

UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE LETRAS
DEPARTAMENTO DE LINGUÍSTICA, PORTUGUÊS E LÍNGUAS CLÁSSICAS-
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA

HUMBERTO BORGES

**A SINTAXE DO SUJEITO NA HISTÓRIA DO PORTUGUÊS EM GOIÁS:
EVIDÊNCIAS OITOCENTISTAS DE UMA LÍNGUA DE SUJEITO NULO PARCIAL**

BRASÍLIA

2019

UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE LETRAS
DEPARTAMENTO DE LINGUÍSTICA, PORTUGUÊS E LÍNGUAS CLÁSSICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA

HUMBERTO BORGES

**A SINTAXE DO SUJEITO NA HISTÓRIA DO PORTUGUÊS EM GOIÁS:
EVIDÊNCIAS OITOCENTISTAS DE UMA LÍNGUA DE SUJEITO NULO PARCIAL**

Tese de doutorado apresentada ao Programa de Pós-Graduação em Linguística da Universidade de Brasília (PPGL/UnB) como requisito obrigatório para a obtenção do título de doutor em Linguística.

Área de concentração: Teoria e análise linguística

Orientadora: Profa. Dra. Rozana Reigota Naves (PPGL/UnB)

Coorientador: Prof. Dr. Acrísio Pires (University of Michigan)

BRASÍLIA

2019

Ficha catalográfica elaborada automaticamente,
com os dados fornecidos pelo(a) autor(a)

BH919s Borges, Humberto
A sintaxe do sujeito na história do português em Goiás: evidências
oitocentistas de uma língua de sujeito nulo parcial / Humberto Borges;
orientador Rozana Reigota Naves; co-orientador Acrísio Pires. --
Brasília, 2019.
273 p.

Tese (Doutorado - Doutorado em Linguística) --
Universidade de Brasília, 2019.

1. Sujeitos nulos. 2. Inversão do sujeito. 3. Sintaxe diacrônica.
4. Movimento do verbo. 5. Se-passivo, se impessoal. I. Naves,
Rozana Reigota, orient. II. Pires, Acrísio, co-orient. III. Título.

HUMBERTO BORGES

**A SINTAXE DO SUJEITO NA HISTÓRIA DO PORTUGUÊS EM GOIÁS:
EVIDÊNCIAS OITOCENTISTAS DE UMA LÍNGUA DE SUJEITO NULO PARCIAL**

Tese de doutorado apresentada ao Programa de Pós-Graduação em Linguística da Universidade de Brasília (PPGL/UnB) como requisito obrigatório para a obtenção do título de doutor em Linguística.

Data da defesa: 30/01/2019.

Conceito: Aprovada.

BANCA EXAMINADORA

Profa. Dra. Rozana Reigota Naves

Universidade de Brasília (UnB) – Presidente/Orientadora

Profa. Dra. Márcia Santos Duarte de Oliveira

Universidade de São Paulo (USP) – Membro Externo

Profa. Dra. Maria Cristina Figueiredo Silva

Universidade Federal do Paraná (UFPR) – Membro Externo

Profa. Dra. Heloísa Maria Moreira Lima de Almeida Salles

Universidade de Brasília (UnB) – Membro Interno

Prof. Dr. Hildo Honório do Couto

Universidade de Brasília (UnB) – Suplente

DEDICATÓRIA

Porque d'Ele, e por Ele, e para Ele são todas as coisas; glória, pois, a Ele eternamente. Amém! (Romanos 11:36). Para as pessoas que mais amo no mundo: *Idalina*, minha avó; *Adriana*, minha mãe; *Isadora* e *Isabela*, minhas irmãs; *Ana Cristina*, *Josué*, *Nathália*, *Ana Vitória* e *Luiz Fernando*, minhas sobrinhas e meus sobrinhos. Ao meu pai, *José Horácio*, por ter sido presente em momentos cruciais.

Fica claro (...) que um novo sistema gramatical – chame-se-o de gramática brasileira ou de dialeto com sua própria configuração uma vez se tratar de uma questão meramente ideológica – emergiu ao final do século XIX, estabelecendo uma nova gramática radicalmente diferente da modalidade lusitana (...) Certamente, estes novos traços gramaticais entraram na língua no final dos anos 1800 porque circunstâncias sociais especiais aconteciam naquele momento da história externa. Isso significa que não descartamos a hipótese que essas mudanças poderiam ter ocorrido antes da virada do século. E isto também não significa que nossa evidência se encontra enviesada pelos dados. Certamente que não! Nosso argumento é que as circunstâncias sociais antes da virada do século podem não ter sido suficientemente satisfatórias para que a pena brasileira começasse a escorrer sua própria tinta. E neste sentido fica comprovado quão importante é o exame de dados linguísticos à luz de evidências sociais. Sem vias de dúvidas, entretanto, pode ser afirmado que o cidadão brasileiro já estava de posse, ao final do século XIX, de sua própria língua/gramática. Fernando Tarallo (1993 [2018], p. 78-79).

AGRADECIMENTOS

Agradeço especialmente a Profa. Dra. Rozana Reigota Naves. Esta tese é, em grande medida, fruto de sua leitura atenta e segura, de seus direcionamentos, de seu comprometimento e de seu espírito conciliador e mediador. Deixo registrado aqui meu muitíssimo obrigado não somente pela orientação desta tese, mas também por ter me orientado desde a graduação, impactando de forma tão positiva minha história de vida.

Agradeço o Prof. Dr. Acrísio Pires por me supervisionar tão atentamente durante meu estágio de doutorando em linguística na Universidade de Michigan, em 2016, e por compartilhar tão generosamente seu conhecimento acadêmico e profissional comigo. Seu entusiasmo e seu comprometimento foram essenciais para o desenvolvimento e a conclusão desta tese. Muitíssimo obrigado.

Agradeço as professoras doutoras Heloísa Maria Moreira Lima de Almeida Salles (UnB), Márcia Santos Duarte de Oliveira (USP) e Maria Cristina Figueiredo Silva (UFPR), membros da banca examinadora desta tese. Os comentários, as sugestões, os direcionamentos e os ensinamentos transmitidos ali foram muito bem-vindos e essenciais para uma melhor versão final do texto. Sinto-me grato e honrado por ter tido meu trabalho avaliado por vocês.

Agradeço o Prof. Emérito Hildo Honório do Couto (UnB) pelas contribuições dadas para o desenvolvimento desta tese durante meu Exame de Qualificação de Doutorado. Foi um momento gratificante e de muito aprendizado para mim.

Agradeço as professoras doutoras Aveliny Mantovan Lima-Gregio, Eloísa Nascimento da Silva Pilati, Helena da Silva Guerra Vicente, Heloísa Maria Moreira Lima de Almeida Salles, Josênia Antunes Vieira e Rozana Reigota Naves, pelas disciplinas ministradas no PPGL/UnB. Todas foram essenciais para meu desenvolvimento como pesquisador e docente.

Agradeço o Prof. Dr. Marcus Lunguinho (UnB) por sempre ter estado disposto a compartilhar seu conhecimento com os alunos do PPGL/UnB.

Agradeço o Prof. Dr. José Ferrari Neto (UFPB) pelas lições iniciais e essenciais em estatística aplicada a dados linguísticos.

Agradeço os amigos e os colegas discentes do PPGL/UnB pelas diversas parcerias nos meus cinco anos de doutorado: Alzira Neves Sandoval, Bruno Pilastre, Cristiany Fernandes, Edite Consuelo, Giovana Santiago, Layane Rodrigues Santos, Stefania Zandomênicó, Wallace Barboza e, em especial, Paula Guedes Baron, Letícia da Cunha Silva e Zenaide Dias Teixeira, esta última minha parceira de café e bate-papos acadêmicos formidáveis.

Agradeço os profissionais do Instituto de Pesquisas e Estudos Históricos do Brasil Central, da Pontifícia Universidade Católica de Goiás (PUC-GO), pela atenciosa recepção e pelo auxílio durante meu trabalho de campo nos anos de 2013 e 2015 para a aquisição digital de manuscritos históricos produzidos em Goiás.

Registro um agradecimento especial a Kallynny Richelly do Amaral Cardoso, Laís Mikeline Crisostomo e Keila Meireles dos Santos, amigas que acompanharam e pacientemente me ouviram falar sobre esta tese e todas as questões que a circunscreveram durante os últimos anos.

Também registro um agradecimento especial aos amigos do Bloco K da Colina-UnB: Andressa Dantas da Silveira, Aniceto Mateus, Carolina Alzate Gouzy, Christe Héliida Montijo, Daniel Sebastián Granda Henao, Flor de Maria Silvestre Estela, Sabrina Magaly Navas Cajamarca e Wildes Souza Andrade.

Agradeço a equipe de professores e de gestores e meus ex-alunos do CEPI José Feliciano Ferreira (Jataí-GO). Se não houvéssemos tido esse encontro, ainda que tão ligeiro, eu teria muito menos a oferecer na escrita desta tese.

ACKNOWLEDGEMENTS

Part of this research was conducted during a one-year visit to the Department of Linguistics at University of Michigan, under the supervision of Professor Acrisio Pires, to whom I am extremely grateful and who I already thanked in Portuguese.

In my time at the University of Michigan, my academic background and this dissertation also benefited from the collaboration of other scholars. Hence, I am deeply indebted to Professor Marlyse Baptista for her theoretical guidance, straightforward answers to my questions and amazing classes on creole languages and the Caribbean identities.

I am also deeply indebted to Professor Samuel Epstein and Professor Sarah Thomason for their unforgettable and excellent classes on minimalist syntax and language contact, respectively. Professor Sarah Thomason also gave me useful insights into my analytical approach. Thank you so much!

I cannot forget that Talisha Winston was so very helpful with my visiting graduate student application and the visa process. Thank you.

Adrian Bokjo and Vinícius Monteiro de Castro were kind friends during my time in Ann Arbor. Thank you.

Finally, I would like to express my gratitude to Christopher Anderson and his family, Ketan Broder and Susan Okolski. You really know how to make America welcoming to foreigners. Many thanks, guys!

FINANCIAMENTO

Os estudos que resultaram nesta tese de doutorado foram majoritariamente financiados pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) por meio da concessão de duas bolsas de estudo:

- i) bolsa de doutorado no Brasil; e
- ii) bolsa de doutorado-Sanduíche no exterior (processo nº: 007549/2015-06).

Após o tempo regulamentar das bolsas CAPES, também houve financiamento por parte da Fundação de Apoio à Pesquisa do Distrito Federal (FAP-DF), com a concessão de bolsa de doutorado no país.

RESUMO

Esta tese descreve as primeiras mudanças atestadas nas propriedades de sujeitos nulos no português brasileiro, analisando dados originais de manuscritos dos períodos colonial e imperial escritos em Goiás e fornece evidência empírica da perda de sujeitos nulos na gramática do português brasileiro em dados históricos do século XIX em comparação com dados do século XVIII, precedendo o período para o qual mudanças foram relatadas em relação a outras variedades do português brasileiro (cf. M. E. DUARTE, 1993 [2018], 2000). Trata, portanto, da evidência mais antiga de uma variedade do português brasileiro com propriedades de uma língua de sujeito nulo parcial. A análise do *corpus*, composto por textos no formato diário, mostra um aumento significativo na realização de sujeitos definidos manifestos de um século para o outro. Além disso, a livre inversão do sujeito, uma propriedade comum às línguas de sujeito nulo, cai de 51,72% no século XVIII para apenas 17,41% no século XIX, ficando restrita a verbos de natureza locativa e existencial. Tomamos isso como evidência para a hipótese de que a perda de sujeitos nulos no português brasileiro em Goiás pode estar crucialmente ligada à perda de inversão livre do sujeito, e não à perda de concordância verbal. Verificamos que um potencial empobrecimento do paradigma verbal parece não ter desempenhado um papel significativo no aumento de sujeitos manifestos: menos de 12% das orações com sujeito manifesto ou nulo no plural não exibiram concordância explícita entre o verbo e o sujeito nos dados do século XIX, e todos os casos de ausência de concordância ocorreram estritamente com verbos inacusativos ou existenciais. De acordo com a distinção tipológica de línguas de sujeito nulo consistentes e línguas de sujeito nulo parcial de Holmberg (2005, 2010a), propomos que um traço de definitude presente no núcleo funcional temporal do português setecentista em Goiás era valorado pelo movimento do verbo para esse núcleo, gerando uma língua de sujeito nulo consistente e a ordenação verbo-sujeito nas orações declarativas dessa língua. Por outro lado, argumentamos que a ausência de um traço de definitude no núcleo funcional temporal do português oitocentista em Goiás foi o responsável pela queda de sujeitos nulos constatada nessa variedade. A partir desses achados, constatamos que o português brasileiro em Goiás emergiu como uma língua de sujeito nulo parcial já na segunda metade do século XIX. Por fim, adotando a proposta teórica de Lightfoot (2017b) para a mudança sintática, que prevê descontinuidades entre as estruturas sintáticas usadas de uma geração para a outra, e os procedimentos analíticos de Thomason (2001) para estabelecer casos de mudança induzida por contato linguístico, sugerimos que a emergência da gramática do PB em Goiás pode ser vista como um resultado

da situação de contato entre a língua portuguesa e línguas nigero-congolesas, especialmente o quimbundo, na América portuguesa, embora essa interpretação não seja exaustiva e estudos mais aprofundados tenham de ser desenvolvidos.

Palavras-chave: Sujeitos nulos. Inversão do sujeito. Sintaxe diacrônica. Movimento do verbo. *Se*-passivo. *Se*-impessoal.

ABSTRACT

This dissertation reports the discovery of a variety of Brazilian Portuguese with properties of a partial null subject language in the second half of the 19th century. That is, this dissertation shows the earliest attested changes in the properties of null subjects in a variety of Brazilian Portuguese. We analyze original data from colonial and imperial period manuscripts written in Goiás, a state located in Brazil's center-west region, and provide empirical evidence of the loss of null subjects in BP grammars in historical data between the 18th and 19th century, preceding the period for which these changes have been reported regarding other varieties of Brazilian Portuguese (cf. M. E. DUARTE, 1993 [2018], 2000). The analysis of our 18th and 19th century corpus – composed of journals – shows an early significant rise in the realization of overt subjects. In addition, the verb-subject (VS) free inversion, a property common to null subject languages, drops from 51,72% in the 18th century to only 17,41% in the 19th century, being restricted to locative and existential verbs. We take this as evidence that the loss of null subjects in BP was primarily linked to the loss of free VS-inversion, and not to the loss of clausal agreement. A potential impoverishment of the verbal paradigm seems not to have played a significant role in this early rise of overt subjects: less than 12% of the clauses with a plural subject in the 19th century data did not show overt agreement between the verb and the plural subject, and all these cases occur strictly with unaccusative or existential verbs. Following a Minimalist approach to null subject languages (HOLMBERG, 2005, 2010a), we propose that a definiteness feature in Tense was valued by the verb movement to Tense in the 18th century, generating a consistent null subject language. We argue that the loss of the requirement of the definiteness feature in Tense was a initial trigger for the partial loss of null subjects and parallel loss of free inversion from the 18th to the 19th century in Goiás Brazilian Portuguese, giving then rise to a partial null subject language and generating the verb-subject order in finite declarative clauses in that language. On the other hand, we argue that the absence of the definiteness feature in Tense was the initial trigger for the loss of null subjects found in the 19th-century Brazilian Portuguese grammar in Goiás. These results provide new empirical evidence for the emergence of the Brazilian Portuguese grammar as a partial null subject language as early as the second half of the 19th century. Finally, following Lightfoot's (2017b) framework for the analysis of syntactic change, which predicts discontinuities between the syntactic structures used by one generation and the next, and Thomason's (2001) framework for establishing contact-induced language change, we suggest that the Brazilian Portuguese grammar in Goiás in the 19th century could be the outcome of

language contact between Niger-Congo languages, especially the Kimbundu, and Portuguese in Portuguese America. However, this interpretation is not exhaustive; thus, further research on this topic is needed.

Keywords: Null subjects. Free inversion. Diachronic syntax. Verb movement. Passive *se*. Impersonal *se*.

SUMÁRIO

1 INTRODUÇÃO	15
2 PRESSUPOSTOS TEÓRICOS: VARIAÇÃO PARAMÉTRICA E MUDANÇA SINTÁTICA	31
2.1 Variação paramétrica.....	33
2.2 Mudança sintática.....	40
2.3 Síntese do capítulo	52
3 TÓPICO DE PESQUISA: A SINTAXE DOS SUJEITOS NULOS	54
3.1 O parâmetro do sujeito nulo na teoria de regência e ligação	54
3.2 A sintaxe dos sujeitos nulos no Programa Minimalista	80
3.2.1 Línguas de sujeito nulo expletivo.....	81
3.2.2 Línguas de sujeito nulo orientadas para o discurso.....	84
3.2.3 Línguas de sujeito nulo consistente	95
3.2.4. Línguas de sujeito nulo parcial.....	107
3.3 Síntese do capítulo	118
4 UM PANORAMA DA SINTAXE DO SUJEITO NO PORTUGUÊS BRASILEIRO	121
4.1 Um mosaico de mudanças na sintaxe do sujeito no PB.....	121
4.1.1 Sujeitos nulos versus sujeitos manifestos.....	121
4.1.2 Questões de ordem: inversão livre do sujeito e movimento do verbo	132
4.2 Panorama das construções com <i>se</i> -passivo e <i>se</i> -impessoal no português.....	139
4.3 Síntese do capítulo	154
5 QUESTÕES DE MÉTODO: CORPUS HISTÓRICO E TRATAMENTO ESTATÍSTICO DE DADOS LINGÜÍSTICOS	156
5.1 A criação de um <i>corpus</i> para um estudo histórico do português em Goiás	156
5.1.1 As fontes escritas.....	160
5.1.2 Estabelecendo os critérios de transcrição, codificação e comparabilidade dos dados	168
5.2 O tratamento estatístico de dados linguísticos.....	170
5.3 Síntese do capítulo	172
6 ASPECTOS DA SINTAXE DO SUJEITO NA HISTÓRIA DO PORTUGUÊS EM GOIÁS: SÉCULOS XVIII E XIX.....	173
6.1 Sujeitos manifestos definidos <i>versus</i> sujeitos nulos definidos.....	176
6.2 Inversão livre do sujeito	185
6.3 A morfologia de concordância.....	191
6.4 Construções com <i>se</i>	197
6.5 Construções impessoais.....	202
6.6 Síntese do capítulo	204
7 UMA HIPÓTESE DE LINGÜÍSTICA DE CONTATO PARA A EMERGÊNCIA DA GRAMÁTICA DO PORTUGUÊS BRASILEIRO EM GOIÁS	206
7.1 Algumas questões sobre linguística de contato e a formação do português brasileiro	206
7.2 Algumas considerações sobre a origem sócio-histórica da gramática do português em Goiás	229
7.3 Síntese do capítulo	241
8 CONSIDERAÇÕES FINAIS.....	243
APÊNDICE	264

1 INTRODUÇÃO

Da teoria de regência e ligação ao programa minimalista da gramática gerativa, Chomsky (1982, 1986a, 1995) propôs que o princípio de projeção estendida (do inglês *Extended Projection Principle* – EPP) requer que, em línguas como o inglês e o francês, um elemento sintático ocupe a posição de especificador de um núcleo funcional flexional (Infl) ou temporal (T). Com base nesse postulado, os estudos gerativistas sobre a sintaxe dos sujeitos nulos buscam descrever e explicar a distribuição desses sujeitos em termos de licenciamento e de identificação, lidando, especialmente, com as restrições de licenciamento de sujeitos nulos e com a identificação desses sujeitos por meio da valoração dos traços- ϕ , ou seja, dos traços formais de pessoa, número e gênero. Os sujeitos nulos caracterizam-se por serem fonologicamente nulos e tacitamente implícitos, sendo compreendidos a partir, principalmente, da morfologia flexional dos verbos finitos ou do contexto discursivo. Nesse sentido, os estudos sobre a sintaxe dos sujeitos têm tido um papel de destaque na linguística gerativa, primeiramente porque eles revelam muito sobre a estrutura subjacente das línguas na mente humana e, segundo, porque dizem muito sobre as interfaces entre sintaxe e morfologia e entre sintaxe e semântica.

Berwick & Chomsky (2016) relatam que, desde suas origens em meados do século XX, a linguística gerativa tem concentrado seus esforços para construir relatos científicos sobre as línguas naturais como um aparato biológico, isto é, como uma propriedade da mente humana. Os autores afirmam que esses esforços capturaram aquilo que é considerado a propriedade básica da linguagem humana: cada língua produz um conjunto infinito de expressões estruturadas hierarquicamente com interpretações sistemáticas em interfaces com dois outros sistemas internos, o sistema articulatorio-perceptual para externalização sonora ou gesto-visual e o sistema conceitual-intencional para inferir, interpretar, planejar, organizar os elementos constitutivos do que é conhecido como pensamento (BERWICK & CHOMSKY, 2016).

Uma língua entendida nesses termos é chamada de língua interna ou língua-I. A teoria de uma língua-I é o estudo de sua chamada gramática gerativa e a teoria geral das línguas-I constitui o estudo da Gramática Universal (GU) (BERWICK & CHOMSKY, 2016). A GU, particularmente, é a teoria do componente genético da faculdade da linguagem, a capacidade que possibilita a aquisição e o uso de língua(s). Um dos objetivos dos estudos da GU é determinar aquilo que é fixo na propriedade básica da linguagem humana e definir as opções de variação. A GU, portanto, determina a classe de procedimentos gerativos que satisfaz a

propriedade básica da linguagem humana e os elementos atômicos que entram na computação da linguagem. Centrados nas línguas-I e na GU, os estudos gerativistas satisfazem, respectivamente, as condições de adequação descritiva e de adequação explicativa da pesquisa em linguística.

Nessa perspectiva biolinguística da linguagem humana, foram estabelecidos três fatores que interagem e determinam o *design* das línguas-I (CHOMSKY, 2005a):

- (1) **1º Fator:** a dotação genética, aparentemente uniforme na espécie humana, o tópico da Gramática Universal. A carga genética interpreta parte do ambiente como experiência linguística e determina o percurso do desenvolvimento da faculdade de linguagem em direção às línguas-I. A interpretação do ambiente como experiência linguística constitui tarefa singular que a criança desempenha de modo reflexivo;
- 2º Fator:** a experiência, que leva à variação, i.e., os dados linguísticos primários (do inglês, *Primary Linguistic Data*). Os dados linguísticos primários permitem a aquisição de língua e conduzem a variação dentro de um escopo bastante estreito;
- 3º Fator:** princípios de arquitetura estrutural e de restrições sem domínio específico, isto é, princípios independentes da linguagem ou até mesmo do organismo. Dissociada de outras habilidades e capacidades cognitivas, a aquisição de uma língua-I baseia-se em restrições da GU e em princípios independentes da linguagem.

Chomsky (2005a) divide o terceiro fator em dois subtipos: (a) princípios de análise de dados que podem ser utilizados na aquisição de língua e (b) princípios da arquitetura estrutural e de restrições de desenvolvimento que operam sobre princípios de eficiência computacional. Nessa direção, o modelo de princípios e parâmetros (doravante P&P), desenvolvido no âmbito da teoria de regência e ligação, consistiu num empreendimento sério de investigação do primeiro fator (CHOMSKY, 2005a). O terceiro fator, por sua vez, assumiu maior destaque em trabalhos recentes dentro do programa minimalista (doravante PM). A tentativa de explicar as propriedades da linguagem humana em termos de considerações gerais de eficiência computacional no PM eliminou parte da tecnologia postulada como específica da linguagem humana em versões anteriores da teoria e passou a fornecer explicações mais fundamentadas dos fenômenos linguísticos, como veremos a seguir (cf. HOLMBERG, 2017; BIBERAUER & ROBERTS, 2017).¹

¹ No Capítulo 2, ao tratarmos de variação paramétrica e mudança sintática, apresentaremos alguns dos principais aspectos do modelo de P&P.

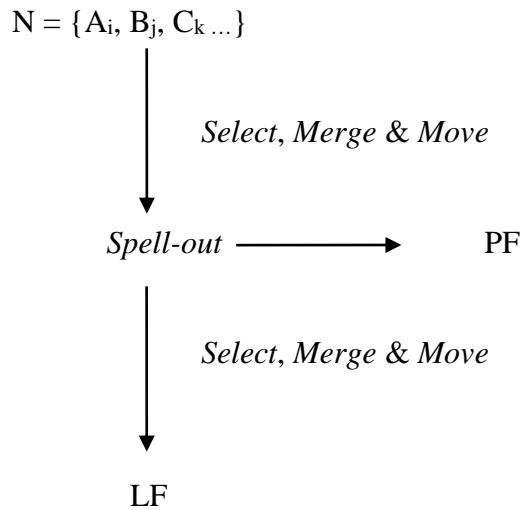
Com o advento do PM, em *The Minimalist Program*, Chomsky (1995) advoga uma teoria derivacional da sintaxe em que a derivação favorece, preferencialmente, relações locais e estruturas simples. Dessa forma, o sistema computacional da linguagem, com sua capacidade infinita de gerar sentenças por meio de recursos finitos, é visto como um sistema perfeito que busca a melhor forma de solucionar a questão das interfaces fonológica e semântica (π, λ) na derivação sintática. O ponto de partida nessa abordagem é que a sintaxe é um sistema combinatório, onde a operação fundamental é *Merge*, responsável por gerar recursivamente um conjunto infinito de expressões que são construídas hierarquicamente. Desse modo, a hierarquia, ou as relações de dependência relativas a c-comando, em princípio, resultam de *Merge* (cf. EPSTEIN *et al.*, 1998).

Chomsky (1995) considera a gramática de uma língua como um procedimento que gera representações em pares (π, λ), interpretados como instruções para os sistemas de desempenho. Assim, o sistema computacional da linguagem interage com os sistemas de desempenho em dois níveis de interface: forma fonética (PF, do inglês *Phonetic Form*) e forma lógica (LF, do inglês *Logical Form*). Os níveis de interface impõem restrições ao sistema computacional, pois cada um interpreta apenas os traços referentes a seu sistema. PF é responsável por transformar as representações sintáticas em instruções interpretáveis na interface articulatorio-perceptual, enquanto LF é responsável por transformar as representações sintáticas em instruções interpretáveis na interface conceitual-intencional. Quando ambos os sistemas licenciam as representações geradas pelo sistema computacional, são atendidas as condições para o princípio de interpretação plena da derivação. O momento em que os traços interpretáveis são enviados para a PF é chamado de *Spell-out*. Os traços remanescentes continuam a derivação até a LF.

Ao iniciar uma derivação, o sistema computacional da linguagem opera a partir da disposição em índice de itens lexicais possuidores de traços fonológicos, semânticos e formais numa numeração (N). Dado um arranjo lexical em N, o sistema computacional efetua operações até gerar uma derivação que converge. Primeiramente, um procedimento denominado *Select* seleciona os itens lexicais disponibilizados em N até reduzir seus índices a zero no espaço derivacional. Posteriormente, *Merge* atua recursivamente na concatenação do léxico selecionado de N, construindo objetos sintáticos. *Merge* consiste em combinar duas categorias (α e β) para gerar uma nova categoria (γ): $\{\gamma, \{\alpha, \beta\}\}$, em que γ identifica a categoria e as propriedades relevantes do objeto formado. Para esse objeto sintático recém-formado ser interpretável nas interfaces de som e de significado, ele precisa de um algoritmo de rotulagem (*Labeling Algorithm*), operação que atribui um rótulo categorial aos sintagmas

derivados por *Merge* (cf. CHOMSKY, 2013, 2015). Esse algoritmo de rotulagem é entendido como um procedimento de busca mínima que detecta o elemento mais acessível em um domínio, por suposição, um elemento mínimo, isto é, um núcleo (CHOMSKY, 2015). Assim que seja formado pelo menos parte do objeto sintático de uma derivação, *Agree*, operação que vincula dois termos – viz. uma sonda (*Probe*) e um alvo (*Goal*) – a distância, executa sua computação valorando os traços- ϕ do especificador em relação ao verbo. Hornstein, Nunes & Grohmann (2005) propuseram um arquétipo, que reproduzimos em (2), para resumir a proposta de arquitetura da gramática do PM.

