Vários, Jornais da Bahia, séc. XIX, Português brasileiro)

- g. A outra nossa tia, a Marquesa de Abrantes, que tinha seu marido e um filho em França, **não obtinha de eles notícias algumas** de o Marquês de Alorna, apesar de lhe as pedir, sempre que se oferecia ocasião.

(ALORNA, Memórias do Marquês da Fronteira e d’Alorna, séc. XIX, Português Europeu)

Nota-se que a pluralidade na inversão negativa é possível, pois temos a sequência [N + algum] no plural sob o escopo do marcador de negação “não” (14 a., c., g.) e da preposição de sentido negativo “sem” (cf. 14b). A possibilidade de flexão no plural é compatível com a análise de “algum” como núcleo da categoria Determinante, conforme iremos propor para a estrutura [N+algum] ainda não gramaticalizada em um IPN, como veremos mais adiante, após observarmos outros fatores relevantes para a proposta.

Ao observarmos os dados extraídos dos textos do Corpus DOViC representativos do PB, notamos a presença da estrutura [N + algum] apenas nas funções de Complemento e Adjunto, não atestamos casos em constituinte sujeito. Os contextos sintáticos em que o NP [N+algum] é complemento ou adjunto, apresentam um contexto de negação para além da negação expressa na inversão do indefinido no interior do NP (15).

(15)

- a. [...] notifiquemos o dito Depozitario, para **não** entregar a **pessoa alguma** a mesma quantia de quinhentos mil reis sem expressa ordem de Justiça, sob pena da Lei dos Depozitarios, de que ficou sciente; ³[...] (Corpus DOViC, Livro 2, carta 15, Séc. XIX)
- b. [...] cujo escravo de hoje em diante fica gozando plena, e inteira liberdade que de hoje em diante lhe transfiro tanto em razão de ser minha cria, como pelos relevantes serviços que me tem prestado; e por isso poderá gozar de inteira liberdade, sem restrição alguma, como se nassese de ventre livre, pois que me o-brigo a sustentar esta carta de liberdade por mim e meus herdeiros ascendentes⁴ [...] (Corpus DOViC, Livro 2, carta 2, Séc. XIX)

O dado “pessoa alguma” apresentado na sentença 15a assume valor negativo no NP, que é objeto indireto, em posição pós-verbal, de uma oração subordinada infinitiva (reduzida) negativa devido à presença do marcador de negação “não”. A presença da negação na sentença – “para **não** entregar a NP” - e a inversão do indefinido no NP complemento - “pessoa alguma” - soam como reforço expressivo da negação sentencial.

³ Texto transcrito na versão original.

⁴ Texto transcrito na versão original.

De maneira semelhante, o dado exposto em 15b possui a expressão “restrição alguma” com valor negativo no NP, que está como adjunto introduzido pela preposição negativa “sem”, carregando um traço de negação, sob o domínio de um IP⁵ (sintagma verbal) positivo – “poderá gozar de...”.

O uso do IPN [restrição + alguma] no interior do sintagma preposicional (PP, do inglês *Prepositional Phrase*) negativo – “sem restrição alguma” - apresenta reforço positivo para as expressões “plena liberdade” e “inteira liberdade”.

Ao observarmos os dados extraídos dos textos do Corpus Tycho Brahe, encontramos os NPs com indefinidos em diferentes funções sintáticas, incluindo a função sujeito e nos deparamos com sentenças como (16) e (17), que parecem questionar a polaridade sempre negativa da estrutura [N + algum] e a impossibilidade de um complemento nominal, ou um adjetivo, ocorrer entre o nome e o indefinido “algum” na posição pós-nominal (cf. MARTINS, 2015).

- (16) Olá, se torno a ouvir de vós **queixa alguma**, juro, pela fé que devo a Balduino meu predecessor, que vos hei-de cozer vivo em uma caldeira, como ele cozeu a outro, que roubou uma viúva pobre. (BERNARDES, séc. XVII, dado coletado no *corpus Tycho Brahe*)
- (17) Porém como esta lei seja voluntária e executada sem rigor, nem **obrigação de justiça alguma**, não querem alguns estar por ela. (GANDAVO, sec. XVII, dado coletado no *corpus Tycho Brahe*)

Os dados analisados, nos textos do CTB, sugerem uma possível ambiguidade na polaridade da estrutura [N + algum], no século XVII, pois, como evidenciado no exemplo 16, mesmo com o indefinido “algum” posposto ao substantivo “queixa”, o sintagma nominal não possui polaridade negativa na sentença, pois no contexto “*se torno a ouvir de vós queixa alguma*” traz a proposição “*tornar a ouvir*”, o que indica que antes se ouviu “*algum queixa*”, e uma promessa de “*...cozer vivo*” caso se volte a ouvir “*queixa*”. A interpretação de “*queixa alguma*”, portanto, deve ser equivalente a “*alguma queixa*”.

⁵ IP (Inflectional Phrase) é um sintagma funcional que possui propriedades flexionais (como tempo e concordância).

Já em 17, o valor de “*obrigação de justiça alguma*” é negativo, porém há um complemento “*de justiça*” que interrompe a adjacência entre o núcleo N e o indefinido, situação esta permitida somente em estruturas sem a incorporação dos Núcleos N e Q(indefinido). No contexto existe uma “*lei*” que é “*voluntária*”, portanto sem nenhuma obrigação da justiça (lei), o núcleo principal negado no enunciado é “*obrigação*” e não “*justiça*”, pois existe uma lei e esta lei é voluntária “*sem obrigação alguma*”.

Mediante esses dados, percebe-se que no século XVII a estrutura [N + algum] não está gramaticalizada uma vez que os dois elementos [N] e [Algum] gozam de certa independência sintática.

Em continuidade, verificamos os indefinidos negativos no português europeu em uma perspectiva histórica comparando as frequências de [N + Algum] com as frequências de [Nenhum + N] e [N + Nenhum]. De acordo com Martins (2015), há um aumento acentuado da frequência da inversão nominal com “*nenhum*” em comparação com o “*algum*” no século XIX, indicando que nesse período, em que, segundo a autora, ocorreria a etapa final do processo de gramaticalização de [N + algum] como um item de polaridade negativa (IPN), a inversão nominal negativa se “*alargou*” de “*algum*” a “*nenhum*” (MARTINS, 2015, p.13). No *Corpus* do Português, utilizado por Martins (2015) para análise, constatou-se que a colocação pós-nominal de “*nenhum*” sobe de 16% no século XVIII para 43% no século XIX.

Partindo desse resultado e com o intuito de verificar o fenômeno em outro *corpus* diacrônico do Português, replicamos a pesquisa de Martins ao *Corpus Tycho Brahe*, com o objetivo de verificar se o fato também pode ser observado nesse *corpus* e se a inversão com “*nenhum*” está associada a gramaticalização do IPN [N + algum] no PE, que, segundo Martins (2016), ocorre entre os séculos XVIII e XIX, ou se a inversão [N + nenhum] já pode ser observada nos séculos anteriores, séculos XVI e XVII em que a não contiguidade entre o núcleo nominal e o indefinido pós-nominal podia ser atestada, indicando uma construção ainda não gramaticalizada, segundo Martins (2016).

A estratégia preferencial para a polaridade negativa nos textos do CTB foi a com o indefinido “*nenhum*” pré-nominal, 602 dados nos textos portugueses (55% do total de 1100 NPs com polaridade negativa) e 90 nos textos brasileiros (47% do total de 191 NPs com polaridade negativa). Os exemplos de 18 a 21 ilustram essa construção.

- (18) O capitão-mor lhe respondeu que os embaixadores tinham seguro para suas pessoas, e licença para dizerem livremente a que eram mandados, pelo que sem **nenhum receio** podia falar o que quisesse. (PINTO, séc. XVI)
- (19) E melhor rezareis assim e não vos fará **nenhum impedimento** dêste modo. (CHAGAS, séc. XVII)
- (20) porque **nenhum homem** quis dizer até agora a loucura de que era branco - dando a atender que, se houvesse um que a proferisse, não faltariam muitos que acressem. (XAVIER, séc. XVIII)
- (21) porém a mana não lhe faz a honra de o pronunciar de **nenhum modo**, conforme as delicadezas do costume, que observa rigidamente. (ALORNA, séc. XIX)

Nota-se que os dados exemplificados, seguem a generalização de Martins sobre a necessidade de concordância negativa quando o NP com “nenhum” pré-nominal segue o verbo e que em posição pré-verbal esta concordância já não é mais obrigatória (exemplo 21).

Atestamos 396 dados em que o indefinido “algum” foi registrado em posição pós-nominal em textos portugueses e 85 dados em textos brasileiros. A grande maioria dos dados (97%) com sentido negativo (apenas 11 dados dos 396 em textos portugueses e 1 dado dos 85 em textos brasileiros possuem interpretação não negativa). A estrutura [N + Algum] com interpretação negativa, pode ocorrer em posição pré ou pós-verbal, sempre em concordância com outro elemento negativo à esquerda do NP, quando este está pós-verbal, como podemos verificar nos exemplos 22 a 25. Também podemos verificar nestes exemplos a possibilidade de complementos e adjuntos no NP e ainda a possibilidade do plural.

- (22) E nos os sete Portugueses que a este tempo, como ja disse, estávamos na praça para nos venderem em leylão, tomamos por remedio mais certo de nossa salvação tornarmo-nos a meter na mazmorra, **sem** que **ministro algum de justiça** ou outrapessoa nos levasse, ou fosse conosco... (PINTO, séc. XVI)
- (23) que **não** faça **caso algum de tudo** o desta vida. (CHAGAS, séc. XVII)

- (24) Discrição, a quem até a idade de vinte anos se **não** tinham feito **civilidades algumas**, viu-se obrigada a aperfeiçoar os talentos naturais para suprir ao defeito em que se achava de tantas prendas, quais eram as que observava em sua irmã. (XAVIER, séc. XVIII)
- (25) A outra nossa tia, a Marquesa de Abrantes, que tinha seu marido e um filho em França, não obtinha de eles **notícias algumas de o Marquês de Alorna**, apesar de lhe as pedir, sempre que se oferecia ocasião. (ALORNA, séc. XIX)

A estrutura com o indefinido algum pós-nominal, com valor negativo, apresenta-se em concordância negativa, podendo flexionar-se, como em 24 e 25, e ter mais elementos no NP como em 22 e 23.

As estruturas negativas [N + algum] representam 36% dos casos (396 casos) contra 55% de [Nenhum + N] (602 casos) e 9% de [N+Nenhum] (102 casos) nos textos portugueses. Nos textos brasileiros, essa distribuição das estratégias de negação no NP é semelhante: 45% dos casos [N + algum] (85 casos) contra 47% de [Nenhum + N] (90 casos) e 5% de [N+Nenhum] (16 casos). Os exemplos 26 a 29 ilustram a opção menos frequente no corpus:

- (26) E assim se partiu sem levar **coisa nenhuma** do que vinha pedir. (PINTO, séc. XVI)
- (27) mas todo o exercício, palavra, letra, meditação, por espiritual que seja, impede a união, que é estar a alma já sem **acção nenhuma sua**, submergida naquele abismo de amor, de glória, de perfeição. (CHAGAS, séc. XVII)
- (28) Este homem não sabe perder ocasião alguma de ostentar os seus talentos, **coisa nenhuma** o prende e **coisa nenhuma** deixa de lhe vir à mão por que ele deixe de pedir. (XAVIER, séc. XVIII)
- (29) Levantou-se minha Mãe e esteve quatro horas fora da cama, sem sentir **abalo nenhum** (ALORNA, séc. XIX)

Esta estrutura, apesar de menos frequente, parece seguir o comportamento sintático de [N] + [algum]. Em posição pós-verbal ocorre com a presença de outro elemento negativo, sem precisar ser

necessariamente a negação sentencial “não”, nos exemplos a preposição “sem” foi atestada (29).

O dado exemplificado em 28 (século XVII) ilustra a possibilidade do possessivo na periferia direita do NP, ou seja, uma estrutura com mais elementos no sintagma, como acontece com [N + algum] nesse período. No século XIX este tipo de construção não foi mais atestada nos períodos posteriores.

Vimos que no *Corpus* do Português, Martins (2015) constatou que a colocação pós-nominal de “nenhum” sobe de 16% no século XVIII para 43% no século XIX, o cálculo da frequência foi feito por Martins com base na soma dos totais de “nenhum” pré e pós-nominal em cada século, o mesmo vale para o cálculo de algum na tabela abaixo.

Tabela 1 - Dados coletados por Martins em *Corpus* do Português (CP)

Período	Nenhum		Algum	
	Pré-nominal	Pós-nominal	Pré-nominal	nominal
Século XVIII	325	63 - 16,2%	2220	391 - 15%
Século XIX	676	504 - 42,9%	8726	812 - 7,6%

Fonte: MARTINS (2015, p.14)

Baseando-nos em Martins (2015), ao fazermos o mesmo tipo de relação e cálculo das ordenações dos indefinidos no corpus Tycho Brahe constatamos que a inversão [N + nenhum] é atestada já no século XVI, com uma frequência semelhante a da inversão de [N + algum], 14% e 12%, respectivamente, em relação à ordem pré-nominal do indefinido. Detectamos um contínuo nos padrões de frequência das inversões desde o século XVI até o século XIX, ambos passaram de um patamar inferior a 20% de frequência nos séculos XVI e XVII a um patamar superior a 20% de estruturas com indefinidos pós-nominais nos séculos XVIII e XIX, como podemos constatar na tabela 2. É importante ressaltar que a tabela 2 considera apenas a posição do indefinido e não o valor polar, seguindo o mesmo tipo de cálculo da tabela 1 reproduzida de Martins (2015).

Tabela 2 - Posição dos indefinidos ALGUM e NENHUM no sintagma nominal (documentos portugueses)

	Algum		Nenhum	
	Pré-nominal [Algum/a + N]	Pós-nominal [N + Algum/a]	Pré-nominal [Nenhum/a + N]	Pós-nominal [N+Nenhum/a]
Século XVI	568	94 - 14%	267	39 - 12 %
Século XVII	551	119 - 16%	172	11 - 6%
Século XVIII	234	137 - 37%	113	30 - 21%
Século XIX	169	47 - 22%	50	23 - 32%

Fonte: Elaboração própria

Uma vez que diferentemente de [Algum + N], [N + algum] pode realizar e normalmente realiza a polaridade negativa do sintagma, juntamente com outro marcador negativo nas fases mais antigas, assim como [Nenhum + N] e [N + Nenhum], excluimos os dados de algum pré-nominal para observar apenas as estratégias de polaridade negativa no tempo e verificamos que a estratégia mais frequente para polaridade negativa no sintagma é sempre [Nenhum + N], sendo a estrutura com posposição de algum mais frequente que a posposição de nenhum, como se pode verificar na tabela 3. Nota-se na tabela que há um aumento da estrutura [N + algum] que passa de 23% no século XVI para 39% e 49% nos séculos XVII e XVIII para voltar a baixar para 39% no século XIX. Já a estrutura [N + nenhum] se mantém abaixo dos 20% em todo período, (10%, 4%, 11% e 19% nos séculos XVI, XVII, XVIII e XIX, respectivamente).

Tabela 3 - Posição dos indefinidos ALGUM e NENHUM no sintagma nominal com valor negativo (documentos portugueses)

	Algum	Nenhum	Total
	Pós-nominal [N + Algum/a]	Pré-nominal [Nenhum/a + N]	
Século XVI	94 - 23%	267 - 67%	400
Século XVII	119 - 39%	172 - 57%	302
Século XVIII	137 - 49%	113 - 40%	280
Século XIX	47 - 39%	50 - 42%	120

Fonte: Elaboração própria

Os dados atestados nos textos brasileiros do século XIX repetem exatamente o mesmo padrão de frequência que os textos portugueses do mesmo período, como se pode verificar na tabela 4.

Tabela 4 - Posição dos indefinidos ALGUM e NENHUM no sintagma nominal com valor negativo (documentos brasileiros)

	Algum	Nenhum		
	Pós-nominal [N + Algum/a]	Pré-nominal [Nenhum/a + N]	Pós-nominal [N+Nenhum/a]	Total
Séc XIX - BRA	85 - 45%	90 - 47%	16 - 8%	191

Fonte: Elaboração própria

A estrutura [N+algum] ganha espaço como preferência de IPN no tempo, sendo seu uso equivalente em termos proporcionais à estratégia mais usada desde o século XVI [Nenhum + N]. A diferença no tempo com relação à frequência é um fator relevante quando somada à qualidade das estruturas. É somente no século XIX que dados com inversão do indefinido contendo flexão para o plural e mais elementos no NP deixam de ser atestados, enquanto dados sem a presença de negação sentencial ou outros operadores negativos são atestados com valor negativo como ilustra o exemplo 30.

- (30) Em 1815, perguntando minha Avó a o General Gomes Freire, diante de mim, se os movimentos militares que deixo referidos seriam devidos a alguma traição, ele respondeu lhe que de **modo algum**, porque a corte de o Príncipe Regente era tão tola que nem para isso tinha capacidade. (ALORNA, séc. XIX)

Não obstante, ao comparar apenas as estruturas com posposição, constatamos que a frequência de [N + algum] é sempre muito superior à frequência [N + nenhum], de podemos verificar na tabela 5, fato que pode corroborar a hipótese de Martins (2015) em relação a origem do IPN com o indefinido “nenhum” na generalização do IPN [N + algum], gramaticalizado no século XIX.

Tabela 5 - Posposição dos indefinidos ALGUM e NENHUM no sintagma nominal com valor negativo (documentos portugueses e brasileiros)

	Algum	Nenhum	
	Pós-nominal [N + Algum/a]	Pós-nominal [N+Nenhum/a]	Total
Século XVI	93 - 71%	39 - 29%	132
Século XVII	119 - 91%	11 - 9%	130
Século XVIII	137 - 82%	30 - 18%	167
Século XIX	47 - 67%	23 - 33%	70
Séc XIX - BRA	85 - 86%	16 - 14%	101

Fonte: Elaboração própria

Apesar da frequência de [N+algum] já ser bastante alta no século XVI (71%) e não apresentar um aumento progressivo e gradual no tempo, e portando não ter se mostrado um argumento suficiente para identificar a mudança no tempo, nos textos do século XVIII e XIX não atestamos mais dados da ordem [N + algum] com valor positivo (não negativo), tão pouco casos de flexão do substantivo e do indefinido, tais ausências são relevantes, pois podem ser justificadas pela gramaticalização da inversão nominal em IPN. Paralelamente a isso, o valor negativo do NP com inversão do indefinido foi atestado sem a presença de um operador de negação no contexto, como vemos no exemplo 30 acima e 31 abaixo, o que pode sugerir que, nessa fase, a ordem do [N + algum] se estabelece como um IPN.

- (31) Se a confusão dos últimos tempos lhe fizeram perder de vista quanto me ordenou, nada mo poderia fazer esquecer a mim, vendo crescer os perigos da minha Pátria e de Vossa Alteza Real Tudo consegui, mas **cousa alguma** se fará se Vossa Alteza Real, por uma carta sua, me não recomenda a Suas Majestades Católicas de um modo tal, que eu não tenha obstáculo para tratar, em nome de Vossa Alteza Real, tudo quanto lhe é preciso e pode desejar. (ALORNA, SÉC. XVIII, dado coletado no *corpus Tycho Brahe*)

Também, em textos brasileiros do CTB foi atestado [N + algum] na função sujeito, com valor negativo e sem este estar sob o escopo de outro elemento negativo.

- (32) Conseqüentemente, entendo que **amigo algum** deixará de concordar comigo, não podendo, como não posso, admitir enquanto investido do cargo de chefe político, que alguém, por ameaças queira impor-se. (CARTAS BRASILEIRAS: CULTOS, séc. XIX)

Este tipo de contexto não foi atestado no corpus DOViC. Nos documentos brasileiros do século XIX do corpus DOViC, a estrutura [N + algum] com valor negativo foi atestada em sentenças não negativas, ou seja, sem a presença da categoria funcional NegP, realizada pelo marcador de negação sentencial “não” mas com presença de um outro item negativo. São licenciadores da valoração negativa do NP com posposição do indefinido nesse corpus: a conjunção “nem” e a preposição “sem”, além da negação sentencial, sendo a posposição do indefinido sempre acompanhada de outra palavra negativa no domínio imediatamente superior, seja ele a sentença (IP), como em (33 a.); ou um sintagma preposicional (PP) com a preposição negativa “sem” dominando o NP, como em (33 b.).

(33)

- a. Primeiramente disse ser casado com Dona Reginalda Maria de Jezus de cujo consorcio **não** teve **filho algum** – (Corpus DOViC, Livro 16, documento 5)
- b. cujo escravo de hoje em diante fica gozando plena, e inteira liberdade que de hoje em diante lhe transfiro tanto em razão de ser minha cria, como pelos relevantes serviços que metem prestado; e por isso poderá gozar de inteira liberdade, **sem restrição alguma**, como se nassese de ventre livre, pois que me o-brigo a sustentar esta carta de liberdade por mim e meus herdeiros ascendentes[...] (Corpus DOViC, Livro 2, carta 2)

O indefinido algum, em contextos como os exemplificados acima, pode ser substituído pelo indefinido “nenhum”, sem que ocorra a perda do sentido nas orações.

Em (33 a.) a estrutura [N + algum], ou seja, filho algum, apresenta-se sob o escopo do marcador de negação sentencial “não”, como objeto direto do verbo “ter” e seu uso na sentença permite negar a existência do

substantivo dentro do sintagma nominal. No exemplo (33 b.) a estrutura [N + algum] (restrição alguma) está contida em uma oração subordinada, em adjunção na posição pós-verbal e precedida pela preposição “sem”, e apesar da sentença subordinada não ser negativa, a preposição núcleo do PP que domina imediatamente o NP [restrição alguma] possui um traço de negação, ainda que não sentencial. A presença desse traço no domínio imediatamente superior ao NP se mostrou relevante para o licenciamento da polaridade negativa do NP [N + algum] nos dados desse corpus.

Assim, podemos concluir que a diminuição, ainda que tênue, da frequência de [N+nenhum] como estratégia de polaridade negativa do NP em detrimento das outras configurações, respectivamente [N + algum] e [N + nenhum] em termos de frequência (tabela 3) somada às alterações nos aspectos qualitativos das estruturas apresentadas neste artigo corroboram a hipótese de Martins (2015, p. 13-14) de que a mudança na frequência de [N+nenhum] ocorreria devido à diferença dos aspectos morfológicos e semânticos que podem estar associadas à gramaticalização da sequência [N + algum] como IPN entre os séculos XVIII e XIX.

Os documentos brasileiros do Corpus DOViC apresentam a inversão N + algum, somente em contextos em que outro elemento negativo está no contexto escopando o NP indefinido, assim como no PCL. Já os documentos brasileiros do CTB apresentam dados em função sujeito livre da presença de outro elemento negativo de maneira semelhante ao PE e essa diferença pode ser explicada pela diferença de tamanho e natureza dos corpora.

4 Proposta de análise: a formulação das hipóteses

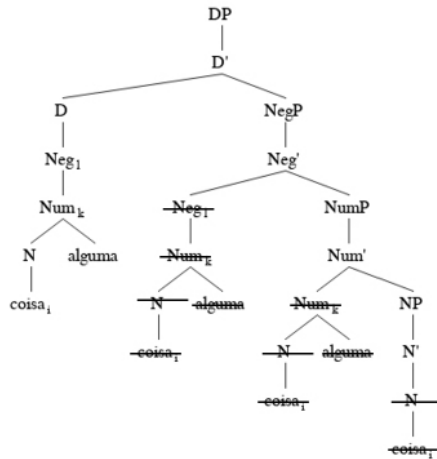
Segundo Martins (2015), o IPN [N + algum] no PE é derivado da incorporação do nome e do quantificador indefinido num núcleo negativo abstrato interno ao DP (Determiner Phrase). A análise formal proposta se apoia na ideia de paralelismo estrutural entre a frase e o sintagma determinante (DP), em particular no que diz respeito à projeção da categoria funcional NegP. A autora afirma que é o movimento cíclico do nome (N) através da estrutura funcional do DP que determina que esse se associe ao quantificador indefinido no caminho para o lugar final de incorporação (34).

(34)

[DP [D" [D [Neg₁ [Num_k coisa_i alguma]]] [NegP [Neg" [Neg_T [Num_k coisa_i alguma-]]] [NumP [Num" [Num_k coisa_i alguma] [NP coisa_i]]]]]]] (baseado em MARTINS, 2015, p. 17)

Para formar o IPN o movimento do núcleo nominal é realizado para o núcleo do sintagma numeral (NumP, do inglês *Number Phrase*) onde, segundo Martins (2015), está o indefinido “algum”, essa operação forma o amálgama [N + algum]. O núcleo “Num” (contendo o amálgama) se move ainda para o núcleo da camada funcional NegP (*Negative Phrase*), do constituinte nominal, para realizar o traço de polaridade negativa e, na sequência, o núcleo Neg, já contendo a incorporação dos núcleos inferiores, se move para D (figura 1).

Figura 1 – Representação arbórea sintática do IPN [animal algum]



Fonte: Elaboração própria, baseada em Martins (2015, p. 17)

A estrutura sintática do IPN, tal como propõe Martins (2015) explica a obrigatoriedade de adjacência entre o nome substantivo e o quantificador indefinido, não possibilitando a presença de um complemento nominal ou adjunto ocorrerem entre o nome e o algum pós-nominal por esta estrutura ser formada em um processo de incorporação de núcleos - um amálgama - formando uma única palavra

negativa (*n-word*), um IPN. Nessa configuração a flexão de número é bloqueada, “coisas algumas” não é possível, é agramatical, logo, o IPN comporta-se como uma palavra invariável, como os pronomes indefinidos “ninguém”/“nada”. A explicação, segundo Martins (2015), se dá pela ausência do núcleo do Pl(ural) na estrutura do DP quando NegP é projetado e pela possibilidade de ocorrer em qualquer posição na sentença sem a presença de um operador de negação para a interpretação negativa.

Essa não deve ser a estrutura da inversão nominal [N + algum] no PCL, uma vez que a polaridade positiva/não negativa para a estrutura [N + algum] bem como a não adjacência estrita entre o nome e o indefinido e a pluralização do sintagma na inversão nominal negativa com o indefinido “algum” foi atestada em textos do *corpus Tycho Brahe*, exemplos retomados em (35).

(35)

- a. Olá, se **torno a ouvir de vós queixa alguma**, juro, pela fé que devo a Balduino meu predecessor, que vos hei-de cozer vivo em uma caldeira, como ele cozeu a outro, que roubou uma viúva pobre. (BERNARDES, sec. XVII, dado coletado no *corpus Tycho Brahe*)
- b. Porém como esta lei seja voluntária e executada sem rigor, **nem obrigação de justiça alguma**, não querem alguns estar por ela[...] (GANDAVO, sec. XVI, dado coletado no *corpus Tycho Brahe*).
- c. Discrição, a quem até a idade de vinte anos se **não tinham feito civildades algumas**, viu-se obrigada a aperfeiçoar os talentos naturais para suprir ao defeito em que se achava de tantas prendas, quais eram as que observava em sua irmã. (XAVIER, sec. XVIII, dado coletado no *corpus Tycho Brahe*)

Vimos que o contexto da sentença “se torno a ouvir de vós queixa alguma” traz a proposição “tornar a ouvir”, o que indica que antes se ouviu “alguma queixa”, e uma promessa de “...cozer vivo” caso se volte a ouvir “queixa”. A interpretação de “queixa alguma” em 35 a., portanto, deve ser equivalente a “alguma queixa”. Já em 35b. o valor de “obrigação de justiça alguma” é negativo, no contexto existe uma “lei” que é “voluntária”, portanto, sem nenhuma obrigação da justiça (lei), o núcleo principal negado no enunciado é “obrigação” e não “justiça”, pois existe uma lei e esta lei é voluntária “sem obrigação alguma”. Nesse caso temos a não adjacência

entre o núcleo nominal e o indefinido *alguma* em um NP com o indefinido pós-nominal e interpretação negativa do sintagma. Por outro lado, a flexão plural é atestada em 35c: “civilidades *algumas*”.

Os fatos reapresentados em 35, somados à dependência da concordância negativa nos textos do PCL constituem evidência de que, nos textos até o século XVIII, a inversão [N + *algum*] não corresponde a um IPN, nos termos de Martins (2015) para o PE.

Também os dados oitocentistas do PB sugerem uma dependência de um outro elemento negativo na sentença para instanciar a polaridade negativa do NP, podendo ser explicado por valoração do traço de negação mediante concordância com o núcleo que subordina o NP com inversão [N + *algum*] (cf. 36 a-b):

(36)

- a. [PP+Neg [P *sem*] [DP ... embargo *algum* ... [NP-embargo]]]
- b. [NegP [Neg *não*] [IP havendo [VP ... [DP embargo *algum*]]]

Para Martins (2015, p. 10), no PCL, [N + *algum*] teria na estrutura subjacente o N em *Neg* e o *algum* em *NumP*, semelhante ao espanhol de hoje.

(37) [DP [D' [e] [NegP [Neg' [animal_i]]k [NumP *algum* [Num' [animal_i]]k [NP animal_i]]]]]]

Nessa estrutura o nome e o quantificador indefinido não estão incorporados, o quantificador realiza o especificador da categoria *NumP* e o substantivo se move para o núcleo *Neg* na parte funcional do NP de maneira independente. Não obstante, a proposta do quantificador indefinido ocupar o especificador da categoria *NumP* não contempla seu comportamento morfosintático de núcleo D.

Brito (2003), sobre o grupo nominal (NP), descreve sua parte funcional e analisa os quantificadores indefinidos “*algum*” e “*nenhum*” como uma classe diferente dos quantificadores universais, dos quantificadores variáveis e dos numerais, uma vez que os quantificadores indefinidos estão em distribuição complementar com os elementos da categoria determinante (artigos e pronomes demonstrativos) e não podem co-ocorrer com os demais elementos da categoria D (38 a), os quantificadores universais antecedem os elementos dessa categoria e os variáveis e numerais os seguem (38 b).

(38)

- a. *Algum um livro | Um algum livro ...
- b. Todos os três livros ...

Brito (2003, p. 356) argumenta que os quantificadores incluem diversos tipos de elementos em três classes diferentes:

- (i) os que exprimem a quantificação existencial (um, uns, uma, umas, algum ...) e realizam o núcleo da categoria funcional denominada “determinante” (D) na teoria X-Barra da Gramática Gerativa.
- (ii) os quantificadores (Qs) discretos, que incluem os numerais que exprimem cardinalidade ou ordem (dois, três, primeiro, ...) e ainda os Qs que indicam pluralidade (inúmeros, muitos, vários ...), estes seguem a categoria D na parte funcional do NP.
- (iii) os quantificadores universais (todos, ambos), estes precedem a categoria D na parte funcional do NP.

O indefinido “algum” em posição pré-nominal é um quantificador usual, que exprime quantidade existencial, conforme exemplifica Brito (2003, p. 356) (39).

(39) Li alguns livros. (BRITO, 2003, p. 356)

Já em posição pós-nominal, possui sentido negativo, como também observa Brito (2003, p. 359) (40).

(40) Livro algum estava sobre a mesa. (BRITO, 2003, p. 359)

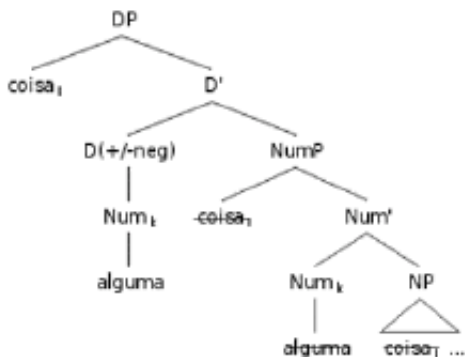
A autora argumenta que os quantificadores existenciais “algum” e “nenhum” são núcleos D (Determinantes), estão em distribuição complementar com os artigos indefinidos, por exemplo, e esta é para nós uma propriedade relevante que guiará nossa análise sobre a estrutura sintática de [N+algum] e a polaridade negativa do DP.

Propomos que a estrutura em que o quantificador se move como núcleo até a camada D da estrutura funcional explica sua distribuição

com os determinantes e quantificadores numerais, inclusive no PCL. Nesse sentido, a derivação em que a parte lexical do sintagma (o NP) se move para o especificador de DP parece ser mais adequada para derivar a inversão [N + algum] no PCL. Tal hipótese tem a vantagem de também explicar a distribuição do indefinido em relação aos demais determinantes (que ocorrem em distribuição complementar com o indefinido), fato que corrobora a hipótese de que o indefinido realiza a categoria D e pode explicar: (i) o valor não negativo da inversão atestado nos dados do PCL e (ii) a possibilidade de não adjacência entre o núcleo nominal e o quantificador, atestada nas fases mais antigas, além de explicar a possibilidade do pronome possessivo à direita do NP como no exemplo (27) “ação nenhuma sua”. Na hierarquia dos núcleos funcionais do NP, a categoria POSSP estaria, segundo Brito (2003), entre D e NUM “os_[D] meus_[poss] três_[num] livros”. A inversão [N + algum], nesse caso, não pode ser um amálgama, a estrutura sintática do sintagma e das palavras no sintagma estão preservadas, portanto não se trata de incorporação dos núcleos N e D. A estrutura do sintagma [N + algum] no PCL seria, portanto, como a representação em (41), ilustrada na figura 2.

(41) [DP [coisa_i] [D’ [D(+/-neg) [Num_k alguma]] [NumP [e_{coisa}_i] [Num’ [Num_k alguma] [NP coisa_i ...]]]]]

Figura 2 – Representação arbórea de [coisa alguma] não especificada quanto ao traço de polaridade no PCL



Fonte: Elaboração própria

Adotamos a estrutura proposta em (41) e representada na figura 2 para a estrutura [N+ algum] no PCL por considerarmos que contempla a hipótese de que nesta fase da língua a posposição do indefinido não é derivada de um processo de incorporação para valorar o traço negativo no NP, pois a inversão só ocorre com a presença de outro marcador de polaridade negativa, como a negação sentencial, as preposições e conjunções negativas. A hipótese nos parece mais econômica em relação à mudança para o PE, uma vez que não envolve alteração na natureza sintagmática do indefinido de categoria máxima (XP), como seria o caso da representação (37) do espanhol, para categoria mínima (X) como seria o caso da representação do IPN em 34. Assumindo a representação em 41 para o PCL, a mudança envolveria apenas o mecanismo de valoração da polaridade no NP, sendo o indefinido sempre núcleo (X) da categoria funcional do NP ligada a sua determinação: a parte funcional do NP passa a espelhar a sentença, instanciando um núcleo de polaridade forte no interior do sintagma nominal, que em 34 (MARTINS, 2015) é denominado NegP, e a valoração da polaridade deixa de ser feita sob escopo/regência/concordância do núcleo de negação no domínio externo ao NP e que o domina (IP[neg] ou PP[neg]) em uma configuração de c-comando do operador de negação em relação ao NP como em (36 a-b), e passa a ser feita pela checagem dos traços polares na camada funcional do NP por palavra negativa. Com essa mudança, a estrutura [N + algum] se gramaticaliza, tornando-se um IPN, como em 34.

Os dados oriundos dos textos brasileiros do século XIX também são compatíveis com a hipótese representada em (41) para o PCL, uma vez que os dados com inversão estão sempre acompanhados do operador de negação sentencial ou outro operador de negação que tem escopo sobre o NP e o c-comanda (como a preposição “sem” ou a conjunção “nem”).

Com relação à mudança gramatical e sua localização no tempo, “as mudanças nas línguas, instanciadas nos documentos históricos, como variação gradual, são reflexos de mudança gramatical que, por uma necessidade teórica, deve ocorrer de modo abrupto na aquisição da linguagem pelo falante (KROCH, 1989 *apud* GALVES; NAMIUTI; PAIXÃO DE SOUSA, 2006, p. 49). A variação de usos nos textos pode ser compreendida como fruto da convivência, no plano do uso, de formas geradas por diferentes gramáticas, formas novas nos textos podem ser pistas para uma mudança gramatical. Segundo Kroch (2001 *apud* GALVES; NAMIUTI; PAIXÃO DE SOUSA, 2006), a mudança

gramatical não afeta apenas uma construção, mas a gramática como um todo, logo se várias alterações ocorrem em um mesmo período de tempo em uma “razão constante”, nos termos de Kroch (1989), temos aí indícios que uma mudança gramatical ocorreu. No caso dos indefinidos, Martins (2015) localizou formas novas no século XIX como a inversão do indefinido com valor negativo sem a presença de um operador de negação, como no caso de “Coisa alguma escapou!” (MARTINS, 2015, p. 12), dados assim também foram atestados por esta pesquisa no século XIX, *corpus Tycho Brahe*. Paralelamente a estes novos fatos, Martins (2015) atestou um aumento da frequência [N + nenhum] no mesmo século o que pode corroborar a hipótese da mudança. O mesmo fato foi por nós observado no *corpus Tycho Brahe*. Sendo assim, com base nos fatos descritos em (35a-c) e (36a-b), também com base nas considerações sobre a mudança gramatical no quadro teórico da gramática gerativa, respondemos às questões levantadas neste trabalho com as seguintes hipóteses:

- (a) A estrutura [N + algum] não pode ser analisada como um IPN no PCL, pois o valor negativo se dá via presença de um operador de negação que precede o NP e desencadeia a valorização do traço negativo do NP, como no PA para as *n-words*, o que evidencia um traço fraco de polaridade no NP no PCL.
- (b) Temos, nos documentos brasileiros do século XIX, um comportamento semelhante ao do PCL em relação às estruturas de inversão dos indefinidos para valorar traço de negação, indicando também um traço fraco de polaridade no NP do PB do século XIX.
- (c) É possível identificar o século XIX como o momento da mudança do PCL para o PE, uma vez que dados novos de [N + algum] foram atestados nos documentos portugueses do século XIX, associados a alterações de frequência de estruturas que podem estar relacionadas à mudança como o aumento de [N + nenhum] em relação a [Nenhum + N] após um aumento de [N + algum] e o surgimento de dados de [N + algum] com valor negativo em um contexto livre da presença de negação sentencial ou outro marcador de negação.

5 Considerações finais

As discussões promovidas neste texto contribuem para a construção do conhecimento sobre as estruturas com indefinidos, colaborando com o avanço das pesquisas sobre a polaridade negativa do NP. Ao realizarmos um estudo comparativo, quantitativo e qualitativo, com base em corpora diacrônicos de textos brasileiros e textos do português europeu, disponibilizamos uma interpretação sobre a composição dessas línguas, bem como estabelecemos comparações entre o PCL, PE e PB, o que nos permitiu atestar possíveis diferenças entre as gramáticas no tempo (século XVI a XIX) e espaço da língua portuguesa (Portugal, Brasil).

Os resultados apresentados possibilitam a percepção de que o indefinido “algum”, na posição pós-nominal, nos séculos XVI e XVII, ainda não estava incorporado ao nome, caracterizando um IPN.

Todavia, localizamos pistas nos dados do século XVIII que sugerem a gramaticalização da estrutura [N + algum] como um IPN, como a possibilidade de [N + algum] ocorrer em estrutura sem a presença de operador de negação e ainda assim carregar valor negativo, e o fato das ocorrências da estrutura com flexão e outros elementos no sintagma deixarem de ser atestadas em textos mais tardios (século XIX).

Sobre os dados do PB, verificamos que a estrutura [N + algum/a] no corpus DOViC sempre é atestada com valor negativo e sempre em um contexto de negação, o que pode sugerir que o valor negativo do NP seja valorado pela negação superior, juntamente com a inversão do algum, como no PCL. Todavia uma investigação envolvendo um corpus maior e mais abrangente se faz necessária para confirmar tal hipótese, uma vez que no corpus-DOViC a construção só foi atestada em posição pós-verbal, na função de complemento ou adjunto, e ainda que há dados, no CTB, em documentos brasileiros escritos por falantes “cultos” que atestam [N+ algum] com valor negativo em contextos pré-verbais sem a presença de operadores de negação.

Declaração de autoria

Este artigo, produzido em coautoria por Cristiane Namiuti e Fernanda Gusmão Silva, traz alguns dos resultados do levantamento de dados da pesquisa de mestrado de Fernanda Gusmão Silva, orientada por Cristiane Namiuti, com proposta de análise e formulação de hipóteses sugeridas e argumentadas pelo primeiro autor.

Agradecimentos

Agradecemos à Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB) pelo financiamento do projeto temático coordenado por Cristiane Namiuti, à CAPES pela bolsa de mestrado de Fernanda Gusmão Silva e ao Programa de Pós-Graduação em Linguística (PPGLIN/UESB).

Referências

BAKER, M. *Incorporation: a theory of grammatical function changing*. Chicago: University Chicago Press, 1988.

BRITO, A. M. Categorias Sintáticas. In: MATHEUS, M. H. M. et al. (orgs). *Gramática da Língua Portuguesa*. Lisboa : Caminho, 2003. p. 325-432.

CHOMSKY, N. *The Minimalist Program*. Cambridge: MIT Press, 1995.

GALVES, C.; NAMIUTI, C.; PAIXÃO DE SOUSA, M. C. Novas perspectivas para antigas questões: revisitando a periodização da língua portuguesa. In: ENDRUSCHAT, A.; KEMMLER, R.; SCHAFFER-PRIEB, B. (orgs.) *Grammatische Strukturen des europäischen Portugiesisch: Synchrone and diachrone Untersuchungen zu Tempora, Pronomina, Präpositionen und mehr*. Tübingen: Calepinus Verlag, 2006. p.45-74.

GALVES, C.; ANDRADE, A.; FARIA, P. *Tycho Brahe Parsed Corpus of Historical Portuguese*. Campinas, s/d. Disponível em: URL: <http://www.tycho.iel.unicamp.br/~tycho/corpus/>

KROCH, A. Reflexes of grammar in patterns of language change. In: LABOV, W.; CACOULOS, R. T. (eds.). *Language Variation and Change*. Cambridge: Cambridge University Press, 1989. p. 199 - 244.

MARTINS, A. M. Aspectos da negação na história das línguas românicas. Da natureza de palavras como nenhum, nada, ninguém. In: CASTRO, I. (org.). *Actas do XII Encontro Nacional da Associação Portuguesa de Linguística: Linguística Histórica e História da Linguística*. Lisboa: Associação Portuguesa de Linguística, 1996. p. 179-210.

MARTINS, A. M. Mudança Sintática. Clíticos, negação e um pouquinho de Scrambling. *Estudos Lingüísticos e Literários*, Salvador, n. 19, p. 129-162, 1997.

MARTINS, A. M. Polarity Items in Romance: underspecification and lexical change. In: PINTZUK, S.; TSOULAS, G.; WARNER, A. (ed.). *Diachronic syntax: models and mechanisms*. Oxford/New York: Oxford University Press, 2000. p. 191- 219.

MARTINS, A. M. Introdução: O português numa perspetiva diacrónica e comparativa. In: MARTINS, A. M.; Carrilho, E. (orgs.). *Manual de Linguística Portuguesa*. Berlin/Boston: De Gruyter, 2016. p. 1-39.

MARTINS, A. M. *Ordem de Palavra e Polaridade: Inversão Nominal Negativa com algum/ alguno e nenhum*. 29. ed. Lisboa: Universidade do Minho, 2015. Disponível em: <http://hdl.handle.net/10451/31174>. Acesso em: 31 ago. 2021.

MEILLET, A. L'”évolution des formes grammaticales. In: MEILLET, A. *Linguistique historique et linguistique générale*. Paris: Champion, 1948. p. 130-148.

MIOTO, C. *Negação Sentencial no Português Brasileiro e Teoria da Gramática*. 1991. 244 f. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, 1991.

NAMIUTI, C. *Aspectos da história gramatical do português: interpolação, negação e mudança*. 2008. 309 f. Tese (Doutorado em linguística) - Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, 2008.

PINTO, C. Para a história da negação: o minimizador “homem” no português antigo. *Estudos de lingüística galega*, v. 7, p. 109-123, 2015. DOI: <http://dx.doi.org/10.15304/elg.7.2335>.

SANTOS, J. V.; NAMIUTI, C. *DOVIC* (Documentos Oitocentistas de Vitória da Conquista e região). Corpora Eletrônico. Vitória da Conquista: UESB, 2014.

SILVA, F. G. *O indefinido “algum” na formação do IPN [N+algum] na história da língua portuguesa: um estudo diacrônico sobre a polaridade negativa no sintagma nominal*. 2021. 98f. Dissertação (Mestrado em Linguística) - Universidade Estadual do Sudoeste da Bahia, 2021.

ZANUTTINI, R. *Syntactic properties of sentential negation. A comparative study of Romance languages*. Philadelphia: University of Pennsylvania, 1991.

Argumentos verbais



Locative adverbs in the subject position in Brazilian Portuguese

Advérbios locativos na posição de sujeito no português brasileiro

Zenaide Dias Teixeira

Universidade Estadual de Goiás (UEG), Luziânia, Goiás / Brazil

zenaide.teixeira@ueg.br

<http://orcid.org/0000-0002-6103-5923>

Humberto Borges

Universidade Federal de Jataí (UFJ), Jataí, Goiás / Brazil

humbertoborges@ufj.edu.br

<http://orcid.org/0000-0001-5492-7264>

Abstract: This article examines the syntactic and semantic properties of locative adverbs in Brazilian Portuguese to verify whether they occupy the subject position. We observe that adverbs can occupy the subject position as well as locative NPs. We propose that, contrary to full DPs, adverbs do not bear the number feature inherently, manifesting only the person feature. Following Baker (2003), we propose that locative adverbs bear a referential index. The conclusion is that locative adverbs in subject position are generated as VP adverbs or locative arguments, bearing the following formal features: [+deictic] and [+ (3rd) person].

Keywords: locative adverbs; position of the subject; Brazilian Portuguese.

Resumo: Este artigo examina as propriedades sintáticas e semânticas dos advérbios locativos no português brasileiro para verificar se eles ocupam a posição de sujeito. Observamos que advérbios podem ocupar a posição de sujeito assim como NPs locativos. Propomos que, contrariamente aos DPs plenos, advérbios não carregam o traço de número inerentemente, manifestando somente o traço de pessoa. Seguindo Baker (2003), propomos que advérbios locativos possuem índice referencial. A conclusão é a de que esses advérbios na posição de sujeito são gerados como advérbio de VP ou argumento locativo, carregando os traços formais: [+dêítico] e [+ (3^a) pessoa].

Palavras-chave: advérbios locativos; posição do sujeito; português brasileiro.

Recebido em 07 de janeiro de 2022

Aceito em 05 de abril de 2022

eISSN: 2237-2083

DOI: 10.17851/2237-2083.30.4.1788-1830

1 Introduction

In this work, we discuss the relationship between locative adverbs and their corresponding locative NPs and PPs, examining their syntactic properties and the hypothesis that they occupy the subject position in Brazilian Portuguese (BP). We formulate the following questions: do locative adverbs occupy the subject position in BP? What are the syntactic and semantic factors responsible for licensing adverbs in the subject position? What factors restrict the presence of other adverbs in this position?

To discuss those issues, in sections 2 and 3, we turn to studies that examine locative inversion and the so-called *bare-NP adverbs*. In section 4, we consider studies that discuss and examine the possibility of locative PPs occupying the subject position in BP and the hypothesis of the locative NP in this position. Those studies are relevant to this work because of the possibility that phrases that share distributional properties may share categorial properties, as evidenced in (1).

- (1) Brazilian Portuguese
- a. *Nessa casa* bate sol.
In-this house hit3SG sun
 ‘This house gets a lot of sunlight.’
- b. *Essa casa* bate sol.
This house hit3SG sun
 ‘This house gets a lot of sunlight.’
- c. *Aqui* bate sol.
Here hit3SG sun
 ‘This house gets a lot of sunlight.’

In section 5, we analyze the locative adverb in the subject position in BP. And, finally, we make our final considerations in section 6.

2 Locative inversion and subject position

In a lexical-functional analysis, Bresnan (1994) investigates structures of English and Chichewa, one of the Bantu languages of Central and Eastern Africa, which involve a preposed locative phrase and a postposed subject. Such structures, illustrated by the English data below, are known as locative inversion (LI).

- (2) a. A lamp was in the corner. (BRESNAN, 1994, p. 75)
 b. In the corner was a lamp.
- (3) a. My friend Rose was sitting among the guests. (BRESNAN, 1994, p. 75)
 b. Among the guests was sitting my friend Rose.
- (4) a. The tax collector came back to the village. (BRESNAN, 1994, p. 75)
 b. Back to the village came the tax collector.

According to the author, locative inversion – as in (2b), (3b) e (4c) – is associated with a theme-locative argument structure. The semantic role of the theme universally alternates between object and subject positions. The theme is the syntactic object of an active transitive verb and the syntactic subject of a passive transitive verb. So, in languages with locative inversion, intransitive verbs have a theme argument as the subject, but allow it to appear in the post-verbal position in LI, in addition to selecting locatives as arguments. Verbal intransitivity, according to Bresnan (1994), is a condition for LI. In Chichewa, intransitive verbs like *-li* (to be), *khala* (to sit) and *bwera* (to come) allow LI, while transitive verbs like *peza* (to find), *thamangitsa* (to pursue) and *tumiza* (to send) do not. Although LI practically only happens with intransitive verbs, not all intransitive verbs allow LI, according to Bresnan (1994).

LI can occur in cases where the subject can be interpreted as the theme of location, change of location or direction expressed by the locative. Thus, intransitive verbs in Chichewa like *-li* (to be), *khala* (to sit) and *bwera* (to come) satisfy the characteristic of having a locative subject. For the author, Chichewa provides evidence that the locative is the grammatical subject in LI. The author points out three properties that commonly lead to the validation of the grammatical subject in languages: the agreement between the subject and the verb, the control of the subject, and the raising of the subject. She explains that finite verbs in Chichewa have a mandatory “subject prefix”, which agrees in gender, number, and person with the grammatical subject. In cases of locative inversion, agreement with the locative is mandatory, as shown in the following data¹.

¹ Regarding the glosses, we refer the reader to the basic clarifications in Bresnan (1994: 76-77): “the locative class markers 16, 17, and 18 of nouns (but not verbs and adjectives) are glossed as particles rather than prefixes in this and subsequent examples, in accordance with the results of Bresnan and Mchombo 1993. (...). Chichewa has eighteen noun classes, which are denoted by Arabic numerals in the glosses, including a class 1A; Roman numerals are used for first and second person; and the following

(5) Chichewa (BRESNAN, 1994, p. 93)

- a. **Pa** m-sikä-pa pá bád-w-a nkhonya.
 16 3-market-16.this 16.SUBJ.IM.FUT-be.born-FV10.fist
 ‘At this Market a fight is going to break out.’
- b. **Ku** mu-dzi ku na-bwér-á a-lëndo.
 17 3-village 17.SUBJ-REC.PST-come-FV 2-visitor
 ‘To the village came visitors.’
- c. **M** nkhalāngo mw a-khal-á mí-kāngo.
 18 9.forest 18.SUBJ-PRF-remain-FV 4-lion
 ‘In the forest have remained lions.’

The lack of any of the three prefixes in those examples would make the sentences ungrammatical. Those prefixes are indistinguishable from the other subject-verb agreement prefixes. Therefore, LI in Chichewa satisfies the first of the generalizations for the grammatical subject: the finite verb necessarily agrees with the subject. As for control of attributive VPs, Chichewa has infinitive verbal forms that can be used as an attributive modifier of NPs, like the participle in English.

(6) Chichewa (BRESNAN, 1994, p. 93)

- a. m-sodzi [VP w-ó-ík-á nsómbá pa m-pando]_{VP}
 1-fisherman 1-ASC.INF-put-FV 10.fish 16 3-chair
 ‘A fisherman putting fish on a chair.’
- b. nsómbá [VP z-ó-ík-ídw-á pá m-pando]_{VP}
 0.fish 1 0-ASC.INF-put-PASS-FV 16 3-chair
 ‘Fish being put on a chair.’

The author notes that the agent role is assigned to DP’s head in (6a), and the patient role, to DP’s head in (6b), thus the agent is the subject of the verb in the active form, and the patient is the subject of the verb in the passive form. Furthermore, the verb carries a prefix that agrees with the controller of the gender class: it agrees with ‘fisherman’ in (6a) and with ‘fish’ in (6b). When the locative inversion occurs in a NP with an attributive VP, the locative role can be assigned to the controller and the adnominal prefix shows locative agreement, as in (7).

abbreviations are also used: NEG= negative, SG= singular, SUBJ= subject, OBJ=object, PROG=progressive, PRF=present perfect, REC PST=recent past, IM FUT= immediate future, PRS HAB=present habitual, APPL=applicative, PASS= passive, FV=final vowel, POSS= possessive pronoun, ASC= associative, and INF = infinitive.”

- (7) Chichewa (BRESNAN, 1994, p. 94)

m-nkhalangó [vp m-ó-khál-á mi-kângo]vp
 18-9.forest 18-ASC.INF-live-FV 4-lion

‘In the forest where there lions live.’

Thus, for the author, those facts indicate that LI satisfies the second grammatical generalization about the subject in Chichewa: the controlled argument of the attributive VP is the subject. The third and last generalization concerns the raising of the subject. In Chichewa, the locative can be raised with a certain class of raising verbs, demonstrating that it is the grammatical subject.

- (8) Chichewa (BRESNAN, 1994, p. 95)

Ku mu-dzi kw-a-yamba — ku-gwá mvûla.
 17 3-village 17.SUBJ-PRF-start — INF-fall 9.rain

‘It has started raining in the village.’ Lit. ‘At the village has started to fall rain.’

Thus, Bresnan (1994) argues that the evidence of agreement, control and raising converge to identify, in LI, the locative as the grammatical subject in Chichewa.

3 Bare-NP adverbs and case assignment

Larson (1985) examines noun phrases that can function as adverbial modifiers, without accompanying prepositions or any other indicator of adjunct status. These are called *bare-NP adverbs*, as illustrated in the following examples:

- (9) a. I saw John [NP that day] or [NP someplace you’d never guess].
 b. John was headed [NP that day].
 c. Max pronounced my name [NP every way imaginable]. (LARSON, 1985, p. 595)

According to the author, modern English exhibits *bare-NP adverbs* in a variety of semantic functions, including temporal and locative modifiers. Many NPs that refer to a point or period of time can function as temporal modifiers, for example: NPs headed by common names that refer to units in calendars such as days, months, and years; NPs that refer to an annual calendar interval or that function as proper names for periods of time; NPs headed by the common noun *time*; the temporal form *then* and the deictic *now*, *yesterday*, *today* and *tomorrow*

(LARSON, 1985, p. 596). On the surface, *bare-NP* adverbs take the form of a simple NP and may be accompanied by several determinants identical to those found in canonical cases of NPs arguments: *some, every, a, the*, etc. They can also be modified by restrictive relative clauses. Furthermore, *bare-NP adverbs* can occupy positions normally filled only by NPs:

- (10) a. Every morning's lecture. (LARSON, 1985, p. 598)
 b. The lecture every morning.
- (11) a. Yesterday's refusal. (LARSON, 1985, p. 598)
 b. The refusal yesterday.

Since the genitive specifier position is only available for NPs, *every morning* and *yesterday* are, for the author, NPs in adverbial usage. Seeking to improve the analysis of *bare-NP adverbs*, the author recalls a characteristic of NPs: they need Case. He then proceeds to consider the sentence below:

- (12) John hit the ball over the fence [_{NP} that day]. (LARSON, 1985, p. 606)

According to the author, *that day* has no potential Case attributor, so the expectation is for the sentence to be ungrammatical; however, it is well formed. He then suggests that *bare-NP adverbs* have a special way of receiving Case: assignment occurs via a special feature [+F], which is carried by these names. This feature is inherited by any NP that has such N as its head, and assigns Oblique Case to the NP (LARSON, 1985, p. 607). This NP Case marking ability [+F] allows these NPs to take Case and satisfy the Case Filter in the absence of an external Case attributor, such as a verb or a preposition. So, what distinguishes NPs that can function as *bare-NP adverbs* is an intrinsic feature, with lexically determined Case marking.

For the author, the relative freedom of syntactic position observed in relation to *bare-NP adverbs* can be attributed to Case Theory. Inherently receiving Case, or not needing Case, these elements are not bound to occur adjacent to some [-N] or [+Tense] ruler; therefore, within the VP, for example, *bare-NP adverbs* can be freely reordered with other independent Case categories, such as PPs (LARSON, 1985, p. 607). The author also considers cases involving the following structures:

- (13) a. That day passed very quickly. (LARSON, 1985, p. 609)
 b. Few places with a view could be found.
 c. We spent that day in New York.
 d. We visited few places with a view.

According to Larson (1985), the NPs [+F] *that day* and *few places with a view* occur as subjects of sentences in (13a-b) and as verbal objects in (13c-d). Given the hypothesis that NPs [+F] inherently receive Oblique Case and given that Nominative Case and Accusative Case are assigned to the subject and object positions, respectively, the expectation is that these examples would be ungrammatical due to a Case clash; however, these sentences are well formed. So, the author proposes that the assignment of Case by feature [+F] is optional. This implies that although the feature [+F] is present in the subject and object NPs in (13), the Oblique Case needs not be assigned and, therefore, there is no Case clash. In summary, Larson (1985) concludes that NPs that carry feature [+F] occur in argument position, so [+F] cannot itself be considered a Case marker: if so, there would be a Case clash, once NP [+F] occurs in the position in which the Nominative Case is assigned.

Although Larson's study is very interesting empirically and theoretically, we consider that the solution of the Case conflict in terms of optionality in the occurrence of the feature [+F] faces a current theoretical problem. How would it be possible to anticipate the position in which *bare-NP adverbs* occur, to define the presence or absence of the feature [+F]? Given that this possibility does not currently exist, we will reinterpret this analysis and look for a workaround for this problem with BP data in section 5.

4 LI and subject-topics in BP

Pilati (2006) observes that VS clauses are restricted in distribution in BP. She argues that this order should be analyzed as a type of LI, in view of the following syntactic characteristics: (i) they are frequent with unaccusative verbs (14a-b); (ii) they generally manifest locative or temporal PPs on the left, with (locative or temporal) deictic interpretation (14c-d), a position that, hypothetically, can be null and anaphorically connected, also occurring in the presence of discursive elements and

focus operators, as *só* (only), *também* (too); (iii) present restrictions to transitive verbs, occurring in concomitant narratives² and idioms (14e-f):

(14) Brazilian Portuguese

a. Aqui chegaram as cartas. (PILATI, 2006, p. 195)

Here arrived the letters.

‘Letters arrived here.’

b.*Aqui avermelhou o urubu. (PILATI, 2006, p. 195)

Here turned.red the vulture

c. Neste hotel dormiu a Maria. (PILATI, 2006, p. 139)

In-this hotel slept the Mary

‘Mary slept in this hotel.’

d.*Em hotéis dormiu a Maria (PILATI, 2006, p. 139)

In hotels slept the Mary

e. Arriscou o chute Diego Tardelli. (PILATI, 2006, p. 20)

Risked the kick Diego Tardelli

‘Diego Tardelli risked the kick.’

f. Hoje tomou posse o novo ministro da Cultura. (PILATI, 2006, p. 20)

Today took office the new Minister of Culture

‘Today the new Minister of Culture took office.’

Based on Bresnan’s (1994) observation that discursive factors interfere with the distribution of LI in Chichewa (and English), Pilati (2006) demonstrates that these factors are present in the VS structures of BP, due to the requirement of a (locative and temporal) deictic element, or of null deictic recovered anaphorically. The author also observes that the presence of (locative and temporal) deictic can be inferred in structures with VS order of BP in which the deictic does not occur on the left, but it is possible to assign a deictic interpretation, as represented by the contrast between the sentences (15a) and (15b):

² According to Pilati (2006, p. 200, translated): “the discursive context of these clauses makes it evident that the interpretation of the locative PP is deictic, as these clauses describe events that occur almost concomitantly with the narration, and it is not necessary for the speaker to verbally express neither the place (which is the soccer field) nor the moment in which the events are taking place (since the narration is concomitant).”

(15) Brazilian Portuguese

a. Morreu Fellini. (PILATI, 2006, p. 199)

Died *Fellini*

Interpretation: Fellini just died (I just heard that Fellini died).

b. Fellini morreu. (PILATI, 2006, p. 199)

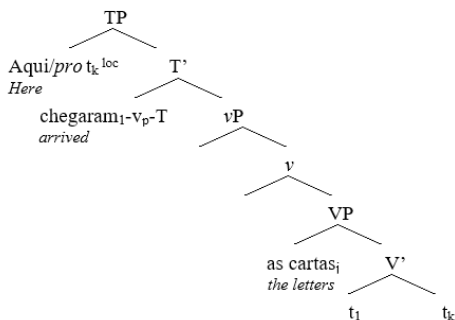
Fellini *died*

Interpretation: Fellini died (some time ago).

This deictic interpretation, related to the moment of enunciation or to a specific place whose reference is shared by the interlocutors, is also found in the VOS structures, in which the so-called concomitant narratives occur, as in (14c). In the analysis of V(O)S structures in BP as cases of LI, Pilati (2006) postulates that the (locative or temporal) deictic element is linked to the subject position. Therefore, the author proposes the following structures, for the V(O)S order in BP:

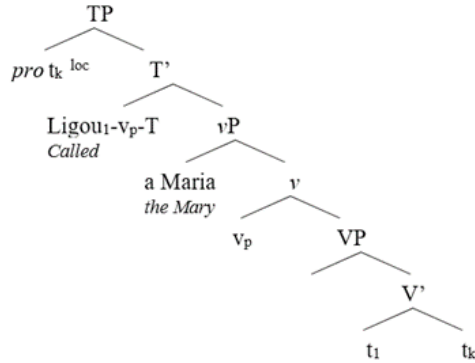
a) Possibilities of ordering with unaccusative verbs, e.g.: *Aqui chegaram as cartas/pro^{loc}* *Chegaram as cartas* (‘Here/*pro^{loc}* arrived the letters’).

(16) Brazilian Portuguese (PILATI, 2006, p. 213)



b) Derivation with unergative verbs, e.g.: *Hoje ligou a Maria/pro^{loc}* *Ligou a Maria* (‘Today/*pro^{loc}* called Maria’).

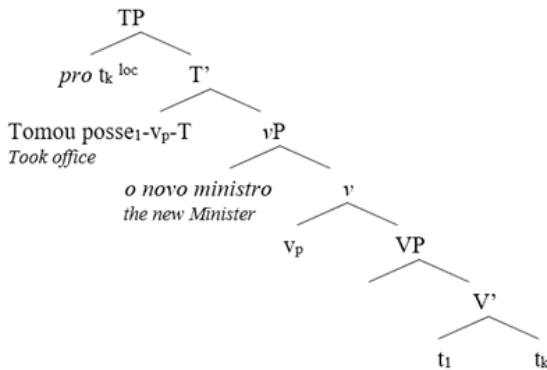
(17) Brazilian Portuguese (PILATI, 2006, p. 213)



Pilati (2006) points out that the adverb ‘*Hoje*’ would occupy the same position as *pro* if it was realized in (17). The same statement is valid for the data in (18).

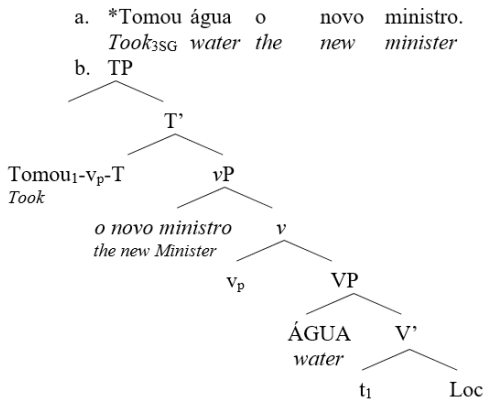
c) Possibilities of ordination with transitive verbs in coinciding narratives or in sentences with light verbs, e.g.: *Hoje tomou posse o novo ministro/pro^{loc} Tomou posse o novo ministro* (‘Today/*pro^{loc}* the new minister took office’).

(18) Brazilian Portuguese (PILATI, 2006, p. 214)



According to Pilati (2006), other transitive verbs cannot present LI for a syntactic reason. Their object receives Accusative Case within the vP, and blocks the possible *probe-goal* relationship between T and Loc, as in (19).

(19) Brazilian Portuguese (PILATI, 2006, p. 215-216)



Pilati (2006) explains that (19) is ungrammatical because there must not be an inactive syntactic object, which has already valued its features, between the probe and the goal. The presence of this inactive element blocks the combination between the probe and the goal in terms of Defective Intervention. She concludes her study noting that the hypothesis that BP VS structures license a null pronoun with locative reference in the pre-verbal position also accounts for other aspects of BP grammar, such as the weakening of the verbal agreement system (cf. DUARTE, 1993) and the appearance of structures referred to as subject-topic (cf. PONTES, 1987).

The hypothesis of the presence of locative PP in the subject position in BP was investigated by Avelar and Cyrino (2008, 2009), Avelar (2009) and Avelar, Cyrino and Galves (2009). According to these authors, LI in BP can be a contribution from languages of the Bantu family (see the characteristics mentioned in the analysis by Bresnan (1994) for Chichewa), given the situation of linguistic contact in the colonization period in Brazil. The authors analyze structures like (20):

(20) Brazilian Portuguese (AVELAR; CYRINO; GALVES, 2009, p. 207)

a. Naquela	loja	vende	muitos	livros.
<i>in-that</i>	<i>shop</i>	<i>sell3SG</i>	<i>many</i>	<i>books</i>

European Portuguese/* Brazilian Portuguese: 'In that shop, (s)he sells books.'

*EP/BP: 'That shop sells many books.'

b. Naquela	fazenda	planta	todo	tipo	de	legume.
<i>in-that</i>	<i>farm</i>	<i>plant3SG</i>	<i>all</i>	<i>kind</i>	<i>of</i>	<i>vegetable</i>

EP/*BP: 'In that farm, (s)he plants all kinds of vegetables.'

*EP/BP: 'In that farm, one plants all kinds of vegetables.' Or: 'In that farm, all kinds of vegetables are planted.'

The authors' initial hypothesis is that, in the absence of the external argument, the locative PP can check the EPP in [Spec-TP]. The authors explain that the locative inversion covers the cases in which locative constituents, which are commonly taken as non-argumental, occur in the position identified as the grammatical position of the subject. However, they emphasize that natural languages behave heterogeneously regarding the properties that involve the phenomenon, distinguishing themselves in terms of agreement properties and the specificity of the argument structure that licenses the inversion.

Analyzing the properties of agreement, the authors point out the possibility of a verb being able to agree with the locative constituent in a pre-verbal position in Chichewa³:

(21) Chichewa (AVELAR; CYRINO, 2009, p. 59)

a. Ku-mu-dzi	ku-na-bwér-á	a-lendô-wo
<i>17-3-village</i>	<i>17SB-REC PST-come-IND</i>	<i>2-visitor-2-those</i>

'To the village came those visitors.'

b. M-mi-têngo	[_{VP} mw-a-khal-a	a-nyāni.]
<i>18-4-tree</i>	<i>18SB-PERF-sit-IND</i>	<i>2-baboons</i>

'In the trees are sitting baboons.'

Regarding the argument structure that licenses locative inversion, Avelar and Cyrino (2009) point out the existence of locative inversion in a wider variety of constructions, as exemplified by Kinyarwanda (Bantu),

³ Data (21) is from Bresnan and Kanerva (1989, p. 1-3, examples (1b) and (4a)).

which allow locative inversion with unergative and transitive verbs in so far as agent and theme do not occur together in the same sentence⁴:

- (22) Kinyarwanda (AVELAR; CYRINO, 2009, p. 60)
- | | | | | |
|-----------|---------------|---------------------------|----------------|--------------|
| a. kw' | iisôko | ha-Ø-guz-w-e | ibi-intu | bi-taandátu. |
| <i>in</i> | <i>market</i> | <i>16-PST-buy-PAS-PRF</i> | <i>8-thing</i> | <i>8-six</i> |
- 'At the market were bought six things.'
- | | | | |
|-----------|-------------|---------------------------|----------------|
| b. Mu | cyumba | ha-Ø-rîi-r-iye | umwáana. |
| <i>In</i> | <i>room</i> | <i>16-PST-eat-APL-PRF</i> | <i>1:child</i> |
- 'In the room, there ate a child.'

Comparing the argument structures that license LI in Bantu languages, Avelar and Cyrino (2009) demonstrate that BP allows LI in constructions with unergative and transitive verbs, as observed in languages such as Kinyarwanda. Thus, in the authors' analysis, the data below are the result of a derivation in which the locative PP is realized in the grammatical position of the subject, in [Spec-TP]:

- (23) LI with unaccusative verbs in Brazilian Portuguese (AVELAR; CYRINO, 2009, p. 61)
- | | | | | | | |
|---------------|--------------|---------------|-------------|------------------|-------------|----------------|
| Na | casa | da | Maria | chegou | algumas | cartas. |
| <i>in-the</i> | <i>house</i> | <i>of-the</i> | <i>Mary</i> | <i>arrive3SG</i> | <i>some</i> | <i>letters</i> |
- 'Some letters arrived at Maria's house.'
- (24) LI with unergative verbs in Brazilian Portuguese (AVELAR; CYRINO, 2009, p. 61)
- | | | | | |
|----------------|-------------|-----------------|----------------|---------------|
| Naquele | quarto | dormiu | várias | peessoas. |
| <i>in-that</i> | <i>room</i> | <i>slept3SG</i> | <i>several</i> | <i>people</i> |
- 'In that room several people slept.'
- (25) LI with ergativized transitive verbs in BP (AVELAR; CYRINO, 2009, p. 61)
- | | | | | | | |
|----------------|---------------------|----------------|---------------|-----------|------------|------------|
| Naquele | bairro | aluga | casa | de | todos | os |
| <i>in-that</i> | <i>neighborhood</i> | <i>rent3SG</i> | <i>houses</i> | <i>at</i> | <i>all</i> | <i>the</i> |
- preços.
prices
- 'In that neighborhood houses are rented at all prices.'
- (26) LI with transitive and unergative verbs without theme and/or agent in BP (AVELAR; CYRINO, 2009, p. 61)

⁴ Data (22a) and (22b) are respectively from Salzmann (2004, p. 51), and Polinsky (1992, p. 298).

Nas cidades do interior não sequestra tanto como nas
 in-the cities of-the interior not kidnap3SG as much
 grandes capitais.

big capitals

‘In small towns there are not as many kidnappings as in the big capitals.’

Avelar and Cyrino (2009) present two tests demonstrating that the locative PP is in Spec, TP in (23) to (26). With the first test, the authors demonstrate the obligatory nature of the locative PP when the subject is post-verbal, in contrast to the optionality of this constituent when the argument subject is pre-verbal⁵:

(27) Brazilian Portuguese (AVELAR; CYRINO, 2009, p. 62)

a. (*Naquele quarto*) várias pessoas dormiram.
 (*in-that room*) several people slept_{3PL}

‘In that room several people slept.’

b. *(*Naquele quarto*) dormiu várias pessoas.
 (*in-that room*) slept_{3SG} several people

Avelar and Cyrino (2009) conclude that the obligatory nature of the locative PP can be adequately explained if we assume that the locative satisfies the EPP condition in contexts where the subject is post-verbal, as in (27b). This situation can also be compared to what occurs in (28a-b), where the presence of the locative licenses the structure in which the subject is not grammatically present.

(28) Brazilian Portuguese (AVELAR; CYRINO, 2009, p. 62)

a. (*Naquela loja*) todos os tipos de livro vendem.
 (*in-that store*) all the kinds of books sell_{3PL}

‘All kinds of books are sold in that store’

b. *(*Naquela loja*) vende todos os tipos de livro.
 (*in-that store*) sell3SG all the kinds of books

Therefore, the authors emphasize the need to assume that the locative is in [Spec-TP], or the obligation of the locative PP in terms of EPP satisfaction cannot be addressed. In a note, the authors note that locative PPs can save sentences also if they are in final position:

⁵ Note that the contrast in (27) and (28) is originally cited in Pilati (2006) to indicate the restrictions on the occurrence of the VS order in BP, as well as to support the hypothesis that such constructions in BP are realized in a locative inversion configuration.

(29) Brazilian Portuguese (AVELAR; CYRINO, 2009, p. 62)

a. Vende muitas coisas naquela loja.
 sell_{3SG} a-lot-of stuff in-that store
 sell_{3SG} a-lot-of stuff in-that store
 ‘A lot of stuff is sold in that store’

In contrast with the data above, the authors analyze the following data:

(30) Brazilian Portuguese (AVELAR; CYRINO, 2009, p. 62)

a. (ele) não quis almoçar hoje, o Roberto_i
 (he) didn't want to-have-lunch today the Roberto_i
 ‘Roberto didn’t want to have lunch today’

b. (ela) vai comprar um carro novo, a Maria_i
 (she) is-going to-buy a car new the Maria_i
 ‘Maria is going to buy a new car’

c. (elas) comeram a comida toda, as crianças_i
 (they) ate the food all the children_i
 ‘The children ate all the food’

They argue that, in (30), the data show that subjects can be licensed in a final position if they are interpreted as a topic by a personal pronoun prepended to the verb. Comparing with the data in (30), the authors note that the locatives in final position can also be in co-occurrence with pronominal adverbial elements prepended to the verb (31), which reinforces the authors’ hypothesis that locative PPs can guarantee the acceptability of the sentence when they enter the position typically occupied by the argument subject.

(31) Brazilian Portuguese (AVELAR; CYRINO, 2009, p. 62)

a. (lá) vende muitas calças, naquela loja_i
 (there) sell_{3SG} a-lot-of pants in-that store_i
 ‘That store sells a lot of pants.’

b. (aí) grava todo tipo de filme, nesse
 (there) record_{3SG} all kinds of movies on-this
 meu DVD_i
 my DVD_i
 This DVD set of mine can record all kinds of movies’

c. (*lá*) trabalha vários amigos meus,
 (*there*) work_{3SG} several friends of-mine,
 naquela loja_i
 in-that store

‘Many friends of mine work at that store.’

This fact reveals the similarity between locatives and nominal subjects regarding the obligation to be moved to [Spec-TP] in raising constructions. Based on this fact, the authors claim that the locative PP can satisfy the EPP condition also in these structures:

(32) Brazilian Portuguese (AVELAR; CYRINO, 2009, p. 63)

a. *Parece *na casa da Maria* chegar muitas cartas.
*seem*_{3SG} *at-the house of-the Maria* *to-arrive many letters*
 b. *Na casa da Maria* parece chegar muitas cartas.
In-the house of-the Maria *seem*_{3SG} *to-arrive many letters*

‘It seems that a lot of letters arrive at Maria’s house.’

(33) Brazilian Portuguese (AVELAR; CYRINO, 2009, p. 63)

a. *Parece *naquele shopping* trabalhar muita gente.
*seem*_{3SG} *in-that mall to-work a-lot-of people*
 b. *Naquele shopping* parece trabalhar muita gente.
*in-that mall seem*_{3SG} *to-work a-lot-of people*

‘It seems that a lot of people work at that mall.’

The authors also analyze the fact that the locative PP does not trigger agreement in the subject position. In Avelar and Cyrino (2009), they highlight that popular BP has an optional subject-verb agreement:

(34) Brazilian Portuguese (AVELAR; CYRINO, 2009, p. 70)

Os menino *comeu/comeram* o bolo.
*The*_{3PL} *boy*_{3SG} *ate*_{3SG/ate}_{3PL} *the cake*

‘The boys ate the cake.’

This variable agreement, for the authors, is related to the possibility that category T can be licensed without ϕ -features. In other words, in BP, constituents whose nuclei are ϕ -featureless, such as prepositions, can be licensed in the subject position since subject-verb agreement is not mandatory.

We consider valid – and specific to BP – the data that illustrate the occurrence of the locative PP in the first position (in relation to

the verb), as presented by the authors. We also consider plausible the hypothesis that, in those cases, the locative PP occupies the position of subject, just like in LI structures, for the cases of VS, as postulated by Pilati (2006). However, the hypothesis that the variable agreement in BP indicates that T can be licensed independently of phi-features does not account for the licensing of a locative PP in the subject position. We will return to this question.

Avelar (2009) advances on the analysis of Avelar and Cyrino (2009), proposing that the inflectional paradigm of BP authorizes agreement relationships between a verb and a prepositioned locative. In addition, the author brings the argument of co-indexing between subjects of coordinated clauses.

(35) Brazilian Portuguese (AVELAR, 2009, p. 236)

[Muita gente] _i	trabalha	naquela	fábrica	e	cvi	mora
[Many people] _i	work	in-that	factory	and	cvi	live
do	outro	lado	da	cidade.		
on-the	other	side	of-the	town		

‘Many people work at that factory and live on the other side of town.’

In (36) it can be noted that if a locative PP is allegedly in Spec, TP, the empty category in the coordinated sentence is co-indexed to it:

(36) Brazilian Portuguese (AVELAR, 2009, p. 236)

[Naquela fábrica] _i	trabalha	muita	gente	e	ainda	assim	
[In-that factory] _i	work _{3SG}	a-lot-of	people	and	even	so	
cv _i	vai	contratar	mais	cem	funcionários	até	o
cv _i	will _{3SG}	hire	another	hundred	employees	by	the
final	do	ano.					
end	of-the	year					

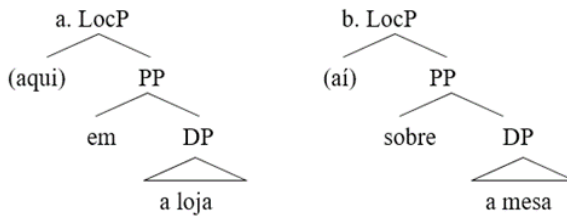
‘A lot of people work at that factory, and it will still hire another hundred employees by the end of the year.’

For the author, the possibility of co-indexing in (36) is another argument that demonstrates that the locative PP is in subject position.⁶ Avelar (2009) treats locative PPs as projections whose nucleus is an adverbial pronoun (like *aqui*/here, *ai*/there and *lá*/there) that can be

⁶ Likewise, the test of the anaphoric resumption of the null subject of the coordinated clause, by the (expressed) subject of the first clause, was used in Pilati (2006) to postulate the realization of the locative PP (in the VS structure) in the subject position.

realized phonologically or not. These pronouns, referred to as category Loc, head a LocP (*Locative Phrase*), and the projection of the locative PP is treated as a complement of Loc, assuming the configurations in (38) for the structure in (37):

- (37) Brazilian Portuguese
 a. (aqui) na loja
 (here) in-the store
 b. (aí) sobre a mesa
 (there) on the table
- (38) Brazilian Portuguese (AVELAR, 2009, p. 238)



With this argument, the author explains the grammatical requirement that authorizes the occurrence of the locative PP in [Spec-TP], as it gives it a nominal status, since the core of the locative phrase in question is not the introductory preposition of the locative PP, but an adverbial pronoun, which, realized phonologically or not, introduces the preposition. Since the adverbial pronoun is a nominal category, the author points out that, in the subject position, it can be realized alone (39a), co-occur with the locative PP (39b) or be a null category (39c):

- (39) Brazilian Portuguese (AVELAR, 2009, p. 241)
- a. *Lá* vende muitos livros.
there sell_{3SG} many books
 ‘There many books are sold.’
- b. *Lá no shopping* vende muitos livros.
there in-the mall sell_{3SG} many books
 ‘There in the mall many books are sold.’
- c. *No shopping* vende muitos livros.
in mall sell_{3SG} many books
 ‘In the mall many books are sold.’

Assuming that it is a nominal category projection, the locative phrase (LocP/PP_{LOC}) can occupy the subject position and the ϕ -features of T can agree with the LocP/PP_{LOC}, since any pronominal form must be able to trigger agreement.

Avelar (2009) points out that cases in which the LocP/PP_{LOC} occurs at the end of the sentence could be a counterargument to the idea that the phrase in question occupies the position of subject. The data below illustrate the situation:

(40) Brazilian Portuguese (AVELAR, 2009, p. 241)

a. Vende muitos livros (lá) naquela shopping.
sell_{3SG} many books (there) in-that mall

‘Many books are sold (there) at that mall.’

b. Dorme criança (aqui) nesse quarto.
sleep_{3SG} child (here) in-this room

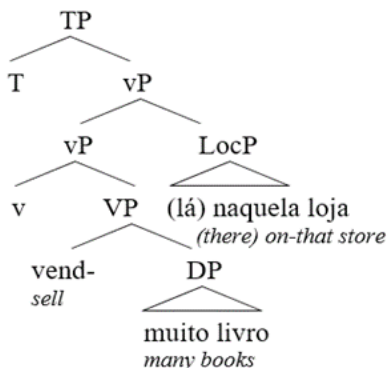
‘Children sleep (here) in this room.’

c. Planta todos os tipos de legume (aí) nessa fazenda.
plant_{3SG} all the kinds of vegetables (there) in-this farm

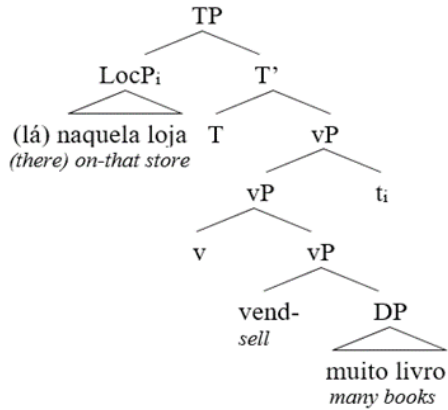
‘All kinds of vegetables are planted (there) on this farm.’

The author explains that, since *Agree* is established at a distance (the interaction between the ϕ -features of T and those of DP happens before the subject is moved to [Spec-TP]), T establishes agreement with the LocP/PP_{LOC} when it is in situ. By presenting an inherently adverbial nature, typical of constituents in an adjunct configuration, the locative is, in principle, an adjunct of VP (41); moreover, the LocP/PP_{loc} can be moved to Spec-TP and receives nominative Case, as in (42).

(41) Brazilian Portuguese (AVELAR, 2009, p. 242)



(42) Brazilian Portuguese (AVELAR, 2009, p. 242)



Avelar's (2009) proposal is not about the contrast between the presence or absence of the preposition in the phrases in question in the subject position, "but between a noun phrase that takes place in the form of a DP (*aquela loja* – 'that store'), and another that is realized as the projection of a deictic adverbial pronoun without phonological realization", the LocP (AVELAR, 2009, p. 245). In this perspective, there is apparently a PP. However, for the syntactic computation, the phrase is introduced by an adverbial pronoun, which can be null.

Avelar (2009) further emphasizes that the inflectional paradigm of BP currently presents the ϕ -features of T in a defective version (T_{DEF}), as the number feature is absent.⁷ Thus, we have a mark for the first person (*eu canto*/'I sing') and another for the others [*tu/você/ele/nós(a gente)/vocês/eles canta* ('you/he/we/you/they sing')]. This characteristic of BP inflectional paradigm can, in his analysis, interfere with the type of category that interacts with the ϕ -features of T in the *Agree* operation. For instance, a category that only has a person marking (but not number) as is the case with adverbial pronouns, can interact with T_{DEF} .

Avelar's (2009) analysis explains the occurrence of locative PPs in the subject position in BP, but, as the author acknowledges, it does not address the cases referred to in the studies by Pontes (1986, 1987), in which the locative PP loses its preposition and occurs as a locative DP, preposed to the verb, being able to agree with the verb:

⁷ The proposal and the first analysis of some consequences of T in a defective version (T_{DEF}) in BP grammar is attributed by Avelar to Ferreira (2000).

(43) Brazilian Portuguese (PONTES, 1987, p. 36)

a. *Essa casa bate sol.*

This house hit_{3SG} sun

‘This house gets a lot of sunlight.’

b. *Essas casas batem sol.*

These houses hit_{3PL} sun

‘These houses get a lot of sunlight.’

In a study that became seminal for the characterization of BP grammar, Pontes (1986) demonstrates that one of the strategies used to fill in the subject in this language is the raising of locative adverbs and/or locative phrases to the subject position. The author analyzes, from a functionalist perspective, structures such as:

(44) Brazilian Portuguese (PONTES, 1986, p. 17-18)

a. *As gavetas não cabem mais nada.*

The drawers not fit_{3PL} anything else

‘The drawers don’t fit anything else.’

b. *Essa casa bate bastante sol.*

This house hit_{3SG} a-lot-of sun

‘This house gets a lot of sunlight.’

The author observes that the DPs in (44) have a prepositioned version, as in (45):

(45) Brazilian Portuguese (PONTES, 1986, p. 17-18)

a. *Nas gavetas não cabe mais nada.*

in-the drawers not fit_{3SG} nothing else

‘In the drawers nothing else can fit.’

b. *Nessa casa bate sol.*

in-that house hit_{3SG} sun

‘This house gets a lot of sunlight.’

For Pontes (1987), sentences in (44) and (45) exemplify a syntactic alternation in which the same phrase is expressed with or without a preposition. The author postulates that such constructions manifest different interpretations. In the prepositioned version, the degree of impersonality is greater, because, regarding non-prepositioned constructions, “the native

speaker feels the first DP as also somehow responsible for what the verb conveys. ‘*Essa casa*’ is a well-built house, and this is one reason why it receives so much sun” (PONTES, 1987, p. 88).

The author observed in a previous study (PONTES, 1986) that verbal agreement and the position before the verb are the most striking characteristics of the subject in BP, which would support the assertion that the phrases highlighted in (46a-b) are the subjects of sentences. For Pontes (1986, p. 18), the verb in sentences like (46a-b) agrees with the first DP and not the second, which makes it impossible to prepose the second DP to the verb, as indicated by the ungrammaticality of (46c-d):

(46) Brazilian Portuguese (PONTES, 1986, p. 18)

- a. *A Belina* cabe 60l de gasolina.
the Belina hold_{3SG} 60l of gasoline
 ‘*Belina holds 60l of gasoline.*’
- b. *Esse carro* cabe 60l de gasolina.
this car hold_{3SG} 60l of gasoline
 ‘*This car holds 60l of gasoline.*’
- c. **A Belina* cabem 60l de gasolina.
the Belina hold_{3PL} 60l of gasoline
- d. **Esse carro* 60l de gasolina cabem.
this car 60l of gasoline hold_{3PL}

As Pontes (1986, p. 19) observes, the occurrence of a subject in Portuguese that cannot trigger agreement would be strange. Hence, we could not think of a postponed subject for structures with locative DPs, as in (46a-b).

Galves (1998), in a generative framework, seeks to characterize the properties of structures such as (46a-b) in BP. The author also presents evidence that locative DPs preposed to the verb perform the function of subject in the sentence.

(47) Brazilian Portuguese (GALVES, 1998, p. 21)

- a. *Bate* muito sol *(n)esta casa.
*hit*_{3SG} a-lot-of sun *(in)-this house
 ‘*This house gets a lot of sunlight.*’
- b. *Nesta casa,* bate muito sol.

- in-this house, hit_{3SG} a-lot-of sun*
 c. Esta casa bate muito sol.
this house hit_{3SG} a lot of sun
 ‘This house gets a lot of sunlight.’
 d. Esta casa, bate muito sol nela.
this house hit_{3SG} a-lot-of sun in-it
 ‘This house gets a lot of sunlight.’

According to the author, in (47a) we have an illustration of a projection of the argument structure, in which the preposition that marks the locative cannot be omitted. The data in (47b-d) illustrate other topicalization structures the BP grammar allows. In (47b), *esta casa* is the complement of a preposition, whereas in (47c-d) we have it is a DP preposed to the verb. (47d) illustrates a DP preposed to the verb that appear in the sentence markedly topicalized by the presence of a resumptive pronoun. In Galves’s analysis, she observes that there is a complementary distribution between the presence of the resumptive pronoun and the agreement between the prefixed DP and the verb:

- (48) Brazilian Portuguese (GALVES, 1998, p. 21)
 a. Estas casas batem muito sol.
these houses hit_{3PL} a-lot-of sun
 ‘These houses get a lot of sunlight.’
 b. *Estas casas batem muito sol nelas.
these houses hit_{3PL} a-lot-of sun on-them

The data in (48) demonstrate that the preposed DP must agree with the verb or be in co-occurrence with a resumptive pronoun. There is also another difference between sentences with resumptive pronouns and without it, which concerns the possibility of agreement between the verb and the postponed DP (GALVES, 1998, p. 21-22):

- (49) Brazilian Portuguese (GALVES, 1998, p. 21-22)
 a. Este carro, cabem muitas pessoas nele.
this car fit3PL a-lot-of people in-it
 ‘This car fits a lot of people.’
 b. ?? Este carro cabem muitas pessoas.
this car fit3PL a-lot-of people
 ‘This car fits a lot of people.’

The ungrammaticality of (49b) can be attributed, according to the author, to the fact that there is no way of licensing the preposed DP: neither agreement nor pronominal resumption. In sum, Galves (1998) defends the hypothesis that locative DPs placed before the verb, without pronominal resumption, are legitimized as a subject in the sentence, and refers to these constructions as ‘subject-topics’. The author thus distinguishes subject-topic sentences from those with pronominal resumption pointing to a lexical-semantic property of the verbs and arguments involved, as there is a restriction on the subject topic construction that is not verified in the construction with pronominal resumption:

(50) Brazilian Portuguese (GALVES, 1998, p. 22)

a. Essa estante, o João põe muita coisa nela.
 this shelf the João puts a-lot-of stuff on-it
 ‘João puts a lot of stuff on this bookcase.’

b. ?? Essa estante o João põe muita coisa.
 this shelf the João puts a-lot-of stuff
 ‘João puts a lot of stuff on this bookcase.’

Since the transitive sentence in (50b) is ill-formed, the hypothesis is that there cannot be a projection of the external argument of the verb in the construction of the subject-topic. According to the Galves (1998), the properties that summarize the subject-topic constructions are as follows: a) there is no resumptive pronoun resuming the preposed NP; b) there is no agreement between the verb and the postponed DP; c) the external argument of the verb is absent.

Munhoz (2011) also discusses the issue of locative DPs in the so-called subject-topic structure. The author assumes the hypothesis of Galves (1998) that in subject-topic structures, the external argument of the verb is absent. The author pays attention to the fact that the class of unaccusative verbs is not homogeneous and, based on the studies of Duarte (2003), analyzes the unaccusative verbs with which the occurrence of the locative as a subject-topic is possible.

Duarte (2003) divides unaccusative verbs into the following semantic classes: a) change of state verbs: with an external cause (some of which would participate in the causative alternation), such as *abrir* (to open), *apodrecer* (to rot), *crystalizar* (to crystallize), *derreter* (to melt), *fritar* (to fry) and *rasgar* (to tear); with an internal cause (not causatively

alternating and non-agent), of a physical or psychic reaction, such as *empalidecer* (to turn pale); of emission, like *explodir* (to explode), and of change of state due to the internal cause, like *crescer* (to grow), *florir* (to bloom) and *morrer* (to die); b) movement verbs: denote inherent direction, such as *cair* (to fall), *chegar* (to arrive)/*partir* (to depart), *descer* (to descend)/*subir* (to climb), *entrar* (to enter)/*sair* (to exit) and *ir* (go)/*vir* (to come); c) existence and appearance verbs, such as existential verbs *constar* (to be present at), *existir* (to exist) and *perdurar* (to endure), locative existential verbs, such as *morar* (to dwell), *residir* (to reside) and *viver* (to live), those that denote absence or lack, such as *escassear* (to run out) and *faltar* (to lack), those that denote apparition, such as *aparecer* (to appear), *brotar* (to sprout) and *surgir* (to arise), those of disappearance, such as *desaparecer* (to disappear) and *sumir-se* (to vanish), and event verbs, such as *acontecer* (to happen), *ocorrer* (to occur) and *passar-se* (to come to pass).

According to Munhoz (2011), state change verbs due to an internal cause in locative subject-topic sentences have the particularity of depending on the definiteness of the post-verbal DP, as in (51), in which the presence of the subject-topic depends on the definiteness of the post-verbal DP:

- (51) Brazilian Portuguese (MUNHOZ, 2011, p. 92)

O campinho de futebol cresceu a grama.
the field of soccer grew the grass
 ‘Grass grew on the soccer field.’

Regarding movement verbs, Munhoz (2011) makes a case for their biargumentality, noting that these verbs also form locative subject-topic constructions:

- (52) Brazilian Portuguese (MUNHOZ, 2011, p. 93)

a. Arroz mexicano vai banana frita no meio.
rice Mexican go_{3SG} banana fried in-the middle
 ‘Mexican rice has fried banana in the mix.’
 b. Cada pacote vem quatro figurinhas.
each package come_{3SG} four stickers
 ‘Each packet comes with four stickers.’

However, the author considers that, despite the verbs above being of displacement, the sense of trajectory seems lost. Rather, these sentences focus on the moment of culmination of the event or present an existential meaning:

(53) Brazilian Portuguese (MUNHOZ, 2011, p. 93)

Em cada pacote, há/tem/existem quatro figurinhas.
In each package, (there) are/have/are four stickers
 ‘In each packet, there are four stickers.’

Munhoz (2011) notes that not all movement verbs produce grammatical sentences in locative subject-topic structures. At first, the restriction seems to occur with verbs that have the source-local information (*partir* (depart), *cair* (fall) and *descer* (descend)):

(54) Brazilian Portuguese (MUNHOZ, 2011, p. 93)

a. Dois navios partiram daquele porto.
two ships depart_{3PL} from-that port
 ‘Two ships left that port.’

b. *Aquele porto partiu dois navios.
that port depart_{3SG} two ships

(55) Brazilian Portuguese (MUNHOZ, 2011, p. 93)

a. Uma fruta caiu no chão.
a fruit fell_{3SG} on-the ground
 ‘A fruit fell to the ground.’

b. *O chão caiu uma fruta.
the ground fell_{3SG} a fruit

(56) Brazilian Portuguese (MUNHOZ, 2011, p. 93)

a. Um gatinho desceu daquela árvore.
a kitten came-down_{3SG} from-that tree

‘A kitten came down from that tree.’

b. ?Aquele árvore desceu um gatinho.
that tree came-down_{3SG} a kitten

Locative subject-topics are also possible with existence and appearance verbs:

(57) Brazilian Portuguese (MUNHOZ, 2011, p. 94)

a. Essa pasta consta todos os documentos necessários.
this folder contains_{3SG} all the documents necessary

‘This folder contains all the necessary documents.’

b. Aquela casa ali mora o Fagner.
that house there live3SG the Fagner

‘Fagner lives in that house over there.’

c. Seu cabelo tá faltando queratina.
your hair be3SG missing keratin

‘Your hair is missing keratin.’

d. O quintal lá de casa apareceu um gatinho.
the backyard there of house appeared_{3SG} a kitten

‘A kitten appeared in my house’s backyard.’

e. O Japão quase aconteceu um desastre nuclear um
 the Japan almost happened_{3SG} a nuclear disaster a
 dia desses.
day of-these

‘Japan almost had a nuclear disaster the other day.’

Munhoz (2011) adds to Duarte’s (2003) list the verbs *cabere* (to fit) (corresponds to a stative predicate and expresses the possibility of a theme being in a place) and *bater* (to hit) (expresses the state of a theme element affecting a place). In addition, Munhoz (2011) also notes that the literature has mentioned the existence of biargumental unaccusative verbs, one corresponding to what exists and the other to the place where this entity exists, noting the possibility that unaccusative verbs, crosslinguistically, can be subdivided and take a locative element as an argument that can occupy the subject position. Thus, the locative would be selected by the verb, since the absence of the locative causes ungrammaticality:

(58) Brazilian Portuguese (MUNHOZ; 2011, p. 76)

a. *Bate bastante sol
hits_{3SG} a-lot-of sun

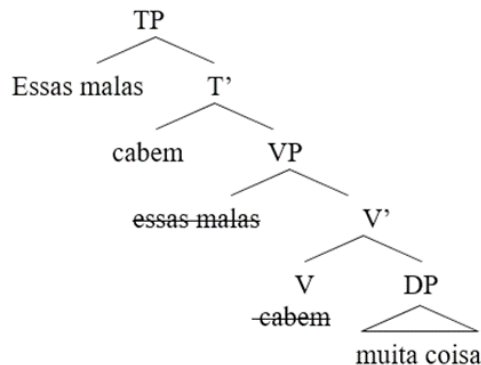
b. *Bastante sol bate.
a-lot-of sun hits_{3SG}

In this sense, Munhoz (2011), based on the findings of different authors that unaccusative verbs have a heterogeneous behavior and that

some unaccusative verbs project a biargumental structure, supports the hypothesis that the locative subject-topic is licensed with biargumental unaccusative verbs, thus defined by selecting two DP's, a theme, and a locative. While the locative argument is raised to the subject position, the *Theme* argument remains internal to the VP.

Munhoz (2011, p. 120) proposes that structures like (59) are the result of derivations like (60):

- (59) Brazilian Portuguese (MUNHOZ, 2011, p. 59)
 Essas malas cabem muita coisa.
these bags fit_{3PL} a-lot-of things
 'These suitcases can fit a lot of stuff.'
- (60) Brazilian Portuguese (MUNHOZ, 2011, p. 59)



The author explains that with the absence of a locative preposition, the phrase '*Essas malas*' structurally values its Case, via *Agree* with T, which results in agreement between the locative argument and the verb – according to the author there is no projection of *vP* in (60) because the verb is unaccusative. Therefore, it does not select an external argument. The hypothesis is that the verb is first concatenated with the theme argument. The second *Merge* will originate the locative in a position from where it will correspond to the closest *target* identified by the *probe* and likely to be raised to the subject position. Thus, the DP '*muita coisa*' does not value its features with T because this category would have already valued its ϕ -features with the closest DP ('*Essas malas*'). For the author, '*muita coisa*' has *default* nominative Case.

As noted by Munhoz (2011), in the genitive and locative subject-topic constructions, there is a distinction regarding the obligatory nature of these arguments, as illustrated in (61a-b) – in which there is a relationship of location (‘sun...(in) the house’) and of possessor-possessed (‘tire...(of) the car’), respectively.

(61) Brazilian Portuguese (MUNHOZ, 2011, p. 62)

a. Essa casa bate sol.
this house hit_{3SG} sun

‘This house gets sunlight.’

b. Meu carro furou o pneu.
my car pierced_{3SG} the tire

‘My car had a flat tire.’

In both constructions the locative and possessor argument fill the subject position, although they have distinct syntactic status: while the suppression of ‘*meu carro*’ in (62) is possible, the suppression of ‘this house’ as in (63) generates ungrammaticality:

(62) Brazilian Portuguese (MUNHOZ, 2011, p. 76)

a. Furou o pneu.
pierced_{3SG} the tire

‘Got a flat tire.’

b. O pneu furou.
the tire pierced_{3SG}

‘The tire got flat.’

(63) Brazilian Portuguese (MUNHOZ, 2011, p. 76)

a. *Bate bastante sol.
hits_{3SG} a-lot-of sun

b. *Bastante sol bate.
a-lot-of sun hits_{3SG}

We observe that this contrast is also manifested in relation to the distribution of adverbs of the ‘*aqui*’ (‘here’) type: while constructions of the locative subject-topic type authorize the realization of the subject position by an adverb of the ‘*aqui*’ (‘here’) type, such a category is not possible in constructions of the genitive subject-topic type, as the data in (64) demonstrate. Our proposal is that the restriction is not only semantic,

but also syntactic, since there is a relationship between the obligatory nature of the locative argument and the possibility of realizing the subject position by a locative deictic (of the ‘*aqui*’ type), modifier of the VP – or conversely, there is a relationship between the fact that the genitive is generated in the structure of the DP and the impossibility of its being realized in the subject position by a modifying constituent of the VP.

- (64) Brazilian Portuguese
- a. *Aqui bate sol.*
here hit_{3SG} sun
 ‘There’s sunlight here.’
- b. **Aqui furou o pneu.*
here pierced_{3SG} the tire
 ‘The tire was pierced here.’

One question that remains open is the fact that, at the VP level, the locative argument can be realized as DP or PP.

5 Locative adverbs in the subject position: the (formal) feature geometry of locative adverbs and their relation to the category name (N)

So far, we have explored the occurrence of locative PPs and NPs in argumental positions. We have also seen that locative adverbs share the syntactic distribution of such NPs and PPs, which suggests that locative adverbs can also occur in the subject position:

- (65) Brazilian Portuguese
- a. *Aqui/ali/lá cabe muita gente.*
Here/there/there fit_{3SG} a-lot-of people
 ‘Many people fit here/there/over there.’
- b. *Aqui/ali/lá bate sol.*
Here/there/there hit_{3SG} sun
 ‘This place/that place gets sunlight.’
- c. *Aqui/ali/lá adoecerá muita gente.*
Here/there/there get_{3SG} sick a-lot-of people
 ‘A lot of people get sick here/there/over there.’
- d. *Aqui/ali/lá desaparece crianças.*

- Here/there/there disappear_{3SG} children*
 ‘Children disappear here/there/over there.’
- e. Aqui/ali/lá brota feijão.
Here/there/there sprout_{3SG} beans
 ‘Beans sprout here/there/over there.’
- f. Aqui/ali/lá nasce muita criança.
Here/there/there be.born_{3SG} a-lot-of children
 ‘A lot of children are born here/there/over there.’

In a study about the status of the adverb category, considered from the point of view of its syntactic distribution in the subject position, Teixeira (2015) investigated structures such as (65). The possibility of locative adverbs occurring in such a position allows us to identify them, in this context, with properties of N. Baker (2003) proposes that there are three lexical categories: nouns, verbs, and adjectives. The author questions the feature system adopted in the tradition of generative theory, which distinguishes lexical categories by binary distinctive features, leading to the identification of four distinct categories.

- (66) a. +N, -V = Noun (BAKER, 2003, p. 21)
 b. -N, +V = Verb
 c. +N, +V = Adjective
 d. -N, -V = Preposition/Posposition

The author emphasizes, however, that such a distinction is not enough to establish the differences between the lexical categories, claiming that such features are not consistent – in the sense that they do not determine natural classes, since not only the pairs of categories formed by identical features {N, A} and {V, P} share properties, but also the category pairs {N, V} and {A, P}, which are formed by opposite features. Thus, it exemplifies the identity of the pair [A, P] with English data, showing that only AP and PP can occur with measure phrases, as in *It is three yards long* and *He went three yards into the water*. Baker then proposes a theory of lexical categories in terms of the features [+N] and [+V], but not as a system of binary oppositions, which results in the following contrast:

- (67) a. Noun is + N = has a referential index (BAKER, 2003, p. 21)
 b. Verb is + V = has a specifier
 c. Adjective is – N, – V (is a default category)
 d. Preposition is part of a different system (functional)

We will only discuss the proposal in (67a) about nouns, and then discuss the possibility that locative adverbs in the subject position have this property. The author proposes, among other properties, the following for N.

- (68) Syntactic version: X is a noun if and only if X is a lexical category and X bears a *referential index*, expressed as an ordered pair of integers. (BAKER, 2003, p. 95)

They also constitute canonical arguments of the sentence, occupying the position of subject, direct object, prepositioned object. Our goal is not to develop the argument, in the sense of detailing Baker's proposal, which encompasses different lexical categories. Although it is a comprehensive proposal, Baker's approach does not consider the specific case of adverbs. In this sense, we will seek to apply some of Baker's ideas to the discussion of the issues investigated here, in relation to the properties of this category, particularly considering locative adverbs.

Our hypothesis is that locative adverbs share with category N the property of manifesting a referential index. In Baker's (2003) words: "(...) the main idea is that only nouns can bear a referential index, because only they have 'criteria of identity' (...) this means that only they can bind anaphors, traces of various kinds, and the theta-roles of verbs, among other things" (BAKER, 2003, p. 21). Although the author is categorical in the sense that the properties mentioned are exclusive to N, it is not difficult to extend them to adverbs (locative adverbs/adverbs of manner), given the understanding that such categories occupy positions introduced by predicates, as in the case of predicates that select locative arguments (*Maria pôs o livro na estante/aqui* – 'Maria put the book on the shelf/ here'). In addition, the other properties mentioned are confirmed, since the locative NPs can be antecedents of *wh* words in relative clauses, as in (69a-b), antecedent in coordinate structure (69c-d):

(69) Brazilian Portuguese

a. Brasília,_p quei chove pouco, é a minha cidade
 predileta.

Brasília,_i which_i rain_{3SG} little, be_{3SG} the my city
favorite

Brasília, where it doesn't rain much, is my favorite city.'

b. Aqui,_p que,_i chove pouco, é a minha

Here, which_i rain_{3SG} little, be_{3SG} the my

cidade predileta.

city favorite

'Here, where it doesn't rain much, is my favorite city.'

c. Esse sitio,_i dá muitas bananas, mas também *e_i* é
this ranch_i give_{3SG} many bananas but also e_i be_{3SG}

cheio de laranjas.

full of oranges

'This ranch produces a lot of bananas, but it also yields a lot of oranges.'

d. Aqui,_i dá muitas bananas, mas também *e_i*

here_i give_{3SG} many bananas but also e_i

é cheio de laranjas.

be_{3SG} full of oranges

'There is a large output of bananas here, but also of oranges.'

The facts regarding the distribution of adverbs in the subject position – particularly in BP, given the requirement to fill this syntactic position by an XP – confirm our hypothesis in the sense of identifying them with category N, in relation to the index manifestation of referentiality. However, it should be noted that this property is, in principle, restricted to one type of adverb – the locative adverb. Considering the analysis of those constructions as locative inversions, such structures demonstrate not only the pronominal character of this category, but also highlight the relevance of this property in the BP grammar. In fact, BP manifests the use of this construction referred to as subject-topic with an innovative character, primarily by the presence of verbal agreement in structures with locative NP (and by the optional pronominal resumption by the adverb in the subject position, in case of realization of the locative argument preposed to the verb).

The idea of assigning nominal character to (locative) adverbs is not recent, as we saw in the analysis of locative inversion structures in BP. We consider that the proposal to assume a configuration headed by a category of the LOC type, with the PP (locative) realized as a complement, is technically complex. We claim that it is possible to generate the structure without resorting to this configuration. A question that arises is what determines the occurrence of this configuration.

We then proceed to explore the possibility of analyzing these cases in terms of the formal properties of the adverb *here* (*'aqui'*), in relation to locative PPs and DPs. As will be shown, we verified some patterns that allowed us to postulate the formal features associated with this category in its pronominal uses.

Thus, we start from the data pointed out by Pontes (1986), assuming with this author, and with Galves (1998), that, in fact, the locative DP controls the verbal agreement, as illustrated below, which constitutes a diagnosis for its realization in the position of subject.

- (70) Brazilian Portuguese
- a. *Esse sítio dá muitas bananas.*
this ranch give_{3SG} many bananas.
 'This ranch produces a lot of bananas.'
- b. *Esses sítios dão muita banana.*
these ranches give_{3PL} a-lot-of banana
 'These ranches yield a lot of bananas'
- c. *Esse sítio e essa fazenda dão muita banana.*
this ranch and this farm give_{3PL} a-lot-of banana
 'This ranch and this farm yield a lot of bananas.'

We also agree with Pilati's (2006) hypothesis that BP manifests the so-called locative inversion with V(O)S structures, the locative being lexically realized or null, the latter recovered anaphorically or by a deictic relationship (locative or temporal). Those structures primarily include clauses with unaccusative verbs and transitive verbs in concomitant narratives, and ergativized transitive structures, as noted by Galves (1998) and Avelar and Cyrino (2009), in which the external argument does not

receive morphosyntactic marking as in Bantu languages. We also follow Munhoz’s (2011) analysis that the locative NP in the subject position is the (internal) argument of the verb in structures with biargumental unaccusative verbs. We also considered structures with meteorological verbs, discussed in Pilati, Naves and Salles (2017), in which the use of locatives seems to manifest the properties postulated for the previous cases – highlighting the possibility of controlling agreement:

- (71) Brazilian Portuguese (PILATI; NAVES; SALLES, 2017, p. 66 and 75)
- a. *Aqui chove muito.*
Here rain_{3SG} a-lot
 ‘It rains a lot here.’
- b. *Essa cidade chove muito.*
this city rain_{3SG} a-lot
 ‘It rains a lot in this city.’
- c. *Essas cidades chovem muito.*
these cities rain_{3PL} a-lot
 ‘It rains a lot in these cities.’

However, when examining locative adverbs in coordination structure, we observed that the verb does not agree with those adverbs, which suggests they do not manifest the number feature, although they do, hypothetically, manifest the interpretable person feature, since they manifest referential index. This contrast is illustrated below:

- (72) Brazilian Portuguese
- a. *Aqui e ali bate sol à tarde.*
here and there hit_{3SG} sun at-the afternoon
 ‘There is sunlight in the afternoon here and there.’
- b. **Aqui e ali batem sol à tarde.*
here and there hit_{3PL} sun at-the afternoon
- (73) Brazilian Portuguese
- a. *Aqui e ali cabe muita coisa.*
here and there fit_{3SG} many things
 ‘A lot of things fit here and there.’
- b. **Aqui e ali cabem muita coisa.*
here and there fit_{3PL} many things

The absence of a number feature in the locative adverb is also referred to in Avelar (2009) as a property that interacts with the defective character of BP agreement, given the phenomenon of variation in verbal inflection, giving rise to locative inversion in BP. As mentioned above, our understanding is that the variable agreement observed in BP should not be taken as a determining factor in relation to the possibility of licensing, in the position of subject, a category such as the locative, seen as defective due to the absence of the number feature. Rather, we consider relevant the proposal of Pilati (2006) to relate the locative inversion in BP to the codification of the deixis in the sentence structure.

In this sense, we propose that the insertion of (lexical or null) locative adverbs/PPs and locative DPs in SpecTP satisfies the (temporal and locative) deixis encoding, which is required in these contexts – in this respect, we extend the proposal in relation to Pilati (2006), which does not include locative PP in that position. The T category, in turn, manifests uninterpretable *phi*-features (of person and number), and the EPP feature, which are verified by the category realized in SpecTP, as indicated below.

In the presence of locative adverbs, which, hypothetically, do not show a number feature, we conclude that the 3rd person feature, considered interpretable for manifesting a referential index, is sufficient to trigger *Agree*, with the number feature being validated as singular, which is the *default* option in the inflectional system of the BP verb.

Considering the specifics of the distribution of locative DP and PP/locative adverbs in SpecTP in locative inversion structures, regarding agreement (full or *default*, respectively), we tentatively propose that locative DP has person and number features, fully validating the *phi*-features and manifesting nominative Case, while PP/locative adverbs manifest the oblique (inherent) Case – which suggests a *Quirky Case*⁸ situation. We consider that the Oblique (inherent) Case, in the case of the adverb, is determined at the lexical/category level, or at the syntax level, in the presence of the preposition.

⁸ According to Boeckx (2000), there are elements that are marked on the surface with *Quirky Case* in the subject position, despite not having Nominative Case, which is the Case canonically associated with this position. Therefore, we suggest the situation of *Quirky Case* in BP in the sense that locatives behave as subjects without triggering *Agree* (in the case of locative PP), or by manifesting incomplete *Agree* (in the case of locative adverbs).

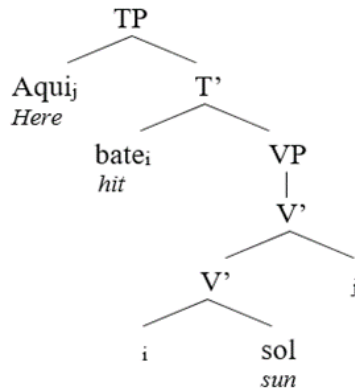
Hypothetically, the oblique Case, in this configuration, does not clash with the *Agree* operation at the TP level, since, in the case of the adverb, only the person feature in T is verified, the number feature being validated as singular, which is the *default* option; in the presence of PPs, there is, by hypothesis, only the licensed EPP-feature, since the features of number and person are not accessible, being, therefore, validated in the *default* option – 3rd person and singular.

We assume that the locative satisfies the requirement of the feature [+Loc] either of the clause structure (in the case of VS and constructions with a generic subject, according to Pilati, 2006) or of the predicate (in the case of subject-topic constructions, according to Munhoz, 2011). For this reason, despite the variable agreement in BP, not all elements without a number feature can occupy this position. Thus, we propose that the locative elements that occupy the position in SpecTP, as illustrated in (75) and (76), have the following formal features:

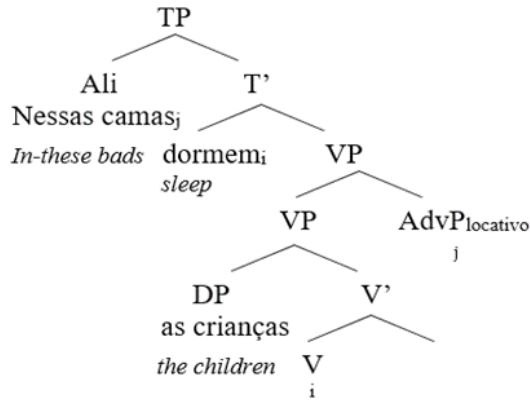
- (74) a. Locative DP: [+Loc], [+Person], [+Case], [+Number];
 b. Locative adverb: [+Loc], [+Person], [+Case];
 c. Locative PP: [+Loc], [+Case].

As the language presents, in the inflectional system of the verb, *default* number and person, the locative adverb and the locative PP occupy, as well as the locative DP, the position in SpecTP, because they satisfy the requirement of the feature [+Loc] of the predicate (in the case of subject-topic constructions), or of the sentence structure (in the case of the VS or generic subject constructions).

(75) Brazilian Portuguese



(76) Brazilian Portuguese



As we have seen, Larson (1985) discusses *bare NPs adverbs*. He proposes that those expressions are marked in the lexicon with the feature [+F], responsible for encoding *time*, *place*, mood, which would make them inherently marked for oblique Case (note that there is the possibility of not manifesting the feature [+F] if they occur in the subject position). The fact that this feature is optional, according to the author, allows *bare-NP adverbs* to occupy an argument position, receiving nominative Case (without conflicting with the oblique Case). Considering the technical difficulty of assuming this optional rule, and the fact that the expressions that satisfy this 'rule' are of different types, we postulate that such elements share the property of manifesting referentiality index, which is encoded by the feature of [person], in syntactic computation. This property is manifested through the following features, in different places in the grammar:

1. In locative or circumstantial adverbs marked for the feature [+deictic] (as *here* and *now*).
2. In nouns denoting days of the week (*Sunday*, *Wednesday*), at the lexicon level; languages can be less restrictive and include a name marked by existential quantification, as in the case of English, which includes 'someplace' (You have lived someplace; Peter put the letter someplace), which confirms the lexical character of the rule, even if it involves syntactic process to mark the quantification.

3. In nouns that denote time, in structures that include categories marked for the feature [+deictic] (demonstrative pronouns), constituting, therefore, at the level of syntax, as in *that day/aquele dia*, in English and Portuguese (*Aquele/*O dia choveu muito – That/*The day rained a lot*) – although a mark is still triggered in the lexicon, as there is a selection of words that denote time.
4. In (barely) any NPs or DP, if they are selected by predicates, as in *Essa casa bate sol* ('This house is sunny') – in this case, the feature [+deictic] is assigned to the DP by the predicate.

We also note that coordinated locative adverbs do not trigger agreement in constructions such as those presented in (72) and (73), but they trigger it in constructions such as those illustrated in the data below:

(77) Brazilian Portuguese

- a. *Aqui e ali são meu refúgio.*
here and there be_{3PL} my refuge
 'I take shelter here and there.' (interpretation: "anywhere")
- b. **Aqui e ali é meu refúgio.*
here and there be_{3SG} my refuge

The question that arises is why coordinate locative adverbs trigger agreement in copula structures, but not in the locative inversion structures discussed above. Our initial hypothesis is that, in this context, the possibility of plural agreement is associated with the fact that the predicate is of the 'equative' type, which denotes the identity between the referential features of the NP 'my refuge' and the locative phrase realized by the adverb. In this type of configuration, the locative adverb does not satisfy properties associated with the verb, which is an auxiliary. The core of the predicate is the NP 'my refuge'. In this configuration, the NP preaches in a distributive way for each adverb, which would explain the possibility of realizing the verb in the plural.

6 Final considerations

To discuss the questions we formulated in this research, we contrasted crosslinguistic data in which adverbs, locative PPs and DPs occupy the subject position. We observed that locative DPs behave differently from locative adverbs. Locative DPs are capable of triggering

agreement; locative adverbs in a coordination structure do not trigger it. We assumed that those DPs manifest an inherent number feature, while locative adverbs do not have such a feature, as we can see in the following data:

(77) Brazilian Portuguese

a. Esse sítio dá muitas bananas.

this ranch give_{3SG} many bananas

‘This ranch produces a lot of bananas.’

b. Esses sítios dão muita banana.

these ranches give_{3PL} a-lot-of banana

‘These ranches produce a lot of bananas.’

c. *Aqui e ali dão muita banana.

here and there give_{3PL} a-lot-of banana

With this in view, we conclude our work with the proposal that the insertion of PP/locative adverb (lexical or null) and of locative DP in Spec-TP satisfies the requirements of deixis coding. In the case of locative adverbs, which, by hypothesis, do not bear a number feature, but manifest the 3rd person feature (referential), we conclude that the number feature in T is validated as singular, which is the *default* option in the BP grammar, as the presence of the person feature is enough to trigger *Agree*. Still, regarding agreement (full or *default*), considering the distribution of locative DP and PP/locative adverbs in Spec-TP in locative inversion structures, we tentatively propose that:

- 1) DP locatives receive nominative Case, since they have the following features: [+Loc], [+Person], [+Number] and [+Case], and check the features of T (complete).
- 2) PP/locative adverbs are marked with the oblique Case (inherent), which, hypothetically, does not conflict with the *Agree* operation, due to the defective character of this operation: in the presence of adverbs, *Agree* validates the person feature in T (but not number, which is validated as singular, which is the *default* option in the inflectional system of the verb);
- 3) in the presence of locative PP, there is no Case marking, and the person and T number features are validated in the *default* option.

We concluded, then, that adverbs and locative PP can occupy the Spec-TP position, as well as locative DP, since they satisfy the requirement of the feature [+Loc] in the predicate. However, we note that the agreement is not triggered with the locative adverb: ‘(*Essas casas batem sol*’/‘**Aqui e ali batem sol*’ (‘These houses are sunny’/‘*Here and there are sunny’). As the DP and the adverb occupy the same position without change of meaning, we propose that locative DPs have a number and person feature, which licenses the triggering of agreement in the plural. In the case of locative adverbs, our hypothesis is that they have a 3rd person feature and do not have a number feature. With this specification, the T probe identifies the interpretable feature of person in the locative phrase, and the *Agree* operation occurs, which allows the checking of the uninterpretable 3rd person feature in T. In the absence of the number feature in the locative adverb, the uninterpretable feature of number in T is validated as singular, which is the *default* option. The property [+deictic] together with the 3rd person feature allows those adverbs to occupy the subject position in BP.

We postulated, then, that in the presence of locative DP, which has the features [+Person], [+Number], [+Loc] and [+Case], the *phi*-features in T are fully valued, and the locative DP checks Nominative Case. In the presence of the adverb that has, by hypothesis, the features [+Loc], [+Person] and [+Case (oblique)], the person feature in T is valued, and the number feature is realized as singular, which is the *default* option in the verb inflection system; in the presence of locative PP which, by hypothesis, has the features [+Loc] and [+Case (oblique)], the features of number and person of T are validated in the *default* option. Those elements can occupy the Spec-TP position because they are able to satisfy the requirement of the feature [+Loc] in the predicate.

With this study, we have contributed to the understanding of the status of the adverb, considering a specific group: the locative adverbs. However, we know that questions answered in this research generate other issues, which can motivate future research. We consider it relevant to investigate the status of the preposition in the structure of the locative PP, which alternates with the locative DP. The possibility of alternating between the PP and the locative DP suggests that the P category manifests properties of a functional head.

References

AVELAR, J. Inversão locativa e sintaxe de concordância no português brasileiro. *Matraga*, v. 16, n. 24, p. 232-252, 2009.

AVELAR, J.; CYRINO, S. Locativos em posição de sujeito: línguas bantas e português brasileiro. In: MORAIS, M. A. T.; ANDRADE, M. L. C. V. O. *História do português paulista*. Campinas: Publiel/Fapesp, 2009. p. 218-249.

AVELAR, J.; CYRINO, S. Locativos preposicionados em posição de sujeito: uma possível contribuição das línguas Bantu à sintaxe do português brasileiro. *Revista de Estudos Linguísticos da Universidade do Porto*, v. 3, p. 55-75, 2008.

AVELAR, J.; CYRINO, S.; GALVES, C. M. C. Locative inversion and agreement patterns: parallelisms between Brazilian Portuguese and Bantu languages. In: PETTER, M.; MENDES, R. B. (org.). *Proceedings of the Special World Congress of African Linguistics*. São Paulo: Humanitas, 2009, p. 207-221.

BAKER, M. *Lexical categories: verbs, nouns and adjectives*. Cambridge University Press, 2003.

BOECKX, C. Quirky Agreement. *Studia Linguistica*, v. 54, n. 3, p. 354-380, 2000. DOI: <https://doi.org/10.1111/1467-9582.00070>.

BRESNAN, J. Locative inversion and the architecture of Universal Grammar. *Language*, v. 70, n. 1, p. 72-131, 1994. DOI: <https://doi.org/10.2307/416741>.

BRESNAN, J.; KANERVA, J. Locative inversion in Chichewa: A case study of factorization of grammar. *Linguistic Inquiry*, v. 20, n. 1, p. 1-50, 1989.

DUARTE, M. E. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In: ROBERTS, I.; KATO, M. (org.). *Português brasileiro: uma viagem diacrônica*. Campinas: Editora da Unicamp, 1993. p. 107-128.

FERREIRA, M. *Argumentos nulos em português brasileiro*. 2000. 113 f. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, 2000.

GALVES, C. Tópicos e sujeitos, pronomes e concordância no português brasileiro. *Cadernos de Estudos Linguísticos*, v. 44, p. 19-31, 1998. DOI: <https://doi.org/10.20396/cel.v34i0.8637048>.

LARSON, R. Bare-NP Adverbs. *Linguistic Inquiry*, v. 16, n. 4, p. 595-621, 1985.

MUNHOZ, A. *A estrutura argumental das construções de tópico-sujeito: o caso dos sujeitos locativos*. 2011. 132f. Dissertação (Mestrado em Linguística) – Instituto de Letras, Universidade de Brasília, 2011.

PEREIRA, B. *A sintaxe cartográfica de 'lá' no português brasileiro: um estudo da periferia esquerda*. 2011, 256f. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade Federal de Minas Gerais, 2011.

PILATI, E. *Aspectos sintáticos e semânticos de orações com ordem verbo-sujeito no português do Brasil*. 2006. 242f. Tese (Doutorado em Linguística) – Instituto de Letras, Universidade de Brasília, 2006.

PILATI, E.; NAVES, R. R.; SALLES, H. M. M. L. Locative DPs and deictic adverbial pronouns in subject position in Brazilian Portuguese. In: FERNÁNDEZ-SORIANO, O.; CASTROVIEJO MIRÓ, E.; PEREZ-JIMÉNEZ, I. (org.). *Boundaries, phases and interfaces: case studies in honor of Violeta Demonte*. Amsterdam: John Benjamins, 2017, p. 64-84. DOI: <https://doi.org/10.1075/la.239.04pil>.

POLINSKY, M. Subject Inversion and Intransitive Subject Incorporation. In: *Papers from the 29th Regional Meeting of the Chicago Linguistic Society*. Volume 1: The Main Session. 1993, p. 343-361.

PONTES, E. *O tópico no português do Brasil*. Campinas: Pontes, 1987.

PONTES, E. *Sujeito: da sintaxe ao discurso*. São Paulo: Ática. Brasília: Fundação Nacional Pró-Memória, 1986.

SALZMANN, M. *Theoretical Approaches to Locative Inversion*. 2001. 317 f. Dissertação revisada (Mestrado em Linguística) – Philosophical Faculty of the University of Zurich, 2004.

TEIXEIRA, Z. D. *Propriedades sintáticas e semânticas dos advérbios no português brasileiro*. 2015. 173 f. Tese (Doutorado em Linguística) – Instituto de Letras, Universidade de Brasília, 2015. DOI: <http://dx.doi.org/10.26512/2015.03.T.18679>.

Authorship statement

Zenaide Dias Teixeira worked in the conception and design of the research, and collected and analyzed the data. Humberto Borges analyzed the data and revised the paper.



Null subjects and null objects in Brazilian Portuguese: correlations and change

Sujeitos e objetos nulos em português brasileiro: correlações e mudança

Gabriel de Ávila Othero

Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Rio Grande do Sul / Brasil

gabriel.othero@ufrgs.br

<http://orcid.org/0000-0002-2060-6312>

Melissa Lazzari

Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Rio Grande do Sul / Brasil

melissaglazzari@gmail.com

<http://orcid.org/0000-0002-7819-9326>

Abstract: Pronominal and null subjects and anaphoric direct objects are amongst the main grammatical phenomena that differentiate Brazilian and European varieties of Portuguese (cf. CYRINO; MATOS, 2016; DUARTE; FIGUEIREDO SILVA, 2016). In Brazilian Portuguese (BP), clauses with overt referential pronominal subjects are preferred over clauses with null subjects – which are restricted to some contexts (cf. AYRES, 2021, DUARTE; REIS, 2018); on the other hand, clauses with null anaphoric direct objects, especially 3rd person direct objects, are preferred over clauses with pronominal direct object (cf. CYRINO, 1993, 1997). The literature relating both phenomena in the history of BP shows that there has been a change in the following direction: null objects and overt pronominal subjects have had an increased frequency throughout the last centuries. Here we argue for a relation connecting both phenomena in BP: both are sensitive to the semantic gender of the referent when it comes to 3rd person pronoun realization. We support our hypothesis by bringing data from our previous work and work from colleagues. We also argue for a crucial difference between both phenomena: even though null objects are stabilized in BP grammar, we present empirical

data that show evidence for change in apparent time in 3rd person pronominal subjects. We have collected and analyzed data from a contemporary oral corpus (LínguaPOA) and here we show that (i) the asymmetry between 1st and 2nd person subjects on the one hand and 3rd person subjects on the other no longer exists; (ii) 1st and 2nd person null subjects are a stabilized grammatical phenomenon; and (iii) 3rd person null subjects are still changing.

Keywords: change; null subject; pronominal subject; anaphoric direct object; Brazilian Portuguese.

Resumo: Sujeitos pronominais e nulos e objetos diretos anafóricos estão entre os principais fenômenos gramaticais que diferenciam as variedades do português brasileiro e europeu (cf. CYRINO; MATOS, 2016; DUARTE; FIGUEIREDO SILVA, 2016). No português brasileiro (PB), as orações com sujeitos pronominais referenciais são preferidas em detrimento das orações com sujeitos nulos – que são restritas a alguns contextos (cf. AYRES, 2021, DUARTE; REIS, 2018); por outro lado, orações com objetos diretos anafóricos nulos, especialmente objetos diretos de 3^a pessoa, são preferíveis a orações com objeto direto pronominal (cf. CYRINO, 1993, 1997). A literatura que relaciona ambos os fenômenos na história do PB mostra que houve uma mudança na seguinte direção: objetos nulos e sujeitos pronominais tiveram uma frequência crescente ao longo dos últimos séculos. Aqui defendemos uma relação ligando ambos os fenômenos no PB: ambos são sensíveis ao gênero semântico do referente quando se trata de realização de pronome de 3^a pessoa. Apoiamos tal hipótese trazendo dados de nossos trabalhos anteriores e de trabalhos de colegas. Também defendemos uma diferença crucial entre os dois fenômenos: embora objetos nulos estejam estabilizados na gramática do PB, apresentamos dados empíricos que mostram evidências de mudança no tempo aparente em sujeitos pronominais de 3^a pessoa. Coletamos e analisamos dados de um *corpus* oral contemporâneo (LínguaPOA) e aqui mostramos que (i) não existe mais a assimetria entre sujeitos de 1^a e 2^a pessoa (por um lado) e sujeitos de 3^a pessoa (por outro); (ii) sujeitos nulos de 1^a e 2^a pessoa são fenômenos estabilizados; e (iii) sujeitos nulos de 3^a pessoa ainda estão em processo de mudança.

Palavras-chave: mudança; sujeito nulo; sujeito pronominal; objeto direto anafórico; português brasileiro.

Recebido em 26 de janeiro de 2022

Aceito em 06 de abril de 2022

1 Introduction

Pronominal and null subjects and anaphoric direct objects are amongst the main grammatical phenomena that differentiate Brazilian and European varieties of Portuguese (cf. CYRINO; MATOS, 2016; DUARTE; FIGUEIREDO SILVA, 2016). In Brazilian Portuguese (BP), clauses with overt referential pronominal subjects are preferred over clauses with null subjects – which are restricted to some contexts (cf. AYRES, 2021; DUARTE; REIS, 2018), thus (1) is ‘preferred’ (or more frequent) in comparison to (2).

(1) A Maria acabou de sair. Ela estava bem feliz.
the Mary finished of leave. She was pretty happy
‘Mary just left. She was pretty happy’

(2) A Maria acabou de sair. Ø Estava bem feliz.
the Mary finished of leave. Was pretty happy
‘Mary just left. She was pretty happy’

On the other hand, clauses with null anaphoric direct objects, especially 3rd person direct objects (3), are preferred over clauses with pronominal direct object (4) (cf. CYRINO, 1993, 1997; SCHWENTER, 2006; SCHWENTER; SILVA, 2003, *inter alia*).

(3) Comprei um livro, mas ainda não tive tempo de ler Ø.
Bought.1p.sg a book, but still no had.1p.sg time of read
‘I bought a book, but I still didn’t have time to read it.’

(4) ?Comprei um livro, mas ainda não tive tempo de ler ele.
Bought.1p.sg a book, but still no had.1p.sg time of read he
‘I bought a book, but I still didn’t have time to read it.’

Literature has related (albeit timidly) both phenomena in the history of BP. An example of such a study is Tarallo’s (1983, 1986) pioneering work, which indicates, through corpus analysis, a change in the expression of pronominal subjects and direct objects in BP. Tarallo highlights the relationship between the two phenomena, presenting diachronic data that show the increase of overt pronominal subject and the simultaneously decrease of pronominal direct objects in BP.

Graph 1 - Overt pronominal subjects and direct objects in BP across time.



Source: adapted from Tarallo (2018, p. 40).

Even though both changes ‘started together’ in BP and they seem to be related, as Tarallo pointed out, only the null direct object is considered a grammatical process whose change is stabilized (cf. CYRINO, 1993, 1997); the change in the parameter of subject marking in BP (from a +pro-drop to a -pro-drop language or, alternatively, to a partial null subject language) is still topic of lively debate among linguists who investigate the history of BP (cf. DUARTE; MARINS, 2021; DUARTE; REIS, 2018; GRAVINA, 2014a, 2014b; HOLMBERG; NAYUDU; SHEEHAN, 2009).

Here we have two main goals: the first goal is to argue for a relation connecting both phenomena in BP (null subjects and null objects); the second one is to argue for a crucial difference between them.

In section 2, we explore their connection. We believe that, since vernacular BP has lost 3rd person clitics (adopting the full pronouns *ele/ela/eles/elas*, ‘he/she/they.masc./they.fem’ along with the possibility of phonetically null objects), an agreement effect has become crucial to understand the distribution of pronouns and null objects. This agreement effect is related to the semantic gender of the referent, in the following sense: referents denoting masculine semantic gender (*o professor*, ‘the male-teacher’, for example) will favor pronominal direct objects with the masculine pronoun (*ele*, ‘he’), referents denoting female semantic gender (*a professora*, ‘the female-teacher’, for example) will favor pronominal direct objects with the feminine pronoun (*ela*, ‘she’) and, finally, referents denoting neutral semantic gender (*a mesa*, ‘the table’, or *o livro*, ‘the

book’) will favor null objects, regardless of their grammatical gender (in BP, nouns can be masculine or feminine, *a mesa* being a feminine noun, *o livro*, a masculine noun).

This idea is not new and is not ours. It dates back to Creus and Menuzzi (2004) and has been explored by recent work (we provide references in section 2). Our idea here is to expand the ‘semantic gender hypothesis’ to investigate null and pronominal subjects. We believe this agreement effect is influencing both phenomena, pronominal direct objects and pronominal subjects, in the sense that a referent denoting masculine or feminine semantic gender will favor the use of a pronoun (in the way previously outlined), whereas a referent denoting neutral semantic gender will favor null subjects and null objects. In the next section, we will not present new data to support this idea; instead, we bring data from previous works (see references in the next section). This agreement effect is, thus, something connecting these two phenomena.

In section 3, we take a different direction. We explore a crucial difference between null subjects and null objects. As we have pointed out, Tarallo’s early investigation showed that both pronominal direct objects and null subjects started to decline in the same period of time in BP history (mid-19th century). Nonetheless, we argue here that, in spite of the stabilization of null and pronominal direct objects, the null subject phenomenon is still a process in change in the grammar of the language. This idea is not new and is not ours either. Duarte (in many works), for example, advocates for the idea that null and pronominal subjects in BP are in constant change since the mid-19th century. Null subjects are residual and are diminishing both in frequency and in the variety of pragmatic and syntactic contexts where they can appear. In other words, this grammatical phenomenon has not yet stabilized in the grammar. Here we bring new unpublished data to support the hypothesis of change. We have investigated a contemporary oral corpus, the LínguaPOA corpus, and found interesting results when we stratified the informants by age groups: we found that the older group (60+) uses considerably more 3rd person null subjects (46% of null subjects, 54% of pronominal subjects) than the younger group (20-39, where we found 11% of null subjects, 89% of pronominal subjects). That ‘change-in-apparent-time’ analysis supports Duarte’s findings about the continuing change involving null and pronominal subjects in BP.

2 Null and overt pronominal subject and anaphoric direct object in BP, a correlation: the semantic gender hypothesis

Both null and overt pronominal subjects and anaphoric direct objects are strongly related (or so we argue) to ‘natural’ or ‘semantic’ gender agreement: BP has a masculine 3rd person pronoun that refers to male beings and referents (‘ele’, he), a feminine 3rd person pronoun that refers to female beings (‘ela’, she) and an empty category that refers to neutral beings and objects (Ø, it), a parallel, so to speak, to the English pronominal system, for example.

All nouns in Portuguese are either masculine or feminine in terms of *grammatical gender* (cf. CAMARA Jr., 2019, 2021), despite referring to inanimate (5) or animate referents (6). They agree in grammatical gender with other elements in the noun phrase, such as determiners and adjectives that modify them:

(5) o livro caro / a casa cara
 the.masc book expensive.masc / the.fem house expensive.fem
 ‘the expensive book / the expensive house’

(6) o menino esperto / a menina esperta
 the.masc boy clever.masc / the.fem girl clever.fem
 ‘the clever boy / the clever girl’

Nouns that refer to inanimate entities, such as *livro* (‘book’), *casa* (‘house’), etc., or animate entities, such as *cônjuge* (‘spouse’), *vítima* (‘victim’), which have grammatical gender, but not semantic gender will be ‘neutral’, i.e. -sem.g. On the other hand, nouns with semantic gender (+sem.g) will either refer to feminine entities, such as *menina* (‘girl’), *gata* (‘female cat’), or masculine entities, *menino* (‘boy’), *gato* (‘cat’), and denote animate entities only.

We argue here (following COELHO *et al.*, 2017; CREUS; MENUZZI, 2004; OTHERO; SPINELLI, 2019a, 2019b) that the pronoun system in BP (at least when it comes to pronoun use as subjects or anaphoric direct objects) is sensitive to the ‘semantic’ gender of the referents. As we said in the previous section, we do not bring new evidence here nor any new theoretical point of view to support this hypothesis, but we present data from our previous investigations and from the works of colleagues.

Table 1 - Distributional tendency of overt and empty pronominal forms in contemporary Brazilian Portuguese.

Semantic gender of referent	Subject and object pronoun
feminine	ela
masculine	ele
neutral	Ø

Source: the authors.

The situation is much clearer when these pronominal forms are anaphoric direct objects. As Cyrino (1993, 1997) and Tarallo (1986) show, the use of 3rd person clitic pronouns *-o/-a/-os/-as* ('him/her/them.masc/ them.fem') as anaphoric direct objects in Brazilian Portuguese started to decline in mid-19th century (these clitic pronouns are accusative pronouns, still used in monitored speech by people with formal education in Brazil, as well as in high standard written language). With the decline of these accusative pronominal clitics in vernacular BP, the grammar explored two other strategies for 3rd person anaphoric direct objects: (i) the use of the nominative 'full' pronouns *ele/ela/eles/elas* ('he/she/they.masc/they.fem') and (ii) the null object Ø, as we can see in the following examples (see also discussion in Nunes, 2018, and Othero; Cardozo, 2017).

- (7) Encontrei -o ontem.
Found.1p.sg. him/it yesterday.
'I found him/it yesterday'.
- (8) Encontrei ele ontem.
Found.1p.sg. he/it yesterday.
'I found him/it yesterday'.
- (9) Encontrei Ø ontem.
Found.1p.sg. yesterday.
'I found it yesterday'.

The original hypothesis that the semantic gender is guiding the distribution of nominative pronouns (8) and null objects (9) comes from

Creus and Menuzzi (2004). As the authors argue, this distribution is guided by an agreement principle (CREUS; MENUZZI, 2004, p. 161)¹:

The hypothesis that associates BP full pronouns to the presence of semantic gender, and null objects to the absence of semantic gender is more natural than the analogous hypothesis based on the distinction of animacy: after all, the basic difference between *ele/ela* ('he'/'she') and null objects is that the former carry gender specifications, while the latter are just unspecified for gender (...). In other words, the choice between null objects and full pronouns would basically result from a process of agreement between the antecedent and the anaphoric form: antecedents with semantic gender favor the use of pronouns because these are the anaphoric forms specified for gender; and antecedents without semantic gender favor the use of null objects precisely because null objects do not have a specification for semantic gender.

Thus, null objects point to referents with no semantic gender, whereas pronominal anaphoric objects preferably point to referents with semantic gender. The pronoun is a way of morphologically (and, thus, phonetically) exponentiating this semantic gender feature. Recent literature has supported this hypothesis. For example, Othero *et al.* (2016) conducted an experiment following this hypothesis and also the remark by Schwenter, to whom BP “displays a ‘split’ system of marking anaphoric DOs [direct objects]. All animate (mainly human) and specific anaphoric DOs are preferentially encoded overtly, while all others are preferentially encoded as null objects” (Schwenter, 2006, p. 26). This

¹ Original text: “a hipótese que associa os pronomes plenos do PB à presença de gênero semântico, e objetos nulos à ausência de gênero semântico é mais natural que a hipótese análoga baseada na distinção de animacidade: afinal, a diferença básica entre as formas *ele/ela* e os objetos nulos é que as primeiras portam especificações de gênero, enquanto que os últimos são justamente não-especificados para gênero (bem como para número, mas nisso os ONs não diferem significativamente dos PrPIs, já que os últimos podem ou não portar a flexão de número). Ou seja, a escolha entre ONs e PrPIs resultaria, basicamente, de um processo de concordância entre antecedente e forma anafórica: antecedentes com gênero semântico favorecem o uso de PrPIs porque estas são as formas anafóricas especificadas para gênero; e antecedentes sem gênero semântico favorecem o uso de ONs precisamente porque ONs não possuem especificação para gênero semântico.”

‘split system’ can be explained by the semantic gender of the referents, as Othero *et al.* (2016) show. According to the authors (2016, p. 10)²,

What our data indicate quite clearly *does not concern the conditioning of the null object itself*, but the conditioning of the use of the full pronoun in the anaphoric resumption of antecedents with semantic gender, [+sem.g]. That is, analyzing the results of our tests, we can see that, if an antecedent has the feature [+sem.g], there is a strong tendency for it to be referred to by a pronoun – and not by a null object (...). On the other hand, if an antecedent does not have semantic gender, its resumption is sometimes made with a pronoun, sometimes with an empty category. This indicates, among other things, that there is a tendency towards complementary distribution, but it is not categorical.

In a later study, Othero and Spinelli (2017) analyzed a speech corpus (VARFUL³, from the Rio Grande do Sul sample) from the 1990s and found more empirical data to support this hypothesis. Here we summarize some of their findings:

Table 2 - Null and pronominal objects in a spoken language corpus from the 1990s (3rd person only).

	+sem.g referents	-sem.g referents
Null objects	10/218 (4,5%)	208/218 (95,4%)
Pronominals	41/61 (67,3%)	20/61 (32,7%)

Source: adapted from Othero and Spinelli (2017, p. 188).

² Original text: “O que nossos dados apontam de maneira bastante clara *não diz respeito ao condicionamento do objeto nulo* em si, mas ao condicionamento do uso do pronome pleno na retomada anafórica de antecedentes *com gênero semântico*, [+gs]. Ou seja, analisando os resultados de nossos testes, podemos perceber que, se um antecedente tem o traço [+gs], há uma forte tendência para que ele seja retomado por um pronome – e não por um objeto nulo (...). Por outro lado, se um antecedente não tiver o gênero semântico marcado, sua retomada ora é feita com um pronome, ora com uma categoria vazia. Isso indica, entre outras coisas, que há uma tendência à distribuição complementar, mas ela não é categórica.”

³ VARFUL is a corpus project that collects speech data, in the form of sociolinguistic interviews, from the southern region of Brazil, cf. Collischonn and Monaretto (2012) and Bisol and Monaretto (2016).

As the authors show, null objects tend to refer to antecedents with no semantic gender (95,4% of the cases), whereas full pronouns tend to refer to antecedents with apparent semantic gender (41 out of the 61 occurrences).

Coelho *et al.* (2017) also analyzed the VARSUL corpus, but from a different region (the sample from Santa Catarina), from the 1990s and 2010s. They reached a similar conclusion (2017, p. 2615)⁴:

After reanalyzing all 322 occurrences of anaphoric direct objects with full pronouns (20 occurrences) and null forms (302 occurrences) in our sample, taking into account the three features of the antecedent NP (animacy, specificity and semantic gender), we conclude that the feature of semantic gender acts as a conditioning factor for the phenomenon: antecedents that have semantic gender favor the anaphoric resumption with the full pronoun; antecedents that do not have semantic gender favor the resumption with null object. (...)

The semantic gender feature, in fact, seems to act not only on the phenomenon of anaphoric direct object of 3rd person, but on the phenomenon of anaphoric direct object in general in BP (involving 1st, 2nd and 3rd grammatical persons, both in pronominal as null objects), as suggested by the works of Schwenter (2006) and Othero *et al.* (2016) (...).

⁴ Original text: “Após reanalisar todas as 322 ocorrências de retomadas anafóricas com pronomes plenos (20 ocorrências) e objetos nulos (302 ocorrências) em nossa amostra, levando em consideração os três traços do SN antecedente (animacidade, especificidade e gênero semântico), concluímos que o traço de gênero semântico atua como fator condicionador para o fenômeno: os SNs antecedentes que têm gênero semântico identificado favorecem a retomada anafórica com o pronome pleno; os SNs antecedentes que não têm gênero semântico identificado favorecem a retomada com o objeto nulo (...)

O traço de gênero semântico, na verdade, parece atuar não somente sobre o fenômeno da retomada de objeto direto anafórico de 3ª pessoa, mas sobre o fenômeno de retomada anafórica de objeto direto em geral em PB (envolvendo pronomes e ONs, de 1ª, 2ª e 3ª pessoas gramaticais), como sugerem os trabalhos de Schwenter (2006) e Othero *et al.* (2016) (...).”

The authors go further (2017, p. 2615)⁵:

If this is indeed the case, we can expect that this feature (semantic gender) is actually behind the distribution of pronouns and empty elements not only in the anaphoric direct object, but also in the subject position. This would allow us to reach an interesting generalization: that referents marked with semantic gender would favor the use of full pronouns in BP, whether as direct objects or as subjects.

Following this suggestion, Othero and Spinelli (2019b) analyzed null and pronominal *subjects* in a spoken language corpus from the 1990s (VARSUL, Porto Alegre sample). They found the same tendency: overt pronominal subjects tended to refer to +sem.g referents, whereas null subjects tended to refer to -sem.g referents. Some of their results can be seen in Table 3:

Table 3 - Null and overt pronominal subjects according to the semantic gender feature of their referent in spoken language corpus from the 1990s.

	+sem.g referents	-sem.g referents
Pronominal subjects	250/336 (74,4%)	86/336 (25,6%)
Null subjects	32/106 (30,2%)	74/106 (69,8%)

Source: Adapted from Othero and Spinelli (2019b, p. 19).

In the same study, the authors investigated a more recent spoken language corpus from the same location as VARSUL, the LínguaPOA corpus, containing sociolinguistic interviews from 2015 to 2018⁶. They found very similar results, strengthening the hypothesis mentioned by Coelho *et al.* (2017, p. 2615), “that this feature (semantic gender) is actually behind the distribution of pronouns and empty elements not

⁵ Original text: “Se esse for realmente o caso, podemos esperar que esse traço (gênero semântico) esteja por trás, na verdade, da distribuição entre pronomes *versus* elementos vazios não apenas na retomada anafórica de objeto direto, mas também na função de sujeito. Isso nos permitiria chegar a uma generalização interessante: a de que os referentes marcados com gênero semântico favoreceriam o uso de pronomes retos em PB, seja na função de objeto direto, seja na função de sujeito.”

⁶ LínguaPOA is an oral corpus composed of sociolinguistic interviews with informants from the city of Porto Alegre, cf. Battisti (2019).

only in the anaphoric direct object, but also in the subject position”. In the table below, we report Othero and Spinelli’s (2019b) results:

Table 4 - Null and overt pronominal subjects according to the semantic gender feature of their referent in a corpus from the 2010’s.

	+sem.g referents	-sem.g referents
Pronominal subjects	236/317 (74,4%)	81/317 (25,6%)
Null subjects	8/86 (9,3%)	78/86 (90,7%)

Source: Adapted from Othero and Spinelli (2019b, p. 22).

However, there are more factors playing a role in the distribution of null and overt pronominal subjects in BP than just the semantic gender feature of the referent. Other factors related to the phenomenon are topic-chain continuity (cf. LAZZARI, 2021; PAREDES SILVA, 2003), linear order or the elements in the clause (cf. KATO, 2020; KATO; DUARTE, 2018) and verbal inflectional morphology (cf. SOARES, 2017; SOARES; MILLER; HEMFORTH, 2019). Our point here is that the semantic gender hypothesis can be applied to the investigation of pronominal subjects and can (help) explain the distribution of 3rd person pronominal direct objects in BP. It is certainly not the only factor at play, but it is certainly one of them (cf. Ayres, 2021, Othero; Spinelli, 2019a, 2019b, Soares; Miller; Hemforth, 2020, for a broad analysis and argumentation). For our purposes here, this is important because unlike topic-chain continuity, linear order and verbal inflection morphology (that are related to the distribution of null and pronominal subjects – but *not* to pronominal direct objects), the semantic gender feature of the referent is a factor playing a role both in the distribution of null and pronominal subjects *and* null and pronominal direct objects in BP. In other words, the semantic gender feature can be pointed out as the factor that connects both phenomena in BP, since we are facing an agreement effect linking pronouns and referents.

We did not present any new data in this section. What we did was bring data from the literature (ours and from colleagues) attesting two points of connection between null subjects and null objects in BP, namely, (i) their frequency started changing in the same period of BP’s history, around mid-19th century, and (ii) their sensitivity to semantic gender information of referents, in the sense that referents denoting neutral

semantic gender favor null forms, whereas masculine and feminine referents (in terms of semantic gender – not grammatical gender) favor pronominal forms.

In the next section, we will explore a crucial difference between the two phenomena. We will present new data to support the idea that the null subject is still changing in contemporary BP. We present a quantitative analysis⁷ of the data we found investigating 5,846 occurrences of null and pronominal subjects in the interviews from the sociolinguistic corpus *LínguaPOA*. We show that there are elements that allow us to pursue the idea of a change in apparent time when it comes to the null subject phenomenon.

3 Change

Although the two phenomena are closely related, as we explored in the previous section, and although null object is a stabilized grammatical phenomenon in the grammar of BP since the late 20th century, we argue here that the null subject phenomenon is still a grammatical process in change. As we have already mentioned, this idea is not new and is not ours. What we do here is bring new data to support this claim. We have analyzed a contemporary oral corpus of vernacular BP, the *LínguaPOA* corpus (cf. BATTISTI, 2019). The corpus is constituted by a series of transcribed sociolinguistic interviews from 2015 to 2018 with informants from the city of Porto Alegre, a capital city in Southern Brazil.

We have analyzed 5,846 occurrences of null and pronominal subjects and found the following distribution:

Table 5 – Null vs. overt pronominal subjects found in the corpus.

Pronominal subjects	4,220 (72%)
Null subjects	1,626 (28%)
Total	5,846 (100%)

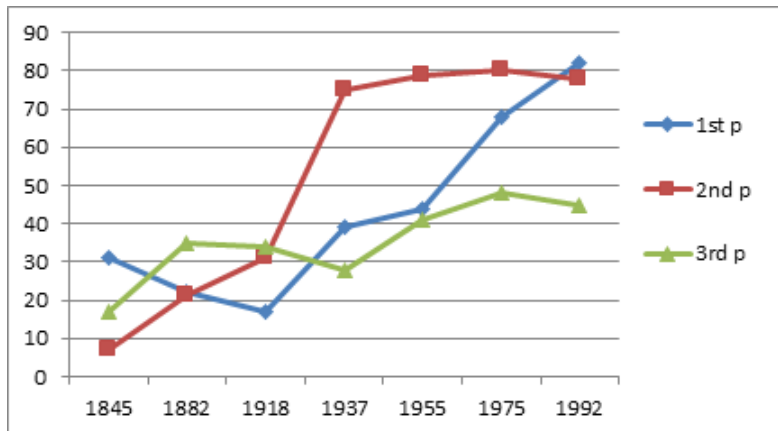
Source: the authors.

⁷ We present a qualitative analysis of the data in Ayres and Othero (forthcoming) and Othero, Lazzari and La Porta (forthcoming).

The data we found, at a first glance, lead us to think that the null subject in BP has stabilized in the late 20th century (the same period of time we find the stabilization of null objects), since we found roughly the same figures that have been long reported in the literature, i.e., around 70% of the anaphoric subjects being filled with an overt pronoun and 30% of null subjects (cf. BERLINCK; DUARTE, 2015, DUARTE, 1995, MONTEIRO, 1994).

Nonetheless, we have made three interesting empirical discoveries in our investigation. The first two are related to 1st and 2nd person subjects versus 3rd person subjects. The data from the 1990s present an asymmetry for 3rd person subjects, when compared to 1st and 2nd persons: Duarte (1993, 1995), analyzing theater plays, found that 1st and 2nd person subjects were majorly overt, whereas 3rd person subjects were split: roughly 45% were pronominal subjects and roughly 55% were null subjects (cf. DUARTE, 1993, p. 117):

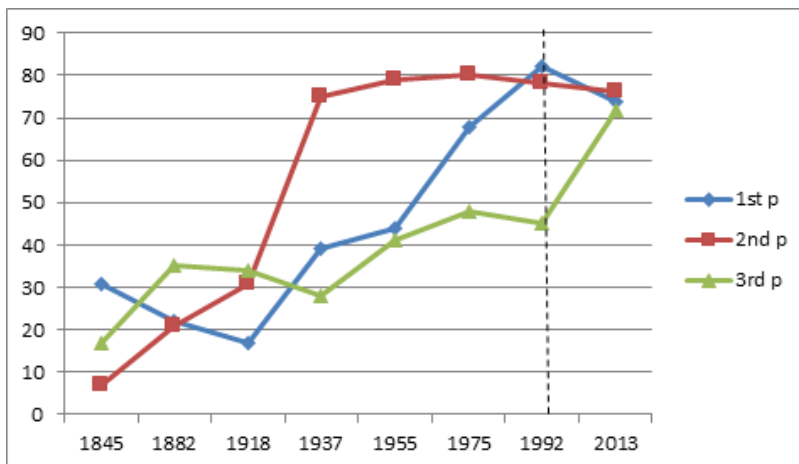
Graph 2 - Null vs. overt pronominal subjects in theatre plays.



Source: adapted from Duarte (1993, p. 117).

In a recent work, Othero and Spinelli (2019a) investigated theatre plays from 2010s (thirty years after the last play analyzed by Duarte) and found out that the asymmetry reported by Duarte (1993) involving the 3rd person no longer exists.

Graph 3 - Null vs. overt pronominal subjects in theatre plays.



Source: adapted from Othero; Spinelli (2019a, p.16).

Here, in our investigation of 3rd person subjects in an oral corpus, we found similar results: overt pronominal subjects are preferred over null subjects, confirming what Othero and Spinelli (2019a) found when investigating their corpus of theater plays from 2010⁸. In our investigation of LínguaPOA, we found 715 (72%) occurrences of 3rd person pronominal subjects versus 281 (28%) of 3rd person null subjects, as we show in Graph 4.

Graph 4 - 3rd person null vs. overt pronominal subjects.



Source: the authors.

⁸ Similar results have been reported by Berlink, Duarte and Oliveira (2015), who found 78% of 3rd person pronominal subjects in their oral corpus investigation.

That means we have found roughly the same distribution (~70/~30) between null and pronominal subjects across 1st, 2nd and 3rd persons, and the asymmetry between 3rd person pronominal/null subjects and 1st and 2nd person subjects no longer exists. The lack of this long-reported asymmetry was the first interesting finding from our corpus investigation, even though it may also lead us to think we are now facing a stable process in the grammar of BP, since all persons (1st, 2nd and 3rd), after all, present nearly the same frequency for overt pronominal subjects – as we show in Table 7:

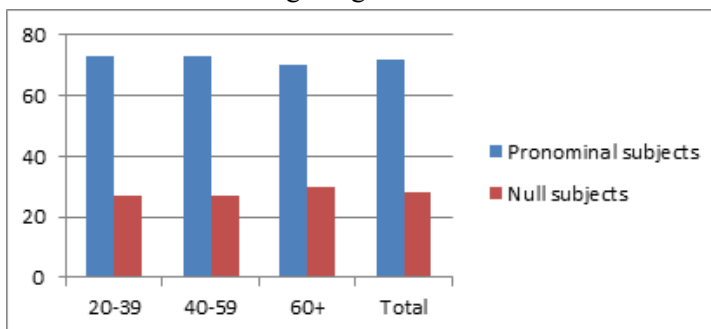
Table 7 – Null vs. overt pronominal subjects found in the corpus.

	1 st and 2 nd person sg.	3 rd person sg.
Pronominal subjects	3,505 (72%)	715 (72%)
Null subjects	1,345 (28%)	281 (28%)
Total	4,850 (100%)	996 (100%)

Source: the authors.

The second finding is another piece of empirical data that could support the idea of nulls subjects being a stable phenomenon in the grammar. It appeared when we decided to stratify the informants according to age group. *LínguaPOA* staff collected personal information from the informants, and we used the three age groups established by the *LínguaPOA* project: 20-39 years old, 40-59 years old and 60+ years old. When we analyzed 1st and 2nd person subjects, we found nearly the same distribution, as we show in Graph 5 (in percentages) and Table 8:

Graph 5 – 1st and 2nd person singular, null vs. overt pronominal subjects, according to age stratification.



Source: the authors.

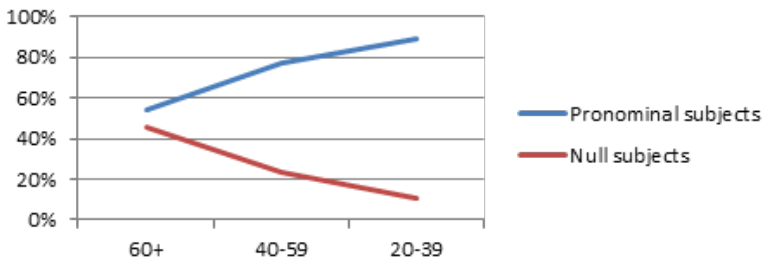
Table 8 – 1st and 2nd person singular, null vs. overt pronominal subjects, according to age stratification.

Age group	20-39	40-59	60+
Pronominal subjects	1,489 (73%)	1,189 (73%)	827 (70%)
Null subjects	537 (27%)	450 (27%)	358 (28%)
Total	2,026 (100%)	1,639 (100%)	1,185 (100%)

Source: the authors.

That means, to our understanding, we have evidence – both from diachronic analyses (such as Duarte, 1993, 1995, 2018) and from our own synchronic analysis – that the null subject is stable *when it comes to 1st and 2nd person singular in BP*. On the other hand, when the 3rd person singular is concerned, things are different – here is our third, and crucial, finding. When we analyzed 3rd person null and pronominal subjects according to age group, we found an interesting distribution that, to our understanding, is evidence of change in apparent time. We present our findings in Graph 6. The horizontal axis corresponds to age stratification (in years); the vertical axis to the application of the rule (in percentage), with the blue line marking the frequency of pronominal subjects and the red line marking the frequency of null subjects.

Graph 6 – 3rd person null and overt pronominal subjects according to age stratification.



Source: the authors.

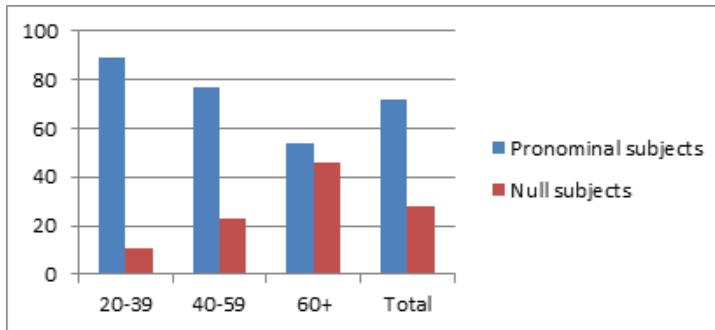
The older group presents a close distribution in the occurrences of 3rd person null and pronominal subjects, whereas the younger group has a clear preference for pronominal subjects. We show this data also in Table 9 and Graph 7:

Table 9 – 3rd person singular, null vs. overt pronominal subjects, according to age stratification.

Age group	20-39	40-59	60+
Pronominal subjects	187 (89%)	362 (78%)	175 (54%)
Null subjects	22 (11%)	101 (22%)	149 (46%)
Total	209 (100%)	463 (100%)	324 (100%)

Source: the authors.

Graph 7 – 3rd person null vs. overt pronominal subjects according to age stratification.



Source: the authors.

If we consider only the total distribution of null and pronominal subjects in the corpus (the last pair of columns in Graph 7), disregarding the age factor, we are misled to think the numbers are virtually the same for 1st, 2nd and 3rd person distribution between pronominal and null subjects (~70/~30, cf. Graph 5). However, analyzing the data more closely and relating the phenomena to the age factor, it seems clear that we are facing a process of change, supporting the argumentation in Duarte and Marins (2021), for example, who have reached a similar conclusion analyzing a very different corpus. We need, however, to keep further investigation in order to find more occurrences of null and pronominal subjects. We should also analyze different kinds of corpora, from different regions of Brazil, with spoken and written varieties.

4 Final remarks

We had two goals here: the first was to review some of the literature relating the phenomena of null and overt pronominal subjects and direct objects in Brazilian Portuguese. We proposed an approximation of these two phenomena via the ‘semantic gender hypothesis’, basing our argumentation on our own previous works and on colleagues’ investigations.

The second goal was to present new empirical data from our investigation of a contemporary transcribed BP oral corpus, the LínguaPOA corpus. What we found and presented in section 3 supports

the ideas that (i) the asymmetry between 1st and 2nd person subjects on the one hand and 3rd person subjects on the other no longer exists; (ii) 1st and 2nd person null subjects are a stabilized grammatical phenomena, since the frequency of null subjects has not changed since the late 20th century and it does not present any variation among different age groups; and (iii) 3rd person null subjects are still changing, since the data show change in apparent time.

Acknowledgments

Some of the ideas here were presented during the *III Encontro de Gramática Gerativa* at Universidade Federal da Bahia. We are grateful for the questions and comments we received during this event, especially the ones made by Maria Eugenia Duarte and Sonia Cyrino. We would also like to thank the colleagues Aline Gravina, Elisa Battisti, Mônica Rigo Ayres and Sergio Menuzzi, for constant dialogue and for reading and discussing some of the ideas we present here. Finally, we are very grateful for the anonymous reviewers who have read our paper and have made important comments on many different parts of the text.

Authorship statement

Gabriel de Ávila Othero analyzed the data and wrote the majority of the paper. Melissa Lazzari analyzed the data and revised the paper.

References

- AYRES, M. R. *Contextos licenciadores de sujeitos nulos em português brasileiro*. 2021. 123f. Tese (Doutorado em Letras) – Instituto de Letras, Universidade Federal do Rio Grande do Sul, 2021.
- AYRES, M. R.; OTHERO, G. A. *Contexts for Null Subjects in Contemporary Brazilian Portuguese*. Forthcoming.
- BATTISTI, E. *O acervo de entrevistas sociolinguísticas LínguaPOA: constituição, possibilidades e desafios*. Rio de Janeiro: e-book Open Access, 2019.
- BERLINK, R. A.; DUARTE, M. E. L.; OLIVEIRA, M. Predicação. In: KATO, M. A.; NASCIMENTO, M. (orgs.). *A Construção da Sentença: Gramática do Português Culto Falado no Brasil*. São Paulo: Contexto, 2015. p. 81-150.

- BISOL, L.; MONARETTO, V. N. O. Prefácio: VARSUL e suas origens, uma história sumariada. *ReVEL*, Novo Hamburgo, edição especial, n. 13, p. 6-11, 2016.
- CAMARA JR. *Problemas de linguística descritiva*. Edição revista e comentada. Petrópolis: Vozes, 2021.
- CAMARA JR., J. M. *Estrutura da língua portuguesa*: edição crítica. Edição, estabelecimento de texto, introdução e notas de Emílio Gozze Pagotto, Maria Cristina Figueiredo Silva, Manoel Mourivaldo Santiago-Almeida. Petrópolis: Vozes, 2019.
- COELHO, I. L.; OTHERO, G. A.; VIEIRA-PINTO, C. A. Reanálise de variáveis semânticas no condicionamento do objeto nulo e do pronome pleno na fala de Florianópolis. *Fórum Linguístico*, Florianópolis, v. 14, n. 4, p. 2606-2617., 2017. DOI: <https://doi.org/10.5007/1984-8412.2017v14n4p2606>
- COLLISCHONN, G.; MONARETTO, V. Banco de dados VARSUL: a relevância de suas características e a abrangência de seus resultados. *ALFA: Revista de Linguística*, v. 56, n. 3, p. 835-853, 2012.
- CREUS, S; MENUZZI, S. O papel do gênero na alternância entre objeto nulo e pronome pleno em português brasileiro. *Revista da ABRALIN*, v. 3, n. 1-2, p. 149-176, 2004.
- CYRINO, S. *et al. História do português brasileiro*: mudança sintática do português brasileiro: perspectiva gerativista. São Paulo: Contexto, 2018.
- CYRINO, S. M. L. *O objeto nulo no português do Brasil*: um estudo sintático-diacrônico. Londrina: UEL, 1997.
- CYRINO, S. M. L. Observações sobre a mudança diacrônica no português do Brasil: objeto nulo e clíticos. In: ROBERTS, I.; KATO, M. A. (orgs.) *Português brasileiro*: uma viagem diacrônica. Campinas: Ed. da Unicamp, 1993. p. 163-185.
- CYRINO, S. M. L.; MATOS, G. Null objects and VP ellipsis in European and Brazilian Portuguese. In: WETZELS, L.; COSTA, J.; MENUZZI, S. (eds) *The handbook of Portuguese linguistics*. Oxford: Blackwell, 2016. p. 294-317.
- DUARTE, I.; FIGUEIREDO SILVA, M. C. The null subject parameter and the structure of the sentence in European and Brazilian Portuguese. In: WETZELS, L.; COSTA, J.; MENUZZI, S. M. (eds.). *The handbook of Portuguese Linguistics*. West Sussex: Wiley Blackwell, 2016. p. 234-253.

DUARTE, M. E. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In: ROBERTS, I.; KATO, M. A. (orgs.). *Português brasileiro: uma viagem diacrônica*. Campinas: Ed. da Unicamp, 1993. p. 107-125.

DUARTE, M. E. L. *A perda do princípio “evite pronome” no português brasileiro*. 1995. 151 f. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Unicamp, 1995.

DUARTE, M. E. L. O sujeito nulo no português brasileiro. In: CYRINO, S.; MORAIS, M. A. T. *História do português brasileiro. Mudança sintática do português brasileiro: perspectiva gerativista*. São Paulo: Contexto, 2018. p. 26-71.

DUARTE, M. E. L.; MARINS, J. Brazilian Portuguese: a ‘partial’ null subject language. *Cadernos de Estudos Linguísticos*, Campinas, v. 63, p. 1-21, 2021. DOI: <https://doi.org/10.20396/cel.v63i00.8661660>

DUARTE, M. E. L.; REIS, E. P. R. Revisitando o sujeito pronominal vinte anos depois. *ReVEL*, Novo Hamburgo, v. 16, n. 30, p. 173-197, 2018.

GRAVINA, A. P. Diacronia e sujeito nulo no português brasileiro: um estudo comparativo. *Filologia e Linguística Portuguesa*, São Paulo, v. 16, p. 199-231, 2014a. DOI: <https://doi.org/10.11606/issn.2176-9419.v16ispep199-231>

GRAVINA, A. P. *Sujeito nulo e ordem VS no português brasileiro: um estudo diacrônico-comparativo baseado em corpus*. 2014. 251 f. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Unicamp, 2014b.