(2) Arquitetura da gramática no programa minimalista (HORNSTEIN *et al.*, 2005, p. 73)



Holmberg (2017) pontua que os dispositivos teóricos necessários no PM incluem muitos daqueles que eram importantes em modelos anteriores da teoria, como, por exemplo: o conjunto de categorias funcionais que constroem projeções verbais e nominais, movimento, incluindo a distinção entre movimentos A e A-barra (em que movimento-A é motivado basicamente por traços- ϕ e Caso estrutural e movimento A-barra, por um traço periférico), papéis- θ , Caso e concordância (*Agree*). O autor aponta, no entanto, que existem algumas diferenças notáveis: a teoria X-barra, por exemplo, não é um postulado para a estrutura de sintagmas no PM, tendo em vista que suas propriedades agora são derivadas por *Merge* e *Labeling* (cf. CHOMSKY, 2013, 2015).²

Holmberg (2017) pontua também que alguns dispositivos teóricos do PM, como *Merge* e *Agree*, foram reformulados. Boeckx (2011), por exemplo, propôs que *Merge* é a

² Remetemos o leitor à obra de Mito, Figueiredo Silva & Lopes (2013) para uma compreensão detalhada de alguns dos dispositivos da teoria de regência e ligação, como a teoria X-barra, por exemplo.

única operação necessária na sintaxe. A proposta de Boeckx (2011) foi impulsionada por uma redefinição do movimento sintático feita por Chomsky (2004), que passa a ser definido como o caso em que um constituinte já concatenado como parte de uma árvore sintática é concatenado novamente. Essa operação foi chamada *internal Merge* e se associa à teoria de movimento por cópia, em oposição a *external Merge*, que é a operação em que uma palavra ou sintagma é concatenado pela primeira vez. Dessa forma, postula-se que quando α é uma parte de β e concatena novamente na raiz de β , ele deixa uma cópia idêntica dentro de β .

Em relação a *Agree*, os estudos mais recentes têm lidado com essa operação nos moldes de Chomsky (2000, 2001). Consoante com esses trabalhos, tem-se que os traços do léxico são de duas naturezas: interpretáveis, que entram na derivação valorados, e não interpretáveis, que entram na derivação não valorados. Os traços interpretáveis são legíveis e interpretáveis na interface conceitual-intencional. Os traços não interpretáveis não são legíveis. Um exemplo de traço não interpretável é o traço de número em um verbo concordando com um DP, que, por sua vez, possui traço interpretável de número.³ Apesar de o traço de número em um verbo ser um marcador evidente da relação existente entre o verbo e o sujeito, ele é não interpretável, pois é redundante e não afeta a semântica do verbo. Dado que uma derivação não pode conter traços não interpretáveis quando as informações são enviadas para a interface conceitual-intensional, *Agree* opera como um mecanismo de valoração dos traços formais de um núcleo funcional (*sonda*) por meio da correspondência desses traços com os traços formais de um elemento nominal (*alvo*). Quando *Agree* apaga os traços não interpretáveis da derivação, *Spell-out* é imediatamente aplicada.⁴

Considerando a localidade e a economia operacional dos constituintes de uma estrutura oracional, Chomsky (2001) propôs um novo modelo de análise linguística denominado de derivação em fases, em que fase é o nome atribuído à limitada quantidade de estrutura oracional com a qual o sistema computacional da linguagem trabalha no percurso de uma derivação. Ao propor esse modelo, Chomsky (2001) reitera que a faculdade da linguagem trabalha com uma pequena quantidade de estrutura em sua memória ativa, o que impõe restrições aos constituintes de uma oração de estabelecerem relações que não sejam locais, simples e econômicas. Na derivação em fases, as expressões sintáticas são construídas num procedimento passo a passo, que, de acordo com Chomsky (2001), demonstra como a sintaxe trabalha como parte de um *design* estrutural da mente humana. Para Chomsky (2001),

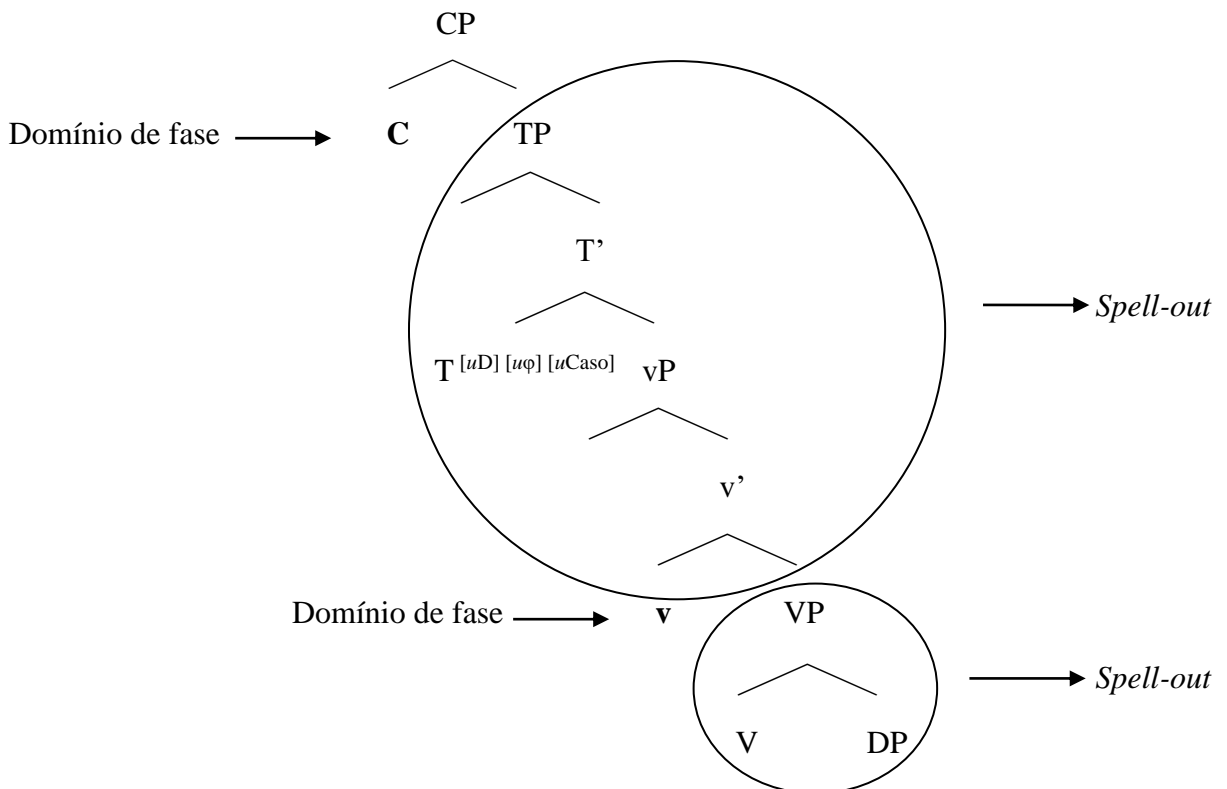
³ Estamos adotando aqui a proposta de Magalhães (2004) de que o DP carrega traços interpretáveis de número, contrária à proposta de Chomsky (2000) de que apenas os nomes possuem traços interpretáveis de número.

⁴ As categorias funcionais C, T e v podem carregar traços não interpretáveis, sendo que as duas últimas obrigatoriamente os carregam (CHOMSKY, 2000).

as informações linguísticas são fornecidas aos sistemas de interface dinamicamente durante a derivação.

Nesse modelo de derivação, por exemplo, após serem formados por *Merge* e terem seus traços não interpretáveis valorados por *Agree*, os objetos sintáticos são enviados aos sistemas de interfaces – esses objetos enviados ciclicamente para *Spell-out*, aliás, são o que Chomsky (2001) denomina de fases. Desse modo, quando todos os itens lexicais de um subarranjo lexical são reduzidos a zero numa estrutura derivada em fases, uma fase está pronta para ser enviada para *Spell-out*. Após ser enviada para *Spell-out*, a fase não retornará para a sintaxe estrita, o que é determinado pela chamada condição de impenetrabilidade da fase, isto é: dada uma configuração $HP = [\alpha [H \beta]]$, em uma fase HP com o núcleo H, o domínio de H (β) não está acessível para operações fora de α ; somente H e sua periferia (α) são acessíveis para tais operações (CHOMSKY, 2001). Não há derivação gramatical (convergente, em termos minimalistas) que contenha traço não interpretável quando as informações são enviadas para as interfaces. Somente após todos os traços não interpretáveis terem sido devidamente valorados, CP, que é uma fase, envia TP, domínio de CP, para *Spell-out*. Em (3), ilustramos uma estrutura oracional derivada em fases.

(3) Configuração de uma estrutura oracional derivada em fases



Ponderando que as operações sintáticas devem envolver uma relação local entre uma sonda (*Probe*) e um alvo (*Goal*), Chomsky (2001) argumenta que a derivação em fases estreita essa relação de localidade, pois reduz a complexidade operativa do sistema computacional, fazendo com que cada estágio da derivação acesse apenas uma parte do arranjo lexical, e considera que todo subarranjo lexical deve conter um item lexical que será rotulado como o núcleo da fase. Nesse sentido, uma característica peculiar do modelo de derivação em fases é a distinção entre as fases de acordo com o valor dos elementos que as compõem: por exemplo, ν P, com estrutura argumental completa, e CP, com um complexo sentencial completo, são considerados fases fortes porque apresentam complexos funcionais completos. Por sua vez, constituintes como o ν P de verbos passivos e inacusativos e as sentenças que não têm CP em sua estrutura, ou seja, que têm um T defeutivo, sem traços de pessoa, são considerados fases fracas. A diferença entre ambos os tipos de fases é que a condição de impenetrabilidade não é aplicada na fase fraca (CHOMSKY, 2001).

Isso posto, esta tese de doutorado – vinculada ao grupo de pesquisa *O Centro-Oeste na história do português brasileiro* – toma como referência os estudos gerativistas sobre sujeitos nulos, com o objetivo primordial de investigar aspectos sintáticos e semânticos desses sujeitos em manuscritos em língua portuguesa produzidos em Goiás nos séculos XVIII e XIX.⁵ A linha temporal proposta para o estudo das fontes escritas foi delimitada, respectivamente, pelo século em que oficialmente houve a integração de Goiás ao empreendimento colonial português na América e pelo século em que o contato mais generalizado da língua portuguesa com as línguas indígenas autóctones e as línguas africanas trasladadas via Atlântico já havia se dissipado nos centros urbanos daquela região e uma gramática particular da língua portuguesa já teria condições de ter se desenvolvido ali. A hipótese primordial desta tese é a de que fontes escritas no século XIX em Goiás constituem a evidência empírica mais antiga – encontrada até a realização dessa tese – do desenvolvimento do português brasileiro como uma língua de sujeito nulo parcial, considerando a proposta tipológica de Holmberg (2010a) de que as línguas de sujeito nulo parcial não possuem um traço-D (do inglês *Definiteness* – definitude) como parte dos traços de T e, por isso, essas

⁵ O grupo de pesquisa *O Centro-Oeste na história do português brasileiro* é coordenado pela Profa. Dra. Eloísa Pilati (coordenadora) e pela Profa. Dra. Heloísa Salles (vice-coordenadora), do Programa de Pós-graduação em Linguística da Universidade de Brasília (PPGL/UnB), e cadastrado no diretório de grupos de pesquisa do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Ressalta-se que, inicialmente, esta tese esteve vinculada ao projeto de pesquisa *Estudos sobre a constituição do português brasileiro* (2012-2018), sob a coordenação da Profa. Dra. Rozana Reigota Naves (PPGL/UnB), que também é membro da equipe do projeto cadastrado no CNPq.

línguas apresentam certas restrições no licenciamento e na distribuição de sujeitos nulos, em contraste com as línguas de sujeito nulo consistente.

Holmberg (2010a) identifica, entre as línguas de sujeito nulo (doravante LSN), aquelas que permitem sujeitos nulos definidos por meio de um traço de definitude – i.e., o traço-D – associado a uma morfologia verbal rica que distingue os traços semânticos de pessoa, número e gênero por meio da flexão dos verbos finitos, como o português europeu e o italiano, daquelas que não possuem um traço de definitude capaz de distinguir uniformemente todas as pessoas gramaticais por meio da morfologia dos verbos finitos, especialmente a terceira pessoa, como no caso do português brasileiro e do finlandês. Nessa direção, Holmberg (2010a) propõe que, tipologicamente, o português europeu (doravante PE) e o italiano são LSN consistentes e o português brasileiro (doravante PB) e o finlandês são LSN parciais. Em sua proposta, Holmberg (2010a) ressalta que há contextos em que sujeitos nulos definidos são opcionais nas LSN parciais, mas obrigatórios nas LSN consistentes, e contextos em que são opcionais nestas, mas obrigatórios naquelas, como podemos verificar pelo exemplo em (4), onde o DP *João* está falando dele mesmo.

- (4) ‘O João₁ disse que (ele₁) queria comprar um carro.’ (HOLMBERG, 2010a, p. 91)
- a. John₁ sa att *(han₁) ville köpa en bil. (Sueco)
 João disse que ele queria comprar um carro
- b. Gianni₁ dice che (*lui₁) vuole comprare una macchina. (Italiano)
 João disse que ele queria comprar um carro
- c. Ram₁ mhanala ki (tyani₁) ghar ghetla. (Marata)
 João disse que ele house comprar
 ‘João disse que (ele) queria comprar uma casa’.

Comparando os dados em (4a) e (4b), Holmberg (2010a) afirma que línguas de sujeito não nulo, como o sueco em (4a), o inglês e o francês, precisam manifestar o sujeito da oração encaixada c-comandado pelo sujeito da oração matriz, ao passo em que LSN consistentes, como o italiano em (4b), o espanhol, o árabe, o grego e o turco, devem deletar o pronome sujeito da oração encaixada cujo referente é o sujeito da matriz. Por fim, Holmberg (2010a) mostra que LSN parciais, como o marata em (4c), o PB e o finlandês, podem opcionalmente manifestar o pronome sujeito de uma oração encaixada cujo referente seja o sujeito da oração matriz. Em resumo, Holmberg (2010a) está a propor que pronomes definidos nulos na 3ª pessoa do singular nas LSN parciais são opcionais em alguns contextos e não podem ser

licenciados a menos que sejam localmente c-comandados por um antecedente. A situação é curiosamente invertida quando se considera o licenciamento de sujeitos nulos indefinidos em sentenças ativas com verbos na 3ª pessoa do singular: eles são comuns nas LSN parciais, mas não nas LSN consistentes (cf. HOLMBERG, 2010a). Holmberg (2010a) ilustra o licenciamento de sujeitos nulos indefinidos na 3ª pessoa do singular contrastando o PB (5b) com o PE (5a):

- (5) a. É assim que *se* faz o doce. (PE/PB) (HOLMBERG, 2010a, p. 92)
 b. É assim que ___ faz o doce. (*PE/PB)

Como mencionado, Holmberg (2010a) postula a ausência de um traço-D em T nas LSN parciais. Assim, para Holmberg (2010a), quando um sujeito nulo singular de 3ª pessoa entra em uma relação de *Agree* com T em uma LSN parcial, esse sujeito só pode ser interpretado indeterminadamente, como em (5b) – ao contrário do que ocorre nas LSN consistentes, que permitem apenas o licenciamento de sujeitos indefinidos marcados abertamente, como em (5a), em que o pronome clítico de 3ª pessoa *se* é o sujeito da sentença.⁶ Holmberg (2010a) distingue dois tipos de leitura genérica para os sujeitos nulos indefinidos nas LSN parciais: (i) genérica inclusiva, na qual o sujeito nulo denota pessoas em geral e inclui o falante e o ouvinte; e (ii) genérica exclusiva, na qual o sujeito nulo denota pessoas em geral, mas exclui o falante e o ouvinte. No Capítulo 4, trataremos mais detidamente da leitura genérica dos sujeitos nulos indefinidos com dados do PB e do finlandês.

O que Holmberg (2010a) propõe, resumidamente, é: (i) a restrição de sujeitos definidos nulos na 3ª pessoa no singular aos contextos em que são localmente c-comandados por um antecedente e o licenciamento de sujeitos nulos indefinidos na 3ª pessoa do singular caracterizam LSN parciais; (ii) a identificação de sujeitos nulos definidos na 3ª pessoa ('ela'/'ele' nulos) por meio de um traço de definitude e de uma morfologia verbal rica e a ausência de sujeitos indefinidos nulos na 3ª pessoa do singular caracterizam LSN consistentes.

Destaca-se que foi Galves (1985) quem primeiro constatou a substituição do clítico *se*-impessoal por um sujeito nulo indefinido de leitura genérica no PB, como ilustrado em (6).

⁶ Nesta tese, chamamos de clítico todo pronome que é fonologicamente dependente (porque é uma forma átona) e, por isso, é adjungido a um hospedeiro, que necessariamente é um verbo, sem, entretanto, apresentar uma posição fixa em relação ao seu hospedeiro (cf. KAYNE, 1975).

(6) Nos nossos dias não __ usa mais saia. (PB) (GALVES, 1985, p. 35)

Foi a partir da observação empírica de dados como em (6), dentre outros que evidenciavam uma diversificação no licenciamento de sujeitos nulos no PB, que pesquisadores gerativistas despertaram sua atenção, principalmente nas décadas de 80 e 90 do século passado, para uma possível mudança na marcação do parâmetro do sujeito nulo no PB (cf. ROBERTS & KATO, 1993 [2018]; KATO & NEGRÃO, 2000; dentre outras obras de referência). Dentre os estudos desenvolvidos nesse período, Galves (1984, 1998) propôs, nos termos de Huang (1984), que o PB era uma LSN orientada para o discurso. A autora defende sua proposta analisando as distinções de licenciamento de categorias vazias e pronomes lexicais em posição de sujeito em orações simples e relativas no PB e no PE (GALVES, 1984) e as chamadas construções com tópicos-sujeitos, inovadoras na gramática do PB (GALVES, 1998). Cotejando dados diacrônicos e sincrônicos, M. E. Duarte (1993 [2018], 1995, 2000) atribuiu a fatores de natureza morfossintática, como a reestruturação dos paradigmas pronominal e verbal, sua análise de que o PB estaria em um processo de mudança que o levaria de uma marcação positiva [+] para uma marcação negativa [-] em relação ao parâmetro *pro-drop*. Figueiredo Silva (1996), por sua vez, ao analisar a distribuição das categorias vazias e dos pronomes lexicais na posição de sujeito em sentenças com verbos finitos no PB, sugere que o PB “é uma língua de sujeito nulo parcial, utilizando-se de estratégias especiais para identificar os sujeitos nulos quando essa posição deve ser interpretada referencialmente” (FIGUEIREIDO SILVA, 1996, p. 120). A autora argumenta que o PB não mais utiliza a morfologia verbal para identificar sujeitos nulos, como ocorre nas LSN como o italiano, mas ainda permite que a posição de sujeito não seja preenchida lexicalmente em determinados contextos, como ilustraremos no Capítulo 4.

Negrão (1999) e Negrão & Viotti (2000), nos termos de Huang (1984), defendem que o PB seria uma LSN orientada para o discurso, sem, contudo, associar essa mudança ao enfraquecimento da concordância verbal. As autoras assumem que o PB seria uma LSN orientada para o discurso, tendo em vista a função informacional dos constituintes da sentença no PB, isto é, as funções de tópico do discurso ou foco, e também o escopo de sintagmas quantificados. As autoras afirmam que essas propriedades discursivas minaram a função que era desempenhada pelo paradigma flexional dos verbos finitos. Desse modo, para as autoras, o empobrecimento da concordância não estaria diretamente relacionado com o preenchimento mais sistemático da posição de sujeito no PB. Estudos mais recentes, como os de C. Rodrigues (2004), Sheehan (2006), Holmberg, Nayudu & Sheehan (2009) e Holmberg (2005,

2010a), têm classificado o PB como uma LSN parcial devido a essa língua licenciar sujeitos nulos definidos (isto é, referenciais) em contextos bastante restritos, como veremos mais detalhadamente nos capítulos 3 e 4 deste trabalho.

Com base na observação empírica de que sujeitos nulos indefinidos na 3ª pessoa do singular eram licenciados em diários goianos oitocentistas, como exemplificado em (7), propusemos que na segunda metade do século XIX em Goiás o PB já teria se desenvolvido como uma LSN parcial, isto é, desde o século XIX, numas das regiões mais isoladas da colônia portuguesa na América, o PB apresentava propriedades sintático-semânticas de uma LSN parcial (cf. H. BORGES, 2014; NAVES & BORGES, 2015; BORGES & PIRES, 2017).

- (7) Português goiano oitocentista (NAVES & BORGES, 2015, p. 83)
- a. Dia 13 eu e Lili [fomos] à missa às 7 horas no Rosário de Santa Luzia, neste ano Nhola deu a Dona Deolinda 1 anel e coco e doce de coco. Dia 14 __ **principiou** fazer estuque aqui na sala.
- b. Dia 17 Mariquinha Alexandra esteve aqui de noite. Dia 18 domingo Mariquinha fez anos. Depois eu e Lili fomos em casa de Maria Póvoa e Iayá Albano. De tarde Manoel Xavier e Nhanhá veio cá. Dia 19 __ **pintou** o estuque da sala.
- c. Dia 21 Lili e Maria e Benedita foram em casa de João Maria Berquó. De tarde Silvina veio aqui. Dia 22 eu Nhola e Mariquinha fomos em casa de Silvestre Santana. Dia 23 __ **esteve** aqui **trabalhando fazendo** soalho.

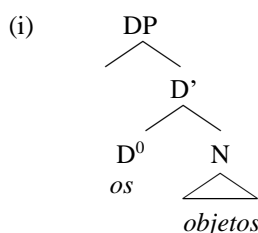
Em (7), os verbos negritados pertencem ao conjunto de verbos das construções com *se*-impessoal do português: o primeiro é inacusativo; o segundo é transitivo com um argumento interno no singular; os últimos verbos, formando uma perífrase progressiva, são inergativo (sem argumento interno) e transitivo (com argumento interno no singular), respectivamente. Chama a atenção, contudo, o fato de que nenhum NP, DP ou pronome (in)definido manifesto pode ser tomado como o referente da categoria vazia que ocupa a posição de sujeito dos verbos em destaque. Essa foi a evidência crucial para postularmos que o PB em Goiás no século XIX era uma LSN parcial, isto é, uma língua sem um traço-D em T. Como vimos nas páginas precedentes, quando um sujeito nulo singular de 3ª pessoa entra em uma relação de *Agree* com T e T não possui um traço-D em T, esse sujeito deve ser

interpretado indeterminadamente, tal como ocorre com os sujeitos dos verbos destacados em (7).⁷

Ademais, contrários à principal premissa da maioria dos estudos sobre a perda de sujeitos nulos na gramática do PB, argumentamos que o empobrecimento da concordância verbal pode não ter desempenhado um papel significativo nos primeiros registros que evidenciam restrições ao licenciamento de sujeitos nulos na gramática do PB em Goiás (cf. BORGES & PIRES, 2017). Uma vez que nossa explicação para a ordem verbo-sujeito (VS) no português setecentista em Goiás, como em (8), fundamenta-se no argumento de que o movimento de V-para-T era suficiente para a valoração dos traços- ϕ de T e a satisfação do traço-EPP (cf. ALEXIADOU & ANAGNOSTOPOULOU, 1998), que, consoante Holmberg (2010a), consiste basicamente na valoração de um traço [uD] em T, propusemos que a ordem sujeito-verbo (SV), como em (11), manifesta-se mais robustamente no PB oitocentista em Goiás devido à perda do traço-D de T na gramática dessa variedade, requerendo a projeção de um elemento sintático de natureza nominal ou pronominal para a posição de especificador de TP [SPEC-TP], a fim de satisfazer o traço-EPP. Estendendo essa análise para o contexto da sintaxe dos sujeitos na gramática oitocentista do PB em Goiás, argumentamos que a perda do traço-D em T impôs restrições à recuperabilidade da referencialidade dos sujeitos nulos no PB oitocentista em Goiás, requerendo também que um elemento sintático de natureza nominal ou pronominal fosse projetado para [SPEC-TP] para satisfazer o traço-EPP.⁸

⁷ Em comunicação pessoal, fomos indagados por Ezekiel Panitz (2015) se os exemplos em (7) não teriam alguma similaridade com o fenômeno *diary pro* do inglês, que consiste na omissão de sujeito referencial de 1ª pessoa apesar de o inglês ser uma língua de sujeito não nulo (cf. HAEGEMAN, 1990). Argumentamos que não porque, além de ocorrer com a 3ª pessoa do singular, (7) não está restrita à escrita nem muito menos a textos do gênero diário, como parece ser o caso do *diary pro*. Por fim, destacamos que nosso estudo consiste na análise de uma gama de aspectos sintáticos do sistema gramatical que operava na língua daquele período, como exemplificamos por meio do contraste em (8) e (9) a seguir com dados sobre a ordem das palavras. As construções com sujeitos nulos indeterminados como em (7) são, portanto, uma das evidências do funcionamento inovador desse sistema.

⁸ Na gramática gerativa, assume-se que sintagmas nominais são nucleados por uma categoria funcional D, conforme ilustrado abaixo.



Verifica-se que os argumentos externos de (8) e (9) são, respectivamente, um DP completo e nominais nus, sendo que apenas esses últimos ocupam a posição de [SPEC-TP].

(8) Ordem VS no português goiano setecentista

Marcha **Sua Excelência** do Julgado de Traíras para o de São Félix. Saiu **Sua Excelência** do Arraial de Traíras no dia trinta de junho pelas sete horas da manhã acompanhado de todos os moradores daquele Arraial, e na retaguarda a Companhia da Cavalaria Auxiliar do Arraial de São José, onde tomou **Sua Excelência** quartel tendo marchado duas léguas. Ali foi **Sua Excelência** recebido do vigário com muita decência que ainda que não há Julgado neste Arraial sempre é freguesia e foi fundado no mesmo tempo que o de Traíras e com as mesmas circunstâncias. Neste Arraial reformou **Sua Excelência** a Companhia da Cavalaria Auxiliar, provendo nela os postos de capitão, tenente, e alferes, que estavam vagos.

(9) Ordem SV no português goiano oitocentista

Dia 25 **Joaquim Fernandez** veio aqui. Dia 26 **Lulu Pitaluga** veio aqui. Dia 27 **Nhola e Lili** foram em casa de Mariana P. Nessa noite **Joaquim Fernandez e Cavalcante** vieram aqui. Dia 28 foi domingo. De tarde **eu e Nhola e Lydia** fomos em casa de Silvina. Dia 29 fiz anos. De noite **Lulu Pitaluga e Tenente Cavalcante** vieram aqui. **Cavalcante** ofereceu-me uns versos muito bonitos. Dia 30 **Joaquim Fernandez** esteve aqui de noite, deu-me um paletó [de] lã. Dia 31 de tarde **Nhola e Lili** foram em casa do senhor Eduardo de Abreu.