HOLMBERG, A.; NAYUDU, A.; SHEEHAN, M. Three partial null-subject languages: a comparison of Brazilian Portuguese, Finnish and Marathi. *Studia Linguistica*, Lund, v. 63, n. 1, p. 59-97, 2009. DOI: <https://doi.org/10.1111/j.1467-9582.2008.01154.x>

KATO, M. A.; DUARTE, M. E. L. Pre-verbal position in BP: a reinterpretation of “avoid pronoun principle”. *Diadorim*, Rio de Janeiro, v. 20 – Especial, p. 610-626, 2018. DOI: <https://doi.org/10.35520/diadorim.2018.v20n0a23293>

KATO, M. *Determinantes prosódicos em mudança sintática*. Abralin ao vivo, 2020. <[http:// https://youtu.be/t3BLRPloZJI](http://https://youtu.be/t3BLRPloZJI)>. Acesso em: 12 nov. 2021.

LAZZARI, M. G. Acessibilidade: o que isso tem a ver com o sujeito pronominal expresso e o sujeito nulo em português brasileiro? *Revista Linguística Rio*, Rio de Janeiro, v. 7, p. 42-58, 2021.

MONTEIRO, J. L. *Pronomes pessoais: subsídios para uma gramática do português do Brasil*. Fortaleza: Edições UFC, 1994.

NUNES, J. Direção de cliticização, objeto nulo e pronome tônico na posição de objeto em português brasileiro. In: ROBERTS, I.; KATO, M. (orgs.). *Português brasileiro: uma viagem diacrônica*. São Paulo: Contexto, 2018. p. 161-174.

OTHERO, G. A.; AYRES, M. R.; SCHWANKE, C.; SPINELLI, A. C. A relevância do traço gênero semântico na realização do objeto nulo em português brasileiro. *Working Papers em Linguística*, Florianópolis, v. 17, n. 1, p. 64-86, 2016. DOI: <http://dx.doi.org/10.5007/1984-8420.2016v17n1p64>

OTHERO, G. A.; CARDOZO, R. W. A ordem pronominal em português brasileiro: da ênclise à próclise, do clítico ao tônico (or There and Back Again, a Word Order's Holiday). *Fórum Linguístico*, Florianópolis, v. 14, n. 1, p. 1717-1734, 2017. DOI: <https://doi.org/10.5007/1984-8412.2017v14n1p1717>

OTHERO, G. A.; LAZZARI, M. G.; LA PORTA, B. *Contextos de resistência do sujeito nulo na análise do corpus LínguaPoa*. Forthcoming.

OTHERO, G. A.; SPINELLI, A. C. Sujeito pronominal expresso e nulo no começo do séc. XXI (e sua relação com o objeto nulo em PB). *Domínios de Linguagem*, Uberlândia, v. 13, n. 1, p. 7-33, 2019a. DOI: <https://doi.org/10.14393/DL37-v13n1a2019-1>

OTHERO, G. A.; SPINELLI, A. C. Um tratamento unificado da omissão e da expressão de sujeitos e objetos diretos pronominais de 3ª pessoa em português brasileiro. *Cadernos de Estudos Linguísticos*, Campinas, v.61, n. 1, p. 1-30, 2019b. DOI: <https://doi.org/10.20396/cel.v61i1.8654211>

PAREDES SILVA, V. L. Motivações funcionais no uso do sujeito pronominal: uma análise em tempo real. In: PAIVA, M. C.; DUARTE, M. E. L. (orgs.). *Mudança linguística em tempo real*. Rio de Janeiro: Contra Capa, 2003. p. 97-114.

PIVETTA, V. *Objeto direto anafórico no português brasileiro: uma discussão sobre a importância dos traços semântico-pragmáticos – animacidade/especificidade vs. gênero semântico*. 2015. 131f. Dissertação (Mestrado) – Faculdade de Letras, Universidade Federal do Rio Grande do Sul, 2015.

SCHWENTER, S. A. Null objects across South America. Selected proceedings of the 8th. In: HISPANIC LINGUISTICS SYMPOSIUM, 2006, Somerville: Cascadilla Press, 2006. p. 23-37.

SCHWENTER, S. A.; SILVA, G. Anaphoric direct objects in spoken Brazilian Portuguese: semantics and pragmatics. *Revista Internacional de Lingüística Iberoamericana*, Frankfurt\Madrid, v. 1, n. 2, p. 99-123, 2003.

SOARES, E. C. *Anaphors in discourse: anaphoric subjects in Brazilian Portuguese*. 2017. 482 f. Tese (Doutorado em Linguística) – Faculdade de Letras, University Sorbonne, 2017.

SOARES, E. C.; MILLER, P.; HEMFORTH, B. The effect of verbal agreement marking on the use of null and overt subjects. *Forum lingüístico*, Florianópolis, v.16, n.1, p. 3579-3600, 2019. DOI: <https://doi.org/10.5007/1984-8412.2019v16n1p3579>

SOARES, E. C.; MILLER, P.; HEMFORTH, B. The effect of semantic and discourse features on the use of null and overt subjects: a quantitative study of third person subjects in Brazilian Portuguese. *DELTA*, São Paulo, v. 36, p. 1-38, 2020. DOI: <http://dx.doi.org/10.1590/1678-460X2020360107>

TARALLO, F. *Relativization strategies in Brazilian Portuguese*. 1983. 270 f. Tese (Doutorado em Linguística) – Faculdade Letras, University of Pennsylvania, 1983.

TARALLO, F. Sobre a alegada origem crioula do português brasileiro: mudanças sintáticas aleatórias. In: ROBERTS, I.; KATO, M. A. (orgs.). *Português brasileiro: uma viagem diacrônica*. São Paulo: Contexto, 2018. p. 29-54.



“Por uma Sociolinguística Românica ‘Paramétrica’” – relendo Tarallo 1987 e virando a página

“*For a ‘Parametric’ Romance Sociolinguistics*” – revisiting Tarallo 1987 and turning the page

Maria Eugenia Lammoglia Duarte

Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro / Brasil

eugenia@letras.ufrj.br

<http://orcid.org/0000-0001-8329-1226>

Eduardo Patrick Rezende dos Reis

Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro / Brasil

eduardorezende@letras.ufrj.br

<https://orcid.org/0000-0002-5049-4200>

Resumo: Este trabalho revisita a proposta de Tarallo, feita em 1987, que observou os resultados de pesquisas variacionistas (WEINREICH; LABOV; HERZOG, 1968) à luz da Teoria de Princípios e Parâmetros (CHOMSKY, 1981), num momento em que esse modelo começava a se desenvolver. A proposta, tida como herética, já que a Teoria Gerativa não olhava, naquele momento, para a mudança linguística (e nem deveria olhar, visto que naquele momento buscava depreender os princípios invariáveis da Gramática Universal), acabou por se mostrar importante no estudo da mudança sintática no português brasileiro, uma vez que os dois modelos teóricos não estavam competindo, mas se complementando. Além de rever a proposta de Tarallo, que tem contribuído para o estudo da mudança sintática, mostramos como se desenvolve uma pesquisa em Sociolinguística utilizando como componente teórico a teoria de Princípios e Parâmetros. Mostramos ainda que os resultados dessa associação, posta em prática no Brasil a partir dos trabalhos de Tarallo e Kato e seus orientandos, e aqui ilustrados por uma análise contrastiva do português europeu e brasileiro contemporâneos, têm sido reconhecidos por eminentes gerativistas interessados no curso da mudança linguística de longa duração.

Palavras-chave: variação e mudança linguística; teoria de princípios e parâmetros; sociolinguística românica paramétrica; Fernando Tarallo; sujeito nulo; análise contrastiva PB-PE.

Abstract: This article revisits Tarallo's (1987) proposal observing variationist results, using the Theory of Language Variation and Change (WEINREICH; LABOV; HERZOG, 1968) in the light of the emerging Principles and Parameters Theory (CHOMSKY, 1981). His proposal, then considered heretical, because generative theory was not concerned about language change then (and it should not have been, because its interest then was then to find invariable principles of Universal Grammar), ended up as an important tool to study syntactic change in Brazilian Portuguese. We would soon realize that the two models were not in competition; on the contrary, they were complementary: the Theory of Language Variation and Change had to be associated with a linguistic theory. As an example of such association, proposed by Tarallo and continued with his work with Mary Kato and their students, we present the advantages of this combination in a contemporary analysis of Brazilian and European Portuguese, recognized by eminent generativists as an important tool to pursue the syntactic change in the long term.

Keywords: language variation and change; principles and parameters theory; parametric romance sociolinguistics; Fernando Tarallo; contrastive analysis BP-EP

Recebido em 26 de fevereiro de 2022

Aceito em 16 de maio de 2022

1 Introdução

O principal objetivo deste artigo é mostrar que diferentes modelos teóricos podem ser usados para o estudo da mudança linguística, a depender do propósito que o analista tem em mente, sem esquecer que um modelo de estudo da mudança não prescinde de uma teoria linguística. Para tanto, fazemos uma releitura do artigo de Fernando Tarallo que saiu nos *Cadernos de Linguística e Teoria da Literatura*, Revista do antigo Departamento de Linguística e Teoria da Literatura da Faculdade de Letras da UFMG, publicada entre 1978 e 1994, certamente uma das pioneiras entre as revistas que tiveram um papel fundamental no desenvolvimento dos estudos linguísticos no Brasil. O volume 13, *Ensaio de Linguística*, publicado em 1987, traz o artigo que será o ponto de partida do nosso trabalho. Na

seção 2, fazemos a nossa releitura do artigo de Tarallo e tentamos situar o texto no momento histórico em que foi escrito, além de mostrar como o texto deveria ser entendido hoje; em seguida, apresentamos, na seção 3, como a pesquisa sobre a mudança sintática nos moldes da Teoria da Variação e Mudança (TVM) é posta em prática, isto é, como se inicia uma investigação dentro da perspectiva paramétrica a que Tarallo se referia e que àquela época ainda não era possível justamente porque a Teoria de Princípios e Parâmetros (TP&P) estava em plena infância, como diz o autor; em 4, mostramos os resultados de uma pesquisa que conjuga a TVM com uma necessária (e óbvia) teoria gramatical, no presente caso, a TP&P; a seção 5 traz alguns efeitos colaterais ou subprodutos da mudança em curso no português brasileiro, que talvez passassem despercebidos sem uma adequada teoria gramatical associada à TVM. Em 6, retomamos a polêmica que envolveu a proposta de Tarallo (1987), para, desta vez, virar mesmo a página, ou seja, deixar de explicar repetidamente, mesmo depois de 40 anos do libelo de Tarallo e de tantos estudos realizados, que as descrições gerativistas são perfeitamente adequadas para um estudo da mudança paramétrica à luz da TVM.

2 Relendo Tarallo (1987)

Em primeiro lugar, é preciso destacar que o texto de Tarallo (1987) deve ser lido e entendido dentro do contexto histórico em que foi escrito: o autor concluíra quatro anos antes sua tese de doutorado em Sociolinguística na Universidade da Pennsylvania e fica bem claro em sua tese que, por trás da aplicação do modelo da Sociolinguística Variacionista (ou Teoria da Variação e Mudança), é possível enxergar uma base teórica gerativista, seja no tratamento da estrutura das orações relativas, seja na discussão sobre o estatuto sintático do pronome relativo, para ele um complementizador,¹ entre outras reflexões. O autor teve, ainda nos anos 1980, a oportunidade de assistir, junto com Mary Kato, a conferências de Rizzi na Universidade de Nova York sobre “a nova sintaxe comparativa

¹ Não entraremos na discussão acerca do estatuto do “que” nas chamadas subordinadas copadoras. Muito se tem escrito a esse respeito. Kato (2018) argumenta em favor do pronome relativo, que teria movimento mais curto na sintaxe, já que retoma um tópico marcado – elemento proeminente na subordinada aparecendo, em geral, sem a preposição quando na estrutura de base é um constituinte oblíquo.

e a teoria de Princípios e Parâmetros”. É verdade que, naquela década, a Teoria de Princípios e Parâmetros (CHOMSKY, 1981) estava “em plena infância”, como diz o autor, e seu texto à época consistiu em **ver resultados de pesquisas variacionistas de fenômenos analisados por outros estudiosos da Teoria da Variação e Mudança (TVM)** e extrair dali generalizações formais sobre os resultados empíricos obtidos para essas línguas românicas à luz das primeiras ideias da Teoria de Princípios e Parâmetros (P&P).

No clássico artigo, o autor observa uma série de resultados de estudos empíricos sobre variáveis fonológicas e sintáticas no grupo românico. No caso das variáveis fonológicas, um exame minucioso dos resultados de análises sobre o espanhol, o francês e o português brasileiro, particularmente sobre a erosão de certos segmentos consonantais, permite ao autor concluir que “por detrás dessa precisão estatística há uma dimensão maior a ser percebida: as pistas que mais frequentemente justificam o encaminhamento de um sistema variável e mutante para uma e não outra direção” (TARALLO, 1987, p. 69), sugerindo claramente “parâmetros” de variação.

No caso das variáveis sintáticas, apesar de muito menos estudadas pela TVM na época do artigo, Tarallo chama a atenção para o fato de ser mais transparente o paralelo entre os avanços do modelo da TVM e o modelo de P&P. Para tal paralelo, o autor toma três estudos, um para cada uma das três línguas românicas em foco, sobre uma das principais propriedades associadas às línguas românicas positivamente marcadas em relação ao Parâmetro do Sujeito Nulo (PSN): a chamada “inversão livre do sujeito” (cf. CHOMSKY, 1981; RIZZI, 1982), presente no espanhol, no português europeu e ainda possível no português brasileiro, com muitas restrições, como mostram inúmeros trabalhos cuja discussão foge aos limites deste artigo, mas ausente no francês, um sistema negativamente marcado em relação ao PSN. O rótulo “inversão livre”, inicialmente usado para nomear uma das propriedades das LSNs, como o italiano, o espanhol e o português europeu, não é muito feliz, uma vez que não se trata de uma inversão propriamente “livre”, mas de uma inversão obrigatória nas LSN prototípicas do grupo românico. Ela ocorre nas sentenças apresentativas, ou seja, aquelas que expressam um juízo tético, como (1), ou um elemento novo, com valor de foco, como em (2):

- (1) Telefonou a Maria (em resposta a “Alguma novidade? / O que aconteceu?”)
- (2) Telefonou a Maria (em resposta a “Alguém telefonou?”)

A ordem VS não sofre restrições severas quanto à transitividade verbal no italiano, no espanhol peninsular e sulamericano e no português europeu, mas, no caso de verbos com mais de um argumento, um deles aparece, em geral, em forma de clítico, o que torna o predicado mais leve:

- (3) Comeu-a a Maria (em resposta a “Quem comeu a torta?”)

Ainda em relação à ordem VS, naturalmente nos sistemas que a licenciam, sujeitos “pesados” tendem a aparecer pospostos (4), um caso que se afasta da inversão livre ilustrada em (1)-(2) acima:

- (4) Foram indicados às bolsas todos os candidatos que tinham obtido notas acima de nove.

Em outras palavras, observando o francês contemporâneo, uma gramática que já não aceita sujeitos nulos nem a ordem VS, ainda raramente licenciados no francês médio, Tarallo observou numa análise de Dubuisson (1981) contextos que ainda permitem VS desde que um elemento desencadeador apareça em primeira posição ou na ocorrência de sujeitos pesados, como ilustrado em (4) acima. Na falta de um desencadeador, geralmente um adjunto, ou de um sujeito pesado, modificado por uma relativa, as sentenças não são gramaticais em francês (TARALLO, 1987, p. 70):

- (5)a. **Au bout** de la rue brille une lumière.
‘No fim da rua brilha uma luz’

b. *Brille uma lumière au bout de la rue.
‘Brilha uma luz no fim da rua’
- (6)a. Sont tenus de présenter d’excuse au professeur tous les enfants qui ont été absents de l’école plus d’une demi-journée.

‘São obrigados a apresentar uma justificativa ao professor todos os alunos que estiveram ausentes da escola por mais de meio turno.

b. *Sont tenus de présenter d’excuse au professeur tous les enfants

Estudos como esses levam Tarallo a perguntar: “até que ponto os resultados obtidos pela pesquisa variacionista permitem identificar uma redefinição do Parâmetro e um realinhamento de suas propriedades?” (TARALLO, 1987, p. 70). De fato, Tarallo estava levantando hipóteses em relação ao conjunto de propriedades que estavam sendo elencadas, na tentativa de buscar os parâmetros de variação entre as línguas, uma vez propostos os princípios invariáveis, comuns a todas. Não se pensava então, no âmbito de TP&P, que um determinado valor paramétrico pudesse mudar ao longo do tempo. Mas, lendo os resultados que lhe caíam às mãos, Tarallo chegou a pensar na redefinição e realinhamento de propriedades associadas a uma determinada marcação paramétrica. Assim, se, algum dia, o francês viesse a se tornar novamente uma língua [+Sujeito Nulo], os mesmos fatores que regulam a ordem VS no italiano, no espanhol peninsular e sulamericano e no português europeu iriam reger a ordem VS no francês. Se, ao contrário, essas três línguas viessem a deixar de ser sistemas [+ Sujeito Nulo], certamente os verbos monoargumentais, a presença de um desencadeador e o peso do sujeito seriam os ambientes mais resistentes à mudança, como mostram os dados do francês, e, podemos acrescentar: como mostram hoje os dados do português do Brasil, em que a ordem VS vem sendo perdida, mas resiste com verbos monoargumentais, particularmente os inacusativos, tanto em sentenças declarativas quanto em interrogativas-Q, desde que o SD sujeito seja lexical (não pronominal). Além disso, a presença de um desencadeador em primeira posição e o peso do SD são igualmente importantes (cf. SANTOS; SOARES DA SILVA, 2012, sobre as declarativas; e DUARTE, 1992; NICOLAU DE PAULA, 2016; PINHEIRO; MARINS, 2012, entre outros, com resultados empíricos em relação às interrogativas-Q.)

A proposta de Tarallo viria a se concretizar no clássico artigo de 1989, reeditado em 2007, que confirma o casamento entre a TVM e a Teoria de P&P, quando Tarallo e Kato publicam “Harmonia trans-sistêmica variação intra e inter-linguística”. O texto de 1987 é aí retomado com maior profundidade e procura compatibilizar a linguística

de “propriedades paramétricas” com a linguística “de probabilidades”, que hoje poderíamos substituir por linguística “de pesos relativos”. O modelo estatístico utilizado pela metodologia variacionista que indica a “probabilidade” de aplicação de uma regra (tendo 0,5 como um ponto neutro) foi substituído por um modelo logístico de “pesos relativos”, capaz de superar os problemas de análise causados pelas frequências brutas. Segundo Naro (2003, p. 20), esse modelo logístico, introduzido por Pascale Rousseau e David Sankoff em 1978, “engloba as boas propriedades dos modelos anteriores, substituindo-os em qualquer análise de dados”. Em princípio, os valores absolutos dos pesos relativos calculados **não têm significância analítica**; o que importa é **a sua ordenação**, sendo justamente por isso que se deve preferir o uso do termo “peso relativo”. Assim, os resultados estatísticos das análises² apresentam (ou deveriam apresentar) a “relação” entre os pesos obtidos - quanto maior a distância entre eles, mais relevante é o que alcança o peso mais alto em relação ao que alcança o mais baixo no que diz respeito ao seu efeito no uso da variante tomada como valor de aplicação. É muito frequente verificar que inúmeros trabalhos deixam de mostrar essa “relação” entre os pesos, apontando apenas o mais alto. Um peso isolado não faz o menor sentido; apenas quando “relacionado” aos demais dentro de um grupo de fatores ele nos diz alguma coisa.

Tarallo nunca afirmou que o modelo da TVM se confundia com uma teoria linguística estruturalista ou gerativista, por exemplo (cf. DUARTE, 2016; 2019a). Essa confusão gerou inúmeras discussões acerca da incomensurabilidade entre o modelo de estudo da mudança, de um lado, e as descrições que essas teorias formais fornecem justamente ao variacionista, que delas se utiliza para pôr de pé uma análise da mudança em curso. Pagotto (2000)³, que retoma a polêmica iniciada com Borges Neto e Müller (1987), e repetida em publicações que não atualizam os

² Apesar da insistência dos que trabalham com a Teoria da Variação e Mudança para que as análises apresentem a “relação” entre os pesos obtidos, a maioria dos trabalhos deixa de mostrar essa “relação” entre os pesos. Não tem sentido mencionar o peso mais alto simplesmente sem apontar sua relação com os demais. Isolado, seja ele qual for, não faz qualquer sentido.

³ O texto de Pagotto (2000), que infelizmente não teve ampla circulação, estará em breve disponível numa edição da Revista Diadorim, n. 23, volume especial – PrInt, 2021. Sua leitura é indispensável aos interessados na questão da (in)comensurabilidade dos dois modelos.

questionamentos à luz de resultados da aplicação da Sociolinguística com a descrição formal gerativista, chama a atenção para o fato de que a TVM e as teorias linguísticas lidam com entidades de outra natureza, dado o caráter imanentista dessas, e “por essa razão, não é possível que a Teoria da Variação e Mudança se coloque como concorrente, seja do Estruturalismo, seja da Teoria Gerativa. Eles estão em mundos diferentes” (PAGOTTO, 2000, p.54). O autor, no entanto, mesmo reconhecendo a incomensurabilidade, mostra que a TVM se apropria do conhecimento acumulado por essas teorias, particularmente quando delas se utiliza para propor os grupos de fatores que podem favorecer ou refrear a variação. Em seu artigo, Pagotto enfatizava que a TVM tinha uma importante função à sua frente: “avançar na sua formulação teórica, definindo mais claramente o estatuto dos grupos de fatores, o que significava, segundo ele, “definir mais claramente o estatuto dos grupos de fatores, a fim de que o nível explicativo da teoria avance com relação ao nível descritivo” (PAGOTTO, 2000, p. 58).

Aliás, a falta de uma explicitação acerca da origem e da justificativa do estabelecimento de grupos de fatores que entram numa análise variacionista é um dos motivos de severas críticas aos estudiosos da variação. Tratando das análises de variáveis sintáticas (mas podendo ter aplicação a quaisquer outras variáveis linguísticas), Alison Henry em 2002 (aqui citada na reedição de 2006) corrobora as palavras de Pagotto:

No âmbito dos estudos variacionistas tem havido pouca discussão sobre **que tipo de fatores** pode afetar a escolha de variantes ou sobre **como determinados fatores são escolhidos** para a análise num caso específico. Tipicamente os fatores selecionados para alimentar uma análise de estatística aparecem sem uma discussão extensiva e não fica claro como, **excetuando as intuições do pesquisador**, se chegou a eles ou se há quaisquer restrições sobre o que pode ser um fator numa determinada análise (HENRY, 2006 [2002], p. 277)⁴.

⁴ Do original: “Within variationist studies, however, there has been little discussion of what type of factors can affect the choice of variants, or of how the particular factors are chosen for analysing any given case. Typically the factors chosen for entry into VARBRUL analysis appear without extensive discussion, and it is not clear how, apart from the intuitions of the researcher, these are arrived at or whether there are any constraints on what can be a factor here” (HENRY 2006, p. 277).

Assim, já estamos entrando num tempo em que a TP&P não se encontra mais em plena infância e a mudança linguística passa fazer a parte da agenda de muitos gerativistas a partir dos estudos de Kroch (1989 e. o.), quando o modelo de competição de gramáticas, no estudo do curso da mudança, vem se contrapor à ideia de que a variação se encontra dentro da mesma gramática. Só o desenvolvimento da TP&P permitiria tal perspectiva.

De todo modo, já não estamos igualmente no momento histórico que levou Tarallo a propor uma “leitura” de diversos trabalhos variacionistas, e a reafirmar sua “intenção de desmistificar alguns pré-conceitos (ou serão eles preconceitos?) de que as análises empíricas teriam o propósito de falsear as análises de modelos psicológicos, de recuperar a sistematicidade na diversidade, de que análises empirista e racionalista de um mesmo fenômeno não poderiam ser compatíveis entre si” (TARALLO, 1986, p. 20).

Diante dos argumentos e fatos exemplos que o autor oferece no texto, as críticas então não procedem. O certo, porém, é que o trabalho do variacionista que utiliza as descrições da TP&P não é mais fazer uma leitura de estudos variacionistas elaborados sob outras perspectivas e dali levantar hipóteses, fazer previsões sobre o que poderia acontecer diante de tais e tais circunstâncias. Não! O variacionista que escolhe estudar uma mudança sintática observada no sistema tem hoje em suas mãos refinadas e abundantes descrições gerativistas, e Tarallo pôde acompanhar o início desse refinamento, que já se faz ver nos trabalhos por ele orientados e em suas parecerias com Kato, aqui mencionadas. Esse refinamento permite ao variacionista responder às questões colocadas por Pagotto (2000) e Henry (2006 [2002]), mostrando o que deu origem às suas hipóteses, de onde vieram seus grupos de fatores, como é o curso da mudança, que evidências do “encaixamento” da mudança aparecem no sistema. Este é um dos problemas mais caros ao modelo da TVM. É sempre bom lembrar as palavras de Weinreich, Labov e Herzog (1968), de acordo com as quais os linguistas desconfiam do surgimento de novas formas no sistema que possam ser atribuídas ao acaso. Uma mudança num determinado ponto da estrutura provoca efeitos colaterais, e somente o analista que se utiliza ou se “apropria” de uma teoria linguística que, de fato, o oriente na identificação de uma nova forma no sistema, que aparece de forma não acidental, poderá relacioná-la a um efeito da mudança. Só assim se pode avançar na formulação teórica e, conseqüentemente fazer com que o nível explicativo avance em relação ao descritivo, como lembra Pagotto.

Para que possamos ilustrar como uma pesquisa sociolinguística evolui a partir de boas descrições formalistas e o quanto seus resultados podem contribuir para o que Tarallo se refere como redefinição e realinhamento de propriedades associadas a uma determinada marcação paramétrica, e até mesmo revelar uma remarcação do valor positivo ou negativo em relação a um parâmetro⁵, vejamos como uma descrição do comportamento do sujeito pronominal numa língua [+Sujeito Nulo] prototípica, como o italiano, pode contribuir, e, nós diríamos, se mostrar indispensável para entender o comportamento das línguas românicas [+Sujeito Nulo], e, a partir desse comportamento, iluminar os caminhos que percorre um sistema do mesmo grupo quando apresenta traços que se afastam do conjunto de propriedades que caracterizam tal marcação paramétrica, que é o caso do PB. Como os passos da pesquisa sociolinguística são conhecidos pelos eventuais leitores interessados neste artigo, passaremos, na próxima seção, a mostrar de onde vêm os grupos de fatores que devem ser levados em conta para pôr em prática a pesquisa em variação e mudança sintática à luz do que conhecemos sobre as línguas românicas de sujeito nulo e não nulo.

3 Alguns pontos de partida para o levantamento dos grupos de fatores linguísticos envolvendo a perda do sujeito nulo no PB

O primeiro passo, nem sempre explicitado numa pesquisa sociolinguística (ou variacionista), é uma descrição, a mais refinada possível, do fenômeno variável em estudo. A alternância entre pronomes nulos e plenos que hoje se vê no português brasileiro não é definitivamente uma característica de línguas *pro-drop* ou [+Sujeito Nulo] do grupo românico. Um texto clássico de Calabrese (1986) chama

⁵ Com o advento do Programa Minimalista (CHOMSKY, 1995), é revisto o conceito de Parâmetro, que passa a ser concebido como traços morfológicos presentes no Léxico, a conjectura *Borer-Chomsky* (BAKER, 2008). A variação paramétrica, portanto, encontra-se associada à presença ou à ausência de traços formais contidos nos núcleos funcionais, conservando a natureza binária desses expoentes (agora) morfológicos. Anos mais tarde, Biberauer e Roberts (2015) propõem uma taxionomia para os Parâmetros, que se subdividem em 4 tipos hierarquicamente organizados: (a) macroparâmetros; (b) mesoparâmetros; (c) microparâmetros; (d) nanoparâmetros. Para uma discussão ontológica e empírica mais aprofundada sobre Parâmetros e variação paramétrica, conferir Roberts (2019).

a atenção para a existência de uma complementaridade entre sujeitos pronominais nulos e expressos em italiano⁶. Segundo o autor, o sujeito nulo é obrigatório quando o referente é esperado (em suas palavras, o Tema de uma predicação); o uso de um pronome pleno em tal caso (a que o autor se refere como “tônico”) implica uma referência disjunta ou a inaceitabilidade da sentença, como mostra (7) (os exemplos de (7) a (12), excetuando (10), que é nosso foram extraídos do texto citado):

(7)a. Quando **Carlo**_i ha pichiato **Antonio**_k Ø_i / **lui**_k era ubriaco.
Quando Carlo bateu no Antonio Ø_i / **ele**_k estava bêbado

b. **Mario**_i si è spaventato dopo che Ø_i / ***lui**_k ha visto quel film.
Maria se assustou quando Ø_i / ***ele**_k viu aquele filme

c. Dopo che Ø_i / ***lui**_i ha visto quel film, **Mario**_i si è spaventato.
Quando Ø_i / ***ele**_i viu aquele filme, Mario_i se assustou

Os exemplos deixam duas propriedades claras: há uma complementaridade entre pronomes nulos e expressos, independentemente de haver c-comando entre os sujeitos, que ocorre em (7b), em que o sujeito da oração matriz precede o da subordinada, e não haver c-comando, como em (7a,c), em que a subordinada precede a matriz. Em outras palavras, um sujeito numa subordinada anteposta não c-comanda o sujeito da principal posposta; isso, porém, não lhe retira o estatuto de Tema – referente esperado – que leva ao sujeito nulo na matriz subsequente.

Da mesma forma, um pronome expresso (como em 7c) e novamente ilustrado em (8) a seguir, não pode preceder seu referente em subordinadas adverbiais antepostas à matriz:

(8) Quando Ø_i / ***lui**_i lavora, Gianni_i non beve.
Quando Ø_i / ***ele**_i trabalha, Gianni_i não bebe.

Se, no entanto, a subordinada não for uma adverbial, o pronome expresso pode preceder seu antecedente sem causar agramaticalidade:

⁶ Tanto Calabrese (1986) quanto Di Eugenio (1990), que será referida ainda nesta seção, utilizam os termos pronomes fracos e fortes, para se referirem a nulos e expressos, respectivamente. Levando em conta estudos posteriores, que viriam a distinguir pronomes nulos, fracos (clíticos) e fortes (cf. CARDINALETTI; STARKE, 1998; KATO, 1999, entre outros) preferimos utilizar aqui apenas pronomes nulos e expressos.

- (9) Le persone che **lui**_i ha aiutato sono convinte che Gianni_i è una buona persona.
As pessoas [que **ele**_i ajudou] estão convencidas de que Gianni_i é uma boa pessoa.

Como a oração subordinada relativa é um adjunto do sujeito da oração matriz, não fica claro o que ocorreria se a relativa viesse após uma matriz completa, tendo no sujeito da matriz um Tema, como em (10a), ou (10b), exemplos nossos, em que o sujeito da relativa tem seu correferente numa oração que se encontra em outro período:

- (10) a. Gianni_i no conosce le persone che Ø_i / **lui**_i há aiutato.
Gianni_i não conhece as pessoas que Ø_i / **ele**_i ajudou.
- b. Gianni mi há dato um libro. Ieri ho trovato il libro que Ø_i / **lui**_i mi há dato
Gianni me deu um livro. Ontem encontrei o livro que Ø_i / **ele**_i me deu.

Os exemplos em (10)⁷ foram apresentados a nativos, que aceitam tanto o sujeito nulo quanto o expresso, embora prefiram um sujeito nulo, porque “o pronome não é necessário”.

Ainda com base em Calabrese, vemos que o sujeito de uma predicação secundária não é um Tema e, por conseguinte, não é um referente esperado, o que implicará o uso do pronome expresso correferente com o sujeito dessa predicação:

- (11) Mentre il dottore_i visitava [**Maria**_k incinta] *pro*_i *_k / *lei*_k *_i canticchiava.
Enquanto o médico_i visitava Maria_k grávida Ø_i *_k / *ela*_k *_i cantava

O mesmo fenômeno observado em sentenças subordinadas pode ser visto operando entre sentenças no discurso. A hipótese de Calabrese para explicar tal semelhança é a de que as sequências de enunciados que

⁷ Agradeço a Carolina Serra, que faz estágio de pós-doc na Universidade de Lecce, Itália, por testar a aceitabilidade de algumas sentenças com falantes nativos do italiano. Eles aceitam tanto o pronome nulo como o expresso nesses contextos, mas parecem preferir o sujeito nulo. A maior variação entre nulos e expressos nesse contexto é atestada por Marins (2009), seguramente a primeira análise variacionista do italiano falado que conhecemos.

compõem o discurso são sintaticamente irmãs umas das outras, como mostra (12):

- (12) **Carlo**_i è entrato. **Mario**_k si è alzato. $\emptyset_{*i/k}$ ha parlato.
Carlo_i entrou. Mario_k se levantou. $\emptyset_{*i/k}$ Falou.

Em 1990, Di Eugenio traz contribuições adicionais ao trabalho de Calabrese. Além da distribuição complementar entre pronomes nulos e expressos, sendo aqueles correferentes com um Tema, como ilustramos acima, a autora aponta outros traços que podem levar ao sujeito nulo além de seu referente ser um tema (um sujeito) de uma predicação em sentença matriz ou na encaixada, ou numa sentença adjacente, como em (12) acima. Com base num manuscrito de 1989, *Towards a computational theory of discourse interpretation*, de Grosz, Joshi e Weinstein, referido como Centering Theory, a autora propõe dar conta dos pronomes nulos e expressos no italiano. A Centering Theory “tem a ver com a coerência local: ela tenta determinar a entidade à qual um enunciado está mais diretamente relacionado” (DI EUGENIO, 1990, p. 271). Segundo os autores citados “a coerência discursiva é uma medida da carga de inferência que uma certa sequência de enunciados de um determinado discurso impõe sobre o ouvinte” (DI EUGENIO, 1990, p. 272). Di Eugenio acrescenta que os autores não deixam claro se sua teoria está mais voltada à produção ou à compreensão do discurso; seu ponto de vista, porém, é a teoria da produção do discurso. Seu exemplo reproduzido em (13) (exemplo (4) do original, p. 273), mostra uma sequência de enunciados:

- (13) E1. Maria_i voleva andare al mare.
Maria_i queria ir à praia
E2. \emptyset_i Telefonò a Giovanni_k.
 \emptyset_i Telefonou para Giovanni_k.
E3. a. \emptyset_i Si arrabiò perche \emptyset_i non lo_k trovò a casa.
 \emptyset_i Se aborreceu porque \emptyset_i **não** o_k encontrou em casa.
b. $\emptyset_i / ? \emptyset_k$ Si arrabiò perche \emptyset_k stava dormendo.
 $\emptyset_i / ? \emptyset_k$ aborreceu porque \emptyset_k estava dormindo.
c. **Lui**_k se arrabiò perche \emptyset_k stava dormendo.
Ele_k se aborreceu porque \emptyset_k estava dormindo.
d. \emptyset_k Si é arrabiato perche \emptyset_k stava dormendo.
 \emptyset_k Se aborreceu porque \emptyset_k estava dormindo

O sujeito nulo no enunciado 2 (E2) é identificado pelo Tema, Maria (E1), e em E3a, Maria continua a ser o elemento ranqueado num nível mais alto do ranking referencial, mantendo a continuação de sua centralidade, e a coerência discursiva é perfeita. Em E3b, a interpretação do sujeito nulo da matriz tendo Maria como correferente é mais natural, enquanto o sujeito nulo da subordinada em E3b, por razões pragmáticas, só pode ser Giovanni. Interpretar o sujeito nulo na oração principal, poderia se referir a Giovanni, mas, pelas mesmas razões pragmáticas, tal correferência soaria menos coerente. Em E3c, o falante realiza uma bem sucedida mudança de centralidade, ao usar o pronome expresso para retomar Giovanni, visto que sua posição não era a mais alta no ranking referencial (esse uso do pronome expresso é normalmente nomeado mudança de referência ‘switch reference’). Considerando agora o enunciado em E3d, o sujeito nulo poderia ser interpretado como em E3b, referindo-se a Maria, mas é, ao contrário, interpretado como se referindo a Giovanni, não por questões pragmáticas, mas por uma questão morfológica: no passado composto, em italiano, o particípio concorda em gênero e número com o sujeito, o que impede a correferência com Maria.

Assim, a autora chega a algumas generalizações sobre o italiano: o falante codifica com um sujeito nulo uma continuação de centralidade (nos termos de Calabrese (1986), a manutenção do Tema); uma mudança na centralidade é codificada com um sujeito expresso; entretanto, questões pragmáticas e/ou questões relacionadas a traços morfológicos ou morfossintáticos podem reverter essa generalização, como vimos em E3b e E3d, respectivamente.

Outro exemplo que revela a influência de um traço morfossintático é brilhantemente ilustrado com o uso de clíticos, como nos exemplos em 5, da autora (p. 274), aqui apresentados em (14):

- (14) E1. Maria_i é arrabiata com Giorgio_k
 Maria está com raiva de Giorgio
 E2. a. Ø_j **non** vuole piú parlargli_k
 Ø_j não quer mais falar com ele
 b. *Ø_j non vuole piú parlarle_i
 (ela) não quer mais falar com ela
 c. Ø_k non **le**_i vuole piú parlare
 (ele) não lhe quer mais falar = não quer mais falar com ela

Enquanto E2a o sujeito nulo é perfeitamente identificado pelo elemento mais alto no ranking de referentes (o elemento central), em E2b, o ouvinte é forçado a interpretar o sujeito como tendo como referente

o clítico no feminino, 3^a. pessoa do singular, o que torna a sentença pragmaticamente incoerente.⁸ Se, porém, o clítico se move (sobe) para a esquerda do auxiliar modal ‘vuole’ (quer), o ouvinte exclui naturalmente Maria como correferente do sujeito nulo.

Outras predições apresentadas pela autora, ainda acerca do papel de elementos morfossintáticos e de questões pragmáticas, que desviam um sujeito nulo do referente “esperado” podem ser vistas no exemplo 6, na p. 275, aqui citado em (15):

- (15) E1. Luisa_i há laciato suo marito_k
 Luísa deixou seu marido
 E2.* $\emptyset_{j,i}$ \emptyset_k picchiava i bambini e si ubriacava
 * $\emptyset_{j,i}$ \emptyset_k batia nas crianças e se embriagava

Na falta de uma pista sintática que licencie o sujeito nulo no enunciado E2, a sua função de explicar por que Luísa abandonou seu marido é que licencia o sujeito nulo.

O que se pode concluir acerca do italiano (e, podemos acrescentar, do espanhol peninsular e do português europeu) é que numa gramática românica [+Sujeito Nulo], muitos fatores, como fatores de natureza morfossintática e pragmática, além do referente esperado no sentido mais estrutural (Tema), apontado por Calabrese (1986), devem ser levados em conta no favorecimento do sujeito nulo. O texto de Di Eugenio, além de apontar todos esses contextos, traz uma informação que é de grande relevância para o sistema pronominal do italiano, além do espanhol e do português europeu: a ausência de pronomes pessoais com o traço [-animado].⁹ Em tais casos, são usados os demonstrativos *esso/essa* “embora não muito usados no italiano corrente” (DI EUGENIO, 1990, p. 271).

⁸ A presença de outros elementos no contexto discursivo, como o clítico aqui apontado, já era citada por Fernandes Soriano (1989) como favorecedores do sujeito nulo no espanhol peninsular, que tornaria o enunciado “mais natural”, sendo um pronome expresso considerado um pronome forte (i. e. com valor de foco):

- (i) É necessario che \emptyset_i t_i vesta bene.
 É necessário que \emptyset_i te vistas bem.

⁹ A ausência de pronomes pessoais no espanhol peninsular e sulamericano é constatada. Quanto ao português europeu, análises empíricas recentes revelam um índice muito baixo de pronomes pessoais com o traço [-animado], como mostraremos neste artigo. Em algumas variedades do espanhol caribenho, entretanto, temos evidências de pronomes pessoais com esse traço, entre outras mudanças envolvendo o Parâmetro do Sujeito Nulo (MARTINEZ-SANZ, 2011).

Di Eugenio termina seu texto mencionando um tipo de sujeito que tem como referente um segmento maior do discurso e sobre o qual sua teoria da centralidade não tem muito a dizer. Ela acrescenta que tais sujeitos, retomados por um demonstrativo ou nulos no italiano, mereceriam a atenção de uma teoria sobre a coerência discursiva. Na verdade, Halliday e Hasan (1979), bem antes do texto de Di Eugenio, trataram desses sujeitos, exatamente em seu livro *Cohesion in English*, nomeando-os como ‘extended reference subjects’, traduzido em português por Paredes Silva (1985), como ‘sujeitos de referência estendida’, justamente para mostrar o uso do verbo *ser* e do demonstrativo em variação com um sujeito nulo em função coesiva no português. Esse sujeito, claramente um sujeito neutro ou proposicional, viria a ser incluído por Cyrino, Duarte e Kato (2000) numa hierarquia de referencialidade, que é altamente relevante para a pronominalização. Utilizando as análises diacrônicas de Cyrino (1993) para o objeto nulo e de Duarte (1993) para o sujeito pronominal, as autoras mostram que o percurso da implementação do sujeito pronominal expresso e o do objeto nulo seguem caminhos opostos segundo a hierarquia proposta: sujeitos com o traço inerentemente [+ani], como os de 1ª. e 2ª pessoas, se tornam mais rapidamente plenos, seguidos dos de 3ª pessoa, quando há a interação de traços [+/-ani], sendo ainda mais resistentes quando seu antecedente é um sujeito neutro (ou proposicional, ou, como foram inicialmente nomeados pelos funcionalistas clássicos, de “referência estendida”. O percurso do objeto nulo segue caminho oposto, começando a se implementar pelo de menor referencialidade, o objeto proposicional.¹⁰

Concluindo esta seção, é importante destacar que o estabelecimento dos grupos de fatores para uma análise da expressão do sujeito pronominal deve contemplar uma grande variedade de propriedades para chegar a uma explicação para a remarcação do valor positivo ou negativo de um parâmetro. Esses grupos de fatores foram sendo construídos a partir da redução do paradigma flexional no PB, com a inserção de **você** e **a gente** no quadro pronominal (DUARTE, 1993). Os padrões estruturais, construídos inicialmente a partir de Calabrese (1986) por Duarte (1993;

¹⁰ Além dos trabalhos pioneiros de Paredes Silva (1985); Paredes Silva e Oliveira (2014), temos as análises diacrônicas numa perspectiva socioparamétrica de Duarte, Mourão e Santos (2012) para peças brasileiras e Guimarães (2021) para uma comparação entre a retomada dos sujeitos neutros em peças brasileiras e portuguesas.

1995) ficaram aquém do refinamento necessário, sem reunir num mesmo grupo todos os padrões ilustrados nesta seção. Evoluíram em Barbosa, Duarte e Kato (2005) e progrediram um pouco mais em Duarte e Rezende dos Reis (2018) e Duarte (2019b), quando a distinção entre c-comando e falta de c-comando foi incorporada. O desenvolvimento de um paradigma pronominal contendo pronomes fracos em substituição à flexão, já apontado nos dados de Duarte (1995), vem encontrar em Kato (1999) uma formalização indispensável à explicação de um subproduto da mudança. Outro efeito colateral da mudança apontado em Duarte (1995) é o surgimento do redobro do sujeito, ou sujeitos deslocados à esquerda e retomados por um pronome fraco, intimamente ligados à emergência o referido paradigma, estruturas não atestadas nas línguas [+Sujeito Nulo] do grupo românico.¹¹ Temos, então os desdobramentos de uma mudança mais profunda na categoria funcional flexão. Falta, porém, diante dos fatores aqui apresentados a partir de Di Eugenio, acrescentar fatores de natureza morfossintática nos grupos de análise empírica. É fato que o PB tem um quadro de clíticos muito pobre, tendo restado à aquisição, um paradigma com os pronomes dêiticos (a primeira e a segunda pessoas). Os anafóricos, acusativo e dativo, e o clítico indefinido, estão em extinção e são parcialmente recuperados pela escrita (cf. DUARTE *et al.* 2021). Não temos uma morfologia verbal rica, seja pelo predomínio de formas de 3^a. pessoa do singular, seja pela ausência de marcas de gênero e número em formas complexas com o particípio (excetuando as passivas perifrásticas), o que pode estar entre os fatores que atuam na perda do sujeito nulo referencial.

A presença de outros elementos no contexto discursivo, capazes de permitir sua identificação, torna o sujeito nulo “mais natural”, como já apontava Fernandes Soriano (1989, p. 233) para o espanhol:

- (16) É necesario che Øti vesta bene.
É necessário que Ø_i te vistas bem.

Considerando que todos os sujeitos referenciais nulos nos exemplos mostrados nesta seção podem ser expressos no português brasileiro, excetuando os casos de mudança de referência ou a interferência

¹¹ Devemos a Pontes (1985) o fato de ter apontado esta como uma das mais frequentes estruturas de tópico marcado no PB e sua relação com a proeminência do tópico no português do Brasil.

de fatores morfossintáticos, podemos concluir que, embora os sujeitos nulos ainda sejam atestados no PB, eles se encontram em variação com os expressos, que já levam vantagem em praticamente todos os contextos estruturais levantados, um panorama esperado num processo de variação em uma mudança no tempo aparente ou real. Observando os exemplos do francês, língua [-Sujeito Nulo], podemos perceber que nenhuma dessas sentenças teria necessariamente um sujeito nulo no PB; todas são gramaticais com um sujeito expresso no PB, obrigatório no francês:

- (17) a. **Je** parle bien le français.
Eu falo bem o francês.
b. **Tu** parles bien le français.
Tu falas/Você fala bem o francês.
c. **Vous** parlez bien le français. (tratamento cerimonioso ou 2pp)
O senhor/A senhora fala bem o francês.
Vocês falam bem o francês.
- (18) a. **Il** / **Elle** parle bien le français (3ª ps)¹²
Ele/Ela fala bem o francês.
b. **Ils** / **Elles** parlent bien le français. (3ª pp)
Eles/elas falam bem o francês.
- (19) a. **Jean**_i a dit qu'**il**_i parle bien le français.
O João_i disse que ele_{i/k} fala bem o francês.
b. [**Les enfants**]_i ont dit q'**ils**_i parlent bien le français
As crianças_i disseram que elas_{i/k} falam bem o francês.
- (20) a. Lorsque **Pierre**_i a frappé **Antoine**_k, **il**_i était saoul.
Quando o Pedro_i bateu no Antônio_k ele_{i/k} estava bêbado.
- (21) a. J'ai lu [**le livre**]_i. **Il**_i est très intéressant.
Eu li o livro_i. Ele_i é muito interessante.

¹² Naturalmente, os sujeitos de 3ª pessoa têm necessariamente um antecedente no contexto discursivo ou são pragmaticamente identificados. Não existem sujeitos nulos de 3ª pessoa "out of the blue" (FRASCARELLI, 2007).

- b. [La Maison]_i a efondré parce qu' elle_i était très vieille.
A casa_i desabou porque ela_i era muito velha.

Feita uma minuciosa descrição do comportamento de uma língua de sujeito nulo prototípica, como é o italiano, e o de uma língua do mesmo grupo mas que passou de mais para menos sujeito nulo, como o francês, passemos aos resultados de uma análise variacionista da mudança em curso no PB. Embora o PB ainda exiba sujeitos nulos em todos os contextos acima ilustrados em variação com sujeitos expressos, eles são cada vez mais raros, como mostraremos a seguir. Os dados nos dizem que preferimos pronomes plenos para representar nossos sujeitos referenciais e nos apontam os contextos mais resistentes aos sujeitos nulos e os que mais rapidamente cedem à implementação da mudança.

4 Aplicando o modelo da TVM para estudar sintaxe comparativa à luz da TP&P

Os resultados de uma análise recente de amostras do português europeu e brasileiro disponíveis no site www.comparaport.letas.ufrj.br, gravadas no Rio de Janeiro e em Lisboa entre 2009 e 2010¹³, são reveladores da importância da associação da TVM com a Teoria de P&P para uma análise de sintaxe comparativa e para acompanhar uma evidente remarcação em curso no valor do PSN no PB.

Esses resultados, publicados em Duarte (2019b; 2020), levantaram grupos de fatores linguísticos e extralinguísticos. Tomando os resultados para os sujeitos nulos de 3ª pessoa¹⁴, 67% para o PE e 28% para o PB, obtivemos na análise logística com o Programa Goldvarb X (SANKOFF; TAGLIAMONTE; SMITH, 2005) a seleção dos mesmos grupos de fatores como significantes para o sujeito nulo, na mesma ordem: o padrão sentencial, os feixes de traços semânticos e a estrutura do CP (Sintagma Complementizador). Isso, entretanto, não significa

¹³ VIEIRA, S. R.; BRANDÃO, S. F. CORPORAPORT: Variedades do Português em análise. Disponível em: www.corporaport.letas.ufrj.br.

¹⁴ A 1ª pessoa exibiu na amostra do PE 57% de sujeitos nulos e 70% na 2ª pessoa; o PB, por outro lado, mostra 20% e 10%, respectivamente. Neste, a segunda pessoa pode ser considerada um caso de mudança concluída, sendo os raros sujeitos nulos identificados pragmaticamente, tal como os nulos encontrados no inglês (“*Quer um café? Want some coffee?*”).

comportamento igual, mas mostra a força relativa de certos fatores em relação a outros dentro de um mesmo grupo, de extrema importância para acompanhar uma gramática em mudança. Os fatores sociais não foram considerados significativos.

Vejam na Tabela 1, os resultados para o primeiro grupo de fatores selecionados, que serão comentados com exemplos retirados das amostras. Os padrões reúnem, até certo ponto, os apontados por Calabrese e Di Eugenio na seção anterior.

Tabela 1 - Sujeito de 3ª pessoa vs padrão sentencial (valor de aplicação: sujeito nulo)

PADRÃO	PE Input: 0,756			PB Input: 0,248		
	N / T	%	P. R.	N / T	%	P. R.
1 (com c-comando)	78/83	94%	0,930	19/46	41%	0,765
1 (sem c-comando)	13/14	93%	0,854	15/116	11,5%	0,197
2 (adjacente)	402/515	78%	0,588	225/586	38%	0,646
3 (outra função)	77/153	50%	0,274	37/175	21%	0,428
4 (distante)	66/183	36%	0,183	35/241	14,5%	0,330
	range		0,747	range		0,435
	Log likelihood = -44,136 Significance = 0,000			Log likelihood = -624,928 Significance = 0,000		

Fonte: Duarte (2019b, Tabela 2, p. 109-110)

O *input* já revela a alta frequência do sujeito nulo no PE em relação ao PB. Na coluna à esquerda se encontram os padrões com e sem c-comando, que já nos revelam uma diferença crucial entre as duas variedades: no PE, a presença do referente na principal anteposta, c-comando, portanto, o sujeito da subordinada posposta, ou na subordinada anteposta, sem relação de c-comando com o sujeito da principal posposta, favorecem amplamente o sujeito nulo, com percentuais de 94% e 93%, confirmando o estatuto de LSN “consistente”, nos termos de Roberts & Holmberg (2010). O único dado com um sujeito expresso em estrutura sem c-comando (24c) parece ser um dado em que temos um sujeito preenchido por questões discursivas, como comentado na seção anterior. No PB, por outro lado, vemos dois resultados importantes. O padrão com c-comando, com 41% de sujeitos nulos, se revela como o fator de

resistência, embora já esteja abaixo da metade das ocorrências, enquanto a falta de c-comando nos mostra o sujeito nulo praticamente extinto, sendo o contexto mais favorável à perda do sujeito nulo, uma questão que nada tem a ver com contiguidade, mas com a acessibilidade sintática. Este é um fator crucial no PB no curso da mudança em direção a um sistema [-Sujeito Nulo]. Os exemplos a seguir ilustram os dois padrões, com sujeitos nulos e expressos para cada variedade. Em (22) e (23) temos as estruturas com c-comando e em (24) e (25), sem c-comando:

- (22) a. [*O pobre*]_i continua com essa mentalidade porque \emptyset_i só pensa nele. (PE) (o pobre = o pobre coitado)
 b. [*O meu filho*]_i vai ter de sair dessa escola porque *ele*_i tá no nono ano. (PE)¹⁵
- (23) a. [*Meu marido*]_i foi quase preso aí no forte porque \emptyset_i foi mergulhar. (PB)
 b. [*Escola pública*]_i nunca é boa opção porque *elas*_i são ruins. (PB)
- (24) a. Quando [*as pessoas*]_i sentem essa união do Espírito Santo, \emptyset_i tornam-se mais alegres mais participativas. (PE)
 b. Se \emptyset_i arranjava um namorado ou \emptyset_k ía com *ela*_i para ali ou para acolá, *ela*_i não tinha problemas. (PE)
- (25) a. Se [*o aluno*]_i tem problema, \emptyset_i vem pra gente conversar. (PB)
 b. Se *ele*_i tem medo, alguma coisa *ele*_i fez. (PB)

O padrão seguinte, em que temos um sujeito com seu antecedente em oração mediatamente adjacente, é o mesmo padrão a que Calabrese se refere como orações sintaticamente irmãs umas das outras, como

¹⁵ Apenas cinco dados com este padrão (6%) exibiam um sujeito expresso no PE. Podemos pensar que se trate de algum tipo de ênfase / contraste. Neste caso, o falante tem dois filhos e pode querer enfatizar que “um” (e não o outro) está no nono ano. Agradecemos ao parecerista que chamou a atenção para essa aparente inconsistência em relação ao nulo em tal contexto. Entretanto, o nulo praticamente categórico afasta o PE do PB, em que os sujeitos expressos alcançam 59%.

mostramos no exemplo em (12)¹⁶. É, como esperado, um contexto que vai atuar nas LSN como favorecedor do sujeito nulo. E é o que vemos para o PE, que exhibe 78% de sujeitos nulos, enquanto o PB alcança 38%, como podemos ver a seguir em (26) e (27):

- (26) a. [*O dono*]_i é um amigo meu. Ø_i Tem tido problemas. (PE)
 b. [*A minha mãe*]_i estava em casa e era completamente diferente. **Ela**_i conseguia perceber quando eu estudava, quando eu não estudava. (PE)
- (27) a. *Ele*_i era bem mais novinho. Ø_i Trouxe o dinheiro no bolso. (PB)
 b. [*César Maia*]_i era mais tranquilo. *Ele*_i investia mais. (PB)

Os percentuais de sujeitos nulos caem à medida que o antecedente é menos acessível, com orações intervenientes ou estando em outra função sintática e sem elementos gramaticais ou discursivos que licenciem o sujeito nulo. Ainda assim, o PE licencia 50% de sujeitos nulos, enquanto o PB alcança apenas 21%:

- (28) a. O meu marido comenta muitas vezes comigo que eles não tinham respeito. Ø_{3pp} Tinham medo d[*o pai*]_i porque Ø_i batia-lhes. (PE)
 b. Se eu (me) pudesse pôr de pé e ir lá, porque eu nunca *o*_i tinha visto quando **ele**_i cá veio. (= o Papa João Paulo II) (PE)

¹⁶ Note-se que não se trata de coordenadas com sujeitos correferentes. São orações separadas por uma curva entonacional descendente, que não caracteriza a coordenação. Neste padrão, encontramos ainda estruturas como:

(i) Eu não acho que [*os portugueses*]_i sejam muito pessimistas; acho [que Ø_i são pouco ambiciosos, não é?] (PE)

O sujeito aparece numa encaixada com o verbo *ser*, mas o verbo da matriz é um verbo epistêmico, e, em geral não obscurece a vinculação do sujeito com seu antecedente. O PB, entretanto, perde sujeitos nulos mesmo nessas estruturas. (Agradeço a Marcello Modesto por ter chamado nossa atenção para tal padrão.)

- (29) a. Porque eu preciso ter a minha linguagem formal com [meu *cliente*]_i. De repente, \emptyset_i é um cara que tem uma condição, ou um cara que usa daquele jeito. (PB)
 b. Eu não posso ter sentado do lado de [*um cara bonitinho*]_i e tal sem saber que *ele*_i era superperigoso. (PB)

Finalmente, vemos estruturas em que o referente está distante e há elementos intervenientes que podem interferir na interpretação de uma categoria vazia. Em tais casos, por razões funcionais (ou de coerência discursiva, como se refere Di Eugenio), temos os mais baixos percentuais de sujeitos nulos no PE, com 36% (31a) e no PB, com 14,5%, com o sujeito expresso amplamente preferido (31b):

- (31) a. Pronto, havia [*amigos meus*]_i, até colegas de escola, que jogavam à bola berlinde na rua. Eu não podia porque o meu pai não deixava, e pronto. Tinha muitas vezes \emptyset_{ips} estava à janela, \emptyset_i chamavam, mas eu não podia. (PE)
 b. [*O meu filho*]_i tava chegando em casa - que nós trabalhamos com festa, como eu te falei, né - e *ele*_i tinha ido comprar bolas. Aí, não tinha as bolas que nós queríamos. *Ele*_i trouxe o dinheiro de volta. (PB)

Observando os pesos relativos para os fatores em cada variedade, vemos que eles são reveladores da significância dos fatores em cada grupo. No PE vemos uma hierarquia descendente, que mostra os pesos mais altos para os fatores mais favorecedores ao sujeito nulo, chegando aos mais baixos no seu desfavorecimento. A distância entre o mais alto e o mais baixo (*range*) é de 0,747, uma manifestação da força dos três primeiros fatores em relação aos dois últimos. No que diz respeito ao PB, o que vemos é uma distância menor (0,435) e uma perda da hierarquia, que coloca em dois extremos as estruturas com c-comando e com o antecedente adjacente, de um lado e, no extremo oposto, nesta ordem, a função diferente do antecedente, o referente distante e a ausência de c-comando. Os pesos relativos muito mais baixos e a menor distância entre eles revela um processo de perda do princípio “Evite Pronome” (CHOMSKY, 1981). O exame da distribuição dos sujeitos nulos no PB não nos permite incluir essa variedade entre as línguas de sujeito nulo “parcial”. Se os percentuais de nulos de 3ª pessoa em encaixadas com c-comando podem dar algum suporte a essa hipótese, ainda temos

sujeitos nulos em todos os padrões, incluindo o padrão com referente em sentença adjacente, o que não ocorre nas LSN parcial. Além disso, o PB ainda exibe sujeitos nulos de 1ª pessoa em todos os padrões (o que não ocorre nas LSN parcial).¹⁷ Quanto à 2ª pessoa, como já dissemos, a mudança parece estar concluída, com os casos atestados serem instâncias de sujeitos pragmaticamente identificados.

Passemos ao segundo grupo selecionado: o feixe de traços semânticos do referente do sujeito. A Tabela 2 exibe esses resultados.

Tabela 2 - Sujeito de 3ª pessoa vs feixe de traços semânticos (valor de aplicação: sujeito nulo)

TRAÇO	PE Input: 0,756			PB Input: 0,248		
	N / T	%	P. R.	N / T	%	P. R.
-ani/-esp	12/12	100%	---	7/12	58%	0,863
-ani/+esp	137/142	96,5%	0,942	73/173	42%	0,692
+ani/-esp	191/246	78%	0,562	62/191	32,5%	0,555
+ani/+esp	308/559	55%	0,307	189/803	23,5%	0,437
	<i>range</i>		0,635	<i>range</i>		0,426
	Log likelihood = -440,725 Significance = 0,000			Loglikelihood = -624,928 Significance = 0,000		

Fonte: Duarte (2019b, Tabela 3, p. 113)

Os resultados para o PE coincidem com o que se disse sobre o italiano na seção precedente – a ausência de pronomes pessoais com o traço [-animado] nas línguas românicas de sujeito nulo. De fato, o sujeito nulo é categórico quando o traço [-ani] interage com o traço [-específico], mas é preciso mencionar que os pronomes pessoais [-ani] não estão ausentes no PE. Houve na amostra cinco ocorrências de pronomes pessoais com esse traço. O exemplo a seguir mostra o SD globalização

¹⁷ Sobre ampla discussão acerca do estatuto de língua de sujeito nulo “parcial” atribuído ao PB e uma comparação minuciosa entre o PB e o finlandês na modalidade oral, remetemos o leitor a Duarte e Marins (2021). Nesse artigo, além de toda a bibliografia relevante à discussão, mostramos que o PB exibe um processo de mudança em curso, não se caracterizando, pois, como um sistema estável, como sugerem os trabalhos ali citados sobre o finlandês. Os nulos de 1ª pessoa ali citados são da modalidade escrita, que, como sabemos, é conservadora.

retomado por um sujeito nulo e por um pronome expresso. Pode-se pensar em ênfase, mas, de todo modo, este não é um comportamento de língua românica de sujeito nulo:

(32) Portanto, [*a globalização*]_i existe. Ø_i Já é uma realidade. A própria tecnologia faz com que *ela*_i exista. (PE)

Em relação aos sujeitos com os traços [+ani/-esp] e [+ani/+esp], encontramos índices de pronomes expressos, que alcançam 78% e 55%, respectivamente, sempre ligados aos padrões sentenciais:

- (33) a. Quando [*as crianças*]_i muitas vezes são deixadas sem orientação, sem pelo menos umas linhas gerais de orientação, Ø_i acabam por escolher caminhos que não são os mais correctos. (PE) (qualquer criança)
 b. Eu costumo dizer: quando há [*uma criança*]_i, alguém tem de ser o pai. *Ela*_i não saiu por obra e graça do espírito santo. (PE) (uma criança específica, um filho)

Quanto ao PB, a tabela nos mostra que o sujeito nulo, como esperado, é mais resistente com o traço [-ani], mas é evidente que o PB já desenvolveu um paradigma de pronomes fracos (KATO, 1999) com o traço [-ani] e exibe 58% de sujeitos nulos se combinado com o traço [-esp] e 42% com o traço [+esp], como ilustram (34) e (35), respectivamente:

- (34) a. É como eu falei: *emprego*_i, eu acho que Ø_i ia melhorar. (PB)
 b. [*Escola pública*]_i nunca é boa opção porque *elas*_i são ruins. (PB)

- (35) a. [*O sistema público*]_i é totalmente diferente de empresas privadas. Ø_i não funciona da mesma maneira. (PB)
 b. Eu acho que [*essa área de programação visual*]_i *ela*_i é um pouco mais difícil. (PB)

A implementação dos pronomes expressos com o traço [+ani] segue sem restrições, com apenas 32,5% e 23,5% de pronomes nulos se a interação ocorre com o traço [-esp] e [+esp], respectivamente:

(36) b. [*A criança*]_i acaba abandonando as escolas. \emptyset_i Perde o interesse. (PB)

a. Acredito que [*a pessoa*]_i tem um talento, tem uma facilidade pra certa coisa e, a partir do momento que *ela*_i desenvolve aquilo, *ela*_i vai ser sensacional.(PB)

(37) a. [*Meu filho*]_i nunca falou um palavrão perto de mim. \emptyset_i Nunca me respondeu. (PB)

b. [*Minha mãe*]_i sempre foi professora. **Ela**_i foi diretora de uma escola do Estado há vinte e cinco anos. (PB)

Os pesos relativos na Tabela 2 revelam a verdadeira força interna entre os fatores, opondo o traço [-ani] ao traço [+ani] no que diz respeito ao PE; em relação ao PB, vemos uma hierarquia que começa com os traços [-ani/-esp], com 0,863, a uma distância significativa dos traços [-ani/+esp] 0,692¹⁸, opondo-se às combinações que interagem com o traço [+ani]. A animacidade, então, se sobrepõe à especificidade, que atua ainda assim no curso da implementação dos pronomes pessoais no PB. A distância entre o peso mais alto e o mais baixo (*range*) é maior no PE, reafirmando a atuação do traço semântico no uso de pronomes pessoais nas LSNs “consistentes”.

O terceiro grupo selecionado é a estrutura do Sintagma Complementizador (CP), o que indica que, mais do que a função sintática da oração, o que importa é a ausência de elemento em CP, a presença de um núcleo (um complementizador – ou conjunção) ou um elemento na posição de especificador, um pronome relativo ou interrogativo. A Tabela 3 mostra a relevância desse grupo.

¹⁸ Segundo Naro (2003), no modelo logístico, uma distância superior a 0,100 entre dois fatores pode ser considerada relevante. Quanto maior a distância, maior a significância.

Tabela 3 - Sujeito de 3ª pessoa vs a estrutura do Sintagma Complementizador (CP) (valor de aplicação: sujeito nulo)

CP	PE Input: 0,757			PB Input: 0,248		
	N / T	%	P. R.	N / T	%	P. R.
Nenhum elemento	438/632	69%	0,578	252/901	28%	0,517
Elemento em C'	161/238	68%	0,422	74/217	34%	0,561
Elem. no espec. do CP	40/81	49%	0,179	05/61	08%	0,130
	<i>range</i>		0,399	<i>range</i>		0,426
Log likelihood = -440,725 Significance = 0,000			Loglikelihood = -624,928 Significance = 0,000			

Fonte: Duarte (2020, Tabela 4, p. 16)

Este grupo confirma não apenas o papel da função sintática da oração, mas também a importância da estrutura do Sintagma Complementizador (CP), que pode não conter qualquer elemento numa oração inicial (seja ela uma oração independente, a primeira de uma sequência de coordenadas ou uma matriz). Em tais casos, a oração pode trazer o CP vazio ou um pronome interrogativo. Sendo uma encaixada desenvolvida, pode conter um complementizador no núcleo do CP ou um pronome interrogativo ou relativo na posição de especificador do CP. A Tabela 3 ainda revela, nos percentuais e nos pesos relativos, que, para o PE, a ausência de elemento ou a presença de um complementizador favorece o sujeito nulo, em oposição à presença de um interrogativo ou relativo encabeçando a oração, como ilustram os exemplos a seguir, com um sujeito nulo numa oração introduzida pelo relativo e numa interrogativa indireta, introduzida por um advérbio interrogativo:

(38) a. E ele_i, passada a crise, pagou tudo aquilo que \emptyset_i não tinha pago durante aqueles três anos. (PE)

b. Daisy_i é amorosa. \emptyset_{2ps} Não vês como ela_i está? (PE)

A julgar pelo exemplo (9) de Calabrese na Seção 2 e pela avaliação dos nativos (cf. nota 5), os pronomes relativos não são fortes condicionadores do sujeito nulo nas línguas como o italiano. Isso pode se dever ao fato de terem uma função sintática, que pode, em certos casos, comprometer a interpretação de um sujeito nulo adjacente.

Quanto ao PB, não surpreende que este seja o contexto mais desfavorável ao sujeito nulo. Enquanto as orações introduzidas por um complementizador lideram o percentual de nulos com parcos 34%, seguidas pelos contextos sem qualquer elemento em CP, com 28%, as orações com um especificador alcançam apenas 8%, um resultado que coincide com o quase desaparecimento do sujeito nulo em interrogativas-Q diretas, tanto na 2ª quanto na 3ª pessoa (DUARTE, 1992; NICOLAU DE PAULA, 2016, entre outros). Na amostra analisada, são apenas cinco os casos de sujeitos nulos, como ilustrado em (39):

- (39) a. Ele_i adorava fazer tudo o que Ø_i não devia [fazer]. (PB)
 b. Adoro [meus filhos]_i. Curto brincar com eles_i. Curto sair com eles_i. É claro que eles_i já tão chegando numa certa idade que eles_i já tão independents. (PB)

De todo modo, os percentuais e pesos relativos reforçam os mesmos efeitos atuando em sistemas com marcações diferentes para o valor do PSN. E, mais importante, permitem ver o efeito de alguns fatores sobre outros, já mostrando uma mudança perto da conclusão no PB.

5 Algumas evidências do “encaixamento” da mudança no PB

Entre os subprodutos que essa perspectiva sintática comparativa nos permite identificar, respondendo à questão do “encaixamento” da mudança, estão os sujeitos indeterminados, quer de referência arbitrária, quer de referência genérica, preferencialmente realizados por pronomes pessoais plenos. Com a quase extinção do clítico *se*, para expressar ambas as referências, a que exclui o falante e a que pode ou não incluí-lo, respectivamente, ganham força, no primeiro caso, o uso do pronome *eles* (ainda em competição com o uso da 3ª p.p com o sujeito nulo), e, no segundo, o uso de *você* (ou *tu*, a depender de fatores diatópicos). Em (40), ilustramos as duas mais frequentes estratégias de indeterminação (arbitrária e genérica), com base em Duarte e Marins (2021), onde aparecem os resultados obtidos para a mesma amostra analisada na seção anterior, com 91,3% e 93,1%, respectivamente:

- (40) a. Agora que *eles*_{arb} ‘tão arrumando tudo...é muito difícil,
 b. Se *você*_{gen} não tiver um sonho, *tu*_{gen} não é nada.

Na referida análise, são apresentados, igualmente, os resultados para o sujeito nulo arbitrário (41a) e o genérico (41b,c) com o verbo na 3ª pessoa do singular, que alcançam, respectivamente 4,6% e 3,7%, ilustrados em (41):

- (41) a. Olha, na televisão Ø_{arb} fala muito isso, né?
 b. Ø_{gen} tá precisando de políticos mais sérios.
 c. Pra beber, Ø_{gen} tem que ter noção!

Os agonizantes clíticos para cada referência alcançam 4,1% e 3,2% e são atestados na fala dos brasileiros mais velhos.

A existência do nulo genérico é o outro argumento usado para incluir o PB entre as LSN “parcial” (ROBERTS; HOLMBERG, 2010). Entretanto, nessas línguas, e aqui citamos o finlandês, o nulo genérico não pode ocorrer em primeira posição, o que é evitado com o uso do expletivo *sita*. Assim, acreditamos que o PB não exhibe contextos específicos e restritos para a ocorrência do sujeito nulo, referencial e genérico, o que torna bem complexa a sua inclusão entre as LSN “parcial”. Para Galves (1987) uma boa explicação para a possibilidade de atribuir uma interpretação genérica (e arbitrária, acrescentamos) para esse sujeito nulo de 3ª pessoa do singular seria justamente a impossibilidade de identificar um sujeito nulo referencial de 3ª pessoa.

Um outro fenômeno atestado na amostra utilizada até aqui para ilustrar as mudanças relacionadas à remarcação do valor do PSN e já atestado em Duarte (1995) é a ocorrência irrestrita de redobro do sujeito (ou sujeitos deslocados à esquerda, apontados por Pontes (1985) entre as construções de tópico marcado mais frequentes no PB, e uma sugestão de que estamos diante de uma língua com proeminência de tópico (LI; THOMPSON, 1976). Essas construções estão ausentes nas LSN do grupo “consistente” (DURANTI; OCHS, 1979), mas presentes no francês (AVANZI, 2011; BARNES, 1986). Trata-se claramente de evidência de um subproduto da mudança. Não há restrições quanto ao estatuto semântico [+ani/-ani], informacional [novo/dado], à definitude [+esp/-esp], podendo o elemento ser inclusive quantificado; da mesma forma, na prosódia, a pausa ou sua ausência não restringem o fenômeno. Alguns exemplos de redobro do sujeito na amostra seguem em (42):

- (42) a. Aí, [*muitas dessas pessoas*], *elas*_i estudavam artes cênicas.
 b. [*O teatro*], *ele* até deu uma encarecida.
 c. Eu acho [*que essa área de programação visual*], *ela*_i é um pouco mais difícil.

Tendo em conta a perspectiva paramétrica que orienta a pesquisa, seria de esperar consequências nas sentenças impessoais. Vejamos então alguns outros efeitos da mudança. O PB não desenvolveu um expletivo lexical, um fenômeno que ocorreu no passado na evolução do francês médio, com o surgimento do expletivo lexical *il* (ROBERTS, 1993; VANCE, 1989), e vem ocorrendo no presente numa variedade do espanhol dominicano (TORIBIO, 1996), com o surgimento de *ello* expletivo, uma mudança de baixo para cima e ainda muito estigmatizada (*ello llueve*). Esse desenvolvimento de um expletivo lexical não é casual, nem no francês nem no espanhol dominicano. É, na verdade, uma consequência da redução do paradigma flexional, que acabou por desencadear a realização plena dos sujeitos referenciais. A consequência foi a emergência do expletivo lexical, um comportamento que, de certa forma, segue o comportamento de uma gramática do grupo das línguas românicas que perderam o sujeito nulo.

E por que o mesmo não teria ocorrido no PB? É aqui que entra a já anunciada orientação para o discurso ou a proeminência de tópico, apontada por Pontes (1985) e já referida neste artigo. Segundo Li e Thomson (1976), línguas com a orientação para o discurso **não têm expletivos lexicais**, ou seja, não há, em línguas como o japonês e o chinês, um elemento foneticamente realizado sem conteúdo semântico. Não queremos dizer que o PB estaria no grupo das línguas de sujeito nulo “radical”, como o japonês e o chinês; desde Pontes (1985), temos observado que o PB tem proeminência de sujeito e de tópico. Entretanto, a mudança pela qual passou o francês no passado, que acabou por desenvolver um expletivo lexical “*il*” (ROBERTS, 1993) e a que ocorre no presente, com uma variedade do espanhol dominicano, que desenvolveu o expletivo “*ello*” (VANCE, 1989), constituem evidência suficiente para explicar a manutenção dos expletivos nulos no PB: sua parcial orientação para o discurso. E é assim que podemos entender as estratégias que o sistema desenvolve para evitar expletivos nulos, lançando mão exatamente de elementos referenciais na posição estrutural do sujeito.

Podemos citar, entre muitas outras estratégias a substituição de **haver** por **ter** existencial no PB (e não no PE). Essa substituição, que mantém além de sentenças impessoais, permite igualmente sentenças com **ter** pessoal, conservando o sentido existencial. É o que mostram Marins (2013) e Duarte e Marins (2021), entre muitos trabalhos que precederam essa investigação, apontando, na amostra aqui analisada, entre 1334 ocorrências, 7% de **haver** versus 93% de **ter**, incluindo sentenças impessoais (43b) e pessoais (43c,d):

- (43) a. Eu acho que \emptyset_{expl} **houve** uma inversão de valores
 b. \emptyset_{expl} **Tinha** esses bares aqui.
 c. **Você tem** ônibus pro shopping, né?
 d. Porque elas são motos importada(s). E **ela não tem** peça nenhuma dela aqui.

Além da análise de Marins sobre as sentenças existenciais, temos, também, com base em *corpus* que utilizou *sites* de reclamação na internet, a análise de Duarte e Fernandes (2016), que pôde computar as três formas variantes em construções com verbos inacusativos apresentando o argumento interno que contém um genitivo, como ilustramos a seguir com o mesmo exemplo. Ele ilustra uma construção com um expletivo nulo em (44a), a anteposição do argumento interno sujeito em (44b) e o alçamento do argumento do genitivo em (44c):

- (44) a. \emptyset_{expl} rachou [a tela do meu celular]
 b. [A tela do meu celular]_i rachou [t]_i
 c. Meu celular rachou [a tela t_i]

Os resultados apontaram ampla preferência pelo alçamento do genitivo em (c), mais uma construção que chamou a atenção de Pontes (1985), que a nomeou “tópico-sujeito”. Enquanto as estruturas com o expletivo nulo exibiram 4% e a anteposição do DP sujeito, em b, 38%, o alçamento do genitivo chegou a 58%. A que se pode atribuir a preferência por c em relação a b? A preferência pela anteposição do genitivo “o celular” sugere que ele tem maior proeminência do que “a tela do meu celular”.

Atestamos ainda expletivos nulos em sentenças com verbos climáticos competindo com o alçamento de adjuntos e construções pessoais, como exemplificamos em (45) e (46), com exemplos retirados

de escrita de redações escolares, de fala espontânea e de comentários na mídia (os exemplos atestados estão em c; os demais foram construídos):

- (45) a. \emptyset_{expl} Chove muito nessas florestas
 b. Nessas florestas \emptyset_{expl} chove muito
 c. Essas florestas chovem muito
- (46) a. Os pombos da Cinelândia somem quando \emptyset_{expl} chove na cidade!
 b. Os pombos da Cinelândia somem quando na cidade \emptyset_{expl} chove!
 c. Os pombos da Cinelândia somem quando **a cidade** chove!

São muitos os casos de alçamento ou inserção de constituintes (como “isso”, “aquilo”, entre outros, que não ilustramos aqui) que permitem evitar um expletivo nulo. Uma pesquisa capaz de fornecer evidências para essa competição, que pode, na verdade, durar muito tempo, precisa lançar mão de diferentes *corpora*, tanto de língua oral quanto da escrita mais espontânea divulgada em meios digitais, ou ainda da observação informal, como fez Eunice Pontes. O fato é que o PB evita V1, como apontado por Kato e Duarte (2003; 2018; 2021), utilizando elementos leves em contextos em que ainda ocorre um sujeito referencial nulo, que, como vimos, é cada vez mais raro (*eu vou, só vou, não vou*). Segundo as autoras, haveria um novo padrão rítmico (ou prosódico) que leva à presença de um elemento em posição inicial. O mesmo novo padrão rítmico explicaria os alçamentos de constituintes para a posição inicial de uma sentença impessoal, evitando um expletivo nulo, como mostramos em (43c, d), (44c, d), (45c) e (46c). E, nesse sentido, compartilhamos a hipótese de Pilati (2006) e Pilati, Salles e Naves (2017).

6 Para concluir

Fechamos este artigo, justamente para virar a página, lembrando um artigo de Tarallo (1986), em que o autor compara o sociolinguista a Zelig, o camaleão – personagem de um filme de Woody Allen, que trazia um psicanalista que assimilava a personalidade de outras pessoas com quem contactava. Esse artigo é citado por Pagotto (2000) no texto a que nos referimos aqui. Pagotto lembra que Tarallo dizia serem os

sociolinguistas meio camaleões, que tinham de assumir os mais diversos matizes teóricos, a fim de dar conta do fenômeno de variação. Mas, se pensarmos que a TVM e as teorias linguísticas estão em campos diferentes, como reitera Pagotto em seu artigo, veremos que Tarallo estava apenas reconhecendo isso. Embora ele não tenha escrito o que estamos aqui dizendo, ele apenas reconhecia que precisava de uma teoria linguística que lhe permitisse fazer sintaxe comparativa. Mas sua postura foi muito criticada, e a metáfora foi muito mal-entendida. Como lembra Pagotto (2000, p. 62), a crítica apontava “uma certa leviandade que trazia uma ausência de rigor científico” e deixava transparecer certa “irracionalidade” (as referências no artigo de Pagotto remetem o leitor interessado à tal polêmica). O irreverente Pagotto conclui seu artigo dizendo pensar que

a imagem do camaleão não é a mais adequada. O sociolinguísta variacionista não é um ser que assume a personalidade deste ou daquele teórico. O sociolinguísta é um predador - um *tiranossaurus rex* - que devora e metaboliza as suas presas teóricas, refundindo e ressignificando as suas carnes preciosas, transformando matéria em energia, tentando perceber, ainda que intuitivamente, que matéria e energia podem ser a mesma ‘coisa’” (PAGOTTO, 2000, p. 62).

Se nos dermos conta do que diz Pagotto (e vimos dizendo isso há já algum tempo) – que a Sociolinguística e as Teorias Linguísticas estão em campos diferentes –, fica evidente que a Sociolinguística terá de lançar mão de uma teoria linguística para pôr de pé a pesquisa formulada em Weinreich, Labov e Herzog (1968), ou deixará de levantar hipóteses, de construir grupos de fatores, de buscar respostas ao “encaixamento” da mudança; deixará, enfim, de iluminar os caminhos da implementação da mudança ao longo do tempo, seja ele “aparente”, observando as diferentes gerações numa mesma sincronia, seja ele de curta ou longa duração. Se bem elaborada, justificando seus “grupos de fatores” e buscando explicações, para além das meras descrições, essa associação necessária certamente rende frutos. Os resultados que as análises empíricas trazem para as teorias linguísticas são inegáveis. Podemos dizer hoje que existe um tipo de retroalimentação, de parte a parte, que só pode contribuir para o conhecimento da mudança linguística. Na verdade, como afirma Roberts (2007, p. 339), ao examinar a mudança em curso no PB, não temos muita clareza sobre o que acontecerá com o sistema, mas “a

dinâmica da mudança corresponde amplamente ao cenário desenhado por Weinreich, Labov e Herzog”. Trata-se de uma situação complexa, “seja do ponto de vista sociolinguístico seja do ponto de vista linguístico”, mas não há dúvida de que as análises apresentadas “mostram como podemos usar as ferramentas da Teoria de Princípios e Parâmetros combinadas com a Teoria Sociolinguística, a fim de observar a mudança sintática em progresso” (ROBERTS, 2007, p.339).

Por que usamos a expressão “virar a página” no título deste artigo? Porque esperamos ver devidamente reconhecido o trabalho de Tarallo e seus orientandos, que continuam a desenvolver os estudos por ele inspirados, e, além disso, ver esclarecida a inutilidade de uma polêmica que, talvez por não se basear na efetiva prática da pesquisa sociolinguística, por não ter se dado conta de que os estudos em fonologia tinham por base teorias linguísticas diversas (entre as quais a de base gerativista), levasse alguns linguistas a proclamar a incomensurabilidade entre a empiria e o formalismo. Desde o libelo de Tarallo “por uma sociolinguística paramétrica”, pudemos ver as primeiras dissertações e teses por ele orientadas, entre as quais as de Duarte (1986), Lopes (1988) e Berlinck (1988), resumidas em Tarallo (1989), seguidas de Nunes (1990) e Pagotto (1992), em Roberts e Kato (1993), as teses de Oliveira (1989) e Ramos (1992), sem mencionar o número de orientações conduzidas por esses e outros pesquisadores por eles formados, que compõem hoje um respeitável grupo que trabalha com variação e mudança sintática não só no português brasileiro como nas outras variedades, além das demais irmãs românicas. É reconhecida a parceria de Tarallo com Mary Kato, que, desde seu falecimento precoce, vem se mantendo nos trabalhos constantes de Kato e Duarte. E a associação entre a variação e a sintaxe formal se deixa ver em todos esses trabalhos.

Declaração de autoria

A primeira autora do artigo cuidou do quadro teórico, da análise, tema de seu projeto financiado pelo CNPq, e da redação do texto. O segundo autor, aluno de IC até fevereiro de 2021, quando passou para o Mestrado, cuidou do levantamento e análise estatística dos dados, e participou ativamente da interpretação e da redação do texto, bem como da implementação das sugestões dos pareceristas.

Agradecimentos

Agradecemos aos pareceristas *ah hoc* pelas inúmeras sugestões no sentido de tornar certos pontos do texto mais claros e assegurar aos leitores que a Sociolinguística é um modelo teórico para estudar a mudança, levando em conta fatores linguísticos e sociais, e que os fatores linguísticos devem necessariamente vir de um modelo teórico gramatical que melhor possibilite o tratamento do fenômeno em questão dentro da teoria com que o pesquisador trabalhe. Todas as falhas remanescentes são de nossa inteira responsabilidade.