Considerando os dados exemplificados em (7), (8) e (9) e as premissas teóricas que embasam esta tese, nossa hipótese de pesquisa pode ser resumida nas seguintes proposições:

- i. manuscritos em língua portuguesa produzidos em Goiás na segunda metade do século XVIII apresentam propriedades de uma LSN consistente, como, por exemplo: alta frequência de sujeitos nulos referenciais em sentenças declarativas; T suficientemente forte para ficar sem um especificador nos casos em que um verbo finito se move para T; predomínio da ordem VS; maior frequência de construções com *se* (passivo e impessoal).
- ii. manuscritos em língua portuguesa produzidos em Goiás na segunda metade do século XIX apresentam propriedades de uma LSN parcial, como, por exemplo: queda no licenciamento de sujeitos nulos referenciais em orações declarativas; sujeitos nulos definidos de 3ª pessoa licenciados localmente por meio de c-comando por um

antecedente; predomínio da ordem SV; declínio das construções com *se* (passivo e impessoal); licenciamento de sujeitos nulos indeterminados na 3ª pessoa do singular.

A partir dessas proposições, elaboramos a seguinte questão de pesquisa para esta tese:

Questão 1: quais mecanismos sintáticos e semânticos envolvidos na sintaxe dos sujeitos nulos nas gramáticas setecentista (segunda metade do século XVIII) e oitocentista (segunda metade do século XIX) do português em Goiás explicariam a mudança paramétrica em tela?

Para responder a essa questão, criamos um *corpus* linguístico composto por amostras de dados de diários produzidos em Goiás nos séculos XVIII e XIX. As amostras dos manuscritos foram transcritas num processador de textos e, posteriormente, num editor de planilhas, a fim de terem suas propriedades em relação à sintaxe do sujeito codificadas. Foram transcritas 1250 sentenças referentes a cada século, totalizando 2500 sentenças transcritas e codificadas. Em suma, o *corpus* é composto pelo total de 2500 sentenças: 1250 referentes ao século XVIII e 1250 referentes ao século XIX (cf. Apêndice para detalhes sobre a codificação dos dados).

Além dessa questão de natureza gramatical, acrescentamos uma questão de pesquisa a esta tese com base na defesa que elaboramos anteriormente sobre o papel de fatores sociolinguísticos e psicolinguísticos inerentes ao contato linguístico no desenvolvimento do PB como uma LSN parcial (cf. H. BORGES, 2014; NAVES & BORGES, 2015). Nesses trabalhos, sugerimos que: (i) a aquisição do português como segunda língua por ameríndios e, sobretudo, africanos e (ii) a consequente transmissão dessa gramática adquirida como L2 a seus descendentes em processo de aquisição de L1 teriam sido o gatilho para a emergência da gramática do PB como uma LSN parcial no século XIX, especialmente em Goiás.⁹ Neste trabalho, retomamos essa questão do papel do contato linguístico na constituição da gramática do PB em Goiás, porém, a reestruturamos como segue:

⁹ A respeito dessa proposta, Sarah Thomason, em comunicação pessoal (2016), afirmou-nos que é teoricamente errado pensar que a aquisição como L1 de uma língua L2 geraria uma reestruturação da L2 quando adquirida como L1. Em outros termos, a pesquisadora chamou-nos a atenção para duas questões: primeiramente, para o fato de que o *input* ser uma L2 não significa que essa L2 não fosse uma língua completa em sua estrutura, afinal, não seria possível transmitir uma língua incompleta(mente). Em segundo, para o fato de que, se houve mudança nesse processo, ela teria ocorrido na L2 dos africanos e/ou ameríndios, não na L1 de seus descendentes. Vamos tratar dessas questões no Capítulo 7 da tese.

Questão 2: por intermédio dos dados linguísticos disponíveis nesta tese, é possível propor um modelo teórico de análise linguística que possa indicar algum papel do contato linguístico no desenvolvimento do PB em Goiás como uma LSN parcial?

Destacamos que, para investigar categoricamente uma mudança paramétrica numa determinada língua (ou numa determinada variedade), é preciso analisar um conjunto de traços e operações sintático-semânticos que devem ter atuado na nova configuração da gramática dessa língua (ou variedade). Destarte, seguimos essa diretriz para propor uma análise linguística sobre o papel do contato linguístico na constituição da gramática do PB como uma LSN parcial, bem como para contribuir com uma melhor compreensão da atuação das propriedades e dos mecanismos da faculdade da linguagem nos processos de mudança sintática, desencadeados ou não por contato linguístico.

Para dar conta dos objetivos e responder às questões de pesquisa desta tese, foram elaborados oito capítulos, incluindo esta Introdução e as Considerações Finais. O Capítulo 2 (*Pressupostos teóricos: variação paramétrica e mudança sintática*) é dedicado a uma sucinta revisão a respeito dos conceitos de variação paramétrica e mudança sintática no âmbito da linguística gerativa. O Capítulo 3 (*O tópico de pesquisa: a sintaxe dos sujeitos nulos*) apresenta um detalhado estado da arte dos estudos sobre a sintaxe dos sujeitos nulos nas línguas naturais e suas propriedades sintático-semânticas dentro do arcabouço gerativista. O Capítulo 4 (*Panorama da sintaxe do sujeito no português brasileiro*) apresenta um estado da arte das principais pesquisas sobre a sintaxe dos sujeitos nulos no PB, especialmente na perspectiva da variação paramétrica e da mudança sintática. O Capítulo 5 (*Questões de método: corpus histórico e tratamento estatístico de dados linguísticos*) apresenta os critérios e os procedimentos metodológicos empregados na seleção das amostras do *corpus* linguístico da tese, bem como fornece uma descrição dos princípios de coleta, codificação e comparabilidade dos dados. O Capítulo 6 (*Aspectos da sintaxe do sujeito na história do português em Goiás: séculos XVIII e XIX*) mostra os resultados estatísticos do *corpus* da tese e analisa aspectos específicos da sintaxe do sujeito na história do português em Goiás, a saber: o licenciamento e a distribuição de sujeitos nulos em orações declarativas, a ordem das palavras, o papel da morfologia de flexão no licenciamento de sujeitos nulos e no movimento do verbo para T, e o comportamento sintático das chamadas construções com *se*-passivo e *se*-impessoal e das construções impessoais de um modo geral. Nesse capítulo, delinea-se uma explicação formal para os achados da tese. O Capítulo 7 (*Uma hipótese de linguística de contato para a emergência da gramática do português brasileiro em Goiás*) trata da

correlação entre mecanismos sociolinguísticos, linguísticos e psicolinguísticos envolvidos nos contextos sociais de contato linguístico para avaliar o possível papel do contato na emergência da gramática oitocentista do PB em Goiás, apresentando as limitações dessa conjectura diante dos achados desta pesquisa e comparando-a com as análises presentes em outros estudos. Nas Considerações Finais, retomamos as questões de pesquisas e resumimos a análise e os achados significativos da tese.

2 PRESSUPOSTOS TEÓRICOS: VARIAÇÃO PARAMÉTRICA E MUDANÇA SINTÁTICA

A variação é uma propriedade inerente às línguas humanas. As línguas diferem umas das outras em seus vários aspectos constitutivos: no vocabulário (léxico); na estrutura dos sons (fonologia); na estrutura das palavras (morfologia); na estrutura de frases (sintaxe); e nos aspectos semânticos e pragmáticos. Definir a extensão e os limites da variação é, contudo, um desafio para a linguística teórica. Na gramática gerativa, por exemplo, com o objetivo de melhor compreender e lidar com o conceito de variação, tal como ele é manifesto na estrutura formal das línguas naturais, são estudadas as maneiras pelas quais a diversidade linguística é codificada no sistema computacional cognitivo que gera estruturas de linguagem, ou seja, que gera gramática(s). Nesse sentido, o pesquisador gerativista lida com o ponto em que a diversidade linguística é restringida pela GU subjacente às línguas humanas e busca responder onde na GU estão as diferentes dimensões que permitem a diversidade linguística.

Outra característica inerente às línguas humanas é que elas são dinâmicas, isto é, elas mudam com o decorrer do tempo; e isso ocorre apesar de empreendimentos sócio-políticos sistemáticos para impedir que as línguas mudem. Como resultado, as propriedades linguísticas de uma dada língua são historicamente instáveis. A mudança sintática é, possivelmente, um dos fenômenos linguísticos mais complexos que existe, pois é difícil precisar uma razão aparente que explique a mudança sintática em uma determinada língua, especialmente quando não há influência externa. Em outras palavras, um dos principais problemas ao lidar com a mudança sintática é que é extremamente difícil determinar o que a causou em um dado estágio da língua. Weinreich, Labov e Herzog (1968) traduziram esse impasse, não o restringindo à mudança sintática, como o problema da implementação:

Quais fatores podem explicar a implementação das mudanças? Por que as mudanças em um traço estrutural ocorrem em uma dada língua particular em um dado momento, mas não em outras línguas com o mesmo traço, ou na mesma língua em outros momentos? (WEINREICH, LABOV & HERZOG, 1968, p. 102, tradução nossa).¹⁰

O problema da implementação impõe aos pesquisadores outras questões cujas respostas, visíveis apenas no eixo diacrônico da língua, são difíceis de decifrar: por qual razão

¹⁰ No original: *The Actuation Problem*: “What factors can account for the actuation of changes? Why do changes in a structural feature take place in a particular language at a particular time, but not in other languages with the same feature, or in the same language at other times?” (WEINREICH, LABOV & HERZOG, 1968, p. 102).

determinada mudança é desencadeada em uma língua em particular em um dado momento histórico? Uma vez iniciado, por que o processo de mudança continua a atuar? Fuß (2016) argumenta que uma resposta ao problema da implementação pressupõe uma compreensão profunda da relação entre a experiência linguística apresentada à criança e a gramática construída com base nessa evidência. O autor pontua, contudo, que nossa compreensão desses fenômenos é bastante limitada, apesar dos avanços em estudos experimentais sobre a aquisição de primeira língua na perspectiva da gramática gerativa.

Apesar do desafio imposto pelo problema da implementação, os pesquisadores em linguística teórica podem lançar mão de duas hipóteses razoáveis para lidar com alguns casos de mudança gramatical, quais sejam: primeiramente, pode-se afirmar que a mudança é espontânea, ou seja, interna, endógena, aquela que ocorre quando algo na gramática de uma dada língua muda sem que haja alguma causa externa aparente. De outro modo, pode-se argumentar que a mudança é induzida por contato linguístico, ou seja, é causada pela transferência de um traço linguístico de uma língua para outra numa dada situação de contato linguístico. Nesse caso, entretanto, não é fácil determinar se, de fato, a mudança em uma dada língua A foi provocada por influência de uma dada língua B.¹¹ Apesar de ontologicamente distintos, em ambos os fenômenos – mudança interna ou mudança induzida por contato – o pesquisador precisa examinar dois estágios distintos da língua: o primeiro é anterior à mudança; e o segundo retrata a língua após a concretização da mudança.¹²

Nessa direção, cabe-nos informar que para identificar os mecanismos sintáticos potencialmente provocadores da mudança sintática nos dados do nosso *corpus* de pesquisa, concentraremos nossa análise nas estruturas inovadoras da gramática do português em Goiás no século XIX (estágio 2) em relação à gramática do português produzida naquela região no século XVIII (estágio 1), potencialmente mais próxima da gramática do PE no que diz respeito à sintaxe do sujeito; em outras palavras, no capítulo de análise do *corpus* da tese (Capítulo 6), interessam-nos as estruturas sintáticas na gramática inovadora do português em Goiás na segunda metade do século XIX em comparação com a gramática do português que chegou no território goiano no século XVIII. É na ruptura daquela gramática com esta que esperamos encontrar respostas satisfatórias para pensar os percursos da mudança interna ou induzida por contato na constituição da gramática do PB em Goiás.

¹¹ No Capítulo 7, trataremos mais detidamente do conceito de mudança induzida por contato linguístico, com base nos postulados teóricos e metodológicos de Thomason & Kaufman (1988) e Thomason (2001).

¹² Nos casos em que a mudança está em curso, o pesquisador pode examiná-la *in loco* e, assim, além dos estágios 1 e 2, tentar capturar os aspectos desencadeadores do novo estágio da gramática. No caso específico dos estudos da mudança em documentos históricos, se os documentos registraram, torna-se mais fácil ao pesquisador tentar reconstituir os aspectos desencadeadores do novo estágio da gramática.

Tendo a variação e a mudança como dois conceitos essenciais para o desenvolvimento da análise linguística desta tese, o objetivo deste capítulo é sintetizar os pressupostos teóricos envolvidos nesses conceitos na perspectiva da gramática gerativa. Desse modo, nossa tarefa aqui consiste em apresentar o tratamento e o desenvolvimento desses conceitos em dois estágios da gramática gerativa: no modelo de P&P e no PM. O capítulo está organizado da seguinte forma: a seção 2.1 trata da variação paramétrica na perspectiva de P&P e do PM; a seção 2.2 trata da mudança gramatical também na perspectiva desses dois estágios da teoria.

2.1 Variação paramétrica

Chomsky (1986a) postulou que a GU é dotada de princípios e parâmetros linguísticos: os princípios são composições sintáticas presentes em todas as línguas humanas e os parâmetros são arranjos gramaticais variáveis valorados no processo de aquisição de língua materna. Ambos compõem o conhecimento gramatical ricamente estruturado da mente humana. O modelo de P&P busca fornecer explicações para o rápido e completo processo de aquisição de primeira língua na infância e, por conseguinte, dar uma resposta à diversidade de línguas no mundo. Nesse modelo, a tarefa de adquirir a gramática de uma dada língua consiste em preencher as lacunas deixadas abertas pelos princípios da GU, isto é, detectar as configurações dos parâmetros que são compatíveis com o *input* (dados linguísticos primários) com o qual a criança é confrontada. Assim, durante a aquisição de primeira língua, a criança testa regras e lança mão de hipóteses, selecionando as estruturas fonológicas, morfológicas e sintáticas que constituem a gramática de sua comunidade linguística. Em suma, a visão de aquisição de primeira língua nessa perspectiva é a de que as crianças, munidas com restrições inatas sobre possíveis gramáticas fornecidas pela GU, são expostas a dados linguísticos primários, isto é, à fala de pessoas, e, assim, desenvolvem sua gramática particular, que é reconhecida em um dado contexto sociocultural como a gramática de uma língua em particular: português, espanhol, inglês etc. Nessa perspectiva, a competência linguística de um falante adulto é vista como o resultado da tradicional dicotomia entre o inato (*nature*) e o adquirido (*nurture*) na teoria gerativa, isto é, entre a faculdade da linguagem e a exposição aos dados linguísticos primários (cf. HUANG & ROBERTS, 2016).

Em P&P, a diversidade linguística é explicada como resultado da valoração de parâmetros distintos para cada língua, exceto no que concerne ao léxico. A variação translinguística é, então, atribuída a diferentes possibilidades de marcação de valores dentro de um conjunto finito de opções paramétricas, fornecidas pela GU durante o processo de

aquisição de primeira língua. Assim, na verdade, tanto a variação quanto a mudança em determinada língua podem ser explicadas em termos de parâmetros (cf. LIGHTFOOT, 1979, 1991, 1999; KROCH, 2001; ROBERTS, 2007).

Um clássico exemplo da relação entre princípios e parâmetros é o princípio de projeção estendida, já mencionado na Introdução desta tese, que requer que em todas as línguas todas as orações tenham um sujeito (CHOMSKY, 1981, 1982, 1986a). Um parâmetro determina se esse sujeito, quando pronominal, é sempre realizado fonologicamente em contextos finitos (e não finitos). Esse é o parâmetro do sujeito nulo, ou parâmetro *pro-drop*. Seus efeitos na sintaxe são ilustrados minimamente em (1), pois, conforme veremos no capítulo seguinte, o parâmetro do sujeito nulo deve, na verdade, capturar um agrupamento de efeitos na sintaxe, incluindo possivelmente: sujeitos nulos, inversão livre do sujeito, o efeito *that-trace*, a distribuição de expletivos etc.¹³

- (1) a. (Ella/él) habla español.
 b. (Lei/lui) parla italiano.
 c. *(Elle/il) parle français.
 d. *(She/he) speaks English.

A abordagem paramétrica da variação translinguística provocou uma intensa agenda de pesquisa sobre as propriedades que os parâmetros manifestam nas línguas naturais. Newmeyer (2013, p. 66) enumerou as seguintes propriedades:

- (2) Parâmetros:
- a. são descritivamente simples;
 - b. têm configurações binárias;
 - c. são pequenos em número;
 - d. são organizados hierarquicamente, tanto para a ordem de aquisição de primeira língua quanto para as generalizações tipológicas;
 - e. são entidades abstratas com uma rica estrutura dedutiva, possibilitando a predição de propriedades de estruturas morfossintáticas;

¹³ Além dos clássicos EPP e parâmetro do sujeito nulo, outros princípios e parâmetros podem ser aqui mencionados: princípio da subordinação, princípio da dependência de estrutura, princípios de ligação; parâmetro do núcleo, parâmetro do movimento *qu-*, parâmetros de movimento do verbo (movimento de V-para-T; V2); etc.

- f. junto com seu conjunto de configurações possíveis, são inatos e, portanto, universais;
- g. têm configurações que são facilmente aprendidas durante o processo de aquisição de língua;
- h. podem sofrer uma mudança rápida (ao contrário dos processos morfológicos e de gramaticalização).

Por décadas, “princípios & parâmetros” era o termo-chave para o *mainstream* dos estudos gerativistas, tendo em vista que o modelo de P&P apresentava uma solução para a tensão entre as adequações descritiva e explicativa e a realidade da diversidade linguística. P&P, aliás, representa até hoje um dos modelos mais sofisticados e elegantes de tratamento da variação translinguística (cf. NEWMAYER, 2013; BIBERAUER & ROBERTS, 2017; HOLMBERG, 2017; entre outros). O desdobramento de P&P no PM por Chomsky (1995), entretanto, eliminou boa parte do construto da teoria paramétrica utilizado para tratar da variação. Com o intuito de simplificar a arquitetura da gramática, a concepção binária de parâmetros deixou de possuir um lugar especial para dar conta da variação entre gramáticas. Diante disso, o PM passa a associar a variação linguística, incluindo a variação no eixo diacrônico, a propriedades lexicais de uma classe fechada de categorias funcionais que desencadeiam operações sintáticas para licenciar seu conteúdo morfológico (abstrato), incluindo C, T, v e D. Nesta visão, a mudança sintática é identificada por meio das mudanças que afetam o conteúdo dos traços formais das categorias funcionais.

Fuß (2016) e Holmberg (2017) chamam a atenção para o fato de que uma vertente radical no PM diz que a derivação de LF, isto é, a sintaxe estreita, é universal, de modo que a variação linguística estaria, então, restrita à derivação de PF, o *input* para a forma falada ou marcada de uma expressão linguística (cf. BERWICK & CHOMSKY, 2011; BOECKX, 2011). Uma vertente alternativa, no entanto, diz que a sintaxe permite alguma variação (cf. ROBERTS & HOLMBERG, 2010; HOLMBERG & ROBERTS, 2014; BIBERAUER & ROBERTS, 2017). Em geral, o tratamento teórico da variação no PM restringiu-se: (i) às diferenças na especificação dos itens lexicais, ou seja, nos traços formais que constituem os itens lexicais, como: Caso, traços- ϕ , traços categoriais, traço-EPP, *edge feature* etc.; e (ii) ao local onde os traços formais dos itens lexicais são interpretados. Nessa perspectiva, o parâmetro do sujeito nulo pode ser formulado como um traço associado a T finito, que interage com os traços de especificação dos pronomes (cf. HOLMBERG, 2005). Esse

tratamento dado à variação no PM está relacionado com a conjectura Borer-Chomsky proposta por Baker (2008).¹⁴

Influenciados pela conjectura Borer-Chomsky, a tendência recente dos pesquisadores gerativistas tem sido concentrar os estudos na variação microparamétrica entre línguas ou variedades. Em vez de propor parâmetros de amplo alcance (macroparâmetros) capazes de dividir as línguas do mundo em classes tipológicas amplas, conforme se vê em Baker (2008), a maioria dos linguistas gerativistas tem concentrado seus estudos na variação minuciosa entre línguas e variedades relacionadas, a qual pode ser caracterizável por pequenas diferenças no inventário lexical ou na organização categorial (cf. BAKER, 2008; GALVES, CYRINO & LOPES, 2012; HUANG & ROBERTS, 2016; para uma discussão mais detalhada sobre micro e macroparâmetros). No âmbito dessa discussão, Biberauer & Roberts (2017) estipulam uma tipologia de parâmetros que não se restringe à dicotomia micro e macro parâmetros. A proposta dos autores, proveniente de trabalhos anteriores, é a seguinte (BIBERAUER & ROBERTS, 2017, p. 149, tradução nossa)¹⁵:

- (3) Para um determinado valor v_i de um traço parametricamente variável F :
- a. macroparâmetros: todos os núcleos tipicamente relevantes, *e.g.* todas as sondas (*probe*), todos os núcleos de fase (*phase heads*) etc. compartilham v_i ;
 - b. mesoparâmetros: todos os núcleos de dada classe natural, *e.g.* [+V], ou uma categoria funcional nuclear compartilham v_i ;
 - c. microparâmetros: uma pequena subclasse lexicalmente definível de núcleos funcionais, *e.g.* verbos auxiliares modais, clíticos sujeitos, compartilha v_i ;
 - d. nanoparâmetros: um ou mais itens lexicais individuais são especificados para v_i .

Biberauer & Roberts (2017) afirmam que os efeitos dos macroparâmetros estão profundamente enraizados no sistema gramatical, de modo que são facilmente detectáveis nos dados linguísticos primários. Diante disso, eles propõem que os macroparâmetros são menos

¹⁴ A conjectura Borer-Chomsky foi proposta por Baker (2008) com base nas seguintes afirmações de Borer e Chomsky, respectivamente: (i) a variação é restrita às possibilidades que o componente flexional disponibiliza (BORER, 1984, p. 3); e (ii) a variação é restrita ao léxico; e no que diz respeito à computação sintática, a uma categoria restrita de propriedades morfológicas, principalmente flexionais (CHOMSKY, 2001, p. 2).

¹⁵ No original (BIBERAUER & ROBERTS, 2017, p. 149):

“For a given value v_i of a parametrically variant feature F :

- a. *Macroparameters: all heads of the relevant type, e.g. all probes, all phase heads, etc., share v_i ;*
- b. *Mesoparameters: all heads of a given natural class, e.g. [+V] or a core functional category, share v_i ;*
- c. *Microparameters: a small, lexically definable subclass of functional heads (e.g. modal auxiliaries, subject clitics) shares v_i ;*
- d. *Nanoparameters: one or more individual lexical items is/are specified for v_i .”*

sujeitos a reanálises por adquirentes de primeira língua em condições normais, de modo que são considerados diacronicamente mais estáveis. Um exemplo de um macroparâmetro citado pelos autores é a harmonia da ordem inicial ou ordem final: a ordem rígida da extremidade do núcleo é estável em muitos idiomas e famílias de idiomas, como, por exemplo, nas línguas dravídicas, no japonês e no coreano. Para Biberauer & Roberts (2017), o contato linguístico extensivo é provavelmente a única maneira pela qual os macroparâmetros mudam.

Os mesoparâmetros, por seu turno, embora sejam difundidos em seus efeitos, são menos definidores do sistema gramatical do que os macroparâmetros e, portanto, são mais suscetíveis a mudança. Os mesoparâmetros, de acordo com os autores, tendem a caracterizar grupos linguísticos como as principais subfamílias de uma grande família de línguas, como o indo-europeu (BIBERAUER & ROBERTS, 2017). Alguns exemplos de mesoparâmetros citados pelos autores são: a ordem VSO nas línguas celtas; o múltiplo fronteamto de sintagmas *qu-* em línguas eslavas; a ordem V2 – em que o verbo finito ocupa a segunda posição na sentença e a primeira posição é ocupada por qualquer outro elemento – em línguas germânicas, com as exceções do gótico e do inglês moderno; e os sujeitos nulos nas línguas românicas, encontrados em todas as variedades românicas, exceto nas do norte da Europa, incluindo o francês, os dialetos do norte da Itália, o reto-romanche, e no PB coloquial (BIBERAUER & ROBERTS, 2017). Os pesquisadores argumentam que os mesoparâmetros parecem mais receptivos à mudança induzida por contato linguístico. Eles defendem que as variedades românicas do norte-europeu, como o francês, perderam os sujeitos nulos devido ao contato com línguas germânicas.

Consoante a definição em (3c), microparâmetros afetam subsistemas relativamente pequenos das gramáticas. Como tais, eles são mais suscetíveis a mudanças e representam propriedades da gramática diacronicamente menos estáveis. Um exemplo de microparâmetros citado por Biberauer & Roberts (2017) são os sistemas de clíticos sujeitos dos dialetos do norte da Itália, restritos a um subgrupo do ítalo-romance e do franco-provençal. Os nanoparâmetros, por fim, são tidos como periféricos ao sistema geral da gramática, representando apenas um punhado de itens lexicais ou apenas um deles. Esses itens mostram o agrupamento de propriedades típicas de todos os parâmetros. Um exemplo é a inversão condicional no inglês moderno (como em *Had I been rich, life would have been great* – tradução: ‘Tivesse eu sido rico, a vida teria sido ótima’), que afeta apenas três verbos auxiliares: *had* (passado de *have*/'ter'), *should* (passado de *should*/'dever') e, marginalmente, *were* (passado de *be*/'ser/estar') (BIBERAUER & ROBERTS, 2017, p. 150). Nanoparâmetros tendem a ser altamente instáveis, de acordo com Biberauer & Roberts (2017).

Biberauer & Roberts (2017) acreditam que a tipologia que eles desenvolvem em (3) pode conciliar a natureza inerentemente abrupta e discreta da mudança paramétrica com o fato de que o registro histórico na maioria das vezes parece mostrar uma mudança gradual. Assim, por exemplo, uma série de mudanças microparamétricas pode dar a aparência de uma mudança mesoparamétrica gradual. Nesses termos, os rótulos macroparâmetro, mesoparâmetro, microparâmetro e nanoparâmetro não se referem a parâmetros distintos, mas ao possível grau de disseminação de um dado traço no sistema gramatical.¹⁶

Por intermédio da tipologia de parâmetros de Biberauer & Roberts (2017), podemos verificar que a variação na sintaxe do sujeito entre o PB e o PE (e as demais línguas românicas) está relacionada à noção de mesoparâmetro. Além disso, verificamos que a afirmação contudente dos autores de que a mudança no nível dos mesoparâmetros é principalmente causada por mudança induzida por contato, citando como exemplo o caso da perda de sujeitos nulos no francês como resultado do contato linguístico, está alinhada com uma das questões de pesquisa desta tese de que pode haver alguma relação entre a emergência da gramática do PB em Goiás e o contato linguístico naquela região nos séculos XVIII e XIX. Retomaremos essas questões nos capítulos 6 e 7 desta tese, a fim de especialmente discutir se o contato linguístico no período colonial brasileiro pode ter tido algum papel na sintaxe do sujeito no PB, assim como o contato do francês com as línguas germânicas teve um impacto na sintaxe do sujeito no francês, distanciando-o das línguas românicas de um modo geral.

Por fim, cabe destacar que, assim como parâmetros, a noção de princípios também passou a ser evitada no PM. Chomsky (1995) propôs, por exemplo, que o EPP se reduzisse a um traço D em T. Nessa perspectiva, Chomsky (2015) estabelece uma análise para o princípio de projeção estendida e para o princípio de categoria vazia (do inglês, *Empty Category Principle* – ECP) em termos de rotulagem (*Labeling*). Chomsky (2015) considera que uma gama de propriedades excepcionais do sujeito, todas relacionadas ao EPP e ao ECP, e sua variação translinguística constituem um enigma de longa data para a teoria gerativa. Desse

¹⁶ Biberauer & Roberts (2017) desenvolvem sua argumentação em torno de uma perspectiva de aquisição de parâmetros chamada por eles de visão emergente dos parâmetros (do inglês, *emergentist view of parameters*). Nessa perspectiva, os parâmetros da GU não são pré-especificados na dotação inata, ou seja, eles não fazem parte do Fator 1 proposto por Chomsky (2005a). Em vez disso, segundo os autores, os parâmetros emergem da interação de todos os três fatores indicados por Chomsky (2005a). Os autores também argumentam que a própria GU deixa certas opções subespecificadas, de modo que as lacunas são preenchidas durante o processo de aquisição de língua pelos falantes, interagindo com os dados linguísticos primários e equipadas com certas estratégias de aquisição de domínio geral.

modo, ele sugere que seria um grande avanço para a teoria, caso seja alcançado, unificar esses princípios por meio de uma única análise.¹⁷

Para Chomsky (2015), existe uma semelhança em ambos os casos EPP e ECP no inglês: na fase de rotulagem, o sujeito deve ser visível, não uma cópia, na posição de SPEC em [C [SPEC-TP]]. Essa observação faz com que Chomsky (2015) sugira que T, em inglês, é fraco para servir como um rótulo, de modo que, com sujeito manifesto, a construção SPEC-TP é rotulada pelos traços de concordância $\{\varphi, \varphi\}$. Assim, desde que T concorde com o sujeito, o movimento do sujeito para [SPEC-TP] estabelece os traços compartilhados de T e o sujeito como o rótulo do conjunto $\{DP, TP\}$. Por meio da rotulagem, Chomsky (2015) fornece uma motivação independente para a razão pela qual um DP sujeito se move para SPEC-T. Nessa visão, o inglês e línguas similares satisfazem o EPP (CHOMSKY, 2015).¹⁸

Retomando o trabalho de Rizzi (1982) sobre as LSN como o italiano, Chomsky (2015) afirma que o ECP parece ser livremente violado nessas línguas e que a diferença paramétrica entre elas e as línguas do tipo do inglês pode ser atribuída à riqueza morfológica daquelas. Desse modo, em termos de rotulagem, o T do italiano, com morfologia verbal rica, pode rotular TP e também $\{SPEC, TP\}$; como vimos, no caso do inglês, língua com concordância verbal fraca, isso não é possível, de modo que SPEC deve estar visível quando o algoritmo de rotulagem se aplica (CHOMSKY, 2015). Assim, Chomsky (2015) conclui que o italiano não se comporta da mesma maneira que o inglês em relação ao EPP e ao ECP.

Chomsky (2015) comenta que se costuma supor que o EPP também é válido para línguas como o italiano e, por essa razão, foi postulado que um sujeito pronominal nulo (*pro*) ocupa a posição de [SPEC-TP] nessas línguas. Chomsky (2015) argumenta, entretanto, que há pouca evidência para essa suposição, que pode ser, na opinião do autor, um acidente histórico nos estudos gerativistas sobre a sintaxe dos sujeitos nulos: como o inglês foi investigado primeiro, naturalmente estenderam-se os princípios descobertos para o inglês para outras línguas. Chomsky (2015) conclui que o italiano não tem EPP, assim como não está em conformidade com o ECP, unificando assim o EPP e o ECP, presentes no inglês, em termos de rotulagem e mantendo a diferença paramétrica do italiano em relação ao inglês em termos de riqueza de concordância. Chomsky (2015) observa, ainda, que o movimento do sujeito para SPEC-T não está bloqueado no italiano; ele só não é necessário. Isso, aponta o autor,

¹⁷ O ECP foi inicialmente postulado na teoria gerativa como uma restrição sintática universal que requer que certos tipos de categorias vazias (isto é, vestígios) sejam apropriadamente regidos (CHOMSKY, 1981).

¹⁸ Epstein & Seely (2006) chamam a atenção para outros fatores como determinantes para o movimento e a projeção do especificador e questionam a postulação do EPP no contexto do minimalismo.

deixa em aberto uma série de questões sobre exatamente onde o sujeito e VP estão quando o sujeito não é projetado para SPEC-T. Essas questões não serão abordadas nesta tese.

2.2 Mudança sintática

Numa perspectiva teórica que considera a gramática das línguas naturais como resultado da fixação de parâmetros em um aparato genético da espécie humana cujos princípios são imutáveis, como em P&P, é natural que se pressuponha que a mudança sintática deva ser um resultado de mudança paramétrica no percurso histórico da língua. Para Lightfoot (1991, 1999), a mudança é iniciada quando uma população de aprendizes converge em um sistema gramatical que difere em pelo menos um valor de parâmetro do sistema internalizado dos falantes cujo comportamento linguístico fornece a entrada para os aprendizes. Como a geração mais jovem substitui a mais antiga, a mudança é realizada através da comunidade de fala. Em resumo, Lightfoot (1991, 1999) assume que a mudança sintática é um aspecto do processo de marcação de parâmetros durante a aquisição de primeira língua e, portanto, está intrinsecamente ligada à língua-I. Certamente, muitos fatores sociais, históricos e culturais influenciam as comunidades de fala e, portanto, o processo de mudança. Do ponto de vista da gramática gerativa, porém, o pesquisador costuma abstrair esses fatores e tenta focar sua análise na mudança como uma relação exclusiva entre sistemas gramaticais (cf. ROBERTS & ROUSSOU, 2003). Nesse caso, o pesquisador ignora os registros históricos até onde lhe for permitido (cf. ROBERTS & ROUSSOU, 2003).

A conexão entre a mudança gramatical e a aquisição de língua foi primeiramente formulada em termos gerativos pela obra seminal de Lightfoot (1979), *Principles of diachronic syntax*. Nessa obra, o autor integrou as preocupações da pesquisa diacrônica com as da gramática gerativa ao postular que a mudança gramatical é impulsionada pela reanálise dos dados linguísticos primários na aquisição da primeira língua, de modo que o resultado pode ser uma nova geração poder gerar uma nova gramática para uma determinada língua. Pela primeira vez, a compreensão da mudança gramatical passou a ser estritamente ligada à elaboração de teorias restritivas da gramática, que poderiam definir os limites da variação e, por consequência, o possível resultado da mudança, dada uma certa experiência desencadeadora. Uma das principais ideias propostas por Lightfoot (1979) foi o princípio de transparência, segundo o qual as derivações transformacionais eram limitadas no grau de opacidade que elas poderiam induzir. Se uma gramática se tornasse muito opaca, o princípio

de transparência forçaria uma reanálise. O autor postulou o princípio da transparência como um princípio da gramática, isto é, uma condição geral sobre a natureza das derivações.

Lightfoot (2017a) exemplifica que quantificadores do inglês (*all, any, each, both, every* etc.) costumavam ter a mesma distribuição que os adjetivos no inglês antigo e médio; ambas as categorias, segundo Lightfoot (2017a), pertenciam a uma supercategoria rotulada de modificadores. Com o passar do tempo, a distribuição dos modificadores que viriam a ser quantificadores permaneceu constante, mas os modificadores que se tornariam adjetivos modernos desenvolveram novas propriedades (cf. LIGHTFOOT, 1979). No final do século XV, modificadores destinados a se tornarem adjetivos tornaram-se mais restritos em sua distribuição: não podiam mais ocorrer livremente diante de um determinante, com um genitivo ou com um nominal. Para o pesquisador, isso sugere que os modificadores acumularam gradualmente três características de exceção, impedindo tais ocorrências. Posteriormente, isto é, no final do século XVI, ocorreram as seguintes mudanças:

- (4)
- a. *all* e *both* aparecem com partitivos;
 - b. obsolescência do determinante-quantificador-substantivo;
 - c. obsolescência de múltiplos quantificadores em sequência;
 - d. obsolescência do adjetivo-determinante-substantivo;
 - e. obsolescência de adjetivos pós-nominais;
 - f. obsolescência de adjetivos usados como nominais. (LIGHTFOOT, 2017a, p. 327)

Lightfoot (2017a) argumenta que, como os modificadores pré-adjetivos e pré-quantificadores passaram a divergir em sua distribuição, havia uma boa razão para os adquirentes de língua postularem que eles eram diferentes categorias: adjetivos ocorriam à esquerda de um substantivo ou seguindo um verbo de cópula, como ilustrado em (5), enquanto quantificadores ocorriam como o especificador de um DP ou em uma posição determinante com um complemento NP ou PP, conforme ilustrado em (6).

- (5) a. NP[_A*red* NP[*books*]] (LIGHTFOOT, 2017a, p. 328)

b. VP[_v*become* _A*angry*]

- (6) a. DP[_{Spec}*all* DP[*his apples*]] (LIGHTFOOT, 2017a, p. 328)

b. DP[_D*both* NP[*books about taxes*]], DP[_D*all* PP[*of the books*]]

Essas estruturas, segundo Lightfoot (2017a), mostram que (4a) pode ser gerado, mas (4b-f) não. A reanálise gramatical aparece aí como um efeito do princípio de transparência: os traços opacos foram eliminados e a nova análise, mais transparente, produziu mudanças.

Ao adaptar sua abordagem inicial sobre a mudança sintática ao modelo de P&P, Lightfoot (1991, 1999) assume que a mudança sintática envolve a redefinição dos valores dos parâmetros. A partir daí, postula que os parâmetros são redefinidos no decurso da aquisição de primeira língua, presumivelmente com base na reanálise dos dados linguísticos primários por adquirentes de primeira língua. P&P representa, assim, um modelo teórico que pode explicar tanto os agrupamentos de propriedades sincrônicas da gramática quanto os agrupamentos de propriedades diacrônicas. Nessa direção, muitos pesquisadores trabalhando com linguística histórica e teoria gerativa estavam convencidos de que todos os principais tipos de mudança sintática podiam ser reformulados em termos de operações de reanálise de parâmetros, as quais parecem ser o principal mecanismo explicativo da mudança sintática (cf. ROBERTS, 2007; MADARIAGA, 2017). A interpretação da mudança gramatical como um exemplo particular de reanálise de parâmetros, isto é, como um fenômeno conduzido pela aquisição de língua, refuta qualquer teoria de mudança que recorre a explicações envolvendo mais de uma geração. No entanto, considerando uma comunidade de falantes monolíngues e homogênea, essa explicação parece levar a um paradoxo: como uma gramática alvo G pode produzir na nova geração um *output* que difere significativamente do *input* que levou à aquisição de G na geração anterior? Isso é, às vezes, chamado de problema lógico da mudança gramatical, a versão gerativista do problema da implementação de Weinreich, Labov e Herzog (1968) (cf. ROBERTS, 2007; WALKDEN, 2017).

Postular que a mudança sintática ocorre durante o processo de aquisição de primeira língua implica considerar que os dados linguísticos primários não são *a priori* suficientes para a marcação de todos os valores paramétricos de uma dada língua – ainda que, paradoxalmente, se deva admitir que, independentemente de os dados serem insuficientes, uma opção paramétrica tem de ser escolhida (ROBERTS, 2007). Ademais, esse postulado reforça o papel do aparato genético da linguagem humana ao considerar as intuições de que uma criança lança mão quando os dados linguísticos primários lhe são apresentados (cf. LIGHTFOOT, 1979, 1991, 1999, 2006; KROCH, 2001, 2005; ROBERTS & ROUSSOU, 2003; ROBERTS, 2007; entre outros). Ao se asseverar que uma opção paramétrica tem de ser escolhida, enfatiza-se também o fato de que o processo de aquisição de língua é, primordialmente, caracterizado pela fixação de parâmetros. Assim, não há possibilidade de não se marcar um parâmetro da GU – e, uma vez que um parâmetro é marcado, não há como

desmarcá-lo (cf. KROCH, 2005; ROBERTS & ROUSSOU, 2003; ROBERTS, 2007; entre outros). Nessa perspectiva, a mudança na gramática de uma língua não é entendida a partir de concepções como frequências, *clines* (gradações), *continua* e tendências, mas, outrossim, como mudança na marcação de opção paramétrica (cf. KROCH, 2005; ROBERTS, 2007).

Diferentemente de propostas que alegam haver alguma ambiguidade estrutural nos dados linguísticos primários para haver a reanálise no processo de aquisição de língua (cf. ROBERTS, 2007), Madariaga (2017) chama a atenção para o fato de que nem todos os estudos gerativistas em perspectiva diacrônica concordam com a explicação da ambiguidade para a reanálise dos dados linguísticos primários, pois é facilmente constatável que uma nova estrutura não pode surgir apenas com base numa ambiguidade que só é detectada uma vez que ocorreu a mudança, ou seja, após a reanálise. Um exemplo que podemos citar nesse sentido é o das construções com *se*-impessoal e com sujeitos nulos indefinidos na 3ª pessoa do singular no PB: esse último caso, ilustrado em (7b), só pôde ser analisado como ambíguo em relação ao primeiro, exemplificado em (7a), porque surgiu abruptamente na gramática da língua.

- (7) a. Vende-se casas
b. Vende casas.

Roberts (2007) e Biberauer & Roberts (2017) defendem, contudo, que um certo grau de ambiguidade dos dados é necessário para que a mudança ocorra. Roberts (2007) recorre à parametrização e à GU para defender seu ponto de vista: a GU restringe a gama de ambiguidades potenciais nos dados linguísticos primários, nos quais diferentes parâmetros possíveis podem ser definidos durante o processo regular de aquisição de primeira língua. Biberauer & Roberts (2017), com base na noção de parâmetro em (3), afirmam que, se um valor definido v_i é inequivocamente expresso para um parâmetro p_i nos dados linguísticos primários, então os adquirentes irão convergir com sucesso em v_i . Ou seja, o parâmetro não será alterado se não houver ambiguidade – esse fenômeno, aliás, é chamado de princípio de inércia: *as coisas ficam como estão a não ser que uma força atue sobre elas* (cf. KEENAN, 2002; LONGOBARDI, 2001; BIBERAUER & ROBERTS, 2017). Por outro lado, a mudança gramatical ocorre quando há um valor inovador v_j de um parâmetro p_k , isto é, uma propriedade lexical, um traço vinculado a determinada categoria funcional, que é por algum motivo mais acessível nos dados do *input* do que o relevante valor v_i que é parte da gramática alvo (FUß, 2016).

Ressalta-se que, nessa perspectiva, a reanálise não é algo que os falantes adultos fazem em suas línguas maternas. A reanálise é um tipo de análise do *input* recebido por um adquirente, ou seja, é uma das maneiras pela qual as crianças adquirem suas gramáticas. A única diferença entre gramáticas sem reanálise e gramáticas reanalisadas é que, no caso das primeiras, a gramática do adquirente converge com a gramática que forneceu os *inputs*, enquanto, no caso das segundas, a nova gramática não converge com aquela que forneceu os dados linguísticos primários (cf. LIGHTFOOT, 1991, 1999; MADARIAGA, 2017). Tanto as gramáticas convergentes quanto as não convergentes são iguais na medida em que devem ser adquiridas novamente de acordo com o procedimento usual de aquisição de língua, ou seja, com base nos dados linguísticos primários, o *input* disponível em um ambiente linguístico (MADARIAGA, 2017). Madariaga (2017) afirma que a não convergência de gramáticas tem sido chamada por linguistas gerativos de *falha na transmissão* ou de *descontinuidade*.

Kroch (2001) é um dos pesquisadores que defende que a mudança gramatical é provocada por uma falha na transmissão linguística durante o processo de aquisição de língua materna. O autor discorre que o entendimento do que vem a ser uma falha na transmissão de traços linguísticos é limitado, pois não se pode precisar a relação entre a evidência apresentada ao aprendiz e a gramática que ele adquire. Argumenta, ainda, que a mudança sintática decorrente de uma falha na transmissão também pode ser provocada por adultos aprendizes de segunda língua, especialmente em casos de contato linguístico. Para o autor, na verdade, nos casos de contato e aquisição de segunda língua, pressupõe-se que a aquisição da nova língua é naturalmente imprecisa.

Kroch (2001) pondera que no percurso diacrônico de uma dada língua a variação pode ser encontrada em textos particulares de um único indivíduo. Ele argumenta que um indivíduo pode ter mais de uma gramática mental para o que é, em termos socioculturais, a mesma língua ou registro. Em outras palavras, um único falante de uma dada língua pode ter mais do que uma língua-I mesmo na ausência de múltiplas línguas-I fornecendo *inputs* durante o processo de aquisição de língua desse indivíduo. Esse falante, inclusive, pode usar essas línguas-I distintas num mesmo registro ou contexto sociolinguístico (KROCH, 2001). Sob esse ponto de vista, diferentes registros da mesma língua que um único falante pode ter dominado podem justificar a postulação de mais de uma língua-I para uma determinada língua, nos mesmos termos assumidos para os casos de bilinguismo.¹⁹

¹⁹ Em Roeper (1999), a ideia de competição de gramáticas é estendida às gramáticas monolíngues de forma mais geral e chamada de *bilinguismo universal*. A assunção fundamental do bilinguismo universal é a de que escolhas aparentemente contraditórias em uma dada língua, dentro do contexto da opcionalidade *versus* a obrigatoriedade,

Kroch (2005) defende a tese de que um aprendiz define um parâmetro “erroneamente” porque certos dados foram ignorados ou não encontrados durante o processo de aquisição de língua. Esse aprendiz, contudo, ainda pode aprender os parâmetros da língua de sua comunidade linguística (isto é, os parâmetros da geração antiga) “corretamente”, mas apenas da maneira com que ele aprende outro idioma ou dialeto, qual seja, adicionando conhecimento ao invés de substituir um parâmetro por outro. O resultado disso é o que Kroch (2005) chama de bidialetismo em relação ao parâmetro aprendido primeiramente. O bidialetismo decorre, portanto, da aquisição imperfeita de um traço linguístico (parâmetro adquirido “erroneamente”) e da posterior aquisição do traço linguístico (parâmetro) “correto”. O bidialetismo apresenta as seguintes características em seu percurso diacrônico:

Bidialectalismo causado por erro de aprendizagem e correção é a fonte, no modelo de transmissão imperfeita aqui desenvolvido, da mudança sintática. Por pelo menos uma razão, o processo de correção de parâmetros adquiridos erroneamente não pode ser infalível. Se uma criança ouve a língua produzida por outra que tenha um parâmetro “incorreto”, então a criança que escuta terá, na verdade, uma evidência positiva da configuração incorreta e pode ser induzida a definir o parâmetro incorretamente. Qualquer viés no sistema que favoreça a configuração incorreta pode, sob circunstâncias apropriadas, favorecer a disseminação eventual da configuração originalmente incorreta. (...) Uma vez que a mudança sintática esteja em andamento, o adquirente ouvirá evidências na comunidade de fala sobre configurações mutuamente inconsistentes e aprenderá ambas. No decorrer do tempo histórico, uma configuração provavelmente vencerá a outra, por razões que ainda não são bem compreendidas. (...) (KROCH, 2005, p. 3, tradução nossa).²⁰

Em suma, nos casos de bidialetismo em perspectiva diacrônica, um cenário linguístico gradualmente ganha espaço sobre o outro. Kroch (2005) defende, portanto, que as formas que dão evidências para ambas as configurações paramétricas aparecem com frequência e todos os aprendizes podem adquirir e usar ambas as configurações. No decorrer do tempo, todavia, a frequência de uma das duas formas cai até, eventualmente, ela ocorrer muito raramente para

são evidências de gramáticas em competição. O autor propôs que cada falante gera um conjunto de minigramáticas que os torna universalmente bilíngues em seu próprio idioma. Nesse sentido, Roeper (1999, p. 173) exemplifica que o inglês, uma língua não *pro-drop*, pode apresentar um registro social informal, conhecido como *diary drop*, em que os sujeitos das orações são omitidos: “*seems like a good idea*”/“*looks good to me*”.

²⁰ No original: “*Bidialectalism due to mislearning and correction is the source, in the imperfect transmission model as developed here, of language change. For at least one reason the process of correction of misset parameters cannot be infallible. If one child hears language produced by another that has misset a given parameter, then the listening child will actually get positive evidence for the incorrect setting and may be induced to itself set the parameter incorrectly. Any bias in the system that favors the incorrect setting may, under appropriate circumstances, favor the eventual spread of the originally incorrect setting. (...) Once language change is in progress, the learner will hear evidence in the speech community for mutually inconsistent settings and learn both. Over historical time, one setting is likely to win out over the other, for reasons that have yet to be well understood. (...)*” (KROCH, 2005, p. 3).

produzir sua aquisição pelas gerações futuras. Desse modo, instaura-se a instabilidade de um dos parâmetros na gramática.

Kroch (2005) defende que, nos casos de aquisição de língua materna, a convergência da gramática da criança com a gramática do adulto é probabilística: “dada uma população de adquirentes, haverá erros com uma distribuição de probabilidade que depende da força da evidência para determinadas propriedades gramaticais e da variação nos dados primários aos quais os aprendizes estão expostos” (KROCH, 2005, p. 1-2, tradução nossa).²¹ Kroch (2005) aponta que o principal problema dessa abordagem é que é impossível saber se os erros de aprendizagem são suficientemente comuns para servir como uma importante causa da mudança.

Kroch (2005) contrapõe o modelo de transmissão imperfeita entre gerações com o modelo de *deriva interna*, segundo o qual a mudança está condicionada a uma mudança gradual na frequência dos usos das formas linguísticas, de maneira que, se a frequência de uma forma linguística cai abaixo do limite capaz de propiciar o aprendizado, a mudança gramatical ocorre (cf. NARO & SCHERRE, 2007, para uma discussão sobre o modelo de deriva interna utilizando a formação do PB como exemplo). Para Kroch (2005), o modelo de transmissão imperfeita explica melhor a mudança, pois se baseia no fato de que ela começa, realmente, com uma mudança na gramática, a qual é responsável por sua propagação na comunidade de fala e pela frequência com que certas formas linguísticas são usadas. Sob esse ponto de vista, a gradualidade nos dados históricos não representa imprecisão (ou seja, nesse modelo, a imprecisão está nos dados linguísticos primários), mas sim duas entidades distintas na mente dos falantes, uma substituindo gradualmente a outra. Estudos nessa perspectiva, por exemplo, foram aplicados à mudança da ordem das palavras na história do ídiche (SANTONIRI, 1992), do grego (TAYLOR, 1994) e do inglês antigo (PINTZUK, 1999).

A abordagem de Kroch (2001, 2005) para a mudança sintática gira em torno da chamada competição de gramáticas, que busca fornecer uma explicação para a opcionalidade vista em textos produzidos pelos falantes de uma dada língua (cf. KROCH 1989, 2001, 2005; SANTORINI, 1989, 1992; TAYLOR, 1994; PINTZUK, 1999; YANG, 2002; e outros). Nessa hipótese, quando uma mudança está ocorrendo em uma dada língua, alguns falantes têm acesso a duas variantes diferentes da mesma forma ou estrutura, uma variante antiga e uma recém-criada, inovadora. Os falantes que exibem as duas variantes, na verdade, exibem duas

²¹ No original: “Given a population of learners, there will be errors with a probability distribution that depends on the strength of the evidence for given grammatical properties and on variation in the primary data that learners are exposed to.” (KROCH, 2005, p. 1-2).

gramáticas diferentes (diglossia) e recorrem à alternância de código enquanto usam uma ou outra, da mesma maneira que os bilíngues fazem com suas diferentes línguas.²²

Biberauer & Roberts (2017) lembram que uma explicação complementar para uma visão de diglossia sobre gramáticas concorrentes é que elas podem coexistir em um falante porque pertencem a diferentes níveis linguísticos. Isto é, a existência de duas variantes nas produções linguísticas pode sinalizar que uma das variantes pertence à competência linguística do falante (língua-I), enquanto a outra é externa a essa competência, uma variante usada episodicamente no conjunto de enunciados produzidos por uma comunidade linguística, também chamada de língua-E (E de externa, extensional). Um exemplo que podemos fornecer nesse sentido é o uso das construções com *se*-passivo, restritas à escrita padrão no PB (cf. MARTINS & NUNES, 2016), em contraste com as construções com *se*-impessoal ou com sujeitos nulos indefinidos na 3ª pessoa do singular, presentes no PB tanto na escrita quanto na fala (8b-c) – mencionadas anteriormente em (7). Nesse sentido, as construções com *se*-passivo no PB, como em (8a), podem ser tomadas como enunciados exclusivos da língua-E.²³

- (8) a. Vendem-se casas. (*Se*-passivo).
 b. Vende-se casas. (*Se*-impessoal).
 c. Vende casa. (Sujeito nulo indeterminado).

De acordo com Fuß (2016), os gerativistas concordam que o objeto apropriado do estudo científico da linguagem humana é a língua-I. A língua-I é geralmente contrastada com a língua-E, definida como o conjunto de expressões reais ou potenciais que estão em uso em uma comunidade linguística. Assim, evidências linguísticas consistem em dados da língua-E. O objetivo do linguista gerativo é detectar as propriedades da língua-I, inspecionando o comportamento linguístico dos indivíduos (FUß, 2016). Qualquer princípio ou regra de gramática que seja postulado pelo linguista gerativo deve ser visto como uma parte da língua-I. A GU é então, como vimos na Introdução, interpretada como uma teoria dos universais formais que identifica o conjunto de possíveis gramáticas das línguas humanas (cf. BERWICK & CHOMSKY, 2016). Adotar essa visão para o estudo da mudança sintática

²² Contrastando Kroch (2001, 2005) com Lightfoot (1979), a mudança de uma gramática para outra na perspectiva de ambos é, respectivamente, gradual e abrupta.

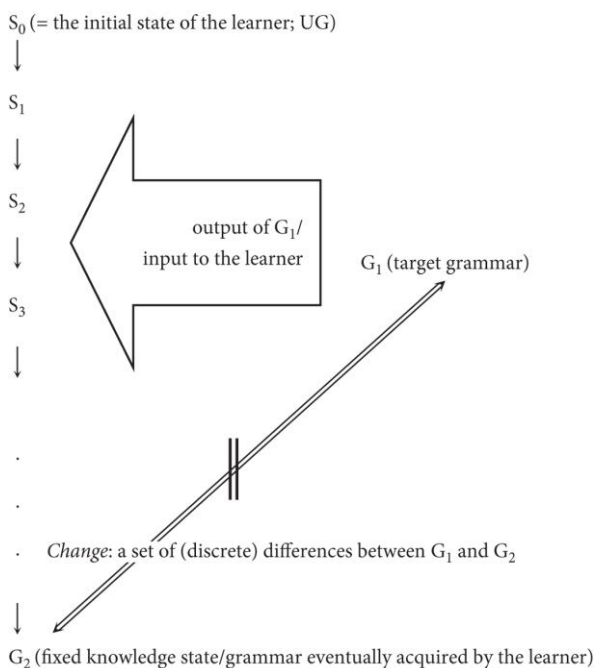
²³ No Capítulo 4 desta tese, apresentamos uma análise detalhada das construções em (8), a fim de evidenciar o papel crucial que elas desempenham na distinção formal entre as gramáticas do PB (LSN parciais) e do PE (LSN consistentes), um dos pilares da sustentação teórica desta tese.

implica afirmar, portanto, que o objeto apropriado da linguística formal histórica deve ser a(s) língua(s)-I ou a(s) gramática(s)-I (Fuß, 2016).