Referências

AVANZI, M. La dislocation à gauche em français spontané. Étude instrumentale. *Le français moderne*, Neuchâtel, n. 2, p. 77-91, 2011. DOI: http://unine.ch/files/live/sites/structuration_periodes/files/shared/new_am/35_.pdf

BARBOSA, P.; DUARTE, M. E. L.; KATO, M. A. Null subjects in European & Brazilian Portuguese. *Journal of Portuguese Linguistic: Studies in the Comparative Syntax of European & Brazilian Portuguese*, Lisboa, v. 4, n. 2, p. 11-52, 2005. DOI: <https://jpl.letras.ulisboa.pt/article/id/5535/>

BARNES, B. K. An Empirical Study of the Syntax and Pragmatics of Left dislocations in Spoken French. In: JAEGGLI, O.; SILVA-CORVALÁN, C. (eds.). *Studies in Romance Linguistics*. Dordrecht: Foris, 1986. p. 207-224.

BERLINCK, R. A de A. *A ordem VSN no português do Brasil: sincronia e diacronia*. 1988. 265 f. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, 1988.

BORGES NETO; PAULA MULLER, A. L. Lingüistas ou Camaleões? Uma resposta a Tarallo. *D.E.L.T.A.* São Paulo, v. 3, n. 1, p. 47-53. 1987. DOI: <https://revistas.pucsp.br/index.php/delta/article/view/43840/29116>

CALABRESE, A. Pronomina: some properties of the Italian pronominal system. *MIT Working Papers in Linguistics*, Cambridge, v. 8, p. 146, 1986. DOI: <http://mitwpl.mit.edu/catalog/mwpl08>

CARDINALETTI, A. ; M. STARKE. The typology of structural deficiency: On the three grammatical classes. In: van RIEMSDIJK, HHenk van Riemsdijk. (ed.). *Clitics in the Languages of Europe*, Empirical Approaches to Language Typology. Berlin: Mouton de Gruyter, 1999. p. 145-233.

CHOMSKY, N. *Lectures on Government and Binding*. Dordrecht: Foris, 1981.

DI EUGENIO, B. Centering Theory and the Italian pronominal System. *COLING'90: Proceedings of the 13th Conference on Computational Linguistics*, V. 2. ACM – Digital Library, 1990. p. 270-275.

CYRINO, S. M. L. Observações sobre a mudança diacrônica no português do Brasil: objeto nulo e clíticos. In: I. ROBERTS, I.; KATO, M. A. (orgs.). *Português Brasileiro: uma viagem diacrônica*. Campinas: Ed. da UNICAMP, 1993. p. 163-184.

DUARTE, M. E. L. *Variação e Sinataxe: clítico acusativo, pronome lexical e categoria vazia no português do Brasil*. 1986. 73 f. Dissertação (Mestrado em Linguística) – Faculdade de Letras, PUC-SP, 1986.

DUARTE, M. E. L. A perda da ordem V(ERBO) S(SUJEITO) em Interrogativas QU- no português do Brasil. *DELTA*. São Paulo, v. 8, n. Especial, p. 37-52, 1992. DOI: <https://revistas.pucsp.br/index.php/delta/article/view/45720/30193>

DUARTE, M. E. L. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português brasileiro. In: ROBERTS, I.; KATO, M. A. (orgs.). *Português Brasileiro: uma viagem diacrônica*. Campinas: ED. DA UNICAMP, 1993. p. 107-128.

DUARTE, M. E. L. *A perda do princípio “Evite Pronome” no português brasileiro*. 1995. 149 f. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Unicamp, 1995.

DUARTE, M. E. L. Sociolinguística “Paramétrica”. In: MOLLICA, M. C.; RFERRAREZI JR, C. (orgs.). *Sociolinguística, sociolinguísticas - uma introdução*. São Paulo: Contexto, 2016. p. 33-44.

DUARTE, M. E. L. A Sociolinguística “paramétrica”: desfazendo alguns equívocos. *GUAVIRA LETRAS*, Três Lagoas, v. 15, p. 124-140, 2019a. DOI: <http://websensors.net.br/seer/index.php/guavira/article/view/868/606>

DUARTE, M. E. L. O sujeito nulo referencial no português brasileiro e no português europeu. In: GALVES, C.; KATO, M. A.; ROBERTS, I. (orgs.). *Português Brasileiro: uma segunda viagem diacrônica*. Campinas, Ed. da Unicamp, 2019b. p. 93-126.

DUARTE, M. E. L. A remarcação em curso no valor do parâmetro do sujeito nulo. *CUADERNOS DE LA ALFAL*, Bogotá, Colômbia; v. 12, p. 71-99, 2020. DOI: https://www.mundoalfal.org/sites/default/files/revista/12_2_cuaderno_005.pdf

DUARTE, M. E. L.; MOURAO, G. C.; SANTOS, H. M. Os sujeitos de 3ª. pessoa: revisitando Duarte 1993. In: DUARTE, M. E. L. (Org.). *O sujeito em peças de teatro (1833-1992): estudos diacrônicos*. São Paulo: Parábola Editorial, 2012. p. 21-44.

DUARTE, M. E. L.; REZENDE DOS REIS, E.P. Revisitando o Sujeito Pronominal Vinte Anos Depois. *ReVEL – Revista Virtual de Estudos da Linguagem*, Porto Alegre, RS, v. 16, n. 30, p. 173-197, 2018. DOI: <http://www.revel.inf.br/files/23f8c093cf2be398414c965bf05f8e75.pdf>

DUARTE, M. E. L.; FERNANDES, U. B. Construções de tópico sujeito em contexto de variação e mudança. *Web-Revista SOCIODIALETO*, Universidade Estadual de Mato Grosso do Sul, v. 6, p. 347-371, 2016.

DUARTE, M. E. L.; MARINS, J. E. Brazilian Portuguese: a ‘partial’ null subject language? *Cadernos De Estudos Linguísticos*, Campinas, v. 63, p. 1-21, 2021. DOI: <https://periodicos.sbu.unicamp.br/ojs/index.php/cel/article/view/8661660/27152>

DUARTE, M. E. L.; SOUSA, A. A. M. de; FERNANDES, U. S. B.; CARDOSO, M. M. C. A redução no quadro de clíticos de terceira pessoa no português brasileiro: um estudo diacrônico. *LABORHISTÓRICO*, Rio de Janeiro, v. 7, p. 154-187, 2021. DOI: <https://revistas.ufrj.br/index.php/lh/article/view/44191/25057>

DUBUISSON, C. L’inversion du SN sujet et la post-position du SN lourd en français. In: SANKOFF, D. ; H. CEDERGREN, H. (eds.). *Variation Omnibus*. Edmonton, Alberta: Linguistic Research, 1981

DURANTI, A.; OCHS, E. Left-dislocation in Italian conversation. In: GIVÓN, T. (ed.), *Syntax and Semantics: Discourse and Syntax*. New York: Academic Press, 1979. p. 377-415.

FERNÁNDEZ SORIANO, O. F. Strong Pronouns in Null Subject Languages and The Avoid Pronoun Principle. *MIT Working Papers in Linguistics*, Cambridge, v. 11, p. 228-239, 1989.

FRASCARELLI, M. Subjects, topics and the interpretation of referential pro. Na interface approach to the linking of (null) pronouns. *Natural Language and Linguistic Theory*, Berlim, v. 25, p. 691–734, 2007. DOI: <https://doi.org/10.1007/s11049-007-9025-x>

GALVES, C. M. C. A sintaxe do português brasileiro. *Ensaio de Linguística*, Belo Horizonte, v. 13, p. 31-50, 1987.

GROSZ, B.; JOSHI, A.; E, WEINSTEIN, S. *Towards a computational theory of discourse interpretation*. 1986. Não publicado.

GUIMARÃES, L. de S. *A representação do sujeito proposicional em peças de teatro portuguesas: um estudo diacrônico*. 2021. 110 f. Dissertação (Mestrado em Letras Vernáculas) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2021.

HALLIDAY, M. A. K.; HASAN, R. *Cohesion in English*. London: Longman, 1979.

HENRY, A. Variation and Syntactic Theory. In: CHAMBERS, J. ; TRUDGILL, P. ; SCHILLING-ESTES, N. Chambers, J., Trudgill, P. And Schilling-Estes, N. (eds.) *The Handbook of Language Variation and Change*. Oxford: Blackwell, 2006. P. 267-282.

KATO, M. A. Strong and weak pronominals in the Null Subject Parameter. *Probus*, Berlin ; New York : Foris Publications, v. 11, p.1-37, 1999. DOI: <https://www.degruyter.com/document/doi/10.1515/prbs.1999.11.1.1/html>

KATO, M. Recontando a História das Relativas em uma Perspectiva Paramétrica. In: KATO, M. (org.). *Português Brasileiro: uma viagem diacrônica*. Campinas: UNICAMP, 1993. P. 175-205.

KATO, M. A. Aspectos morfofonológicos nos paradigmas dos pronomes fortes e fracos do Português brasileiro. *Revista da ANPOLL*, Brasília, v. 1, p. 13-23, 2018. DOI: <https://doi.org/10.18309/anp.v1i46.1084>.

KATO, M. A.; DUARTE, M. E. L. Semantic and phonological constraints on the distribution of null subjects in Brazilian Portuguese. Comunicação apresentada no NWAV31, Univ. da Pennsylvania, 2003. DOI: <https://www.ling.upenn.edu/NWAVE32/abs-pdf/kato-duarte.pdf>

KATO, M. A.; DUARTE, M. E. L. Pre-verbal position in Brazilian Portuguese: a reinterpretation of the ‘avoid pronoun’ principle. *DIADORIM*, Rio de Janeiro, v. 20, p. 610-626, 2018. DOI: <https://revistas.ufrj.br/index.php/diadorim/article/view/23293/15252>

KATO, M. A.; DUARTE, M. E. L. Prosodic Determinants in Syntax. *Cadernos de Linguística*, v. 2, p. 1-18, 2021. DOI: <https://cadernos.abralin.org/index.php/cadernos/article/view/350/441>

KROCH, A. Reflexes of Grammars in Patterns of Language Change. *Language Variations and Change*, Cambridge, v. 1, p.199-244, 1989. DOI: <https://www.ling.upenn.edu/~kroch/papers/reflexes.pdf>

- LI, C.; THOMPSON, S. Subject and topic: a new typology of language. In: LI, C. (ed.). *Subject and Topic*. New York: Academic Press. 1976. p. 457-489.
- LOPES, R. E. V. *Passivas nos discursos oral e escrito: sintaxe e variação*. 1988. 97f. Dissertação (Mestrado em Linguística), Faculdade de Letras, PUC-SP, 1988.
- MARINS, J. *O Parâmetro do Sujeito Nulo: uma análise contrastiva entre o português e o italiano oral*. 2009. 111 f. Dissertação (Mestrado em Linguística), Faculdade de Letras, UFRJ, 2009.
- MARINS, J. E. *As repercussões na marcação do parâmetro do sujeito nulo: um estudo diacrônico das sentenças existenciais com ter e haver no PB e no PE*. 2013. 166 f. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2013.
- MARTÍNEZ-SANZ, C. *Null and overt subjects in a variable system: the case of Dominican Spanish*. 2011. 484 p. Thesis (PhD in Spanish) – Faculty of Arts, University of Ottawa, 2011.
- NARO, A. J. Modelos quantitativos e tratamento estatístico. In: MOLLICA, M. C. M.; BRAGA, M. L. (orgs.). *Introdução à Sociolinguística: o tratamento da variação*. São Paulo: Contexto, 2003. p. 15-25.
- NICOLAU DE PAULA, M. *A ordem VS/SV e as interrogativas-Q no PE e no PB: uma análise diacrônica*. 2016. 147 f. Tese (Doutorado em Letras Vernáculas) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2016.
- NUNES, J. M. *O famigerado SE: Uma análise sincrônica e diacrônica das construções com SE apassivador e indeterminador*. 1990. 172 f. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, UNICAMP. 1990.
- OLIVEIRA, D. *Constituintes sentenciais: preenchimento, queda e ordenação*. 1989. 150 f. Tese (Doutorado em Linguística) - Faculdade de Letras, PUC-SP, 1989.
- PAGOTTO, E. G. *A posição dos clíticos em Português: um estudo diacrônico*. 1992. 179 f. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, UNICAMP, 1992.
- PAGOTTO, E. G. De Camaleão a Tiranossaurus Rex - Linguística e Sociolinguística. *Laços - Revista da Associação de Estudos da Linguagem do Rio de Janeiro*, Rio de Janeiro, v. 1, p. 47-57, 2000.

PAREDES SILVA, V. L. *É isso aí: verbo ser e demonstrative em função coesiva no Português*, 1985, não publicado.

PAREDES SILVA, V. L.; OLIVEIRA, A. *É isso aí: a variação na referência estendida em diferentes gêneros da escrita*. In: PAIVA, M. C.; GOMES, C. A. (orgs.) *A Dinâmica da Variação e Mudança na Fala e na Escrita*. Rio de Janeiro: Contracapa, 2014. p.45-68.

PILATI, E. *Aspectos sintáticos e semânticos da ordem verbo-sujeito no português*. 2006. 242 f. Tese (Doutorado em Linguística) – Instituto de Letras, UnB, 2006.

PILATI, E.; NAVES, R.; SALLES, H. M. M. L. On the Syntax of Subjects in Brazilian Portuguese: Using the “SPLIT” Pronominal System as the Basis for an Alternative Analysis. *DIADORIM*, Rio de Janeiro, v. 19, p. 99-139, 2017. DOI: <https://revistas.ufrj.br/index.php/diadorim/article/view/13579/15376>

PINHEIRO, D. O. R.; MARINS, J. E. A trajetória das interrogativas QU-clivadas e não clivadas no português brasileiro. In: DUARTE, M. E. L. (org.). *O sujeito em peças de teatro (1833-1992)*: Estudos diacrônicos. São Paulo: Parábola Editorial, 2012. p. 161-180.

PONTES, E. *O Tópico no Português do Brasil*. Campinas: Editora Pontes, 1987.

RAMOS, J. M. *Marcação de Caso e Mudança Sintática no Português*. 1992. 380 f. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Unicamp, 1992.

RIZZI, L. *Issues in Italian Syntax*. Dordrecht : Foris, 1982.

ROBERTS, I. *Verbs and Diachronic Syntax*. Dordrecht: Kluwer, 1993.

ROBERTS, I. *Diachronic Syntax*. Oxford: Oxford University Press, 2007.

ROBERTS, I. *Parameter hierarchies and universal grammar*. Oxford: Oxford University Press, 2019.

ROBERTS, I.; HOLMBERG, A. Introduction: parameters in Minimalist theory. In: BIBERAUER, T. *et al.* (eds). *Parametric Variation: null subjects in Minimalist theory*. Cambridge: Cambridge University Press, 2010. p. 1-57.

ROBERTS, I.; KATO, M. A. *Português Brasileiro: uma viagem diacrônica*. Campinas: Ed. da Unicamp. 1993.

ROUSSEAU, P.; SANKOFF, D. Singularities in the analysis of binomial data. *Biometrika*, Oxford, v. 65, n. 3, p. 603–608, 1978. DOI: <https://doi.org/10.1093/biomet/65.3.603>

SANKOFF, D.; TAGLIAMONTE, S. SMITH, E. *Golbvarb X Programs*. Toronto: University of Toronto, 2005.

SANTOS, D. R; SOARES DA SILVA, H. A ordem V-DP/DP-V com verbos inacusativos. In: DUARTE, M. E. L. (org.). *O sujeito em peças de teatro*. São Paulo: Parábola, 2012. p. 121 - 142.

TARALLO, F. *A Pesquisa Sociolinguística*. São Paulo: Ática. 1986.

TARALLO, F. Por uma Sociolinguística Românica Paramétrica: Fonologia e Sintaxe. *Ensaio de Linguística*, Belo Horizonte, v. 13, p. 51-84, 1987.

TARALLO, F. *Fotografias Sociolinguísticas*. Campinas: Editora da Unicamp. 1989.

TARALLO, F.; KATO, M. Harmonia trans-sistêmica: variação intra- e inter-linguística. *DIADORIM*, Rio de Janeiro, n. 2, p. 13-42, 2007. DOI: <https://revistas.ufrj.br/index.php/diadorim/article/view/3849/2827>

TORIBIO, J. Dialectal Variation in the licensing of null referential expletive subjects. In: PARODI, C.; QUICOLI, C.; SALTARELLI, M.; ZUBIZARRETA, M. L. (Eds.) *Aspects of Romance Linguistics*. Washington: Georgetown University Press, 1996. p. 409-432.

VANCE, B. S. *Null Subjects and Syntactic change in Medieval French*. 1989, Tese (Doutorado em Linguística), Cornell University, 1989.

VIEIRA, S. R.; BRANDÃO, S. F. *CORPORAPORT: Variedades do Português em análise*. Rio de Janeiro: Faculdade de Letras-UFRJ. Disponível em: <www.corporaport.letras.ufrj.br> Acesso em: 14 fev 2022.

WEINREICH, U.; LABOV, W.; HERZOG, M. Empirical foundations for a theory of language change. In: LEHMAN, W.; MALKIEL, Y. (eds.). *Directions for historical linguistics*. Austin: University of Texas Press, 1968. p. 97-195.



On Partial Null Subject Languages: Why Pro-Drop in Brazilian Portuguese and Russian Became Similar But Not Identical

*Sobre as línguas de Sujeito Parcialmente Nulo:
Por que o pro-drop em português brasileiro e em russo se
tornaram semelhantes, mas não idênticos*

Nerea Madariaga

University of the Basque Country (UPV/EHU), Vitoria-Gasteiz / Spain

nerea.madariaga@ehu.es

<https://orcid.org/0000-0003-4732-4144>

Abstract: In this paper, I claim that a parametric view on change in pro-drop does not contradict the fact that not all the Partial Null Subject (PNS) languages display identical properties. I show that the contingent nature of diachronic change is the reason for the slight differences between PNS languages. Modern Russian (MR) and Brazilian Portuguese (BP) are two PNS languages that developed from Consistent Null Subject antecedents (Old Russian and European Portuguese) independently from each other. I account for the change in pro-drop experienced by these two languages, analyzing the properties usually related to the null subject parameter (verbal inflection, clitics, null objects, embedded and arbitrary null subjects), and show that the final parametric setting in both MR and BP was almost identical, with small differences that can be attributed to the different initial conditions for the change.

Keywords: Pro-drop; null subjects; null objects; clitics; Consistent Null Subject languages; Partial Null Subject languages; Brazilian Portuguese; European Portuguese; Modern Russian; Old Russian.

Resumo: Neste artigo, afirmo que uma visão paramétrica sobre a mudança em pro-drop não está em conflito com o fato de que nem todas as línguas de sujeito nulo parcial (PNS) apresentam propriedades idênticas. Mostro que a natureza contingente da mudança diacrônica é a razão para as pequenas diferenças entre as línguas PNS. O russo moderno

(MR) e o português brasileiro (BP) são duas línguas PNS que se desenvolveram a partir dos antecessores de sujeito nulo consistente (antigo russo e português europeu) independentemente uma da outra. Relato a mudança de pro-drop experimentada por essas duas línguas, analisando as propriedades geralmente relacionadas ao parâmetro de sujeito nulo (inflexão verbal, clíticos, objetos nulos, sujeitos nulos subordinados e arbitrários), e mostro que a configuração paramétrica final em ambas MR e BP foi quase idêntica, com pequenas diferenças que podem ser atribuídas às diferentes condições iniciais para a mudança.

Palavras-chave: Pro-drop; sujeitos nulos; objetos nulos; clíticos; línguas de sujeito nulo consistente; línguas de sujeito nulo parcial; português brasileiro; português europeu; russo moderno; russo antigo.

Recebido em 16 de fevereiro de 2022

Aceito em 28 de março de 2022

1 Initial remarks

A recurrent topic in the literature in recent years is the discussion about the nature of so-called Partial Null Subject (PNS) languages, since Holmberg (2005), Holmberg; Nayudu; Sheehan (2009), Biberauer; Roberts; Holmberg; Sheehan (2010), etc. On the one hand, there seem to be salient similarities between the languages making up this group, but at the same time, not all the features characterizing PNS languages are common to all of them.

This paper aims to contribute to a more accurate characterization of PNS languages, through the analysis of the current status and historical development of Null Subjects (NSs) in two Indo-European languages, Brazilian Portuguese and Russian, which shifted from a Consistent Null Subject (CNS) pattern into a PNS pattern independently from each other. The comparison of the diachronic development of these two languages will shed light on the reasons why PNS languages are different and similar at the same time. The similarities, it will be argued, stem from commonalities in parameter settings, while the differences will be proven to be a consequence of the contingent circumstances that surround change each time in each language (LIGHTFOOT, 1999).

In Section 2, I introduce PNS languages and some basic assumptions. In Section 3, I compare the pro-drop system in Brazilian Portuguese (BP) and Modern Russian (MR), pointing out their differences with respect to European Portuguese (EP) and Old Russian (OR), respectively.¹ In Section 4, I show that the emergence of the PNS status in these two languages correlates with the emergence of restrictions in licensing NSs, and explain why that happened. In BP, the old CNS system decayed with the “weakening” of the person inflectional paradigm, triggered by the reorganization of the pronominal system (cf., i.a., DUARTE, 1993, 2000; MODESTO, 2000; NUNES, 2011). In Middle Russian, however, the breaking point was the loss of V-to-T movement, due to the whole rearrangement of the verbal system around new aspectual distinctions, which replaced the old tense distinctions (JUNG, 2018; MADARIAGA, 2022a, 2022b).

2 Consistent Null Subject (CNS) languages and Partial Null Subject (PNS) languages

In recent years, the classical view on the so-called pro-drop parameter has been called into question after observing that its clustering effects, described by, i.a., Rizzi (1982), are less regular across languages than predicted by the standard parametric theory (cf. NEWMAYER 2005). To give an example, the correlation between the existence of NSs and the morphological richness of verbal agreement does not hold in a straightforward way. In addition, “morphological richness” (as opposed to “morphological poverty”) has been shown to be difficult to characterize in a precise way (cf. BIBERAUER, 2008 for an overview).

Recently, Roberts (2010); Biberauer; Roberts; Holmberg; Sheehan (2010); Holmberg; Roberts (2014); and subsequent work, have revisited the properties of pro-drop and formulated new hypotheses about its parametric options. These authors emphasize the fact that PNS languages are especially heterogeneous, which in principle could be considered a problem if we want to propose a common parameter setting for all of them.

¹ The change in pro-drop in Russian took place in the 16th century (between early and late Middle Russian), so I include early Middle Russian under the denomination OR, and late Middle Russian in MR. As for the synchronic description of BP, I follow the convention of Portuguese / Brazilian scholars (BARBOSA; DUARTE; KATO, 2005; DUARTE; SILVA, 2016. i.a.) and consider EP a comparable antecessor of BP.

However, in this paper, I will claim that the heterogeneous nature of PNS languages is not a problem for parametric setting. More specifically, I will show that, if two unrelated languages adopt the same parametric setting of pro-drop, the basic pro-drop properties will be essentially the same in both languages. However, some clustering properties, which are contingent on previous stages of the language, can differ between the two languages. In the following pages, I will prove this hypothesis by comparing the change in pro-drop experienced by two unrelated PNS languages, BP and MR. First, however, let us introduce some basic facts and assumptions about CNS and PNS languages, implied in the change analyzed in this paper.

In CNS languages (e.g., Romance languages, such as Spanish, European Portuguese, and Italian), non-emphatic, non-stressed, non-contrastive pronoun subjects are typically dropped; cf. Peninsular Spanish (1a-b), while stressed, focused, or contrastive subjects must be overtly realized (1c):

(Peninsular Spanish)

- (1) a. – ¿Cómo (*tú) quieres (*tú) la sopa?
 h ow (*you) want (*you) the soup
 – (*Yo) la quiero caliente.
 (*I) cl want hot
 ‘– How do you want the soup? – I want it hot.’
- b. – ¿Ha pasado Juan por casa?
 – Sí, (*él) vino y se llevó el ordenador.
 yes (*he) came and SE took the computer
 ‘– Did Juan come home? – Yes, he came and took the computer away.’
- c. – ¿Se ha llevado Juan el ordenador?
 – No, se lo ha llevado María, *(él) sólo lo ha mirado.
 not cl cl aux taken Maria *(he) just cl aux looked
 ‘– Did Juan take the computer? – No, Maria took it away; he just took a look at it.’

PNS languages, on the other hand, form a heterogeneous group and are not easy to define. In general, the baseline realization of pronominal subjects in PSN languages is overt, but they can be dropped under certain conditions. Depending on the person or tense/mood, PNS languages vary their pro-drop pattern. For example, in Hebrew, Standard Finnish, Jakaltek, and Kenga, 1st and 2nd person NSs are available, as opposed to 3rd person (cf. VAINIKKA; LEVY, 1999), while the reverse

pattern, i.e., availability of 3rd person NSs, as opposed to 1st and 2nd person, is found in Shipibo, historical varieties of Germanic, Dinka, and Lamani (CAMACHO; ELÍAS-ULLOA, 2010; WALKDEN, 2012). The first contrast is illustrated for Finnish in (2a-b):

- (2) a. (Sinä) puhut englantia.
 (you) speak.2sg English
 ‘You speak English’
 b.*(Hän) puhuu englantia.
 ((s)he) speaks.3sg English
 Intended: ‘She/he speaks English.’ (Finnish - Holmberg, 2005, p. 539)

In this paper, I will follow Holmberg (2005, 2010), Holmberg; Nayudu; Sheehan (2009), Roberts (2010), and subsequent work, in that the relevant parametric option for setting a CNS *versus* a PNS language is the acquisition of a [+D] feature on T. I will also follow these authors by assuming that NSs are deficient minimal ϕ Ps with unvalued interpretable ϕ -features ($\phi P_{[i\phi]} : _$), which can in principle receive or not a referential value, and whose behavior varies according to the specific parametric setting on T in the language. The available structures for PNS *versus* CNS languages are described in detail in Madariaga (2022a, 2022b), and summarized here:

(i) In CNS languages, T is endowed with a [+D] feature, and a NS (a deficient ϕ P), located at [Spec,T], automatically receives a referential interpretation by D-matching with T. In these languages, any Topic at CP (null or overt) is able to allow identifying the specific reference of the NS; more specifically, the referential index of a null or overt Topic at [Spec,CP] is copied by the unvalued D-feature of T, and, then, through Agree, by the NS, which at the same time matches its unvalued ϕ -features. The D-feature of T has no morphological expression of its own, and is “spelled out” by realizing person and number features, resulting in rich verbal morphology (HOLMBERG; NAYUDU; SHEEHAN, 2009).

- (3) $[_{CP} \text{Topic}_i C [_{TP} \phi P_i T_D : _ [_{VP} \dots]]]$

In the Spanish examples (1a-b) above, referential NSs are naturally licensed by virtue of the D-feature on T and their topical nature, while a [+focus] or [+contrastive] feature at CP forces the overt realization of the subject (1c).

(ii) In PNS languages, T lacks a D-feature, so that a NS (a deficient ϕ P) is able to match its ϕ -features with T, but does not necessarily raise to [Spec,TP], and does not automatically receive a referential index through T. In that case, the NS gets a generic or arbitrary interpretation. However, there can be some independent mechanism in the language, which is able to endow the deficient ϕ P with a referential index (HOLMBERG; NAYUDU; SHEEHAN, 2009). In languages like Finnish, MR, and BP, there seem to be at least two alternative mechanisms to achieve this referentiality:

(a) The first mechanism is related to pragmatics: NSs can receive a referential interpretation by copying the index of some null topic at C. This null topic is part of the discourse common ground, shared by the speaker and the hearer, i.e., it can convey either deictic (situational / contextual) features or logophoric features (for the 1st and 2nd person), which allow the reference of the NS to be recovered (SIGURÐSSON, 2011; TSEDRYK, 2015). Depending on the particular conditions of the specific PNS language, it can be the case that referential NSs are licensed only in one or two grammatical person(s), or that all persons are licit in different pragmatic situations (TSEDRYK, 2022).

(4) [_{CP} Topic_i C [_{TP} ϕ_i T... [_{VP} V...]]]

A subtype within pragmatically licensed referential NSs are NSs licensed as a part of a “topic chain” (FRASCARELLI, 2007), that is, as successive occurrences of a previous overtly expressed Sentence Topic (REINHART, 1981). This kind of topic chain is characteristic of MR, albeit in a very restrictive way, and licenses the presence of 3rd person NSs in matrix clauses, as we will see in Section 3.1.

(b) The second mechanism available in PNS languages involves embedded clauses: an embedded NS can enter an anaphoric relation with respect to a c-commanding antecedent in the matrix clause, and thus receive a referential interpretation. This phenomenon is also known as finite control; see, i.a., Landau (2004) for Hebrew; Rodrigues (2004), Boeckx; Hornstein; Nunes (2010) for BP; Livitz (2014); Tsedryk (2015) for MR. The embedded NSs in these contexts must be c-commanded by a coreferent antecedent in strict locality.

(5) [DP_i ... [CP Op_i C [TP_θ P_i T ... [VP V ...]]]]

In Section 4, I will revisit the structures introduced in this section, according to the specific languages and stages analyzed.

3 A compared characterization of NSs in EP-BP and OR-MR

In this section, I will revise the “bundle of properties” that often go hand-in-hand with pro-drop, and which tend to cluster in a similar way in PNS languages (BIBERAUER, 2008). I will check these properties in BP and MR, and compare them with their CNS “ancestors”, EP, and OR, respectively (cf. fn. 1). The properties reviewed are the following: (i) pragmatically licensed referential NSs; (ii) finite referential embedded NSs; (iii) NSs with arbitrary and generic readings; (iv) null objects; (v) clitic pronouns; (vi) verbal inflection; and (vii) the role of infinitives.

3.1 Pragmatically licensed referential NSs

MR licenses 1st and 2nd person NSs related to logophoric topical features, on a hearer-speaker discourse basis (6a), in a similar way to several Germanic languages, analyzed by e.g., Sigurðsson (2011). According to Duarte (1993), written BP allows this kind of drop in restricted cases, e.g., with negation or in a verbal locution, as in (6b), although some authors indicate that NSs in root sentences are not generally licensed in BP (MODESTO, 2000), or that they are on their way to being lost (DUARTE; SILVA, 2016):

- (6) a. Privet! Xorošo, što (vy) prišli!
 hi well that (you) came.pl
 ‘Hi! So good that you came!’ (MR - MADARIAGA, 2022a, p. 79)
- b.(Eu) não posso mais ficar aqui a tarde toda.
 (I) not can.1sg more stay here the evening whole
 ‘I cannot stay here the whole evening any more’ (BP - DUARTE, 1993, p. 119)

In MR, this kind of dropping is blocked by any fronted referential overt argumental NP introduced between the CP and the NS (LIVITZ, 2014; TSEDRYK, 2015). In example (7), the fronted object blocks referential subject drop, and the only way to interpret the NS is as a generic in an impersonal sentence (see Section 3.3 below):

- (7) Svetu *pro* nedavno videli na rynke.
Sveta.acc recently saw.pl at market

‘Sveta was seen recently at the market.’ / ‘*We saw Sveta recently at the market.’ (MR - TSEDRYK, 2015, p. 349)

Additionally, MR licenses 3rd person NSs related to Given or Familiar Topics or as successive occurrences of a Sentence Topic in a topic chain (LIVITZ, 2014; TSEDRYK, 2015), as in (8a). As in the previous case, any fronted referential NP can block subject drop (8b).

- (8) a. Onii idut na ozero. *pro*_i Nadejutsja tam vstretit’ Ivanaj.
they.nom go.3pl to lake hope.3pl there see Ivan.acc
*(On)_j obeščal im peredat’ ključi.
he.nom promised.m.sg them pass keys

‘They are going to the lake. (They) hope to see Ivan there. He has promised them to pass the keys.’ (MR - TSEDRYK, 2022, p. 42)

- b. Jai tol’ko čto videl Svetuj. Mnei *(onaj) skazala,
I.nom just saw Sveta.acc me.dat she.nom said.f.sg
čto naš dom uže prodan.
that our house already sold

‘I have just seen Sveta. She told me that our house had already been sold.’ (MR - TSEDRYK, 2015, p. 349)

Notice that Sentence Topics themselves cannot be dropped in MR; thus, in example (8a), the subject of the second sentence in this sequence is the first identical occurrence of the initial subject *oni* “they” in a topic chain and, therefore, can be dropped. However, the topic chain finishes here, because the third sentence introduces new propositional content; i.e., it shifts from “us going to the lake and doing something there” to “information about Ivan”. Thus, the new subject, *Ivan*, qualifies as a Sentence Topic and cannot be dropped. In example (8b), the presence of a fronted indirect object (*mne*) blocks subject drop, and the pronominal subject must be overtly realized.

In BP, it seems that a subject can be dropped only in those cases in which it is topically very prominent, as in question-answer pairs (9a). Barbosa; Duarte; Kato (2005) also report examples of 3rd person referential NSs in BP tied to Topics; they state that it is required “that the antecedent is in an adjacent sentence”, as shown in (9b). The latter are very similar to those NSs inserted in topic chains in MR (8a).

- (9) a. – E o João? – *pro*_i Viajou.
 and the Joao? travelled.3sg
 ‘– What about Joao? – He travelled.’ (BP - DUARTE; SILVA, 2016, p. 239)
- b. O homem_i finge que é um certo tipo de homem
 the man pretends that is a certain type of man
 para escrever. Ou seja, *pro*_i trai o homem.
 to write or be betrays the man
 ‘A man pretends to be a certain type of man in order to write. In other words,
 he betrays himself.’ (BP – BARBOSA; DUARTE; KATO, 2005, p. 26)

As for OR and EP / early BP, as well-behaved CNS languages, they show no locality effects in licensing referential NSs. Example (10a) shows that, in OR, a NS can take its reference not from the closest subjectual antecedent, i.e., the patriarch, but from a previous Sentence Topic, Olga. In MR, there would need to be an overt subject “she” in the place of the NS; otherwise, the antecedent of the NS would have to be interpreted as the patriarch, the last overt subject previously mentioned. Similarly, in EP an overt pronoun tends to be avoided unless its reference is impaired (BARBOSA; DUARTE; KATO, 2005), but the lack of proximity, or the syntactic position of the antecedent, as well as the presence of intervening elements, as in (10b), are not a problem for a NS to be licensed. According to Duarte (1993), in early BP texts, the rate of NSs was similar to that of EP, so we can assume that the situation was similar in both languages.

- (10) a. I bl(ago)s(lo)vi jui patreargxj. I *pro*_i ide
 blessed.3sg her.cl.acc.f patriarch.nom and went.3sg
 s miromъ въ svoju zemlju. (OR - *Laurentian Chronicle* 18)
 with peace to own land
 ‘And the patriarch blessed her (=Olga). And (she) went to her land in peace.’
- b. As coisas não têm corrido tão bem como [as Nações Unidas],
 the things not have run so well as the nations united
 queriam, (...). A violência não pode ser dinamizadora
 wanted.3pl the violence not can be propeller
 da mudança. Mas muito rapidamente *pro*_i concluíram que
 of the change but very quickly concluded.3pl that
 todas as condições estavam reunidas.
 all the conditions were gathered
 ‘Things are not going as well as the United Nations expected (...). Violence cannot be the propeller of change. But very quickly they (=the UN) concluded that all the conditions were gathered.’ (EP – BARBOSA; DUARTE; KATO, 2005, p. 27.)

3.2 Coreferent NSs in embedded finite clauses

As in other PNS languages (e.g., Finnish, Hebrew, and Marathi; cf. HOLMBERG; NAYUDU; SHEEHAN, 2009; LANDAU, 2004), both BP and MR display coreferent NSs in embedded finite clauses (i.a., LIVITZ, 2014; MODESTO, 2009; NUNES, 2010; TSEDRYK, 2015). In both cases, these NSs are licensed only in strict conditions of control, i.e., c-commanded by the closest antecedent in a sort of “control” relation (11b, 12b).²

- (11) a. Vrač_i skazal, čto pro_{i,ej} primet bol'nyx_j.
 doctor said that will receive.3sg sick people
 ‘The doctor said that he will see the patients.’ (MR – MADARIAGA, 2022a, p. 83)
- b. *Majja_i boitsja, čto roditelij dumajut, [čto proi ne
 Maia fears that parents think that not
 pridët vovremja].
 will come.3sg on time
 Intended: ‘Maia is afraid that her parents think that (Maia) won’t arrive on time’ (MR – MADARIAGA, 2022a, p. 85)
- (12) a. Só o João_i acha que proi/*j vai ganhar a corrida.
 only the João thinks that goes win the race.
 ‘Only João thinks that he (=João) will win the race.’ (BP – NUNES, 2011, p. 333)
- b. *O paii [da Mariaj] acha [que proj está grávida]
 the father of the Maria thinks [that is pregnant.f]
 Intended: ‘Maria’s father thinks that she is pregnant.’ (BP – NUNES, 2011, p. 333)

As in the case of matrix clauses, in OR and EP, embedded NSs are freely licensed in non-stressed non-contrastive positions, and do not display locality or c-commanding effects. In (13a), I show the lack of a local c-commanding antecedent of an embedded NS in OR; example (13b) illustrates the same lack of locality in EP:

² Some authors analyze these instances as finite Obligatory Control (BOECK; HORNSTEIN; NUNES, 2010; LIVITZ, 2014; MADARIAGA, 2018; TSEDRYK, 2013). Other hypotheses for this structure have been put forward by, e.g., Modesto (2009), Landau (2004), and Sheehan (2018). I will leave aside this discussion and stick to the fact that embedded finite NSs in both languages require a local c-commanding antecedent.

- (13) a. Slyšavše že derevljanei jako proj opjat' idet'ь...
 earing.pl part Drevlians that again come.3sg
 ‘When the Drevlians heard that (he = prince Igor) was coming again...’ (OR - Laurentian Chronicle 14R)
- b. O amigo_i do Pedro_k disse que pro_{ij/k} ganhou na loto.
 the friend of the Pedro says that won in the lottery
 ‘Pedro’s friend says that he (=Pedro, his friend or someone else) won the lottery.’ (EP – Modesto, 2000, p. 151)

3.3 NSs with generic or arbitrary reading

According to Holmberg (2010), 3rd person singular NSs automatically get an indefinite reading in PNS languages, as opposed to CNS languages, which resort to other mechanisms to achieve arbitrariness or genericity of the NS, like the Romance *se*-strategy. This is true for BP (14a-b), which patterns with Finnish and Marathi in this respect, contrasting with EP (14c).

- (14) a. Aqui vende sapato.
 here sells.3sg shoes
 ‘Shoes are sold here.’ (BP - DUARTE; SILVA, 2016, p. 242)
- b. É assim que faz o doce.
 is thus that makes.3sg the sweet
 ‘This is how one makes the dessert.’ (BP – HOLMBERG, 2010, p. 92)
- c. É assim que se faz o doce.
 is thus that se makes.3sg the sweet
 ‘This is how one makes the dessert.’ (EP – HOLMBERG, 2010, p. 92)

In MR, however, only 3rd person plural NSs are available in impersonal sentences with an indefinite subject (15a), i.e., 3rd person singular is not available (TSEDRYK, 2022). For impersonal sentences with a generic reading, impersonal sentences of the type of (15b), with an adverbial plus an infinitive, are generally used.³

³ A further available variant is a generic NS displaying 2nd person singular agreement, which is disregarded here; for an analysis in MR, cf. TSEDRYK (2022).

- (15) a. Tut prodajut / *prodaēt obuv'. (MR)
 here sell.3pl / *sell.3sg shoes
 'Shoes are sold here.'
- b. Na stule bylo udobno / možno bylo sidet'.
 on chair past.3sg comfortably / possible past.3sg sit.inf
 'It was comfortable / possible to be sitting on the chair.'

OR functioned in the same way as MR; NSs could be used with an arbitrary reading with a 3rd person plural verb (16a) (BORKOVSKIJ, 1978), while generic NSs did not exist and impersonal clauses of the “adverbial + infinitive” type were used instead (16b).

- (16) a. Ože ubĭjutъ novgorodca posla za morem...
 if kill.3pl Novgorodian consul beyond sea
 'If someone kills a delegate from Novgorod in a foreign land...' (OR – Novgorod letter 1189-1199, ap. BORKOVSKIJ, 1978, p. 221)
- b. Ne bĕ lzĕ slyšati pĕnĭje vo plači.
 not past.3sg possible hear.inf song in crying
 'It was not possible / one could not hear the chants because of the crying.' (OR – Laurentian Chronicle 68)

As stated before, EP does not license 3rd singular person generic or arbitrary NSs (like OR / MR), and uses either 3rd person plural to convey arbitrariness of the subject or a *se*-impersonal strategy for generic NSs (DUARTE; MATOS; GONÇALVES, 2005; DUARTE; SILVA, 2016). The first strategy, illustrated in (17a), is fully available in BP. The second strategy, with the *se* pronominal (17b and 14c above), was not completely lost in BP, but it gave rise to a sort of diglossia as it also developed into an impersonal construction with a 3sg NS (14b).

- (17) a. Dizem que o governo vai aumentar os impostos.
 say.3pl that the government goes raise the taxes.
 'People say that the government will raise taxes.' (EP - DUARTE; SILVA, 2016, p. 242)
- b. Trabalha-se demasiado.
 work.3sg-se too much
 'One works too much.' (EP – DUARTE; MATOS; GONÇALVES, 2005, p. 119)

3.5 Existence of clitic pronouns

Clitic pronouns were completely lost in MR, but not in BP.⁴ However, in regular conditions, in BP, cliticization displays proclisis (19a), while EP resorts to enclisis (19b).

- (19) a. Me diz uma coisa! (BP / *EP)
 me say one thing
 b. Diz-me uma coisa! (EP / *BP)
 say me one thing

‘Tell me something!’ (NUNES, 2011, p. 343)

In OR and EP, clitics display the typical behavior of CNS languages, that is, they are phonologically deficient pronouns used in non-stressed positions, with a non-contrastive or non-focal value (like NSs). I will assume that they are generated in VP and then raise higher to find a proper phonological support, depending on the specific conditions in each language (BLEAM, 1999; URIAGEREKA, 1995, among many others). In the case of OR and EP, the neutral position of object clitics is enclisis (19b, 20a-b), although in both languages, some specific elements, such as *wh*-words, sentential negation, complementizers, etc., trigger clitic proclisis (21a-b):

- (20) a. poneže ljublju tja pače brati tvoeje.
 because love.1sg you.cl.acc more brothers yours
 ‘Because I love you more than your brothers.’ (OR - Laurentian Chronicle 72)
 b. O João disse-o à Maria.
 the João told-cl.acc.3sg to Maria
 ‘João told it to Maria.’ (EP – DUARTE; MATOS; GONÇALVES, 2005, p. 122)
- (21) a. na čto mja este pribavili, ose esmъ.
 for what me.cl.acc aux.2pl call.pl here am.1sg
 ‘Why did you call me? Here I am.’ (OR – Laurentian Chronicle 92)

⁴ However, in the specific case of 3rd person accusative clitics in BP, Corrêa (1991, apud NUNES, 1993, 2011), claims that they are not “native,” i.e., are not acquired by children as a part of their core grammar, but are instead acquired artificially via schooling. As we will see in Section 4, according to Nunes (1993, 2011), the factual loss of 3rd person accusative clitics in colloquial BP was due to a change in the directionality of cliticization with respect to EP. As Nunes (2011) shows, even in formal BP these clitics differ from 1st / 2nd person clitics and from their counterparts in EP.

b. O João não lhe mandou a carta.
 the João not cl.dat.3sg sent the letter

‘João did not send him/her the letter.’ (EP – DUARTE; MATOS; GONÇALVES, 2005, p. 128)

3.6 Verbal inflection

EP / early BP and present-day BP differ with regard to the richness of their inflectional patterns. In most early BP finite paradigms, just as in EP, leaving aside the polite forms for the 2nd person, all the six forms available are unambiguous, as shown in Table 1. In present-day colloquial BP, finite inflection displays only one unambiguous form out of three total forms; in some moods or tenses, there are no unambiguous forms out of just two total forms (DUARTE, 1993, 2000; NUNES, 2011).

In contrast, both Old and Modern Russian display full-fledged forms in non-past tenses, as shown for MR present tense in Table 1. This pattern remained essentially the same as in OR. However, MR has defective inflection (only gender and number, no person) in past tense.

Table 1 – present tense paradigm in EP and BP (adapted from Duarte 1993, p. 109), and present and past tense in MR. Verb *cantar / pet’* ‘to sing’

Person and number	Early BP = EP	Colloquial BP	MR non-past	MR past
1st sg	Eu canto	Eu canto	Ja poju	Ja/ty/on pel (m.) Ja/ty/ona pela (f.) (Vy peli - polite)
2nd sg	Tu cantas	Você/(tu) canta	Ty poěš’	
2nd sg polite	Você canta	Você canta	Vy poěte	
3rd sg	Ele/ela canta	Ele/ela canta	On/ona poët	My/vy/oni peli (pl.)
1st pl	Nós cantamos	A gente canta	My poëm	
2nd pl	Vós cantais	Vocês cantam	Vy poěte	
2nd pl polite	Vocês cantam	Vocês cantam	Vy poěte	
3rd pl	Eles/elas cantam	Eles/elas cantam	Oni pojut	

As for OR, all the tenses (present, future, and past) were fully-inflected. The past perfect tense was a compound form with distributed features: a participle (so-called *l*-form) inflected for number and gender plus an auxiliary clitic form inflected for person and number (21a, 22). As we will see, the loss of the clitics in the language, including auxiliaries, conveyed the loss of person agreement only in past perfect tense (rendering the MR defective inflectional past pattern illustrated in Table 1). A second (synthetic) past form, the aorist, also fully-fledged, existed in early OR, but had decayed by the 14th century. Here is an early

- [*pro*_i počti c(ěza)rja vaše(go)] (. . .) da budutŝ.
 honour.inf tsar your let be.pl
 ‘If you need to gather an army (lit. to go to war), and these (=the Russians) want to join your
 tsar (. . .), so be it.’ (OR – *Laurentian Chronicle*, 18R)
- (25) a. Vižŝ sego ty ježe esi хотěŝ.
 look this you what aux.2.sg wanted.m.sg
 ‘See, this is what you (i.e., but not me) wanted.’ (OR – *Laurentian Chronicle*, 23R,
 MADARIAGA, 2018, p. 181)
- b. Počto *pro*_i ideši opjats, *pro*_i poimaŝ esi
 what for go.2.sg again took.m.sg aux.2.sg
 vsju danŝ.
 whole tax
 ‘Why did you come back? You collected the tribute already’. (OR – *Laurentian Chronicle*,
 14R, MADARIAGA, 2018, p. 181)

Overt infinitive subjects alternating in this way are not found in EP or early BP. As we will see in Section 4.1, this property of OR will be significant for the rise of embedded control in MR.

3.8 Interim summary

In Table 2, I offer a revised summary of the similarities and differences between the four relevant languages with respect to the features usually ascribed to the pro-drop parameter:

Table 2 – clustering properties ascribed to the pro-drop parameter in OR, MR, EP, and BP

	OR	MR	EP	BP
1. Referential NSs	free	restricted	free	restricted
2. Finite control	no	yes	no	yes
3. Generic 3sg NSs	no	no	no	yes
4. Null objects	no	yes	restricted	yes
5. Clitic objects	yes	no	yes	restricted
6. Finite verbal person inflection	yes	yes (only non-past tenses)	yes	very defective
7. Inflected infinitivals	no	no	yes	very defective

In view of these correlations, we can affirm that the change in Russian pro-drop affected defective referential pronouns (referential null subjects, including finite control, and clitics / null objects), but it did not have a strong impact on generic/arbitrary NSs or verbal inflection.

In BP, all defective subjects (referential or generic) changed, together with a clear impoverishment in verbal inflection (finite and non-finite), and perhaps minor consequences concerning defective objects (null or clitic). In the following section, I offer a unified account for these facts.

4 An explanation for the change in pro-drop in BP and MR

In this section, I will develop the idea that variation among PNS languages is the effect of historical change, whose contingent properties can partially shape grammars.⁵ The fact that pro-drop is triggered by the confabulation of a few recurrent factors, which can be diachronically modified in different ways, gives us the slight differences between BP and MR with respect to their pro-drop properties, rendering thus a satisfactory explanation for the parametrical “mismatches” we observe in these PNS languages.

I will show: (i) how the final setting of the PNS parametric stage was essentially the same in both BP and MR; and (ii) in which way the triggers of the change were different in the two languages, and why the clustering ways of the properties reviewed above also differed at some points.

4.1 Structures and change in NS licensing in Russian

The most striking fact about referential NSs in Russian (27a-b), contrasting with Hebrew (26a-b) or Chamorro, is that their conditions of licensing, described in Section 3.1, are completely equal regardless of the richness or the poverty of personal inflection on the verbal form. That is, whether the verbal form displays morphologically overt personal features (in non-past tenses, endowed with overt personal morphology, 27b) or not (with past impoverished *l*-forms, lacking personal morphology, 27a), the

⁵ I assume a Lightfootian (LIGHTFOOT, 1999, and subsequent work) approach to change. In neutral conditions, any modification of the Primary Linguistic Data (PLD) can lead learners to diachronic change, i.e., parse a different grammar with respect to previous generations. However, change is contingent on the unpredictability of the external factors conditioning it. The most frequent triggers modifying the PDL that learners receive are the following: (i) changes that have taken place previously in the language and can condition the “clustering” effects, like the ones we are considering in this paper (traditionally termed “chains of changes” in historical linguistics); (ii) phonological erosion or attrition (which can have an impact on morphological material); (iii) language contact and other sociolinguistic factors.

conditions for licensing referential NSs apply equally; in example (27), logophoric features license dropping of any of the subjects. In Hebrew, however, rich agreement licenses subject dropping (26a), while defective non-personal agreement does not (26b):

- (26) a. Etmol šama't harca'a.
 yesterday hear.past.2sg.f lecture
 'Yesterday you heard a lecture.' (Hebrew – DORON, 1988, p. 205)
- b. *Axšav šoma'at harca'a.
 now hear.pres.sg.f lecture
 Intended: 'You are hearing a lecture now.' (Hebrew – DORON, 1988, p. 207) (MR)
- (27) a.(Ty) v magazin xodil, ili doma sidel ves' den'
 you to shop go.past.m.sg or home sit.past.m.sg
 all day
 pered televizorom?
 before TV
 'Have you gone to the shop or were you sitting in front of the TV all day long?'
- b. (Ty) v magazin pojděš' ili doma budeš' sidet'
 you to shop go.fut.2sg or home fut.2sg sit
 pered televizorom?
 before TV
 'Will you go shopping or are you going to be sitting in front of the TV?'

The diachronic process that gave rise to the PNS character of Russian evidences the fact that the role of verbal inflection in the change was marginal (cf. MEYER, 2011; JUNG, 2018, *contra* MÜLLER, 2006); changes in verbal inflection did not constitute the initial trigger for the change and, in fact, verbal inflection only changed in past tenses (as a side-effect of a previous change, as we will see), whereas present and future tenses preserved their inflection paradigms completely untouched.

Let us briefly review the process of loss of the referential NSs in Russian step-by-step. The following description is based on the works by Borkovskij (1978), Borkovskij; Kuznecov (1965), Ivanov (1990), Jung; Migdalski (2015), Kibrik (2013), Meyer (2011), Migdalski (2013), and Zaliznjak (2008).

- (1) Old Russian was a well-behaved CNS language (cf. Section 3). Referential subjects were dropped in pragmatically “neutral” conditions, regardless of the type of topic characterizing the NS, as shown in example (10a, 25b) above. Referential subjects had to be overtly realized as pronouns only in emphatic positions (foci or contrastive topics, ex. 25a). This was the situation until approximately the 14-15th century.

The verbal system was inflectionally rich, there existed 4 past forms, a present, and several forms to express the future. However, in OR, some tense paradigms started to fall into disuse: in early Slavic there were two synthetic past forms (the imperfect and the aorist) and two analytic forms (perfect and pluperfect). Analytic verbs were formed by an auxiliary conveying person and number morphemes, and a participle (an *l*-form) conveying the lexical verbal content plus gender and number morphemes (examples 21a and 22 above). The auxiliaries corresponding to the 3rd person singular and plural had already been lost by early OR, and the synthetic past forms were in decay, too: the imperfect was almost inexistent, and the aorist was archaic, restricted to literary language. The fall into disuse of most tense distinctions led to the eventual rearrangement of the Russian verbal system in favor of aspectual distinctions (the MR pattern) rather than tense distinctions.

Deficient pronouns other than subjects, as well as auxiliary verbs, were formally clitics. Auxiliary verbal forms, just like other clitics, were raised to a high position in the sentence, as in examples (21a) and (25a) above. Synthetic verbal forms were also raised higher (compare the OR example (28) and the later Middle Russian example (29b) below). This high verbal position is common to other CNS languages and is usually considered to reflect the existence of V-to-T movement in these languages.

- (28) i *pro* poklaneju ti se.
and bow.1sg you.cl.dt refl.cl.acc

‘I greet (lit. bow) you.’ (OR – Birch bark letter 798, 12th c., JUNG, 2018, p. 105)

Pronominal clitics in OR raised to be adjacent to the raised verb or auxiliary (MIGDALSKY, 2013; JUNG, 2018). They were used in the same pragmatic situations as NSs, that is, in non-emphatic non-stressed positions (28), while strong pronouns were used in contrastive or focused positions.

(2) Early Middle Russian: by the 14-15th century, verbal auxiliaries started to surface in a lower position (29a), and synthetic verbs also remained low in the structure (29b); compare the position of the synthetic verbal form in the OR example (28) and the Middle Russian one (29b). In other words, long verb movement got lost.

- (29) a. *estʹ u menei edinʹ s(y)nʹj doma menšii, a pro_i sʹ četyrmi*
 is at me one son at home young and with four
 esmʹ vyšelʹ, a onʹj doma.
 aux.1sg left.m.sg and that at.home
 ‘I have a little son, and I came here with my other four children, and the young one stayed at home.’
 (Middle Russian – *Hypathian Chronicle* 46, 14th c.)
- b. *jěza tobe koloneju-sę.*
 I.nom you.strong.dt bow.1sg-refl
 ‘I greet you.’ (Middle Russian – Birch bark letter 501, 14th c., JUNG, 2018, p. 105)

Very soon, between the 15th and 16th centuries, verbal auxiliaries (of 1st and 2nd person; remember that 3rd person auxiliary had been lost long ago) were lost in the language, after a short period of lowering. At the same time, pronominal clitics were also being lost, and being replaced with strong pronominal forms. Auxiliary clitics were not eroded phonologically or progressively worn out; the whole forms, all of them disyllabic (sg. *esmʹ, esi*, pl. *esme, este*), were lost altogether (30). Almost simultaneously, overt pronouns of 1st and 2nd person were realized in neutral positions (non-emphatic, non-stressed), as in (30) and (29b), in the places where formerly only NSs were used (28).

- (30) A az stal v dolu s polkom,
 and I.nom stood.l-form.m.sg in valley with army,
 a Vasilju prikazal...
 and Vasili.dt ordered.l-form.m.sg
 ‘And I stood in the valley with the army and I ordered Vasili...’ (Middle Russian – *Pervoe pis'mo Vasilija Grjaznogo Ivanu IV Groznomu* 1576, 20)

This rendered a new pattern with “impoverished” personal morphology, but only in past tense, which happened to be the one formerly including a clitic auxiliary (which, in turn, happened to be the part of the verbal complex bearing personal morphology). At the same time, present and future personal morphology remained intact but,

nonetheless, showed similar rates of overt non-emphatic pronominal subjects as past impoverished forms (MEYER, 2011). In other words, overt non-emphatic pronouns were realized in the place of NSs at a fixed rate regardless of the “richness or poverty” of agreement in the specific verbal form in the sentence.

- (3) Late Middle Russian: by the 16-17th centuries, when the modern system of 1st and 2nd pro-drop was already well-established, 3rd person sentences experienced the same shift as 1st and 2nd person. 3rd person verbal auxiliaries in perfect > past tense had already been lost in early OR (approx. 5-6 centuries before), but no personal pronoun existed in the language for the 3rd person. In stressed or emphatic positions, when the pronominal subject had to be overtly realized, various demonstrative pronouns were realized instead. One of these demonstratives, *онъ* ‘that’, was reanalyzed in Late Middle Russian as the personal pronoun of 3rd person and immediately spread to non-emphatic and non-stressed positions (31), following the pattern of the 1st and 2nd person pronouns.

- (31) Prosi y nego na brašna deneg, i on tebe dast sto rublej.
ask from him for food money and he you will give hundred rubles
‘Ask him for money for food, and he will give you a hundred rubles.’ (Middle Russian – *Story about Karp Sutulov*, 17th century)

As for pronominal clitics, 1st and 2nd person object clitics were lost by the 15-16th century, while 3rd person clitics were lost approximately a century later. The delay in the loss of 3rd person clitics as compared to 1st and 2nd person parallels the development of 3rd person vs. 1st / 2nd person overt pronominal pronouns in the place of the old NSs.

Null objects are found in texts roughly by the 16th century, and became more common later on:

- (32) A ženix_i po nevestu_j ne ezdit, a privezet pro_j družka da svaxa.
and groom for bride not goes but carries best man and matchmaker
‘The groom does not go for the bride; the best man and the matchmaker bring (her).’ (Middle Russian – *Putešestvija russkix poslov*, BORKOVSKIJ, 1978, p. 314)

In Table 3, I summarize the development of the change in the relevant pro-drop features in Russian:

Table 3 – diachronic succession of changes in the pro-drop system between OR and MR

When	Changes		
11-12 th c	Loss of 3 rd person auxiliary	Loss of past tenses in favor of aspectual distinctions	
15 th c	Loss of long verb movement = lowering of auxiliaries and synthetic verbs		
15-16 th century	Loss of 1 st /2 nd auxiliaries	Extension of 1 st /2 nd overt pron. subjects	Loss of 1 st /2 nd pron. object clitics
16-17 th century	Extension of <i>онъ</i> as 3 rd p. overt pronoun	Loss of 3 rd person object clitics	Extension of null objects

Now let us interpret the change in formal terms (cf. MADARIAGA, 2022a; 2022b). As I explained in Section 2, in OR, the minimal ϕ P subject, located at [Spec,T], automatically received a referential interpretation from a higher topic by D-matching. The referential index was copied by the unvalued D-feature of T, and, then, through Agree, by the NS, which at the same time matched its unvalued ϕ -features. The D-feature of T was “spelled out” in the form of rich verbal morphology. Overt V-to-T movement conveyed the presence of rich agreement morphology high in the sentence, which constituted a clear cue for learners to establish the presence of a [+D] feature on T.

- (33) [_{CP} Topic_i C [_{TP} ϕ _i TD: _+V [_{VP} V...]]]

The loss of 3rd person auxiliaries and the obsolescence of most past forms in early OR led to a reorganization of the tense – aspect system in Russian. Because of the loss of 3rd person verbal auxiliaries, analytic past forms in 3rd person could seem to speakers like low “synthetic” *l*-forms, marked for gender and number (not person), which did not raise to T. V started to be interpreted as undergoing short movement (to Asp or a similar intermediate node, as it does nowadays (cf. GRIBANOVA, 2013); confirmed by experimental work by Kallestinova; Slabakova (2008)), and the old tense-based system disappeared in favor of a new system based on aspect distinctions. As for non-past synthetic forms, they preserved person marking but, analogically to past *l*-forms, they stopped moving to T. 1st and 2nd person auxiliaries also displayed a low position at this time (ZALIZNJAK, 2008). In other words, V-to-T movement was lost.

The loss of long verb movement modified the cue that learners needed to posit an unvalued D-feature on T, that is, the overt

morphological realization of the operation of D-feature valuing, which was person agreement on T (cf. a similar idea in Jung, 2018, who proposes D-feature lowering rather D-feature loss). The loss of overt morphology at T prevented the acquisition of T as playing a role in D-feature transmission together with Agree. In addition, once V remained in AspP, V's low position preempted the establishment of the direct syntactic relation between T and C the way it did before. The loss of the D-feature on T was completed between the 15th and 16th centuries, with the total loss of personal auxiliaries in the language.

After T's D-linking ability was lost, pronominal subjects had to be overtly realized in order to be interpreted as referential. However, learners of Middle Russian kept on receiving referential subject gaps in their linguistic input in the form of NSs generated by older generations of speakers, and subject gaps in control infinitive clauses (PRO). As a common output in historical processes, residual structures and elements after diachronic change can die out or be "recycled" (reanalyzed) with a new value. The old NSs in Russian experienced the second path. Speakers of Middle Russian, in the absence of a D-linking T, had two ways to acquire the finite subject gaps they received: (i) as generic or arbitrary NSs (they retained 3rd person plural, as in OR), or (ii) as referential NSs when they were able to find a proper mechanism of index transmission, whether a prominent or logophoric null topic at CP, or a higher c-commanding referential antecedent in the case of embedded NSs.

As for other non-subjectual deficient pronominals, namely, pronominal clitics, the change reviewed here also had a direct impact on them. I will assume the classic proposal by Cardinaletti; Starke (1999); namely that prosodic, phonological, and morphological deficiency of clitics correlates with less syntactic structure, as compared to full pronouns. Thus, a clitic is a minimal noun projection, i.e., just an agreement projection. In our terms, clitics would have the same basic structure as NSs, i.e., minimal ϕ Ps with unvalued interpretable ϕ -features (ϕ P[$i\phi$: _]). Cf., i.a., BLEAM, 1999; DUARTE; MATOS; GONÇALVES, 2005; ROBERTS, 2010; BOŠKOVIĆ, 2016, for a similar idea.⁶

⁶ OR deficient pronouns qualify as ϕ P clitics, rather than mere agreement markers (cf. contrastive properties of the two types of clitics in Bleam, 1999; Ormazabal; Romero, 2013; and references therein). Accusative clitics in Slavic, including its earlier periods,

As in the case of NSs, OR clitic ϕ Ps are not D-complete elements, so they must have their reference defined in the course of the derivation. As discussed before, the reference of a minimal ϕ P is determined by some (null or overt) Topic at CP. Getting a reference from this high Topic conveys also ϕ -feature valuation on the clitic, as in the case of NSs. In the case of CNS languages, licensing minimal ϕ Ps (NSs and clitics) is tied to the presence of an unvalued D-feature on T (HOLMBERG; NAYUDU; SHEEHAN, 2009), which copies the referential index of the higher Topic, and transmits it downwards during the Agree operation of ϕ -feature matching by the ϕ P. In the case of subjects, as a result of Agree, the valued features of T are “spelled out” in the shape of rich verbal morphology, and NSs can be realized as null. However, objectual ϕ Ps can receive their reference, but they lack the Agree relation with T described above for subjects, so they must be overtly realized, albeit in a prosodically deficient way.

In the specific case of OR and other Slavic languages, following, i.a., Progovac (1999); Migdalski (2016); Bošković (2016), clitic licensing is parasitic on verb movement, like *pro* subjects in CNS languages. Being prosodically deficient, in the specific case of OR, objectual ϕ Ps had the requirement to be adjacent to the verb (JUNG; MIGDALSKI, 2015). Thus, when the verb raised in OR, the ϕ P moved together with the verb up to their final common landing site, in which V attached to T. In this pattern, both elements moved “for a reason”; the V to get its inflectional suffix, and the clitic to get prosodic support and referential interpretation.

$$(34) \quad [_{CP} \text{Topic}_1 \text{Topic}_2 \dots C [_{TP} \phi P_1 (=NS) T-v-V \phi P_2 (=clitic) [_{VP} \phi P_4 v+V-\phi P_2 [_{VP} V-\phi P_2]]]]$$

After V-to-T movement was lost, clitics could not raise, and had to remain low in the structure, together with their prosodical host, the verb. The clitics’ new low position, however, led to their unavoidable loss, as they were too far from C, and could not get a suitable reference. Therefore, very soon clitics were completely replaced by either (i) full, phonologically independent, and syntactically complete pronouns, or (ii) null objects, which, following the new reanalyzed pattern of the other null ϕ Ps in the language, the NSs, could raise as much as they needed

cannot double DPs or override Principle B, in contrast to clitics in other languages, which are better analyzed as agreement elements.

(being silent, they did not need phonological support), and thus could be licensed by a suitable topic at C.

As for embedded finite “control”, as explained in detail in Madariaga (2018), Russian learners started to identify embedded referential subject gaps in finite and non-finite clauses, whenever the embedded NS was c-commanded by an antecedent in the matrix clause (in coreference). In other words, embedded finite referential coreferent NSs were reanalyzed as “controlled”, following the model of their embedded non-finite counterpart, PRO. The reason for learners to identify embedded finite and non-finite coreferent subject gaps was the loss of the alternation existing in OR between coreferent emphatic overt pronouns and coreferent non-emphatic null pronouns in embedded infinitive clauses (24a-b), which had been thus far the same as in finite clauses (25a-b). When the old CNS pattern of “free” referential NSs was lost, and overt pronouns started to be realized in non-emphatic positions, the old alternation holding in both finite and non-finite clauses disappeared, and a suitable way to preserve referential subject gaps in embedded finite contexts was by reanalyzing coreferent subject gaps as in non-finite sentences, following the model of PRO. As a second side-effect of the loss of this alternation, in Late Middle Russian, all the embedded non-finite subject gaps that did not fit the pattern of PRO, i.e., non-coreferent overt subjects in infinitive sentences, were completely lost (BORKOVSKIJ, 1978).

Finally, generic/arbitrary NSs did not play a role in the change process. Impersonal infinitive clauses and 3rd person plural arbitrary sentences remained essentially the same in MR as they were in OR. The only difference, derived from the loss of the D-feature on T, is that a NS in a 3rd person plural sentence is automatically interpreted as arbitrary in MR, unless there is a proper D-linker in the sentence, whereas in OR, those sentences were ambiguous between impersonal and referential.

As for the role of rich agreement, verbal inflection by itself was marginal in the change reviewed. The crucial trigger was a general reorganization of the tense – aspect system in the language, derived from the falling into disuse of certain past forms, and which led to the real trigger of the change: the loss of long verb movement. It just so happened that, in OR past forms, personal morphology was conveyed by the auxiliary, which happened to be a clitic element, lost in the language after the loss of V-to-T movement. The features overtly expressed on the

l-form (the former participle), i.e., gender and number, remained the same. Non-past forms were synthetic forms and included all the morphology, including person. Thus, these forms remained intact in the history of Russian; they just stopped undergoing V-to-T movement.

4.2 Structures and change in licensing NSs in BP

Unlike MR with respect to OR, BP displays severely deficient inflectional verbal paradigms, as compared to EP. In fact, it seems that most authors accept the hypothesis that the change in the pro-drop character of BP was triggered by the weakening of its verbal morphology (cf. i.a., BARBOSA; DUARTE; KATO, 2005; DUARTE, 1993, 2000; GALVES, 1993; NUNES, 2011). According to these authors, NSs in BP have dropped by 60% in the last 150 years, and the process is still ongoing, which is evidenced by the fact that younger speakers tend to drop referential subjects less often than older speakers (DUARTE, 1993; 2000).

It also seems, similarly to MR, that the referential NSs that survived in BP have been relegated to special contexts, in which their reference is recoverable thanks to certain special strategies (SILVA, 2000), that is, finite control (in embedded contexts), logophoric 1st-2nd person, and prominent topics for 3rd person.

As noted previously in Section 3.1 (Table 1), the final stage of the impoverishment of verbal inflection in BP renders a system with almost no person distinctions, except for the 1st person singular, while the old system displayed rich person agreement. According to Galves (1993), the simplification experienced by BP verbs concealed the loss of the semantic value of the person feature, while only a grammatical value of person was left in the language. Her proposal about the distribution of features in BP is as follows:

- (35)
- | | | |
|--------------|---------|----------|
| a. Ending –o | +person | –plural |
| b. Ending –ø | –person | –plural |
| c. Ending –m | –person | + plural |

The simplification of person inflection was parallel in finite and infinitive paradigms (cf. Table 1), and it was not due to phonological decay or erosion of the morphological endings, as happened in English or French, but to the reorganization of the pronominal system in the language, presumably, for contingent extra-linguistic / sociolinguistic

reasons. As happened in varieties of Spanish in the Americas, in BP, too, the 2nd person inflection collapsed with the 3rd person, after the polite forms *você / vocês* ‘you’ replaced the former familiar *tu / vós* ‘you’ for the 2nd person. In colloquial BP, additionally, the 1st person plural was conflated with the 3rd person singular when the corresponding personal pronoun *nós* ‘we’ was replaced with the generic *a gente* ‘the people, one’. Let us see the change step-by-step, following Duarte (1993):

- (1) First period of change (1910-1970): Duarte (1993) shows that the decreasing use of NSs in BP occurred at the same time as the simplification of person inflection: until approximately the 1920-30’s, BP displayed a pro-drop system, which was similar to that which EP has nowadays. The breaking point for the change was precisely the 1930’s, when two things happened: (i) the use of the second direct person (non-polite) decayed in favor of the polite forms; in fact, mixed treatments addressed to the same person are used in some passages in texts of this period, cf. (36); (ii) freely licensed referential NSs started to be replaced by overt pronouns in certain conditions. This process was particularly clear in the case of 2nd person NSs. On the other hand, 1st person singular and plural still show person agreement at this point (*canto / cantamos*), and 1st person still shows a high rate of NSs in this period.

(36)	Não	digas	tolices,	menino (...)	Não	seja	bobo,
	Not	say.2sg	nonsense	boy	not	be.3sg(=2pol)	silly
	menino!	Sabe		que dia	é	hoje?	
	boy	know.3sg(=2pol)	what	day	is	today	

‘Don’t say nonsense, boy (...) Don’t be silly, boy! Don’t you know which day is today?’ (BP – *A vida tem três andares* (1951), p. 74, ap. DUARTE, 1993, p. 114)

- (2) Second period of change (1970-): from the 1970s on, a further change in the pronominal system has become prominent, in which the 1st person plural *nós* is replaced in colloquial language by the generic *a gente*, and the decay of 1st person NSs becomes clear only at this point (DUARTE, 1993).

As for 3rd person NSs, their number decayed by just 25% in the whole period analyzed by Duarte, contrasting with subject drop in 1st

and 2nd person; NSs in these persons experienced a decrease of 80% and 70%, respectively, during the whole period. In addition, the loss of 3rd person NSs was progressive, without any sharp drop or visible breaking point, while 2nd person, for example, dropped from 69% to 25% only in the first period. Duarte (1993) attributes the availability of 3rd person NSs in BP to the presence of a prominent topic (“tema” in her words), which is able by itself to license the reference of such NSs, as in example (37).

- (37) – O que é que o nosso anjo_i tem hoje?
 what is that the our angel has today
 – pro_i Tá com essa cara desde que pro_i chegou do ginásio.
 is.3sg with that face since that arrived.3sgfrom gym
 ‘– What’s the matter with our angel today? – (He) has been with that face since he returned from the gym.’ (BP – *No coração do Brasil* (1992), ap. DUARTE, 1993, p. 118)

Galves (1993), citing Tarallo (1989), discovers a further correlation: null objects increase in the texts approximately in parallel with the loss of NSs. In 1880, the rate of overt pronominal subjects was 32.7%, while overt pronominal objects constituted 60.2%. A century later, in 1981, the figures were almost reversed: overt pronominal subjects represented 79.4%, and objects, 18.2%. Cyrino (1990), in turn, observes a decrease in the use of object clitics during the second half of the 20th century (from 81.6% to 47.3%), a general loss of 3rd person clitics, and the total replacement of clitic enclisis by clitic proclisis with respect to the verb; see examples (19b) and (19a) in Section 3.5.

The conditions of licensing null objects became more flexible with time; null objects in EP cannot appear within a syntactic island (38), which points to the fact that they can be just traces left by null operator movement (RAPOSO, 1986). Galves (1989), Cyrino (1993) and Kato (1993) claim that, in contrast, null objects in BP do not obey such grammatical restrictions (38), but correlate instead with the presence of a topic that allows their reference to be recovered (like in MR), which reveals that they are real null pronominal categories (*pro* or a null clitic), in our terms, a ϕP .⁷

⁷ There is an ongoing discussion about the nature of null arguments in the generative literature, whether they are *pro* or instances of ellipsis (cf. i.a., CYRINO; ORDÓÑEZ, 2018 specifically for BP).

- (38) Eu arquitei o livroi sem ler øi (*EP / OKBP)
 I filed the book without read.inf
 ‘I filed the book without reading it.’ (KATO, 1993, p. 225)

Finally, Nunes (1991) detects an increase in the lack of verbal agreement in *se*-sentences (from 62% in the 19th century to 84% in the 20th century). He shows that what really happened was the loss in colloquial BP of the passivizing value of *se* (39a), and the gain of an indeterminative value (indefinite or arbitrary reading), illustrated in (39b), which correlates with the lack of verbal agreement.

- (39) a. Alugam-se casas.
 rent.3pl-se houses.pl
 b. Aluga-se casas.
 rent.3sg-se houses.pl
 ‘Houses are rented.’ (BP – NUNES, 1991, p. 33)

As for the emergence of 3rd person singular constructions with an indefinite or arbitrary reading in BP, as in (14a-b) in Section 3.3, Nunes (1991) shows that, by the time in which *se*-constructions became indefinite (in the 19th century), the *se* clitic started to be suppressed (in just 9% of the cases in the texts), but the suppression of *se* in impersonal sentences has overwhelmingly increased in present-day colloquial BP, reaching nowadays 78% of the instances.

In Table 4, I summarize the change in the relevant pro-drop features in BP.

Table 4 – changes in the pro-drop system between early BP and present-day colloquial BP (approximate dates)

When	Changes					
1910-1970	Loss of 2 nd p. agreement	Decay of 2 nd p. NSs		Proclisis of 1 st -2 nd p. clitics replaces enclisis	Change in null objects	Loss of 3 rd person clitics
1970-	Loss of 1 st p. pl. agreement	Decay of 1 st p. NSs	Loss of <i>se</i> in arbitrary 3 rd sg sentences			

Now let us interpret all these data in formal terms, keeping in mind the assumptions about NSs and null objects adopted so far. In BP, as a result of the reorganization of the pronominal system (for extralinguistic reasons), the verbal agreement paradigm was severely impoverished and

learners received too little morphological evidence from finite verbs to be able to detect a D-feature on T anymore. As a consequence, the D-feature on T was lost, and pronominal subjects had to be overtly realized, unless speakers were able to relate the subject gap to a proper logophoric or very prominent topic feature endowing it with a proper reference. Thus far, the picture is very similar to Russian, except that the initial triggers of the change were different, namely, the change from a tense-based verbal system into an aspect-based system in OR *versus* the rearrangement of the pronominal system in BP. However, the crucial parametric setting (the loss of a D-feature on T) was common to MR and BP.

A further difference with respect to Russian involves infinitives and embedded NSs. According to Nunes (2011), the availability of inflected infinitives in BP and, most importantly, the parallel process of the impoverishment of person values in both finite and infinitive verbal forms led speakers to identify -partially inflected- finite forms with -partially inflected- infinitives. Learners of BP, unlike learners of European Portuguese, receive clear cues that infinitivals are ambiguous between being ϕ -complete or not (i.e., *cantar* specified only for singular number; *cantarem* specified only for plural). The little inflection present in finite and non-finite T signaled the presence of the same defective set of ϕ -features on both types of T and forced learners to reanalyze embedded referential NSs as controlled elements in finite embedded clauses too, generalizing from the already existing controlled NSs in infinitive clauses (PRO). In Russian, inflected infinitives never existed, and the important cue to identify embedded finite and non-finite subject gaps was the loss of the old alternation existing in OR between coreferent emphatic overt (dative) pronouns and coreferent non-emphatic NSs in infinitive clauses, paralleling the situation of finite clauses, as explained in Section 4.1. All in all, BP speakers “saved” referential embedded subject gaps in the same way as learners of Middle Russian did -albeit for different reasons-, by reanalyzing those gaps and identifying them with the other embedded referential subject gaps that were preserved in the language, i.e., coreferent infinitive subjects (PRO).

Further, Galves (1994) shows that BP lacks long verb movement, contrasting with EP, which does display long verb movement (cf. i.a., MODESTO, 2009, p. 103; PIRES, 2005; TESCARI NETO, 2012). Therefore, long verb movement was lost at some point in the history of BP, and verbs started to undergo short movement (to Asp or a similar

intermediate position, as in Russian), as a side-effect of the loss of rich person agreement. It seems that the change in pro-drop in BP and in Russian represented two sides of the same coin: in Russian, the loss of V-to-T movement modified the cue that learners needed to posit an unvalued D-feature on T, the overt morphological realization of the operation of D-feature valuing, i.e., person agreement overtly realized in a high position (at T). Thus, T was considered not to take part in D-feature valuing any more, and the D-feature of T was lost. In BP, on the other hand, the overt morphological realization of D-feature valuing, person agreement, was lost altogether, and this automatically led to the loss of the D-feature on T. At the same time, the loss of person agreement had the side-effect that V-to-T movement could be perceived by BP learners as unmotivated and decayed as a result.

Clitic objects did not change in BP as much as in Russian, in which clitics were totally replaced by strong pronouns or newly created null objects. In BP, 1st and 2nd person clitics were preserved, but the low position of the verb brought with it a change in the cliticization direction of clitics. Duarte; Matos; Gonçalves (2005) claim that variation in clitic order, and its different patterns in EP and BP depend on the ability of T and Asp to attract V *versus* T's capacity to check uninterpretable features through Agree without attracting V. In EP, clitics are always enclitic, whether the preceding word is a verb or not (NUNES, 2011); when the verb is attracted to T, the clitic “remains” to its right, as in (19b) above, and in case there is some element at C (a wh-word, negation, etc.), leftward cliticization takes place with respect to the C-element (21b). On the other hand, in BP the verb remains low and the clitic is spelled-out “to its left”; thus, learners interpret this pattern as proclitic cliticization. In this sense, the change from enclisis to proclisis in the 1st and 2nd person clitics in BP could be a further consequence of the change in the position of V after the loss of rich person agreement.

With regard to the loss of 3rd person clitics, Nunes (1993) argues that it was due to the change of directionality in cliticization in BP. The new system of proclisis made impossible for learners to acquire 3rd person clitics, whose syllabic onset was incompatible with rightward cliticization.⁸ The loss of 3rd person clitics led to the emergence of weak

⁸ Nunes (1993) notes that 3rd person clitics in Portuguese (*o, a, os, as*) lost the onset of their syllable (the initial *l-*, present in other Romance languages, like the Spanish *lo, la, los, las*). He

pronouns in object position, as well as the expansion of the already existing null object constructions.

As was the case in Russian, the new ways of licensing referential NSs in pragmatically motivated specific situations in BP paved the way to reanalyze the old null objects, traces of *wh*-movement in EP (RAPOSO, 1986), as real null pronouns (GALVES, 1989, i.a.) by analogy with the new NSs. Thus, null objects in BP became more unrestricted from the grammatical point of view than in EP, as they only needed to be pragmatically motivated. In that sense, the final result of the change in MR and BP, being initially quite different, resulted in a similar output.

Finally, arbitrary sentences in BP retained the old patterns (3rd person plural agreement and the *se*-construction), and created a new pattern with 3rd person singular agreement, which arose after the loss of the *se* clitic element (NUNES, 1991), probably in relation to the loss of 3rd person clitics in general in the language. This process differs from Russian, because in OR, unlike EP, impersonal sentences were never formed with a *sja* clitic (equivalent to Romance *se* clitic), and thus the corresponding arbitrary construction in MR did not have a suitable ancestor with 3rd person singular agreement, on the basis of which learners could build the construction that arose in BP.

5 Conclusion

In this paper, I have shown that the contingent nature of change determines the fact that one and the same parametric change can be biased by slightly different historical conditions, which can render slightly different results. However, the final parametric setting will be very similar, as proven by the detailed examination of the change in pro-drop, experienced by MR and BP independently from each other.

These two languages developed from a CNS stage into a PNS stage for different reasons and according to different pathways of change, but ended up having basically the same properties. From a parametric

argues that the onset of the syllable of 3rd person clitics is licensed via phonological cliticization to the preceding word as happens in EP, in which clitics are always enclitic to whatever word they need to be. However, in BP cliticization is always rightward, which prevented the onset of the syllable of 3rd person clitics to be phonologically licensed, and led to their disappearance. 1st and 2nd person clitics (*me*, *te*, *nos*, *vos*) did not lose their syllabic onset, so they were preserved in BP, albeit in the new position.

point of view, both languages lost the D-feature on T, associated with a series of common properties (short verb movement, reanalysis of NSs as pragmatically licensed, automatic interpretation of a NS as arbitrary or generic, finite control, expansion of null objects, weakening of clitics, reanalysis of embedded NSs), even if the exact pathways of change differed between the two languages. The small differences in the final results of the change have been attributed to different initial conditions in the CNS stage of pro-drop; namely, the existence or not of impersonal *se* constructions, the role of the impoverishment of agreement, the differences in non-finite embedded clauses, and the initial trigger of the change (reorganization of the tense–aspect system vs. rearrangement of the pronominal system).

Acknowledgements

This research has been supported by the PGC2018-096870-B-100 and PGC2018-098995-B-I00 projects, funded by the Spanish Ministry of Science, Innovation, and Universities, the Spanish Research Agency, and the European Regional Development Fund (FEDER), and the IT1344-19 research group, funded by the Basque Government. The author is especially grateful to Maia Duguine for advice and help with this paper, and to two anonymous reviewers at the *Revista de Estudos da Linguagem* for their comments and suggestions. I also want to thank Xavier Cruz and Pavel Grashchenkov for help with European Portuguese and Russian data, respectively.

References

- BARBOSA, P.; DUARTE, M. E. L.; KATO, M. A. Null Subjects in European and Brazilian Portuguese. *Journal of Portuguese Linguistics*, Lisboa v. 4, n. 2, p. 11-52, 2005. DOI:10.5334/jpl.158
- BIBERAUER, T. Introduction. In: BIBERAUER, T. (ed.). *The limits of syntactic variation*. John Benjamins, 2008. p. 1-72.
- BIBERAUER, T.; ROBERTS, I.; HOLMBERG, A.; SHEEHAN, M. Introduction: parameters in minimalist theory. In: BIBERAUER, T.; ROBERTS, I.; HOLMBERG, A.; SHEEHAN, M. (eds.). *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge: Cambridge University Press, 2010. p. 1-57.

BLEAM, T. *Leista Spanish and the syntax of clitic doubling*. PhD Dissertation, University of Delaware, 1999.

BOECKX, C.; HORNSTEIN, N.; NUNES, J. *Control as Movement*. Cambridge: Cambridge University Press, 2010.

BORKOVSKIJ, V. I. *Istoričeskaja grammatika ruskogo jazyka: sintaksis - prostoe predloženie*. Moskva: Nauka, 1978.

BORKOVSKIJ, V. I.; KUZNECOV, P. S. *Istoričeskaja grammatika ruskogo jazyka*. Moskva: AN SSSR, 1965.

BOŠKOVIĆ, Ž. On second position clitics crosslinguistically. In: MARUŠIĆ, F.; ŽAUCER, R. (eds.). *Formal studies in Slovenian syntax: In honor of Janez Orešnik*. Amsterdam - Philadelphia: John Benjamins, 2016. p. 23-53.

CAMACHO, J.; ELÍAS-ULLOA, J. Null subject systems in Shipibo switch-reference. In: CAMACHO, J.; GUTIERREZ-BRAVO, R.; SANCHEZ, L. (eds.). *Information structure in languages of the Americas*. Berlin: Mouton de Gruyter, 2010. p. 65-85.

CARDINALETTI, A.; STARKE, M. The typology of structural deficiency: A case study of the three classes of pronouns. In: VAN RIEMSDIJK, H. (ed.). *Clitics in the languages of Europe*. Berlin: Mouton de Gruyter, 1999. p. 185-234.

CORRÊA, V. *O Objeto Direto Nulo no Português do Brasil*. MA thesis, Universidade Estadual de Campinas, 1991.

CYRINO, S. M. L. Observações sobre a mudança diacrônica no português do Brasil: Objeto nulo e clíticos. In: ROBERTS, I; KATO, M. A. (eds.). *Português Brasileiro: Uma Viagem Diacrônica*, Campinas: Editora da UNICAMP, 1993. p. 163-184.

CYRINO, S.; ORDOÑEZ, F. Null objects in Brazilian Portuguese and DOM in Spanish: similarities and differences. In: PARODI, T. (ed.). *Proceedings of the VIII Nereus International Workshop: Referential Properties of the Romance DP in the Context of Multilingualism*. Kontanz: Fachbereich Sprachwissenschaft, Universität Konstanz, 2018. p. 103-120.

DUARTE, M. E. L. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In: ROBERTS, I; KATO, M. A. (eds.). *Português Brasileiro: Uma Viagem Diacrônica*, Campinas: Editora da UNICAMP, 1993. p. 107-128.

DUARTE, M. E. L. The loss of the 'avoid pronoun' principle in Brazilian Portuguese. In: KATO, M. A.; NEGRÃO, E. V. (eds.). *Brazilian*

Portuguese and the Null Subject Parameter, Madrid & Frankfurt am Main: Iberoamericana & Vervuert, 2000. p. 17-36.

DUARTE, M. E. L.; SILVA, M. C. F. The Null Subject Parameter and the Structure of the Sentence in European and Brazilian Portuguese. In: WETZELS, W. L.; COSTA, J.; MENUZZI, S. (eds.). *The Handbook of Portuguese Linguistics*, Malden, MA: Wiley-Blackwell, 2016. p. 234-253.

DUARTE, I.; GONÇALVES, A.; SANTOS, A. L. Infinitivo flexionado, independência temporal e controlo. In: COSTA, A.; FLORES, C.; ALEXANDRE, N. (eds.). *Textos Seleccionados do XXVII Encontro Nacional da Associação Portuguesa de Linguística*, Lisboa: APL, 2012. p. 217-234.

DUARTE, I.; MATOS, G.; GONÇALVES, A. Pronominal Clitics in European and Brazilian Portuguese. *Journal of Portuguese Linguistics* Lisboa, v. 4, p. 113-141, 2005. DOI: 10.5334/jpl.161

FRASCARELLI, M. Subjects, topics and the interpretation of referential pro. *Natural Language and Linguistic Theory*, Dordrecht, v. 25, p. 691-734, 2007. DOI: 10.1007/s11049-007-9025-x

GALVES, C. O objeto nulo no português brasileiro: Percurso de uma pesquisa. *Cadernos de Estudos Linguísticos*, Campinas, v. 17, p. 65-90, 1989. DOI: XXXXX

GALVES, C. O enfraquecimento da concordância no português brasileiro. In: ROBERTS, I.; KATO, M. A. (eds.). *Português Brasileiro: Uma Viagem Diacrônica*. Campinas: Editora da UNICAMP, 1993. p. 387-403.

GALVES, C. V-movement, levels of representation and the Structure of S. *Letras de Hoje*, Porto Alegre, v. 29, n. 2, p. 35-58, 1994. DOI: XXXXX

GRIBANOVA, V. Verb-stranding verb phrase ellipsis and the structure of the Russian verbal complex. *Natural Language and Linguistic Theory*, Dordrecht, v. 31, n. 1, p. 91-136, 2013. DOI: 10.1007/s11049-012-9183-3

HOLMBERG, A. Is There a Little Pro? Evidence from Finnish. *Linguistic Inquiry*, Cambridge, MA, v. 36, p. 533-564, 2005. DOI: 10.1162/0024389057744464322

HOLMBERG, A. Null Subject Parameters. In: BIBERAUER, T; HOLMBERG, A.; ROBERTS, I.; SHEEHAN, M. (eds.). *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge: Cambridge University Press, 2010. p. 88-124.

HOLMBERG, A.; NAYUDU, A.; SHEEHAN, M. Three partial null-subject languages: a comparison of Brazilian Portuguese, Finnish and Marathi. *Studia Linguistica*, Hoboken, New Jersey, v. 63, p. 59-97, 2009. DOI: 10.1111/j.1467-9582.2008.01154.x

HOLMBERG, A.; ROBERTS, I. Parameters and the three factors of language design. In: PICALLO, C. (ed.). *Linguistic Variation in the Minimalist Framework*. Oxford: Oxford University Press, 2014. p. 61-81.

IVANOV, V. V. *Istoričeskaja grammatika ruskogo jazyka*. Moskva: Prosveščenie, 1990.

JUNG, H. Null subjects and person in Old North Russian. In: GRKOVIC-MAJOR, J.; HANSEN, B.; SONNENHAUSER, B. (eds.). *Diachronic Slavonic syntax*. Berlin: De Gruyter, 2018. p. 95-123.

JUNG, H.; MIGDALSKI, K. On the degrammaticalization of pronominal clitics in Slavic. In: SZAJBEL-KECK, M.; BURNS, R.; KAVITSKAYA, D. (eds.). *Formal approaches to Slavic linguistics: The first Berkeley meeting 2014*. Ann Arbor: Michigan Slavic Publications, 2015. p. 143-162.

KALLESTINOVA, E.; SLABAKOVA, R. Does the verb move in Russian?. In: ANTONENKO, A.; BAILYN, J.; BETHIN, C. (eds.). *Formal approaches to Slavic linguistics 16: The Stony Brook meeting 2007*. Ann Arbor: Michigan Slavic Publications, 2008. p. 199-214.

KATO, M. A. The distribution of pronouns and null elements in object position in Brazilian Portuguese. In: ASHBY, W.; PERISSINOTTO, M. M. G.; RAPOSO, E. P. (eds.). *Linguistic Perspectives on the Romance Languages*. Amsterdam/Philadelphia: John Benjamins., 1993. p. 225-235.

KIBRIK, A. A. Peculiarities and origins of the Russian referential system. In: BAKKER, D.; HASPELMATH, M. (ed.). *Languages across boundaries: studies in memory of Anna Siewierska*. Berlin : De Gruyter, 2013. p. 227-262.

LANDAU, I. The scale of finiteness and the calculus of control. *Natural Language and Linguistic Theory*, Dordrecht, v. 22, p. 811-877, 2004. DOI: 10.1007/s11049-004-4265-5

LIGHTFOOT, D. *The Development of Language: Acquisition, change and evolution*. Oxford: Blackwell, 1999.

LIVITZ, I. *Deriving silence through internal reference: focus on pronouns*. Ph.D. Dissertation, New York University, 2014.

MADARIAGA, N. Diachronic change and the nature of pronominal null subjects: the case of Russian. In: CASALICCHIO, J.; COGNOLA, F. (eds.). *Null Subjects in Generative Grammar: A synchronic and diachronic perspective*. Oxford: Oxford University Press, 2018. p. 171-98.

MADARIAGA, N. Referential null subjects in Russian: a synchronic and diachronic overview. In: DALMI, G.; TSEDRYK, E.; CEGŁOWSKI, P.

(eds.). *Null subjects in Slavic and Finno-Ugric: Licensing, structure, and typology*. Berlin: Mouton de Gruyter, 2022a. p. 75-103.

MADARIAGA, N. Deficient pronouns in the history of Russian: null subjects and object clitics. Ms. UPV/EHU, 2002b.

McSHANE, M. *A Theory of Ellipsis*. Oxford: Oxford University Press, 2005.

MEYER, R. *The history of Null Subjects in East Slavonic. A corpus based diachronic investigation*. Habilitation Thesis, University of Regensburg, 2011.

MIGDALSKI, K. Diachronic source of two cliticization patterns in Slavic. In: SALVESEN, C. M.; HELLAND, H. P. (eds.). *Challenging Clitics*. Amsterdam: John Benjamins, 2013. p. 135-158.

MODESTO, M. Null Subjects without “rich” Agreement. In: KATO, M. A.; NEGRÃO, E. V. (eds.). *Brazilian Portuguese and the Null Subject Parameter*, Madrid & Frankfurt am Main: Iberoamericana & Vervuert, 2000. p. 147-174.

MODESTO, M. Null subjects in Brazilian Portuguese: a critique of two possible analyses. In: TORRES MORAIS, M. A.; ANDRADE, M. L. (eds.). *História do Português Caipira*. Campinas: editora da UNICAMP, 2009. p. 99-123.

MODESTO, M. Inflected infinitives in Brazilian Portuguese and the theory of Control. In: SANTOS, A. L.; GONÇALVES, A. (eds.) *Complement Clauses in Portuguese: Syntax and acquisition: Issues in Hispanic and Lusophone Linguistics*, vol. 17. Amsterdam: John Benjamins, 2018. p. 59-100.

MÜLLER, G. Pro-drop and impoverishment. In: BRANDT, P.; FUß, E. (eds.). *Form, Structure, and Grammar*. A Festschrift presented to Günther Grewendorf. Berlin: Akademie Verlag, 2006. p. 93-115.

NEWMAYER, F. Against a Parameter-setting approach to language variation. *Linguistic Variation Yearbook*, Amsterdam, v. 4, p. 181-234, 2005. DOI: 10.1075/livy.4.06new

NUNES, J. *Se apassivador e se indeterminador: o percurso diacrônico no português brasileiro*. *Cadernos de Estudos Lingüísticos*, Campinas, v. 20, p. 33-57, 1991.

NUNES, J. Direção de cliticização, objeto nulo e pronome tônico na posição de objeto em português brasileiro. In: ROBERTS, I; KATO, M. A. (eds.). *Português Brasileiro: Uma Viagem Diacrônica*. Campinas: Editora da UNICAMP, 1993. p. 207-222.

NUNES, J. A Note on Wh-islands and Finite Control in Brazilian Portuguese. *Estudos da Língua*, Vitória da Conquista, v. 8, n. 2, p. 79-103, 2010. DOI: 10.22481/el.v8i2.1129

NUNES, J. On the diachronic reanalysis of null subjects and null objects in Brazilian Portuguese: triggers and consequences. In: RINKÉ, E.; KUPISCH, T. (eds). *The Development of Grammar: Language acquisition and diachronic change*. John Benjamins, 2011. p. 331-354.

ORMAZABAL, J.; ROMERO, J. Object clitics, agreement and dialectal variation. *International journal of Latin and Romance linguistics*, Berlin, v. 25, n. 2, p. 301-344, 2013. DOI: 10.1515/probus-2013-0012

PIRES, A. Verb movement and clitics: Variation and change in Portuguese. In: BATLLORI, M.; HERNANZ, M-L; PICALLÓ, C; ROCA, F. (eds.) *Grammaticalization and Parametric Variation*. Oxford: Oxford University Press, 2005. p. 48-59.

PROGOVAC, L. Eventive *to* and the placement of clitics in Serbo-Croatian. In: KENESEI, I. (ed.). *Crossing boundaries: Advances in the theory of Central and Eastern European languages*. Amsterdam: John Benjamins, 1999. p. 33-44.

RAPOSO, E. P. On the null object in European Portuguese. In: JAEGGLI, O.; SILVA-CORVALAN, C. (eds.). *Studies in Romance Linguistics*. Dordrecht: Foris, 1986. p. 373-390.

RAPOSO, E. P. Case Theory and Infl-to-Comp: The Inflected Infinitive in European Portuguese. *Linguistic Inquiry*, Amsterdam, v. 18, p. 85-109, 1987.

REINHART, T. Pragmatics and linguistics: an analysis of sentence topics. *Philosophica*, Valparaíso, v. 27, n. 1, p. 53-94, 1981. DOI: 10.21825/philosophica.82606

RIZZI, L. *Issues in Italian Syntax*. Dordrecht: Foris, 1982.

ROBERTS, I. A deletion analysis of null subjects: French as a case study. Ms, University of Cambridge, 2010.

RODRIGUES, C. *Impoverished morphology and A-movement out of case domains*. PhD Dissertation, University of Maryland, 2004.

SHEEHAN, M. On the difference between exhaustive and partial control. In: CASALICCHIO, J.; COGNOLA, F. (eds.). *Null Subjects in Generative Grammar: A synchronic and diachronic perspective*. Oxford: Oxford University Press, 2018. p. 141-170.

SIGURÐSSON, H. A. Conditions on argument pro-drop. *Linguistic Inquiry*, Amsterdam, v. 42, n. 2, p. 267-304, 2011. DOI: 10.1162/ling_a_00042

SILVA, M. C. F. Main and embedded null subjects in Brazilian Portuguese. In: KATO, M. A.; NEGRÃO, E. V. (eds.). *Brazilian Portuguese and the Null Subject Parameter*. Madrid & Frankfurt am Main: Iberoamericana & Vervuert, 2000. p. 127-146.

TARALLO, F. *Fotografias Sociolingüísticas*. Campinas: Pontes, 1989.

TESCARI NETO, A. *On Verb Movement in Brazilian Portuguese: A Cartographic Study*. PhD Dissertation, Università Ca' Foscari di Venezia, 2012.

TSEDRYK, E. Deriving null pronouns: A unified analysis of subject drop in Russian. In: SZAJBEL-KECK, M.; BURNS, R.; KAVITSKAYA, D. (eds.). *Formal approaches to Slavic linguistics: The first Berkeley meeting 2014*. Ann Arbor: Michigan Slavic Publications, 2015. p. 342-361.

TSEDRYK, E. Null subjects and their overt counterparts in East Slavic root clauses: referential and non-referential readings. In: DALMI, G.; TSEDRYK, E.; CEGŁOWSKI, P. (eds.). *Null subjects in Slavic and Finno-Ugric: Licensing, structure, and typology*. Berlin: Mouton de Gruyter, 2022. p. 35-74.

URIAGEREKA, J. Aspects of the syntax of clitic placement in Western Romance. *Linguistic inquiry*, Amsterdam, v. 26, n. 1, p. 79-123, 1995.

VAINIKKA, A.; LEVY, Y. Empty subjects in Finnish and Hebrew. *Natural Language and Linguistic Theory*, Dordrecht, v. 17, p. 613-671, 1999. DOI: 10.1023/A:1006225032592

WALKDEN, G. Refining the “null argument cycle”: the place of partial null argument languages. In: 14TH INTERNATIONAL DIACHRONIC GENERATIVE SYNTAX CONFERENCE, Lisbon, 2012. Available at: <http://walkden.space/Walkden_2012_Lisbon.pdf>. Last accessed: 23 March 2022.

ZALIZNJAK, A. A. *Drevnerusskie ènklitiki*. Moskva: Jazyki slavjanskoj kul'tury, 2008.

Semântica: Tempo e aspecto, léxico



**A aquisição de Tempo e Aspecto: uma investigação
sobre a emergência de estruturas verbais e a influência do
aspecto semântico sobre o aspecto gramatical no processo de
aquisição da linguagem**

***The Tense and Aspect acquisition: an investigation
on the emergence of verbal structures and the influence of the
semantic aspect on the grammatical aspect in the language
acquisition process***

Arabie Bezri Hermont

Pontifícia Universidade Católica de Minas Gerais (PUC-MG), Belo Horizonte, Minas Gerais / Brasil

arabie@uol.com.br

<http://orcid.org/0000-0003-2551-6145>

Kelly Cesário de Oliveira

Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Minas Gerais / Brasil

kellycesario@gmail.com

<http://orcid.org/0000-0002-6262-5976>

Lucas Segantini Brito

Pontifícia Universidade Católica de Minas Gerais (PUC-MG), Belo Horizonte, Minas Gerais / Brasil

lucassegantini2812@gmail.com

<http://orcid.org/0000-0002-2860-9323>

Resumo: Este artigo apresenta resultados de uma pesquisa que objetivou verificar a aquisição de formas verbais relacionadas a tempo e a aspecto, demonstrando resultados de uma investigação realizada após o estudo do surgimento de estruturas verbais (na forma simples e na forma perifrástica – auxiliar e verbo principal) efetivamente produzidas por cinco crianças monolíngues, entre 2 anos de idade e 5 anos e 11 meses de idade, no processo de aquisição de linguagem. O *corpus* utilizado foi catalogado

por dois integrantes do grupo eLinC - Estudos em Linguagem e Cognição (a sair). Objetivou-se entender, além disso, em que medida as noções de aspecto semântico/lexical influenciam ou não no surgimento das marcações aspectuais gramaticais. Este trabalho tem duas hipóteses: (i) na gramática infantil, a noção de aspecto surge, de forma mais recorrente, antes da noção de tempo e (ii) quando as formas verbais no presente e no pretérito perfeito surgem, o que está expresso é o aspecto. A partir de análises qualitativas, entendemos que o aspecto é uma categoria que surge antes de tempo, mas, mesmo assim, crianças mais jovens apresentam a categoria tempo em suas produções espontâneas, o que vai ao encontro do que preconiza a Teoria Gerativa, quadro teórico adotado para a realização desta pesquisa.

Palavras-chave: aquisição de linguagem; aspecto; tempo; psicolinguística; teoria gerativa.

Abstract: This article presents the results of a research that aimed to verify the acquisition of verbal forms related to tense and aspect, demonstrating the results of an investigation carried out after the study of the emergence of verbal structures (in simple form and in periphrastic form – auxiliary and main verb) effectively produced by five monolingual children, between 2 years old and 5 years and 11 months old, in the process of language acquisition. The *corpus* used was cataloged by two members of the eLinC group – Studies in Language and Cognition (in press). The objective was to understand, in addition, to what extent the notions of semantic/lexical aspect influence or not on the appearance of grammatical aspect marks. This work has two hypotheses: (i) in children's grammar, the notion of aspect appears, more recurrently, before the notion of tense and (ii) when the verbal forms in the present tense and in the perfect past tense appear, the aspect is expressed. Based on qualitative analyses, we understand that the aspect is a category that appears before tense, but, even so, younger children present the tense category in their spontaneous productions, which is in line with what the Generative Theory advocates, which is the theoretical framework adopted to carry out this research.

Keywords: language acquisition; aspect; tense; psycholinguistics; generative grammar.

Recebido em 28 de janeiro de 2022

Aceito em 09 de março de 2022

1 Considerações iniciais

O estudo da aquisição da linguagem teve seu início a partir de registros em diários de dados da fala da criança e foi seguido por muitas pesquisas à luz do behaviorismo. Foi somente depois dos anos 50 do século XX que vimos estudos de aquisição sob a ótica da teoria gerativa,

que concebe uma gramática mental, o que faz com que as pesquisas possam ser realizadas a partir de dados coletados de fala espontânea de crianças em intervalos regulares. Nesse raciocínio, este trabalho versa sobre a aquisição das categorias gramaticais tempo e aspecto em um viés Psicolinguista, uma vez que investiga fenômenos linguísticos na aquisição de linguagem, com motivação cognitiva, pois entender como as categorias tempo e aspecto emergem na gramática infantil é inserir-se em uma perspectiva linguística e psicológica.

O trabalho da criança ao adquirir a sintaxe de uma língua estaria ligado à aquisição de um léxico, constituído de palavras de classe aberta (por exemplo, nomes, adjetivos, verbos) e de classe fechada (por exemplo, as categorias funcionais, como determinantes, complementizadores e as categorias de tempo e aspecto). Para alguns estudiosos, os elementos funcionais, de um modo geral, estariam ausentes da fala da criança. Entretanto, vários autores demonstram que as categorias funcionais já estariam presentes na produção e na compreensão da linguagem da criança desde muito cedo, como Wexler (1996, 1998).

Nos últimos anos, houve um crescimento expressivo de pesquisas nas áreas que investigam a produção e a compreensão das categorias funcionais na gramática de crianças em fase normal de aquisição da linguagem e duas categorias que têm recebido atenção são a de tempo e a de aspecto. Tempo, nessa perspectiva, situaria o momento de ocorrência de uma situação: no passado, no presente ou no futuro. Aspecto seria uma noção ligada a diferentes formas de verificar a constituição temporal interna da situação, como, por exemplo, a sua duração ou sua conclusão ou não.

Este trabalho adotará os pressupostos da Teoria Gerativa, que assume que o conhecimento linguístico pode ser explicado levando-se em conta a existência de um conhecimento inato, codificado biologicamente, que é a Gramática Universal. Há muito tempo e com diversas reflexões, os estudos de Princípios e Parâmetros – P&P – (CHOMSKY, 1981, 1995) vêm promovendo a investigação das categorias funcionais e, diante disso, declaramos que o objetivo geral deste artigo é entender como as categorias tempo e aspecto estão representadas nas gramáticas de crianças em processo de aquisição de linguagem. Assim sendo, esta pesquisa, apoiada na Hipótese da Primazia do Aspecto (ANDERSEN; SHIRAI, 1996), objetivou verificar a aquisição de formas verbais relacionadas a tempo e a aspecto e demonstrar resultados de uma investigação realizada após o estudo do surgimento de estruturas verbais (na forma simples e na

forma perifrástica – auxiliar e verbo principal) efetivamente produzidas por cinco crianças monolíngues no processo de aquisição de linguagem. Além disso, havia um objetivo mais específico que é entender em que medida as noções de aspecto semântico/ lexical influenciam ou não no surgimento das marcações aspectuais gramaticais.

Duas hipóteses são estabelecidas: (i) na gramática infantil, a noção de aspecto surge, de forma mais recorrente, antes da noção de tempo e (ii) quando as formas verbais no presente e no pretérito perfeito surgem, o que está expresso é o aspecto.

Esta pesquisa foca-se em análise *off-line*, já que trabalha com dados já transcritos com dados de falas espontâneas de criança sem medição de tempo na produção. O *corpus* utilizado foi catalogado por dois integrantes do grupo eLinC – Estudos em Linguagem e Cognição¹ –, e constitui-se de gravações, realizadas entre 2016 e 2017, de falas espontâneas de crianças no período de aquisição de linguagem, que moram na cidade de Belo Horizonte e região metropolitana. Assim, selecionaram-se dados de algumas fases de aquisição de linguagem de cinco crianças com idade entre dois a cinco anos. A fim de obter um quadro mais completo da expressão de tempo e aspecto na fala dessas crianças, foram selecionados todos os verbos e perífrases verbais (auxiliar e verbo principal) e foi verificado como as categorias gramaticais de tempo e de aspecto se apresentam na produção linguística de cada criança. Foram realizadas, por fim, análises qualitativas.

Este artigo está organizado da seguinte forma: inicialmente, apresentaremos uma breve abordagem de principais pressupostos da Teoria Gerativa. Na seção seguinte, demonstraremos distintos conceitos de tempo e de aspecto, objeto de estudo deste trabalho. Na seção 3, trazemos um aporte teórico sobre aquisição de linguagem, detendo-nos em Wexler (1998), inspirador para a explicação dos dados obtidos em nossa pesquisa e em algumas pesquisas que tratam da aquisição de aspecto. Nas seções seguintes, traremos a metodologia e a descrição e análise de dados. Por fim, traçamos as considerações finais.

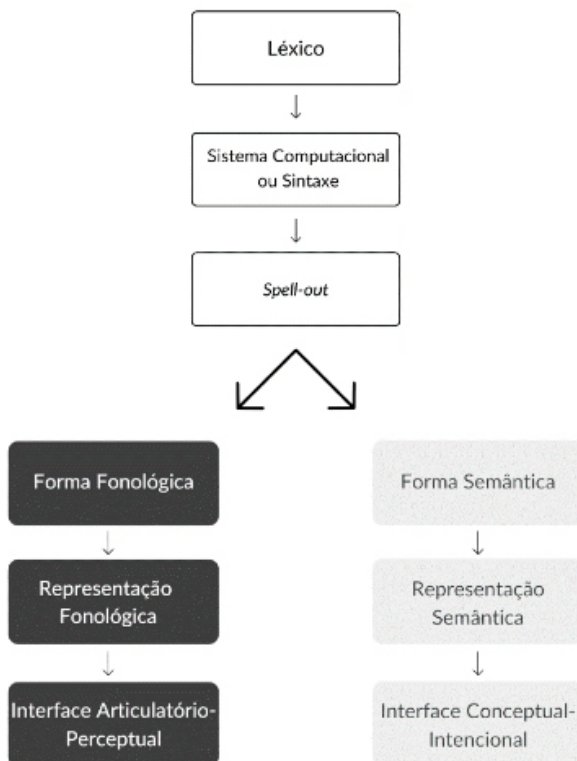
2 Teoria gerativa

Pesquisas sobre aquisição e desenvolvimento da linguagem à luz do modelo de Princípios & Parâmetros assumem a existência de uma faculdade da linguagem inconsciente e inata, também conhecida como

¹ Leôncio e Miranda, no prelo.

Gramática Universal (de agora em diante, GU). A Teoria Gerativa, no modelo do Programa Minimalista (CHOMSKY, 1995, 1998, 1999, 2001), considera a linguagem um sistema perfeito com um *design* ótimo. Assim sendo, as gramáticas das línguas naturais criariam estruturas a serem enviadas às interfaces, de acordo com a figura 1.

Figura 1 – Esquema da arquitetura de linguagem segundo o Modelo Minimalista



Fonte: elaborada pelos autores.

De acordo com Hermont e Morato (2014, p. 215),

As palavras estariam em um componente chamado léxico (com traços fonológicos, semânticos e formais) e seriam selecionadas para formar as sentenças. Essa seleção recebe o nome *select* e forma uma numeração. A numeração seria o conjunto de itens

lexicais selecionados. A proposta é que, na maioria das vezes, os itens já viriam da numeração com traços definidos, como traços de gênero, de número, de categoria, de caso etc. Após a seleção de palavras no léxico, ocorreria, no sistema computacional, a formação das sentenças. Para que a sentença formada no sistema computacional seja expressa e compreendida, uma operação chamada *spell-out* entraria em ação. É ela que permitiria a divisão da estrutura sintática para as interfaces articulatório-perceptual e conceptual-intencional.

A sentença formada passaria, então, pela forma fonológica, ganhando representação para ser enviada à interface articulatório-perceptual e haveria articulação dos sons da fala. Também há a forma semântica, diretamente ligada ao mundo das intenções e significados. Seria a parte em que a mente compreenderia a sentença e daria *input* para interface conceptual-intencional (HERMONT; MORATO, 2014, p. 215).

Também ancorados em Hermont e Morato, (2014, p. 215-216), passemos, agora, a explicitar sobre as categorias funcionais.

No modelo da Teoria Padrão (CHOMSKY, 1965) e da Teoria Padrão Estendida (CHOMSKY, 1972), considerava-se que a estrutura de uma sentença simples seria constituída por um SN e um SV, ambos ligados a um nó, denominado S. A formalização seria a seguinte: $S \rightarrow SN SV$. Mas, a fim de acomodar dados em que havia a presença de um verbo auxiliar, a proposta para a descrição da sentença passou a ser a seguinte: $S \rightarrow SN Aux SV$. Em Aux, haveria informação de tempo e de concordância.

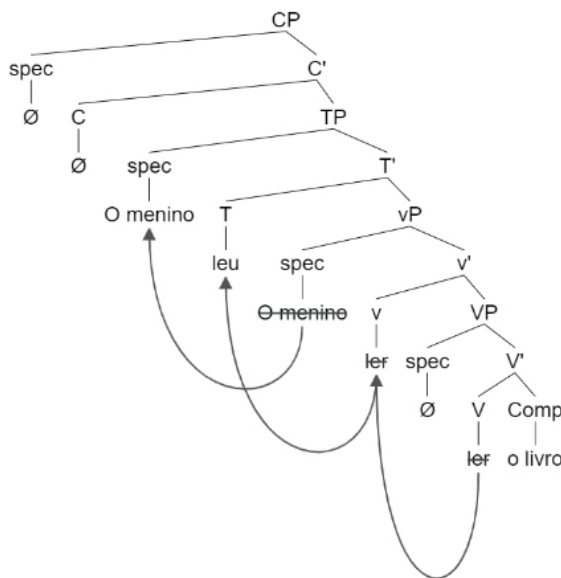
Joseph Emonds (1976) propôs a denominação de INFL para Aux e indicou uma marcação binária para o nó, tal como se faz com as categorias lexicais. Então, uma sentença finita teria INFL [+T, +AGR], e uma sentença no infinitivo teria INFL [-T, -AGR]. A regra sintagmática passou a ser $S \rightarrow SN INFL SV$ e era válida para todos os tipos de sentenças.

Com o advento da Teoria da Regência e Ligação (CHOMSKY, 1981), a proposta era de que INFL fosse o núcleo da sentença. Sendo assim, teríamos $IP_3 \rightarrow I SV$, em que a ideia geral seria a de que tanto verbos auxiliares (como *should*, *can*, etc.) finitos quanto o *to*, característico de formas não finitas em inglês, ocupassem o núcleo de IP. Além disso, o SN sujeito ocuparia a posição de especificador de I.

Pollock (1989), após realização de um estudo de comparação de movimento de verbo (em relação a advérbios, quantificadores e partículas de negação) nas formas finitas e não finitas em inglês e em francês, propôs que deveria haver um nó de Concordância (AGR) e um de Tempo (convencionado I – de *Inflection*) (HERMONT; MORATO, 2014, p. 215 -16).

Em 1995, Chomsky sinaliza que AGR já não seria mais um nóculo constituindo um sintagma, pois, diferentemente de T², AGR não teria interpretabilidade semântica na Forma Semântica e, portanto, há sugestões de que tal categoria não deveria ser o núcleo de uma projeção. Resumindo, o que era INFLP passou a ser IP e, depois, a TP, onde Tempo é marcado. Em TP, haveria também os traços ϕ , isto é, pessoa, número e gênero do DP³ sujeito; além disso, no especificador de TP, há um traço de EPP, que atrai o DP sujeito. Na figura 2, apresentamos um exemplo da representação de uma dada sentença.

Figura 2 – Representação arbórea



Fonte: elaborada pelos autores via jsSyntaxTree.

Em Hermont (2005) e Hermont e Morato (2014), a ideia é que a camada flexional seja dividida em Tempo e Aspecto. Outros trabalhos, como de Cinque (1999), sustentam que a camada flexional seja cindida

² AGR, neste trabalho, advém de *Agreement*, que significa “concordância”. T, por sua vez, advém de *Tense*, em português, tempo.

³ DP-sujeito é a sigla de sintagma determinante, uma terminologia usada atualmente para designar o que é comumente chamado, em linguística geral, SN sujeito.

em várias categorias, tais como modalidade, modo, tempo e aspecto. Neste artigo, vamos nos deter nas duas categorias – tempo e aspecto, sobre as quais trataremos a seguir.

3 As categorias tempo e aspecto

As categorias de tempo e de aspecto são cognitivamente muito próximas. Segundo Comrie (1976), tempo é uma categoria dêitica, pois relaciona um determinado fato a um ponto no tempo. Aspecto, por sua vez, “são diferentes maneiras de ver a constituição interna de uma situação” (COMRIE, 1976, p. 3). Ainda segundo o autor, tempo verbal só existe se contrastado a três momentos de fala distintos: passado, presente e futuro. Esses três tempos verbais descrevem o momento anterior, o simultâneo e o posterior. Para Travaglia (2006)⁴, tempo pode ser dividido de três maneiras.

Tempo 1: categoria verbal (corresponde às épocas passado, presente e futuro) (...). Tempo 2: flexão temporal. Estaremos nos referindo então aos agrupamentos de flexões da conjugação verbal: presente do indicativo, pretérito imperfeito do indicativo, futuro do presente, futuro do subjuntivo, etc. Falaremos então em **tempos flexionais** (grifo do autor); Tempo 3: a ideia geral e abstrata de tempo sem consideração de sua indicação pelo verbo ou qualquer outro elemento da frase (...) (TRAVAGLIA, 2006, p. 38).

O texto de Vendler (1957) é clássico na discussão sobre aspecto. O autor propõe uma classificação de tipos de verbos, que poderiam ser divididos em verbos de estado, de atividade, *accomplishments* e *achievements*. A proposta de Vendler é a de que o uso dos verbos pode sugerir diferentes esquemas de tempo e, para explicitar a ideia do autor, usaremos exemplos nossos. Verbos como *escrever* e *construir uma casa* seriam caracterizados como processos que decorrem no tempo. A diferença é que, no último exemplo, há um processo que se encaminha para um ponto final determinado, que é o da casa construída, e *escrever* seria um processo cujo término é indefinido. Verbos como *escrever* foram

⁴ Ainda que os dois autores fujam do escopo da teoria gerativa, quadro eleito para este trabalho, apresentamos Comrie (1976) por ser um texto clássico internacionalmente e Travaglia (2006) por ser uma obra com ampla explicitação da categoria aspectual em língua portuguesa do Brasil.

classificados como atividades, e verbos que constam em expressões como *construir uma casa* foram classificados como *accomplishments*. Aos verbos de atividade e *accomplishments*, Vendler contrapôs verbos como *perceber* e *amar*, que não envolvem sucessão de fases ao longo do tempo. Enquanto *perceber* pode ser caracterizado como um instante de tempo único e definido, *amar* pode ser concebido como um período de tempo indefinido. Para verbos do tipo *perceber*, Vendler classificou como *achievements*; para os do tipo *amar*, verbos de estado.

Alguns autores incrementaram as noções de Vendler (1957), apresentando-as em termos de traços distintivos, como Smith (1997), que trabalha com os traços [estático], [durativo] e [télico]. O traço [estático] diz respeito à ausência de mudança de estado e, desta forma, se isola das demais classificações. O traço [durativo] estaria ligado à quantidade de tempo que cada eventualidade gasta. Esse traço separa a classe dos *achievements* das demais. Já o traço [télico] agruparia *accomplishments* e *achievements*, de um lado, e estados e atividades, de outro. Uma eventualidade télica diz respeito ao encaminhamento para um ponto final, ao passo que os eventos atélicos teriam ponto final arbitrário.

Comrie (1976) também apresenta uma classificação de oposições aspectuais, em que assinala que perfectividade indicaria a perspectiva de uma situação como um todo único, sem distinção das várias fases separadas que fazem aquela situação. Já o imperfectivo estaria relacionado à estrutura interna da situação. O autor (COMRIE, 1976, p. 3) apresenta dois exemplos: “Ele leu” e “Ele estava lendo”. A diferença, nessas duas sentenças, efetivamente, não é de tempo, pois as duas encontram-se no passado. A diferença básica é de aspecto, sendo a primeira sentença marcada pelo aspecto perfectivo e a segunda marcada pelo aspecto imperfectivo. Em língua portuguesa, a noção de (im)perfectividade se dá nos tempos passados de forma clara.

Ancorando-nos em Travaglia (2016, p. 131), que assinala “O presente do indicativo apresenta a situação sempre com aspecto não acabado”, assumimos as formas verbais no pretérito perfeito como perfectivas e as formas verbais no pretérito imperfeito e no presente como imperfectivas.

Dito isso, verificamos que, embora tempo e aspecto sejam categorias muito próximas, conseguimos diferenciá-las. Além disso, por ser a categoria aspecto bastante complexa, uma possível sistematização é conceber que há aspecto semântico/lexical e aspecto gramatical. O aspecto

semântico/ lexical seria inerente aos núcleos verbais, complementos e adjuntos. Dessa forma, os possíveis sentidos atribuídos às raízes verbais seriam responsáveis pela realização do aspecto semântico/ lexical. Olhar para seus complementos e adjuntos (como os advérbios) também seria uma forma de vislumbrar o aspecto semântico/ lexical, que seria caracterizado pelos traços de estatividade, duratividade e telicidade, nos moldes de Smith (1997). Por exemplo, em ‘Construí a casa’, temos uma situação com duração de tempo e telicidade. Já em ‘Caminho pelo bosque’, temos duratividade da atividade, mas não telicidade. Assim, designar a denominação de aspecto semântico/ lexical seria uma forma de vislumbrarmos tal categoria. O aspecto gramatical estaria relacionado, em português, a um conjunto de distinções morfológicas que caracterizariam a circunstância descrita pelo verbo como acabada ou em curso. Carregaria a noção do evento envolvendo distinções semânticas que podem ser interpretadas por meio de verbos auxiliares ou morfemas flexionais e tal noção pode ser separada em perfectivo e imperfectivo.

Isso posto, reiteramos que a nossa pesquisa tem o objetivo maior de compreender como as categorias tempo e aspecto são adquiridas pela criança e, de modo mais específico, entender em que medida as noções de aspecto lexical (semântico) influenciam ou não no surgimento das marcações aspectuais gramaticais. A seguir apresentamos uma abordagem sobre a aquisição de linguagem.

4 Aquisição de linguagem

As crianças, de um modo geral, dos 18 aos 30 meses, demonstram progresso significativo na aquisição de sua língua nativa e é fácil perceber que as crianças fazem isso com muita facilidade. Elas começam a esboçar algumas palavras de caráter lexical e a lista de palavras inicial é bem restrita; em pouco tempo, cerca de um ano e meio, já dominam um vocabulário extenso que é organizado por uma gramática. Esse processo de rápido desenvolvimento dos conhecimentos da língua a que a criança é exposta é evidência de que o ser humano é dotado de um mecanismo inato capaz de guiar a aquisição da linguagem. Nesta perspectiva, as categorias funcionais têm especial destaque, pois parecem ser as responsáveis pela formação da sintaxe, ou seja, são elas que permitem o estabelecimento de inúmeras relações entre estruturas linguísticas que o indivíduo é capaz de produzir.

Assim, à luz do arcabouço da Teoria Gerativa, a aquisição e desenvolvimento da linguagem são interpretados como um processo de fixação de parâmetros a partir do estado inicial da Faculdade da Linguagem. Os parâmetros da GU, segundo Chomsky (1995), estão relacionados não com o sistema computacional, mas, sim, com o léxico. E isso poderia ser interpretado da seguinte forma: cada parâmetro deve referir-se a propriedades de elementos específicos do léxico ou a categorias de itens lexicais. Nesse contexto, existiria apenas uma língua humana e a aquisição da linguagem seria a determinação das idiossincrasias lexicais, ou seja, os elementos substantivos (por exemplo, nomes, verbos etc.) seriam retirados de um vocabulário universal invariante, e só os elementos funcionais seriam parametrizados. Na acepção dos Princípios e Parâmetros, a criança parte de uma base pré-moldada à experiência (que corresponde aos princípios) e vai adequando os dados da língua que está adquirindo a essa base (que seriam os parâmetros).

A partir do trabalho de Borer (1984), a visão de parâmetros é relacionada exclusivamente com as categorias funcionais. A ideia é que, se as línguas naturais variam no que diz respeito à sua morfologia, por conseguinte os elementos responsáveis pela morfologia é que devem estar relacionados aos parâmetros, isto é, os elementos morfológicos, que, em tese, são as categorias funcionais, são aqueles responsáveis pela diferença de fixação de parâmetro. As categorias funcionais, por exemplo, seriam as responsáveis pela fixação, nas línguas naturais distintas, da ordem sujeito-verbo ou posição do verbo em uma sentença. Como as categorias funcionais pertencem a uma classe fechada, constituem um conjunto pequeno de elementos. Por conseguinte, a criança, ao adquirir uma língua, faz escolhas dentro de um conjunto muito pequeno de opções, o que explicaria a rapidez do processo de aquisição da língua, que é uma das preocupações centrais da Teoria Gerativa.

Para Wexler (1996), a criança passa pelo “Estágio do Infinitivo Opcional” ou Estágio OI. Nessa etapa, a criança ora produz, ora omite tempo verbal. Dessa forma, a criança oscilaria no uso de formas finitas e formas não finitas. Em 1998, o autor sofisticou sua explicação, sustentando que crianças, desde muito cedo, já têm os parâmetros básicos e as propriedades gramaticais de sua língua realizados corretamente e defendendo duas hipóteses, conforme nos assinalam Hermont e Morato (2014):

VEPS (*Very Early Parameter-Setting*), que aponta que os parâmetros básicos são realizados corretamente em estágios iniciais, quando a criança tem por volta de 18 meses de idade; e VEKI (*Very Early Knowledge of Inflection*), que sugere que as crianças, a partir de 18 meses, já conhecem propriedades fonológicas e gramaticais de alguns elementos importantes de sua língua (WEXLER, 1998 apud HERMONT; MORATO, 2014, p. 219).

Wexler (1996) adota a explicação já desenvolvida em Schütze e Wexler (1996) para as sentenças em que ora ocorre flexão de tempo e ora ocorre flexão de concordância. O modelo, conhecido como ATOM (Modelo de Omissão de Concordância e Tempo no Estágio OI), é consequência da atuação de uma Restrição de Checagem Única, que vem a ser uma propriedade do Estágio OI (infinitivo opcional) e indicaria que a gramática da criança, em alguns momentos, impede que o DP-sujeito cheque⁵ seus traços mais de uma vez durante a derivação de uma sentença. De acordo com os autores, isso deve ocorrer devido ao fato de a gramática da criança “enxergar”, algumas vezes, o DP-sujeito como tendo traços [-interpretáveis]. Quando tal gramática “enxerga” o DP-sujeito como tendo traços [+interpretáveis], a valoração dupla pode ocorrer e a sentença tem tempo e concordância, tal como a fala do adulto com a gramática intacta.

Wexler (1998) adotou a noção de dois nódulos, tempo e concordância, neste trabalho, o pesquisador analisou os dados sob a ótica da Morfologia Distribuída. As sentenças produzidas pelas crianças com atribuição apropriada de Caso e de Tempo Finito são gramaticais e perfizeram um total de 95%. Seriam sentenças como em *He likes ice-cream*. Assim, na gramática da criança haveria a seguinte representação: [+AGR, +T], em que Concordância e Tempo seriam especificados. Para os 54% de uso da forma verbal no infinitivo com os pronomes *he* e *she* que surgiram na pesquisa, a explicação é que, se há esses pronomes, ainda que faltando o morfema caracterizador de tempo, provavelmente há concordância na gramática da criança porque houve a designação do caso nominativo, isto é, foram produzidos os pronomes *he* e *she*, e não os de caso acusativo, *him* e *her*. Concordância, nesta perspectiva, estaria ligada a caso nominativo. As sentenças produzidas pelas crianças seriam

⁵ O autor usava na época esta terminologia. Depois, passou a se chamar ‘valorar’ (os traços) no lugar de ‘checar’.

parecidas com *He like ice-cream* e a representação mental seria [+AGR, -T], em que teríamos, como especificação de Tempo [+passado], em que o morfema é *-ed*, competindo com [-passado], representado pelo morfema \emptyset . Se Tempo estivesse especificado para passado, o morfema *-ed* seria inserido, porque tem mais especificações que o morfema de Tempo presente. Se Tempo estivesse marcado para presente, o morfema de presente \emptyset seria inserido. Como na sentença há concordância, teríamos que ter a marca *-s*. Como não há o morfema de concordância – (morfema de concordância que só ocorre no presente), temos fortes sugestões de que não há marcação de tempo presente na sentença, ou seja, a especificação, neste exemplo, seria [+AGR, -T]. No que diz respeito às sentenças em que houve 46% de uso de verbos na forma infinita com os pronomes *him* e *her*, as sentenças teriam uma estrutura parecida com a de *Him like ice-cream*, e a representação mental seria [-AGR, -T]. A explicação para o não surgimento de morfema relativo a tempo é a mesma feita para a última situação explicitada. Já o caso de 5% de produção de formas verbais flexionadas e sujeitos com caso acusativo, por parte de crianças no período OI, a explicação está relacionada à falta de concordância, pois essa categoria é que atribuiria caso nominativo.

Mas a proposta de ATOM, por si só, não consegue explicar o motivo de ora haver só Tempo e ora só AGR na gramática das crianças. Wexler (1998) sugere, então, que deve haver uma restrição no momento de checagem do DP sujeito na gramática da criança. Para explicar como se dá essa restrição, ancoremo-nos na leitura de Wexler (1998) feita por Hermont e Morato (2014). Wexler adota as propostas de Chomsky (1995) e assume que os DPs-sujeito, que têm um traço [+interpretável⁶], passariam a movimentar-se a fim de checar o traço D [-interpretável] em Tempo. Wexler (1998), admitindo a presença da projeção de AGR, assume que tanto T quanto AGR têm um traço D [-interpretável]. Se tanto AGR quanto T têm um traço D [-interpretável], um traço D [+interpretável] deve deslocar-se para checar aqueles [-interpretáveis]. De acordo com o

⁶ Faz-se importante explicitar a diferença entre traços interpretáveis e traços não interpretáveis. Os primeiros têm importância na computação de representações semânticas e não devem ser apagados (e nem o podem) do sistema. Já os traços não interpretáveis, ou seja, que não têm interpretabilidade semântica e que devem estar circunscritos ao componente computacional, como, AGR (na acepção de Chomsky, 1995), devem ser eliminados do sistema, a fim de que a estrutura possa convergir.

pesquisador, dentro do SV há um DP sujeito com traços [+interpretáveis]. Portanto, esse DP sujeito é que deve mover-se para apagar os traços nas categorias funcionais, que têm traços [-interpretáveis], os quais devem ser apagados a fim de que a derivação não fracasse. Na gramática do adulto, o DP sujeito então vai para o especificador de Tempo e checa o traço D [-interpretável], em seguida faria o mesmo no especificador de AGR e checa o traço D [-interpretável]. O que parece ser complicado para a criança em processo de aquisição da linguagem, segundo Wexler (1998), é o fato de serem necessárias duas checagens. Desse modo, o autor coloca que haverá duas representações para a criança, cuja gramática revelará que ou AGR ou T terá os traços [-interpretáveis] checados, gerando a omissão de morfema de uma ou outra categoria na fala da criança.

A ideia de Wexler (1998) será usada para explicar os dados obtidos em nossa pesquisa, usando a noção de aspecto gramatical no lugar de concordância. O raciocínio é o seguinte: concordância, sendo caracterizada como [-interpretável] na interface semântica, não haveria necessidade de resguardar-se um nódulo específico para tal categoria. Entretanto, existem noções da estrutura cognitiva imanescentes aos verbos das línguas naturais que são bastante relevantes, tais como tempo e aspecto. Como essas duas noções são fenômenos intrincados, mas distintos na interface semântica, pode-se pensar que, no lugar de concordância, podemos ter um nódulo de aspecto. Assim sendo, passaremos a trazer dados de pesquisas que investigaram a aquisição da categoria aspectual.

4.1 Abordagens que tratam da aquisição de aspecto verbal

Alguns trabalhos que visaram a descrever e a analisar a aquisição de aspecto verbal foram realizados na década de 70 e 80 e tiveram o objetivo de validar a hipótese estabelecida por Jakobson (1971), que propôs que aspecto verbal é adquirido antes de tempo verbal. Para o entendimento das propostas a serem apresentadas, é interessante que se tenham em mente as duas formas de conceber aspecto já delineadas neste trabalho: o aspecto semântico/lexical e o aspecto gramatical. Passemos agora à explicitação dos trabalhos de Bronckart e Sinclair (1973) e Bloom, Lifter e Hafitz (1980), para os quais a emergência de flexões verbais na fala das crianças seria uma influência da semântica verbal e não exatamente da aquisição da morfologia de tempo.

Bronckart e Sinclair (1973) realizaram um estudo para verificar a relação semântica dos verbos e o surgimento de flexões verbais no processo de aquisição do francês por parte de crianças de dois a oito anos. Os pesquisadores observaram que, até os seis anos, as crianças tendiam a usar, nos experimentos, o passado composto para ações cujo resultado era muito claro. Já para as ações que não tinham a ideia de fim, tais crianças produziam formas verbais com morfemas de imperfectividade, ou seja, os resultados dos experimentos desenvolvidos, de acordo com os autores, sugeriam que as crianças usavam flexões verbais para indicar diferenças aspectuais, e não temporais.

Resultados semelhantes foram obtidos por Bloom, Lifter e Hafitz (1980) em língua inglesa. Esses pesquisadores verificaram que verbos como ‘brincar’ e ‘escrever’, caracterizados pelo traço de duratividade, surgiam, na fala das crianças, de forma recorrente, com o morfema -ing. Já verbos como ‘encontrar’, ‘cair’ e ‘quebrar’, marcados claramente pela não duratividade, ocorriam na fala infantil com o morfema -ed. Para os autores, o uso seletivo de diferentes morfemas era determinado pelo aspecto inerente de cada verbo.

Em Hermont e Morato (2014), buscou-se compreender como as categorias Tempo e Aspecto estão representadas na gramática de crianças em fase de aquisição de linguagem e na gramática de pessoas com Déficit Específico de Linguagem (DEL). A pesquisa desenvolvida dentro do arcabouço teórico da teoria gerativa propôs uma dissociação do sintagma de flexão em um sintagma de tempo e sintagma de aspecto e sugeriu influência do aspecto semântico sobre o aspecto gramatical. Foram comparadas estruturas verbais produzidas por uma criança em fase normal de aquisição de linguagem e por uma pessoa com DEL. Neste trabalho, também se expressou que o sistema computacional, a sintaxe de uma língua, deve ser sensível às características aspectuais do SV, considerando, portanto, a natureza aspectual lexical/ semântica dos verbos.

Castro e Hermont (2017) examinaram a relação entre o aspecto lexical e o aspecto gramatical na fala de uma criança em fase de aquisição da linguagem, tendo por base a classificação aspectual proposta por Vendler (1957). As autoras verificaram que as formas verbais produzidas pela criança investigada na pesquisa sugerem uma associação do aspecto lexical/ semântico com traço [-durativo], evidenciada em verbos caracterizados por serem *achievement*, e aspecto gramatical perfectivo. Além disso, Castro e Hermont (2017) constataram uma relação entre

o aspecto semântico/ lexical com traço [+durativo] – encontrado em verbos de estado, atividade e *accomplishment* – e o aspecto gramatical imperfectivo, o que indica que o aspecto semântico/ lexical pode influenciar o surgimento do aspecto gramatical.

Vale dizer que Castro e Hermont (2017) ancoraram-se na Hipótese da Primazia do Aspecto e na Generalização de Li-Shirai. Segundo Andersen (1989) e Shirai e Andersen (1996), entre outros, a Primazia do Aspecto seria um fenômeno que limitaria o marcador morfológico de tempo/ aspecto a uma determinada classe de verbos de acordo com seu aspecto inerente, o que quer dizer que as flexões verbais utilizadas pelas crianças em processo inicial de aquisição da linguagem denotariam, na verdade, a noção de aspecto lexical. Assim, aspecto semântico/ lexical comandaria o uso da flexão verbal.

O que é importante apontar sobre todos os trabalhos apresentados que versam sobre a aquisição da categoria aspectual é que consideram que a noção de aspecto, no início da aquisição da linguagem, é que guia a expressão temporal. Assim, a criança, na verdade, ao produzir um verbo com morfema temporal, estaria realizando a noção aspectual. E isso se daria por uma relação muito próxima do que denominamos aspecto semântico/ lexical e aspecto gramatical. No caso da língua portuguesa, a detecção de qual categoria, de fato, está sendo produzida é mais difícil de ser feita quando a forma verbal é simples. Isso porque a desinência que expressa modo e tempo é a mesma que denota aspecto. Um exemplo: em ‘am-á-va-mos’, em -va-, temos as três noções expressas. Para expressar aspecto gramatical, deve-se ‘produzir’ o -va-, que é a mesma categoria que expressa a noção temporal (e modo). Quando se observam as formas perifrásticas, temos, com clareza, a realização do tempo, que se dá no auxiliar, aliada ao verbo principal que ocorre no infinitivo, gerúndio ou particípio. Para essas duas últimas formas, temos a noção de imperfectividade e perfectividade, respectivamente.

Finalizando esta seção, podemos assinalar que os resultados de Bronckart e Sinclair (1973), de Bloom, Lifter e Hafitz (1980), de Hermont e Morato (2014), de Castro e Hermont (2017), cada um com objetivos específicos, são consistentes com a hipótese de aquisição de aspecto antes do tempo por parte das crianças em fase de aquisição de linguagem, em que aspecto semântico/ lexical pode influenciar o surgimento do aspecto gramatical.

Com base neste arcabouço teórico, reafirmamos que o objetivo principal deste trabalho é entender como as categorias – tempo e aspecto – estão representadas nas gramáticas de crianças em processo de aquisição de linguagem. Para tal, apresentamos a seguir a metodologia adotada para este trabalho.

5 Metodologia

Uma pesquisa realizada no arcabouço da Teoria Gerativa pressupõe, conforme já assinalado, a existência de princípios universais sobre a gramática mental, sendo que esses podem ser observados nas produções orais e compreensões linguísticas de cada indivíduo. Esta pesquisa detém-se na análise de dados advindos de produções de linguagem por parte de crianças em idade de aquisição típica de linguagem.

Assim, pode-se também dizer que os pesquisadores dessa área partem da análise do desempenho linguístico para compreender a competência dos falantes e, nesta perspectiva, utilizam o método dedutivo. Faremos uso desse método para a análise do surgimento de morfemas relacionados às categorias de tempo e de aspecto na fala das crianças. O método dedutivo tem como ponto de partida um problema, isto é, de um conjunto de fatos para o qual não se tem uma explicação aceitável ou uma explicação que ainda não foi consolidada. Algumas hipóteses são apresentadas como possíveis explicações ou soluções provisórias para o problema formulado. As hipóteses inicialmente criadas servem de impulso para o pesquisador buscar fatos complementares, o que faz com que apareçam novas hipóteses. Dentre as hipóteses enumeradas, o pesquisador optará por aquela que lhe parecer a mais provável submetendo-a a testes e experiências, no caso. Para este trabalho, estabelecemos as seguintes hipóteses, aqui repetidas: (i) na gramática infantil, a noção de aspecto surge, de forma mais recorrente, antes da noção de tempo; (ii) quando as formas verbais no presente e no pretérito perfeito surgem, o que está expresso é o aspecto.

Para a realização da pesquisa, realizamos um estudo transversal, a fim de refletir acerca de dados coletados de uma amostra em um determinado período de tempo e examinar a relação entre as ocorrências verbais, usando nossas lentes nas realizações das categorias de tempo e de aspecto. Uma das características mais importante do estudo transversal é que se torna possível comparar diferentes amostras de diferentes períodos de tempo.

Nesta perspectiva, passemos a apresentar a amostra do nosso trabalho. O *corpus* é constituído por gravações espontâneas de cinco crianças em fase de aquisição de linguagem, com idades variando entre dois e cinco anos. Tais gravações foram catalogadas, como já anunciado neste artigo, por Ana Lúcia Barros Leôncio (LEÔNCIO, 2018, p. 114) e Washington Gomes de Miranda (MIRANDA, 2018, p. 136), membros do grupo eLinC. A pesquisa aqui apresentada analisou os dados já transcritos. Durante a gravação, os pais estavam presentes e interagindo com as crianças. As pesquisas foram submetidas e aprovadas pelo Comitê de Ética em Pesquisa envolvendo Seres Humanos da Pontifícia Universidade Católica de Minas Gerais. Utilizamos todas as gravações realizadas para cada criança, configurando o seguinte quadro:

Quadro 1 – Faixa etária das crianças

CRIANÇA	IDADE
A	2 anos
B	2 anos e 1 mês
B	2 anos e 2 meses
B	2 anos e 3 meses
C	3 anos e 1 mês
C	3 anos e 2 meses
D	3 anos e 11 meses
D	4 anos e 2 meses
E	5 anos e 5 meses
E	5 anos e 11 meses

Fonte: elaborado pelos autores.

Para a criança A, temos somente uma gravação, quando ela tinha dois anos. Para a criança B, usamos três gravações, com diferença de um mês: dois anos e um mês, dois anos e dois meses e dois anos e três meses. Em relação à criança C, tínhamos duas gravações. Nas ocasiões, a criança tinha três anos e um mês e três anos e dois meses. Já para a criança D, conseguimos duas gravações em que a criança tinha três anos e onze meses e outra, em que a criança tinha quatro anos e dois meses. Por fim, para a criança E, obtivemos duas gravações em que a criança tinha cinco anos e cinco meses e cinco anos e onze meses.

A realização desta pesquisa contou com três momentos de investigação. No primeiro, observamos, em cada sentença produzida

pelas crianças, quais foram os tempos manifestados nas formas verbais: seja na forma simples, seja nas perífrases verbais, isto é, verificamos se o tempo estava no passado, no presente ou no futuro. No segundo momento, buscamos a identificação do aspecto semântico/ lexical detendo-nos especialmente no traço de telicidade e, em seguida, verificamos o surgimento de formas imperfectivas e perfectivas no repertório dessas crianças. Deve-se salientar que são consideradas imperfectivas as formas que estão no presente e no pretérito imperfeito. São designadas como perfectivas as formas no pretérito perfeito. Ao depararmos com formas perifrásticas, verificamos se o verbo principal estava na forma participial ou no gerúndio, os quais revelam, de um modo geral, a forma aspectual. Observamos também o tempo gramatical expresso no auxiliar. No terceiro momento, contabilizamos as produções das formas verbais relacionando-as às idades das crianças. Em seguida, procedeu-se à análise dos dados observados, os quais passam a ser descritos na seção posterior.

6 Descrição e análise de dados

Lembremo-nos dos objetivos traçados para este trabalho: entender como as categorias tempo e aspecto estão representadas nas gramáticas de crianças em processo de aquisição de linguagem a partir da análise de estruturas efetivamente produzidas por cinco crianças no processo de aquisição de linguagem, além de compreender em que medida as noções de aspecto semântico/ lexical influenciam ou não no surgimento das marcações aspectuais gramaticais. Após a coleta de dados, constatamos o seguinte (quadro 2):

Quadro 2 – Dados da Pesquisa

Criança	Idade	Aspecto Imperfectivo				Aspecto Perfectivo	Total de produções
		Presente	Gerúndio	Perífrase Verbal	Pretérito Imperfeito	Pretérito perfeito	
A	2 anos	-	-	1 (1C)	-	1	2
B	2 anos e 1 mês	-	-	-	-	1	
B	2 anos e 2 meses	9	-	4 (1C)	-	14	
B	2 anos e 3 meses	6	1	-	-	9	44
C	3 anos e 1 mês	7	-	-	-	1	
C	3 anos e 2 meses	10	-	1	1	3	23
D	3 anos e 2 meses	5	-	5 (5C)	-	-	
D	3 anos e 2 meses	127	2	29 (3C)	9	36	213
E	5 anos e 5 meses	14	1	26 (1C)	5	17	
E	5 anos e 11 meses	66	2	30 (8C) ^{<-?>}	5	11	177

Fonte: elaborado pelos autores.

Observando o quadro anterior, assinalamos que cada número corresponde ao total de formas verbais caracterizadas por determinado tempo e aspecto, sendo que (C) diz respeito a verbos controle presentes em perífrases verbais. Após a análise do quadro, verificamos que as crianças de dois a três anos produzem menos formas verbais que as crianças com mais de quatro anos. Constatamos também que as formas verbais no presente e no pretérito perfeito são mais frequentes na fala da criança nas primeiras fases da vida (dois a três anos) e, próximo aos quatro anos, surgem, de forma recorrente, as perífrases verbais (constituídas de auxiliar e os verbos principais) e as formas verbais no pretérito imperfeito.

Ancorando-nos na literatura apresentada neste artigo, podemos afirmar que as formas verbais que surgem inicialmente na produção linguística das crianças denotam aspecto e, não, tempo verbal. Para explicar isso, estabelecemos uma relação com o estudo de Wexler (1998) e os nossos dados e verificamos uma semelhança. O sistema de flexão proposto pelo autor é constituído por Tempo e Concordância e, para ele, ora há morfemas expressando traços de tempo, ora de concordância. Amparando-nos na assunção de que concordância não é mais considerada uma projeção dentro da arquitetura de linguagem do modelo minimalista atual, podemos substituir o raciocínio com a adoção de um nóculo aspectual e, assim sendo, declaramos que, na gramática infantil, ora parece haver a valoração dos traços de tempo (com menor recorrência), ora parece haver a valoração dos traços de aspecto (em maior quantidade). Como a maior ou menor recorrência pode ser explicada?

A ideia é que o tempo está presente, com mais clareza, nas formas perifrásticas, que ocorrem de forma mais produtiva dos quatro anos em diante, tais como em:

- (1) eu **tô pedindo** a minha mãe pra abri porque:: (Criança D – 4 anos e 2 meses)
- (2) Cê **tá colorindu** di duas coris aqui vovô. (Criança E – 5 anos e 11 meses)
- (3) **Tô imitando**. (Criança E – 5 anos 11 meses)

Os auxiliares “tô” (estou), “tá” (está) e “tô” expressam tempo e ‘pedindo’, ‘colorindo’ e ‘imitando’, todos na forma de gerúndio, expressam o aspecto. A noção aspectual veiculada pelo auxiliar não será discutida aqui, pois o objetivo, neste momento, é demonstrar que tempo é veiculado no auxiliar nas perífrases verbais e tais estruturas ocorrem mais tarde no processo de desenvolvimento da linguagem. O que se deve salientar aqui é que o auxiliar manifesta a noção temporal nas perífrases verbais e elas são largamente encontradas de forma produtiva na produção linguística das crianças estudadas após os quatro anos de idade.

Mas em relação às formas verbais simples, que ocorrem desde muito cedo, como saber se temos as duas noções – aspecto e tempo – manifestadas? Essa pergunta se justifica porque o morfema que designa tempo também codifica os traços aspectuais. Vejamos o quadro 3.

Quadro 3 – Inspirado no sistema flexional do português, segundo Câmara Jr. (1970)

Classificação	Forma verbal	Radical + vogal temática	Desinência de modo, tempo e aspecto	Desinência de número e pessoa
3ª pessoa sing. Presente do Indicativo	(ele) Canta	Canta-	- Ø -	- Ø
3ª pessoa plural Presente do Indicativo	(eles) Cantam	Canta-	- Ø -	- m
3ª pessoa sing. Pretérito Imperfeito do Indicativo	(ele) Cantava	Canta-	- va -	- Ø
3ª pessoa plural Pretérito Imperfeito do Indicativo	(eles) Cantavam	Canta-	- va -	- m
3ª pessoa sing. Pretérito Perfeito do Indicativo	(ele) Cantou	Canto-	- Ø -	- u
3ª pessoa plural Pretérito Perfeito do Indicativo	(eles) Cantaram	Canta-	- ra -	- m

Fonte: elaborado pelos autores.

O morfema de tempo em formas simples acumula noção de modo, mas também acumula a noção de aspecto, seja de perfectividade ou de imperfectividade. Portanto, a ideia é que, quando a criança produz formas verbais simples, ela está expressando ora noção temporal, ora noção aspectual. Associando as primeiras formas verbais simples àquelas perifrásticas, em que, claramente, há tempo marcado no auxiliar, a sugestão é que majoritariamente há mais denotação de aspecto nas formas simples no início do processo de aquisição da linguagem.

Esse raciocínio pode ser realizado nos moldes de Wexler (1998). Assim sendo, poderíamos adotar a proposta de Restrição de Checagem Única feita por Wexler (1998), para explicar os dados relativos a Tempo e a Aspecto presentes ou ausentes na fala das crianças pesquisadas, argumentando que há a possibilidade de tais crianças terem as seguintes representações na gramática mental a serem escolhidas para valoração de traços: ora ocorreria a valoração dos traços de aspecto, ora ocorreria a valoração de traços de tempo.

De forma interessante, quando se observa o grande surgimento de formas perifrásticas (auxiliar e verbo principal), verifica-se que ocorrem, ainda que em baixo número, as primeiras formas verbais simples no

pretérito imperfeito. Ora, se nesta fase da aquisição da linguagem a criança já produz o auxiliar com o tempo demarcado claramente em auxiliares nas formas perifrásticas, a ideia é que, ao produzir as formas verbais no pretérito imperfeito, já ocorreria, na gramática mental da criança, a valoração tanto de traços de aspecto, quanto de tempo, tal como ocorre na gramática mental de um adulto.

Realizando um cruzamento das noções de aspecto semântico/lexical com as noções de aspecto gramatical das formas verbais produzidas pelas crianças A, B e C, que são as mais novas, verificamos que há uma forte relação entre as informações aspectuais semântico/lexical e gramaticais, conforme se pode averiguar no quadro 4.

Quadro 4 – Cruzamento das noções de aspecto semântico/lexical e aspecto gramatical

Criança A – 2 produções	Criança B – 42 produções	Criança C – 23 produções
Atélico + Imperfeito: 1 Télico + Perfectivo: 1	Atélico + Imperfeito: 18 Atélico + Perfectivo: 2 Télico + Perfectivo: 22	Atélico + Imperfeito: 14 Atélico + Perfectivo: 2 Télico + Imperfeito: 5 Télico + Perfectivo: 2

Fonte: elaborado pelos autores.

Observando o quadro anterior, podemos ter uma ideia clara da influência do traço de telicidade determinando o surgimento do morfema perfectivo (por exemplo, “Isso qui comi mimi boi mom” (Criança A), “papai igô” (Criança B),) e do traço de atelicidade desencadeando a emergência de morfemas relacionados à imperfectividade (por exemplo, “Quelia amarrá” (Criança A), “Eu tinhu medu di altura quanu eu *tava nadanu* isso é baxu” (Criança C)). Para as crianças A e B, o resultado é bastante coerente como é previsto neste trabalho. Para a criança C, ainda que os verbos caracterizados pela telicidade (e que surgiram em número muito pequeno) não tenham ocorrido, na maior das vezes, sob a forma perfectiva ((Ela) *casca* e Ela *brinca*, ambos os verbos marcados pela telicidade sob a forma imperfectiva), podemos verificar que quase todos os verbos marcados pela atelicidade, ocorreram na forma imperfectiva. Esses resultados, portanto, de um modo geral, parecem ir ao encontro do que preconiza a Hipótese da primazia do Aspecto proposta por Andersen & Shirai (1996) e Castro & Hermont (2017).

7 Considerações finais

Este trabalho delineou dois principais objetivos: verificar a aquisição de formas verbais relacionadas a tempo e a aspecto e entender em que medida as noções de aspecto semântico/ lexical influenciam ou não no surgimento das marcações aspectuais gramaticais no processo de desenvolvimento da linguagem.

As duas hipóteses estabelecidas foram confirmadas. Pelo raciocínio desenhado e ancorando-nos em uma literatura que trata do tema, parece que aspecto é adquirido antes de tempo na gramática infantil. Além disso, observando os dados obtidos na pesquisa, as únicas formas verbais nas fases iniciais do desenvolvimento da linguagem são o presente e o pretérito perfeito, caracterizados pelos traços da imperfectividade e da perfectividade, respectivamente.

Ainda que estejamos evidenciando que aspecto parece surgir antes de tempo na linguagem infantil, esta última categoria parece também já estar disponível para a criança desde os dois anos em consonância com os pressupostos de Wexler (1996, 1998), VEPS, (*Very Early Parameter-Setting*), confirmando que os parâmetros básicos são realizados de forma adequada pela criança em estágios iniciais, e VEKI (*Very Early Knowledge of Inflection*), que assevera que as crianças de um ano e meio já conhecem propriedades fonológicas e gramaticais de alguns elementos importantes de sua língua. Tanto é que surgiram formas como “tô siguino”, por parte da criança B aos dois anos e dois meses, e “Eu tinhu medo di altura quanu eu tava nadanu”, por parte da criança C com três anos e dois meses, ou seja, a categoria temporal expressa no auxiliar está presente desde a mais tenra idade, embora de forma escassa.

Ao lado dos resultados obtidos, podemos ainda verificar algo bastante interessante e que enseja futuras investigações, pois, no processo da aquisição da linguagem, a forma verbal no pretérito imperfeito surge mais tarde que a forma verbal no presente e no pretérito perfeito. Essa forma apresenta, tal como as outras duas, um morfema que acumula várias categorias. A única diferença é que tal morfema é foneticamente exposto. Além disso, o surgimento da forma verbal no pretérito imperfeito se dá na mesma ocasião em que aparecem as perífrases verbais, expressões essas que, claramente, denotam tempo.

Declaração de autoria

Arabie Bezri Hermont realizou a orientação da pesquisa, averiguação da análise dos dados e a redação do texto. Kelly Cesário de Oliveira realizou a pesquisa e a redação do texto. Lucas Segantini Brito realizou a pesquisa e a redação do texto.

Referências

- ANDERSEN, R. W. The acquisition of verbal morphology. Los Angeles University of California. Published in Spanish as 'La adquisición de la morfología verbal'. *Linguística*, Los Angeles, v.1, p. 89-141, 1989.
- ANDERSEN, R. W.; SHIRAI, Y. Primacy of Aspect in First and Second Language Acquisition: The pidgin/creole connection. In: RITCHIE, W.; BHATIA, T. (eds.). *Handbook of second language acquisition*. San Diego: Academic Press, 1996. p. 527-570.
- BLOOM, L.; LIFTER, K.; HAFITZ, J. Semantics of verbs and the development of verb inflection in child language. *Language*, Washington, DC, v. 56, n. 2, 1980. DOI: <https://doi.org/10.1353/lan.1980.0001>
- BORER, H. *Parametric syntax*. Dordrecht: Foris Publications, 1984.
- BRONCKART, J.; SINCLAIR, H. Time, Tense and Aspect. *Cognition*, Genebra, n. 2, p. 107-130, 1973. DOI: [https://doi.org/10.1016/0010-0277\(72\)90032-7](https://doi.org/10.1016/0010-0277(72)90032-7)
- CÂMARA JR., J. M. *Estrutura da Língua Portuguesa*. Petrópolis: Vozes, 1970.
- CASTRO, G. G.; HERMONT, A. B. A relação entre o aspecto lexical e o aspecto gramatical em contexto de aquisição da linguagem. *PERcursos Linguísticos*, Vitória, v. 7, n. 14, p. 405–420, 2017. Disponível em: <<https://periodicos.ufes.br/percursos/article/view/15620>>. Acesso em: 20 mar. 2021.
- CHOMSKY, N. *Aspects of theory of syntax*, Cambridge: MIT Press, 1965.
- CHOMSKY, N. *Beyond explanatory adequacy*, Ms. Cambridge: MIT, 2001 (não publicado).
- CHOMSKY, N. Derivation by Phase. In: KENSTOWICZ, M. (ed.). *Ken Hale: A life in language*. Cambridge: MIT Press, 2001. p. 1-52.

CHOMSKY, N. *Knowledge of language: Its nature, origin, and use*. Cambridge: MIT Press, 1986.

CHOMSKY, N. *Language and mind*. New York: Harcourt Brace Jovanovich, 1972.

CHOMSKY, N. *Lectures on government and binding*. Dordrecht: Foris Publications, 1981.

CHOMSKY, N. *Minimalism Inquiries: the framework*. Cambridge, MA: MIT Press, 1998.

CHOMSKY, N. *The minimalist program*. Cambridge: MIT Press, 1995.

CINQUE, G. *Adverbs and functional heads: a cross-linguistics perspective*. New York: Oxford University Press, 1999.

COMRIE, B. *Aspect*. Cambridge: Cambridge University Press, 1976.

EMONDS, J. *A transformation approach to syntax*. New York: Academic Press, 1976.

HERMONT, A. B. *Aquisição de tempo e aspecto no déficit especificamente linguístico*. 2005. 284 f. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2005.

HERMONT, A. B.; MORATO, R. A. Aquisição de tempo e aspecto em condições normais e no déficit específico de linguagem. *Revista Lingüística*, Rio de Janeiro, v. 10, p. 213-233, 2014. DOI: <https://doi.org/10.31513/linguistica.2014.v10n1a4588>

JAKOBSON, R. Shifters, verbal categories, and the Russian verb. In: JAKOBSON, R. *Selected writings*. Vol. II. Mouton: The Hague, 1971. p.130-147.

LEÔNCIO, A. L. B. *O objeto nulo na aquisição da linguagem*. 2018. 149 f. Dissertação (Mestrado em Estudos Linguísticos) – Programa de Pós-Graduação em Letras, Pontifícia Universidade Católica de Minas Gerais, 2018.

LEÔNCIO, W.; MIRANDA, A.L. *Corpus de produção de linguagem infantil*, no prelo.

MIRANDA, W. G. de. *Advérbios de tempo e aspecto no processo de aquisição da linguagem*. 2018. 174 f. Dissertação (Mestrado em Estudos Linguísticos) - Programa de Pós-Graduação em Letras, Pontifícia Universidade Católica de Minas Gerais, 2018.

POLLOCK, J. Verb movement, universal grammar and the structure of IP. *Linguistic Inquiry*, v. 20, n.3, p. 365-425, 1989.

SCHÜTZE, C.; WEXLER, K. Subject case licensing and English root infinitives. In: STRINGFELLOW, A.; CAHANA-AMITAY, D.; HUGHES, E.; ZUKOWSKI, A. (eds.). *Proceedings of the 20th Boston University Conference on Language Development*. Somerville: Cascadilla Press, 1996. p. 670-681.

SMITH, C. S. *The Parameter of Aspect*. 2. ed. Dordrecht: Springer, 1997.

SOLÀ, J. Morphology and word order in germanic languages. In: ABRAHAM, W.; EPSTEIN, S.; THRAINSSON, H.; ZWART, J-W. (eds.). *Minimal Ideas: syntactic studies in the minimalist framework*. Amsterdam: John Benjamins, 1996. p. 217-251.

TRAVAGLIA, L. C. *O aspecto verbal no português: a categoria e sua expressão*. 4. ed. Uberlândia: EDUFU, 2006.

TRAVAGLIA, L. C. *O aspecto verbal no português: a categoria e sua expressão*. 5.ed. Uberlândia: Universidade Federal de Uberlândia, 2016.

VENDLER, Z. Verbs and Times. *Philosophical Review*, v. 66, n. 2, p. 143-160, 1957. DOI: <https://doi.org/10.2307/2182371>

WEXLER, K. The development of inflexion in a biologically based theory of language acquisition. In: RICE, M. L. *Toward a genetics of language*. Mahwah: Lawrence. Erlbaum Assoc., 1996. p. 113-144.

WEXLER, K. Very early parameter setting and the unique checking constraint: a new explanation of the optional infinitive stage. *Lingua*, Cambridge, MA, v. 106, p. 23-79, 1998. DOI: [https://doi.org/10.1016/S0024-3841\(98\)00029-1](https://doi.org/10.1016/S0024-3841(98)00029-1).



Entre accomplishments e atividades: mudanças direcionais e o caso dos verbos de editoração

Between accomplishments and activities: directed changes and the case of publishing verbs

Gabriela Vilela Souza Martins

Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Minas Gerais / Brasil

gabrielavsm@ufmg.br

<https://orcid.org/0000-0002-5602-2845>

Luana Lopes Amaral

Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Minas Gerais / Brasil

luanalopes@ufmg.br

<https://orcid.org/0000-0002-4290-1208>

Resumo: Este trabalho se propõe a analisar e a classificar aspectualmente os verbos de editoração do inglês e do português: *to edit/editar* ou *editar*, *to copy edit/preparar*, *to proofread/revisar*, *to typeset/diagramar* e *to format/formatar*. Esses verbos se apresentam como um interessante objeto de análise no âmbito do aspecto lexical porque possuem tanto características de *accomplishments* quanto características de atividades. Esse tipo de comportamento dúbio foi amplamente estudado na literatura sobre aspecto e é normalmente atribuído a propriedades do VP. No caso dos verbos de editoração, porém, a variabilidade da classificação aspectual parece não estar estritamente atrelada ao tipo de complemento, mas à incrementalidade da ação descrita pelo verbo. Assim, para a análise desses verbos, adotamos o modelo bidimensional de representação do aspecto lexical, numa abordagem cognitivo-funcional, propondo que esses verbos possuem propriedades aspectuais que estão na interseção entre *accomplishments* e atividades. Nossos resultados indicam que os verbos de editoração especificam um tipo de mudança direcional. Assim, eles possuem potencial para serem (re)construídos como mudanças direcionais télicas (*accomplishments*) ou atélicas (atividades), a depender da forma como o falante pretende descrever o evento de editoração.

Palavras-chave: aspecto lexical; verbos de editoração; incrementalidade; telicidade; mudança.

Abstract: This work aims to analyze and aspectually classify the publishing verbs in English and Portuguese: *to edit/editar* or *editorar*, *to copy edit/preparar*, *to proofread/revisar*, *to typeset/diagramar* and *to format/formatar*. These verbs present themselves as an interesting object of analysis in terms of the lexical aspect because they have characteristics of both accomplishments and activities. This type of dubious behavior has been extensively studied in the aspect literature and is usually attributed to VP properties. In the case of publishing verbs, however, the variability of the aspectual classification seems not to be strictly linked to the type of complement, but to the incrementality of the action described by the verb. Thus, for the analysis of these verbs, we adopt the two-dimensional model of representation of the lexical aspect, in a cognitive-functional approach, proposing that these verbs have aspectual properties that are at the intersection between accomplishments and activities. Our results indicate that the publishing verbs specify a type of directed change. Thus, they have the potential to be (re)construed as telic (accomplishments) or atelic (activities) directed changes, depending on how the speaker intends to describe the publishing event.

Keywords: lexical aspect; publishing verbs; incrementality; telicity; change.

Recebido em 24 de fevereiro de 2022

Aceito em 11 de maio de 2022

1 Introdução

Desde Vendler (1974), a classificação aspectual dos verbos tem sido uma importante tarefa de análise linguística. Os verbos são as palavras por excelência que descrevem os eventos; o aspecto lexical (também chamado de acionalidade) é a categorização da forma como esses eventos se desenrolam no decorrer do tempo. Esse tipo de classificação dos eventos tem se mostrado relevante para a organização gramatical (CANÇADO; AMARAL, 2016; CROFT, 2012; LEVIN; RAPPAPORT HOVAV, 2005).

Vendler (1974) propõe, a princípio, uma classificação dos verbos em quatro classes aspectuais (ou acionais): estados, atividades, *accomplishments* e *achievements*. Porém, o próprio autor e muitos que

o sucederam (DOWTY, 1979; ROTHSTEIN, 2004; SMITH, 1991; VERKUYL, 1989; entre outros) apontaram que há verbos que não apresentam uma classificação estanque, podendo transitar pelas classes de acordo com o contexto. Uma das explicações apresentadas na literatura para esse fenômeno é que o aspecto não é estritamente verbal, mas é derivado da combinação do verbo com seu complemento. Assim, alterando-se o tipo de denotação do complemento (nominais definidos *versus* nominais nus ou plurais, por exemplo), altera-se também o aspecto (WACHOWICZ; FOLTRAN, 2006).

Os verbos de editoração do inglês e do português (*to edit/editar* ou *editar*, *to copy edit/preparar*, *to proofread/revisar*, *to typeset/diagramar* e *to format/formatar*) são um exemplo interessante de um conjunto de verbos que apresentam essa propriedade. Os “verbos de editoração” (ou *publishing verbs*) são assim denominados em Martins (2022) por descreverem o que acontece com o livro (ou outro tipo de material manuscrito, como artigos) dentro de uma editora, no setor de editoração. Segundo o guia do processo de publicação da Bloomsbury¹, a editoração envolve: i) o ato de polir a história; ii) melhorar a coerência e a coesão; iii) verificação de informações, correção de erros gramaticais, esclarecimento de dúvidas com o autor; iv) padronização de destaques e nomes; v) organização de elementos gráficos e aplicação de estilos; e vi) verificação de erros despercebidos e da aplicação dos padrões de estilo.

Esses verbos podem se comportar como atividades e como *accomplishments*, a depender do contexto em que ocorrem. Apesar disso, não há nenhum trabalho de classificação que considere os verbos de editoração, nem no português, nem em inglês (até onde sabemos). Assim, este trabalho se propõe a investigar os verbos de editoração do inglês e do português buscando suprir essa lacuna nos estudos sobre as classes de verbos e, de forma especial, analisar seu comportamento aspectual e entender por que eles podem se distribuir em duas classes distintas². Tendo em vista que os cinco verbos descrevem os procedimentos que acontecem no setor editorial, assumimos, seguindo Martins (2022), que eles fazem parte do mesmo *frame*, ou domínio, sendo eles um grupo semanticamente coerente de itens lexicais.

¹ Disponível em: <https://www.bloomsbury.com/us/discover/bloomsbury-academic/authors/a-guide-to-the-publishing-process/>. Acesso em: 21 fev. 2022.

² Este artigo apresenta uma reanálise dos dados apresentados em Martins (2022), disponibilizados nos Anexos.

Os dados desta pesquisa foram retirados de Martins (2022), que buscou uma lista de verbos de editoração em inglês e em português e suas ocorrências em *sites* diversos, na sua maioria *sites* especializados sobre o mercado editorial (*posts* de editoras, manuais *online*, páginas que oferecem os serviços de edição, preparação, revisão e formatação, etc.). Esse tipo de *corpus* foi selecionado levando-se em conta que estes domínios teriam mais propriedade intelectual para discorrer sobre o tema, além de apresentarem mais exemplos de ocorrências com os verbos em análise. Os dados investigados na presente pesquisa encontram-se dispostos nos Anexos A e B.

Para analisar os verbos de editoração e suas propriedades aspectuais, adotamos o modelo bidimensional de representação do aspecto lexical, proposto em Croft (2012; 2015), numa perspectiva cognitivo-funcional. A partir da análise dentro desse modelo teórico, propomos que esses verbos descrevem mudanças direcionais, possuindo potencial para serem (re)construídos de formas diferentes: a mudança direcional pode atingir um ponto de culminação (caso em que o verbo se comporta como um *accomplishment*) ou a mudança direcional pode indicar o progresso da ação, sem atingir necessariamente um ponto de culminação (caso em que o verbo se comporta como uma atividade). Interessantemente, a presença da culminação não está atrelada estritamente à definitude do complemento, mas à incrementalidade da mudança direcional.

Este artigo está dividido da seguinte maneira: a Seção 2 apresenta o referencial teórico, o modelo bidimensional de representação do aspecto lexical, proposto em Croft (2012; 2015); a Seção 3 traz a análise e a discussão dos dados; e na Seção 4 terminamos com nossas conclusões e considerações finais.