Fuß (2016) e Lightfoot (2017b) afirmam que uma perspectiva de língua-I sobre a mudança sintática evidencia o fato de que a aquisição de língua entre uma geração de falantes e outra é necessariamente descontínua – termo que, conceitualmente, difere da noção de transmissão (imperfeita) e se associa a uma propriedade externa das línguas humanas: entende-se descontinuidade como a emergência de uma nova gramática-I desencadeada por novas evidências da língua-E, ou seja, por novos dados linguísticos primários (LIGHTFOOT, 2017b).²⁴ Nessa perspectiva, durante o processo de aquisição de primeira língua, as crianças não têm acesso direto às propriedades (abstratas) da gramática alvo; em vez disso, elas constroem uma gramática baseada nos dados linguísticos primários que recebem, os quais são potencialmente diferentes e novos em relação àqueles que geraram a gramática da geração anterior. Desse modo, a mudança sintática expressa na gramática da nova geração ocorre não porque traços linguísticos deixaram de ser transmitidas no decorrer da aquisição de língua, mas porque os dados linguísticos primários eram novos em relação àqueles que geraram a gramática da geração anterior. Vejamos o modelo de aquisição de língua e de mudança sintática na Figura 1.

Figura 1: Modelo de aquisição e mudança linguística.

A model of language acquisition and change



Fonte: Fuß (2016, p. 463). Legenda: G1 (gramática 1); G2 (gramática 2).

²⁴ Dizer que a *descontinuidade* é uma propriedade externa às línguas humanas a vincula ao conjunto de princípios e restrições que integram o terceiro fator do *design* das línguas humanas apresentado na Introdução.

Em relação à Figura 1, Fuß (2016) explica que o aprendiz – partindo de S_0 , o estado inicial da gramática, geralmente considerado uma expressão dos genes, que pode ser modelado em termos de um sistema de princípios abstratos (GU) – constrói um número de estágios intermediários de conhecimento durante o processo de aquisição de primeira língua com base nas evidências fornecidas pelo *input*, onde S_n é revisado para S_{n+1} sempre que o aprendiz toma conhecimento das evidências necessárias para ativar uma certa propriedade da gramática, que não é necessariamente uma propriedade da gramática da geração anterior. Finalizando, o processo de construção gramatical chega a um estado de conhecimento fixo que representa a gramática adquirida pelo aprendiz no decorrer do processo de aquisição de língua (Fuß, 2016).

Lightfoot (2017b) argumenta que a descontinuidade não constitui um evento especial, mas o estado normal das coisas que confronta qualquer criança durante o processo de aquisição de primeira língua. Desse modo, as línguas-E são o *locus* da mudança no processo de aquisição de primeira língua, e as línguas-I, em particular, são geradas em cada nova geração e em cada indivíduo. Nessa perspectiva, finalmente, o pesquisador pode vincular aspectos particulares das línguas-E às novas línguas-I. Para Lightfoot (2017b), se o trabalho sobre mudança fornece uma visão sobre a vinculação de aspectos específicos das línguas-E às novas línguas-I, isso constitui uma contribuição importante para a compreensão da aquisição de línguas, o que, para o autor, ainda não foi alcançado pela pesquisa experimental.

Lightfoot (2017b) afirma ser impossível que duas crianças recebam os mesmos dados linguísticos primários, pois elas naturalmente ouvem coisas distintas. No entanto, apesar da variação na experiência, as crianças vão muitas vezes atingir a mesma língua-I em termos do conjunto de estruturas sintáticas. As experiências individuais podem variar indefinidamente, mas as línguas-I dentro de uma comunidade linguística mostram estabilidade estrutural e variam apenas de forma limitada. Ou seja, conforme salienta Chomsky (2015), numa dada comunidade linguística, as línguas-I podem não ser inteiramente idênticas, mas se sobrepõem substancialmente. Além disso, Lightfoot (2017b) destaca que os dados linguísticos primários podem variar em diferentes modos e, ainda assim, não conceber novas línguas-I.

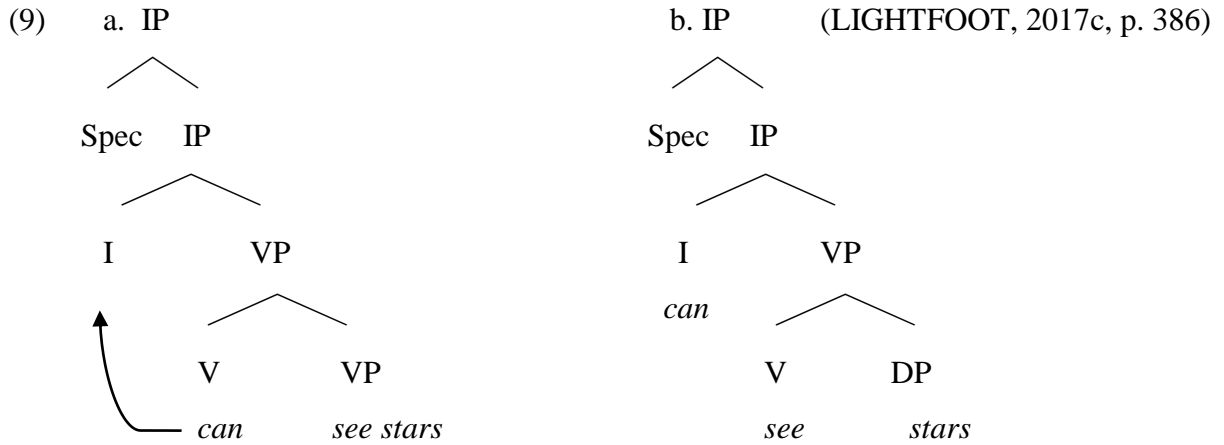
De acordo com Lightfoot (2017b, 2017c), a partir do conceito de descontinuidade, surge uma nova compreensão nos estudos gerativistas sobre a mudança gramatical, que passa a ser entendida também como um fenômeno mais individual que pode afetar a experiência linguística de outros e levar a uma mudança no comportamento de toda a comunidade linguística. Esse seria o caso, por exemplo, dos falantes bilingues de inglês escandinavo que

teriam desencadeado uma mudança generalizada na gramática do inglês médio, qual seja, a perda da morfologia verbal (cf. LIGHTFOOT, 2017b, 2017c). Nessa perspectiva, novas línguas-E também resultam do contato linguístico e do bilinguismo, e línguas-E resultantes do contato linguístico e do bilinguismo também geram novas línguas-I.²⁵

Referenciando o estudo de O’Neil (1978), Lightfoot (2017b, 2017c) menciona a perda maciça da morfologia verbal no inglês médio como um dos resultados do estreito contato dos falantes de inglês com os falantes escandinavos no norte da Inglaterra, mais especificamente pelo bilinguismo inglês-escandinavo, no qual os indivíduos com as línguas-I inglesa e escandinava apresentavam inicialmente um sistema morfológico rico para cada língua, semelhantes entre si, mas diferentes e não aprendidos como um único sistema (LIGHTFOOT, 2017b). A partir desse período de contato linguístico e bilinguismo, registra-se que, diferentemente da maioria das línguas europeias, o inglês mudou drasticamente seu paradigma flexional, restando apenas a conjugação *-s/-eþ* da 3ª pessoa do singular do tempo presente – para Lightfoot (2017b), ainda não se tem uma explicação consistente sobre como esse morfema sobreviveu no paradigma flexional da língua.

Lightfoot (2017c) mostra que o inglês primitivo tinha propriedades morfológicas complexas e que naquele período do idioma era possível encontrar, por exemplo: (a) as conjugações *fremme*, *fremst*, *fremþ* e *fremmaþ* no tempo presente e *fremed*, *fremedest*, *fremede* e *fremedon* no passado de *do* ‘fazer’; (b) as conjugações *sēo*, *sieħst*, *sieħþ* e *sēoþ* no tempo presente de *see* ‘ver’; (c) e as conjugações *rīde*, *rītst*, *rītt* e *rīdaþ* no tempo presente e *rād*, *ride*, *rād* e *ridon* no pretérito de *ride* ‘montar’. Lightfoot (2017c) mostra também que, nos primórdios do inglês moderno, verbos auxiliares modais do tipo *can*, *could*, *must*, *may*, *might*, *will*, *would*, *shall*, *should* e *do*, perderam o movimento do verbo para Infl e foram recategorizados como itens de Infl, conforme exemplificado, respectivamente, em (9a) e (9b). O autor menciona que, por determinado período, ambos os sistemas coexistiam e possivelmente competiam entre si: verbos como *can* eram gerados em VP e moviam-se para Infl ou eram gerados diretamente em Infl. Com o tempo, entretanto, (9a) tornou-se rara, e (9b) mais produtiva, vindo a se estabilizar.

²⁵ Considerando que o conceito de descontinuidade abarca, de certo modo, a mudança induzida por contato, inferimos que a mudança ocorrida na gramática do PB em Goiás nos idos de oitocentos – nosso objeto empírico de pesquisa – pode ser resultante do contato linguístico. Trataremos dessa questão detidamente no Capítulo 7.



Com base nessas evidências, Lightfoot (2017c) supõe que diacronicamente as línguas-E que as crianças ouviram mudaram de tal forma que auxiliares modais modernos como *can*, *may*, *shall* etc. vieram a ser morfologicamente distintos de outros verbos, até porque eles nunca possuíram a única característica restante da morfologia verbal do tempo presente, a terminação *-s/-es* da 3ª pessoa do singular. Além disso, Lightfoot (2017c) mostra que as formas do tempo passado desses verbos (*could*, *would*, *might*, etc.) tinham significados que não eram tempo passado, refletindo antigos usos subjuntivos: “*They might/could/would leave tomorrow*”. Para Lightfoot (2017c), essas evidências indicam que esses verbos modais foram recategorizados na gramática interna dos falantes porque se tornaram formalmente distintos de outros verbos como resultado da simplificação morfológica dos verbos. Assim, houve o que Lightfoot (2017c) chama de *efeito dominó*: mudanças no que as crianças ouviram e a morfologia do verbo recentemente reduzida levaram a uma categorização diferente de certos verbos. Em outras palavras, as novas línguas-E resultantes do contato e do bilinguismo entre falantes de inglês e escandinavo acionaram novas línguas-I com verbos modais do tipo *can* recategorizados como itens de Infl, como ilustrado em (9b). Aí constam, de acordo com Lightfoot (2017b), a descontinuidade e sua explicação, ambas evidenciando que não há aprendizado imperfeito ou transmissão imperfeita, apenas novas línguas-E acionando novas línguas-I.²⁶

²⁶ Kroch, Taylor & Ringe (1995) apresentaram um estudo com conclusões similares às de Lightfoot (2017b, 2017c). A partir da constatação de que, no inglês médio, havia uma distinção entre a sintaxe V2 do sul e a sintaxe V2 do norte e do leste da Inglaterra, esses autores propõem que a sintaxe V2 inovadora surge como uma consequência da perda das flexões verbais no inglês médio do norte – que levou também à perda do movimento de V-para-T, tal como teria ocorrido nas línguas escandinavas continentais quando houve a perda das flexões verbais nessas línguas. A perda da concordância rica, por sua vez, é considerada um resultado da aquisição imperfeita do inglês pelos invasores escandinavos, que formavam uma proporção suficientemente grande da população do norte naquele período da língua para impor efeitos de substrato ao dialeto do norte.

Cabe ressaltar que a partir desses dados envolvendo a morfologia verbal e a recategorização de verbos modais em Infl no inglês exemplificados por Lightfoot (2017b, 2017c), podemos inferir que, se a morfologia flexional fonologicamente expressa for tomada para refletir o conteúdo de traços abstratos de categorias funcionais, é possível construir uma correlação entre mudança sintática e mudança morfológica, isto é, fornecer uma explicação baseada na observação tradicional de que mudanças que afetam a morfologia flexional de uma dada língua andam de mãos dadas com a mudança sintática (cf. FUß, 2016). Esse é o caso, por exemplo, dos estudos que explicam a mudança no licenciamento e na identificação de sujeitos nulos por meio dos aspectos morfofonológicos do paradigma flexional dos verbos finitos no PB (cf. GALVES, 1984, 1998; M. DUARTE, 1993 [2018], 1995; FIGUEIREDO SILVA, 1996; entre muitos outros). Como veremos mais adiante, essa não é, contudo, a direção em que conduziremos a análise desta pesquisa. Nossa hipótese de trabalho é sustentada pela suposição que dados de língua-E no período colonial teriam levado a uma mudança formal na gramática do PB em Goiás no século XIX. O gatilho dessa mudança teria sido a perda do traço-D de T, provocando uma queda na distribuição de sujeitos nulos; esse traço, contudo, não possui um expoente morfofonológico no paradigma flexional da língua, de modo que sua perda, num primeiro momento, não necessariamente provocaria uma alteração no paradigma flexional dos verbos finitos no PB. Retomaremos essa questão no Capítulo 6.

2.3 Síntese do capítulo

Neste capítulo, lidamos com três propriedades inerentes às línguas humanas, quais sejam: elas são inatas, variam e mudam com o tempo. Fundamentados em uma teoria que enxerga a capacidade de a espécie humana adquirir línguas como uma dotação genética – a teoria gerativa –, tratamos mais detidamente dos conceitos de variação e mudança. Vimos que, na abordagem de P&P, a variação translinguística esteve relacionada às idiossincrasias do léxico e às diferentes marcações paramétricas disponíveis no processo de aquisição de primeira língua. A mudança sintática, por outro lado, era vinculada a diferentes marcações paramétricas reanalisadas durante o processo de aquisição de primeira língua de uma nova geração. Mostramos que a noção de mudança sintática como objeto da gramática gerativa restringe a mudança a um conjunto de diferenças discretas entre a gramática alvo e a gramática adquirida pelo aprendiz. Argumentamos que a partir dessa perspectiva se pode

melhor compreender o funcionamento da aquisição de língua e descobrir restrições a possíveis mudanças impostas pelas propriedades da GU.

Considerando a abordagem do PM, mostramos que os pesquisadores gerativistas passaram a enxergar o *locus* da variação nas diferentes especificações dos itens lexicais e no local onde os traços formais desses itens lexicais são interpretados. Vimos que a mudança sintática é extremamente sistemática, em contraste com outros tipos de mudança, que são esporádicos e geralmente confinados a um único elemento lexical (cf. LEDGEWAY & ROBERTS, 2017; i.e., os artigos publicados na obra). Vimos também que a pesquisa diacrônica gerativista mais recente se firma no pressuposto de que a mudança gramatical ocorre devido a uma descontinuidade naturalmente provocada pela aquisição de língua por meio de dados língua-E. Quer dizer, atualmente, alguns estudiosos concordam com a proposta de que a mudança gramatical deve ser atribuída à descontinuidade, caracterizada por contingências impostas pelo contexto social. Esse pressuposto vincula a mudança gramatical e o processo de aquisição de língua e presumivelmente condiciona seus resultados ao imperativo epistemológico de que todas as especificações gramaticais das línguas humanas se adaptam às especificações da faculdade da linguagem. Nessa perspectiva, nem as línguas-E nem as línguas-I são transmitidas, imperfeita ou perfeitamente. Paralelamente, mostramos que o contato linguístico também tem um papel importante na mudança sintática, ocasionando mudanças estruturais mais sistemáticas, porém mais difíceis de serem detectadas.

3 TÓPICO DE PESQUISA: A SINTAXE DOS SUJEITOS NULOS

O objetivo deste capítulo é revisar alguns pontos fundamentais dos principais estudos gerativistas sobre a sintaxe dos sujeitos nulos e suas propriedades correlatas nas línguas naturais, apresentando discussões relevantes sobre essa postulação teórica por meio de evidências empíricas. Com isso, munimos nossa tese dos principais argumentos teóricos e as principais evidências empíricas em torno da sintaxe dos sujeitos nulos nas línguas naturais, contribuindo para uma melhor compreensão dos estudos gerativistas sobre a sintaxe dos sujeitos nulos no PB a serem apresentados no Capítulo 4, bem como para uma melhor compreensão da proposta de análise dos dados linguísticos desta tese a ser apresentada no Capítulo 6. Este capítulo estrutura-se da seguinte forma: a seção 3.1 explora a abordagem clássica do parâmetro *pro-drop* desenvolvida no âmbito da teoria de regência e ligação, bem como trata da correlação estabelecida entre morfologia verbal rica, movimento do verbo e sujeitos nulos nos estudos gerativistas. A seção 3.2 descreve os principais aspectos sintático-semânticos da tipologia das LSN no âmbito do PM (LSN expletivo, LSN orientadas para o discurso, LSN consistentes e LSN parciais). A seção 3.3 sintetiza o conteúdo do capítulo.

3.1 O parâmetro do sujeito nulo na teoria de regência e ligação

Na postulação do parâmetro do sujeito nulo no contexto da teoria de regência e ligação, estão em contraste: o EPP, que requer que em toda sentença a posição de sujeito seja sintaticamente preenchida por um NP/DP, um pronome referencial ou expletivo (CHOMSKY, 1982), como em (1) no inglês; e a observação empírica de que certas línguas omitem sistematicamente o sujeito de suas orações, como em (2).²⁷

- (1) Inglês
- | | | |
|----|---------------------|--------------|
| a. | <i>John sings</i> | <i>well.</i> |
| | João cantar-3SG-PRS | bem |
| | ‘João canta bem.’ | |
| b. | <i>He sings</i> | <i>well.</i> |
| | Ele cantar-3SG-PRS | bem |
| | ‘Ele canta bem.’ | |

²⁷ A primeira formulação do EPP, inicialmente chamado Princípio P, ocorreu em Chomsky (1981, p. 27) e se resumia basicamente ao requerimento estrutural de que certas configurações sintáticas deviam ter um sujeito.

c. *It is too late to come back.*

EXP ser-3SG-IMP-PRS muito tarde para voltar atrás

‘É muito tarde para voltar.’

d. *It rained.*

EXP chover-3SG-IMP-PASS

‘Choveu.’

(2) Italiano

a. *Giovanni canta bene.*

João cantar-3SG bem

‘João canta bem.’

b. ___ *Canta bene.*

(ele) cantar-3SG bem

‘Ele canta bem.’

c. ___ *È troppo tardi per tornare.*

EXP ser.3SG-PRS muito tarde para voltar

‘É muito tarde para voltar.’

d. ___ *È piovuto.*

EXP ser-3SG-IMP chover-3SG-IMP.PASS

‘Choveu.’

Antes mesmo da formulação do parâmetro do sujeito nulo por Chomsky (1981, 1982), estudos gerativistas pioneiros associavam a possibilidade de uma língua permitir sujeitos não realizados fonologicamente em orações finitas, entre outros fenômenos, a uma morfologia verbal rica que distingue pessoa e número (cf. PERLMUTTER, 1971; TARALDSEN, 1980): para esses estudos, a omissão do sujeito na sentença em (2b) ocorre devido à possibilidade de recuperá-lo pela morfologia do verbo, o que não é possível em seu contraexemplo em (1b), que precisa de um sujeito manifesto para satisfazer o EPP. Em outras palavras, o EPP nas línguas que permitem sujeitos nulos consistentemente, como o italiano, é satisfeito pelos traços formais de concordância presentes nas flexões do verbo. Além disso, essas línguas licenciam expletivos nulos (cf. RIZZI, 1982, 1986; CAMACHO, 2013), como ilustramos por meio dos contrastes em (1c-2c) e (1d-2d).

A título de ilustração, contrastamos a seguir o paradigma dos verbos finitos no presente do indicativo de línguas que apresentam morfologia flexional rica – distinguindo

todas as pessoas do discurso – e, portanto, licenciam sujeitos nulos (3), e o paradigma dos verbos finitos no presente do indicativo de línguas que apresentam morfologia verbal pobre, restringindo o licenciamento de sujeitos nulos nessas línguas (4). Observa-se em (4) que em inglês há apenas marca de concordância na 3ª pessoa do singular e que em francês o paradigma da 1ª e da 3ª pessoa do singular (*chante/chante*) e o paradigma verbal da 3ª pessoa do plural (*chantent*) do verbo ‘cantar’ (*chanter*) são fonologicamente indistintos. Tanto o inglês quanto o francês são, portanto, línguas com morfologia verbal pobre e de sujeito não nulo.

(3) Paradigma no presente do indicativo em PE/espanhol/italiano: *cantar/cantar/cantare*

Pessoa	Singular	Plural
1ª	<i>canto/canto/canto</i>	<i>cantamos/cantamos/cantiamo</i>
2ª	<i>cantas/cantas/canti</i>	<i>cantais/cantáis/cantate</i>
3ª	<i>canta/canta/canta</i>	<i>cantam/cantan/cantano</i>

(4) Paradigma do verbo *cantar* no presente do indicativo em inglês/francês (*sing/chanter*)

Pessoa	Singular	Plural
1ª	<i>sing/chante</i>	<i>sing/chantons</i>
2ª	<i>sing/chantes</i>	<i>sing/chantez</i>
3ª	<i>sings/chante</i>	<i>sing/chantent</i>

Para além da riqueza morfológica de línguas como em (3), capazes de recuperar a referência do sujeito por meio de suas terminações flexionais, Chomsky (1981) propôs que um conjunto de propriedades gramaticais específicas era geralmente encontrado em línguas com marcação positiva para o que ele inicialmente denominou de parâmetro *pro-drop*, a saber: (i) a omissão do sujeito da sentença, como em (5a); (ii) a inversão livre do sujeito em orações simples, como em (5b); (iii) o movimento longo do sujeito a partir de uma ilha *qu-*, como em (5c) – cabe destacar que as restrições de ilha representam configurações estruturais por meio das quais a extração/o movimento de certos elementos não deriva sentenças gramaticais (cf. CHOMSKY, 1986a); (iv) a presença de um pronome resumptivo/lembrete vazio em orações encaixadas, como em (5d); e (v) a aparente violação do filtro *that-trace* –

proposto por Chomsky & Lasnik (1977), segundo o qual o sujeito é impedido de se mover por cima de um complementador lexicalmente realizado –, como em (5e).²⁸

- (5) a. *Ho trovato il libro.* (Italiano)
 ter-1SG-AUX encontrado-PRF o livro
 ‘Encontrei o livro.’
- b. *Ha mangiato Giovanni.*
 ter.3SG-AUX comido-PRF João
 ‘O João comeu.’
- c. *L’uomo_i [che mi domando [chi t_i abbia visto]].*
 O=homem_i que me pergunto quem t_i havia-3SG-AUX visto-PRF
 ‘O homem que me pergunto quem tenha visto.’
- d. *Ecco la ragazza_i [che mi domando [chi crede [che t_i possa fare questo]]].*
 Aqui a garota que me pergunto quem acredita que t_i possa fazer isso
 ‘Eis a garota que eu me pergunto quem acredita que (ela) possa fazer isso.’
- e. *Chi_i credi [che t_i partirà].*
 Quem crê-2SG [que t_i partirá-3SG-FUT]
 ‘Quem você pensa que partirá?’ (CHOMSKY, 1981, p. 240)

Kayne (1980) foi quem primeiro indicou que as línguas que permitem a omissão do sujeito, como em (5a), também permitem a inversão da ordem sujeito-verbo (SV), como em (5b), por meio da constatação de que o italiano permite a ordem verbo-sujeito (VS), mas o inglês e o francês exigem que um pronome expletivo ou que um elemento *qu-* ocupe a posição pré-verbal do sujeito nos casos restritos de ordem VS, como ilustrado em (6) com um dado de inversão estilística no francês.

- (6) Francês – inversão estilística
Où te caches-tu?
 Onde REFL-3SG escondes=tu
 ‘Onde tu te escondes?’

²⁸ Em relação a (iii), Chomsky (1986b) distinguiu as restrições de ilha por meio dos conceitos ilhas fortes e fracas. Aquelas restringem o movimento de argumentos e adjuntos, e estas constituem configurações em que a extração de adjuntos é restrita, mas não a de argumentos.

Safir (1982), contudo, desfez essa correlação entre sujeitos nulos e a inversão livre do sujeito ao trazer para o debate evidência de dois dialetos italianos (trentino e modenese) em que sujeitos nulos não são permitidos, mas a inversão livre é permitida. Safir (1982) exemplifica que, assim como no francês, orações com inversão do sujeito clítico em trentino se restringem a orações matrizes (7) e a contextos em que o sujeito clítico ocupa a projeção mais alta do verbo, como (9), o que corrobora a teoria do autor de que sujeitos clíticos em orações com inversão são marcadores de concordância que aparecem em línguas cujo paradigma flexional dos verbos é fraco.

- (7) Ausência de inversão livre no trentino em comparação com o francês (SAFIR, 1982, p. 375)
- a. *Ho domanda se (el Mario) el magna a casa.*
 ‘Pergunto-me se Mario come em casa’.
- b. * *Ho domanda se (el Mario) magnelo a casa.*
- c. * *Je m’ai demandé si Mari mange-t-il chez lui.*
- (8) Ordem-SV em trentino em comparação com o francês (SAFIR, 1982, p. 375)
- a. *Alò magna?*
- b. *A-t-il mange?*
 ‘Ele comeu’.
- (9) Inversão livre no trentino em comparação com o francês (SAFIR, 1982, p. 375)
- a. *La magnelo?*
- b. *La mange-t-il?*
 ‘Ele come isso?’

Safir (1982) mostra ainda que sujeitos clíticos em orações com inversão em trentino não são licenciados simultaneamente com sua forma pré-verbal, de modo que (10a) só pode ser considerada gramatical se o pronome definido *el* for interpretado como um objeto clítico homófono.

- (10) Inversão livre no trentino em comparação com o francês (SAFIR, 1982, p. 375)
- a. *El magnelo?*
- b. **Il mange-t-il?*
 ‘Ele come?’

A despeito desses achados (cf. SAFIR, 1982, p. 373-393, para uma discussão mais detalhada), Rizzi (1982) propôs que, nos casos de inversão livre em LSN, a posição vazia de sujeito está coindexada com o NP e é legitimada por INFL, como exemplificado em (11).

- (11) Derivação da ordem VS no italiano (RIZZI, 1982, p. 132)
e_i INFL_i [VP [VP ha telefonato] Gianni_i]

Línguas como o inglês e o francês não geram construções como em (5c), tendo em vista que, considerando as restrições de ilha, o movimento longo do elemento *qu-* resulta agramatical nessas línguas. Essas línguas também não permitem que o sujeito de uma oração encaixada como em (5d) seja nulo. Como o sujeito nulo do verbo *possa* em (5d) não pode ser resultado de movimento de acordo com restrições de localidade impostas pelo princípio de subjacência²⁹, Chomsky (1981) propõe que ele é gerado na base como um pronome lembrete. (5e) aparentemente viola o conceito de filtro *that-trace*. Chomsky (1981), contrariando a generalização de Taraldsen (1980) de que os traços semânticos do sujeito nulo nos casos de aparente violação do filtro *that-trace*, como em (5e), são regidos (se o sujeito for anafórico) pela morfologia de concordância do verbo, argumenta que a sentença encaixada com o sujeito nulo nesses casos não viola o filtro *that-trace*, pois o movimento *qu-* realizado nesse casos provém de uma posição pós-verbal, como ilustrado em (12b), de modo que as línguas de sujeito nulo, de fato, obedecem ao filtro *that-trace*.

- (12) Italiano (CHOMSKY, 1981, p. 254)
- a. *Chi pensi che verrà?*
 Quem pensa-2SG que virá-3SG-FUT
 ‘Quem você pensa que virá?’
- b. *Pensi [Š che [S α verrà chi]]*
- c. *Pensi [Š che [S α chi verrà]]*

²⁹ O princípio de subjacência é definido conforme segue:

- (i) Nenhuma regra pode relacionar α e β na configuração
 ... α ... [Y ... [δ ... β ...] ...] ... α ...
 se Y e δ são nódulos/barreiras.

Cabe mencionar que barreira é uma categoria que serve como um bloqueio para a regência e/ou o movimento de outra (cf. CHOMSKY, 1986b).

Rizzi (1982) corrobora a proposta de Chomsky (1981) por meio da evidência de que no italiano quantificadores negativos licenciados como anáfora em orações encaixadas, como em (13b-c), não podem ser legitimados por INFL.

- (13) Italiano (RIZZI, 1982, p. 124-125)
- a. *non pretendo che tu arresti nessuno.*
 NEG espero que você prenda ninguém
 ‘Eu não espero que você prenda ninguém’
- b. *non pretendo che nessuno ti arresti.*
 NEG espero que ninguém você prenda
 ‘Eu não espero que ninguém prenda você.’
- c. [*non pretendo [che nessuno_i [e_i ti arresti]]]*
 [+ neg]

Em (13a), o objeto quantificador *nessuno* tem leitura existencial de escopo largo, porque, segundo Rizzi (1982), é licenciado por INFL. Se o mesmo fosse verdade para o sujeito encaixado em (13b), a interpretação de escopo largo deveria ser igualmente possível, mas não o é. Para Rizzi (1982), esse comportamento assimétrico sugere que a posição a partir da qual o sujeito encaixado deve ser extraído é pós-verbal, a mesma para os sujeitos *qu-* e para os quantificadores negativos. Nesse sentido, Rizzi (1982) propõe que os casos de aparente violação do filtro *that-trace* somente ocorrem em línguas de sujeito nulo porque essas línguas permitem a inversão livre do sujeito. Em outras palavras, de acordo com Rizzi (1982), se a extração ocorre através da inversão livre do sujeito, línguas sem essa inversão não podem ter extração. Seguindo as premissas de P&P, Rizzi (1982) foi quem primeiro propôs um modelo formal para o parâmetro do sujeito nulo que agrupasse propriedades sintático-semânticas desse fenômeno das línguas humanas.

- (14) Estatuto pronominal de INFL nas línguas de sujeito nulo (RIZZI, 1982, p. 143)
- a. INFL pode ser especificado [+ pronome]
- b. INFL [+pronome] pode ser referencial

Em (14a), Rizzi (1982) distingue as línguas de sujeito nulo das línguas de sujeito não nulo; em (14b), assinala uma discreta distinção entre línguas de sujeito nulo que permitem todo tipo de sujeito nulo (referencial e não referencial), como o italiano, e aquelas que

permitem somente sujeitos nulos não referenciais (isto é, sujeitos nulos não temáticos), como ocorre em variedades do holandês. Rizzi (1982) enfatiza o papel da riqueza morfológica na sintaxe dos sujeitos nulos ao postular que INFL tem estatuto pronominal. Com isso, Rizzi (1986, p. 519-524) reforça a existência de *pro*, um argumento pronominal fonologicamente nulo que ocupa a posição de sujeito em determinadas línguas e tem coindexada a ele a especificação gramatical dos traços do núcleo X que o licencia, como ilustrado em (15).³⁰

(15) *pro* na teoria de regência e ligação

- a. Seja X o núcleo licenciador de uma ocorrência de *pro*: então, *pro* tem as especificações gramaticais de X coindexada a ele.
- b. *pro* é regido por X_y^0 ;
- c. *pro* recebe marcação de Caso de X_y^0 .

Os postulados em (15) predizem que em uma dada língua um núcleo flexional pode diferir em relação à sua capacidade de licenciar e identificar *pro*. De acordo com Rizzi (1986), a *operação de licenciamento* controla as condições sintáticas dos constituintes não realizados e a *operação de identificação* regula o conteúdo semântico do argumento não realizado. Dessa forma, a interpretação de *pro* como referencial ou não referencial é legitimada pela forma com que *pro* é identificado, isto é, por meio dos traços formais presentes na morfologia de concordância em X_y^0 . Rizzi (1986) afirma que *pro* somente pode ser referencial se é licenciado por um núcleo flexional que identifique pessoa e número. Para os outros contextos de identificação, argumenta que *pro* quase argumental é licenciado por número, enquanto *pro* expletivo é licenciado pela ausência de qualquer marca de identificação de pessoa ou número.

Quadro 1: Traços de identificação de *pro*.

Interpretação	Pessoa	Número
<i>pro</i> referencial	+	+
<i>pro</i> quase argumental	-	+
<i>pro</i> expletivo	-	-

Fonte: Rizzi (1986).

³⁰ Rizzi (1986), na verdade, desenvolve um estudo sobre o objeto no italiano ser licenciado como *pro*. A hipótese de *pro* na posição de sujeito provém de sua análise sobre condições gerais de licenciamento de *pro* presente nesse estudo.

As propostas de Rizzi (1982, 1986) e de outros pesquisadores na fase de P&P foram em parte confirmadas pelo extenso estudo tipológico de Gilligan (1987). Em uma amostra composta por 100 línguas tipologicamente distintas, Gilligan (1987) identificou quatro propriedades relevantes para a caracterização das LSN, quais sejam:

- (16) a. a existência de sujeitos nulos temáticos/referenciais;
- b. a existência de sujeitos nulos não temáticos/não referenciais;
- c. a possibilidade de inversão livre da ordem SV; e
- d. a existência do filtro *that-trace* na extração de sujeitos de orações encaixadas.

Em sua expressiva amostra, Gilligan (1987) não encontrou nenhuma LSN que permitisse sujeitos nulos referenciais (16a), mas não permitisse sujeitos nulos não referenciais (16b). Isso confirma um dos achados de Rizzi (1982, 1986). Essa correlação, contudo, perde o respaldo com dados registrados no espanhol dominicano e no napolitano, duas LSN que expressam expletivos manifestos em detrimento de expletivos nulos, como nos exemplos a seguir em que o expletivo manifesto é obrigatório (cf. CAMACHO, 2013; LEDGEWAY, 2009).

- (17) Espanhol dominicano (CAMACHO, 2013, p. 43)

Ello hay personas que lo aprenden bien el inglés.

EXP há pessoas que o aprendem bem o inglês

‘Há pessoas que aprendem bem inglês.’

- (18) Napolitano (LEDGEWAY, 2009, p. 291-294)

a. *E chello mò so’ quase l’unnece.*

E EXPL agora são quase o-onze

‘E agora são/é quase onze.’

b. *Chello fino a mò è mancata ’a corrente.*

EXPL até a agora está perdida a eletricidade

‘Não há eletricidade até agora.’

Ainda a respeito de (12), Gilligan (1987) contesta a teoria de Rizzi (1982) de que a inversão livre do sujeito exerce um papel crucial na extração de sujeitos de orações encaixadas nas LSN. O pesquisador exhibe como contraargumento dados do papiamento,

língua nativa das ilhas caribenhas de Aruba, Curaçao e Bonaire, em que a inversão livre do sujeito não é permitida (19), ao passo que a extração do sujeito de uma sentença encaixada é possível (20).

(19) Papiamento (GILLIGAN, 1987, p. 141)

- a. *Maria a dispase.*
 Maria PRF desaparecer
 ‘Maria desapareceu.’
- b. * *a dispase Maria*

(20) Papiamento (GILLIGAN, 1987, p. 141)

- a. *ken_i bo kere ku e_i a bini.*
 quem você pensa que PRF vir
 ‘Quem você pensa que veio?’
- b. *e homber_i ku ta evidente ku e_i a bini.*
 o homem que PROG evidente que PRF vir
 ‘O homem que é evidente que veio.’

Gilligan (1987) arremata seu argumento afirmando que a única predição em relação a esses dois fenômenos é que se uma língua tem inversão livre do sujeito, ela deve permitir a extração do sujeito de uma oração encaixada, mas nunca o contrário. Gilligan (1987), assim como Huang (1982, 1984), contesta o papel da morfologia verbal no licenciamento e na identificação de sujeitos nulos, pois verifica em seu estudo a presença de sujeitos nulos referenciais em línguas que não apresentam riqueza morfológica, como o chinês.

Para Huang (1984), línguas como o chinês são orientadas para o discurso e, portanto, os sujeitos nulos nessas línguas podem ser contextualmente identificados, ao contrário das línguas orientadas para a sentença, que não legitimam a identificação contextual e, portanto, precisam de morfologia de concordância para identificar um sujeito referencial nulo. De acordo com Huang (1984), uma categoria vazia, em princípio, pode estar vinculada a um NP em uma posição A ou a uma variável discursivamente ligada a um tópico nulo, ou tópico zero, na posição inicial da sentença. Em outras palavras, Huang (1984), seguindo Li & Thompson (1976) e Tsao (1977), propõe uma teoria sobre argumentos nulos com base em dois parâmetros independentemente motivados. Um parâmetro distingue as línguas com tópico

nulo das línguas de tópico não nulo; o outro distingue as línguas *pro-drop* das línguas não *pro-drop*. Assim, de acordo com a proposta de Huang (1984), temos a seguinte tipologia: (i) línguas que não possuem tópico nulo nem sujeito nulo (inglês e francês); (ii) línguas que não permitem tópico nulo, mas permitem o sujeito nulo (italiano e espanhol); (iii) línguas que licenciam tanto tópico nulo quanto sujeito nulo (japonês e chinês); e (iv) línguas que licenciam tópico nulo, mas não licenciam sujeitos nulos (alemão).

Huang (1984) fornece evidências bem contundentes do comportamento predito para as línguas voltadas para a sentença e para aquelas voltadas para o discurso por meio, respectivamente, de dados do pashto, uma das línguas nacionais do Afeganistão, e do chinês. Como uma língua de ergatividade cindida, o pashto exhibe alinhamento nominativo-acusativo na conjugação dos verbos transitivos e intransitivos no tempo presente e alinhamento ergativo no tempo passado. Isso prediz que no presente do indicativo o verbo transitivo ou intransitivo exhibe concordância com o argumento externo, enquanto no pretérito o verbo transitivo exhibe concordância com o argumento interno e o verbo intransitivo concorda com o argumento interno. Essa predição é confirmada, respectivamente, em (21a-b) e em (22a-b), a seguir.

(21) Pashto (HUANG, 1984, p. 535)

a. *Jān ra-z-i.*

João DIR-vir-3-M.SG

‘João vem.’

b. *ze man a xwr-əm.*

eu maçãs comer-1-M.SG

‘Eu como maçãs.’

(22) Pashto (HUANG, 1984, p. 535)

a. *Jān ra-g -ay.*

João DIR-vir-3-M.SG

‘João veio.’

b. *ma mana wə-xwar-a.*

eu maçãs PRF-comer-3.F.SG

‘Eu comi as maçãs.’

No primeiro exemplo, tanto com o verbo intransitivo em (21a) quanto com o verbo transitivo em (21b), a morfologia verbal no tempo presente concorda com o sujeito. Isso

também ocorre com o verbo intransitivo no pretérito em (22a). Entretanto, em (22b), o argumento interno *as maçãs* desencadeia a concordância com o verbo, fazendo-o expressar seus traços semânticos de pessoa e gênero. Seguindo a generalização de Taraldsen (1980) sobre o papel da morfologia de concordância nas LSN, Huang (1984) prediz que somente os sujeitos de (21a-b), (22a) e o objeto de (22b), ou seja, os argumentos que desencadeiam a concordância com o verbo, podem ser omitidos em pashto, conforme ilustrado em (23) e (24).

(23) Pashto (HUANG, 1984, p. 536)

a. *e ra-z-i.*

DIR-vir-3SG

‘Ele vem.’

b. *e mana xwar-əm.*

maçãs comer-1-M.SG

‘Eu como as maçãs.’

(24) Pashto (HUANG, 1984, p. 536)

a. *e ra-g -ay.*

DIR-comer-3.M.SG

‘[Ele] veio’

b. *ma e wə-xwar-a.*

eu PRF-comer-3.F.SG

‘Eu comi as maçãs.’

Conforme apontado por Huang (1984), o argumento interno *as maçãs*, em (23b), e o sujeito *eu*, em (24b), não podem ser omitidos em pashto – o que é exemplificado em (25) a seguir.

(25) Pashto (HUANG, 1984, p. 536)

a. **zə e xwr-əm.*

Eu comer-1.M.SG

b. **e mana wə-xwar-a.*

maçã PRF-comer-3.F.SG

Em síntese, nas construções com verbo (transitivo ou intransitivo) conjugado no tempo presente, a morfologia flexional do verbo em pashto concorda com o sujeito, não com o objeto, daí o sujeito, não o objeto, poder ser nulo. Em contrapartida, nas construções com verbo transitivo no pretérito, a morfologia flexional do verbo em pashto concorda com o objeto, não com o sujeito, daí o objeto, não o sujeito, poder ser lexicalmente nulo (cf. HUANG, 1984). O pashto pode ser tomado, portanto, como uma forte evidência para as teorias que associam a morfologia verbal ao licenciamento e à identificação de argumentos nulos e para a proposta de Huang (1984) de distinção entre LSN orientada para a sentença e LSN orientada para o discurso. Huang (1984) caracteriza as LSN orientadas para o discurso como línguas em que se pode omitir um argumento na posição de sujeito em contextos discursivos apropriados e sem qualquer marca flexional no verbo para identificá-lo. É o caso de muitas línguas asiáticas, como o chinês, ilustrado em (26).

(26) Chinês

(HUANG, 1984, p. 533)

Falante A: *Zhangsan kanjian Lisi le ma?*

Zhangsan ver Lisi LE Q

‘Zhangsan viu Lisi?’

Falante B: a. *ta kanjian ta le.*

ele ver ele ASP

‘Ele o viu/Ele viu ele.’

b. *e kanjian ta le.*

[ele] ver ele ASP

c. *ta kanjian e le.*

ele ver [ele] ASP

d. *e kanjian e le.*

[ele] ver [ele] ASP

e. *wo cai [e kanjian e le].*

eu achar ver ASP

‘Eu acho [que] [ele] viu [ele]’

f. *Zhangsan shuo [e kanjian e le].*

Zhangsan dizer ver ASP

‘Zhangsan disse que [ele] viu [ele]’

Em (26), o falante B omite não somente o sujeito em (26a), mas também o objeto em (26b) e ambos em (26c) e (26d); (26e) mostra que toda uma oração subordinada pode ser omitida em chinês. Consoante a proposta de Huang (1984), o sujeito vazio em (26b) não pode ser pronominal porque um pronome nulo precisa ser identificado de acordo com o princípio da recuperabilidade (*Principle of Recoverability*), segundo o qual só se pode apagar da derivação material que seja redundante, e não há nada na sentença que possa identificar seu conteúdo. (26b), no entanto, pode ser uma variável vinculada a um tópico nulo. Desse modo, a sentença é gramatical com a categoria vazia interpretada como um tópico nulo. Em (26c), a categoria vazia não pode ser pronominal, pois, caso fosse, teria que ser coindexada com o elemento nominal mais próximo (*ta* ‘ele’), em violação à condição de referência disjunta (*Condition of Disjoint Reference*), segundo a qual um pronome deve ser livre na categoria de regência (HUANG, 1984, p. 552). Como resultado, (26c) é gramatical somente com a interpretação de tópico nulo do objeto nulo. (26d) é gramatical porque ambos os argumentos nulos são variáveis ligadas a tópicos nulos. De acordo com a proposta de Huang (1984), em todos os casos em que a categoria vazia é interpretada como uma variável ligada a um tópico nulo, a relação de ligação entre o tópico zero e a variável pode ser obtida por movimento sintático ou coindexação. Huang (1984) argumenta que todos os elementos omitidos em (26) são tópicos.

Huang (1984) postula, portanto, que, em línguas de tópico proeminente, os sujeitos estruturais não são um requisito básico da sentença, uma vez que o tópico é mais uma noção do discurso do que os sujeitos, que são uma noção sintática. Em outras palavras, nessa proposta, *pro* e tópico nulo têm pouco em comum: *pro* está sujeito a um princípio sintático, o que o obriga a ser coindexado com Agr ou um NP para a sua interpretação, enquanto um tópico nulo está sujeito a um princípio do discurso e não depende de concordância para a sua interpretação.

Para lidar com as LSN orientadas tanto para o sujeito quanto para o discurso em um único arcabouço teórico, Jaeggli & Safir (1989) formularam uma hipótese para as línguas de sujeito nulo com base no que eles denominaram de uniformidade morfológica:

- (27) Hipótese da uniformidade morfológica (JAEGGLI & SAFIR, 1989, p. 29)
 Sujeitos nulos são permitidos em todas (e somente n)as línguas com paradigmas flexionais morfológicamente uniformes.

(28) Uniformidade morfológica (JAEGGLI & SAFIR, 1989, p. 30)

Um paradigma flexional P em uma língua L é morfológicamente uniforme se e somente se P tem somente formas flexionais não derivadas ou somente formas flexionais derivadas.

Se são considerados os dados anteriormente apresentados neste capítulo, a hipótese de Jaeggli & Safir (1989) parece estar na direção correta: línguas com todas as formas flexionais derivadas, como o PE, o espanhol, o italiano e o pashto, licenciam sujeitos nulos temáticos da mesma forma que línguas com flexões não derivadas, como o chinês. (27) e (28) também predizem corretamente que línguas não uniformes morfológicamente, como o inglês e o francês, não permitem sujeitos nulos. Jaeggli & Safir (1989) reconhecem, contudo, que a hipótese em (27) parece não resistir a dados empíricos de línguas escandinavas continentais, como o islandês, que possui cinco paradigmas flexionais distintos, como no exemplo em (29), mas não licencia sujeitos nulos temáticos, apenas expletivos nulos. Em (29), nota-se que, no presente singular, a 1ª pessoa no sistema flexional islandês geralmente tem uma forma que a distingue da 2ª e 3ª pessoa. No passado singular indicativo, assim como no presente e no passado ambos no modo subjuntivo, a 2ª pessoa apresenta uma forma que a distingue da 1ª e da 3ª pessoa. No plural, entretanto, existem diferentes formas verbais para 1ª, 2ª e 3ª pessoas, tanto no tempo presente quanto no pretérito, quer seja indicativo, quer seja subjuntivo.

(29) PARADIGMA FLEXIONAL EM ISLANDÊS (PLATZACK, 1987, p. 386)

	Presente do indicativo 'tomar' 'dizer'		Presente do subjuntivo 'tomar' 'dizer'		Passado indicativo 'tomar' 'dizer'		Passado subjuntivo 'tomar' 'dizer'	
Singular	1 <i>tek</i>	<i>segi</i>	<i>taki</i>	<i>segi</i>	<i>tók</i>	<i>sagði</i>	<i>tæki</i>	<i>segði</i>
	2 <i>kastar</i>	<i>segir</i>	<i>takir</i>	<i>segir</i>	<i>tókst</i>	<i>sagðir</i>	<i>tækir</i>	<i>segðir</i>
	3 <i>tekur</i>	<i>segir</i>	<i>taki</i>	<i>segi</i>	<i>tók</i>	<i>sagði</i>	<i>tæki</i>	<i>segði</i>
Plural	1 <i>tökum</i>	<i>segjum</i>	<i>tökum</i>	<i>segjum</i>	<i>tökum</i>	<i>sögðum</i>	<i>tækjum</i>	<i>segðum</i>
	2 <i>takið</i>	<i>segið</i>	<i>takið</i>	<i>segið</i>	<i>tókuð</i>	<i>sögðuð</i>	<i>tækjuð</i>	<i>segðuð</i>
	3 <i>taka</i>	<i>segja</i>	<i>taki</i>	<i>segi</i>	<i>tóku</i>	<i>sögðu</i>	<i>tækju</i>	<i>segðu</i>

Os autores explicam que, sendo o islandês uma língua V2, essa língua tem T e AGR em nós separados. Já que em islandês T está em C, AGR não pode governar uma categoria nula, pois ela não poderia ser identificada. Em outros termos, os autores argumentam que em islandês, assim como no alemão, tempo e concordância são núcleos dissociados e, por isso, essas línguas não licenciam sujeitos nulos temáticos, mas, sim, expletivos nulos.³¹ Jaeggli & Safir (1989) lidam, ainda, com esse impasse propondo uma implicação unidirecional para a hipótese da uniformidade morfológica: se uma língua tem sujeitos nulos, então, ela é morfológicamente uniforme.

Verificamos que as análises sobre a sintaxe dos sujeitos nulos no modelo de P&P seguiram duas direções: de um lado, associaram a morfologia verbal ao licenciamento e à identificação dos sujeitos nulos por meio do argumento de que a própria morfologia de concordância representa o sujeito da oração, como é o caso do espanhol e do italiano; de outro lado, destacaram o papel do discurso na identificação de sujeitos nulos, especialmente para dar conta do licenciamento de argumentos nulos em línguas asiáticas que não possuem morfologia verbal capaz de recuperar a referencialidade dos argumentos elididos, como é o caso do japonês e do chinês. Jaeggli & Safir (1989) representaram uma tentativa de redirecionar essas duas perspectivas para um único caminho.

Em relação ao papel da morfologia na sintaxe, cabe destacar que, na transição de P&P para o PM, uma corrente de estudos que defende que a morfologia verbal repercute na sintaxe passa a fornecer análises interessantes no âmbito das assimetrias do movimento do verbo, especialmente no grupo das línguas germânicas e escandinavas, com implicações (in)diretas para a sintaxe dos sujeitos nulos. Essa vertente de estudos é aprimorada sob o escrutínio do que Bobaljik (2002) denominou de a hipótese da concordância rica (do inglês *The Rich Agreement Hypothesis*).

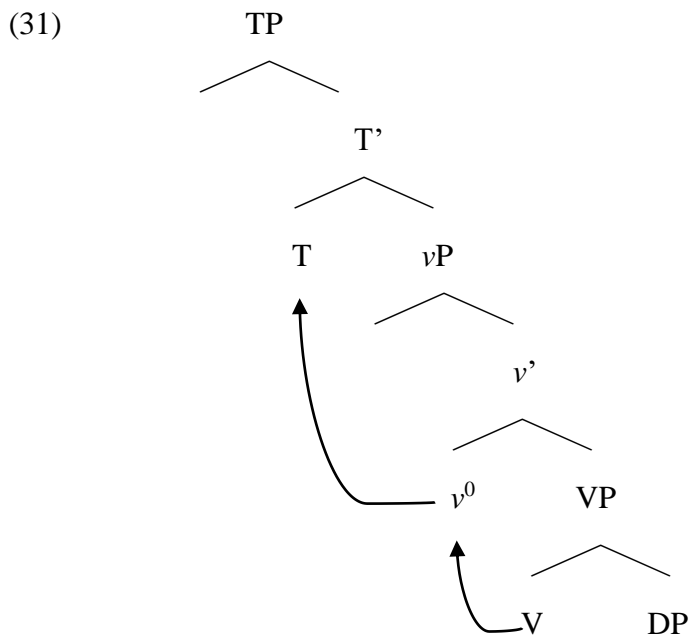
(30) Hipótese da concordância rica

Concordância verbal rica provoca movimento de V-para-T

Corolário: a perda da concordância rica implica a perda de movimento do verbo

³¹ Consoante com a observação de Jaeggli & Safir (1989), veremos na subseção 3.2.1 que nem o islandês nem o alemão têm expletivos nulos ou manifestos licenciados por T/AGR, mas por C (BIBERAUER, 2010).

Em geral, os estudos em torno da hipótese da concordância rica apontam que a riqueza das distinções no paradigma das flexões verbais se correlaciona com uma posição específica do verbo na oração, ou ainda com o licenciamento de sujeitos nulos (cf. ROHRBACHER, 1994, 1999). Mais especificadamente, esses estudos observam que, se o verbo em uma língua tem um paradigma flexional rico para estabelecer a concordância com o sujeito, o verbo nessa língua deve sempre aparecer à esquerda de advérbios como resultado do movimento de V-para-T (cf. EMONDS, 1978; POLLOCK, 1989; ROHRBACHER, 1994, 1999; VIKNER, 1997; BOBALJIK & THRÁINSSON, 1998; dentre outros). Desse modo, o movimento de V-para-T é a operação que resulta da correlação entre morfologia flexional rica e movimento do verbo.³² Na configuração em (31), representamos sucintamente o movimento de V-para-T.



Uma das abordagens pioneiras no escopo da hipótese em (30) é a de Rohrbacher (1994, 1999), que assegura que a concordância verbal rica e o movimento de V-para-T em uma dada língua têm implicações no licenciamento de sujeitos nulos. Para o autor, se e somente se os traços de 1ª e 2ª pessoas são distintamente marcados pela morfologia de concordância, os afixos de concordância são listados separadamente no léxico e projetam seus próprios núcleos na sintaxe, onde eles atraem o verbo para posições de especificador e

³² Ressalta-se que, além da morfologia de concordância, há uma variedade de razões pelas quais o verbo se move para T. O movimento V2, por exemplo, é gerado por um gatilho independente, cujo resultado é a colocação obrigatória do verbo na segunda posição da oração, independentemente da posição de outros constituintes. Isto é, o verbo nas construções V2 também aparece mais à esquerda, de modo que, se uma língua tiver concordância rica e V2, constitui tarefa do pesquisador explicar se o verbo que aparece mais à esquerda é resultado da concordância rica, de V2, ou de ambos os fenômenos.

permitem que as posições de especificador sejam preenchidas por vários elementos (não) realizados fonologicamente (ROHRBACHER, 1999).³³

Rohrbacher (1999) assume que todas as línguas têm um TP nucleado pelos traços abstratos de tempo, que são essenciais para a interpretação da sentença em LF e que, portanto, são sempre estruturalmente representados na sintaxe. As características abstratas de concordância entre sujeito e verbo, por sua vez, não são essenciais para a interpretação em LF e são, portanto, estruturalmente representadas na sintaxe apenas se o paradigma que as expressa for referencial e listado no léxico, por meio do qual núcleos sintáticos são projetados em razão de marcar distintamente os traços de concordância referencial de 1ª e 2ª pessoas. Para Rohrbacher (1999), se a distinção morfológica estiver presente em ao menos um tempo verbal, então o movimento de V-para-T ocorre sistematicamente em todos os tempos, independentemente dos outros tempos marcarem as distinções relevantes de pessoa. Rohrbacher (1999) argumenta que os traços de pessoa na morfologia do verbo repercutem de modo especial na sintaxe porque eles têm a habilidade especial de se referirem a entidades no discurso: *“os traços formais de pessoa por si só estabelecem se o sujeito se refere ao(s) falante(s), ao(s) destinatário(s) ou a outro(s). Outras características frequentemente expressas pela concordância entre sujeito-verbo [e.g. número e gênero] não têm essa habilidade.”* (ROHRBACHER, 1999, p. 128, tradução nossa).³⁴

A proposta de Rohrbacher (1999) prevê que apenas dois tipos de línguas podem existir dadas essas duas variáveis: um tipo no qual a concordância é pobre e na qual os adjuntos de VP não podem intervir entre o verbo e seu objeto direto; e um tipo no qual a concordância é rica e, portanto, sujeitos nulos são licenciados e os adjuntos de VP devem intervir entre o verbo e seu objeto direto. Para explicar a correlação que faz entre morfologia verbal rica, movimento de V-para-T e sujeitos nulos, Rohrbacher (1999) toma como exemplos diferenças entre as gramáticas do PE e do PB. O autor mostra, em (32), que os paradigmas flexionais do PE diferem quanto ao fato de que os traços-φ referenciais de 2ª pessoa são minimamente marcados, o que não ocorre no PB (ROHRBACHER, 1999, p. 222).