2 Um modelo bidimensional para o aspecto lexical

Vendler (1974) foi pioneiro no estudo do aspecto lexical (também chamado de acionalidade), que tem sido um dos principais campos de análise linguística dos verbos e dos eventos que eles descrevem. O aspecto lexical é uma propriedade semântica que expressa a forma como o evento descrito por um verbo se desenrola no decorrer do tempo. Essa categoria se difere tanto da categoria de tempo, por não ser dêitica, como da categoria de aspecto gramatical, por tratar de características próprias do sentido dos verbos, que não são marcadas morfossintaticamente

(CANÇADO; AMARAL, 2016). Esse tipo de classificação dos eventos tem se mostrado relevante para a organização gramatical (LEVIN; RAPPAPORT HOVAV, 2005; CROFT, 2012; CANÇADO; AMARAL, 2016). No entanto, os complementos dos verbos, e suas características, também se mostram relevantes para o aspecto, principalmente no caso dos *accomplishments*, para diferenciá-los de outras classes aspectuais, como explicaremos em breve.

De acordo com propriedades do aspecto lexical, Vendler (1974) agrupa os verbos em quatro classes: estados, atividades, *achievements* e *accomplishments*. Segundo o autor, há três pares básicos de características que dividem esses verbos, que foram esquematizados por Comrie (1976) (conforme explicam CANÇADO; AMARAL, 2016): estaticidade x dinamicidade; pontualidade x duratividade; telicidade x atelicidade.

A primeira dupla de características, estaticidade e dinamicidade, distingue eventos estativos de eventos dinâmicos. Verbos de característica estativa descrevem eventos que não apresentam fases que progridem temporalmente, descrevem estados que não se alteram ao longo do tempo. Exemplos são *amar*, *ter* e *existir*. Já um verbo que demonstra traços dinâmicos apresenta uma sucessão de intervalos ou fases em que há alteração do estado inicial. Exemplos são *construir*, *conversar* e *estourar*.

O segundo par, pontualidade x duratividade, diferencia um evento pontual de um evento durativo. Um verbo de característica pontual descreve um evento que ocorre instantaneamente, enquanto um verbo de propriedade durativa descreve um evento que se prolonga no decorrer do tempo. Nos exemplos citados anteriormente, o evento descrito por *estourar* é pontual, enquanto eventos descritos por verbos como *construir*, *conversar* e os de qualidade estativa (*amar*, *ter* e *existir*) são durativos.

Já o último par, telicidade x atelicidade, trata do ponto final, do resultado, que um evento pode ou não atingir. Quando o evento descrito por um verbo é télico, como *construir* e *estourar*, ele apresenta um resultado final; se for atélico, como *amar*, *ter*, *existir* e *conversar*, não há naturalmente no evento um resultado final bem definido. O Quadro 1 resume a classificação de Vendler (1974).

Quadro 1 – Classes aspectuais

Classe aspectual	Propriedades	Exemplos
Estados	Estativos, durativos, atélicos	<i>amar, ter, existir</i>
Atividades	Dinâmicos, durativos, atélicos	Conversar
<i>Accomplishments</i>	Dinâmicos, durativos, télicos	Construir
<i>Achievements</i>	Dinâmicos, pontuais, télicos	Estourar

Fonte: adaptado de Vendler (1974).

A proposta de Vendler (1974) tem sido a base para a classificação aspectual dos verbos em diversas análises linguísticas. Porém, apesar de amplamente utilizada e reafirmada, essa proposta não abarca outros tipos aspectuais, encontrados em estudos mais recentes. Alguns exemplos de novas classes aparecem em diferentes trabalhos. Dowty (1979) menciona a existência dos *degree achievements*, como *esquentar*, que se parecem com *achievements*, mas são durativos. Em Smith (1991), encontramos os semelfactivos, como *pular*, que são pontuais, porém atélicos. Ainda, Carlson (1979) indica a existência de diferentes tipos de estados, os permanentes, como *João é brasileiro*, e os transitórios, como *a janela está aberta*.

Além disso, vários autores, desde o próprio Vendler (1974) e especialmente Verkuyl (1989), apontam que o aspecto lexical sofre alterações a depender do tipo de complementos com que o verbo se combina. A telicidade, de forma especial, pode ser impactada pela natureza dos nomes na posição de complemento. Rothstein (2004, p. 92) afirma que os predicados de *accomplishment* geralmente encabeçam um sintagma verbal télico, mas se comportam como predicados de atividade (atélicos) quando o seu complemento é um nominal nu ou massivo (denotando tipos, e não indivíduos). Nesse contexto, mesmo que um verbo como *construir* possa ser *a priori* classificado como um *accomplishment* (por ser télico), ele terá comportamento de atividade quando combinado a certos complementos, como em *construir casas* ou *construir morada* (que tornam o evento atélico).

Essa relação entre verbo e complemento é relevante também na classificação dos *accomplishments* devido à noção de incrementalidade, que deixa claro que os complementos são “usados” aos poucos enquanto o evento expresso pelo verbo progride (ROTHSTEIN, 2004). Sendo assim, o complemento pode determinar a progressão ou até mesmo a duração

do evento. Retomando o nosso exemplo, temos para o verbo *construir* um objeto incremental, pois um objeto só pode ser construído parte por parte; em cada etapa do evento, ocorre a construção de uma parte até a construção do objeto estar completa ao final.

Como mostraremos na Seção 3, os verbos de editoração aqui analisados não se encaixam perfeitamente na classe dos *accomplishments*, à semelhança do caso de *construir*. Porém, o comportamento desses verbos não pode ser explicado puramente a partir do tipo de denotação de seus complementos (tipos ou indivíduos). Por isso, adotamos como referencial teórico o modelo bidimensional para o aspecto lexical, delineado em Croft (2012, 2015). Esse modelo leva em conta dois conceitos fundamentais para a análise desses verbos: a noção de potencial aspectual e a incrementalidade. A essas noções está associada a ideia de “mudança direcional”³, que é capaz de explicar o comportamento de verbos que parecem estar na interseção entre a classe dos *accomplishments* e das atividades.

O modelo bidimensional de Croft (2012, 2015) para o aspecto lexical parte da perspectiva cognitivo-funcional da linguagem. Para o autor, seguindo a perspectiva cognitiva, as expressões linguísticas são produto de um processo cognitivo de conceptualização. A conceptualização de um evento explicita como a realidade é percebida e (re)construída pelos falantes no uso da língua. Seguindo a perspectiva funcionalista, tais conceptualizações, expressas nas estruturas linguísticas, são motivadas por objetivos comunicativos. Isso significa que o aspecto lexical não é inerente às ações no mundo, mas que cada tipo de evento pode ser (re)construído aspectualmente de diferentes formas, de acordo com os objetivos comunicativos. Assim, cada tipo de evento que os falantes vivenciam no mundo tem potencial para ser (re)construído de formas alternativas.

Como exemplo, apontamos os eventos descritos pelo verbo *saber*. O verbo *saber* descreve um tipo de estado mental de um participante animado, que tem um certo conhecimento a respeito de uma situação no mundo. Se o falante quer falar do evento de *saber* como algo estático, durativo, pode-se construir o evento como estativo (e o verbo apresenta propriedades da classe aspectual dos estados):

³ O conceito de mudança direcional é essencialmente o que Beavers (2008) e Rappaport Hovav e Levin (2010) chamam de *scalar change* “mudança escalar”.

- (1) Qualquer dona-de-casa **sabe** [que os bifes não fritam direito se forem colocados um sobre os outros]⁴.

O mesmo verbo, porém, pode também descrever o ponto em que alguém passa a saber de determinada situação (passa do estado mental de não conhecimento para um estado mental de conhecimento de algo). Se o falante quer falar do evento de *saber* como algo dinâmico, pontual, pode-se construir o evento como uma mudança (e o verbo apresenta propriedades da classe aspectual dos *achievements*):

- (2) Aliás, só quando teve conhecimento do fax que lhe foi enviado ontem pelo FBLP é que **soube** [que a primeira reunião deste organismo era tão urgente]⁵.

Dentro dessa perspectiva, podemos afirmar que o verbo *saber* tem potencial de (re)construir o evento descrito como um estado ou como um *achievement*, em combinação com outras expressões linguísticas no enunciado. Sendo uma teoria de base cognitiva, no modelo bidimensional, esses tipos de variações aspectuais são impactadas por construções gramaticais de tempo e aspecto, a exemplo das flexões de presente e pretérito perfeito nos exemplos em (1) e (2), respectivamente.

No modelo bidimensional, como o nome indica, o aspecto é definido em duas dimensões de uma estrutura geométrica: a dimensão do tempo (*t*) e a dimensão dos estados (*q*). O aspecto lexical compreende, então, a forma como os participantes de um evento alteram (ou mantêm) seus estados (*q*) ao longo do tempo (*t*) (funções de intervalos em *t* a intervalos em *q*). Tipos de intervalos nas dimensões *t* e *q* definem fases aspectuais (também chamadas de subeventos); as fases, por sua vez, definem as classes ou os tipos aspectuais. Um ponto em *t* ou em *q* é um intervalo que não contém subintervalos; um intervalo com subintervalos é chamado simplesmente de intervalo. As fases são definidas como:

- Estados: fases definidas como um ponto em *q* e um intervalo em *t*. Ou seja, ao longo de um intervalo de tempo, o estado dos participantes não se altera;

⁴ Exemplo coletado no conjunto de *corpora* disponibilizado pela Linguateca (linguateca.pt/ACDC). Acesso em: 21 fev. 2022.

⁵ Exemplo coletado no conjunto de *corpora* disponibilizado pela Linguateca (linguateca.pt/ACDC). Acesso em: 21 fev. 2022.

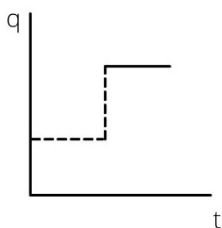
- **Processos:** fases definidas como intervalos em q e em t . Ou seja, ao longo de um intervalo de tempo, o estado dos participantes se altera;
- **Transições:** fases que derivam da combinação de duas outras fases: um ponto resultante (culminação) e um ponto de início (inceptivo) em q , sem intervalos intervenientes (um ponto em t). Ou seja, em um ponto no tempo, o estado dos participantes se altera.

Uma sequência de fases determina o tipo aspectual do evento, composicionalmente. Por sua natureza composicional, Croft (2012) afirma que este modelo consegue explicar todos os tipos aspectuais já descritos na literatura, inclusive aqueles que fogem da classificação tradicional de Vendler (1974). Alguns desses tipos (mais relevantes para a presente pesquisa) são explicitados a seguir:

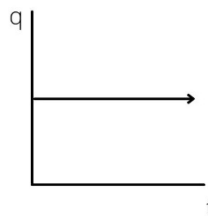
- **Estados:** tipos aspectuais compostos por uma fase de estado (ponto em q e intervalo em t). Diferentes tipos de estados podem ser definidos a partir de transições pressupostas para/de o estado em questão (que são indicadas em linhas pontilhadas). Estados transitórios, como *a porta está aberta*, pressupõem uma mudança de estado em q , indicada em (a) na Imagem 1 em linhas pontilhadas (para chegar ao estado de aberta, a porta passou por uma mudança de estado q). Já os estados permanentes, como *João é brasileiro*, não pressupõem nenhum tipo de mudança, nem inicial e nem final, e são representados, portanto, como um ponto em q e um intervalo contínuo em t (como mostra a estrutura em (b) na Imagem 1).

Imagem 1 – Tipos de estados

a. Estados transitórios



b. Estados permanentes



Fonte: Croft (2012, p. 58).

• *Achievements*: todos os *achievements* incluem uma fase de transição, uma mudança em q que ocorre em um único ponto em t . Mas *achievements* também podem ser de diferentes tipos. Os *achievements* direcionais, como *o balão estourou*, apresentam dois estados diferentes em q como pontos inicial e final, já que o evento pressupõe uma mudança irreversível no participante (em (a) na Imagem 2). Já os *achievements* cíclicos (também conhecidos como semelfactivos), em (b) na Imagem 2, como *o menino tossiu*, não indicam mudanças irreversíveis, por isso as duas fases de estado como pontos inicial e final são idênticas, ou seja, após uma mudança repentina, pressupõe-se que o participante retorna ao estado inicial.

Imagem 2 – Tipos de *achievements*

a. *Achievements* direcionais

b. *Achievements* cíclicos



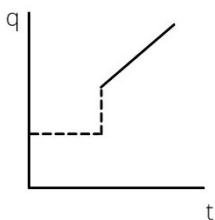
Fonte: Croft (2012, p. 60).

• *Atividades*: atividades se compõem a partir de uma fase de processo, sem transições (ou apenas com transições iniciais pressupostas, indicadas nas linhas pontilhadas). Atividades direcionais incluem a classe dos *degree achievements*, como *a água esquentou*, e atividades como *a babá empurrou o carrinho*. As atividades direcionais são aquelas que progridem em q , alterando os estados do participante ao longo do processo, sem indicar um estado final ou ponto de culminação (função monotônica de um intervalo em t para um intervalo em q); são representadas em (a) na Imagem 3. Atividades não direcionais (também chamadas de cíclicas), como *os meninos dançaram a noite toda* e *as crianças brincavam*, por exemplo, são mudanças que alteram o estado q do participante sem uma progressão (função não monotônica de um intervalo em t para um intervalo em

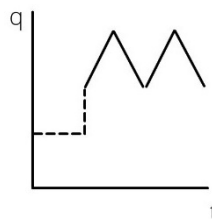
q), e também sem um ponto de culminação; são representadas em (b) na Imagem 3.

Imagem 3 – Tipos de atividades

a. Atividades direcionais



b. Atividades não direcionais



Fonte: Croft (2012, p. 61).

• *Accomplishments*: um *accomplishment* compreende uma fase de processo e suas transições – inicial e final. Tipos diferentes de *accomplishments* são definidos pelo tipo de fase de processo: direcional ou não direcional, correspondendo aos tipos de atividade. É exatamente nesse ponto que podemos encontrar a interseção entre atividades e *accomplishments*. *Accomplishments* incrementais, como *construir uma casa*, são compostos por uma fase de processo de mudança direcional (funções monotônicas de t para q) e uma transição, como visto em (a) na Imagem 4; *accomplishments* não incrementais, como *consertar um computador*, são compostos por uma fase de processo de mudança não direcional (funções não monotônicas de t para q) e uma transição, como em (b) na Imagem 4. Estados iniciais e finais pressupostos são indicados nas linhas pontilhadas. As transições correspondem às culminações dos eventos, ou ao ponto télico.

Imagem 4 – Tipos de *accomplishments*

a. *Accomplishments* incrementais

b. *accomplishments* não incrementais



Fonte: Croft (2012, p. 62).

A transição para um estado final, a culminação ou ponto télico, nos *accomplishments* incrementais está ligada à extensão do objeto ligado ao verbo, pois só acontece quando este é “usado” completamente, de forma especial, pelos participantes que “medem” o evento, os temas incrementais. No modelo bidimensional, temas incrementais são participantes em um evento de *accomplishment* que passam por mudanças direcionais em q , durante um intervalo t e que são delimitados em t por uma fase de transição final.

A partir da definição dos tipos aspectuais, portanto, é possível observar como o modelo bidimensional define formalmente as noções de direcionalidade e incrementalidade. Nas estruturas em (a) das Imagens de 2 a 4, que representam eventos direcionais, existem dois pontos distintos em q , que simbolizam estados iniciais e finais diferentes do evento; nas estruturas em (b) nas Imagens de 2 a 4, que representam eventos não direcionais, apesar de haver mudanças em q , há um retorno aos estados iniciais dos eventos. Ou seja, há mudanças, mas não há uma mudança em uma direção (monotônica) em q , apenas mudanças que retornam ao seu estado de base (não monotônicas). Portanto, a direcionalidade é definida como uma mudança de estados em q , de forma que o estado inicial e o estado final sejam diferentes.

Retomando alguns dos exemplos já mencionados, atividades direcionais, como *a água esquentou* e *a babá empurrou o carrinho*, não possuem pontos de culminação em q . Porém, as ações são direcionais em q , no sentido de que o estado dos participantes se altera progressivamente ao longo do tempo (mesmo que não haja uma temperatura específica final, a água muda de estado; mesmo que não haja uma meta final do movimento, o carrinho se desloca). Contrastivamente, em eventos não

direcionais, como os descritos por verbos como *brincar*, no caso das atividades, e *consertar*, no caso dos *accomplishments*, as mudanças que ocorrem não progridem em q , ou seja, as alterações que ocorrem ao longo do evento causam mudanças nos participantes, mas os participantes retornam aos estados de base. A brincadeira das crianças é composta por uma série de mudanças dinâmicas, mas não há um progresso ou mudança de estado final. No caso de *consertar* (um exemplo traduzido de CROFT, 2012, que é originalmente de ROTHSTEIN, 2004), Rothstein (2004) afirma que consertar algo não envolve uma afetação gradual parte por parte, mas uma série de diferentes ações não diretamente relacionadas, que culmina em um estado final (de concerto). Com base nessa ideia, Croft (2012, p. 64) explica que “a dimensão q para consertar um computador não consiste de uma escala que representa um progresso mensurável e contínuo em direção ao estado resultante”⁶.

Ainda, a partir da direcionalidade, podemos definir o que, nesse modelo, se entende por incrementalidade. Temas incrementais são participantes em um evento de *accomplishment* que passam por mudanças direcionais em q , durante um intervalo t e que são delimitados em t por uma fase de transição final. A incrementalidade, portanto, é definida a partir do conceito de direcionalidade: só é incremental o evento que é direcional, ou seja, que possui alterações direcionais em q . Ademais, o evento incremental deve ser durativo, ocorrer ao longo de um intervalo de tempo em t , e deve também ser télico, pois é delimitado em t por uma fase de transição. Tal definição também exclui da noção de incrementalidade as mudanças direcionais pontuais, como os *achievements* direcionais (já que a incrementalidade é definida como uma função monotônica de um intervalo em t para um intervalo em q).

A partir dessa caracterização, é possível observar com clareza os pontos de interseção entre as atividades e os *accomplishments*, o que explica o comportamento dúbio de alguns verbos, que podem ser classificados em um tipo e em outro. De forma similar, Rothstein (2004, p. 21) já afirmava que “um *accomplishment* é uma atividade que progride em direção a um ‘ponto final’ ou ‘um ponto de término’ ou ‘culminação’

⁶ Do original: “The q dimension for repairing a computer does not consist of a scale representing continuous measurable progress toward the result state” (tradução nossa).

ou ‘ponto télico’⁷, ou seja, é uma situação que apresenta as características de uma atividade, com o acréscimo das fases de transição. De acordo com as propriedades de Comrie (1976), apenas a telicidade distingue atividades de *accomplishments*. No modelo bidimensional, a telicidade é representada pela transição final de um evento de *accomplishment*. O ponto comum entre essas classes no modelo bidimensional, portanto, é a fase de processo, sendo a fase de transição final o ponto de divergência entre elas.

Como já apontamos, por ser uma teoria de base cognitiva, o modelo bidimensional assume que o aspecto é uma conceptualização e que os verbos possuem diferentes potenciais de (re)construção de seu tipo aspectual. Assim, sendo atividades e *accomplishments* tão intimamente relacionados, elementos externos ao verbo e diferentes construções de tempo e aspecto podem acrescentar, retirar, realçar ou desfocar fases (especialmente as de transição, que indicam a telicidade), fazendo com que certos verbos se encontrem no limiar entre um tipo aspectual e outro. A partir desse modelo teórico, apresentaremos na próxima seção uma análise dos verbos de editoração em inglês e em português, mostrando que a noção de mudança direcional é a mais relevante para a caracterização aspectual desses verbos.

3 Uma análise aspectual dos verbos de editoração em inglês e em português

Ao apresentarmos uma análise aspectual dos verbos de editoração do inglês e do português – *to edit/editar* ou *editar*⁸, *to copy edit/preparar*, *to proofread/revisar*, *to typeset/diagramar* e *to format/formatar* –, partimos da hipótese central de que as características das ações complexas realizadas no contexto editorial, e descritas por esses verbos, são importantes para a caracterização aspectual. Ainda, como esses verbos se encontram no limiar entre os tipos aspectuais *accomplishments* e atividades, propomos que o mais relevante para sua

⁷ Do original: “an accomplishment is an activity which moves toward a finishing point, or ‘set terminal point’, or ‘culmination’ or ‘telic point’” (tradução nossa).

⁸ Como em inglês encontramos apenas um correspondente (*to edit*) para os verbos *editar* e *editar*, nossa análise para o português será baseada apenas no verbo *editar*. O verbo *editar* possui o mesmo comportamento que o verbo *editar*.

caracterização aspectual é a ideia de mudança direcional (que pode ou não ter um ponto télico).

Os dados dos verbos de editoração aqui investigados foram retirados de Martins (2022). A autora coletou tais dados em sites específicos do mundo editorial, como mencionado na Introdução, desde editoras reconhecidas mundialmente até artigos cujo objetivo era explicar e diferenciar os cinco verbos. Assim, por pertencerem ao mesmo meio, consideramos não apenas o domínio intelectual de especialistas na área, como também a ocorrência dos verbos dentro de um contexto específico, relacionado à editoração de livros. Após a coleta dos dados, dispomos as ocorrências em quadros separados, um para cada verbo, e os analisamos sintática e semanticamente, conforme Martins (2022). A lista completa de verbos e enunciados analisados, com suas respectivas fontes, encontra-se em Martins (2022) e também nos Anexos A e B deste artigo.

A partir das análises, os verbos de editoração se apresentam nos dados como verbos transitivos diretos cuja estrutura argumental revela um sujeito que recebe o papel temático de agente, e um objeto direto que pode ser, *grosso modo*, definido como um paciente. Então, classificá-los como verbos de *accomplishment*, considerando a definição desse tipo aspectual, se adequa, em uma primeira análise, aos dados que encontramos e analisamos neste trabalho. A análise dos dados evidencia que os cinco verbos, em inglês e em português, denotam processos realizados por alguém ao longo de um determinado período de tempo, com a culminação sendo o livro, ou manuscrito, que passa a estar em determinado estado.

(3) We edited the novel [...].

‘Nós editamos o romance.’

(4) I fact checked and copy edited the text [...].

‘Eu chequei os fatos e preparei o texto [...]’.

(5) We proofread the book [...].

‘ Nós revisamos o livro [...]’.

(6) We typeset the book, formatted the book, designed the cover and distributed the book in all formats globally.

‘Nós diagramamos o livro, formatamos o livro, desenhamos a capa e distribuimos o livro em todos os formatos globalmente.’

(7) I just formatted the e-book edition of my short story collection, *Twenty Miles West of Branch, Texas and other stories*.

‘Eu acabei de formatar a edição e-book da minha coleção de contos, *Twenty Miles West of Branch, Texas and other stories*.’

(8) Professora Rose Mary editou livro sobre semiologia cardiovascular.

(9) A autora está revisando o livro.

(10) Ao questionar os/as preparadores/as se eles/as conhecem as línguas dos originais dos textos traduzidos que preparam e se acreditam que isso seja importante [...].

(11) Ele escreveu, editou, diagramou e até ilustrou a capa do volume.

(12) [...] a coordenadora Raiza Tonon formatou o livro e devolveu as [*sic*] crianças para que ilustrassem.

Nos exemplos de (3) a (12), temos a mesma estrutura argumental para todas as ocorrências dos verbos. A única distinção entre os cinco verbos é o significado específico de cada um. *To edit/editar* é a primeira etapa pela qual o livro passa no setor de editoração de uma editora. O responsável pode ser o editor, o coordenador de textos, enfim, há vários títulos para quem realiza o mesmo trabalho: a edição do manuscrito. Editar significa adequar o livro, torná-lo coeso de modo que a história faça sentido internamente, além de também trabalhar ao lado do autor a fim de que o resultado final seja o que foi originalmente pensado. Então, quando se diz *We edited the novel*, como em (3), também se está dizendo que o romance já passou por modificações textuais, ainda que superficiais. Algo importante de salientar é que o editor é o responsável pelo andamento do livro dentro da editora, é ele, ou ela, quem tem a palavra final (desde que em acordo com o autor). E “editar” um livro abarca todo o processo editorial, desde a seleção do manuscrito até a criação da capa. Para quem conhece a produção de livros, a edição só acaba no momento que a última prova é encaminhada para a gráfica – e

isso acontece depois da liberação do editor, que conta com o trabalho minucioso do preparador e do revisor.

To copy edit/preparar é o processo seguinte após o primeiro trabalho do editor. Nele, realiza-se a leitura detalhada e cuidadosa do manuscrito, a procura de erros gramaticais e ortográficos, incoerências que o editor não encontrou, desorganização das ideias, até mesmo verificação dos fatos utilizados pelo autor, como acontece na sentença em (4). Uma boa preparação de texto leva, no mínimo, quatro semanas para ser realizada, a depender do tema, da quantidade de páginas, da qualidade textual, de como alguns itens serão padronizados (por exemplo, termos em língua estrangeira devem aparecer entre itálico ou entre aspas? Em quais casos o negrito deve ser utilizado? Na maioria dos casos, cada editora tem seu manual de padronização e é nele que o preparador se baseia) e de quantos precisarão de padronização. Após o fim da preparação textual, o manuscrito é enviado ao autor para conferência das sugestões e das dúvidas, e só chega ao diagramador após tudo estar resolvido.

O processo que sucede a preparação de texto é a diagramação, no entanto, como o livro é finalizado pelo formatador, seguindo para a gráfica depois disso, decidimos deixar *to typeset/diagramar* e *to format/formatar* por último. Sendo assim, falaremos mais detalhadamente agora de *to proofread/revisar*.

Como os maiores problemas textuais são resolvidos pelo preparador, sobra para o revisor de provas encontrar o que passou despercebido, pequenos detalhes (como conjugação errada, ausência de itálico em termos que deveriam estar assim demarcados, pontuação, letras sobrando ou faltando, hifenização, etc.). A revisão é um processo menos demorado em comparação à preparação, justamente pelo fato de o texto estar “limpo” e mais organizado. Outra diferença entre os processos é que, enquanto a preparação é normalmente feita em formato digital, a revisão de provas é feita no papel, através de marcações nas margens com sinais convencionais do meio. Após a primeira leitura, o livro é encaminhado ao diagramador, que realiza as emendas solicitadas, retorna ao revisor que confere as alterações e pode retornar ao diagramador caso algo tenha ficado para trás – esse vaivém pode demorar, a depender de vários fatores. Por último, se houver dúvidas, elas são enviadas para o autor, voltam para o revisor que, novamente, faz as emendas e manda para o diagramador que, após realizá-las, mostra a provável última prova para verificação. O livro só é liberado para a gráfica quando não houver

mais nenhuma emenda a ser realizada, e outro revisor já tiver verificado se não falta nada no livro.

To typeset/diagramar significa organizar o manuscrito visualmente. Essa é a etapa da produção em que o designer gráfico irá transformar um texto em um livro, de acordo com a coleção em que foi incluído pelo editor. Cada coleção apresenta um design diferente, seja a fonte da capa, do miolo (a parte escrita do livro), até os subtítulos, a capa e a lombada (a parte lateral do livro). Nesta fase, o foco principal é a estética do livro que envolve a aplicação de estilos, a verificação do espaço entre as fontes, se tem apenas uma palavra na última linha da página ou na primeira linha, a hifenização, as notas de rodapé, os recuos, os itálicos, as imagens, se houver. É o diagramador quem irá dispor o texto nas páginas, controlar o tamanho das letras e a qualidade das imagens para que o resultado seja o melhor possível na hora da impressão. Na diagramação também acontece a formatação. É o diagramador quem irá formatar (*to format/formatar*) o texto seguindo as indicações do preparador textual e do manual da editora. Como mencionado anteriormente, o livro pode voltar para o diagramador quantas vezes forem necessárias até que todas as emendas estejam resolvidas.

A partir da descrição semântica dos eventos analisados, por meio das especificações dos processos descritos por cada um desses verbos, fica evidente que *to edit/editar*, *to copy edit/preparar*, *to proofread/revisar*, *to typeset/diagramar* e *to format/formatar* apresentam múltiplas fases (ou múltiplos subeventos), já que há etapas distintas para cada um deles, além de um produto final do trabalho. Eventos que possuem mais de uma fase são caracterizados no modelo bidimensional como *accomplishments*, assim como eventos que possuem subeventos são também classificados como *accomplishments* em outras perspectivas teóricas (CANÇADO; AMARAL, 2016). Na proposta teórica do modelo bidimensional (CROFT, 2012; 2015), apresentada na Seção 2, os *accomplishments* são os únicos tipos aspectuais compostos por mais de uma fase (uma fase de processo e uma fase de transição). Isso nos permite afirmar que os verbos de editoração são télicos, outra característica dos *accomplishments*, além de também evidenciarmos a duratividade de todos os cinco verbos analisados.

Para comprovar essa classificação aspectual, aplicamos dois testes para aspecto lexical encontrados na literatura: o paradoxo do imperfectivo e a ambiguidade com *almost/quase* (DOWTY, 1979). Em (13)-(22), o teste do paradoxo do imperfectivo evidencia que as sentenças em (a), com

o verbo auxiliar *estar* mais a forma do verbo de editoração no gerúndio, não acarretam o sentido de evento finalizado descrito em (b), mas sim a ação em andamento, que foi interrompida por tempo indeterminado⁹:

- (13) a. *We were editing* the novel.
'Nós *estávamos editando* o romance.'
b. *We edited* the novel.
'Nós *editamos* o romance.'
- (14) a. I *was* fact checking and *copy editing* the text.
'Eu *estava* checando os fatos e *preparando* o texto.'
b. I fact checked and *copy edited* the text [...].
'Eu chequei os fatos e *preparei* o texto.'
- (15) a. *We were proofreading* the book [...].
'Nós *estávamos revisando* o livro [...].'
b. *We proofread* the book [...].
'Nós *revisamos* o livro [...].'
- (16) a. *We were typesetting* the book.
'Nós *estávamos diagramando* o livro.'
b. *We typeset* the book [...].
'Nós *diagramamos* o livro [...].'
- (17) a. I *was* just *formatting* the e-book edition of my short story collection.
'Eu *estava formatando* a edição e-book da minha coleção de contos.'
b. I just *formatted* the e-book edition of my short story collection [...].

⁹ As sentenças utilizadas para os testes aspectuais na análise são dados de introspecção, construídos por nós com base nos dados de *corpora* de Martins (2022), disponibilizados nos Anexos.

‘Eu *formatei* a edição e-book da minha coleção de contos [...].’

- (18) a. A Professora Rose Mary *estava editando* o livro sobre semiologia cardiovascular.
 b. A professora Rose Mary *editou* livro sobre semiologia cardiovascular.
- (19) a. A autora *estava revisando* o livro.
 b. A autora *revisou* o livro.
- (20) a. Ao questionar os/as preparadores/as se eles/as conhecem as línguas dos originais dos textos traduzidos que *estavam preparando* e se acreditam que isso seja importante [...].
 b. Ao questionar os/as preparadores/as se eles/as conhecem as línguas dos originais dos textos traduzidos que *prepararam* e se acreditam que isso seja importante [...].
- (21) a. Ele *estava* escrevendo, editando, *diagramando* e até *ilustrando a capa do volume*.
 b. Ele escreveu, editou, *diagramou* e até *ilustrou a capa do volume*.
- (22) a. [...] a coordenadora Raiza Tonon *estava formatando* o livro e devolveu as [*sic*] crianças para que ilustrassem.
 b. Depois de escrito a coordenadora Raiza Tonon *formatou* o livro e devolveu as [*sic*] crianças para que ilustrassem.

Esse é o resultado esperado no teste do paradoxo do imperfectivo para um verbo télico, já que esses verbos, quando usados no progressivo, não indicam a realização completa do evento¹⁰. Para verbos atélicos, em qualquer ponto após o início do evento há o acarretamento de ação finalizada (já que não existe uma culminação esperada). Por exemplo, para o verbo de atividade *brincar*, temos que se *as crianças estavam brincando*, podemos afirmar que *as crianças brincaram*. Para os verbos

¹⁰ Na intuição de um dos pareceristas “Paulo estava revisando o texto” acarreta que “Paulo revisou o texto”. Apesar de essa intuição diferir da nossa, vemos essa possibilidade de interpretação como mais uma evidência do comportamento dúbio desses verbos, entre *accomplishments* e atividades.

de editoração em (13)-(22), o que podemos afirmar é que parte do material textual foi editado, preparado, revisado, diagramado e formatado, mas não o material completo, até a última página.

Em (23)-(24), o teste com *almost/quase* reforça a decomposição das ações descritas por verbos de editoração em fases ou subeventos. O acréscimo de *almost/quase* em sentenças com verbos de editoração gera ambiguidades de escopo. Essa ambiguidade pode recair em qualquer um dos subeventos, significando que 1) o agente tinha a intenção de iniciar a ação, mas não o fez ou 2) o agente iniciou a ação, mas não atingiu o ponto de culminação esperado. Como mostram as paráfrases em (b) e (c), o sentido das sentenças em (a) é ambíguo, em relação ao escopo de *almost/quase*: o agente pode ter iniciado a ação e não a ter terminado ou nem ter começado a ação:

(23) a. I *almost* edited / copy edited / proofread / typeset / formatted the manuscript [...].

‘Eu quase editei / preparei / revisei / diamei / formatei o manuscrito.’

b. What I *almost* did was edit / copy edit / proofread / typeset / format the manuscript [...].

‘O que eu quase fiz foi editar / preparar / revisar / diagramar / formatar o manuscrito.’

c. What I did was *almost* edit / copy edit / proofread / typeset / format the manuscript [...].

‘O que eu fiz foi quase editar / preparar / revisar / diagramar / formatar o manuscrito.’

(24) a. A Professora Rose Mary *quase* editou / preparou / revisou / diagramou / formatou o livro.

b. O que a Professora Rose Mary *quase* fez foi editar / preparar / revisar / diagramar / formatar o livro.

c. O que a Professora Rose Mary fez foi *quase* editar / preparar / revisar / diagramar / formatar o livro.

As diferentes possibilidades de escopo do advérbio indicam as diferentes fases ou subeventos sobre os quais ele pode operar. Novamente, verbos atélicos, que não apresentam subeventos (ou que apresentam subeventos idênticos), também não apresentam essa ambiguidade.

Retomando o verbo *brincar*, se *as crianças quase brincaram*, elas simplesmente não iniciaram a ação.

A partir dessa análise, podemos afirmar que os verbos de editoração são classificados aspectualmente como *accomplishments*, ou seja, eventos dinâmicos, durativos e télicos. No modelo bidimensional, podemos caracterizá-los como *accomplishments* incrementais, considerando que apresentam uma fase de processo que é uma mudança direcional, além de uma fase de transição final.

Em outros contextos, no entanto, esses verbos parecem se comportar como atélicos. Verificamos em nossos dados que, ao aplicar um terceiro teste proposto na literatura, a inserção de expressões temporais como *for a time/por a tempo* (DOWTY, 1979), os verbos demonstram características esperadas dos de atividade. De acordo com Wachowicz e Foltran (2006), os verbos télicos se combinam com expressões temporais como *in a time/em a tempo*, mas não com expressões do tipo *for a time/por a tempo*. Segundo Rothstein (2004, p. 24), a expressão adverbial *for a time/por a tempo* denota blocos de intervalos em que “x P-ou por duas horas é verdade se a cada subintervalo de um período de duas horas, x P-ou for verdade”¹¹. Por isso, não seria de se esperar que um verbo de *accomplishment* aceitasse esse tipo de modificação, uma vez que as ações contidas em cada subintervalo são, a princípio, distintas entre si.

- (25) a. Polly Kummel skillfully edited / copy edited / proofread / typeset / formatted the text *for four days* [...].

‘Polly Kummel habilmente editou / preparou / revisou / diagramou / formatou o texto *por quatro dias*.’

- b. Polly Kummel skillfully edited / copy edited / proofread / typeset / formatted the text *for four days before her vacation*.

‘Polly Kummel habilmente editou / preparou / revisou / diagramou / formatou o texto *por quatro dias antes das suas férias*.’

- (26) a. Luís Garnier [...] editou/ revisou / preparou / formatou / diagramou mais de seiscentos autores brasileiros *por vinte anos*.

¹¹ Do original: “x P-ed for two hours is true if at every subinterval of a two-hour period, x P-ed is true” (tradução nossa).

b. Luís Garnier [...] editou / revisou / preparou / formatou / diagramou mais de seiscentos autores brasileiros *por cinco meses antes de ir para Paris*.

Dowty (1979) já reconhecia que verbos de *accomplishment* que denotam mudança de estado podem aceitar essas expressões de duração, mas o autor afirma que a duração descrita por tais expressões terá escopo sobre o resultado final do evento, não sobre o evento em si. Porém, observando os exemplos em (25)-(26), veremos que os enunciados não indicam que as ações foram concluídas. Em todos os exemplos, a interpretação é de que a expressão temporal está ligada à ação como um todo, não ao resultado final. A expressão temporal demonstra que nenhuma das ações chegou à culminação. Sendo assim, a duração expressa no modificador não pode estar relacionada à duração do estado resultante. Segundo Basso (2011), o que ocorre na combinação de *por a tempo* com eventos télicos é um processo de “detelicização”, indicando que o ponto de culminação do evento não foi atingido.

Além da ocorrência com expressões do tipo *for a time/por a tempo*, o próprio teste do paradoxo do imperfectivo indica nesses verbos algumas propriedades de atividades. Se considerarmos apenas parte de um objeto textual, haverá acarretamento da culminação expressa pelo pretérito perfeito, conforme os exemplos em (27) e (28).

(27) a. [...] it's the publisher that was formatting *part of the book* for the author.

‘[...] é a editora que estava formatando *parte do livro* para o autor.’

b. [...] it's the publisher that formatted *part of the book* for the author.

‘[...] é a editora que formatou *parte do livro* para o autor.’

(28) a. [...] um agradecimento especial ao Rodrigo Franco, que estava revisando e formatando *parte desse playbook*.

b. [...] um agradecimento especial ao Rodrigo Franco, que revisou e diagramou *parte desse playbook* [...].

Por fim, conforme apontam Levin (1993) e Rappaport Hovav e Levin (1998), verbos de atividade aceitam a construção de sentenças sem a expressão do seu complemento (como *ela varria o chão/ela varria*),

contrariamente a verbos típicos de *accomplishment* (como *ela quebrou o jarro*/**ela quebrou*). Como é possível observar nos enunciados em (29) e (30), os verbos de editoração podem ocorrer em construções intransitivas¹², sem a expressão de seu complemento.

(29) a. Authors themselves rarely are trained how to typeset [...].

‘Os próprios autores raramente são treinados para diagramar.’

b. There are some techniques you can use to proofread efficiently and effectively [...].

‘Existem algumas técnicas que você pode usar para revisar eficientemente e efetivamente.’

c. When they copy edit, these individuals provide an objective set of eyes that ensures grammar and spelling are correct [...].

‘Quando eles preparam, esses indivíduos proporcionam um olhar objetivo que garante que gramática e ortografia estão corretos [...].’

d. I know some writers who edit as they go [...].

‘Eu conheço alguns autores que editam enquanto escrevem [...].’

e. Don’t locally format.

‘Não formate localmente.’

¹² A ocorrência desses verbos em construções de estrutura argumental intransitivas parece ser mais produtiva em inglês que em português, mas ocorre em ambas as línguas. Esse fenômeno precisa ainda ser melhor investigado e não será abordado em detalhes neste trabalho.

- (30) a. Estava editando quando 28% do Pantanal foi embora. [...].
- b. Quem revisa, portanto, só tem a ganhar, ainda mais quando se trata de textos acadêmicos para defesa ou concursos [...].
- c. Olhando assim, pode parecer difícil pensar que existam técnicas e dicas para diagramar como um profissional, certo?

Partindo dessas observações, é necessário olharmos com mais atenção os objetos diretos dos verbos de editoração. *Book/livro, manuscript/manuscrito, text/texto* são materiais que apresentam início, meio e fim. Por causa disso, as ações que os envolvem no processo editorial também apresentam intervalos organizados temporalmente. Um revisor não pode (pode, mas não deve) começar a revisar um livro a partir do quarto capítulo, por exemplo, porque começará do desenvolvimento da narração, quando o estilo de escrita do autor já é perceptível (e, portanto, não mais estranho aos olhos do revisor) e o tema já está bem estabelecido aos olhos não acostumados do revisor. Por isso, podemos considerar os objetos dos verbos de editoração como incrementais já que são afetados ao longo da ação, iniciada a partir da primeira página e concluída apenas ao final da edição, preparação, revisão, etc., do livro, na última página.

Segundo Rothstein (2004, p. 94), a noção de incrementalidade é essencial para definir os verbos de *accomplishment*, portanto a relação entre verbo e seu complemento pode ser usada para explicar como a incrementalidade funciona. Considerando o que foi dito no parágrafo acima, o complemento *book/livro* ou *manuscript/manuscrito* é afetado aos poucos pelas ações de editar, revisar, diagramar, etc. Então, esses argumentos são modificados enquanto houver páginas a serem editoradas. Dessa forma, a extensão e a progressão do evento podem ser mensuradas olhando o que acontece aos materiais textuais.

Retomando os exemplos em (27) e (28), argumentamos que, mesmo que as ações descritas por esses verbos tenham um progresso e uma culminação esperada, os subeventos que compõem essas ações podem ser caracterizados como realizados sobre partes dos objetos textuais (certo número de páginas, por exemplo). Assim, quando verbos de editoração são combinados a expressões do tipo *for a time/por a tempo*, a interpretação resultante é de que várias etapas do processo foram realizadas por um período de tempo α , mas a culminação final daquela etapa da editoração não foi concluída. Podemos imaginar que se o editor preparou o livro por dois meses, ele passou por várias páginas

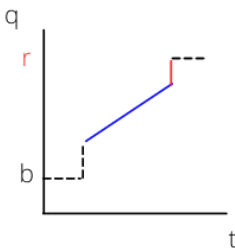
do texto, mas não chegou ainda à página final. O mesmo ocorre no caso das sentenças no progressivo. Mesmo que o evento esteja ainda em andamento, como caracteriza o progressivo, a ação pode ser vista como concluída se esta for descrita como realizada sobre partes dos objetos textuais, como a revisão de um capítulo apenas.

Portanto, as características do objeto incremental estão diretamente relacionadas ao comportamento de atividade desses verbos. A partir do modelo bidimensional do aspecto lexical (CROFT, 2012; 2015), analisamos essa incrementalidade como associada a mudanças direcionais (uma mudança incremental é uma mudança direcional durativa e tética). Isso significa que, mesmo que as ações descritas possam ser durativas e atélicas, elas progredem no sentido de que o objeto textual é editado, formatado, preparado, revisado e diagramado e muda suas propriedades (pelo menos em parte), ou seja, são atividades direcionais. O ponto de interseção entre a interpretação de *accomplishment* incremental e a de atividade direcional é a fase de processo, que é, no caso desses verbos, uma mudança direcional. A evidência da direcionalidade dessa mudança é que tais objetos, sofrendo alterações ao longo dos processos editoriais, possuem diferentes nomenclaturas para cada etapa: *manuscrito*, *original*, *prova*, *livro*, etc. A direcionalidade, porém, não impede que o processo seja repetido (o que de fato ocorre frequentemente nas editoras), nem garante que uma culminação seja atingida, mas indica que os objetos incrementais sofrem alterações em seu estado ao longo do tempo.

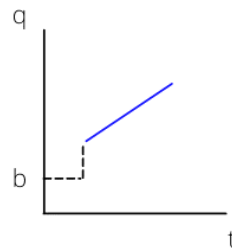
Partindo dessa ideia, propomos representar o aspecto lexical dos verbos de editoração a partir do que se vê na Imagem 5, conforme o modelo bidimensional.

Imagem 5 – O aspecto lexical dos verbos de editoração segundo o modelo bidimensional

a. Evento de editoração tético



b. Evento de editoração atélico



Fonte: elaborado pelas autoras, com base em Croft (2012).

A estrutura em (a) na Imagem 5 representa o evento télico (do tipo *accomplishment* incremental), que é composto por uma fase de processo (em azul) e por uma transição final (em vermelho). As linhas pontilhadas indicam os estados inicial e final do objeto. O ponto *b* é o estado de base em *q*, o estado inicial, o manuscrito da forma como foi entregue pelo autor (que poderíamos chamar de *original*). O ponto *r* é o estado resultante, o manuscrito na etapa final do processo descrito pelo verbo (que poderíamos chamar de *livro*, para o processo de edição como um todo). Cada ponto em *q* entre *b* e *r* é uma etapa distinta do processo e afeta uma parte do tema incremental. A estrutura em (b) na Imagem 5 representa o evento atélico (do tipo atividade direcional), que é composto por uma fase de processo (em azul) apenas. As linhas pontilhadas indicam o estado inicial do objeto, representado como *b*. A estrutura explicita a ideia de progressão da ação atélica. Mesmo sem um estado *r* definido, o objeto textual muda de estado em *q* ao longo do tempo. Apesar de todos os verbos de editoração apresentarem a mesma estrutura aspectual, para cada verbo os estados de *q* diferem. Ou seja, para cada tipo de processo editorial, as mudanças que ocorrem no material textual são distintas e é isso que determina as diferenças semânticas entre os verbos, que explicamos com detalhes no início desta seção.

Portanto, o que podemos afirmar a respeito do aspecto lexical dos verbos de editoração é que eles descrevem mudanças direcionais, com o potencial de serem télicas ou não, dependendo da forma como os falantes (re)constróem os eventos para seus objetivos comunicativos. A vantagem do modelo bidimensional na análise desses verbos é que, nessa teoria, atividades podem ser direcionais, mesmo não sendo télicas. Ou seja, a atividade de editar um livro, por exemplo, pode ter etapas e um progresso, mesmo que o ponto final de culminação não seja atingido ou não seja expresso. Por isso, esses verbos apresentam características de atividades e de *accomplishments*, conforme a classificação tradicional. É importante salientar que esse comportamento dos verbos de editoração ocorre até mesmo em casos em que o objeto direto é um sintagma nominal de referência singular definida (cf. exemplos em (25)-(26)). Isso indica que a característica aspectual decorre da incrementalidade da ação (da mudança direcional, mais especificamente), e não de outras propriedades dos complementos, externas ao verbo, como o tipo de denotação e a quantificação (nomes nus, plurais, etc.).

Em contextos em que os verbos de editoração se combinam com complementos que denotam tipos, como nominais nus ou plurais, eles seguem o padrão dos verbos de *accomplishment*, se comportando como atividades, com VPs atélicos. Rothstein (2004) afirma que os verbos de *accomplishment* se comportam como verbos de atividade quando o objeto direto é um nominal nu ou massivo (com denotação de tipo, não de indivíduo). Segundo Rothstein (2004, p. 149), se “o objeto direto é de massa ou contém um plural simples, então o VP é atélico; no entanto, se for um substantivo singular contável (seja definido ou indefinido), ou um nominal com um determinante numérico, então é télico”¹³.

(31) He has edited and written *columns* for the opinion pages of The Indian Express [...].

‘Ele editou e escreveu *colunas* para as páginas de opinião do The Indian Express.’

(32) I have copy edited and proofread: *fiction, creative non-fiction* [...].

‘Eu preparei e revisei: *ficção e não ficção criativa*.’

(33) Sorrel Packham has typeset *a variety of publications* for Walker Books [...].

‘Sorrel Packham diagramou *uma variedade de publicações* para Walker Books.’

(34) [...] ele aceitou editar *os livros* por sua própria conta [...].

(35) Eles tiveram que revisar *artigos* de vários tipos [...].

(36) Também diagramamos *e-books* para estratégias de inbound marketing.

¹³ Do original: “the direct object is mass or contains a bare plural, then the VP is atelic; however, if it is a singular count noun (either definite or indefinite), or a nominal with a numerical determiner, then it is telic” (tradução nossa).

Concluimos, portanto, que os verbos de editoração estabelecem um ponto de interseção entre os tipos aspectuais atividades e *accomplishments*, que pode ser definido no modelo bidimensional como uma fase de processo direcional, a mudança direcional. A telicidade, que diferencia esses tipos aspectuais, pode ou não estar presente no evento e é representada no modelo bidimensional como uma transição para um estado final. Nossa hipótese inicial era de que características das ações realizadas no domínio da editoração são responsáveis por tais propriedades aspectuais. Corroboramos essa hipótese por meio da análise das fases dos eventos de editoração e da natureza incremental do complemento desses verbos (os materiais textuais). Ademais, na nossa análise comparativa, observamos que os verbos de editoração do inglês e do português se comportam da mesma forma em relação ao aspecto lexical.

4 Considerações finais

Verbos com comportamentos aspectuais múltiplos já são bastante conhecidos. Até mesmo Vendler (1974), em sua proposta original, já reconhecia que certos verbos poderiam ter múltiplas classificações aspectuais. Neste artigo, nossa proposta foi abordar esse problema a partir da análise de 5 verbos específicos em português e inglês (os verbos de editoração), que podem ser classificados como *accomplishments* ou como atividades, dependendo do contexto. Esses verbos, apesar de apresentarem essa dupla classificação aspectual, não aparecem em análises anteriores. Portanto, a contribuição desta pesquisa consiste em apresentar uma descrição e uma análise aspectual dos verbos que descrevem as ações do universo da editoração.

Ainda, a partir do modelo bidimensional para o aspecto lexical, de Croft (2012; 2015), propomos uma explicação para o comportamento dos verbos de editoração. Sugerimos que esses verbos se comportam como atividades ou *accomplishments* porque descrevem mudanças direcionais que têm o potencial de serem construídas como télicas ou como atélicas. Mesmo com objetos diretos definidos (com denotação de indivíduo e não de tipo), os verbos de editoração podem ser atélicos em alguns contextos. Explicamos que isso ocorre por causa da natureza incremental desses verbos, que descrevem ações realizadas sobre um material textual, que é editado, preparado, revisado, diagramado e formatado parte por parte. Assim, analisando um pequeno grupo de

verbos, em inglês e em português, demonstramos que as variações no seu comportamento aspectual se explicam a partir da natureza semântica do domínio semântico a que esses verbos pertencem, o universo editorial.

Apesar de tratarmos de poucos verbos, acreditamos que essa análise possa ser estendida a outros tipos semânticos de verbos que também têm temas incrementais e parecem estar entre as atividades e os *accomplishments* (alguns exemplos são os verbos de criação, como *construir*, os verbos como *comer*, e outros). Uma extensão dessa análise para outros verbos é uma proposta de trabalho para o futuro.

Declaração de autoria

Gabriela Vilela Souza Martins: métodos, coleta e análise de dados, conceito da proposta original de análise dos verbos estudados; Luana Lopes Amaral: refinamento da proposta de análise, perspectiva teórica e explicações teóricas.

Agradecimentos

As autoras agradecem as valiosas contribuições dos pareceristas anônimos que avaliaram a primeira versão submetida deste artigo. Os erros que permanecem são, claramente, de nossa responsabilidade.

Referências

BASSO, R. M. Uma proposta para semântica dos adjuntos ‘em X tempo’ e ‘por X tempo’. *Alfa*, São Paulo, v. 55, n. 1, p. 113-134, 2011. Disponível em: <https://periodicos.fclar.unesp.br/alfa/article/view/4170>. Acesso em: 21 fev. 2022.

BEAVERS, J. Scalar complexity and the structure of events. In: DÖLLING, J.; HEYDE-ZYBATOW, T.; SCHÄFER, M. (eds.). *Event structures in linguistic form and interpretation*. Berlin: De Gruyter, 2008. p. 245-265.

CANÇADO, M.; AMARAL, L. *Introdução à Semântica Lexical: papéis temáticos, aspecto lexical e decomposição de predicados*. Petrópolis: Editora Vozes, 2016.

CARLSON, G. N. Generics and atemporal when. *Linguistics and Philosophy*, v. 3, n. 1, p. 49-98, 1979. DOI: <https://doi.org/10.1007/BF00578448>. Disponível em: <https://link.springer.com/article/10.1007/BF00578448>. Acesso em: 21 fev. 2022.

COMRIE, B. *Aspect: an introduction to the study of verbal aspect and related problems*. New York: Cambridge University Press, 1976.

CROFT, W. Force dynamics and directed change in event lexicalization and argument realization. In: ALMEIDA, R. G. de; MANOUILIDOU, C. (eds.). *Cognitive science perspectives on verb representation and processing*. New York: Springer, 2015. p. 103-30.

CROFT, W. *Verbs: aspect and causal structure*. Oxford: Oxford University Press, 2012.

DOWTY, D. *Word Meaning and Montague Grammar*. Dordrecht: D. Reidel, 1979.

LEVIN, B. *English verb classes and alternations: a preliminary investigation*. Chicago: University of Chicago Press, 1993.

LEVIN, B.; RAPPAPORT HOVAV, M. *Argument realization*. Cambridge: Cambridge University Press, 2005.

MARTINS, G. *Uma análise sintático-semântica dos verbos de editoração do inglês*. 2022. 82f. Monografia (Bacharel em Linguística do Inglês) - Faculdade de Letras, Universidade Federal de Minas Gerais, 2022.

RAPPAPORT HOVAV, M.; LEVIN, B. Building verb meanings. In: BUTT, M.; GEUDER, W. (eds.). *The projection of arguments: lexical and syntactic constraints*. Stanford: CSLI Publications, 1998. p. 97134.

RAPPAPORT HOVAV, M.; LEVIN, B. Reflections on manner/result complementarity. In: RAPPAPORT HOVAV, M.; DORON, E.; SICHEL, I. (eds.). *Lexical semantics, syntax, and event structure*. Oxford: Oxford University Press, 2010. p. 21-38.

ROTHSTEIN, S. *Structuring events: a Study in the Semantics of Lexical Aspect*. Oxford: Wiley-Blackwell, 2004.

SMITH, C. *The Parameter of Aspect*. Dordrecht: Kluwer, 1991.

VENDLER, Z. *Linguistics in philosophy*. New York: Cornell University, 1974.

VERKUYL, H. J. *Aspectual classes and aspectual composition*. *Linguistics and Philosophy*, v. 12, n. 1, p. 39-94, 1989. DOI: <https://doi.org/10.1007/BF00627398>

WACHOWICZ, T. C.; FOLTRAN, M. J.. Sobre a noção de aspecto. *Cadernos de Estudos Linguísticos*, Campinas, v. 48, n. 2, p. 211-232, 2006. DOI: <https://doi.org/10.20396/cel.v48i2.8637179>

ANEXO A

Dados do inglês (retirados de MARTINS, 2022)

To edit
This ensures that the copyeditor isn't spending time editing content that may be deleted or rearranged after a line edit. https://nybookeditors.com/2016/05/whats-the-difference-between-copyediting-and-proofreading/ .
I have been revising and editing my book for a few days now [...]. https://jerichowriters.com/typeset-own-book/ .
10 Simple Ways to Edit Your Own Book. https://thewritelife.com/self-editing-basics/ .
You want to try to forget everything you've written so that when you do come back to self-edit, the book almost seems as if someone else wrote it. https://thewritelife.com/self-editing-basics/
[...] which allows them more time to edit your actual words instead of tweaking your formatting. https://thewritelife.com/self-editing-basics/ .
I'm currently editing my second novel and this came at a perfect time for me. https://thewritelife.com/self-editing-basics/
[...] the only way to stand out against even more competition is by ferociously editing your own book until it's as crisp and clean as possible. https://thewritelife.com/self-editing-basics/ .
Hands down, one of the hardest things you'll ever do as a writer is edit your own work. https://nybookeditors.com/2016/08/diy-edit-10-tips-shape-manuscript/ .
I know some writers who edit as they go [...]. https://nybookeditors.com/2016/08/diy-edit-10-tips-shape-manuscript/ .
Tempted as you may be to go back and start editing what you just wrote [...]. https://nybookeditors.com/2016/08/diy-edit-10-tips-shape-manuscript/ .
[...] don't immediately start editing. https://nybookeditors.com/2016/08/diy-edit-10-tips-shape-manuscript/ .

Consider editing away from your computer. https://nybookeditors.com/2016/08/diy-edit-10-tips-shape-manuscript/ .
[...] why not edit in a completely different place than your creative zone. https://nybookeditors.com/2016/08/diy-edit-10-tips-shape-manuscript/ .
Before you go, don't miss these extra resources to help you edit your own manuscript. https://nybookeditors.com/2016/08/diy-edit-10-tips-shape-manuscript/ .
[...] make sure you've thoroughly revised and edited your work. https://www.scribbr.com/language-rules/what-is-proofreading/ .
When the material is nearly a finished product, meaning it has been edited, laid out, and designed [...]. https://www.grammarly.com/blog/whats-the-difference-between-copy-editing-and-proofreading/ .
[...] why, exactly, would you want to find out who edited the book you just finished reading? https://bookriot.com/how-to-find-the-editor-of-a-book/ .
For example, The Best American Short Stories 2019 was edited by Anthony Doerr. https://bookriot.com/how-to-find-the-editor-of-a-book/ .
Townsend Library classics have been edited to make them more accessible to today's readers. https://www.townsendpress.com/editing-process .
All of our books have been carefully edited to make them more readable. https://www.townsendpress.com/editing-process .
He has edited and written columns for the opinion pages of The Indian Express [...]. https://www.ipea.gov.br/revistas/index.php/rtm/article/view/235 .
He has co-edited the book In Search of Stability, Security and Growth: BRICS and a new world order. https://www.ipea.gov.br/revistas/index.php/rtm/article/view/235 .
When you've edited your manuscript once through with these questions in mind, pat yourself on the back. https://www.lisatener.com/2010/12/how-to-edit-a-book-how-many-times-should-i-edit/ .
Once you've edited a book for these three areas [...] you can likely move to a copy edit of the book. https://www.lisatener.com/2010/12/how-to-edit-a-book-how-many-times-should-i-edit/ .
I have a book that I've edited more times than I can recall. https://www.lisatener.com/2010/12/how-to-edit-a-book-how-many-times-should-i-edit/ .
Every book you've ever read has been edited. https://www.lisatener.com/2010/12/how-to-edit-a-book-how-many-times-should-i-edit/ .
The more famous the book is, the more likely it is that it's been edited up to a dozen times. https://www.lisatener.com/2010/12/how-to-edit-a-book-how-many-times-should-i-edit/ .

We edited the novel [...]. http://booksbywomen.org/how-i-kept-going-in-the-face-of-rejection-and-how-you-can-too/ .
I am editing a book that was shortlisted for a prize back in 2016 [...]. http://booksbywomen.org/how-i-kept-going-in-the-face-of-rejection-and-how-you-can-too/ .
We edited the book, designed and laid it out [...]. https://ukbookpublishing.com/examples/page/3/ .
We edited Calum's book [...]. https://ukbookpublishing.com/examples/page/3/ .
To copy edit
When they copy edit, these individuals provide an objective set of eyes that ensures grammar and spelling are correct [...]. https://opentextbc.ca/selfpublishguide/chapter/how-to-copy-edit/ .
Determine which sections and elements of the textbook should be copy edited. https://opentextbc.ca/selfpublishguide/chapter/how-to-copy-edit/ .
[...] decide if the unchanged original text and other elements of the book should be copy edited in addition to the new/changed text by the adapting author. https://opentextbc.ca/selfpublishguide/chapter/how-to-copy-edit/ .
[...] even if the original text has been copy edited to ensure that the new/changed text is consistent with the adapted work. https://opentextbc.ca/selfpublishguide/chapter/how-to-copy-edit/ .
[...] if the original text was not copy edited or poorly copy edited. https://opentextbc.ca/selfpublishguide/chapter/how-to-copy-edit/ .
[...] an experienced professional may copyedit around 1,000 to 3,000 words per hour on a typically straightforward text. https://www.ciep.uk/about/faqs/what-is-copyediting/ .
The author would like to thank Katharina Hoppe, Franziska von Verschuer [...] for their comments on an earlier version of this article and Gerard Holden, who copyedited the text. https://www.tandfonline.com/doi/abs/10.1080/1600910X.2017.1373686 .
Polly Kummel skillfully copyedited the text [...]. http://www.amazinphrasin.com/whattheysay.htm .
I copy-edited the text and proofread the display panels and object captions for the museum's reworked galleries [...]. http://www.archeloneditorial.co.uk/ .
I have copy edited and proofread: fiction, creative non-fiction [...]. https://www.eganediting.com/ .
[...] I fact checked and copy edited the text [...]. https://www.eganediting.com/ .
Have you copy-edited the manuscript and references carefully? https://academic.oup.com/jpubhealth/pages/instructions_for_authors .

<p>A manuscript that has been copy edited [...] by a professional ensures that you are prepared for the submission process [...]. https://www.econtentpro.com/blog/why-your-work-needs-to-be-copy-edited/171.</p>
<p>Copy edit your work to ensure the smoothest submission process possible! https://www.econtentpro.com/blog/why-your-work-needs-to-be-copy-edited/171.</p>
<p>A document that has been copy edited presents a much more attractive submission. https://www.econtentpro.com/blog/difference-between-copy-editing-and-scientific-and-scholarly-editing/172.</p>
<p>[...] the one that has been copy-edited will make it to print ahead of the other [...]. https://www.editage.com/insights/sites/default/files/Copyediting%20of%20research%20papers_who%20and%20why%20and%20why%20not.pdf.</p>
<p>Proofreading is the process of examining the final draft of a document or text – after it has been copy-edited – to ensure there are absolutely no errors. https://www.expressproofreading.com/what-is-proofreading/</p>
<p>Once the manuscript has been copy edited [...], a sample draft of the final product is made up by a typesetter or designer. https://hankers4words.com/copy-editing-v-proofreading-the-differences/</p>
<p>[...] as it's a production-level edit that is only appropriate after a manuscript destined for publication has been copy edited. https://www.editorialdepartment.com/services-2-2/line-copy-editing/.</p>
<p>Asking for a proofread of a document that has not been copy-edited will either lead to refusal or a longer, more expensive process than you were expecting. https://www.athenacopy.com/post/2016/11/07/copy-editing-and-proofreading-whats-the-difference.</p>
<p>We usually recommend a proof read once the book has been copy edited [...]. https://www.thebookrefinery.com/editing-proofreading/.</p>
<p>Vultee then copy-edited the articles to fix any mistakes in style, word usage, or grammar [...]. https://www.americanpressinstitute.org/publications/research-review/the-value-of-copy-editing/</p>
<p>[...] with the title of your manuscript as the subject, and it will be copy edited [...]. https://publishers.forewordreviews.com/editing/copy/#chapter-preview.</p>
<p>[...] a significant number of NPR digital stories are not copy edited prior to publication or at all [...]. https://www.npr.org/sections/publiceditor/2020/06/25/882829255/in-choosing-the-wrong-photo-npr-editors-paved-the-way-for-partisan-attack.</p>
<p>This column was corrected to indicate that the story in question was copy edited [...]. https://www.npr.org/sections/publiceditor/2020/06/25/882829255/in-choosing-the-wrong-photo-npr-editors-paved-the-way-for-partisan-attack.</p>

<p>[...] the story was not copy edited at all. https://www.npr.org/sections/publiceditor/2020/06/25/882829255/in-choosing-the-wrong-photo-npr-editors-paved-the-way-for-partisan-attack.</p>
<p>I copy edited the manuscript to correct for sentence-level errors and flow [...]. https://www.theeditingco.com/blog/10306/ask-the-editor-how-do-i-get-an-agent.</p>
<p>Cathie Bucci copy-edited the manuscript for my second novel [...]. https://www.redpeneditorialmaine.com/testimonials.</p>
<p>To proofread</p>
<p>Please ensure that the final manuscript matches our House Style, has been proof-read and is as free from error or omission as you can make it. https://www.bloomsbury.com/us/academic/for-authors/a-guide-to-the-publishing-process/.</p>
<p>[...] you might try to proofread your own work. https://nybookeditors.com/2016/05/whats-the-difference-between-copyediting-and-proofreading/</p>
<p>[...] you can choose to proofread the text yourself or to hire a professional. https://www.scribbr.com/language-rules/what-is-proofreading/.</p>
<p>Often a text will go through several stages of editing before it is proofread. https://www.scribbr.com/language-rules/what-is-proofreading/.</p>
<p>[...] there are some techniques you can use to proofread efficiently and effectively [...]. https://www.scribbr.com/language-rules/what-is-proofreading/.</p>
<p>[...] Only proofread once you've got a completed final draft that you're happy with. https://www.scribbr.com/language-rules/what-is-proofreading/.</p>
<p>[...] you might not realize just how long it will take to thoroughly proofread your text. https://www.scribbr.com/language-rules/what-is-proofreading/.</p>
<p>The easiest way to proofread is to read your writing aloud. https://www.bbc.co.uk/bitesize/topics/zr6bxyc/articles/z7vhsrd.</p>
<p>You can also enlist a friend you trust to proofread your work. https://www.masterclass.com/articles/how-to-proofread-your-writing.</p>
<p>Writing Center instructors won't proofread your papers, but they'll be glad to explain mistakes [...]. https://writing.wisc.edu/handbook/grammarpunct/proofreading/.</p>
<p>The entire accuracy and integrity of the research represented and discussed in a paper can hinge on whether it was proofread. https://www.proofreadnow.com/blog/6-types-of-writing-you-always-want-a-proofreader-to-review.</p>
<p>But every white paper should still be proofread [...]. https://www.proofreadnow.com/blog/6-types-of-writing-you-always-want-a-proofreader-to-review.</p>

<p>Not everything you write may have to be proofread. https://www.proofreadnow.com/blog/6-types-of-writing-you-always-want-a-proofreader-to-review.</p>
<p>But everything you write that's important [...] should certainly be proofread. https://www.proofreadnow.com/blog/6-types-of-writing-you-always-want-a-proofreader-to-review.</p>
<p>The text was proofread by a native or advanced English speaker. https://en.wikisource.org/wiki/Template:Translation.</p>
<p>This text needs to be proofread by a native or advanced speaker of the original language [...]. https://en.wikisource.org/wiki/Template:Translation.</p>
<p>[...] and the copyediting was done well and we proofread, the book should come through with a minimum level of errors [...]. https://www.johnhuntpublishing.com/publishing-guide/chapter-8-editorial/proofs/.</p>
<p>We have proofread the manuscript again and made some modifications. https://www.mdpi.com/2073-4425/10/11/886/review_report.</p>
<p>Authors should not make changes to textbook chapters once they have been proofread as this can undo the proofreader's work. https://opentextbc.ca/selfpublishguide/chapter/how-to-proofread/.</p>
<p>[...] this post aims to give a lot more detail and advice on how to check your document after it has been proofread and returned to you. http://www.technicalix.co.uk/top-tips-for-reviewing-your-document-after-it-has-been-proofread/.</p>
<p>Also, the guidelines below only relate to material that I have proofread [...]. http://www.technicalix.co.uk/top-tips-for-reviewing-your-document-after-it-has-been-proofread/.</p>
<p>Getting your manuscript or paper back after it has been proofread/edited is an exciting time. http://www.technicalix.co.uk/how-to-respond-to-editorial-comments-and-queries/.</p>
<p>Lots of my clients ask me whether their manuscript will be "perfect" or "free from errors" after I've edited or proofread it. http://www.technicalix.co.uk/the-search-for-perfection/.</p>
<p>We proofread the book, laid it out and designed a cover for the hardback book [...]. https://ukbookpublishing.com/testimonial/the-history-gate-adventures/.</p>
<p>We proofread the book and converted it into Kindle [...]. https://www.karenaundersassoc.com/case-study/.</p>
<p>In other words, write first and then proofread. https://www.relevance.com/how-to-use-proofreading-tool-for-error-free-writing/.</p>
<p>We systematically proofread any project before delivery [...]. https://intrawords.com/translation-edition-and-proofreading/?lang=en/.</p>
<p>[...] we proofread the originals for any possible errors or difficulties [...]. https://intrawords.com/translation-edition-and-proofreading/?lang=en/.</p>

We created the interior page layout design and then proofread the typeset PDF file. https://www.brisbaneselfpublishing.com.au/our-authors/ .
[...] we designed the cover and interior pages, proofread the typeset pages and wrote the blurb for the back cover. https://www.brisbaneselfpublishing.com.au/our-authors/ .
To typeset
All changes are then incorporated into the electronic text before being typeset. https://www.bloomsbury.com/us/academic/for-authors/a-guide-to-the-publishing-process/ .
[...] and the book is typeset in InDesign. https://www.bloomsbury.com/us/academic/for-authors/a-guide-to-the-publishing-process/ .
[...] writers can now typeset their manuscripts themselves. https://jerichowriters.com/typeset-own-book/ .
Authors who typeset their own work often don't justify their text. https://jerichowriters.com/typeset-own-book/ .
Whether you end up typesetting your own book or not, you'll be surprised how often some of these typesetting skills will crop up in your writing career. https://jerichowriters.com/typeset-own-book/ .
Authors themselves rarely are trained how to typeset [...]. https://www.graphicdesigndegreehub.com/faq/what-is-typesetting/ .
We edited Calum's book, [...] typeset the book and set it up for print on demand. https://ukbookpublishing.com/examples/page/3/ .
ABU performed statistical analysis and typesetted the manuscript [...]. http://www.emtoscipublisher.com/index.php/ija/article/html/2196/ .
During this phase, our production designer will typeset the manuscript. https://fitpublishing.com/information_for/lifecycle-manuscript .
Do not attempt to typeset the manuscript or use any special formatting. https://journals.sagepub.com/pb-assets/cmscontent/JHC/JHC_Instructions_to_Authors_Revised_JUNE_2017.pdf .
[...] we typeset the manuscript and send you final proofs in PDF format. https://www.openbookpublishers.com/section/139/1 .
You typeset the manuscript, or hire it done. https://goteenwriters.com/2019/09/09/an-overview-of-the-self-publishing-process/ .
I typeset the manuscript on the compositor of the ANU Students' Association [...]. https://documents.uow.edu.au/~bmartin/pubs/80nk/writingnk.html .
[...] we typeset a document that would show how the various design factors actually looked on paper. https://www.tug.org/TUGboat/tb12-3-4/tb33weiss.pdf .
While typesetting the manuscript, we kept a print-out of the source code for the sample pages handy for reference purposes. https://www.tug.org/TUGboat/tb12-3-4/tb33weiss.pdf .

<p>The company also typeset advertisements for large grocery stores chains throughout the Mid-Atlantic. https://carterprinting.com/who-we-are.</p>
<p>Carter Mat Service had typeset for many religious magazines and other retail clients as well. https://carterprinting.com/who-we-are.</p>
<p>I had learned the very basics of the inDesign software [...] and had typeset a couple books as part of that course. https://blog.pshares.org/an-interview-with-book-designer-kate-hargreaves/.</p>
<p>In 2013 he married Mary Merry, who had typeset Berlin's books since 1996 [...]. http://calculus.wolf.ox.ac.uk/~hardy/lists/unpublished/auto-obituary.html.</p>
<p>But no [...] he had typeset it himself. https://www.rferl.org/a/Was_Pasternaks_Path_To_The_Nobel_Paved_By_The_CIA/1496794.html.</p>
<p>[...] it ended up containing the most fonts of any book Sorrel had typeset! https://laurenejames.co.uk/2015/08/21/behind-the-book-typesetter-sorrel-packham/.</p>
<p>Sorrel Packham has typeset a variety of publications for Walker Books [...]. https://laurenejames.co.uk/2015/08/21/behind-the-book-typesetter-sorrel-packham/.</p>
<p>We typeset the front matter (title page, copyright page, dedication, etc.) as well as the first few pages and send it back. https://typerightediting.com/typesetting/.</p>
<p>[...] we'll typeset the rest of the book. https://typerightediting.com/typesetting/.</p>
<p>To format</p>
<p>I designed and created my own thumbnail cover, formatted the book to their specs, and uploaded it. https://www.thebookdesigner.com/2011/08/independent-publishing-thats-evolution/.</p>
<p>This is your opportunity to ensure we've formatted the book correctly and fix any typos or omissions [...]. https://design.bookmobile.com/book-design/.</p>
<p>Often it's the publisher that formatted the book for the author. https://www.firstediting.com/trying-to-edit-a-document-in-pdf-vs-word-format/.</p>
<p>Kitty Buchner [...] formatted the book so it could be sold on Amazon.com in paperback and Kindle editions [...]. http://library.cityofalbany.net/wp-content/uploads/2011/11/Kathy-Hering-with-book-ad-final.pdf.</p>
<p>Many authors hand distribution over to the self-publishing company that edited and formatted the book [...]. https://www.1106design.com/2019/03/25/difference-between-ingramspark-and-kindle-direct-publishing/.</p>

We typeset the book, formatted the book, designed the cover and distributed the book in all formats globally. https://publishingpush.com/ .
I just formatted the e-book edition of my short story collection, Twenty Miles West of Branch, Texas and other stories [...]. https://www.kcknouse.com/my-initial-experience-using-kindle-create .
The book was formatted using the Kindle template I downloaded two years ago. https://www.kcknouse.com/my-initial-experience-using-kindle-create .
The title pages of chapters/stories were easily formatted with drop caps [...]. https://www.kcknouse.com/my-initial-experience-using-kindle-create .
I chose Modern but experimented with all of them once I had the book formatted. https://www.kcknouse.com/my-initial-experience-using-kindle-create .
These steps should only be completed after all chapters have been formatted in the manuscript [...]. https://www.thebookdesigner.com/2019/11/preparing-your-manuscript-for-publication-part-3/ .
Format the book as normal, completing the addition of all chapters [...]. https://www.thebookdesigner.com/2019/11/preparing-your-manuscript-for-publication-part-3/ .
[...] and you've formatted it to look the way you want. https://www.harvard.com/images/uploads/diy.pdf .
Whether you hire someone to help you or you format the book on your own, here are some helpful guidelines. https://www.harvard.com/images/uploads/diy.pdf .
When formatting your document, each PDF page should equal one book page. https://www.harvard.com/images/uploads/diy.pdf .
Don't locally format. https://www.thebookdesigner.com/2011/03/4-top-book-formatting-mistakes-to-avoid/ .
Learn to format right. https://www.thebookdesigner.com/2011/03/4-top-book-formatting-mistakes-to-avoid/ .
Write first, format later. https://www.self-pub.net/blog/common-mistakes-made-when-writing-a-book-in-microsoft-word/ .
[...] and they can format it accordingly. https://www.self-pub.net/blog/common-mistakes-made-when-writing-a-book-in-microsoft-word/ .
Imagine a whole book being formatted and you receive the first draft [...]. https://www.indiepublishinggroup.com/what-is-book-formatting/ .
Once the book has been formatted you'll get your first draft back. https://www.indiepublishinggroup.com/what-is-book-formatting/ .

This book has been formatted to be readable by visually-impaired and non-impaired readers and writers.

<https://bookshop.org/books/a-quick-guide-to-formatting-your-book-or-ebook-for-the-dyslexic-and-visually-impaired-reader-formatted-for-all-readers-and-writers/97817920-05381>.

Proofing happens after a book has been formatted [...].
<https://bookalysr.com/editing/>.

ANEXO B

Dados do português (retirados de MARTINS, 2022)

Editorar
Editorou e ilustrou vários livros [...]. http://mapacultural.saocaetanodosul.sp.gov.br/historico/151777/ .
[...] editorou e desenvolveu infográficos para mais de 40 soluções. https://murilhas.com.br/portfolio/folheteria-totvs .
Ilustrou e editorou diversos livros “Os Vivos, o morto e o peixe frito”, do autor angolano Ondjaki [...]. https://www.coletivoteatrodecafonico.com/vania.html .
Editorou 19 livros [...]. https://sigaa.ufla.br/sigaa/public/docente/portal.jsf?siape=395992 .
Em 2017 editorou o projeto transmídia “Cartografia Cultural do Crato” junto a Secult-Crato e a URCA [...]. https://mapacultural.juazeiro.ce.gov.br/agente/26920/ .
Charles Pessanha não é apenas um nome central na história de DADOS, periódico que ele editorou por mais de três décadas [...]. https://blog.scielo.org/blog/tag/programa-scielo/#.YH3NGu5KjIU .
Lönnrot os editorou para obter uma narrativa coerente, e o resultado foi um longo poema épico [...]. https://www.jundiaqui.com.br/memoria/origem-e-cura-da-pesto-na-mitologia-finlandesa/ .
A S2C e Secco Editora editorou e imprimiu a obra. https://www.imprensaolivros.com.br/escolas?lightbox=dataItem-k3hji1wv .
Além de revisar textos médicos e científicos, dissertações e teses, ainda editorava os artigos [...]. https://www.researchgate.net/publication/317815730_A_revisao_de_portugues_da_ASTEC_CAISM .
Fora essas jornadas, eu editorava periódicos do Programa Pós-Graduação em Artes da UFES [...]. https://notamanuscrita.com/2020/09/17/transcricao-npc-01/ .