³³ Duas propostas alternativas à teoria de Rohrbacher (1994, 1999) são Vikner (1997) e Koenenman (2000). Vikner (1997, p. 207) diz que “uma língua SVO tem movimento de V-para-T se e somente se a morfologia de pessoa é encontrada em todos os tempos”. Koenenman (2000, p. 72) afirma que “os afixos em um paradigma são listados lexicalmente (e, portanto, forçam o movimento V-para-T) se as características de [1ª] e [2ª] e do [singular] forem cada uma delas contrastadas em algum lugar nesse paradigma”.

³⁴ No original: *“the person features by themselves establish whether the subject refers to the speaker(s), the addressee(s), or other(s). Other features often expressed by subject-verb agreement [e.g. number and gender] do not have this ability.”* (ROHRBACHER, 1999, p. 128, tradução nossa).

(32)	a. PE: comprar		b. PB: falar	
	SG	PL	SG	PL
	1 ^a compr- o	compr- amos	1 ^a fal- o	fal- a
	2 ^a compr- as	compr- am	2 ^a fal- a	fal- am
	3 ^a compr- a	compr- am	3 ^a fal- a	fal- am

A 2^a pessoa do plural *vós* (*comprais, falais*) perdeu-se em ambas as línguas, restando apenas a 2^a pessoa do plural *vocês*, cuja morfologia de flexão é indistinguível da 3^a pessoa do plural. Com base em Duarte (1993 [2018]), estudo pioneiro sobre a perda de sujeitos nulos no PB, Rohrbacher (1999) menciona que, ao contrário do PE, o PB também perdeu a distintamente marcada 2^a pessoa do singular *tu* (*compras, falas*), que foi substituída pela forma pronominal inovadora *você* (*compra, fala*) por volta do início do século XX, restando apenas a forma pronominal inovadora, que é morfologicamente indistinguível da 3^a pessoa do singular. O autor enfatiza, ainda, que o PB, assim como o francês, teria substituído a 1^a pessoa do plural (*nós falamos*) pela forma impessoal (*a gente fala*). Assim, se, conforme postulado por Rohrbacher (1999), o movimento de V-para-T depende de marcação morfológica distintiva mínima de todas as características de referência da 1^a e 2^a pessoas, prediz-se que o PE tem movimento do verbo para o núcleo flexional/temporal mais alto, enquanto o PB não o tem. Rohrbacher (1999) considera que suas predições parecem estar certas: a literatura gerativista na época já vinha atestando que o PE tem movimento de V-para-T enquanto o PB tem *V in situ* ou movimento do verbo para uma posição intermediária, mas crucialmente não mais alta que o núcleo flexional/temporal (cf. BIANCHI & FIGUEIREDO SILVA, 1993, *apud* ROHRBACHER, 1999).

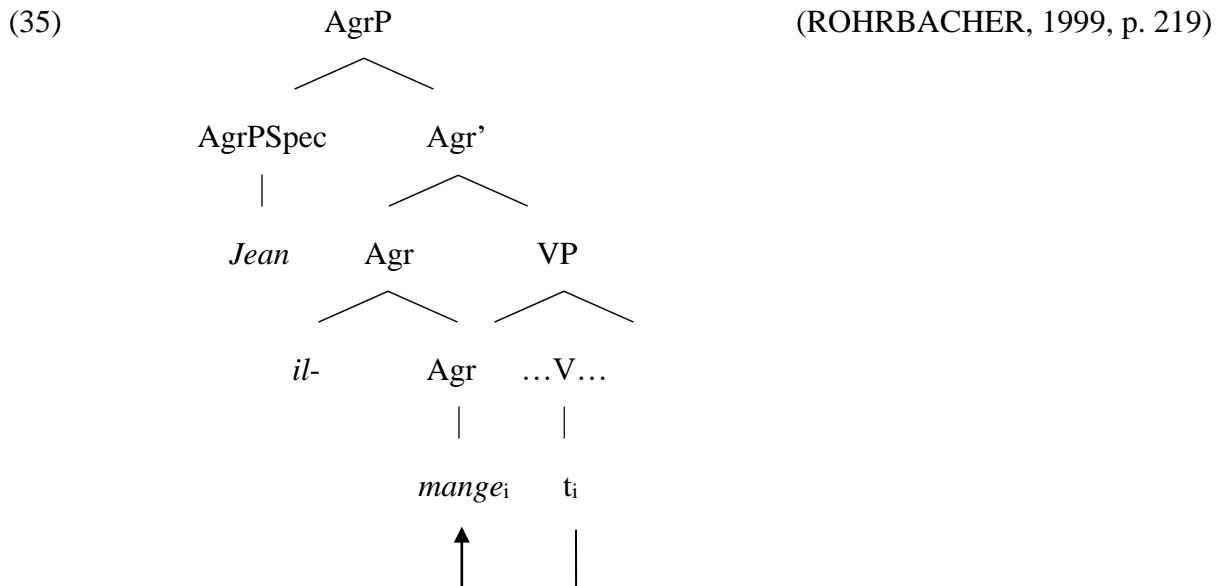
Conforme vimos em (4), o francês é considerada uma língua de morfologia pobre, porém, apresenta movimento de V-para-T, o que o torna um desafio para a proposta de Rohrbacher (1994, 1999).³⁵ A esse respeito, todavia, o pesquisador argumenta que dados de duplicação do sujeito no francês coloquial podem ser tomados como exemplos de concordância entre sujeito e verbo em vez de deslocamento de clítico à esquerda; isto é, para Rohrbacher (1999), clíticos sujeitos em francês, como em (33) a seguir, formam *um paradigma completo* (ROHRBACHER, 1999, p. 218). Rohrbacher (1999) afirma que a

³⁵ Línguas e dialetos escandinavos, como uma variedade do ferões, o dialeto sueco de Kronoby e o dialeto norueguês de Tromsø, também constituem contraexemplos à proposta de Rohrbacher (1999). Em estudo sobre variação nas línguas escandinavas, Thráinsson (2003) apresenta um detalhado estudo minimalista a esse respeito.

duplicação de clíticos é obrigatória com os pronomes em francês (34a) e que, no geral, os clíticos são encontrados em quase todas as orações finitas, inclusive com NP (34b).

(33)	a. Pronomes sujeitos no francês	b. Clíticos sujeitos no francês
	SG PL	SG PL
	1 ^a .p. <i>je on (nous)</i>	1 ^a .p. <i>moi nous</i>
	2 ^a .p. <i>tu vous</i>	2 ^a .p. <i>toi vous</i>
	3 ^a .p. MASC <i>il ils</i>	3 ^a .p. MASC <i>lui eux</i>
	FEM <i>elle elles</i>	FEM <i>elle “</i>
	IMP <i>ce/ça ces</i>	IMP <i>ça –</i>

- (34) a. *Lui il mange.* (Francês) (ROHRBACHER, 1999, p. 218)
 Ele CLITIC.3.MASC.SG *come*
 ‘Ele está comendo.’
- b. *Jean il mange.*
 João CLITIC.3.MASC.SG *come*
 ‘João está comendo.’



Diante desse conjunto de dados, Rohrbacher (1999) argumenta que é bastante improvável que o fenômeno dos clíticos sujeitos no francês coloquial possa ser reduzido a um simples deslocamento do pronome à esquerda, principalmente porque não requer ênfase ou

ruptura entonacional. Ele sugere, portanto, que os clíticos sujeitos funcionam como um paradigma flexional rico e, por isso, exibem movimento de V-para-T, como ilustrado em (35).

De acordo com Rohrbacher (1999), a morfologia flexional francesa é um sistema misto: parte dela, isto é, os clíticos, é lexical e, portanto, sintaticamente relevante e a outra parte dela, isto é, os sufixos verbais, é introduzida por PF e, portanto, sintaticamente irrelevante. Rohrbacher (1999) enfatiza que não há nada na teoria que refute a postulação de um sistema misto para a morfologia flexional. Uma questão, contudo, permanece: o autor não explica a ausência de sujeitos nulos no francês, tendo em vista que sua proposta teórica correlaciona concordância rica e movimento do verbo para T ao licenciamento de sujeitos nulos.

Uma vez que os estudos em Morfologia Distribuída postulam que a morfologia é um componente pós-sintático na arquitetura da gramática (cf. HALLE & MARANTZ, 1993; entre outros), a análise de Rohrbacher (1999), que propõe que estruturas sintáticas são conduzidas morfologicamente, é rejeitada por autores como Bobaljik e Thráinsson (1998) e Bobaljik (2002). Esses estudiosos defendem que as condições externas da variação morfológica entre as línguas são resultantes de uma variação sintática subjacente. “Uma vez que a variação sintática ocorre na ausência de variação morfológica evidente, essa não pode ser a única causa daquela” (BOBALJIK, 2002, p. 23, tradução nossa), tem-se que a morfologia não pode conduzir a sintaxe nos termos de Rohrbacher (1999).³⁶

Bobaljik e Thráinsson (1998) e Bobaljik (2002) sugerem que a generalização bidirecional não pode ser sustentada, pois o verbo também pode se mover para essa posição mais alta na ausência de concordância rica. Bobaljik e Thráinsson (1998) e Bobaljik (2002) resolvem esse impasse com o seguinte postulado:

(36) Hipótese da concordância rica (versão fraca)

Se uma língua tem uma morfologia suficientemente rica, então ela tem movimento do verbo.

Diferentemente da versão estrita da hipótese da concordância rica em (30), a generalização em (36) prediz apenas uma relação entre flexão rica e movimento de V-para-T. Isto é, ela não afirma que morfologia pobre corresponde à ausência de movimento do verbo ou que perda de morfologia rica provoca, diacronicamente, perda do movimento do verbo

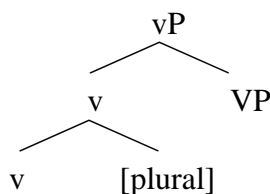
³⁶ No original: “*Since syntactic variation occurs in the absence of overt morphological variation, the latter cannot be the only cause of the former*” (BOBALJIK, 2002, p. 23).

para T. É importante notar, contudo, que ambas as abordagens da hipótese da concordância rica, a versão estrita e a versão fraca, preveem uma correlação unidirecional entre o movimento de V-para-T e a riqueza da morfologia de concordância.

Koeneman & Zeijlstra (2014) retomam o debate em torno da hipótese da concordância rica juntamente com a predição diacrônica de que a perda de morfologia verbal desencadeia a perda de movimento de V-para-T, tendo como principal argumento empírico a perda generalizada da morfologia de concordância no paradigma verbal finito na variedade padrão das línguas escandinavas ao mesmo passo em que também houve a perda do movimento de V-para-T. Koeneman & Zeijlstra (2014) fundamentam-se na aquisição de língua materna para explicar a perda diacrônica do movimento de V-para-T nessas línguas. Para esses autores, uma interação indireta entre os componentes sintático e morfológico ocorre durante a aquisição de língua e, embora a sintaxe preceda a morfologia, a morfologia modela a sintaxe por meio de *inputs* durante o processo de aquisição de língua. De acordo com essa análise, a bidirecionalidade da correlação entre concordância rica e movimento de V-para-T é refletida na sintaxe porque a ausência de movimento do verbo também é bidirecionalmente correlacionada com a ausência de uma projeção funcional que o desencadeie.

A análise bidirecional da hipótese da concordância rica proposta em Koeneman & Zeijlstra (2014) implica a proposição de que a distinção tipológica com relação à posição do verbo entre línguas com morfologia de concordância rica e pobre surge da presença ou ausência de traços distintivos no paradigma de flexões de concordância durante o processo de aquisição de primeira língua. Isto é, se não houver traços distintivos suficientes de concordância verbal no *input*, uma criança não pode postular projeções funcionais adicionais acima do sintagma verbal, dando origem a uma gramática como na configuração em (37), na qual não há posições para o verbo se mover (KOENEMAN & ZEIJLSTRA, 2014). Consequentemente, se a morfologia de concordância é fonologicamente dependente do verbo, aparecendo como um afixo no verbo, então a criança postula a presença de características formais no morfema de concordância que precisam ser verificadas pelo verbo.

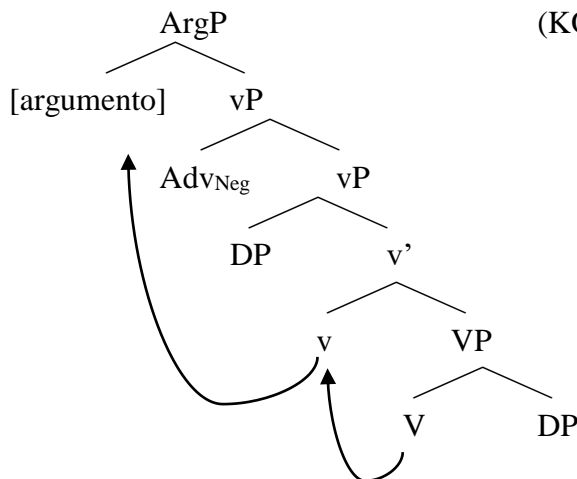
(37)



(KOENEMAN & ZEIJLSTRA, 2014, p. 597)

Em contrapartida a (37), Koeneman & Zeijlstra (2014) propõem a configuração em (38) a seguir e argumentam que uma projeção ArgP, na notação dos autores a projeção para a qual o verbo pode se mover, é esperada nas línguas que exibem um paradigma verbal suficientemente rico para refletir as seguintes diferenças semântica de pessoa e número: [\pm falante], [\pm ouvinte] e [\pm plural]. Koeneman & Zeijlstra (2014) argumentam que, somente quando as características semântica de pessoa e número: [\pm falante], [\pm ouvinte] e [\pm plural] são refletidas no paradigma verbal finito, as línguas têm uma projeção ArgP para a qual o verbo deve se mover para sustentar fonologicamente a morfologia de concordância.³⁷ Cabe destacar que, nessa visão, para Koeneman & Zeijlstra (2014), tanto o PB quanto o francês são línguas de concordância rica, pois agregam em sua morfologia verbal os traços distintivos de ([\pm falante], [\pm ouvinte] e [\pm plural]).

(38) (KOENEMAN & ZEIJLSTRA, 2014, p. 597)



Um contraexemplo à proposta bidirecional de Koeneman & Zeijlstra (2014) é apontado por Heycock & Sundquist (2017): os autores trazem evidência de um dos *corpora* da tese de Sundquist (2002), composto por cartas diplomáticas, cartas pessoais, sermões, memórias e excertos de livros escritos em dinamarquês, norueguês e sueco, que mostra que no início do dinamarquês moderno (1500-1700) o movimento de V-para-T era recorrente nos dados, apesar de dois séculos antes ter ocorrido a perda total da morfologia de concordância na língua. Na tese e em trabalho posterior, a questão central da proposta de Sundquist (2002, 2003) é a de que os dados linguísticos referentes ao início do dinamarquês moderno são um

³⁷ A partir de uma perspectiva de aquisição de língua, Koeneman & Zeijlstra (2014) propõem que se os traços distintivos mencionados ([\pm falante], [\pm ouvinte] e [\pm plural]) estão presentes no *input*, uma criança postula a projeção ArgP, que hospeda a morfologia da concordância.

sério desafio para qualquer hipótese que supõe uma relação direta entre mudança morfológica e mudança sintática na história de uma língua.

Com dados retirados de textos divididos em quatro períodos (1500–1550, 1550–1600, 1600–1650 e 1650–1700), Sundquist (2003) mostra que o movimento do verbo em dinamarquês médio diminui gradualmente em frequência da primeira para a segunda metade do século XVI (de 45% para 33%) e permanece opcional na segunda metade do século XVII (quando alguns textos ainda apresentam movimento do verbo em até 12% dos casos), mais de 250 anos após o esgotamento do paradigma verbal naquele idioma. Sundquist (2003) interpreta a opcionalidade de V-para-T como a competição entre duas gramáticas, uma com o movimento do verbo visível e a outra com movimento coberto do verbo. Sundquist (2003) afirma que, embora haja discordância entre estudiosos sobre o período da perda das distinções de número no tempo presente do dinamarquês médio, é consensual afirmar que as distinções morfológicas de pessoa desapareceram completamente no século XIV (cf. (39)).

(39)	Dinamarquês médio (1350): <i>dømae</i> ‘julgar’	(SUNDQUIST, 2003, p. 244)
	Presente	Passado
1 ^a p. sg.	<i>dømaer</i>	<i>dømdæ</i>
2 ^a p. sg.	<i>dømaer</i>	<i>dømdæ</i>
3 ^a p. sg.	<i>dømaer</i>	<i>dømdæ</i>
1 ^a p. pl.	<i>dømae</i>	<i>dømdæ</i>
2 ^a p. pl.	<i>dømae</i>	<i>dømdæ</i>
3 ^a p. pl.	<i>dømae</i>	<i>dømdæ</i>

Para uma proposta teórica baseada em paradigmas flexionais como a de Rohrbacher (1994, 1999), a distinção entre singular e plural não é suficiente para a definição de morfologia rica. O sincretismo para a morfologia de pessoa difundido no paradigma flexional do dinamarquês, como ilustrado em (39), todavia, apresenta-se como uma problemática para o postulado teórico de Rohrbacher (1994, 1999) e de outros similares ao dele, como o de Vikner (1997) e o de Koeneman (2000). Em sua análise da ordem de palavras em orações encaixadas no dinamarquês médio, Sundquist (2002, 2003) coletou principalmente orações que exibem dois tipos de ordem de palavras: no exemplo em (40a), dado que a negação é geralmente usada como um diagnóstico para o movimento de V-para-T, a negação sentencial sem itálico marca a periferia à esquerda do VP, e o verbo finito em negrito está numa projeção mais alta; no exemplo em (40b), o verbo finito permanece em VP (*in situ*), numa posição pós-adverbial.

- (40) Dinamarquês médio (SUNDQUIST, 2003, p. 236)
- a. *att wii kunde icke komme diid till dennom.*
 que nós poderíamos não ir lá para eles
 ‘que nós não poderíamos ir lá para eles’
- b. *at hun icke kunde leffue offuer en dag.*
 que ela não poderia viver mais um dia
 ‘que ela não poderia sobreviver outro dia’

Sundquist (2003) argumenta que uma alta frequência de padrões ambíguos de ordem de palavras, especialmente em orações encaixadas, é a principal responsável pelo declínio do movimento de V-para-T em dinamarquês. O autor mostra, em um grande conjunto de exemplos com o verbo aparentemente *in situ* em orações encaixadas com sujeitos pronominais, que é difícil determinar se tais dados são as primeiras evidências da ordem moderna das palavras no dinamarquês, ou se envolvem a cliticização do sujeito pronominal ao complementador e o chamado fronteamento estilístico (*Stylistic Fronting*) dos adverbiais, que, de acordo com o autor, provoca ambiguidade e parece depender mais de um autor particular do que de um período específico. Ele exemplifica esse impasse com os dados em (41), nos quais um sujeito pronominal é escrito junto com o complementador nos manuscritos.³⁸

- (41) Dinamarquês médio (SUNDQUIST, 2003, p. 238)
- a. *adttuu jicke glemmer adt fly thet salt.*
 que-você-CLITIC não esqueça de dar aquele sal
 ‘que você não se esqueça de dar esse sal.’
- b. *attj altid kunde retteligen bruge dem.*
 que-vocês-CLITIC sempre possa legalmente usar lhes
 ‘que todos vocês possam sempre usá-los legalmente.’

Sundquist (2003) informa que 69% (232/336) das orações encaixadas analisadas em seu *corpus* apresentam sujeitos pronominais. O autor, contudo, supõe que os dados em (41) não são exemplos inequívocos do surgimento da ordem das palavras nas orações encaixadas

³⁸ O fronteamento estilístico nas línguas escandinavas é tradicionalmente visto como o movimento de várias categorias para uma posição de sujeito vazia em orações finitas, geralmente o sujeito de orações relativas, mas também o sujeito de orações encaixadas em perguntas e orações completivas com uma lacuna na posição de sujeito (SUNDQUIST, 2003).

do dinamarquês moderno. Assim, ele argumenta que não teria sido a perda de paradigmas flexionais, mas a presença de padrões sintáticos ambíguos os responsáveis pela perda do movimento de V-para-T no dinamarquês, pois, embora não seja certo que essas orações envolvam cliticização, criando a lacuna do sujeito necessária para o fronteamto estilístico, não se pode descartar essa possibilidade à luz das evidências ortográficas (SUNDQUIST, 2003). Além disso, Sundquist (2003) sugere que as sentenças com a ordem [[pronome relativo]-[advérbio]-[verbo finito]] também são ambíguas e podem ser interpretadas estruturalmente como uma oração sem alçamento do verbo ou como fronteamto estilístico com um advérbio. Ele exemplifica isso com dados do dinamarquês moderno, como em (42), e argumenta que a frequência de fronteamto estilístico durante os séculos XVI e XVII e a alta frequência da ordem [[advérbio]-[verbo finito]] proporcionaram aos adquirentes de primeira língua a possibilidade de que essas orações pudessem ser interpretadas com o verbo *in situ*.

(42) Dinamarquês moderno (SUNDQUIST, 2003, p. 246)

- a. *som icke kan skriffuiss paa denne gang*
 que não pode escrito-ser em esse momento
 ‘que não pode ser escrito nesse momento.’
- b. *som alltijd handlede med min fader*
 que sempre negociou com meu pai
 ‘quem sempre negociou com meu pai.’

Sundquist (2003) afirma que, sob o espectro dos dados dinamarqueses, a versão fraca da hipótese da concordância rica se sai melhor do que a versão estrita, pois ela prediz que uma língua pode permanecer tendo movimento do verbo para T, obrigatório ou opcional, ao mesmo tempo em que carece de morfologia rica.³⁹ Sundquist (2003) parece reconhecer que certos fenômenos sintáticos podem estar associados com a morfologia. Se as línguas exibem sincronicamente correspondências estreitas entre morfologia e sintaxe, é então razoável que tais relações, sob certos aspectos, influenciem a mudança sintática. Todavia, uma ligação

³⁹ A proposta de Sundquist (2003) coincide com os estágios da mudança no dinamarquês de Haeberli (2004, p. 12-13), que também defende a versão fraca da hipótese da concordância rica:

- Estágio (i): concordância rica e gramática A na qual Agr é sintaticamente representada. A ordem XP-sujeito é licenciada devido à presença de AgrP.
- Estágio (ii): perda da concordância rica.
- Estágio (iii): competição entre a gramática A e a nova gramática B, na qual Agr é ausente. A gramática A ainda pode derivar a ordem XP-sujeito. A gramática B, no entanto, não gera a ordem XP-sujeito e fenômenos relacionados.
- Estágio (iv): perda da gramática A, possivelmente após vários séculos de competição.

direta entre a morfologia de concordância e a ordem das palavras, bem como o licenciamento de sujeitos nulos, conforme apontada na proposta de Rohrbacher (1999), não se sustenta.

É importante ressaltar que a possível ausência de sustentação de uma teoria como a de Rohrbacher (1999) não repercute nas análises minimalistas sobre sujeitos nulos em que os traços- ϕ nos verbos finitos são nominais o suficiente para satisfazer o traço-EPP (cf. P. BARBOSA, 1995; MANZINI & SAVOIA, 1997; ORDOÑEZ, 1997; ALEXIADOU & ANAGNOSTOPOULOU, 1998; KATO, 1999; entre outros). Nessas propostas, seguindo Chomsky (1995), o EPP é universal e consiste em um traço D em T; assim, nas línguas em que os traços- ϕ dos verbos finitos são suficientemente nominais, o movimento de V-para-T é meramente capaz de satisfazer o EPP, licenciando sentenças com sujeitos nulos e também com a ordem VS. Por outro lado, um argumento nominal ou um expletivo nominal deve ser inserido em [SPEC-TP] para satisfazer o traço-EPP nas demais línguas, ou seja, nas línguas sem traços- ϕ em T suficientemente nominais. Na seção 3.2.3 deste capítulo, trataremos dessa proposta para o licenciamento de sujeitos nulos em línguas com paradigma verbal rico, bem como discutiremos as análises que argumentam que os sujeitos nulos são pronomes deletados (HOLMBERG, 2005, 2010a; ROBERTS, 2010).

3.2 A sintaxe dos sujeitos nulos no Programa Minimalista

Línguas como inglês e francês requerem que o especificador da projeção flexional ou temporal seja preenchido por um NP/DP, embora outras formas, como a concatenação de um expletivo, também sejam possíveis.⁴⁰

- (43) a. **Is a linguist in the garden.* (Inglês)
 b. *There is a linguist in the garden.*
 ‘Há um linguista no jardim’.
 c. *A linguist is in the garden.*
 ‘Um linguista está no jardim’.

Essas são as línguas de sujeito não nulo, em oposição às LSN, que não requerem que o especificador do núcleo funcional flexional ou temporal seja preenchido e que têm sido estudadas dentro do PM de acordo com a seguinte proposta tipológica de Roberts &

⁴⁰ Conforme Biberauer (2010) menciona, parece que apenas línguas com o movimento obrigatório de um DP para [SPEC-TP] apresentam sistematicamente sujeitos expletivos manifestos, como exemplificado em (43b).

Holmberg (2010): LSN expletivo, LSN orientadas para o discurso, LSN consistentes e LSN parciais. Nas subseções seguintes, apresentamos, respectivamente, algumas das propriedades de cada um desses tipos de LSN, com o intuito principal de fornecer subsídios para explicar e refutar algumas das análises correntes sobre a sintaxe dos sujeitos nulos no PB (Capítulo 4), bem como para melhor estruturar os argumentos de análise dos dados da tese (Capítulo 6).

3.2.1 Línguas de sujeito nulo expletivo

A diferença crucial entre as LSN expletivo e as LSN consistentes é que aquelas apresentam a restrição de os sujeitos nulos terem de ser apenas não referenciais, enquanto nestas tanto sujeitos nulos referenciais quanto não referenciais são licenciados (cf. ROBERTS & HOLMBERG, 2010). Nessa direção, línguas como o holandês (44) e o alemão (45) licenciam sujeito nulo expletivo, teoricamente um pronome fraco,⁴¹ mas não licenciam sujeitos nulos referenciais. No dado alemão em (45a), o pronome expletivo *es* não pode ser expresso, enquanto em (45b) o mesmo pronome na mesma posição sintática, mas com interpretação referencial, deve ser manifesto.

(44) *gisteren werd (*er) door het hele dorp gedanst.* (Holandês)
 ontem foi EXPL pela a toda vizinhança dançado
 ‘Ontem, houve dança por toda a vizinhança.’ (GILLIGAN, 1987, p. 80)

(45) a. *Gestern wurde (*es) getanzt.* (Alemão)
 Ontem estava-3SG-IMP EXPL dançado
 ‘Ontem houve dança.’
 b. *Gestern war *(es) geschlossen.*
 Ontem estava-3SG-IMP (ele) fechado.
 ‘Ontem estava fechado.’ (CARDINALETTI, 1900, p. 5-6)

Huang (2000) distingue três comportamentos dentre aquelas que ele considera serem as LSN expletivo: (i) permitir apenas a omissão do expletivo não argumental, isto é, a omissão de expletivos que não têm papel- θ , de modo que argumentos de verbos meteorológicos devem ser explicitamente realizados, esse é o caso do alemão e do holandês;

⁴¹ Na linguística gerativa, assume-se geralmente que pronomes fortes e fracos ocupam uma posição XP, contudo, estes não são coordenáveis e não se restringem semanticamente a uma leitura humana (CARDINALETTI & STARKE, 1999).