<p>[...] e, como editor, foi responsável pelo chamado “Caderno B” e também editorava o “Informe JB”.</p> <p>http://monumentosdorio.com.br/monu/antigo/br/esculturas/038/013.htm.</p>
<p>[...] Monteiro Lobato editorava a revista [...].</p> <p>https://historiapt.info/feira-de-santana-ba.html?page=24.</p>
<p>Edward Cave era o editor do periódico e após conhecer seu portfólio, deu uma história [...] além de alguns outros desenhos para um livro que [ele] estava editorando.</p> <p>https://chamaescoteira.wordpress.com/2012/11/01/norman-rockwell-o-artista-escoteiro/.</p>
<p>[...] ligações de madrugada para fazer a correção de uma palavra em um livro que eu estava editorando. https://www.julianetorres.com.br/business-dicas-de-etiqueta-para-usar-o-whatsapp-comercial/.</p>
<p>Conheci-o pessoalmente quando estava editorando a revista Last Call for Beer [...].</p> <p>https://www.brejas.com.br/forum/mercado/41750-o-meio-ervejeiro-e-machista/58789?limitstart=48.</p>
<p>O almanaque foi editorado e entregue na escola para os alunos [...].</p> <p>https://ojs.unesp.br/index.php/revista_proex/article/view/601.</p>
<p>O volume, de 368 páginas, foi editorado pela Profa. Darcília Simões.</p> <p>http://www.filologia.org.br/filologo/pub_filologo11.html.</p>
<p>O livro foi editorado e organizado pelo Instituto Quorum [...].</p> <p>https://livrosporlivria.medium.com/a-sutileza-e-a-imposi%C3%A7%C3%A3o-do-tempo-83694b1bfcf5.</p>
<p>“Este livro foi editorado durante o mês de junho de 1979 pelos alunos do quinto semestre de Editoração da Escola de Comunicação e Artes da Universidade de São Paulo” [...].</p> <p>https://revistas.pucsp.br/index.php/pontoevirgula/article/viewFile/14054/10356.</p>
<p>O livro foi editorado e a arte-finalização feita por Jonas de Sene Pinto, da editora Occasio.</p> <p>https://gebio.com.br/2018/08/12/livro-sobre-monitoramento-de-adultos-da-broca-da-cana-sera-lancado-no-6o-tecnobio-cana/.</p>
<p>O livro foi editorado com o apoio da Fundação de Apoio ao Colégio Estadual Júlio de Castilhos [...].</p> <p>http://www.poa.ifrs.edu.br/index.php?option=com_content&view=article&id=869:pibid-do-campus-porto-alegre-lanca-livro-que-relata-vivencias-docentes-dos-bolsistas&catid=17&Itemid=121.</p>
<p>Para registrar a memória das palestras apresentadas no evento, foi editorado o livro “SAVANAS: desafios e estratégias para o equilíbrio entre sociedade, agronegócio e recursos naturais”.</p> <p>https://www.grupocultivar.com.br/noticias/artigo-savanas-desafios-e-estrategias-para-o-equilibrio-entre-sociedade-agronegocio-e-recursos-naturais.</p>

Editar
<p>Professora Rose Mary editou livro sobre semiologia cardiovascular. https://www.medicina.ufmg.br/professora-lanca-livro-sobre-metodos-clinicos-exames-e-principais-sindromes-da-cardiologia/.</p>
<p>Luís Garnier [...] editou mais de seiscentos autores brasileiros [...]. https://historiapt.info/feira-de-santana-ba.html?page=24.</p>
<p>[...] Lobato editou pela revista o seu livro “Urupês” [...]. https://historiapt.info/feira-de-santana-ba.html?page=24.</p>
<p>Fátima Silva editou livro sobre as raízes de Virginia Moura. http://public.vivacidade.org/informacao/fatima-silva-editou-livro-sobre-as-raizes-de-virginia-moura/.</p>
<p>A publicação está em fase de revisão e deverá ser lançada pela Matrix Editora, a mesma que editou livro sobre a história não-autorizada de Suzane Von Richthofen. https://www.jornaldacidadeonline.com.br/noticias/26025/em-livro-cunha-promete-revelacoes-bombasticas-sobre-maia.</p>
<p>Para enriquecer este acervo, a Editora Bertrand Brasil editou “Livro dos Sith” do original “Book of Sith” [...]. https://www.universodosleitores.com/2014/04/o-livro-dos-sith.html.</p>
<p>Associação ‘Leões Da Beira’ editou livro sobre as lendas e tradições de Rio de Loba. https://www.noticiasdeveiu.com/associacao-leoes-da-beira-editou-livro-sobre-as-lendas-e-tradicoes-de-rio-de-loba/.</p>
<p>Irene editou o seu primeiro livro de memórias fotográficas: Negro em Preto e Branco [...]. https://www.jornaldocomercio.com/_conteudo/especiais/reportagem_cultural/2020/10/761084-como-o-sentimento-da-saudade-se-revela-na-criacao-artistica-contemporanea.html.</p>
<p>Christopher Tokien, filho do escritor, editou livro com história escrita em 1917. https://oglobo.globo.com/cultura/livros/novo-livro-de-jrr-tolkien-vai-ser-lancado-no-ano-que-vem-20333429#ixzz4Nk66JcOP.</p>
<p>Quando os autores ficavam fora de uma dessas três casas, estavam sendo editados em Portugal: a Livraria Chardron [...]. https://periodicos.fclar.unesp.br/itinerarios/article/view/2488/2086.</p>
<p>[...] ele aceitou editar os livros por sua própria conta [...]. https://periodicos.fclar.unesp.br/itinerarios/article/view/2488/2086.</p>
<p>Foi um dos poucos poetas a viver exclusivamente da sua arte: escrevia, editava, [...] e comercializava suas criações. http://www.casaruibarbosa.gov.br/interna.php?ID_S=52&ID_M=4122.</p>
<p>Vindo da França, no Rio de Janeiro construiu a Livraria Garnier, que editava seus livros no Brasil [...]. https://www.literaturabrasileira.ufsc.br/autores/?id=18437.</p>

<p>Confira neste tutorial como usar a ferramenta de backup do Word para recuperar um arquivo que você estava editando e acabou perdendo porque não foi salvo [...]. https://canaltech.com.br/software/como-recuperar-facilmente-arquivos-perdidos-no-word/.</p>
<p>Estava editando quando 28% do Pantanal foi embora. https://cgn.inf.br/noticia/400112/%C2%91existe-no-pais-um-projeto-ecocida%C2%92-diz-araquem-alcantara.x</p>
<p>Antes de 1500 a Europa estava editando livros a um ritmo de 1000 títulos por ano. https://www.passeidireto.com/arquivo/40149645/didatica.</p>
<p>Desde então, Aline se dedicou a revisá-la e a editá-la. https://livrosporlivria.medium.com/a-sutileza-e-a-imposi%C3%A7%C3%A3o-do-tempo-83694b1bfcf5.</p>
<p>Nela foi editado o primeiro jornal da colônia americana: a Gazeta do Rio de Janeiro. http://www.multirio.rj.gov.br/index.php/estude/historia-do-brasil/rio-de-janeiro/52-o-rio-de-janeiro-como-a-capital-do-reino/2483-a-imprensa-regia.</p>
<p>O “Quarto de despejo” foi editado e publicado por Audálio Dantas [...]. https://brainly.com.br/tarefa/40443013.</p>
<p>Livro sobre o Brasil na ordem global, editado por Oliver Stuenkel, é resenhado no Journal of Latin American Studies. https://ri.fgv.br/noticias/2017-10-30/livro-sobre-o-brasil-na-ordem-global-editado-por-oliver-stuenkel-e-resenhado-no-.</p>
<p>O livro, publicado em 2015 pela Palgrave Macmillan, foi editado pelos professores Matthew M. Taylor [...]. https://ri.fgv.br/noticias/2017-10-30/livro-sobre-o-brasil-na-ordem-global-editado-por-oliver-stuenkel-e-resenhado-no-.</p>
<p>Este número foi editado, diagramado e publicado com o apoio do auxílio financeiro concedido pelo PIPEq – PUC-SP [...]. https://revistas.pucsp.br/index.php/IGISP/issue/archive.</p>
<p>Livro foi editado pela George Allen & Unwin. https://www.correiodopovo.com.br/artegenda/o-hobbit-obra-que-apresentou-a-terra-m%C3%A9dia-%C3%A0-literatura-completa-80-anos-1.241746.</p>
<p>O segundo livro, em 553 páginas, [...] foi editado em 1953, 30 anos após a primeira edição do Volume I [...]. http://www.aambm.org.br/slivros.php.</p>
<p>O livro “Congresso Nacional: A Construção do Espaço da Democracia” foi editado em três línguas [...]. https://www.jornaldoeste.com.br/geral/camara-lanca-livro-com-historia-do-congresso-nacional/.</p>
<p>Preparar</p>
<p>Além de preparar o corpo do artigo, na forma IMRD, há outras exigências dos editores de periódicos. https://www.scielo.br/pdf/ress/v26n3/2237-9622-ress-26-03-00661.pdf.</p>

<p>Além disso, ele dá noções de como preparar o livro para enviar para a Apple [...]. https://blogdoiphone.com/livros/conheca-um-passo-a-passo-de-como-publicar-seus-livros-na-ibookstore-sem-editoras/.</p>
<p>[...] busquei o apoio de historiadores/as, professores/as e profissionais de edição que discutem a leitura e seus suportes, a construção do livro e, também, o próprio ato de preparar o texto em uma editora. https://repositorio.ufsc.br/bitstream/handle/123456789/221302/PGET0500-D.pdf?sequence=-1.</p>
<p>Preparar um texto significa colocá-lo “em condições adequadas para ser transformado em livro”, assim conclui Plínio Martins Filho (2016, p. 151) [...]. https://repositorio.ufsc.br/bitstream/handle/123456789/221302/PGET0500-D.pdf?sequence=-1.</p>
<p>Preparar o texto em uma editora é, portanto, normalizá-lo, padronizá-lo, corrigir eventuais ‘erros’ de ortografia e gramática [...]. https://repositorio.ufsc.br/bitstream/handle/123456789/221302/PGET0500-D.pdf?sequence=-1.</p>
<p>A primeira etapa, “batida de parágrafos”, aplica-se quando o texto a ser preparado é uma tradução [...]. https://repositorio.ufsc.br/bitstream/handle/123456789/221302/PGET0500-D.pdf?sequence=-1.</p>
<p>[...] há divergências entre teóricos e práticos da edição de textos quanto à necessidade do conhecimento, por parte do/a preparador/a, da língua do texto fonte da tradução a ser preparada. https://repositorio.ufsc.br/bitstream/handle/123456789/221302/PGET0500-D.pdf?sequence=-1.</p>
<p>Ao questionar os/as preparadores/as se eles/as conhecem as línguas dos originais dos textos traduzidos que preparam e se acreditam que isso seja importante [...]. https://repositorio.ufsc.br/bitstream/handle/123456789/221302/PGET0500-D.pdf?sequence=-1.</p>
<p>Acredito que essa seja a forma ideal de se preparar um texto literário traduzido [...]. https://repositorio.ufsc.br/bitstream/handle/123456789/221302/PGET0500-D.pdf?sequence=-1.</p>
<p>[...] pensei, amedrontado, ao receber da Carol Leocadio, minha subeditora, a incumbência de preparar o texto de A sexta extinção. https://www.intrinseca.com.br/blog/2015/09/notas-sobre-o-mercado-editorial-e-a-extincao-em-massa/.</p>
<p>Enquanto preparava o texto do livro, colhi algumas informações que podem ajudar você na excelente leitura de A sexta extinção. https://www.intrinseca.com.br/blog/2015/09/notas-sobre-o-mercado-editorial-e-a-extincao-em-massa/.</p>

<p>O revisor atua normalmente nas empresas jornalísticas e nas editoras, preparando o texto até que ele se transforme em sua versão final, pronta para ser publicada. https://www.infoescola.com/profissoes/revisor-de-textos/.</p>
<p>No planejamento gráfico, a primeira coisa que o produtor deve levar em conta ao preparar um original é a finalidade e o público a que se destina [...]. https://repositorio.ifpb.edu.br/bitstream/177683/907/1/TCC%20-%20%20Adino%20Saraiva%20Bandeira.pdf.</p>
Revisar
<p>Quem revisa, portanto, só tem a ganhar, ainda mais quando se trata de textos acadêmicos para defesa ou concursos [...]. https://www.keimelion.com.br/2014/07/importancia-de-revisar.html.</p>
<p>O texto que revisarmos não vai ter crítica da banca. https://www.keimelion.com.br/2014/07/importancia-de-revisar.html.</p>
<p>Todo texto nasce pronto para ser revisado. https://www.keimelion.com.br/2014/07/importancia-de-revisar.html.</p>
<p>Ser capaz de melhorar esse tipo de texto, revisando-o, para efeitos de transmissão direta, coerente e eficaz de informação [...]. https://www.keimelion.com.br/2017/12/revisao-textos-procedimentais.html.</p>
<p>Revisar um manual de produto, é também uma tarefa em si [...]. https://www.keimelion.com.br/2017/12/revisao-textos-procedimentais.html.</p>
<p>Portanto, há também a possibilidade de se ler um texto para revisá-lo. https://www.keimelion.com.br/2017/11/estrategias-revisao.html.</p>
<p>[...] as características particulares do texto que é revisado [...]. https://www.keimelion.com.br/2017/11/estrategias-revisao.html.</p>
<p>Eles tiveram que revisar artigos de vários tipos [...]. https://www.keimelion.com.br/2017/03/precos-servicos-revisao.html.</p>
<p>De fato, estudos indicam que escrever ou revisar consiste em representações mentais das quais uma parte é mais ou menos acessível e, portanto, verbal. https://www.keimelion.com.br/2017/12/revisao-de-texto-em-tempo-real.html.</p>
<p>No contexto de comunicação específico que descrevemos, o autor explica a si as atividades cognitivas que realiza, ou seja, escrever e revisar [...]. https://www.keimelion.com.br/2017/12/revisao-de-texto-em-tempo-real.html.</p>
<p>As escolhas que os autores fazem ao autorrevisar em tempo real indicam que as questões que os perturbam estão relacionadas tanto à forma como ao significado [...]. https://www.keimelion.com.br/2017/12/revisao-de-texto-em-tempo-real.html.</p>
<p>Em geral, pode-se dizer que quem relê seu texto e faz ajustes ortográficos, semânticos ou de estilo o está revisando. https://www.keimelion.com.br/2021/03/terminologia-revisao-tese.html.</p>

<p>Temos repetido isso exaustivamente: ninguém pode revisar o que escreveu. https://www.keimelion.com.br/2021/03/terminologia-revisao-tese.html.</p>
<p>Revisávamos uma tese para universidade cubana [...]. https://www.keimelion.com.br/2021/03/terminologia-revisao-tese.html.</p>
<p>O texto foi revisado e correções sugeridas foram incluídas no corpo do manual. http://www2.ufac.br/ppgespa/menu/ppgespa_old/noticias/5.html.</p>
<p>O texto foi revisado a partir da tradução para o italiano [...]. https://www.e-publicacoes.uerj.br/index.php/revistaceaju/article/view/39507.</p>
<p>A autora está revisando o livro. https://www.atelie.com.br/livro/angola-mocambique-experiencia-colonial-territorios-literarios/.</p>
<p>Se você está insatisfeito com ele, recomece, reescreva, revise, refaça. https://medium.com/carinavd/querida-pessoa-que-quer-ser-autor-vamos-conversar-19482b56da52.</p>
<p>Revisando o livro pela 20ª vez. https://medium.com/carinavd/querida-pessoa-que-quer-ser-autor-vamos-conversar-19482b56da52.</p>
<p>Diagramar</p>
<p>A editora chamada Amauta não quer um tostão pelos seis livros que diagramou, traduziu, ilustrou, revisou, imprimiu [...]. https://www1.folha.uol.com.br/fsp/ilustrad/fq0704200511.htm.</p>
<p>Ele escreveu, editou, diagramou e até ilustrou a capa do volume. https://www.folhape.com.br/cultura/publicacao-independente-de-poesia-a-cidade-vence-jabuti-de-livro-do-an/87010/.</p>
<p>E não só publicou como escreveu, ilustrou, diagramou e imprimiu cada um [...]. https://lovelyhouse.com.br/publicacao/gogmagog-morris-cox-e-sua-gogmagog-press-gustavo-piqueira/.</p>
<p>Por fim, um agradecimento especial ao Rodrigo Franco, que revisou e diagramou esse playbook [...]. https://openaccess.blucher.com.br/article-details/07-21555.</p>
<p>Criativo e entusiasmado, Breno Ilan escreveu a obra, criou capa e diagramou o texto. http://www.jornaldacidade.net/cultura/2018/08/302752/jovem-autor-lanca-livro-pela-edise.html.</p>
<p>Nessa peça, as ilustrações utilizadas não foram produzidas pelo artista, que diagramou e projetou as mesmas. https://www.laionpessoa.com/one-page-ipuranga.</p>
<p>Olhando assim, pode parecer difícil pensar que existam técnicas e dicas para diagramar como um profissional, certo? https://comunidade.rockcontent.com/como-fazer-uma-diagramacao/.</p>

Um mesmo conteúdo pode ter um resultado completamente diferente de acordo com a maneira como foi diagramado. https://comunidade.rockcontent.com/como-fazer-uma-diagramacao/ .
Também diagramamos e-books para estratégias de inbound marketing. https://www.diagramado.com.br/ .
Permite aproveitar melhor os espaços do material a ser diagramado. http://www.lynxcomunicacao.com.br/blog/o-que-e-diagramacao/ .
[...] recomendamos que verifique sempre o seu layout para ver se seu livro foi diagramado corretamente [...]. https://www.viegaseditora.com/diagramacao-miolo .
[...] é só pegar o texto que já foi diagramado para o livro físico e converter no formato de e-book, não é mesmo? https://www.editorabookmarks.com/post/quarta-p%C3%A1gina-por-que-a-diagrama%C3%A7%C3%A3o-do-e-book-%C3%A9-diferente-do-livro-f%C3%ADsico .
O livro virtual, lançado pelo Conif com foco nos parceiros internacionais, foi diagramado e publicado pela Essentia Editora do IFF. http://portal1.iff.edu.br/@@search?advanced_search=True&sort_on=sortable_title&b_start:int=20&Subject:list=Livro .
O texto da obra foi diagramado em caixa alta para contemplar as crianças da alfabetização [...]. https://www.araquari.sc.gov.br/noticia/4142/professora-do-cei-crianca-bela-no-itinga-lanca-livro-infantil .
Para descobrir como diagramar bem, você precisa compreender essas demandas e trabalhar de acordo com elas. https://deolhonofuturo.uninter.com/como-diagramar/ .
Nesse sentido, é importante saber como diagramar um livro [...]. https://printstore.com.br/4-dicas-de-como-diagramar-um-livro-de-sucesso/ .
Para diagramar um livro com eficiência, é fundamental observar alguns aspectos, como público-alvo [...] entre outras questões envolvidas na edição da obra. https://printstore.com.br/4-dicas-de-como-diagramar-um-livro-de-sucesso/ .
Como vimos, saber como diagramar um livro pode ser um fator decisivo para o seu sucesso. https://printstore.com.br/4-dicas-de-como-diagramar-um-livro-de-sucesso/ .
[...] pois era ela quem diagramava a revista [...]. https://www.scielo.br/scielo.php?script=sci_arttext&pid=S0370-44672014000100001 .
No início, em 2010, acordava de madrugada para distribuir a edição feita no dia anterior, que [eu] também diagramava. https://www.lenoticias.com.br/axe-schettini/1157/o-futuro-do-jornalismo-no-IE-notfcias .
No início não diagramava, porque meu computador não tinha memória suficiente (é sério) [...]. http://www.seguinte.inf.br/mobile/noticias/colunas/8898_Historias-de-jornalista-Y-2 .

<p>Quando eu tinha 9 anos, a professora Maria do Carmo, que tinha me dado aula na 2a série [sic], descobriu que eu escrevia historinhas em casa, sozinho, e que diagramava em formato de livrinho [...].</p> <p>https://daguitorodrigues.com/2016/06/24/um-recado-com-mais-de-26-anos/.</p>
<p>No entanto, é possível conseguir seus objetivos diagramando livros no Word [...].</p> <p>https://www.wellingtondemelo.com.br/site/2013/02/10/diagramando-livros-no-word-parte-1/.</p>
<p>Se você diagramar em espaçamento duplo tamanho A4 seu livro [...].</p> <p>https://www.wellingtondemelo.com.br/site/2013/02/10/diagramando-livros-no-word-parte-1/.</p>
<p>[...] quando atuava como estagiária de programação visual, diagramando livros com uso do software PageMaker 7.0. https://labs.dualpixel.com.br/mulheres-e-indesign.</p>
<p>Lembro-me que passei momentos difíceis com arquivos corrompidos pelo Pagemaker, sendo obrigada a rediagramar do zero o livro já feito [...].</p> <p>https://labs.dualpixel.com.br/mulheres-e-indesign.</p>
<p>Formatar</p>
<p>Este exemplo demonstra como você pode formatar texto em um documento HTML.</p> <p>http://www.clem.ufba.br/tuts/html/c05.htm.</p>
<p>Ele nos pergunta como formatar o livro que ele acabou de criar para que tenha mais chances de ser lido pela turma da editora.</p> <p>https://carreirasolo.org/respostas/editorial/como-formatar-o-original-de-um-livro#YJA6K9VKjIU.</p>
<p>Formate o seu livro com a mais alta qualidade e menor custo do mercado.</p> <p>https://www.editorauniv.com.br/diagramacao-de-livro.</p>
<p>Não passe vergonhas formatando o seu livro no Word a não ser que ele seja impresso em impressoras comuns. https://www.editorauniv.com.br/diagramacao-de-livro.</p>
<p>[...] além de formatar o original [...] para um programa que seja aceito pelas impressoras offsets de gráficas profissionais, elevando a qualidade do material.</p> <p>https://www.editorauniv.com.br/diagramacao-de-livro.</p>
<p>Um diagramador de livros experiente irá formatar uma ou duas páginas do livro para sua aprovação.</p> <p>https://www.editorauniv.com.br/diagramacao-de-livro.</p>
<p>Não se deve, porém, formatar o livro com essa ferramenta [...].</p> <p>https://www.editorauniv.com.br/diagramacao-de-livro.</p>
<p>Quanto tempo é necessário para formatar / diagramar um livro?</p> <p>https://www.editorauniv.com.br/diagramacao-de-livro.</p>
<p>Com essa ferramenta, o escritor só precisa escrever, formatar o livro nos moldes do Kindle e publicar na Amazon [...]. https://tecnoblog.net/238106/como-publicar-um-livro-na-amazon/.</p>

<p>Antes de publicar um livro na Amazon é preciso formatá-lo, para que se enquadre nos formatos de leitura do Kindle e possa ser lido em outros aparelhos [...]. https://tecnoblog.net/238106/como-publicar-um-livro-na-amazon/.</p>
<p>O ideal é já formatar a página no tamanho da tela do Kindle (9 cm x 12 cm) e sem margens. https://contaumahistoria.com.br/2013/01/internet-ferramenta-eficaz-para-tornar-obras-publicaveis/.</p>
<p>[...] o ideal é formatar o livro com margens e com uma definição razoável (imaginando o livro como um livro físico realmente). https://contaumahistoria.com.br/2013/01/internet-ferramenta-eficaz-para-tornar-obras-publicaveis/.</p>
<p>Verdadeiramente é o melhor programa para planejar, escrever e formatar seu livro. https://melhores-aplicativos.com/3152/scrivener.</p>
<p>[...] basta compilar em um arquivo básico do Word e formatar o livro sozinho [...]. https://melhores-aplicativos.com/3152/scrivener.</p>
<p>Depois de escrito a coordenadora Raiza Tonon formatou o livro e devolveu as [sic] crianças para que ilustrassem. http://www.saosebastiao.sp.gov.br/noticia.asp?id=N12122018173946.</p>
<p>Meriellen editou esses desabafos que estava escrevendo durante a pandemia e formatou o livro “Crônicas de uma mãe desesperada” [...]. https://ocp.news/entretenimento/professora-jaraguense-lanca-livro-infantil-sobre-educacao-financeira.</p>
<p>Há alguns erros de digitação, ou de quando formatou o livro para o aparelho. [comentário] https://www.amazon.co.jp/-en/Prof-Carlos-Eduardo-Rocha/dp/1087219809.</p>
<p>Quando o texto é formatado com os outros elementos da linguagem HTML (sem utilizar o elemento PRE), é possível levar em consideração o tamanho da janela [...]. http://www.nce.ufrj.br/ginape/cursos/html/conteudo/textos/pre.htm.</p>
<p>Caso o escritor opte por formatá-lo por conta própria, a Amazon oferece um guia de formatação para e-books. https://mesadoescritor.com/como-publicar-um-livro-na-amazon/.</p>
<p>Primeiro, de um pequeno escritório de Desktop Publishing, a MAK Desktop Publishing, onde basicamente formatava trabalhos de graduação nas normas das faculdades. https://marciokarsten.pro.br/?p=11.</p>
<p>Utilizando análises e entrevistas, o livro foi formatado como uma linha do tempo que apresenta justamente a trajetória da animação brasileira. https://www.correiobraziliense.com.br/impresso/2017/07/2737421-100-anos-de--animacao.html.</p>
<p>Este livro foi formatado para simplificar o aprendizado de programação [sic] em computadores [...]. https://tudosobrelivro.com.br/livro-curso-de-programacao-em-linguagem-c#sobre.</p>
<p>Esse terceiro volume do livro foi formatado mais uma vez pela Plena Comunicação [...]. http://canindesoares.com/letras-e-imagens-do-bem.</p>

Arretche admitiu que o projeto do livro foi formatado em um momento de euforia com o crescimento da economia brasileira [...].

<https://agencia.fapesp.br/diminuicao-das-desigualdades-no-brasil-e-debatida-em-washington/22754/>.

O livro foi formatado com a colaboração de parceiras incansáveis [...]

https://issuu.com/editorarubio/docs/issuu_vitaminas__mineraiis_e_eletr_/11.

É importante formatar o seu manuscrito de acordo com os requisitos do periódico desejado [...].

<https://www.springer.com/br/authors-editors/authorandreviewertutorials/writing-a-journal-manuscript/formatting-your-manuscript/12011920>.

Isso pode lhe poupar muito tempo, já que você não terá de reformatar um artigo já escrito depois de selecionar o periódico!

<https://www.springer.com/br/authors-editors/authorandreviewertutorials/writing-a-journal-manuscript/formatting-your-manuscript/12011920>.



A gradação em português brasileiro e a variação translinguística na expressão da comparação

Degree constructions in Brazilian Portuguese and the translinguistic variation in the expression of comparative constructions

Luisandro Mendes de Souza

Universidade Federal do Paraná (UFPR), Curitiba, Paraná / Brasil

luisandro@ufpr.br

<http://orcid.org/0000-0002-4499-3820>

Roberta Pires de Oliveira

Universidade Federal de Santa Catarina (UFSC, CNPq), Florianópolis, Santa Catarina / Brasil

ropiolive@gmail.com

<http://orcid.org/0000-0002-4946-7205>

Lara Frutos

Universidade Estadual do Oeste do Paraná (UNIOESTE), Marechal Cândido Rondon, Paraná / Brasil

larafrutosg@gmail.com

<http://orcid.org/0000-0001-5456-3536>

Kayron Beviláqua

Instituto Federal de Santa Catarina (IFSC), São José, Santa Catarina / Brasil

kayronbevilaqua@gmail.com

<http://orcid.org/0000-0002-6627-8932>

Recebido em 31 de janeiro de 2022

Aceito em 09 de abril de 2022

Resumo: O artigo discute a hipótese dos parâmetros para expressão da comparação e das construções graduais analisando dados do português brasileiro (PB). Beck *et alii* (2009) e Beck, Oda e Sugisaki (2004) estipulam três parâmetros: i) a língua tem predicados graduais (DSP – Degree Structure Parameter); ii) a língua tem ligação de variáveis na sintaxe (DAP –Degree Abstraction Parameter); e iii) a língua tem a posição sintática de grau preenchida (DegPP – Degree Phrase Parameter). Línguas com marcação negativa no DSP não exibem qualquer construção gradual, enquanto as marcadas positivamente exibem uma variedade de construções graduais, especialmente comparativas tradicionais com algo semelhante a “mais”. O DAP dá conta de estruturas que envolvem ligação de variáveis de grau na sintaxe, como comparativas que exibem ambiguidades de escopo. Por fim, também há a impossibilidade de se preencher o argumento de grau dos adjetivos com sintagmas de medida (cf. *2 metros alto), um fato explicado pela marcação negativa no DegPP. Mostramos que o PB recebe marcação positiva nos dois primeiros parâmetros e negativa no último. Discutimos estruturas interrogativas graduais, bem como sintagmas de medida como *2 metros de altura*. O artigo conclui que essas estruturas precisam ser melhor discutidas para entendermos sua relação com os parâmetros.

Palavras-chave: Semântica; Universais Semânticos; Modificadores Graduais; Orações Comparativas.

Abstract: The paper discusses the hypothesis of parameters for the expression of comparison and degree constructions discussing data from Brazilian Portuguese (BP). Beck *et alii* (2009) and Beck, Oda and Sugisaki (2004) propose three parameters: i) the language has degree predicates (DSP – Degree Structure Parameter); ii) the language has or does not have binding of degree variables in the syntax (DAP –Degree Abstraction Parameter); and iii) the language has a syntactic degree argument filled (DegPP – Degree Phrase Parameter). Languages with negative setting in the DSP do not show any degree construction at all, especially comparative constructions with an expression like “more”. The DAP explains structures with binding of degree variables in the syntax, like comparative clauses with scope ambiguities. At last, there is the impossibility of filling the syntactic degree argument of gradable adjectives with measure phrases (cf. *2 metros alto/2 meters tall). This fact is explained by a negative setting in the DegPP. It is shown that BP receives positive setting in the first ones but negative setting in the last. It is also discussed degree interrogatives and measure phrases like *2 metros de altura* “2 meters of height”. The paper concludes that these structures should be better discussed to understand their place in the degree parameter hypothesis.

Keywords: Semantics; Semantic Universals; Degree Modifiers; Comparative Clauses.

1 Introdução

Dentro da abordagem referencial do significado, a busca por universais semânticos alcança novas fronteiras dia a dia. Dos universais na expressão da quantificação (cf. BARWISE; COOPER, 1981; CHIERCHIA, 2003; e referências lá citadas) aos parâmetros no domínio nominal (cf. CHIERCHIA, 1998; ROTHSTEIN; PIRES DE OLIVEIRA, 2011; entre outros), essa busca também atingiu o domínio dos predicados graduais e da expressão da comparação com os trabalhos de Beck, Oda e Sugisaki (2004) e Beck *et alii* (2009).

Nesse domínio, desde os trabalhos tipológicos de Ultan (1972) e de Stassen (1984), outros estudos (e.g. BOCHNAK, 2013; KENNEDY, 2007) fizeram considerações universalizantes, buscando compreender por que as línguas expressam a relação de comparação de forma diversa e por que, pelo que sabemos, nem todas as línguas aparentam ter modificadores graduais (como advérbios de gradação/intensificadores ou interrogativas graduais).

Expressar alguma forma de comparação parece ser universal, mas o fato é que nem todas as línguas o fazem através de um item lexical ou de um morfema que signifique alguma coisa como “mais”. Por exemplo, em português brasileiro (PB) e inglês temos o *mais ... (do) que, -er/more ... than*, ou seja, um item expressa a relação (de superioridade, inferioridade ou igualdade) e outro introduz o termo comparado – o padrão das línguas indo-europeias. Assim, a falta de um item linguístico que signifique “mais” ou “menos” coloca em questão a universalidade desse conceito.

Além disso, o estatuto sintático do constituinte *do que X* é matéria de discussão. Há quem defenda que estamos diante de uma conjunção subordinada (adverbial ou relativa) e há quem defenda que estamos diante de uma estrutura de coordenação¹. No plano semântico o valor de verdade da oração principal é dependente do valor de verdade da subordinada; o que poderia ser um argumento a favor da subordinação, já que na coordenação eles são independentes. Em (1b), por exemplo, temos uma sentença comparativa formada pelo adjetivo relativo *alto*, em que não há acarretamento de que João e Beto sejam altos ou baixos. Para que (1b) seja verdadeira, é preciso que João esteja acima de Beto na escala de altura. O mesmo pode ser afirmado para a sentença do inglês em (1a).

¹ cf. Marques (2003); Brito e Matos (2003); Souza (2010) e Dinis (2018) para discussão sobre o português.

- (1) a. John is taller than Bill.
 b. João é mais alto (do) que Beto.

Ampliando o escopo empírico, há várias línguas no mundo em que a comparação se dá em estruturas como (2a), do motu (Papua Nova Guiné), e (2b), do hixkaryana (Brasil, AM – família caribe), que envolvem coordenação no plano sintático, e, aparentemente, também no semântico. No plano estrutural, vemos algo muito similar às línguas indo-europeias, especialmente a presença de predicados graduais e o estabelecimento de uma relação entre dois indivíduos; contudo, a diferença está no fato de os predicados serem de polos opostos da mesma dimensão, como em (2a), ou a afirmação de um e a negação de outro, como em (2b). E, especialmente, a falta de um item que expresse “mais”.

- (2) a. *Mary na lata, to Frank na kwadoḡi.* (BECK et alii, 2009, p. 4)
 Mary TOP alto, mas Frank TOP baixo
 b. *Kaw-ohra naha Waraka, kaw naha Kaywerye* (KLEIN, 1991, p. 675)
 alto-não ele-é Waraka, alto ele-é Kayweryve

Partindo desse cenário, Beck *et alii* (2009) compararam mais de uma dezena de línguas, tendo num extremo o inglês, como a língua com o maior número de estruturas graduais (nesse conjunto está também o alemão) e no outro o motu, que aparenta não ter nenhum tipo de modificação gradual ou estrutura de comparação, embora tenha predicados vagos e dependentes de contexto. Beck *et alii* (2009) hipotetizaram três parâmetros para explicar por que nem todas as línguas investigadas apresentavam as estruturas atestadas em inglês. Assim, nossa meta nesse trabalho é investigar como o português brasileiro (PB, doravante) se comporta em relação a esses parâmetros, contribuindo para a discussão sobre as estruturas graduais.

Os parâmetros são divididos em três conjuntos e as línguas podem apresentar valor positivo ou negativo para eles.

- a) língua com grau: e com grau disponível para operações sintáticas como quantificação;
 b) língua com grau: mas sem possibilidade de preenchimento da variável de grau;
 c) línguas sem grau: a língua não tem nenhuma estrutura gradual tradicional.

Beck *et alii* (2009) fazem algumas suposições teóricas, assumindo que graus são entidades primitivas da ontologia, ao lado de indivíduos e valores de verdade.²

Para cumprir nosso objetivo principal, a primeira seção do artigo apresenta as estruturas graduais típicas no inglês e no japonês, comparando-as com o PB. A segunda seção apresenta os parâmetros. Por fim, trazemos considerações sobre os testes e a comparação com outros estudos que replicaram e eventualmente questionaram a hipótese inicial dos parâmetros, como Sanchez-Mendes (2016) fez para o caritiana e para o PB.

2 Inventário das estruturas graduais e comparativas

2.1 Rudimentos da semântica dos predicados graduais

Se graus existem como entidades na ontologia das línguas é uma questão aberta na semântica contemporânea (MORZYCKI, 2016). A tipologia básica na semântica formal (cf. HEIM; KRATZER, 1998; entre outros) prevê dois domínios básicos, o das entidades, D , de tipo e e o dos valores de verdade, de tipo t ; e a partir daí a existência de funções de mapeamento de um domínio ao outro, funções de tipo $\langle et \rangle$, $\langle e, \langle et \rangle \rangle$ etc. Por exemplo, *brasileiro* é um tipo de função que tem como domínio os indivíduos do universo de discurso e como contra-domínio os valores de verdade, $\{0,1\}$, isto é, uma função de tipo $\langle e, t \rangle$. Posteriormente outros domínios foram sendo estipulados, como domínios para eventos D_v , no estudo dos predicados de eventos, e mundos D_s , no estudo da modalidade.

Analisar predicados graduais como *alto/baixo* ou *perto/longe* como simples funções de indivíduos a valores de verdade não captura uma série de propriedades linguísticas dessas expressões, como a vagueza, a dependência contextual, a polaridade e as diversas modificações que podem sofrer. Para dar conta dessa variedade, alguns autores (BURNETT, 2014; KLEIN, 1980 entre outros) propuseram que esses predicados são funções parciais e que denotam domínios organizados, enquanto outros (BECK, 2011; HEIM, 2001; KENNEDY, 1997; 2007; von STECHOW,

² Abordagens alternativas, como Klein (1982), Burnett (2014), *inter alia*, assumem que adjetivos são funções parciais de indivíduos a valores de verdade, sem propor que entidades como graus façam parte da ontologia. Mais detalhes na seção 1.1.

1984) assumem que essa classe de predicados denota relações entre indivíduos e graus, ou seja, postulando entidades de tipo *d* na ontologia. Em (3) temos a formalização da entrada lexical de *alto* de acordo com cada visão. Em (3a), *alto* denota o conjunto daqueles indivíduos que estão na extensão positiva de *alto* no contexto dado. Ao mesmo tempo, o adjetivo também tem uma extensão negativa (aqueles indivíduos que não são altos na situação) e uma lacuna extensional (aqueles indivíduos que não são nem altos nem baixos). Em (3b), o adjetivo é uma relação entre um indivíduo e um grau na escala de altura. Embora em (3c) não tenhamos um grau explícito na entrada lexical, o adjetivo é essencialmente uma relação entre um indivíduo e uma posição na escala de altura, um grau.

- (3)a. $alto = \text{pos.alto}(x)$ (KLEIN, 1980)
 b. $alto = \text{ALTO}(x) \geq d$ (von STECHOW, 1984; HEIM, 2001)
 c. $alto = \text{ALTURA}(x)$ (KENNEDY, 1997, 2007; KENNEDY; MCNALLY, 2005)

Essas análises têm como objetivo principal dar conta da semântica de orações simples como (4a) e comparativas canônicas, como (4b), bem como da relação composicional entre elas. Afinal, esperamos que a semântica assumida para os adjetivos em orações simples seja a mesma nas suas variações modificadas por outras construções.

- (4)a. João é alto.
 b. João é mais alto (do) que Beto.

Abaixo comparamos a análise de Klein (1980) com a de Kennedy (2007) para as estruturas em (4), simplificando algumas suposições. (5a), partindo do visto acima, é verdadeira se e somente se João está na extensão positiva de *alto*. Note que a comparação nessa abordagem precisa que os indivíduos estejam dispostos em duas extensões diferentes. Assim, (6a) é verdadeira se e somente se João está na extensão positiva de *alto* e Beto na extensão negativa. Na abordagem de graus, em (5b), temos a formalização da intuição de que para que *João é alto* seja verdadeira é preciso que sua altura ultrapasse o padrão para ser considerado alto no contexto. Já em (6b), *João é mais alto que Beto* é verdadeira se e somente se João está acima de Beto na escala de altura.

- (5) a. *João é alto* = pos.alto(j) (KLEIN, 1980)
 b. *João é alto* = ALTURA(j) \geq d_{padrão em C} (KENNEDY, 2007)
- (6) a. *João é mais alto que Beto* = pos.alto(j) & neg.alto(b)
 b. *João é mais alto que Beto* = ALTURA(j) > ALTURA(b)

A literatura tem argumentado que a opção por graus tem maior cobertura empírica (ver BECK, 2011; KENNEDY, 1997, 2007), mas há novas tentativas de aplicar uma análise kleiniana a fenômenos que desafiam a abordagem de delineação, especialmente a polaridade adjetival, os adjetivos absolutos, ambiguidades de escopo e o escopo de quantificadores (BURNETT, 2014).

2.2 Inventariando as estruturas

Como dissemos, as estruturas que as línguas exibem, e que envolvem algum tipo de manipulação semântica sobre o predicado gradual, são muito variadas. Beck *et alii* desenvolveram um questionário que objetiva mapear o inventário de construções que a língua exhibe. Tomando o inglês como parâmetro, as seguintes estruturas mostrariam que temos construções que operam sobre o grau. Vamos aproveitar e introduzir as estruturas equivalentes no PB, conforme avançamos. Todos os exemplos do inglês e do japonês são de Beck *et alii*.

Traremos também alguns dados do japonês comentados por Beck *et alii* (2009), pois foi a partir de um estudo prévio (BECK; ODA; SUGISAKI, 2004) que surgiu o questionamento sobre as diferenças entre as línguas. A comparativa canônica em japonês tem a estrutura em (7) e a interpretação parafraseada logo abaixo. Beck, Oda e Sugisaki (2004) propõem que a interpretação mais adequada seria algo próximo a (7b) e não uma comparação padrão, como representada no PB em (7c).

- (7) a. Sally-wa Joe-yori se-ga takai.
 Sally-Top Joe-yori atrás-Nom alto
 b. Comparada ao Joe, Sally é mais alta.
 c. Sally é mais alta do que Joe.

- Interrogativa gradual

A interrogativa questiona o grau que o indivíduo exhibe da propriedade ou a sua posição na escala, como vemos na parafrase da interrogativa (8a) em (8b). Em (8c) estão as condições de verdade da

interrogativa, isto é, o conjunto de respostas possíveis para a pergunta em (8a). Em (8d) vemos que a interrogativa gradual no japonês não permite a mesma estrutura.

- (8) a. How tall is Captain Apollo?
 b. Para qual d: o Cap. Apollo é d-alto.
 c. {CA is 6ft tall; CA is very tall; CA is taller than me, ...}
 d. *John-wa dore-kurai kasikoi no?
 John-Top qual-grau esperto Q

Quanto ao PB, parece haver interrogativas graduais como o inglês, mas a questão a ser investigada é se ela é interpretada da mesma forma que sua contraparte no inglês. A resposta é parcialmente positiva. Enquanto em inglês o conjunto de respostas de (8c) pode envolver um sintagma de medida, no PB não pode (cf. 9c). Talvez esse seja um problema menor, mas interrelacionado, já que poder preencher o argumento de grau com uma expressão que denota um grau não é uma possibilidade atestada amplamente nas línguas.

- (9) a. Quão alto é o João?
 b. Para qual d: o João é d-alto.
 c. {O João é muito alto; O João é bem alto; O João é mais alto do que eu, ...}

Outro ponto importante é que essa não é a estrutura preferencial para se formular a interrogativa no PB coloquial. Tendemos a usar (10a) ou (10b). Poder-se-ia argumentar que as perguntas em (9a) e (10a-b) são distintas apenas superficialmente; no entanto, supondo que as condições de verdade de uma interrogativa sejam o conjunto das suas respostas, as sentenças em (9a) e as em (10) não podem ser vistas como variantes estilísticas. As interrogativas em (10a-b) só podem ter uma resposta que seja um intervalo, ou um sintagma de medida, como (10c). Já (9a) pode ser respondida com qualquer uma das opções em (9c), menos com um sintagma de medida. *O João tem 1,75m de altura* não pode ser uma resposta para (9a) e pode para (10a-b). Assim, é como se essas interrogativas tivessem respostas que estão em distribuição complementar.

- (10) a. Quanto que o João tem de altura?
 {quanto é a medida d na escala de altura: João tem d de altura}

- b. Qual é a altura do João?
 {o único grau d: João tem d de altura}
 c. O João tem 1,80m de altura.

No inglês, a estrutura da interrogativa gradual é relativamente uniforme, independentemente do predicado, mas no PB não. Por exemplo, como idade é um parâmetro contado numericamente, a interrogativa sobre a idade no PB é feita com *quantos*. No inglês ela é feita com *how*. A estrutura típica para interrogativas utilizando tal escala no PB é (11b). Comparando com (11c), note que ela interroga sobre qual o intervalo de graus em que um indivíduo está inserido na escala positiva de ‘idoso’ (muito? pouco?), implicando que o indivíduo é idoso. O mesmo vale para (11d). Isso é diferente da pergunta (11b) em que não há a pressuposição desse tipo.³

- (11) a. How old are you?
 b. Quantos anos você tem?
 c. Quão idoso você é? [buzzfeed.com]
 d. Quão jovem com alma de velho você é? [buzzfeed.com]

Embora dados como (11c) e (11d) não sejam tão amplamente utilizados na fala vernacular e parecem não ser aceitos universalmente por falantes de PB, encontramos diversos registros de sentenças similares e consultamos informantes que aceitam seu uso. Não se trata, portanto,

³ Um dos pareceristas sugeriu que embasássemos essa afirmação. Pelo contraste entre o termo negativo e positivo de um par de adjetivos polares, percebemos que o adjetivo negativo é marcado na interrogativa. Isso não vale de forma geral. Compare os pares abaixo. No par *alto/baixo*, *baixo* é a forma marcada e a interrogativa (ib) pressupõe que João é baixo. Agora no par *quente/frio*, embora denotem regiões diferentes de uma mesma escala, temperatura, ambos trazem o pressuposto de que a sopa está fria em (iia) e quente em (iib).

- (i) a. Quão alto João é? (ii) a. Quão fria está essa sopa?
 b. Quão baixo João é? b. Quão quente está essa sopa?

Como estamos diante de uma interrogativa, é difícil testar sua pressuposição. Podemos usar, nesse caso, uma versão do teste de von Stechow (2004), que adaptaremos para *perai* (‘wait a minute’, no original). (ia) pode ser respondida com *muito alto*, *bem alto*, *não muito alto*, *bem baixinho* etc.; já (iib) não pode ser respondida com *muito alto*. Além disso, se um interlocutor responde a essa interrogativa com *Peraí! Eu não sabia que ele era baixo* demonstra que não faz parte do seu ‘common ground’ uma proposição que parece ser verdade compartilhada entre os outros interlocutores. Autores como Rett (2020) assumem que interrogativas como (ib) trazem uma pressuposição, embora sem muita discussão; a ideia também aparece em Cruse (1986).

apena de uma tradução literal do inglês, mas de uma possibilidade estrutural do PB. Ou seja, podemos concluir que essa diferença se deva apenas a como as línguas lexicalizam as escalas. O PB opta por uma nominalização nas interrogativas, primordialmente, pois o uso do adjetivo para fazer a interrogativa é marcado com leituras específicas.

- Sintagmas de medida

Sintagmas de medida são entendidos como expressões que denotam um grau, ou uma posição na escala. Como a literatura mostra (CORVER, 2009; SCHWARZSCHILD, 2006), há muita variabilidade no preenchimento do argumento de grau de predicados graduais, especialmente se considerarmos que isso depende de termos medidas culturais para os graus das escalas. Mesmo assim, há línguas que não permitem essa possibilidade, caso do PB. A construção no japonês, embora bem formada, tem apenas a interpretação de uma comparação implícita, não de que 5cm seja a altura do indivíduo em (12c).

- (12) a. Captain Apollo is exactly 1.74m tall.
 b. O grau máx d tal que Cap. Apollo é d-alto é 1,74m.
 c. Sally-wa 5cm se-ga takai.
 Sally-wa 5cm atrás-Nom alto.
 Sally é 5cm mais alta/*Sally é 5cm alta.
- (13) a. *João é 1,75m alto.
 b. João tem 1,75m de altura.

Diferentemente do inglês, no PB o argumento sintático-semântico do adjetivo não pode ser preenchido por um sintagma de medida, como em (13a). Contudo, podemos supor que a nominalização em (13b) tenha condições de verdade similares à sentença inglesa em (12a), qual seja: o grau máximo d tal que João é exatamente d-alto é 1,80m. O que diferencia as duas sentenças nesse caso é como essa proposição é expressa sintaticamente. Acreditamos que esse é um tema para investigação potencial e que o parâmetro não consegue explicar totalmente, já que embora o PB requeira uma estrutura com nominalização como em (13b), o inglês também exhibe estruturas com nominalizações de escalas (cf. *This man has 2m of height*).

Não há uma explicação consensual para essa proibição. Já que o argumento está disponível, por que línguas como o PB (e as outras românicas) não permitem que esse argumento seja preenchido? A resposta de Beck *et alii* é a estipulação de um traço ligado a um dos parâmetros, como veremos na próxima seção.

- Comparativas de igualdade

As estruturas comparativas de igualdade, também chamadas de equativas, são possíveis no PB, assim como estruturas similares em inglês. Do ponto de vista interpretativo, o ponto importante é que em inglês não temos acarretamento para o positivo, (14a) não acarreta que Starbuck e Captain Apollo sejam altos. O mesmo acontece com a sentença em (15).

- (14) a. Starbuck is as tall as Captain Apollo is.
b. O grau máximo de altura que a Starbuck possui é pelo menos tão alto quanto o grau máximo de altura que o Capitão Apollo possui.

- (15) O João é tão alto quanto o pai dele.

- Superlativo relativo

Comparando as sentenças em (16a) e (17), o superlativo no PB tem estrutura diferente do inglês, mas as mesmas condições de verdade. Embora o superlativo relativo no PB seja feito com *mais*, a interpretação é de que estamos comparando implicitamente a altura de João com a de outros membros de algum conjunto implícito dado contextualmente. A diferença dessa para as estruturas de comparação canônicas é que o sintagma [*o mais A*] é uma descrição definida, dando a esse sintagma uma função identificacional, como aponta Marques (2003).

- (16) a. Helo is the tallest.
b. O grau máximo de altura que Helo possui excede o grau máximo de altura que qualquer outra pessoa relevante possui.

- (17) João é o mais alto (da turma/da família).

- Comparativas diferenciais

Abaixo temos os exemplos de comparativas diferenciais do inglês em (18a) e do japonês em (18c). O PB está em (19). A principal característica dos diferenciais é medir a diferença entre a altura de dois indivíduos que estão sendo comparados em uma comparativa de inferioridade ou de superioridade. Note que no PB também podemos usar um sintagma de medida nessa posição, bem como outros graduadores (como *muito*, *pouco*, *um pouco*, *bem* etc.).

- (18) a. Helo is 8 cm taller than Starbuck is.
 b. O grau máximo de altura que Helo possui é 8cm maior do que o grau máximo de altura que Starbuck possui.
 c. Sally-wa Joe-yori 5cm se-ga takai.
 Sally-Top Joe-yori 5cm atrás-Nom alto
Sally é 5cm mais alta do que Joe
- (19) O João é {5cm/bem/muito} mais alto do que eu.

- Outras comparativas

As estruturas comparativas também podem ser atributivas (20a), adverbiais (20b) ou oracionais (20c). As estruturas equivalentes no PB estão em (21). Note que as orações em (a) comparam quantidades, não graus no sentido tradicional. Por sua vez, as orações em (b) tem como predicado de comparação um advérbio, embora envolvam elipse de um sintagma verbal, enquanto em (c) vemos que o constituinte complemento do marcador do padrão é explicitamente oracional.

- (20) a. Mr Bingley keeps more servants than Mr Bennet does.
 b. Colonel Fitzwilliam behaved more amiably than his cousin did.
 c. Colonel Fitzwilliam behaved more amiably than I had expected.
- (21) a. Brás Cubas tem mais serviçais do que Quincas Borba.
 b. Brás Cubas se portou mais amigavelmente do que seu primo.
 c. Brás Cubas se portou mais amigavelmente do que Virgília esperava.

Outra construção importante é a chamada subcomparativa, exemplificada em (22a). Esse tipo de comparativa tem como característica envolver dois predicados diferentes de escalas similares, mas a comparação, na verdade, é entre a diferença que cada um exibe em relação ao grau padrão positivo de cada predicado, interpretação que a paráfrase logo abaixo busca capturar. O japonês não tem essa construção, cf. (22b). A estrutura no PB está em (23).

(22) a. Your shoes are longer than this cupboard is deep.

O grau máximo de comprimento do teu sapato ultrapassa o grau máximo de profundidade do balcão.

b. *Kono tana-wa [ano doa-ga hiroi yori (mo)] (motto) takai.

this shelf-Top [aquela porta-Nom largo yori (mo)] mais alto

(23) Essa estante é mais alta do que a porta é larga.

Aqui nossa intuição variou. Houve quem julgasse a sentença em (23) na interpretação pretendida como, no mínimo, estranha, mas a maioria de nós a julgou gramatical e interpretável. De qualquer modo, seria importante a testagem desse tipo de construção no PB para maior clareza de sua gramaticalidade e quais interpretações são possíveis. Vale dizer, ademais, que Sanchez-Mendes (2016) também julgou a sentença gramatical. Vamos assumir que ela tem condições de verdade como aquelas atribuídas para (22a): comparamos o quanto cada indivíduo desvia do padrão para estar no intervalo positivo de cada predicado. Ou, em outras palavras, seguindo o padrão de (22a): o grau máximo de altura da estante ultrapassa o grau máximo de largura da porta.

- Interação com outros operadores

De acordo com a visão tradicional, a oração comparativa denota uma descrição definida de graus, *o grau máximo...*, é essa expressão que em tese deveria entrar em interação de escopo com outros operadores que possuem força quantificacional. Para Heim (2001), (24) é o tipo de exemplo que nos mostra que comparativas podem entrar em interação de escopo com outros operadores; nesse caso, a expressão modal *is required*.⁴

⁴ Ver também Morzycki (2016) e Beck (2011).

(24) This draft is 10 pages long. The paper is required to be exactly 5 pages longer than that.

a. O tamanho que o artigo tem em todos os mundos que preenchem os requisitos é 15 páginas. = o tamanho mínimo necessário para o artigo é 15 páginas.

b. Em todos os mundos que preenchem os requisitos, o tamanho do artigo é 15 páginas. = o artigo precisa ter exatamente 15 páginas.

A sentença tem duas leituras. A relevante é a parafraseada em (24a). Ela é a leitura importante porque nos mostra que a descrição definida de graus que a oração comparativa *than X* denota pode ter escopo sobre a expressão modal *is required*. Qualquer situação em que o artigo tem mais de 15 páginas já é uma situação que torna essa leitura verdadeira. Já a leitura (24b) só é verdadeira se o artigo tiver exatamente 15 páginas e não mais do que isso. Em todas as situações é o caso que há um único tamanho para o artigo, 15 páginas.

Esse é um tema potencialmente controverso, pois mesmo para o inglês os julgamentos sobre a existência da ambiguidade não são unânimes. A nossa intuição em relação ao equivalente no PB, em (25), também variou. Contudo, para a maioria de nós a leitura principal é a equivalente a (24a) para o exemplo em inglês. Reforçando, o fato crucial aqui é que a descrição definida parece ter escopo acima da expressão modal *tem que*.

(25) Esse manuscrito tem 10 páginas. O artigo tem que ser exatamente 5 páginas mais longo do que isso.

No caso do japonês, a leitura de escopo amplo da descrição definida não existe para o exemplo (26):

(26) Sono ronbun-wa sore yori(mo) tyoodo 5 peeji nagaku-nakerebanaranai.

Esse artigo-Top que yori(mo) exatamente 5 páginas longo-necessário

Esse artigo tem que ser exatamente 5 páginas mais longo do que isso.

- Efeito da ilha negativa

Um fenômeno importante que as comparativas em inglês exibem é o chamado efeito da ilha negativa. Basicamente, a presença de uma negação sentencial dentro da oração sob *than* torna a sentença agramatical,

como vemos em (27a). Mas no japonês uma estrutura equivalente é gramatical, como temos em (27b), embora, como a interpretação sugere, as condições de verdade sejam diferentes do inglês. A tradução da mesma sentença no PB é agramatical, cf. (28).

- (27) a. *Mary bought a more expensive book than no boy did.
b. John-wa [dare-mo kawa-naka-tta no yori] takai hon-o katta.
John-Top [qualquer comprar-Neg-Past No yori] caro livro-Acc comprar
Mary comprou um livro mais caro do que o livro que nenhuma outra pessoa comprou.

- (28) *Maria comprou um livro mais caro do que nenhum garoto comprou.

Esse efeito é explicado por abordagens que usam graus postulando que o sintagma que traz o padrão de comparação denota um conjunto de graus que não serve de padrão para a comparação, pois não há um grau máximo tal que não é o caso que existe um garoto que comprou um livro d-caro. Essa é a denotação do sintagma, como vemos em (29). Como esse grau não tem “referente”, a comparação é anômala.

- (29) *than no boy did* = λd . nenhum garoto comprou um livro d-caro.

- Forma positiva

A forma mais simples do uso dos adjetivos graduais é o seu uso como predicado nominal, como temos em (30-31).

- (30) a. Hello is tall.
b. A altura da Hello excede o padrão contextual de altura.

- (31) João é alto.

A importância dessa sentença simples está na sua dependência contextual e na sua vagueza. Embora relacionados, esses dois aspectos são diferentes. Compare com sentenças com indexicais, por exemplo, que precisam de informação contextual para que alcancemos a proposição

que elas expressam, mas que, uma vez resolvida a indexicalidade, deixam de lançar mão do contexto. Algo similar acontece em orações como (30-31): o valor de verdade delas depende do que julgamos como *alto* no contexto em que a sentença é proferida e o que ela expressa é que o sujeito ultrapassa esse padrão contextual, que é estabelecido através da consideração de uma classe de comparação (isto é, o conjunto de indivíduos contra os quais comparo a altura do sujeito da oração). Além da dependência contextual, outra característica da vagueza são os casos limítrofes, isto é, casos em que podemos vacilar se o predicado se aplica ou não ao indivíduo (KENNEDY, 2007).

Na prática é o que ocorre com sentenças similares em línguas como o motu, que não têm modificadores, mas o valor de verdade de uma sentença como *x é alto*, em motu, também depende de um valor contextual.

Feita essa apresentação das estruturas relevantes, passemos à discussão dos parâmetros.

3 Os parâmetros

A comparação entre o inglês e o japonês levou Beck, Oda e Sugisaki (2004) a criar a hipótese de que um parâmetro explicaria as diferenças atestadas pelos autores entre as línguas. Embora sejam três os parâmetros, o mais importante é o parâmetro da estrutura gradual (DSP): a língua tem ou não tem predicados graduais. O motu não tem qualquer tipo de comparativa usual, nem modificadores graduais (como intensificadores), portanto recebe marcação negativa. Já o inglês e o japonês apresentam marcação positiva. Contudo, há ainda diferenças entre as línguas, como vimos. As diferenças são então explicadas por outros dois parâmetros: o parâmetro da abstração dos graus (DAP) e o parâmetro do sintagma de grau (DegPP). Vamos a eles e às estruturas envolvidas.

DAP: Parâmetro da abstração dos graus (Degree abstraction parameter) – A língua (não) tem ligação de variáveis na sintaxe.

A falta de um sistema de ligação de variáveis de grau explicaria a ausência da leitura de escopo amplo no japonês, a falta do efeito da ilha negativa e a interrogativa de grau. É menos claro para nós que o parâmetro explique também a ausência das estruturas subcomparativas (de comparação de desvio) e a ausência de sintagmas de medida preenchendo o argumento de grau, pois essa explicação depende de se

supor que a estrutura no japonês não permita comparar dois conjuntos de graus (de dimensões diferentes, é importante notar) e que o sintagma de medida denote algum tipo de quantificação sobre graus e não apenas seja ele mesmo um grau.

A investigação translinguística objetiva verificar se outras línguas apresentam o inventário de estruturas que o inglês apresenta. Assim, esse levantamento tem três partes:

- (a) Inventário das estruturas I: graus tradicionais do adjetivo (positivo, comparativo de superioridade, inferioridade e igualdade e superlativo), interrogativas graduais e outros modificadores, como intensificadores e construções de excesso (como *demais*).
- (b) Inventário de estruturas II: verifica se a língua tem estruturas de comparação atributiva, adverbial e oracional, pois se a língua não tem comparativas oracionais não faz sentido questionar se existem o efeito da ilha negativa e a estrutura de comparação de desvio.
- (c) Inventário de estruturas III: presença de graus explícitos.

A questão subjacente a toda a investigação é a manipulação de graus. Nesse sentido, é importante verificar se a língua pode fazer comparação com um grau explícito e se tem diferenciais, como temos em (32-33).

- (32) a. Captain Apollo is taller than 1.70m
b. Helo is 8cm taller than Starbuck is.

- (33) a. Pedro é mais alto do que 1,70cm.
b. Pedro é 5cm mais alto do que Gabriel.

Note que, embora (33a) seja gramatical, a estrutura preferida para fazer essa comparação no PB é (34). Contudo, como comentamos para os casos com sintagmas de medida, podemos supor que (34) tem as mesmas condições de verdade que (33a).

- (34) Pedro tem {mais do que/de} 1,70m (de altura).

Assim, a investigação empírica translinguística originou o parâmetro principal, formulado como abaixo:

DSP (Degree semantics parameter) – a língua (não) tem predicados graduais, ou predicados que introduzam graus.

Esse parâmetro corresponde à existência de estruturas típicas para expressão de grau (como as listadas no inventário I em (a) acima, nesta seção). Se a língua tem predicados graduais, esse grau está disponível para ser modificado ou ligado por essas expressões que estabelecem relações entre graus, como a comparação e outros tipos de modificadores graduais. Nesse aspecto, o item (c) também é relevante, pois verifica a presença de estruturas que manipulam graus.

Motu é o exemplo típico da marcação negativa do parâmetro. É uma língua com comparação expressa via conjunção, como vimos. Ela não tem qualquer item para realizar modificação gradual, muito menos comparativas com diferenciais ou com um grau como o padrão. (35) exemplifica a estrutura agramatical, com o significado pretendido de “Maria tem mais de 1,70m de altura”.

(35) * Mary na lata 1.70m.
Mary TOP alta, 1.70m

Beck *et alii* sugerem que os predicados como *lata* ‘alto’ em motu devem ser analisados como se analisam preposições locativas lexicais, que, embora sejam vagas e dependentes de contexto, não são predicados com um argumento semântico de grau. Como sugerem: “Perhaps degrees and scales are a level of abstraction above context dependency that a language may or may not choose to develop.” (BECK *et alii*, 2009, p. 28)

Há línguas que apresentam um verbo significando algo como “exceder”. Em yorubá (África ocidental) vemos uma comparativa típica. Mas essa língua também tem comparações com graus, que vemos em (36-37), e diferenciais nas orações comparativas, em (38a) e (38b).

(36) a. Owó osù rè ju ti e lo
dinheiro mês dela excede que teu ir
O salário dela excede o teu salário.
b. Ade ga jo Isaac lo.
Ade é alta excede Isaac ir
Ade é mais alta do que Isaac.

- (37) a. *Kathy fi esebata kan ga ju Sandra lo.*
 Kathy com pé um alta excede Sandra ir
Kahty é um pé mais alta do que Sandra.
- b. *Kathy ga ju esebata marun ataabo lo.*
 Kathy é alta exceed pé cinco e meio ir
Kathy é tem mais de cinco pés e meio.

Embora o operador de comparação seja um verbo, isso não quer dizer que não tenhamos graus. Contudo, Beck *et alii* supõem que o verbo *ju* denote uma relação de superioridade. Lembrando que o crucial é a presença do diferencial e da comparação de grau, para que a língua tenha a marcação positiva no parâmetro.

DegPP (Degree Phrase Parameter) – O argumento de grau do predicado pode estar (ou não) explicitamente preenchido.

Esse parâmetro não quer dar conta da presença de sintagmas de medida como argumentos semânticos/sintáticos do predicado gradual. Na verdade, sua elaboração é mais sintática, no sentido que está preocupado com o preenchimento da posição de especificador do sintagma adjetival (SA). Essa posição pode ser preenchida por operadores na sintaxe encoberta, que é o que acontece com as interrogativas graduais (preenchida por um *Qu-*) e nas subcomparativas (preenchida por um operador). Ambos são elementos que se movem em forma lógica (FL).⁵ Em (38) vemos os operadores na posição em que devem nascer na estrutura, como especificadores. Em (38a) o operador é uma expressão interrogativa que se move de forma visível; em (38b) o especificador do SA está preenchido pelo *mais* na oração principal e por um operador (OP) na subordinada.

- (38) a. [_{SA} *quão* [_A *alto*]]
 b. A mesa é mais alta do [que a porta é [_{SA} OP [_A *larga*]]]

⁵ Lembrando a distinção entre esses dois tipos de movimento na abordagem gerativa: movimentos visíveis são movimentos com consequências fonológicas, como o movimento de expressões interrogativas; enquanto o movimento encoberto é o movimento de quantificadores e operadores em forma lógica, sem consequências visíveis na superfície sonora.

Esse parâmetro objetiva explicar por que línguas como o russo não apresentam estruturas como interrogativas graduais, subcomparativas e sintagmas de medida.

Não temos uma ampla discussão desses aspectos na literatura sobre o PB. Mas assumindo, com Martinho (2007), que o sintagma de grau (DegP), ou nesse caso o especificador do SA, também pode ser preenchido na sintaxe pelos operadores em discussão, nossa língua teria marcação positiva no parâmetro, pois vimos que as estruturas em questão são possíveis.

A tabela abaixo resume alguns exemplos típicos de línguas, algumas mencionadas aqui. O inglês está num extremo da tabela (possuindo todas as estruturas) e o motu no extremo oposto (sem nenhuma), considerando apenas um subconjunto das estruturas graduais, aquelas que se mostraram relevantes para estabelecer as diferenças entre as línguas. Os subconjuntos das estruturas são aqueles que mostraram se comportar de forma parecida.

Tabela 1 – Resumo dos parâmetros nas línguas

Língua	Dif	CGrad	Escopo	Ilha-N	Qu-G	SM	SubC
Inglês	sim	sim	sim	sim	sim	sim	sim
Alemão	sim	sim	sim	sim	sim	sim	sim
Espanhol	sim	sim	sim	sim	(não)	(não)	(não)
Guarani	sim	sim	sim	sim	não	não	não
Chinês	sim	sim	não	não	não	não	não
Motu	não	não	não	n.a.	não	não	n.a.
PB	sim	sim	sim	sim	não	não	(sim)

Fonte: Beck et alii (2009) (adaptado com a inclusão do PB)

(): dúvida; n.a.: não se aplica

Note que os dados cruciais envolvem três estruturas que geraram algumas dúvidas sobre a aceitabilidade no PB (interrogativas graduais e subcomparativas) estão organizados dentro de um mesmo grupo. Além disso o PB não permite o preenchimento do argumento sintático/semântico de grau com um sintagma de medida, o que nos faria pensar que nossa língua está próxima do inglês nesse parâmetro. Note que em relação ao espanhol, língua de mesma família que o PB, há um “(não)”, representando que há dúvida sobre a aceitabilidade dos dados. O espanhol

permite, como o PB, estruturas subcomparativas. Os dados abaixo são de Beck *et alii* (2009, p. 57, nos apêndices). Note que a estrutura gramatical envolve a inversão sujeito-predicado. Ou seja, a estrutura é possível⁶.

- (39) a. *La mesa es más alta que la puerta es ancha.
b. La mesa es más alta que ancha es la puerta

Resumindo, temos três parâmetros (DSP, DAP e DegPP), cada um deles está relacionado com um conjunto de estruturas na tabela 1. Note que se a língua tem marcação negativa para o primeiro, não terá os demais. Se a marcação é positiva para o primeiro, ela pode ter ou não o segundo, cuja marcação negativa impede que tenha o terceiro. E a marcação positiva no segundo, leva a possibilidade de ter marcação positiva ou negativa no terceiro. Assim, o quadro geral que se revela é bastante elegante.

4 Discussão

Há várias discussões teóricas sobre a semântica e a sintaxe da comparação, como mencionamos na introdução: a) qual é o nexos sintático entre as proposições?; b) todas as comparativas envolvem base oracional ou há línguas com estruturas ambíguas?; c) quais são as estratégias linguísticas de comparação? O trabalho de Beck *et alii* (2009) contribui para estimular a discussão da estipulação de parâmetros semânticos, dentro de uma abordagem referencial do significado, olhando particularmente para a questão (c).

O quadro geral revelado pela revisão do trabalho de Beck *et alii* (2009) nas seções anteriores nos mostrou que o PB parece se comportar parcialmente como o inglês, embora haja algumas dúvidas sobre a presença de algumas estruturas, especialmente as relacionadas ao parâmetro do sintagma de grau (DegPP), que envolve o preenchimento na sintaxe de uma posição estrutural por um operador ou por um grau (caso das interrogativas graduais e dos sintagmas de medida). Veja que isso poderia nos levar a crer que o PB esteja mais próximo do romeno e do espanhol, línguas em que a gramaticalidade dessas estruturas também é restrita.

Beck *et alii* não são claros sobre o que explica a impossibilidade de se preencher o argumento de grau com sintagmas de medida. Sanchez-

⁶ Ver também Reglero (2007).

Mendes (2016) também discute essa questão e aponta que uma solução seria reformular esse parâmetro, dado que no PB e no caritiana esse preenchimento não é possível. Sua sugestão é que esse aspecto fosse retirado do quadro geral de definição do parâmetro, pois para ela o sintagma de medida é um predicado de grau, não um elemento que ligaria uma variável de grau.

Nossa sugestão será na manutenção do sintagma de medida como parte do parâmetro. Como vimos anteriormente, o sintagma de medida pode ser analisado como um grau, ou um nome para um intervalo. Se essa análise está correta, podemos entender que as condições de verdade de uma sentença do inglês como (40a) e a do PB em (40b) são apenas diferentes lexicalizações ou realizações sintáticas da mesma proposição. Isto é, suas condições de verdade são as mesmas: são verdadeiras se e somente se John tem 6 ft de altura.

- (40) a. John is 6ft tall.
 b. John tem 6 pés de altura.
 c. *John tem 6 pés de altura* = $\exists d[\text{ALTURA}_{\text{alto}}(j) \geq d \ \& \ d = 6 \text{ pés}]$

Ou seja, esse parâmetro poderia ser reformulado como um subparâmetro ou um microparâmetro⁷: há línguas que preenchem sintaticamente o adjetivo com um grau, enquanto outras optam por preencher o nome da escala, que não deixaria de ser analisado como um predicado de graus ou uma função de medida.

- (41) a. A altura do prédio assusta.
 b. Os 100m de altura do prédio me assustam.

Em resumo: a expressão do sintagma de medida ou se realiza como um modificador de uma nominalização da escala ou como um argumento sintático.

⁷ Supomos que talvez esse seja um caminho, dado que microparâmetros acomodariam especificações lexicais e ter o argumento de grau preenchido por sintagmas de medida depende, em alto grau, de se ter algum sistema de medida para a dimensão. E mesmo isso não garante a possibilidade da estrutura. No inglês há um sistema de medição de temperatura do ambiente, mas não é gramatical uma estrutura como **today is 30° C hot*. Ou seja, há bastante irregularidade na distribuição desses sintagmas de medida.

O mesmo raciocínio poderia ser aplicado no caso das questões graduais. Contudo, note que a resposta para interrogativas como (42) só pode ser um sintagma de medida, enquanto a interrogativa com o *quão* pode ser com outras expressões graduais. Tomamos isso como um indício de que a semântica dessas interrogativas não é a mesma. Isto é, a interrogativa gradual com adjetivos dimensionais “tradicional” não é apenas uma variante arcaica, como sugere Sanchez-Mendes (2016), embora, certamente, seja uma forma mais formal de interrogativa gradual.

- (42) a. Qual é a altura do João?
 b. Quantoⁱ João tem tⁱ de altura?/O João tem quanto de altura?
 c. {d: 1,80m}
- (43) a. Quão alto é o João?
 b. {muito alto, bem alto, mais alto do que o Pedro}

Considerações Finais

Na primeira parte do artigo apresentamos as principais estruturas graduais e comparativas que as línguas podem exibir ou não. Essas diferentes estruturas são organizadas em três parâmetros, de acordo com a proposta de Beck *et alii* (2009). Esses parâmetros são de natureza diversa. O primeiro envolve a existência de predicados de graus no sentido clássico, isto é, a língua tem (ou não) predicado com um argumento sintático e semântico de tipo *d* (DSP). Se a língua tem esse tipo de predicado, ela exibe as estruturas comparativas que são relevantes para o segundo parâmetro, o parâmetro da abstração de graus. Esse parâmetro (DAP) explica diferenças relacionadas com a manipulação de graus por processos sintáticos (como as interrogativas graduais) e operações de escopo, processos que lidam com variáveis de grau, ligação sintática e movimento em forma lógica. O parâmetro do argumento de grau (DegPP) é essencialmente semântico e verifica a disponibilidade do preenchimento do argumento de grau dos adjetivos por sintagmas de medida em línguas como o inglês (cf. *John is 6ft tall.*). Na discussão dos parâmetros, analisamos a relação do PB com o inglês. A princípio a diferença significativa é em relação ao parâmetro do preenchimento do grau (DegPP). Nesse aspecto, nossa língua se comporta como o espanhol e o chinês, não apresentando o preenchimento do argumento de grau por

sintagmas de medida, nem interrogativas graduais similares à estrutura interrogativa típica do parâmetro.

Na discussão enfatizamos que é importante considerar também outras possibilidades linguísticas de expressão desses conteúdos. No PB temos interrogativas graduais de dois tipos que parecem estar em distribuição complementar. *Quão alto é o João?* interroga vagamente sobre a altura de João, enquanto *Qual é a altura de João?* interroga especificamente sobre a sua medida na escala de altura. Em inglês há uma interrogativa que cumpre as duas funções. A proposta de Beck *et alii* não considera o papel dessas outras estruturas. Notamos que elas exibem no PB a particularidade de envolverem nominalizações, como *altura*, e que essa nominalização também se manifesta no licenciamento de sintagmas de medida, como em *João tem 1,80m de altura*. Acreditamos que isso não é apenas coincidência. Nossa expectativa é de mais línguas que se comportem como o PB e o espanhol nesse aspecto: ao não permitir estruturas como sintagmas de medida e interrogativas graduais que operem sobre o argumento semântico de grau do adjetivo, a língua oferece outro mecanismo sintático, a nominalização da propriedade gradual. Note que em comum, tanto interrogativas quanto sintagmas de medida, se lexicalizam via estruturas desse tipo.

Declaração de Autoria

O artigo é o resultado de várias discussões que os autores travaram e a sua redação é fruto de trabalho conjunto. Para fins formais, seguem informações mais detalhadas:

Luisandro Mendes de Souza: discussão do objeto de estudo e análise. Revisão teórica. Planejamento e escrita inicial do texto. Adequações das sugestões dos pareceristas. Construção de análises. Redação e revisão geral do artigo.

Roberta Pires de Oliveira: discussão do objeto de estudo e análise. Levantamento teórico. Escrita complementar do artigo. Construção de análises e exemplos. Adequações das sugestões dos pareceristas. Revisão geral do artigo.

Lara Frutos: discussão do objeto de estudo e análise. Busca de dados. Criação de piloto para interpretação de sentenças limítrofes. Redação e revisão do artigo. Revisão final de normalização do texto.

Kayron Campos Beviláqua: discussão do objeto de estudo e análise. Busca de dados. Criação de piloto para interpretação de sentenças limítrofes. Escrita e revisão do artigo. Adequações das sugestões dos pareceristas.

Agradecimentos

Ao CNPq pelo apoio financeiro através de bolsa PQ a Roberta Pires de Oliveira. O artigo surgiu das discussões de um seminário sobre semântica dos predicados e modificadores graduais que Luisandro Mendes de Souza e Roberta Pires de Oliveira ministraram na pós-graduação em Letras da UFPR durante o segundo semestre de 2017. Agradecemos aos participantes do seminário, Denise Mazocco, Kayron Beviláqua e Lara Frutos. Agradecemos aos participantes do Workshop Universals and Variation in Semantics and Pragmatics, parte da programação do V Congresso Internacional de Linguística Histórica, onde esse trabalho foi apresentado e aos pareceristas anônimos pelas sugestões e leitura atenciosa do artigo.

Referências

- BARWISE, J.; COOPER, R. Generalized quantifiers and natural language. *Linguistics and Philosophy*, Nova York, v. 4, n. 2, p. 159–219, 1981. DOI: <https://doi.org/10.1007/BF00350139>. Acesso em: 31 jan. 2022.
- BECK, S.; KRASIKOVA, S.; FLEISHER, D.; GERGEL, R.; HOFSTETTER, S.; SAVELSBERG, C.; VANDERELST, J.; VILLALTA, E. Crosslinguistic variation in comparison constructions. *Linguistic Variation Yearbook*, Amsterdam, v. 9, p. 1-66, 2009. DOI: <https://doi.org/10.1075/livy.9.01bec>.
- BECK, S.; ODA, T.; SUGISAKI, K. Parametric Variation in the Semantics of Comparison: Japanese vs. English. *Journal of East Asian Linguistics*, Nova York, v. 13, p. 289-344, 2004. DOI: <https://doi.org/10.1007/s10831-004-1289-0>
- BECK, S. Comparison Constructions. In: MAIENBORN, C.; von HEUSINGER, K.; PORTNER, P. (eds.) *Semantics: an international handbook of natural language meaning*. Berlin: De Gruyter, 2011. p. 1343-1389.
- BOCHNAK, R. The non-universal status of degrees: evidence from Washo. In: KEINE, S.; SLOGGET, S. (eds.). *Proceedings of North-East Linguistic Society (NELS)*, 42, Amherst: GLSA, 2013. p. 79-92.
- BRITO, A.; MATOS, G. Construções de graduação e comparação. In: MIRAMATEUS, M. H. *et alli* (orgs.). *Gramática da língua portuguesa*. 5. ed. Porto: Caminho, 2003. p. 729-766.

BURNETT, H. A Delineation solution to the puzzles of absolute adjectives. *Linguistics and Philosophy*, Nova York, v. 37, n. 1, p. 1–39, 2014. DOI: <https://doi.org/10.1007/s10988-014-9145-9>. Disponível em: <<http://link.springer.com/10.1007/s10988-014-9145-9>>. Acesso em: 18 set. 2020.

CHIERCHIA, G. Reference to Kinds Across Languages. *Natural Language Semantics*, Nova Iorque, v. 6, p. 339-405, 1998. DOI: <https://doi.org/10.1023/A:1008324218506>

CORVER, N. Getting the (syntactic) measure of Measure Phrases. *The Linguistic Review*, Berlin, v. 26, n. 1, p. 67–134, 2009. DOI: <https://doi.org/10.1515/tlir.2009.003>. Disponível em: <<https://www.degruyter.com/document/doi/10.1515/tlir.2009.003/html>>. Acesso em: 31 jan. 2022.

CRUSE, D. *Lexical Semantics*. Cambridge: Cambridge University Press, 1986.

DINIZ, V. R. *Construções comparativas no português brasileiro : uma análise sintática*. 2018. 143 f. Dissertação (Mestrado em Linguística) - Universidade Federal do Paraná, 2018.

von FINTEL, K. Would you believe it? The King of France is back! Presuppositions and truth-value intuitions. In: REIMER, M.; BEZUIDENHOUT, A (eds.). *Descriptions and Beyond*. Oxford: Oxford University Press, 2004. p. 269-296.

HEIM, I. Degree Operators and Scope. *Semantics and Linguistic Theory (SALT)*, Nova York, v. 10, n. 0, p. 40–64, 2000. DOI: <https://doi.org/10.3765/salt.v10i0.3102>

HEIM, I.; KRATZER, A. *Semantics in Generative Grammar*. Oxford: Blackwell, 1998.

KENNEDY, C. *Projecting the Adjective: the Syntax and Semantics of Gradability and Comparison*. 1997. 243 f. Ph.D. dissertation, UC Santa Cruz, 1997.

KENNEDY, C. Vagueness and grammar: the semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy*, Nova York. v. 30, n. 1, p. 1–45, 2007. DOI: <https://doi.org/10.1007/s10988-006-9008-0>. Acesso em: 11 set. 2020.

KENNEDY, C. Modes of Comparison. In: Malcolm Elliott, James Kirby, Osamu Sawada, Eleni Staraki, and Suwon Yoon (eds) *The Proceedings of the 43th annual meetings of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society. p. 141-165.

KENNEDY, C.; MCNALLY, L. Scale structure, degree modification, and the semantics of gradable predicates. *Language*, Baltimore, v. 81, n. 2, p. 345–381, 2005. DOI: 10.1353/lan.2005.0071

KLEIN, E. A Semantics for Positive and Comparative Adjectives. *Linguistics and Philosophy*, Nova York, v. 4, p. 1-45, 1980. DOI: <https://doi.org/10.1007/BF00351812>

KLEIN, E. Comparatives. In: STECHOW, A.; WUNDERLICH, D. (eds.). *Semantics*. An international handbook on contemporary research. Berlin: de Gruyter. 1991. p. 673-691.

MARQUES, R. R. *Para uma semântica das construções comparativas em português*. 2003. 318 f. Tese (Doutorado em Linguística) - Universidade de Lisboa, 2003.

MORZYCKI, M. *Modification*. Cambridge: Cambridge University Press, 2016.

REGLERO, L. On Spanish Comparative Subdeletion Constructions. *Studia Linguistica*, Hoboken, v. 61, p. 130-169, 2007. DOI: <https://doi.org/10.1111/j.1467-9582.2007.00133.x>

RETT, J. Manner implicatures and how to spot them. *International Review of Pragmatics*, Leida, v. 12, n. 1, p. 44–79, 2020. DOI: <https://doi.org/10.1163/18773109-01201105>. Disponível em: <https://brill.com/view/journals/irp/12/1/article-p44_2.xml>. Acesso em: 8 abr. 2022.

ROTHSTEIN, S.; PIRES DE OLIVEIRA, R. Bare singular noun phrases are mass in Brazilian Portuguese. *Lingua*, Amsterdam, v. 121, n. 15, p. 2153–2175, 2011. DOI: <https://doi.org/10.1016/j.lingua.2011.09.004>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0024384111001744>>. Acesso em: 20 jul. 2021.

SCHWARZSCHILD, R. The Role of Dimensions in the Syntax of Noun Phrases. *Syntax*, Hoboken, v. 9, n. 1, p. 67–110, 2006. DOI: <https://doi.org/10.1111/j.1467-9612.2006.00083.x>

SOUZA, L. M. *Comparativas quantitativas no português brasileiro: sintaxe e semântica*. 2010. 157 f. Tese (Doutorado em Linguística) - Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina, 2010.

SANCHEZ-MENDES, L. Degree Constructions in Brazilian Portuguese and in Karitiana: a challenge to universal parameters. In: BUI, T.; IVAN, R.R. (eds.). *SULA 9: Proceedings of the Ninth Conference on the Semantics of Under-Represented Languages in the Americas*. Santa Cruz: GLSA, 2016. p. 141-154.

STASSEN, L. *Comparison and Universal Grammar: An Essay in Universal Grammar*. Oxford: Blackwell, 1985.

von STECHOW, A. Comparing Semantic Theories of Comparison. *Journal of Semantics*, Oxford, v. 3, p. 1-77, 1984. DOI: <https://doi.org/10.1093/jos/3.1-2.1>

ULTAN, R. Some features of basic comparative constructions. *Working Papers on Language Universals*, Palo Alto, n. 9, p. 117-162, 1972.



Mudança semântica e *word embeddings*: estudos de caso na diacronia do português

Semantic change and word embeddings: case studies on the diachrony of Portuguese

Lucas Fonseca Lage

Universität des Saarlandes (UdS), Saarbrücken, Saarland / Alemanha

flage.lucas@gmail.com

<https://orcid.org/0000-0003-1141-4236>

Evandro Landulfo Teixeira Paradela Cunha

Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Minas Gerais / Brasil

cunhae@ufmg.br

<https://orcid.org/0000-0002-5302-2946>

Resumo: De acordo com Givón (2001) o léxico é um repositório de conceitos relativamente estáveis no tempo, compartilhados socialmente e bem codificados, além de ser organizado em forma de rede, onde conceitos similares estão agrupados próximos uns aos outros. Em viés similar, o lexicólogo Georges Matoré propõe que palavras estabelecem relações associativas entre si e define os conceitos de campos nocionais e palavras-testemunho, elementos em torno dos quais o léxico se organiza. Com o uso de técnicas computacionais como *word embeddings*, que permitem a representação de palavras como vetores em um espaço vetorial, é possível analisar palavras agrupadas pelos mesmos traços semânticos. Este trabalho se propõe investigar a viabilidade de tais métodos para análise de mudança semântica. Para isso, foram analisadas ocorrências das formas “deus”, “homem”, “mulher”, “pai”, “mae” e “terra” no *corpus* Tycho Brahe do português. Através do algoritmo Skip-gram foram gerados Word Embeddings, e, posteriormente, visualizações para a rede de relações semânticas de cada palavra em três diferentes recortes temporais. Através das visualizações foram observadas evidências da organização semântica do léxico, além de sua reorganização através do tempo.

Palavras-chave: Linguística Computacional; Estudos Diacrônicos; Processamento de Língua Natural; Mudança Linguística; Vetorização de Palavras.

Abstract: According to Givón (2001), the lexicon is a repository of concepts which are relatively stable in time, socially shared and well encoded. They are well organized in a network where similar concepts are grouped next to each other. On a similar note, the lexicographer Georges Matoré proposes associative relationships between words and defines the concepts of notional field and testimonial words, which are organizational elements of the lexicon. Using computational techniques such as Word Embeddings, which represent words as vectors in a vector space, it is possible to analyze groupings of words based on their semantic features. This paper aims to explore the viability of such methods in semantic change. The occurrences of the word forms “deus”, “homem”, “mulher”, “pai”, “mae” and “terra” were analyzed in the Tycho Brahe *corpus* for Portuguese. Word Embeddings were created using the Skip-gram algorithm, and visualizations for a semantic feature network were created for each word in three different time slices. Evidence of the semantic organization of the lexicon and its reorganization was observed through the generated visualizations.

Keywords: Computational Linguistics; Diachronic Studies; Natural Language Processing; Linguistic Change; Word Embeddings.

Recebido em 27 de fevereiro de 2022

Aceito em 31 de maio de 2022

1 Introdução

Muitas são as propostas a respeito do que causa a mudança de significado de itens lexicais. Uma delas é a de Givón (1995), que caracteriza o léxico como um repositório de conceitos relativamente estáveis no tempo, compartilhados socialmente e bem codificados. De acordo com sua perspectiva, esses conceitos são interconectados em rede, de forma que a ativação de um conceito leva à ativação de conceitos vizinhos. A linguagem teria evoluído em paralelo com mecanismos cognitivos, com a organização sociocultural e com as habilidades comunicativas dos homínídeos. Assim, em uma sociedade na qual evoluções tecnológicas e culturais são a norma, a possibilidade de transmitir conhecimento e habilidades é de grande valia, mas, com a

variância de conceitos relevantes socialmente, algumas formas tornam-se mais frequentes e outras caem em desuso (GIVÓN, 1995).

Antes mesmo de Givón propor suas ideias e, com um olhar voltado para a lexicografia, Georges Matoré apresentou em seu trabalho os conceitos de *palavra-testemunho* e *campo nocional*, que seriam usados para descrever termos social e culturalmente relevantes. Matoré chega a apresentar redes associativas para certos termos e busca, também, analisar como essas redes se alteram temporalmente. Entretanto, devido às fortes críticas à sua metodologia, os estudos em lexicologia social foram negligenciados, tendo sido retomados, mais recentemente e com uma abordagem renovada (a lexicologia sócio-histórica), por Cambraia e colaboradores (CAMBRAIA, 2013; DORES; TOLEDO, 2018; RAFAEL; SIMIÃO, 2019; entre outros).

Tendo em vista os recentes desenvolvimentos de técnicas na área da computação, em especial as técnicas de *word embeddings* (vetorização de palavras), novas categorias de análise linguística têm se tornado possíveis. Por meio de estudos na área de linguística de *corpus*, que têm disponibilizado dados de qualidade em abundância, tornam-se viáveis novas metodologias de pesquisa que permitem analisar uma grande quantidade de dados (BERBER SARDINHA, 2004).

As técnicas de manipulação de dados muitas vezes não precisam ser sofisticadas, como se observa no trabalho de Michel *et al.* (2011), no qual, por meio de medidas de frequência e análise de entidades culturalmente relevantes, são expostos fenômenos sobre a evolução da gramática, relevância cultural e até mesmo censura. Alguns exemplos são a constante competição entre formas regulares e irregulares do passado no inglês, a redução no tempo em que entidades famosas permanecem relevantes, e a supressão de nomes judeus em livros alemães durante a década de 1930. Essas mesmas medidas podem auxiliar tanto na decisão de inclusão de novos termos em um dicionário quanto na remoção de termos irrelevantes (MICHEL *et al.*, 2011).

A frequência de uso de palavras também pode ser vista como evidência para fenômenos de mudança. Givón afirma que a forma deve cumprir uma função – logo, quando a função se torna pouco útil, a forma cai em desuso. Similarmente, Bochkarev, Solovyev e Wichmann (2014) sugerem que a mudança lexical é favorecida não só por mudanças em um ambiente social e natural, como também em um ambiente linguístico particular. Esses autores utilizam um *corpus* diacrônico para analisar a

taxa de mudança lexical de acordo com a frequência de ocorrência das 100.000 palavras mais frequentes, e mostram como o inglês britânico e o inglês americano estão convergindo, apesar da inicial separação. Eles afirmam que as duas variedades linguísticas se tornaram mais próximas dado ao advento da mídia de massa, que cresceu exponencialmente nos séculos XX e XXI (BOCHKAREV; SOLOVYEV; WICHMANN, 2014; GIVÓN, 2001).

Usando técnicas computacionais mais complexas, Hamilton, Leskovec e Jurafsky, (2016) utilizam vetorização de palavras não só para buscar formas que sofreram mudança de significado, mas também para validar mudanças de significado já conhecidas. Com sua análise, é possível, ainda, buscar as formas que passaram por maiores mudanças semânticas. Após o uso dessas técnicas e com a validação de seus próprios trabalhos, os autores chegam a apresentar leis para a mudança semântica (HAMILTON; LESKOVEC; JURAFSKY, 2016).

Buscando otimizar o trabalho de Hamilton, Leskovec e Jurafsky (2016), foi proposto por Yao *et al.* (2018) uma nova metodologia de aprendizado de vetores de palavras, capaz de codificar também o componente tempo. Esses autores, entretanto, não analisam o significado de formas específicas, mas as palavras associadas a elas. Portanto, o algoritmo desenvolvido é capaz de codificar associações como “apple” e “strawberry” em um primeiro ponto no tempo, e em outro ponto, “apple” e “iphone” (YAO *et al.*, 2018).

O presente artigo visa avaliar as mudanças semânticas sofridas por expressões ou palavras por meio de técnicas desenvolvidas na área de Processamento de Língua/Linguagem Natural (PLN). A princípio, utiliza-se a análise das frequências de ocorrências das palavras de um *corpus* diacrônico do português; em seguida, são gerados vetores de palavras para três recortes temporais distintos desse *corpus*; e, finalmente, são geradas imagens nas quais se é possível visualizar as redes de relações semânticas ao longo do tempo.

A fundamentação teórica para esta pesquisa se pautará em uma abordagem lexicológica e funcional, tendo em vista que os conceitos aplicados de forma prática pelos algoritmos de vetorização de palavras foram fundamentados, indiretamente, por autores como Georges Matoré e Talmy Givón (CAMBRAIA, 2013; GIVÓN, 1995).

Como fonte de dados, é utilizado o Corpus Histórico do Português Tycho Brahe, que abrange textos do século XIII até o século XX (DE SOUSA, 2014; GALVES; ANDRADE; FARIA, 2017).

2 O léxico em Givón e o léxico em rede

Segundo Cunha (2008), diferentemente das correntes estruturalistas e gerativistas, que buscam uma separação clara entre a língua como sistema (*langue*, competência) e a língua em uso (*parole*, desempenho), as abordagens funcionalistas tratam a estrutura gramatical da língua em relação aos seus contextos de uso. Dessa forma, as pesquisas dessa vertente se diferenciam nos métodos utilizados, nos dados considerados relevantes e, mais profundamente, nos objetivos da análise linguística (CUNHA, 2008).

Dentro de uma perspectiva funcionalista, uma sentença é analisada sempre de acordo com o contexto em que ela foi produzida. Dessa forma, pode-se afirmar que a metodologia de análise parte de um método indutivo, analisando os dados, criando generalizações e, só então, testando essas generalizações. Portanto, de acordo com essas abordagens, os dados analisados devem ser obtidos a partir de produções reais de fala e escrita (CUNHA, 2008).

Na perspectiva funcionalista, destacam-se os estudos de Givón (1995, 2001), que desenvolveu uma gramática propriamente funcionalista. Em sua obra, ele traz conceitos de outros campos, como a biologia, para justificar os caminhos tomados pela evolução das línguas. Segundo o autor, a evolução biológica é cercada por inúmeros fatores, muitas vezes envolvendo fenômenos aleatórios, e a forma que persevera é a forma que melhor realiza uma função específica. Essa evolução, considerada funcional, ocorre de forma similar nas línguas naturais. Givón ainda afirma que, apesar de os pontos abordados por ele não serem novos, como, por exemplo, a capacidade humana de processamento de linguagem ser uma evolução do sistema de processamento de imagens visuais, a abrangência de fatos influenciados por essas conclusões não foi ainda estudada (GIVÓN, 1995, 2001).

Em seu livro *Functionalism and Grammar*, o autor lista os componentes funcionais para a comunicação humana e os divide em dois módulos principais, que interagem entre si (GIVÓN, 1995). São eles o Sistema de Representação Cognitiva e os Sistemas de Codificação.

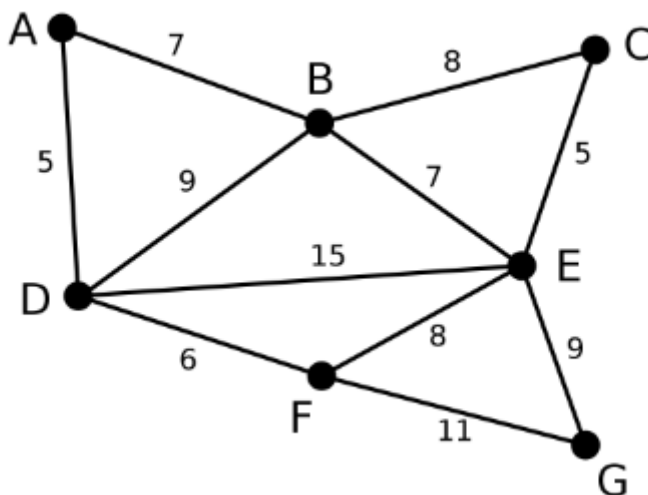
Dentro do Sistema de Representação Cognitiva, tem-se três componentes: o léxico conceitual, a informação proposicional e o discurso multi-proposicional. Aqui, interessa principalmente a definição de léxico conceitual, pois é a partir dela que Givón relaciona, inicialmente, o léxico ao meio e às experiências humanas (GIVÓN, 1995).

O léxico humano é definido na mesma obra como um conjunto de conhecimentos que, quando tomados juntos, constituem um mapa cognitivo do nosso universo de experiências como seres humanos. Esse universo de experiências se refere aos meios externo-físico, ao universo sociocultural, e ao nosso universo mental-interno. Além disso, os conceitos que compõem o léxico são definidos por Givón como estáveis no tempo, compartilhados socialmente e bem codificados. Nessa visão, ser estável no tempo significa que as palavras e os conceitos associados a elas não estão em um fluxo rápido – por exemplo, o termo “cavalo” provavelmente possuirá o mesmo significado daqui a alguns anos. Dizer que os conceitos são compartilhados socialmente significa que as palavras possuem aproximadamente o mesmo significado para os outros membros de sua comunidade de fala. E, por fim, ser bem codificado quer dizer que cada parte da informação armazenada no léxico é, em partes, fortemente associada a apenas um código, ou etiqueta perceptual. Ou seja, cada parte do conhecimento lexical possui apenas um correspondente no código (GIVÓN, 2001).

Com essas características do léxico, Givón (2001) conclui que ele está organizado por meio de nós e arestas, e que cada nó corresponde a uma palavra. A ativação de um nó-palavra ainda seria responsável pela ativação de outros nós-palavra que possuam uma relação íntima com o primeiro. Na Figura 1 tem-se um exemplo de como se organiza a rede descrita por Givón, que se assemelha a um grafo. Grafos são uma abstração matemática para se representar objetos e as relações entre eles, os quais possuem nós conectados por arestas. As arestas de um grafo podem possuir pesos, o que pode corresponder à distância entre dois pontos, por exemplo. Na Figura 1, observam-se os nós de “A” a “G” e as arestas com diferentes pesos conectando-os. Dentro de uma rede léxico-semântica, os nós correspondem a conceitos individuais, cada um com seu próprio significado e código-etiqueta. As conclusões de Givón são justificadas pelo trabalho de Swinney (1979), que analisa o tempo de reconhecimento de palavras, quando apresentadas a um leitor em contextos ambíguos, e conclui que, quando uma palavra é percebida,

todos os possíveis significados dela são também ativados na mente da pessoa (GIVÓN, 2001; GRIFFIN, 2017; SWINNEY, 1979).

Figura 1 - Exemplo de grafo. Cada letra do alfabeto corresponde a um nó, enquanto os números indicam os pesos das arestas



Fonte: Griffin (2017, p. 47).

Os conceitos lexicais são as experiências humanas armazenadas de forma convencional e genérica, e não pontos específicos para cada experiência. Por serem genéricos, eles presumem um padrão de ativação para os conjuntos interconectados de nós. Um conceito lexical pode se referir a uma entidade relativamente estável no tempo, como um objeto, uma cidade, um local, animal ou até a conceitos abstratos – essa entidade corresponderia, então, a um substantivo. Pode se referir, ainda, a uma ação temporária, um processo ou relação, ou seja, um verbo. E por fim, pode representar uma qualidade estável no tempo ou temporária, como um adjetivo (GIVÓN, 2001).

A ideia do léxico em rede descrita por Givón, como ele mesmo coloca, não é necessariamente nova. Outros autores já haviam buscado trabalhar o léxico de forma sistemática. Um deles é Georges Matoré, que trabalhou com o léxico como uma rede, buscando uma lexicologia social. De acordo com Cambraia, através de seu livro *La lexicologie sociale*,

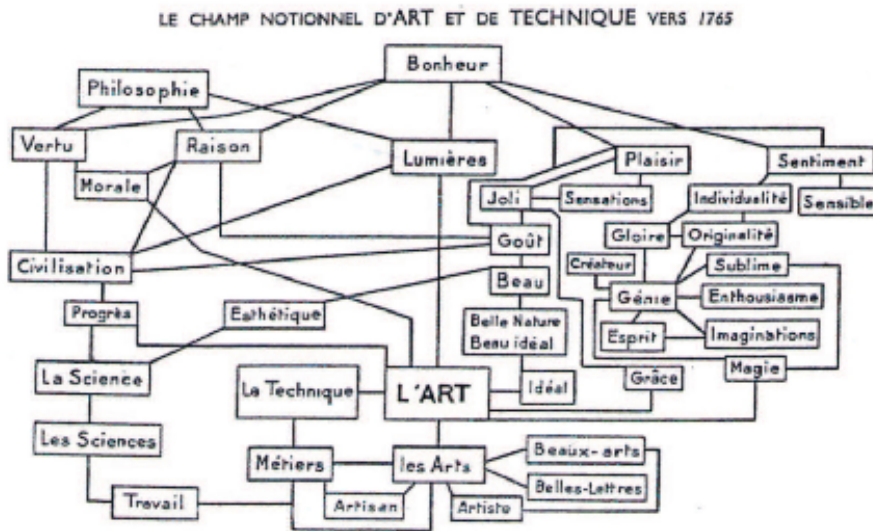
Georges Matoré (1949) cita a influência de fatores sociais no estudo do léxico. Ele propõe uma série de princípios para uma nova lexicologia, denominada lexicologia social. O primeiro desses princípios propõe que forma e conceito são indissociáveis. Segundo o autor, a formação de uma palavra equivale à formação de um conceito e esse processo criativo, apesar de individual em seu início, é seguido de uma socialização, que difunde e coletiviza a palavra e o conceito. Portanto, existe um caráter social da palavra e é por esse aspecto da significação que a lexicologia deveria se interessar (CAMBRAIA, 2013).

Matoré mantém uma visão sistêmica do comportamento do vocabulário, ou seja, admite que as palavras estabelecem relações recíprocas na consciência. As palavras podem se relacionar com suas vizinhas, através de relações sintagmáticas, ou com palavras similares, através de forma ou sentido, estabelecendo relações associativas. Matoré ainda afirma ser impossível extrair o fator tempo de suas análises, pois o momento de criação da palavra faz parte do conjunto de operações mentais que a produziram (CAMBRAIA, 2013).

São encontradas algumas similaridades entre o trabalho de Matoré e os princípios estabelecidos por Saussure para o estruturalismo, mas o lexicólogo se desvencilha dessa corrente. Ele o faz discordando de Saussure a respeito da organização morfológica do léxico e atribui ao fator social o principal papel na organização do léxico (CAMBRAIA, 2013).

A metodologia de estudo do francês define que se façam recortes temporais que levem em conta a noção de “geração”, cuja definição é uma faixa de tempo de 30 a 36 anos. Em seguida, devem ser identificados os *campos nocionais*, baseados no parentesco sociológico dos elementos. Esses campos são compostos por *palavras-testemunho*, que são elementos importantes em torno dos quais a estrutura lexicológica, sua hierarquia e sua coordenação são estabelecidos. Com base nesses métodos, Matoré exemplifica seu estudo através dos campos nocionais de Arte e Técnica por volta de 1765, Figura 2, e o campo nocional de Artista entre os anos de 1827 e 1834, Figura 3 (CAMBRAIA, 2013).

Figura 2 - Campo nocional de “Arte” e “Técnica” em 1765, segundo Georges Matoré



Fonte: Cambraia (2013, p. 163).

Os métodos defendidos por Matoré foram muito criticados na época. Algumas críticas eram voltadas a partes mais técnicas do trabalho, como a definição arbitrária de uma geração de 30 a 36 anos, ou a imprecisão na definição dos termos “palavras-testemunho” e “campos nocionais”. Além disso, uma consideração muito importante foi realizada por Robin. A autora considera que os estudos de Matoré não poderiam refletir a sociedade como um todo, mas apenas os grupos aos quais pertencem as pessoas cujos textos foram analisados. Apesar dos problemas da metodologia proposta pelo lexicólogo, permanece em destaque a importância de se considerar a influência do social na organização do léxico (CAMBRAIA, 2013; ROBIN; DE MENESES BOLLE, 1977).

3 Semântica vetorial e *word embeddings*

O conceito de *word embeddings*, ou vetorização de palavras, assim como as propostas de Givón, também se desenvolveu a partir de conceitos evolucionários. Da mesma forma que espécies diferentes desenvolvem estruturas corporais similares por evoluírem em ambientes similares, palavras que ocorrem em contextos similares devem possuir significados similares. Essa hipótese, denominada hipótese distribucional, foi proposta por linguistas como John R. Firth, na década de 1950, quando se percebeu que palavras sinônimas ocorrem no mesmo ambiente/contexto. Firth afirma que “Você conhece uma palavra pela sua companhia”¹ (FIRTH, 1957; JURAFSKY; MARTIN, 2008).

Enquanto muitas palavras não possuem sinônimos, a grande maioria das palavras possui outras que são muito similares. Por exemplo, apesar de as palavras “cão” e “gato” não serem sinônimas, há muitas semelhanças entre elas e entre os contextos onde ocorrem. Ambas são substantivos, referem-se a animais domésticos, de quatro patas, alimentados pelos seus donos, mas um mia e outro late. Essas semelhanças semânticas têm como consequência seu surgimento em contextos sintáticos semelhantes. A similaridade entre palavras, sentenças ou documentos é bastante útil em diversas tarefas de PLN como sistemas de resposta a perguntas, paráfrase e sumarização (JURAFSKY; MARTIN, 2008).

Além de relações entre palavras como sinonímia, polissemia, antonímia e similaridade, as palavras também possuem um caráter afetivo. O caráter afetivo, ou conotação, refere-se aos aspectos do sentido de uma palavra que estão relacionados aos sentimentos do falante ou do ouvinte. Assim, as palavras podem ter uma conotação positiva (feliz, bom, amor) ou uma conotação negativa (triste, mal, ódio). Um dos primeiros trabalhos sobre o sentido afetivo de palavras foi o de Osgood e colegas, no qual são criados três eixos a fim de avaliar o sentido afetivo de uma palavra e então associa-se um valor numérico a cada eixo. Na Figura 4 observam-se as valorações das palavras *courageous*, *music*, *heartbreak* e *cub* em três eixos. O eixo de valência está relacionado à agradabilidade do estímulo gerado, o de excitação diz respeito à intensidade da emoção provocada pelo estímulo e o eixo de dominância se refere ao grau de controle exercido pelo estímulo (JURAFSKY; MARTIN, 2008; OSGOOD; SUCI; TANNENBAUM, 1957).

¹ “You shall know a word by the company it keeps” (FIRTH, 1957, p.11).

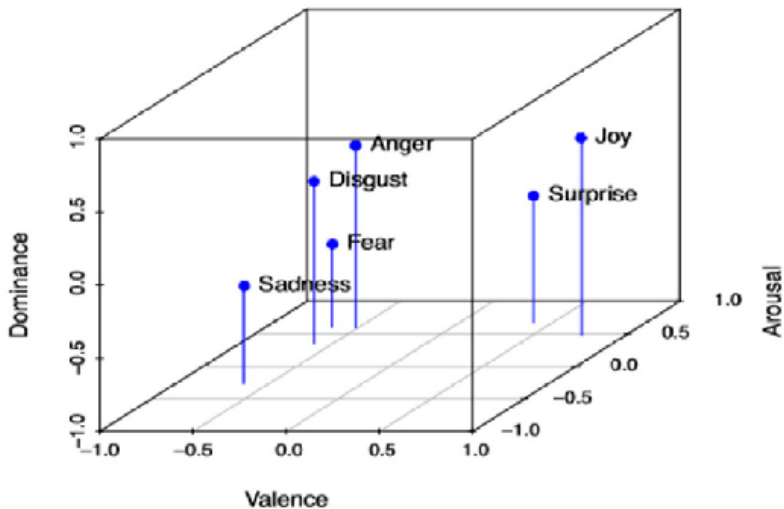
Figura 4 - Variação quantitativa do sentido afetivo em três eixos

	Valence	Arousal	Dominance
courageous	8.05	5.5	7.38
music	7.67	5.57	6.5
heartbreak	2.45	5.65	3.58
cub	6.71	3.95	4.24

Fonte: Jurafsky; Martin, (2008, p. 106)

A grande contribuição de Osgood para o campo foi a percepção de que as palavras poderiam ser representadas em um espaço vetorial, criando assim, um espaço tridimensional para localização espacial das palavras. A Figura 5 mostra um exemplo de como palavras relacionadas a sentimentos estão posicionadas nesse espaço vetorial (nesse caso, de três dimensões). As palavras mostradas são *anger*, *disgust*, *fear*, *joy*, *sadness* e *surprise* (“raiva”, “repulsa”, “medo”, “alegria”, “tristeza” e “surpresa”).

Figura 5 - Representação afetiva de palavras, segundo Osgood et al. (1957)



Fonte: Bălan *et al.*, (2020, p. 4).

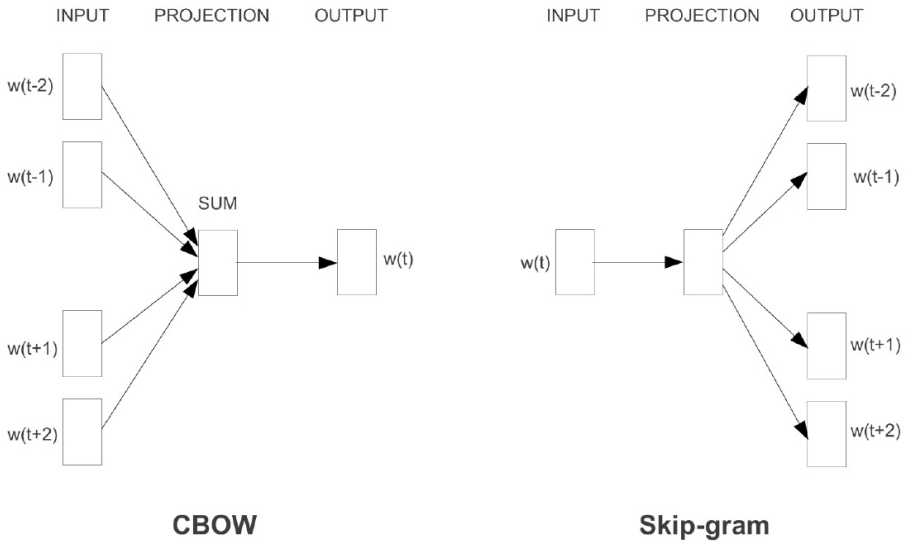
Após esses estudos iniciais, os métodos para obtenção de representações vetoriais evoluíram até chegarem a usar métodos de aprendizado automático, como os modelos *Continuous Bag-of-Words* e *Skip-gram*.

Um dos maiores avanços em vetorização de palavras é a publicação de Mikolov *et al.* (2013). Nela, é proposta uma arquitetura de redes neurais capaz de criar uma representação vetorial para cada palavra apresentada ao modelo, e, em seguida, é possível reproduzi-las em um espaço vetorial. O mérito desse trabalho está relacionado à resolução de questões sobre a complexidade da arquitetura e o tempo necessário para o treinamento dessas redes. Além disso, também foram desenvolvidas métricas de validação de modelos que não só são capazes de determinar se palavras estão próximas entre si, como também de quantificar o grau de similaridade entre as palavras (MIKOLOV *et al.*, 2013).

As arquiteturas propostas pelos autores foram denominadas de *Continuous Bag-of-Words (CBOW)* e *Continuous Skip-gram (Skip-gram)*. A primeira é criada a partir de uma tarefa de predição, em que uma palavra é prevista dado seu contexto, ou palavras vizinhas, como entrada em uma rede neural. O contexto, no caso, deve ser entendido como as palavras em posições anteriores e posteriores a palavra a ser predita. A partir dos valores de entrada, um classificador log-linear calcula a palavra mais provável de ocorrer naquele contexto; caso a predição esteja correta, a rede realiza operações dentro de si para reforçar seu aprendizado. Caso a predição esteja errada, ela altera valores dentro de si para buscar acertar nas próximas tentativas. É importante ressaltar que a ordem das palavras dentro da janela de entrada não é um fator relevante para a predição (MIKOLOV *et al.*, 2013).

O modelo Skip-gram possui uma arquitetura similar ao modelo CBOW, mas ao invés de prever uma palavra dado seu contexto, ele realiza a tarefa inversa, isto é, prediz o contexto a partir de uma palavra. Na Figura 6 tem-se uma representação da arquitetura dos modelos aqui discutidos (MIKOLOV *et al.*, 2013).

Figura 6 - Esquema da arquitetura dos modelos CBOW e Skip-gram



Fonte: Mikolov *et al.* (2013, p. 5)

Para a validação dos modelos criados, os autores desenvolveram tarefas de predição baseadas em relações sintáticas e semânticas. Esses testes são usados como uma métrica de desempenho dos modelos, permitindo, assim, compará-los quantitativamente.

Como exemplo de similaridade sintática, são utilizadas as relações entre os adjetivos do inglês em sua forma base, comparativa e superlativa. Essas relações podem ser preditas conforme os valores obtidos para os vetores após o treinamento do modelo. Assim, pode-se encontrar que a relação entre *big* e *bigger* é a mesma que entre *small* e *smaller*. Essa relação pode então ser reescrita através de uma operação vetorial como $\text{vetor}(\text{"big"}) + \text{vetor}(\text{"bigger"}) - \text{vetor}(\text{"small"}) = \text{vetor}(\text{"smaller"})$ (MIKOLOV *et al.*, 2013).

Como exemplo de relação semântica, Mikolov et al. criaram tarefas para determinar as relações entre nomes de países e suas capitais. Assim podemos traduzir a relação "Paris está para França assim como Berlim está para Alemanha" como uma operação vetorial: $\text{vetor}(\text{"Paris"}) + \text{vetor}(\text{"França"}) - \text{vetor}(\text{"Alemanha"}) = \text{vetor}(\text{"Berlim"})$ (MIKOLOV *et al.*, 2013).

Dentro dessas tarefas, o modelo Skip-gram atingiu a melhor acurácia total quando comparado ao modelo CBOW e a outros modelos

de vetorização de palavras. O desempenho geral do modelo CBOW foi a princípio ruim, mas teve o terceiro melhor desempenho dentro das tarefas de relações sintáticas. A grande vitória de ambos os modelos, entretanto, é no tempo gasto para seu treinamento. Para as mesmas condições de treino, mesmo tamanho de *corpus* e mesmas capacidades de processamento, os modelos CBOW e Skip-gram demoraram 2 e 2,5 dias respectivamente para finalizarem o treinamento, enquanto as arquiteturas comparadas precisaram de 14 dias de treino para criar o seu modelo. A redução no tempo de treinamento reduziu o custo computacional de forma que os modelos de vetorização de palavras se tornaram mais populares ainda (MIKOLOV *et al.*, 2013).

Como anteriormente dito, *word embeddings* são capazes de capturar características semânticas e sintáticas, com base nisso, Hartmann (2016), os utiliza como *feature* para tarefas de similaridade semântica. Em seu trabalho é combinado um método já difundido, o TF-IDF (*term frequency-inverse document frequency*), com *word embeddings*, obtidos a partir de um corpus jornalístico do português brasileiro. O modelo apresentado pelo autor apresenta resultados superiores em comparação com o modelo base e o modelo TF-IDF em tarefas de similaridade semântica. (HARTMANN, 2016)

Tendo em mãos uma técnica não só capaz de capturar relações sintáticas e semânticas, como também de menor custo computacional, falta então a capacidade de se trabalhar com um *corpus* diacrônico. Ao propor leis estatísticas para a mudança semântica, Hamilton, Leskovec e Jurafsky (2016) utilizaram-se de vetores de palavras para obter seus resultados.

As leis propostas pelos autores são a Lei da Conformidade e a Lei da Inovação. A primeira diz que a velocidade com que uma palavra muda seu sentido é inversamente proporcional a uma função exponencial da frequência de palavras. Já a segunda alega que, dentre palavras com frequência de ocorrência similar, as palavras polissêmicas mudam seu sentido mais rapidamente (HAMILTON; LESKOVEC; JURAFSKY, 2016).

Para chegar a essa conclusão, os autores utilizam três diferentes arquiteturas de *word embeddings* e *corpora* diacrônicos que englobam quatro línguas diferentes, sendo elas inglês, alemão, francês e chinês. Foram então criados modelos de *word embeddings* que abrangiam diferentes períodos de tempo e, após alinharem os modelos para cada período, foi criada uma representação para palavras cuja mudança semântica é conhecida. As palavras escolhidas foram *broadcast*, *gay* e *awful* e a partir

das mudanças sofridas obteve-se uma representação gráfica visível na Figura 7 (HAMILTON; LESKOVEC; JURAFSKY, 2016).

Figura 7 - Deslocamento vetorial das formas broadcast, gay e awful entre 1800 e 1990.



Fonte: Hamilton; Leskovec; Jurafsky (2016)

A escolha das palavras pelos autores não foi aleatória, pois foram escolhidas palavras que pudessem validar a metodologia descrita por eles. Assim, esperava-se que a palavra *broadcast* estivesse ligada a termos relacionados a agricultura em um primeiro momento e, em seguida, estivesse próxima de termos relacionados a notícias, jornais, televisão e rádio (HAMILTON; LESKOVEC; JURAFSKY, 2016).

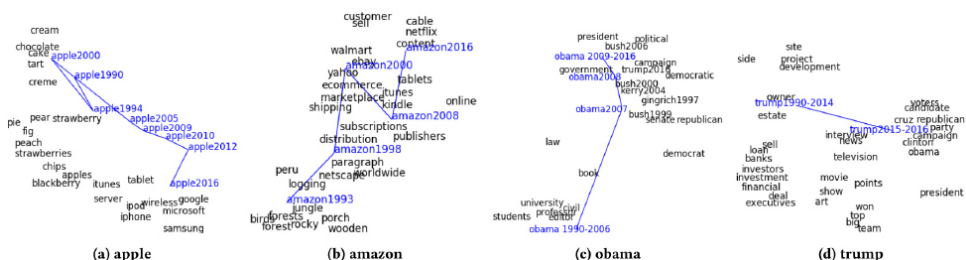
O estudo conseguiu, em um primeiro momento, validar mudanças de significado já conhecidas e foi em seguida usado para buscar as palavras que sofreram maior mudança semântica, aqui considerado o maior deslocamento no espaço vetorial ao longo dos períodos analisados (JURAFSKY; MARTIN, 2008).

Buscando otimizar a visualização de vetores em *corpora* diacrônicos, Yao *et al.* (2018) apresentam um novo modelo capaz de aprender vetores de palavras diacrônicos em um único passo. O maior problema ao se realizar esse tipo de vetorização em *corpora* diacrônicos é alinhar os eixos dos modelos. Devido às operações realizadas nos vetores para a obtenção de um modelo como o Skip-gram, as mesmas palavras podem ser geradas em pontos diferentes do espaço vetorial. Isso não altera a distância entre elas dentro do mesmo modelo, mas entre modelos diferentes não é possível analisar-se o deslocamento da palavra no intervalo de tempo entre os dois modelos. O método proposto por Hamilton, Leskovec e Jurafsky (2016) é constituído de dois passos: primeiro cria-se os vetores de palavras, em seguida alinha-se esses vetores

em um mesmo eixo. A proposta de Yao et al. (2018) busca realizar a codificação do fator tempo paralelamente ao treinamento do modelo (HAMILTON; LESKOVEC; JURAFSKY, 2016; YAO *et al.*, 2018).

Por fim, os autores apresentam redes associativas temporais para as palavras *apple*, *amazon*, *obama* e *trump*, vistos na Figura 8.

Figura 8 - Trajetórias de nomes através do tempo



Fonte: Yao *et al.* (2018)

As associações aprendidas mostram como, por exemplo, o termo “*amazon*”, inicialmente associado a termos do campo da natureza, se torna associado a termos do campo da tecnologia. Assim, mostra-se que é possível analisar diacronicamente e a partir de métodos computacionais as redes de associações dentro do léxico.

4 Metodologia

A partir das visões do léxico em rede (como as de Givón e aquela proposta por Matoré), propõe-se, no presente trabalho, buscar a representação da mudança semântica de palavras relevantes para o português. Para isso, será utilizado o Corpus Histórico do Português Tycho Brahe, elaborado por De Sousa (2014).

Buscando uma metodologia similar à de Hamilton, Leskovec e Jurafsky (2016), o *corpus* foi dividido em períodos relevantes para o estudo da mudança semântica. Entretanto, a determinação de um período relevante se mostra bastante imprecisa. Além disso, apesar da importância do *corpus* Tycho Brahe, a quantidade de palavras no *corpus* (em torno de 8 milhões de tokens) é muito menor quando comparada aos *corpora* utilizados por Hamilton, Leskovec e Jurafsky (2016), que possuem 850 bilhões de tokens. Há ainda uma escassez de textos nos períodos mais antigos da língua, em especial no século XIV. A fim de minimizar

o impacto desses obstáculos, os textos foram agrupados por século, garantindo uma quantidade minimamente significativa de tokens para cada século e mantendo um recorte de tempo padronizado (HAMILTON; LESKOVEC; JURAFSKY, 2016).

Primeiramente foi realizada uma análise exploratória a fim de encontrar as formas mais frequentes, a distribuição das frequências das palavras e quais delas são relevantes social e culturalmente. As palavras foram escolhidas e analisadas através da perspectiva de *palavras-testemunho* e *campos nocionais* de Matoré. Dessa forma, buscaram-se as redes de relações para os seguintes termos: *homem, mulher, pai, mãe, terra e deus*. As palavras aqui mencionadas foram escolhidas por dois motivos: primeiramente, são centrais para a articulação de valores socioculturais, são termos geralmente carregados e podem evidenciar vieses; segundo, possuem frequência relativamente alta dentro do *corpus* e possibilitam uma melhor qualidade de resultados. A partir daí foram buscadas associações que possam dar pistas sobre a percepção e os conceitos culturais que permearam esses termos durante a história da língua portuguesa.

As análises foram desenvolvidas utilizando a linguagem de programação Python, além das diversas bibliotecas para PLN existentes como Spacy e NLTK (Natural Language Toolkit). O *corpus* escolhido para a análise foi o *corpus* Anotado do Português Histórico Tycho Brahe (CTB), pioneiro no que concerne à língua portuguesa, que permanece hoje como o maior *corpus* eletrônico anotado de textos históricos em português. Hoje, o conjunto de dados inclui textos escritos por autores portugueses, brasileiros e africanos, nascidos entre 1380 e 1845, publicados entre os séculos XIV e XX. Segundo De Sousa, as anotações realizadas nos textos têm como objetivo principal possibilitar a recuperação de informações filológicas e linguísticas dos textos (DE SOUSA, 2014).

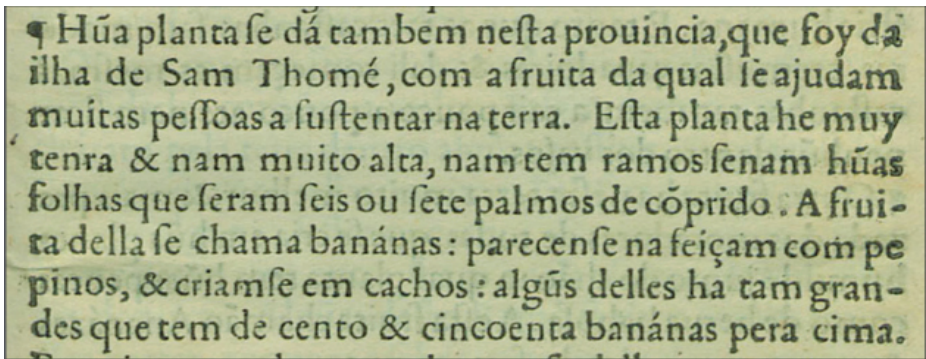
O *corpus* é composto de 88 textos, totalizando 3.544.628 palavras, sendo 58 textos anotados morfologicamente e 27 textos anotados sintaticamente. O *corpus* possui uma variedade de gêneros textuais, sendo eles: cartas, atas, textos narrativos, textos dissertativos, gramáticas, gazetas e jornais e textos de dramaturgia (DE SOUSA, 2014).

Como muitos textos provêm de séculos passados, o seu processamento deve envolver uma adaptação para que possam ser lidos e analisados hoje. O processamento do texto a partir da obra original foi feito por três camadas: uma camada de edição, uma camada

morfossintática e uma camada sintática. As anotações acontecem de forma incremental, ou seja, cada uma depende do resultado da etapa anterior (DE SOUSA, 2014).

A primeira etapa é a anotação de edição, que codifica informações relativas às decisões editoriais e à estrutura do texto (quebras de linha, parágrafos, seções, etc.) e também lida com intervenções interpretativas, como atualização grafemática, expansão de abreviaturas e atualização ortográfica. Na Figura 9, tem-se um exemplo de como os textos originais se encontram (DE SOUSA, 2014).

Figura 9 - Exemplo de trecho original antes de ser adaptado para o *corpus* Tycho Brahe.



Fonte: De Sousa (2014, p. 57).

A segunda etapa de anotação é a anotação morfossintática, que consiste na identificação e codificação das classes de palavras. A terceira e última etapa é a anotação sintática. Nessa etapa, é realizada a identificação e codificação da estrutura sintagmática da sentença, que foi realizada através de um parser sintático automático. Para realizar a anotação sintática do *corpus*, foi desenvolvido um parser sintático a partir do sistema Penn-Treebank. O parser foi treinado ao ser realimentado com seus resultados corrigidos por pesquisadores até que seu desempenho foi considerado satisfatório (DE SOUSA, 2014).

Dadas essas diversas características do *corpus* Tycho Brahe, fez-se então uma análise exploratória preliminar do *corpus*. Para este trabalho, foram feitos recortes temporais nos textos do *corpus*. O agrupamento foi realizado a partir de uma adaptação da classificação proposta por Bechara (1985). A delimitação proposta por Bechara tem início na fase

arcaica, que compreende o século XIII até o final do século XIV. Essa fase compreende o período chamado de galego-português, em que os documentos escritos existentes são de variedade culta e erudita. Alguns dos fenômenos encontrados nessa variedade são:

- possessivos femininos de formas proclíticas (ma, ta, as) ao lado de formas normais (mha, mia; tua, sua), que eram empregados sem muito rigor quanto sua função;
- o -d- etimológico da desinência de 2ª pessoa plural: amades, fazedes, queredes, seeredes, leixedes, fazede, etc.;
- terminação -on (-om) nas formas verbais oriundas de -unt: amáron (amárom), quiseron (quiserom), etc.

A próxima fase é a arcaica média, que corresponde ao intervalo entre a 1ª metade do século XV até a 1ª metade do século XVI. O autor a caracteriza como uma fase de transição, mas destaca a queda do -d- da desinência de 2ª pessoa do plural como essencial para se delimitar esse período (BECHARA, 1985).

A terceira fase proposta por Bechara (1985) é a fase moderna, que vai da 2ª metade do século XVI até o final do século XVII. Alguns dos fenômenos dessa fase destacados pelo autor são:

- A fixação do plural dos nomes em -ão (mãos, cães, leões) e do feminino dos adjetivos em -ão (são/sã);
- A presença obrigatória do pronome demonstrativo antes do pronome relativo em construções como *eu sou o que, tu és o que, nós somos os que*, etc. (persistindo até final do séc. XVIII).
- A progressiva ação analógica do radical do infinitivo sobre o radical da 1ª pessoa de muitos verbos, como *senço/sinto, menço/minto, arco/ardo*, etc.

A quarta e última fase definida por Bechara (1985) é a fase contemporânea, que compreende o século XVIII até hoje. Alguns dos fenômenos característicos desse momento são:

- a progressiva eliminação do pronome vós;
- fixação da oposição lhe singular/lhes plural, quando não combinados com os pronomes o, a, os, as;
- o desaparecimento de formas de indeterminação do sujeito como *homem e um*;
- o emprego das preposições *per* e *por* é unificado na forma única *por*.

5 Análise exploratória

A partir desses conhecimentos, foi realizada uma análise exploratória do CBT com a intenção de avaliar quais termos atenderiam às exigências do trabalho. As palavras a serem analisadas devem possuir ocorrência significativa para serem gerados vetores de qualidade e, também, serem relevantes social e culturalmente, possibilitando uma análise de seus contextos de uso e de formas relacionadas.

Para a análise exploratória do *corpus* foram feitas as seguintes etapas de pré-processamento:

- Tokenização²;
- Padronização em caixa baixa;
- Remoção de *stopwords*;
- Remoção de acentuação.

O texto foi inicialmente tokenizado e, em seguida, foi padronizado completamente em caixa baixa, evitando assim que tokens iguais sejam considerados diferentes devido à escrita em maiúscula. Dessa forma, os tokens “Portanto” e “portanto” correspondem ao mesmo token, “portanto”.

² A tokenização consiste em segmentar o texto em pedaços menores que possuam relevância para a análise. Esses pedaços podem ser frases, palavras, símbolos gráficos ou numéricos, desde que sejam relevantes. O resultado da tokenização é o token, que aqui refere-se ao nível da palavra.

A retirada de *stopwords* é um processo comum em tarefas de PLN. As *stopwords* são palavras gramaticais ou consideradas pouco relevantes semanticamente, como pronomes, preposições e artigos. Podem também ser palavras filtradas por não serem de interesse do trabalho em questão. Sendo assim, frequentemente elas são ignoradas em análises de PLN. Existem diversas listas de Stopwords disponíveis com diferentes critérios, e neste trabalho foi usada a lista fornecida por padrão pela biblioteca NLTK (BIRD; KLEIN; LOPER, 2009).

Após realizados esses passos, foi obtido o número de ocorrências das palavras mais frequentes, o que pode ser visto na Tabela 1.

Tabela 1 - Dez palavras mais frequentes do *corpus* após remoção de *stopwords* e acentuação

Palavra	Número de Ocorrências
senhor	7433
bem	6005
deus	4602
grande	4600
dom	4589
assim	4584
tempo	4265
tudo	4126
pois	3922
fazer	3738

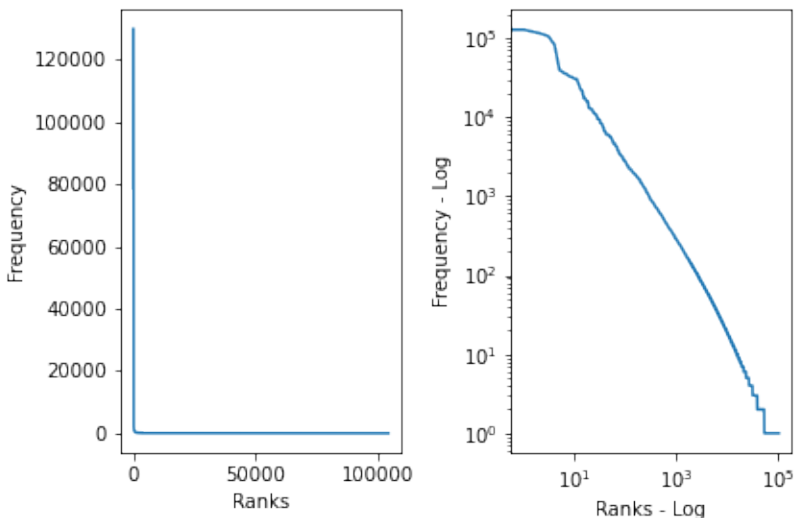
Fonte: Elaboração própria.

Na Figura 10 à esquerda, vê-se a distribuição de frequências para as palavras do *corpus*. Aqui, se vê como poucos termos possuem uma frequência alta, enquanto a grande maioria do *corpus* possui uma frequência similar, mais baixa. Já o gráfico à direita mostra o mesmo gráfico em escala logarítmica, em que se tem uma relação linear entre as potências das frequências e seu posicionamento na ordem de frequência.

A Figura 10 é um exemplo da Lei de Zipf. A Lei de Zipf é uma constatação empírica que nos diz que, a posição, ou ranque, de uma palavra em uma tabela de frequências ordenada de forma decrescente, é inversamente proporcional a sua frequência no *corpus* analisado. Assim, a segunda palavra mais frequente em um *corpus* possui frequência aproximadamente duas vezes menor que a palavra mais frequente e assim

por diante. Em geral, as palavras que aparecem com maior frequência são palavras de função gramatical.

Figura 10 - Distribuição de frequências das palavras do corpus.



Fonte: Elaboração própria.

Por fim, foram encontradas as frequências e o ranque, após a retirada de *stopwords*, para as palavras selecionadas para a análise, indicadas na Tabela 2.

Tabela 2 - Frequência das palavras a serem analisadas e seu ranque

Palavra	Frequência	Ranque
deus	4602	3
homem	2506	32
terra	2209	45
pai	1312	133
mulher	1229	145
mãe	798	240

Fonte: Elaboração própria.

Para as formas “pai”, “mae”, “deus”, “homem”, “mulher” e “terra” foram encontradas palavras de frequência consideradas satisfatórias e

que possivelmente carregam vieses em seus contextos de uso, portanto foram elas as escolhidas para serem analisadas.

A partir de uma inspeção visual inicial, constatou-se que os números de textos presentes nas fases Arcaica e Arcaica Média aparentavam ser muito menores que os números de textos nas demais fases. Assim, já em um primeiro recorte temporal do *corpus* foram unidas as fases Arcaica e Arcaica Média em um só período de tempo, aqui denominado Período I. Vê-se na Tabela 3 que o número de *tokens* para o Período I, mesmo constituindo-se da união de dois outros períodos, ainda é baixo em comparação com os demais períodos. Foi mantida essa separação apesar do desbalanceamento temporal aqui encontrado.

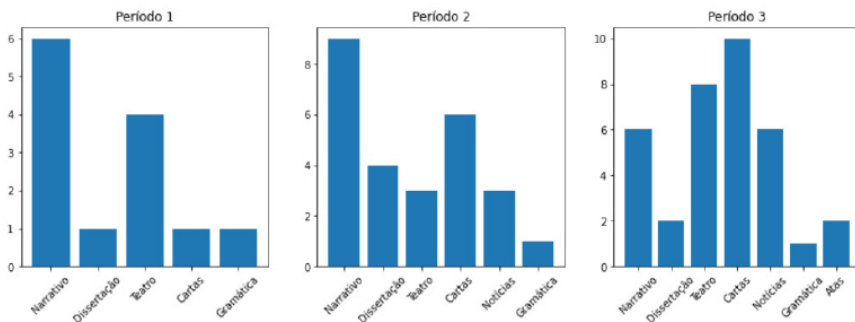
Tabela 3 - Fases do português (adaptado de Bechara (1985)) e respectivo número de *tokens* no *corpus* Tycho Brahe.

Fases	Séculos	Nr. de <i>Tokens</i>
Período I (Arcaica/Arcaica Média)	Até a 1ª metade séc. XVI	632.907
Período II (Moderna)	2ª metade séc. XVI até o fim séc. XVII	1.230.507
Período III (Contemporânea)	Séc XVII até início do séc. XX	1.439.397

Fonte: Elaboração própria.

Com o intuito de se analisar o balanceamento do *corpus* para os diferentes gêneros textuais encontrados, obtém-se a Figura 11, que indica o número de textos para cada gênero textual.

Figura 11 - Gêneros textuais presentes



Fonte: Elaboração própria.

Após serem analisados os balanceamentos do *corpus* e agrupados os textos em recortes temporais, foi feita a limpeza e processamento do *corpus* para a posterior criação dos modelos *Skip-gram*.

6 Limpeza e treinamento dos modelos

Após o agrupamento dos textos, foram realizados os seguintes passos de pré-processamento:

- Tokenização;
- Padronização do texto em caixa baixa;
- Retirada de espaços em branco em excesso;
- Retirada de *Stopwords*;
- Lematização;
- Retirada de acentos gráficos.

A lematização é um processo também comum em tarefas de PLN e consiste em deflexionar uma palavra, retornando-a para sua forma base, dicionarizada. O resultado desse processo são nomes no singular e no masculino, e verbos no infinitivo. Como exemplo de lematização, tem-se a palavra “professoras”, forma plural e feminina, que, após lematizada, torna-se a palavra “professor”, forma singular e masculina. Esse processo é realizado a fim de manter formas que carregam os mesmos significados agrupadas, assim as formas “andei” e “andou” seriam representadas pelo mesmo token “andar”.

A retirada de espaços em branco em excesso se dá para padronização dos espaçamentos e facilitar o processamento dos arquivos de texto.

Todos os passos de limpeza e pré-processamento foram feitos através da biblioteca Spacy³, baseada na linguagem de programação Python em sua versão 3.1.1.

Os modelos Skip-gram foram treinados, um para cada período, através do código fornecido por Mikolov.⁴

Os hiperparâmetros são as condições de treinamento utilizadas, e foram mantidas no padrão. O motivo dessa escolha é justificado pela diminuição nos ganhos com a alteração dos parâmetros, como mencionado por Mikolov *et al.* (2013).

Os hiperparâmetros mais relevantes de treinamento podem ser vistos na tabela 4.

Tabela 4 - Hiperparâmetros de treinamento

Hiperparâmetro	Valor
Tamanho do vetor (size)	300
Janela (window)	8
Amostragem negativa (negative)	25
Amostra (sample)	1e-4
Binário (binary)	1
Iterações (iter)	25

Fonte: Elaboração própria.

O parâmetro *size* diz respeito ao tamanho do vetor, ou número de dimensões que o vetor de cada palavras possuirá após o treinamento. O

³ Disponível em: <https://spacy.io/>

⁴ Disponível em: <https://github.com/tmikolov/word2vec>

parâmetro *window* se refere à janela de treinamento: o seu valor determina o número de *tokens* antes e depois da palavra alvo, para um intervalo total de $16+1$. O parâmetro *negative* corresponde a amostragem negativa, que é o número de exemplos negativos gerados para o treinamento. Um exemplo negativo é, neste caso, uma sequência de palavras que não ocorre no *corpus*. Esses exemplos são gerados ao se substituir uma palavra em exemplo por uma palavra aleatória do *corpus*. Pode-se ver um exemplo negativo na Figura 12, onde a forma *apricot* é emparelhada em contextos reais na coluna da esquerda e em contextos não existentes no *corpus* na coluna da direita. O valor fornecido corresponde à razão entre o número de amostras negativas e positivas: neste caso, tem-se 25 vezes mais amostras negativas que positivas. Por fim, o parâmetro *iter* diz respeito ao número de iterações necessárias para o treinamento: neste caso, o treinamento foi repetido 25 vezes.

Figura 12 - Exemplos positivos e negativos no corpus

Exemplos positivos +		Exemplos Negativos -	
Palavra (n)	Contexto (n+1)	Palavra (n)	Contexto (n+1)
deus	ser	deus	cachorro
deus	governar	deus	sete
deus	irado	deus	meu

Fonte: Elaboração própria

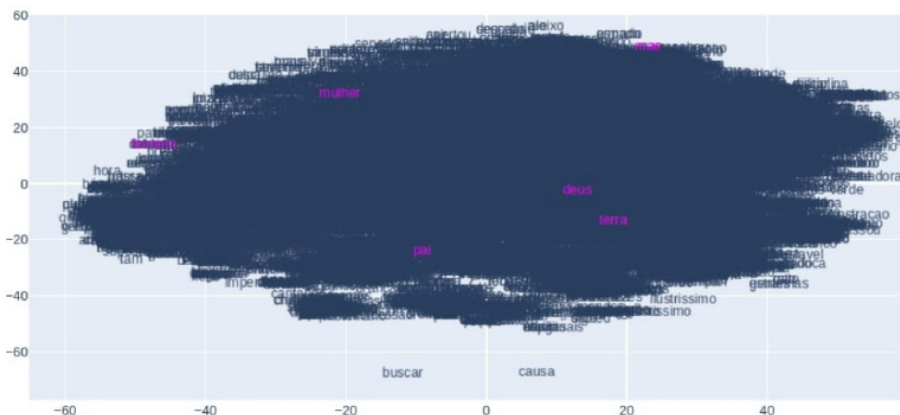
Após a obtenção dos vetores de palavras, resta apenas gerar a visualização gráfica.

7 Visualizações de redes de relações semânticas

Assim como Osgood et al. representam palavras em três dimensões, os modelos de vetorização de palavras aqui utilizados geram vetores com uma dimensão definida no momento do treinamento. Para este trabalho foram criados vetores de 300 dimensões e, por isso, deve-se reduzir essas dimensões a apenas duas para que possam ser visualizadas em um espaço bidimensional. Para isso, foi utilizada a biblioteca Gensim 4.0.1 para Python.

Com ela, foi usado o algoritmo T-SNE para a redução dos vetores de 300 dimensões a apenas duas dimensões e, por fim, os pontos foram plotados em um espaço bidimensional, como mostra a Figura 13. As imagens finais foram obtidas após a devida ampliação da imagem.

Figura 13 - Exemplo de representação gráfica da rede de relações semânticas

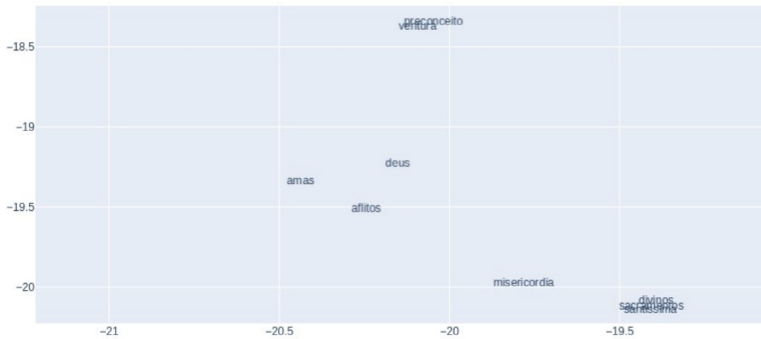


Fonte: Elaboração própria.

Como primeiro resultado, obtiveram-se as redes de relações semânticas das palavras escolhidas inicialmente para cada um dos períodos delimitados no *corpus*.

Primeiramente, o gráfico gerado para a palavra “deus” no período I mostra um contexto similar ao esperado, com palavra como “devocao”, “louvores”, “devotamente” (Figura 14), em sua proximidade. Já para o período II, vê-se palavras como “inspiracoes”, “converte”, “perdi” e “maldade” (Figura 15) próximas. Por fim, no período III têm-se as formas “amas”, “aflitos”, “misericordia”, “divinos”, “preconceito”, “santissima” (Figura 16).

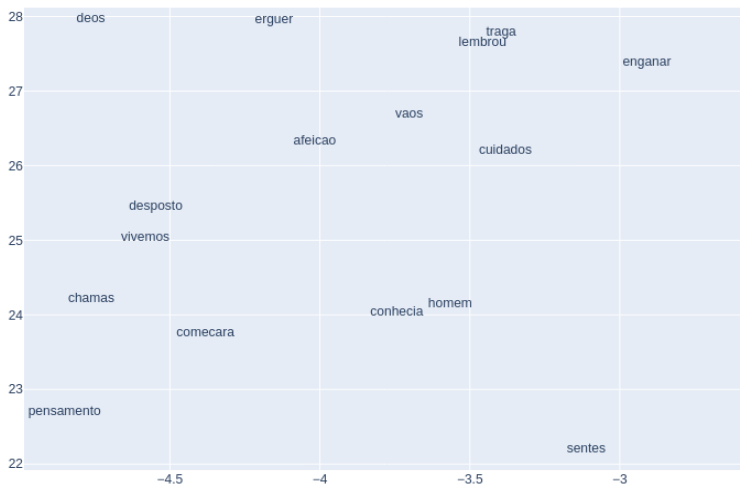
Figura 16 - Rede de relações semânticas da palavra “deus”, período III



Fonte: Elaboração própria

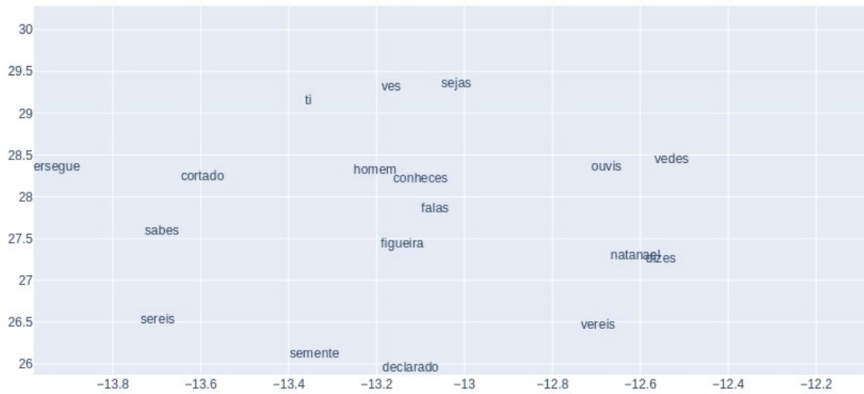
Agora, as relações semânticas para a forma “homem”, no período I vemos as palavras “conhecia”, “cuidados”, “sentes”, “comecara” mais próximas (Figura 17). Já para o período II têm-se as palavras “conheces”, “falas”, “figueira”, “ouvis”, “ves”, “sejas” (Figura 18). No Período III vê-se as palavras “miseravel”, “ateu”, “ímpio”, “solene”, “livrar” (Figura 19).

Figura 17 - Rede de relações semânticas da palavra “homem”, período I



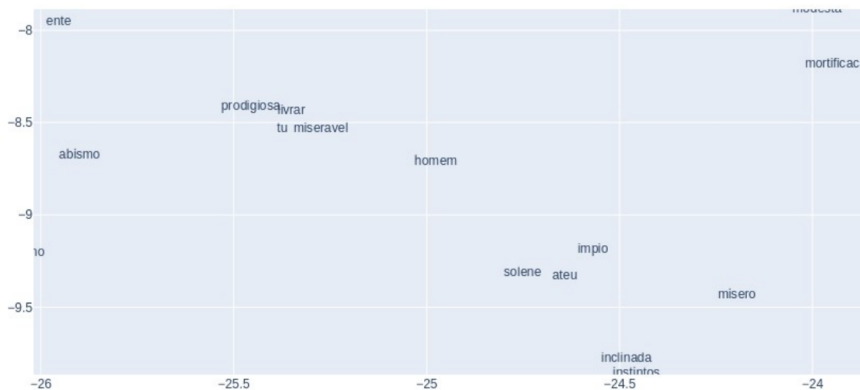
Fonte: Elaboração própria.

Figura 18 - Rede de relações semânticas da palavra “homem”, período II



Fonte: Elaboração própria.

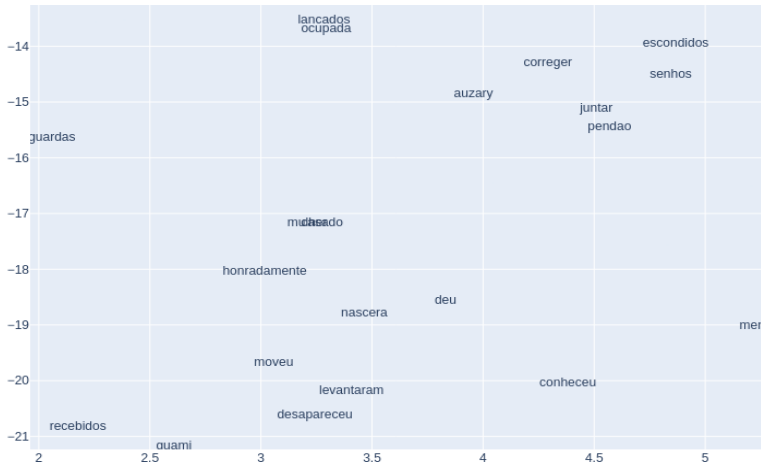
Figura 19 - Rede de relações semânticas da palavra “homem”, período III



Fonte: Elaboração própria.

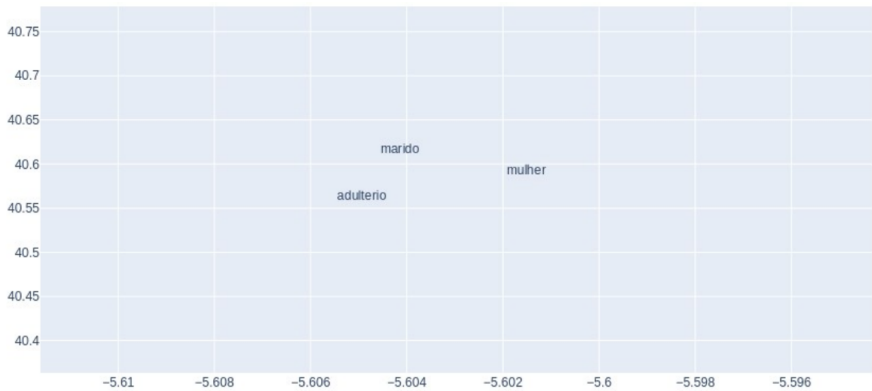
Para a rede da palavra “mulher”, vemos no período I, que ela ficou bastante próxima da palavra “casado”, seguida das palavras “nascera”, “deu”, “nascera”, “honradamente” (Figura 20). Já no período II, temos as palavras “marido” e “adultério” muito próximas (Figura 21). Finalmente no período III, temos as palavras “desgraçada”, “coitadinha”, “marido” e “margarida” na proximidade da palavra analisada (Figura 22).

Figura 20 - Rede de relações semânticas da palavra “mulher”, período I



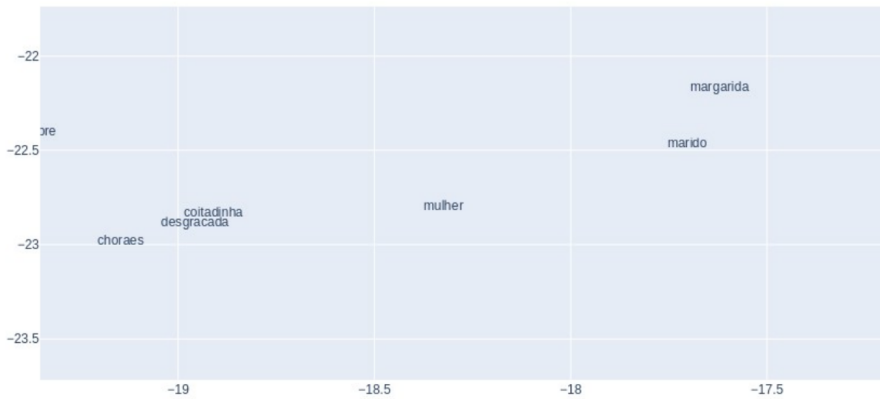
Fonte: Elaboração própria

Figura 21 - Rede de relações semânticas da palavra “mulher”, período II



Fonte: Elaboração própria.

Figura 22 - Rede de relações semânticas da palavra “mulher”, período III



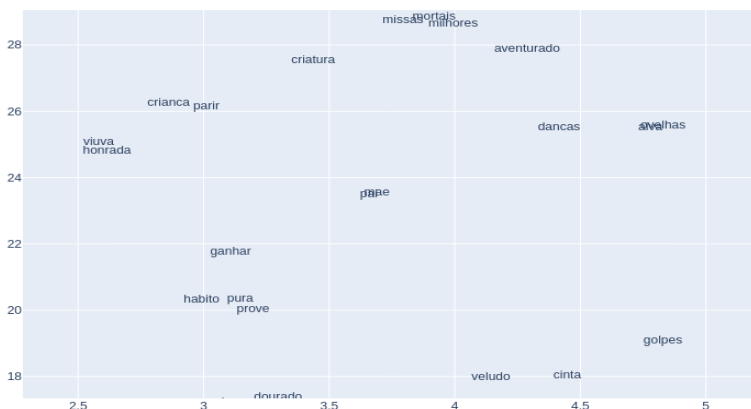
Fonte: Elaboração própria.

As formas “pai” e “mae” foram analisadas em conjunto devido à proximidade que elas se encontram nos resultados do período I. Nele, elas se sobrepõem, como mostrado na Figura 23. Temos as palavras “criatura”, “ganhar”, “cinta”, “pura”, “missas”, “parir”, “crianca”, “viuva” na proximidade das palavras analisadas.

Já para o período II, as palavras foram analisadas separadamente. A palavra “pai” possui em sua proximidade as palavras “testemunho”, “conheceis”, “credes”, “guardado”, “enviou” em sua vizinhança, como mostra a Figura 24. Já a palavra “mae” se encontra próxima de expressões como “casados”, “bodas”, “legítimo”, “casal”, “embaracos” e “ajustada”, como visto na Figura 25.

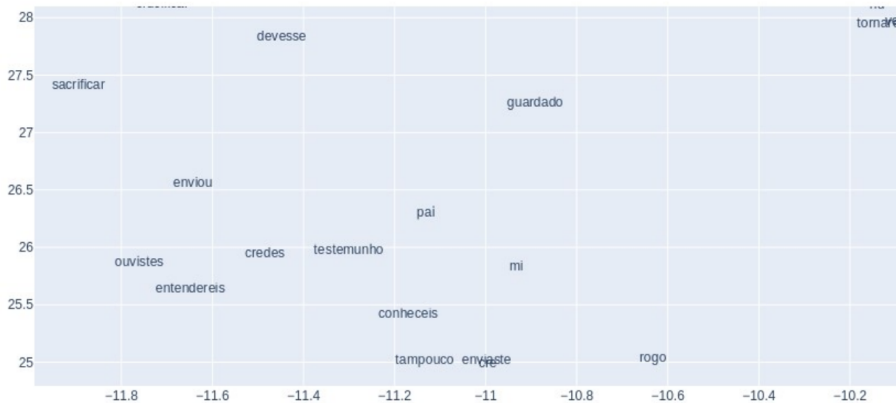
No período III, a rede da palavra “pai” mostra proximidade com “consolar”, “paterno”, “filho”, além do verbo “tourear”, Figura 26. Já o campo de “mae”, mostra as palavras “virtuosa”, “filha”, “carinhosa”, vide Figura 27.

Figura 23 - Rede de relações semânticas das palavras “pai” e “mãe”, período I.



Fonte: Elaboração própria.

Figura 24 - Rede de relações semânticas da palavra “pai”, período II



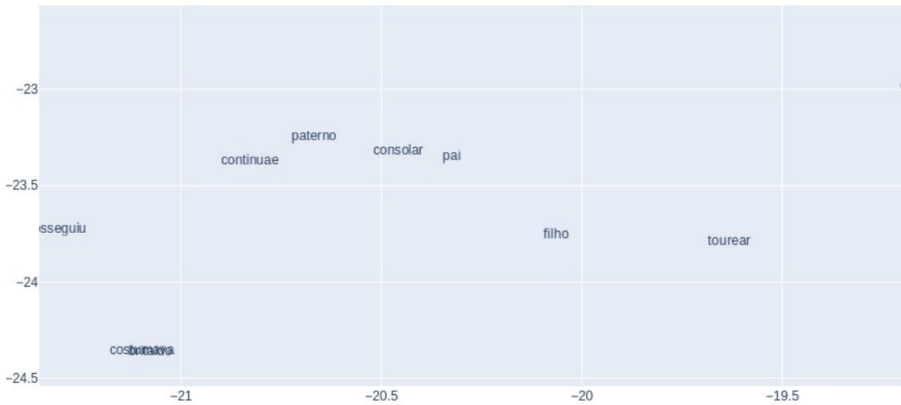
Fonte: Elaboração própria.

Figura 25 - Rede de relações semânticas da palavra “mae”, período II



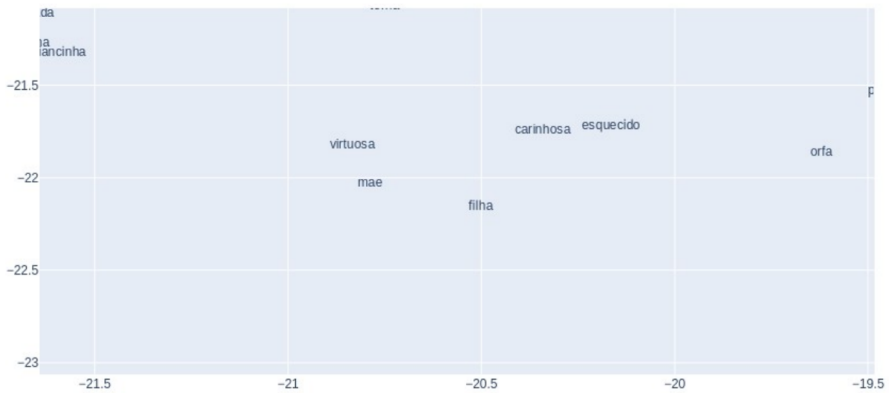
Fonte: Elaboração própria

Figura 26 - Rede de relações semânticas da palavra “pai”, período III



Fonte: Elaboração própria.

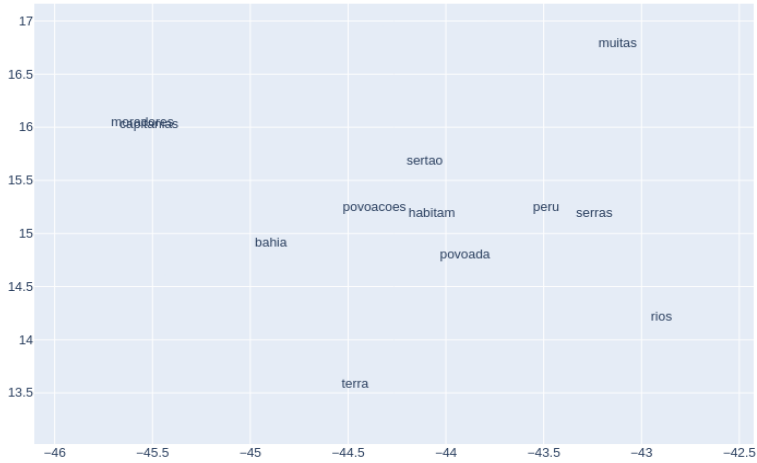
Figura 27 - Rede de relações semânticas da palavra “mae”, período III



Fonte: Elaboração própria.

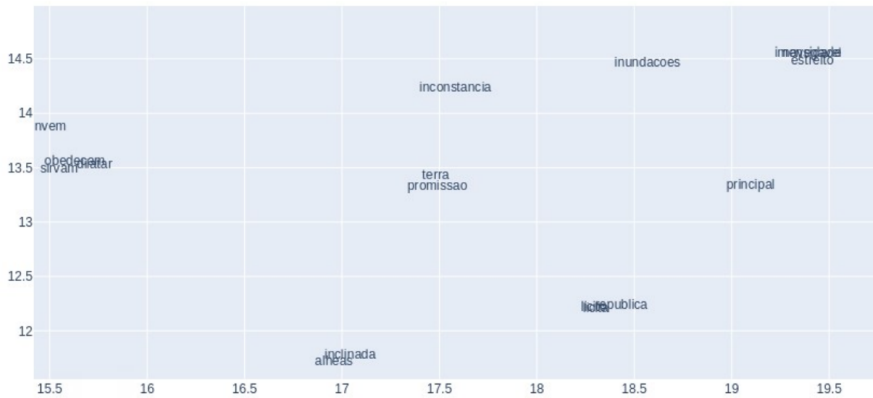
Para o período I, a palavra “terra” possui em sua vizinhança as palavras “sertao”, “habitam”, “povoada”, “bahia”, “rios”, conforme indicado na Figura 28. Já para o período II a Figura 29 mostra as palavras “inconstancia”, “promissao” e destaca-se a palavra “republica”. Finalmente o período III apresenta palavras como “campinas”, “montanhas”, “ribeiras”, “ventos” e “tempestades”, como visto na Figura 30.

Figura 28 - Rede de relações semânticas de “terra”, período I



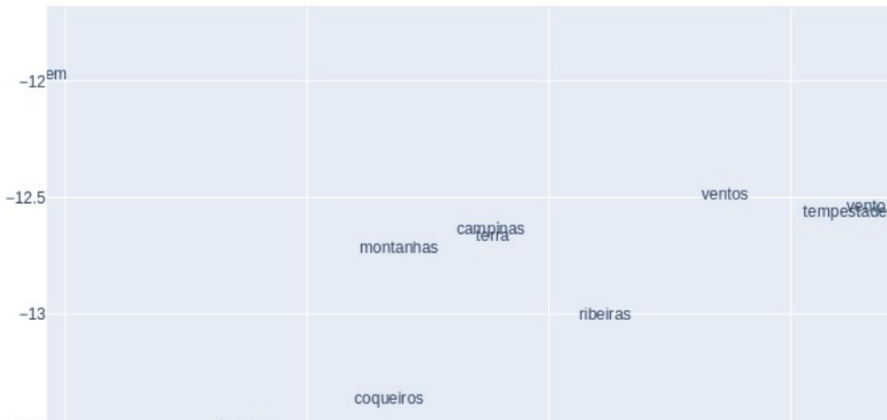
Fonte: Elaboração própria

Figura 29 - Rede de relações semânticas de “terra”, período II



Fonte: Elaboração própria

Figura 30 - Rede de relações semânticas para “terra”, período III



Fonte: Elaboração própria.

8 Discussão dos resultados

Inicialmente, vê-se que as redes de relações passaram por mudanças ao longo dos períodos analisados. Observa-se que a palavra “deus” surge em contextos ligados a religião, o que é um resultado esperado.

A palavra “homem” não vem acompanhada, em um primeiro momento, de palavras que aparentam estar fortemente ligadas a ela. Já no segundo período observa-se essa palavra associada a verbos sensoriais como “vedes”, “ouvis”, “falas”. Essa ausência de palavras fortemente ligadas à forma “homem” pode ser uma consequência de seu uso como forma de indeterminação do sujeito nesses períodos, como informado por Bechara (1985). Já no último momento, vê-se a forma “homem” próxima a palavras de viés negativo, como “miserável” e “ímpio”, e também à palavra “ateu” e à palavra com viés positivo “solene”.

A palavra “mulher” aparece inicialmente em contexto não muito definido, com as palavras “honradamente”, “nascera”, “deu”, “moveu”. As poucas palavras encontradas para esse contexto podem se dar pela preferência por palavras como “rapariga” para se referenciar a mulher jovem. Já para o período II, é interessante notar a relação próxima dos termos “marido”, “mulher” e “adulterio”. Por fim, a palavra “mulher” continua próxima da palavra “marido”, mas também na vizinhança de termos como “coitadinha” e “desgraçada”. Apesar da falha do lematizador

em deflexionar essas expressões, elas apareceram próximas à palavra “mulher”, mostrando a relação com o “gênero”.

As palavras “pai” e “mãe” aparecem bastante próximas em sua rede de relações. Vê-se palavras relacionadas a família em suas proximidades, como “criança”, “viúva” e também o verbo “parir”. Em seguida, para o segundo período, a palavra “pai” aparenta estar mais relacionada a um contexto religioso, como nas palavras “testemunho”, “credes”, “rogo” em sua proximidade. Já “mae”, no mesmo período, mostra palavras relacionadas a família e casamento, como “casados”, “casal”, “prole”, “bodas”, “primogenito”. Por fim, no período III, a palavra “pai” aparece ligada a palavras relacionadas ao contexto familiar, como “paterno” e “filho”. Já a palavra “mãe” também aparece ligada a contextos familiares, mostrando a palavra “filha” e “orfa” em sua proximidade. Também se vê os adjetivos “virtuosa” e “carinhosa” próximos.

Por último, a palavra “terra” apresenta, no período I, uma proximidade maior com termos relacionados a contextos geográficos, como “bahia” e “sertão”. Já no período II, a palavra parece ser encontrada em contextos diferentes, tendo em vista o surgimento da palavra “republica”. E por fim ela se encontra novamente relacionada a termos geográficos como “campinas”, “montanhas”, “ventos”, “areia”.

Os resultados apresentados mostram-se de qualidade variável. Em alguns casos, como no da palavra “terra”, apesar de apresentar apenas 2209 ocorrências no *corpus*, foi possível visualizar uma mudança no seu uso dentro dos três períodos analisados. Já a forma “deus” não apresentou variação notável em sua rede. Apesar disso, o fato de essa forma estar sempre presa ao contexto religioso surge como forma de validar o modelo, o que nem sempre é possível se fazer de forma quantitativa. Por fim, a palavra “mulher” encontra-se associada a formas como “honrada”, “nasceu”, “coitadinha”, “desgraçada” e “adultério”, mostrando de certa forma as diferentes percepções ao longo dos períodos analisados.

Os processos de mudança (ou manutenção) semântica analisados anteriormente estão em conformidade com as propostas de Givón: palavras de sentido semelhantes foram agrupadas em regiões próximas nos modelos. Não foi possível, entretanto, verificar alguma mudança drástica de sentido, até porque essas palavras não sofrem necessariamente uma mudança de sentido ao longo do tempo. O que pode ser analisado,

porém, é a vizinhança dessas palavras e, a partir disso, examinar como as formas estão organizadas no léxico disponível no *corpus*.

Em um caráter mais técnico, o processo de lematização não foi eficiente. Encontram-se formas verbais flexionadas, como “vedes”, “sejas”, “sabes”, “amas” e formas nominais apresentando o gênero feminino e flexão de grau como em “coitadinha”, “desgraçada”, enquanto eram esperadas suas formas dicionarizadas. O lematizador utilizado é fornecido pela biblioteca Spacy, que realiza o processo automaticamente e possui acurácia relativamente baixa (76%). Além disso, a lematização do *corpus* diacrônico sofreu também por possuir formas desconhecidas ao modelo, como a palavra “molher” e o verbo “cazar”, já que o modelo utiliza um conjunto de regras para gerar os lemas.

9 Conclusão

O presente trabalho buscou analisar palavras significativas e os diferentes contextos semânticos em que elas surgem em diferentes períodos de um *corpus* diacrônico. A análise foi realizada por meio do uso da técnica de PLN conhecida como *word embeddings* (vetorização de palavras), que permite agrupar palavras de sentido próximas em um espaço vetorial e visualizar seus vizinhos.

Destaca-se, aqui, a importância do processo de lematização, que permite agrupar palavras flexionadas dentro da mesma forma, tornando, assim, a análise mais precisa. Esse processo é de extrema importância para línguas de morfologia rica, como o português, e não recebe tanta atenção devido aos sistemas de PLN.

serem desenvolvidos principalmente para o inglês, que possui morfologia mais pobre.

Os obstáculos para um melhor resultado se dão principalmente pelo tamanho do *corpus*. Os resultados apresentados por Hamilton, Leskovec e Jurafsky (2016) foram obtidos por meio do uso de um *corpus* que possui mais de 410 milhões de tokens, totalizando mais de 100 milhões de tokens por período estudado – um valor muito distante da quantidade de tokens obtidos com o uso do *corpus* Tycho Brahe. Apesar de os autores citarem outras formas de obtenção de *word embeddings* e, também, recomendarem seu uso para *corpora* menores, não há uma forma definitiva de se determinar o que é um *corpus* “pequeno” ou “grande”, sendo esse conceito determinado pela tarefa a se realizar.

Por fim, o caráter misto do *corpus* também influenciou os resultados. O *corpus* Tycho Brahe possui tanto textos de cartas, poemas e peças de teatro, quanto textos jornalísticos. Além disso, esses diferentes gêneros encontram-se desbalanceados quanto à sua representação no *corpus*.

Considerando as limitações encontradas nesta análise, propõe-se para estudos futuros o uso de formas alternativas de obtenção de *word embeddings*, comparando os resultados com os aqui obtidos. Além disso, sugere-se o uso de outro lematizador, que possa fornecer um resultado mais satisfatório. Pode-se, também, realizar novos recortes temporais e analisá-los a fim de buscar diferentes relações semânticas.

Todos os desafios citados anteriormente decorrem do caráter pioneiro do presente estudo para o português, não existindo, até o momento e até onde os autores constataram, uma análise quantitativa que utilize metodologia similar para dados diacrônicos nessa língua.

Agradecimentos

Os autores agradecem à Profa. Adriana Pagano (FALE/UFMG) e ao Prof. Bruno Rocha (FALE/UFMG) pela participação na banca de avaliação da monografia que levou a este artigo e pelas valiosas sugestões ao trabalho.

Declaração de autoria

Ambos os autores conceberam e planejaram o estudo. Lucas F. Lage realizou o pré-processamento dos dados, gerou as análises, interpretou os resultados e escreveu o texto. Evandro L. T. P. Cunha orientou a realização do trabalho, auxiliou na interpretação dos resultados e revisou o texto. Ambos aprovaram a versão final.

Referências

BĂLAN, O. *et al.* Emotion classification based on biophysical signals and machine learning techniques. *Symmetry*, [s.l.], v. 12, n. 1, 2020. DOI: <<https://doi.org/10.3390/sym12010021>>. Acesso em: 31 jan. 2022.

BECHARA, E. *As fases históricas da língua portuguesa: tentativa de proposta de nova periodização*. Niterói: Universidade Federal Fluminense, 1985.

BIRD, S.; KLEIN, E.; LOPER, E. *Natural language processing with Python: analyzing text with the natural language toolkit*. [s.l.]: O'Reilly Media, Inc., 2009.

BOCHKAREV, V.; SOLOVYEV, V.; WICHMANN, S. Universals versus historical contingencies in lexical evolution. *Journal of The Royal Society Interface*, [s.l.], v. 11, n. 101, 2014. DOI: <<http://dx.doi.org/10.1098/rsif.2014.0841>>. Acesso em: 31 jan. 2022

CAMBRAIA, C. N. Da lexicologia social a uma lexicologia sócio-histórica: caminhos possíveis. *Revista de Estudos da Linguagem*, Belo Horizonte, v. 21, n. 1, p. 157–188, 2013. DOI: <<http://dx.doi.org/10.17851/2237-2083.21.1.157-188>>. Acesso em: 31 jan. 2022.

CUNHA, A. F. DA. Funcionalismo. In: MARTELOTTA, M.E. (org.). *Manual de linguística*. São Paulo: Contexto, 2008. p. 157–176.

DE SOUSA, M. C. P. O *corpus* Tycho Brahe: contribuições para as humanidades digitais no Brasil. *Filologia e Linguística Portuguesa*, São Paulo, v. 16, n. esp., p. 53–93, 2014. DOI: <https://doi.org/10.11606/issn.2176-9419.v16issep53-93>

DORES, M. V. P. das; TOLEDO, C. V. S. De “lepra” a “hanseníase”: uma análise lexicológica de base sócio-histórica. *Diacrítica*, [s.l.], v. 32, n. 1, p. 179–208, 2018.

FIRTH, J. R. A synopsis of linguistic theory, 1930-1955. In: FIRTH, J. R. *Studies in linguistic analysis*. Oxford: Blackwell, 1957. p. 1-32.

GALVES, C.; ANDRADE, A. L.; FARIA, P. *Tycho Brahe Parsed Corpus of Historical Portuguese*. Disponível em <<http://www.tycho.iel.unicamp.br/~tycho/corpus/texts/psd.zip>>. Acesso em: dez. 2017.

GIVÓN, T. *Functionalism and grammar*. [s.l.]: John Benjamins Publishing, 1995.

GIVÓN, T. *Syntax: an introduction*. v. 1. [s.l.]: John Benjamins Publishing, 2001.

GRIFFIN, C. *Graph Theory*: Penn State Math 485 Lecture Notes. 2017. Disponível em: <<https://www.personal.psu.edu/cxg286/Math485.pdf>>. Acesso em: 22 fev. 2022.

HAMILTON, W. L.; LESKOVEC, J.; JURAFSKY, D. Diachronic word embeddings reveal statistical laws of semantic change. In: 1ST INTERNATIONAL CONFERENCE ON LEARNING

REPRESENTATIONS, 2016. Disponível em: <<https://arxiv.org/abs/1605.09096>>. Acesso em: 31 jan. 2022.

HARTMANN, N. S. Solo queue at ASSIN: Combinando abordagens tradicionais e emergentes. *Linguamática*, [s.l.], v. 8, n. 2, p. 59-64, 2016

JURAFSKY, D.; MARTIN, J. H. *Speech and Language Processing: An introduction to speech recognition, computational linguistics and natural language processing*. Upper Saddle River: Prentice Hall, 2008.

LABOV, W. Some principles of linguistic methodology. *Language in Society*, [s.l.], v. 1, n. 1, p. 97-120, 1972. DOI: <<https://doi.org/10.1017/S0047404500006576>>

MATORÉ, G. La lexicologie sociale. *L'Information Littéraire*, Paris, n. 2, mar./abr. 1949.

MICHEL, J. B. *et al.* Quantitative analysis of culture using millions of digitized books. *Science*, [s.l.], v. 331, n. 6014, p. 176-182, 2011. DOI: <[10.1126/science.1199644](https://doi.org/10.1126/science.1199644)> Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3279742/>>. Acesso em: 31 jan. 2022.

MIKOLOV, T. *et al.* Efficient estimation of word representations in vector space. In: 1ST INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS, ICLR 2013.

OSGOOD, C. E.; SUCI, G. J.; TANNENBAUM, P. H. *The measurement of meaning*. [s.l.] University of Illinois Press, 1957.

RAFAEL, G. C. R. A.; SIMIÃO, D. P. Aidético e soropositivo: análise sócio-histórica da concorrência entre qualificadores utilizados em referência a portadores do HIV. *Inventário*, n. 23, p. 45-68, 2019.

ROBIN, R.; DE MENESES BOLLE, A. B. *História e lingüística*. São Paulo: Editora Cultrix, 1977.

SWINNEY, D. A. Lexical access during sentence comprehension: (Re) consideration of context effects. *Journal of verbal learning and verbal behavior*, v. 18, n. 6, p.645-659, 1979. DOI: [https://doi.org/10.1016/S0022-5371\(79\)90355-4](https://doi.org/10.1016/S0022-5371(79)90355-4)

ZIPF, G. K. *Human behavior and the principle of least effort: An introduction to human ecology*. [s.l.] Books, 2016.