(ii) permitir a omissão dos expletivos não argumentais e também dos chamados quase argumentais, ou seja, os argumentos dos verbos meteorológicos, esse é o caso do islandês e do ídiche; e (iii) permitir a omissão de expletivos não argumentais, quase argumentais e, sob certas circunstâncias restritas, sujeitos nulos referenciais, esse é o caso, segundo o autor, do finlandês, do hebraico e do ucraniano, por exemplo. De fato, Holmberg & Nikanne (2002) mostram que o finlandês apresenta sujeitos nulos expletivos, como em (46).

- (46) a. (*Sitä*) *On ilmennyt ongelmia.* (Finlandês) (HOLMBERG & NIKANNE, 2002)
 EXP tem surgido problemas
 ‘Tem surgido problemas’
- b. (*Sitä*) *Sattui onnettomuus.*
 EXP ocorreu (um) acidente
 ‘Ocorreu um acidente’
- c. (*Sitä*) *Tuli kiire.*
 EXP veio pressa
 ‘Nós/eles estamos/estão com pressa.’
- d. (*Sitä*) *Sataa vettä.*
 EXP chove água
 ‘Chove.’
- e. (*Se*) *Oli hauskaa että tulit käymään.*
 EXP foi legal que veio-2SG visitar
 ‘Foi legal que você veio visitar.’ (HOLMBERG & NIKANNE, 2002)

Holmberg & Nikanne (2002) argumentam que os expletivos *sitä* e *se* (nominativo como o *it* do inglês) são opcionais em finlandês, porque nessa língua há uma posição de tópico (FP) que licencia esses sujeitos e cujo traço EPP é opcional. A proposta dos autores tem como objetivo explicar não somente os expletivos obrigatórios no finlandês, mas também construções com tópico proeminente, no sentido de que o argumento externo não precisa ser o sujeito frásico, mas pode ser qualquer categoria que possa servir de tópico da sentença. Esses exemplos, enfatizam os autores, distinguem o finlandês das LSN canônicas, como o italiano.

Biberauer (2010) apresenta um extensivo estudo das LSN expletivo, especialmente do alemão e islandês, e enfatiza que as línguas com o comportamento citado por Huang (2000) em (iii) não podem ser consideradas um subtipo dentro das LSN expletivo. O finlandês, por exemplo, apesar de ter sujeitos nulos expletivos, tem sido amplamente caracterizado como

uma LSN parcial devido aos contextos restritos e sistemáticos em que licencia sujeitos nulos (cf. HOLMBERG, 2005, 2010a). Diante disso, Biberauer (2010) segue a proposta original de Rizzi (1986) de que as LSN expletivo são divididas em dois tipos: (a) aquelas que licenciam expletivos nulos quase argumentais e não argumentais; e (b) aquelas que licenciam expletivos nulos não argumentais, mas não licenciam expletivos nulos quase argumentais.

Em seu estudo, Biberauer (2010) propõe que são precisamente as especificações características de T e dos itens lexicais mais as categorias funcionais com as quais os itens lexicais e T interagem que determinam se um dado sistema gramatical exibirá ou não o que parece ser um sujeito nulo expletivo. Considerando Holmberg (2005, 2010a) e Roberts (2010), a proposta que a pesquisadora desenvolve centra-se em diferentes composições de T e em diferentes maneiras de satisfazer a exigência de EPP. Ela sugere, por exemplo, que em algumas dessas línguas [SPEC-TP] pode ser preenchido por sintagmas não expletivos como vP e, às vezes, pode até não ser projetado.

Biberauer (2010) atribui a variação no licenciamento de expletivos nulos e manifestos no alemão e no islandês a uma diferença nos inventários lexicais dessas línguas. Essa questão passa a ser ilustrada nos exemplos a seguir, em que o islandês parece licenciar *pro* quase argumental (47), e o alemão não (48).

(47) *I' gær* *rigndi* (*það) (Islandês) (BIBERAUER, 2010, p. 158)
ontem choveu EXPL
'Ontem choveu.'

(48) a. *Es* *schneit heute* (Alemão) (BIBERAUER, 2010, p. 159)
EXPL neva hoje
'Está nevando hoje.'

b. *Heute* *schneit *(es)*
hoje snows it
'Hoje está nevando.'

Contrastando os dados, Biberauer (2010) enfatiza que o alemão (48a-b) difere do islandês (47) ao exigir de forma consistente um expletivo manifesto com verbos meteorológicos. Por meio de dados como em (47) e (48), bem como por meio da análise das propriedades de T nas orações com sujeitos impessoais e nominais nessas línguas, Biberauer (2010) postula que no islandês há apenas um expletivo *það*, que é, nos termos da autora, um

expletivo verdadeiro e está vinculado a CP. Já no alemão há dois expletivos *es* distintos: um expletivo *es* quase argumental, como em (48), que é um expletivo manifesto fundido a Spec-vP; e um expletivo verdadeiro, como o expletivo nulo em (45a). Em ambas as línguas parece não haver um expletivo fundido em Spec-TP, e o expletivo verdadeiro em ambas, possivelmente codificando apenas pessoa (cf. CHOMSKY, 2000), é um tópico expletivo associado a CP devido às propriedades V2 dessas línguas (BIBERAUER, 2010).⁴² Biberauer (2010) mostra, com isso, que as LSN expletivo não formam uma classe unificada.⁴³

3.2.2 Línguas de sujeito nulo orientadas para o discurso⁴⁴

As LSN orientadas para o discurso, também conhecidas como línguas de sujeito nulo radical (i.e., *radical pro-drop languages*), como o chinês, o japonês e o coreano, licenciam sujeito nulo, sem, contudo, exibir uma gramática em que as terminações flexionais dos verbos recuperem qualquer pessoa do discurso. Nesse sentido, as LSN orientadas para o discurso diferem das LSN consistentes devido ao fato de o argumento omitido ser recuperado pelo discurso, e não pela gramática (HUANG, 1984). Tomioka (2003) associa a possibilidade de essas línguas licenciarem argumentos nulos (sujeito e objeto) à possibilidade de elas permitirem NPs nus robustos como argumentos, propondo a seguinte generalização.

(49) Generalização de Tomioka (2003) para as LSN orientadas para o discurso

As línguas que permitem sujeitos nulos discursivos também permitem NPs nus como argumento.

De acordo com Tomioka (2003), *pro* nas LSN orientadas para o discurso é um NP nulo, cujo conteúdo é pragmaticamente recuperado. Ele afirma que as mesmas ferramentas semânticas usadas para interpretar NPs plenos, como representado na configuração em (50b), são usadas para interpretar *pro* em línguas como o chinês e o japonês. Tomioka (2003) sugere, portanto, que o que constitui a base das LSN orientadas para o discurso é o fato de que as línguas naturais permitem, quase universalmente, NPs nus anafóricos. De acordo com

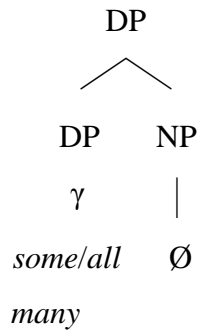
⁴² Biberauer (2010) também atesta a opcionalidade de expletivos verdadeiros em holandês e afrikaans.

⁴³ No Capítulo 4, retomaremos Biberauer (2010) quando apresentarmos a proposta de Naves & Pilati (2018) para a natureza dos sujeitos nulos genéricos e expletivos no PB.

⁴⁴ Toda a discussão teórica desenvolvida aqui em 3.2.2 restringe-se às línguas asiáticas orientadas para o discurso. Ressaltamos, contudo, que essa discussão servirá para melhor compreender as teorias que associam a elipse de DPs/NPs ao licenciamento de argumentos nulos de um modo em geral (cf. subseções seguintes deste capítulo) e as análises que associam a gramática do PB à orientação discursiva (cf. Capítulo 4).

Tomioka (2003), há, portanto, duas maneiras de derivar argumentos nulos: sob anáfora e sob elipse de NP nulo, como representado nas configurações em (50a) e (50b), respectivamente.

(50) a. Línguas-DP (e.g. inglês)



b. Línguas-NP (e.g. japonês) (elipse de NP)



(TOMIOKA, 2003, p. 336)

Sob esse esquema, Tomioka (2003) defende que os argumentos fonologicamente nulos e as elipses do NP têm o mesmo resultado em línguas com NPs como argumentos. A hipótese do autor, portanto, é de que se trata do mesmo fenômeno, ou seja: pronomes nulos nas LSN orientadas para o discurso são simplesmente o resultado da elipse de um NP. Nessa direção, o EPP de uma língua como o japonês requer somente um NP, ao invés do DP completo requerido na posição de sujeito nas línguas de sujeito não nulo. Se o núcleo funcional D é a posição onde os traços- ϕ e, portanto, pronomes manifestos e nulos são concatenados, então a ausência desse núcleo minimiza a condição de que sujeitos nulos devam ser licenciados por meio de concordância rica.⁴⁵

Contrários à generalização de Tomioka (2003), Neeleman & Szendrői (2007) propõem uma nova generalização que rege a distribuição de sujeitos nulos nas línguas *pro-drop* radical e rejeitam (49), com base nos seguintes argumentos: primeiramente, para os autores, não está claro porque somente D bloqueia *pro-drop* radical, haja vista que várias LSN desse tipo exigem que certos núcleos funcionais estejam presentes na projeção nominal estendida. Dois exemplos são o japonês e o coreano, cujos NPs devem estar acompanhados por uma partícula de caso, salvo raríssimas exceções. Para Neeleman & Szendrői (2007), caso a generalização de Tomioka (2003) estivesse certa, sujeitos nulos em japonês deveriam ser licenciados apenas

⁴⁵ Em conformidade com a generalização de Tomioka (2003), Saab (2016) e P. Barbosa (2018) desenvolvem diferentes versões da ideia de que os sujeitos nulos são elididos na derivação. Comparando dialetos do norte da Itália e (a mudança gramatical do) PB, Saab (2016) explora a ideia de que sujeitos nulos em LSN são instâncias de elipse morfológica. P. Barbosa (2018), por sua vez, estende a proposta de Tomioka (2003) para as LSN parciais, com o argumento de que as LSN parciais e as LSN orientadas para o discurso têm propriedades comuns que as diferenciam das LSN consistentes. A autora ilustra esse contraste comparando, principalmente, o PE e o PB. Nas subseções seguintes deste capítulo, apresentaremos mais detidamente essas propostas.

pela elisão do NP, como em (51a), em que o NP é elidido, mas uma partícula de caso nominativo permanece. A sentença só é gramatical se a partícula também for elidida (51b).

- (51) a. * \emptyset -ga subete-no hon-o yon-da. (Japonês)
 NOM todo-GEN livro-ACC ler-PASS
 b. \emptyset subete-no hon-o yon-da.
 todo-GEN livro-ACC ler-PASS
 ‘pro leu todo livro.’ (NEELEMAN & SZENDRÖI, 2007, p. 678)

Neeleman & Szendrői (2007) acrescentam que a generalização de Tomioka (2003) não está isenta de exceções, pois existem línguas que permitem sujeitos nulos na ausência de concordância e que exigem que NPs referenciais sejam acompanhados por determinantes. Nesse sentido, os autores citam a sintaxe de Cheke Holo, uma das línguas oceânicas falada nas Ilhas Salomão, para mostrar que essa língua permite que todo argumento seja omitido, ainda que o predicado não concorde com ele. Em (52), por exemplo, o sujeito pode ser descartado em (52a), embora nenhum marcador de concordância correspondente apareça no verbo; por outro lado, Cheke Holo tem determinantes, mesmo em nomes próprios, como mostra (52b).

- (52) a. Wasi gu \emptyset pohe are. (Cheke Holo)
 lavar ENF roupas aquelas
 ‘[Ela] lava as roupas.’
 b. Richard *(na) e tusu mei radio *(na) ka iara. (Cheke Holo)
 Richard ART PM entregar ir rádio ART para mim
 ‘Ricardo entregou o rádio para mim.’ (NEELEMAN & SZENDRÖI, 2007, p. 678)

Neeleman & Szendrői (2007) postulam, portanto, que sujeitos nulos radicais iguais aos encontrados no japonês e no chinês são licenciados não devido à morfologia flexional dos verbos, mas devido à morfologia dos pronomes. Os autores afirmam que esses sujeitos ocorrem apenas em línguas cujos pronomes são aglutinantes de Caso, número ou outro traço nominal. Nesse sentido, línguas de sujeito não nulo não podem omitir os pronomes livremente embora possam apresentar concordância *pro-drop*, ao passo que as demais LSN permitem o licenciamento de sujeitos nulos sob a presença de concordância rica.

- (53) Generalização de Neeleman e Szendrői (2007) para *pro-drop* radical
Pro-drop radical requer aglutinação morfológica de pronomes.

Como visto em (53), os autores sustentam que uma língua pode apagar seus pronomes se ela tiver pelo menos alguma morfologia pronominal aglutinante. Na ausência de tal morfologia, *pro-drop* é bloqueado (NEELEMAN & SZENDRÖI, 2007). Os autores mostram que, em japonês, a morfologia de casos é aglutinante: o radical pronominal *kare* em (54a) é acompanhado por morfemas de casos separados (-*ga* e -*o*). Já no chinês há um marcador plural -*men*, que se liga a radicais pronominais como *ta* ‘ele’ em (54b).

- (54) a. *Kare-ga kare-o settokusuru.* (Japonês)
 ele-NOM ele-ACC persuade
 ‘Ele persuade ele.’
 b. *Ta-men kanjian ta le.* (Chinês)
 he-PL ver he le
 ‘Eles viram ele.’ (NEELEMAN & SZENDRÖI, 2007, p. 679)

Neeleman & Szendrői (2007) explicam que a correlação entre a morfologia dos pronomes e a disponibilidade de *pro* em *pro-drop* radical pode ser derivada de três operações independentemente motivadas. Simplificadamente essa suposição pode ser assim resumida: primeiramente, Neeleman e Szendrői (2007) tomam os argumentos nulos como pronomes regulares não pronunciados na PF, em vez de instanciações de um item lexical silencioso especial (*pro*). Em segundo lugar, argumentam que as regras de *Spell-out* para os pronomes podem ter como alvo nós não terminais, assim como nós terminais. Por fim, sugerem que uma potencial competição entre as diferentes regras de *Spell-out* é regulada por condições como o *Elsewhere Principle* (cf. KIPARSKY, 1973).⁴⁶

Sato & Kim (2012) fornecem evidências do inglês coloquial de Singapura, língua de base lexical inglesa resultante do contato linguístico dessa língua com línguas siníticas, como o mandarim e o cantonês, como contraexemplo à generalização de Neeleman e Szendrői (2007) de que *pro-drop* radical requer morfologia aglutinante nos pronomes pessoais. Sato & Kim (2012) afirmam que o inglês coloquial de Singapura exhibe padrão de língua *pro-drop* radical, mas seu paradigma pronominal é claramente fusional como no inglês padrão. Os

⁴⁶ Sugerimos a leitura de Neeleman & Szendrői (2007) para uma melhor compreensão dessas motivações.

autores concordam com Huang (1984) em sua proposta de que a característica central de uma língua de tópico proeminente (*pro-drop* radical) é o licenciamento de *pro-drop* e tópico nulo. Citam como característica da sintaxe do sujeito no inglês coloquial de Singapura a ausência de pronomes expletivos, o que é ilustrado pela construção existencial em (55a), em contraste com a construção expletiva do inglês padrão em (55b).

- (55) a. *Got people in the classroom.* (SATO & KIM, 2012, p. 865)
 b. *There are people in the classroom.*
 ‘Há pessoas na sala de aula.’

Ademais, Sato & Kim (2012) mostram que o inglês coloquial de Singapura permite construções de tópico nulo em cadeia, como ilustrado em (56).

- (56) A: *So, you can cycle now, can you?* (SATO & KIM, 2012, p. 865)
 B: *Yeah, [Ø]_{Top} can cycle, not very well, but [Ø]_{Top} can cycle, ah. [Ø]_{Top} knocked myself against the pillar, but then [Ø]_{Top} managed to pick up cycling.*
 ‘A: Então, você pode pedalar agora, você pode?
 B: Yeah, [Ø]_{Top} posso pedalar, não muito bem, mas [Ø]_{Top} posso pedalar, ah. [Ø]_{Top} Bati-me contra o pilar, mas depois [Ø]_{Top} consegui pedalar.’

Em (56), a pergunta do falante A apresenta o falante B como o tópico do discurso (SATO & KIM, 2012). Todas as orações subsequentes na resposta do falante B se referem à mesma entidade, que é o tópico já estabelecido pela pergunta do falante A, e, por isso, o deletam fonologicamente, formando o que Huang (1984) chama de cadeia de tópicos, na qual os tópicos vazios são definidos pelo tópico original (SATO & KIM, 2012). O inglês coloquial de Singapura também apresenta construções com tópico fonologicamente marcado no início da sentença, como em (57), bem como exibe assimetrias de argumentos nulos na posição de sujeito e de objeto, como ilustrado em (58), conformando-se à proposta de Huang (1984) de que uma língua de tópico proeminente exibe duas propriedades essenciais: (i) livre omissão de argumentos e (ii) assimetrias na referência de argumentos nulos nas posições de sujeito e de objeto.

- (57) a. *Australia, I've been to Perth.* (SATO & KIM, 2012, p. 866)
 ‘Austrália, eu estive em Perth.’

b. *Local food, you must try chicken rice.*

‘Comida local, você deve provar o arroz com frango’.

- (58) a. [...] [*Sar Che and Sar Ee*]_i said [*e*_{i/j} *go to the airport to fetch you and Li Sa*].
 b. [...] then [*my mum*]_i said [*e*_{i/j} *must call her sister*].
 c. [...] *they*_i said [*they will send e*_{*i/j} *to you*]. (SATO & KIM, 2012, p. 866)

O dado em (58a) é citado em um diálogo de uma peça de teatro de Singapura, e o antecedente do sujeito vazio, nesse exemplo, pode ser o sujeito da oração matriz (*Sar Che and Sar Ee*), ou outra pessoa no discurso (SATO & KIM, 2012). A mesma operação vale para o sujeito vazio em (58). O objeto nulo em (58c), por outro lado, não pode tomar o pronome *they* da oração matriz como seu antecedente, mas deve referir-se a alguém compreendido no discurso. Essa assimetria interpretativa, presente no inglês coloquial de Singapura, é exatamente o que Huang (1984) observou no chinês (SATO & KIM, 2012).

Com base nos dados do inglês coloquial de Singapura, Sato & Kim (2012) argumentam que não há como negar que essa língua emergente do contato linguístico é de proeminência de tópico, o que não é de surpreender dadas as influências sistemáticas sobre a gramática dessa língua, oriundas de línguas siníticas com tópico proeminente, como, por exemplo, o cantonês e mandarim. Os autores evidenciam, contudo, que, nos casos em que o verbo concorda morfologicamente com o sujeito no inglês coloquial de Singapura (mais restritamente, no tempo presente da 3ª pessoa do singular, como em (60)), sujeitos nulos não são permitidos.

(59) Falante A: *Zhangsan see Lisi already?* (Inglês coloquial de Singapura)

‘*Did Zhangsan see Lisi?*’ (Inglês padrão)

‘*Zhangsan já viu a Lisi?*’

Falante B: a. *e see him already.*

[Ele] viu ele já

b. *He see e already.*

Ele viu [ele] já

c. *e see e already.*

[Ele] viu [ele] já

d. *Lisi_i say [e_{i/j} see him already].*

Lisi_i disse [ele]_{i/j} viu ele já

e. *Lisi_i say [he see e*_{i/k} already].*

Lisi_i disse he viu [ele]*_{i/j}. (SATO & KIM, 2012, p. 868)

(60) Falante A: *Zhangsan meets Xu all the time or not?* (Inglês coloquial de Singapura)

‘*Does Zhangsan meet Xu all the time?*’ (Inglês padrão)

‘Zhangsan encontra-se com Xu o tempo todo ou não?’

Falante B: a. **e meets him all the time!*

[Ele] encontra ele todo o tempo

b. *He meets e all the time!*

Ele encontra [ele] todo o tempo

c. **e meets e all the time!*

[Ele] encontra [ele] todo o tempo

d. * *Lisi_i say [e_{i/j} meets him all the time!].*

Lisi_i disse [ele]_{i/j} encontra o todo o tempo

e. *Lisi_i say [he meets e*_{i/k} all the time!].*

Lisi_i diz ele encontra [o]*_{i/j} todo o tempo (SATO & KIM, 2012, p. 868)

Como se pode verificar em (59), a categoria vazia do inglês coloquial de Singapura comporta-se como o chinês na ausência total de morfologia verbal (cf. os dados em (26) neste capítulo, p. 60). Por outro lado, em (60a-e), enquanto a categoria vazia de objeto no inglês coloquial de Singapura permanece intacta na presença de concordância entre sujeito e verbo na 3ª pessoa do singular do tempo presente, uma categoria vazia na posição de sujeito é sistematicamente restrita na presença de concordância entre sujeito e verbo. Para lidar com esse impasse, os autores integram a proposta de Huang (1984) ao trabalho de Speas (1994, 2006), que afirma que a possibilidade de argumentos nulos é bloqueada por concordância escassa (*meager agreement*). Speas (1994, 2006) propõe que, quando T contém traços- ϕ que não possuem uma especificação completa, seus traços devem ser valorados por meio da relação especificador-núcleo. Assim, como não possuem traços- ϕ , as línguas com concordância escassa não permitem a inserção de *pro* na posição do sujeito quando algum traço de concordância é acionado. Por outro lado, quando uma língua tem uma concordância rica em T, que é totalmente especificado, a língua permite sujeito nulo. E, por fim, quando uma língua não tem nenhuma concordância, não há nada que exclua a inserção do *pro* na posição do sujeito (cf. SATO & KIM, 2012; SPEAS, 1994, 2006).

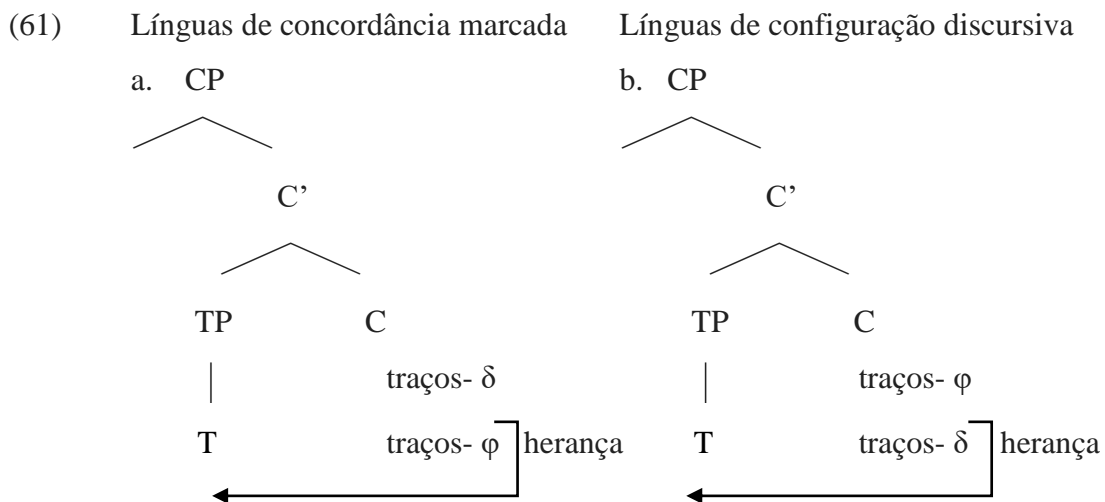
Como resultado da integração da proposta de Huang (1984) e de Speas (1994, 2006), Sato & Kim (2012) argumentam que a presença de concordância escassa no inglês coloquial de Singapura, como ilustrado em (60a) e (60c), bloqueia a possível omissão do sujeito. Mais precisamente, os autores postulam que em uma língua de proeminência de tópico com concordância escassa opcional, como é o caso do inglês coloquial de Singapura para a 3ª pessoa do singular, *pro* é descartado em duas situações: (i) quando seu conteúdo não é recuperável por meio da concordância escassa e (ii) quando ele não pode fornecer os traços formais para o estabelecimento da concordância escassa. Em contrapartida, um tópico nulo é descartado, não porque seu conteúdo não pode ser recuperado por meio da concordância escassa, mas porque não pode fornecer os traços- ϕ necessários para serem valorados por meio de concordância escassa (SATO & KIM, 2012). Para as demais línguas orientadas para o discurso, isto é, sem morfologia escassa, a proposta de Huang (1984), com as devidas adaptações no uso dos dispositivos teóricos impostas pela abordagem minimalista, parece ser ainda a teoria mais adequada na visão de Sato & Kim (2012) para explicar *pro-drop* em línguas como o chinês e o japonês, ainda que essas línguas apresentem padrões bem distintos no licenciamento e na identificação de seus argumentos nulos de um modo geral.⁴⁷

Miyagawa (2010, 2017) apresenta uma teoria sobre concordância nas línguas naturais em que as características de tópico e foco desempenham o papel dos traços- ϕ no discurso configuracional de línguas como o chinês e o japonês. Ele argumenta que toda língua manifesta algum tipo de concordância, algumas na forma do sistema familiar de pessoa, número e gênero, e outras na forma dos chamados traços discursivos configuracionais, como tópico e foco. O cerne de sua teoria é o de que o sistema computacional da sintaxe lida com uma gama de tipos de concordância de maneira uniforme, como se houvesse apenas um sistema, e uma parte integrante desse cálculo acaba sendo movimento. Sob o escopo de sua teoria da Uniformidade Forte (*Strong Uniformity*), segundo a qual todas as línguas compartilham o mesmo conjunto de características gramaticais e todas as línguas manifestam abertamente essas características, Miyagawa (2010) defende que as línguas com e sem concordância morfológica marcada são geradas por um sistema único e unificadas sob a perspectiva de um conjunto de traços gramaticais que inclui tanto os traços- ϕ quanto os

⁴⁷ Sato & Kim (2012) afirmam que sua análise se assemelha à proposta de Saito (2007), que argumenta que *pro* em *pro-drop* radical é introduzido na derivação por cópia em LF. Esse processo, no entanto, é bloqueado quando *pro* é inserido em uma posição que entra em concordância com núcleos funcionais, como T e *v*, que precisam ter seus traços- ϕ não interpretáveis valorados. Nesse sentido, em japonês, *pro* é copiado em qualquer posição sintática porque essa língua não tem concordância sintática (SAITO, 2007). A semelhança nos postulados dos autores está mais precisamente no que concerne à observação de que sujeitos nulos em línguas de tópico proeminente são permitidos somente quando a concordância é totalmente ausente, afinal, o tópico nulo é permitido quando T carece de traços- ϕ , não sendo licenciado quando T possui esses traços.

traços configuracionais do discurso. A hipótese da Uniformidade Forte de Miyagawa (2010) é alicerçada no Princípio de Uniformidade (*Uniformity Principle*) de Chomsky (2001), que diz que, na ausência de evidência convincente do contrário, assume-se que as línguas são uniformes, com variedade restrita a propriedades facilmente detectáveis nos enunciados.

Mais recentemente, Miyagawa (2017) estende sua teoria da Uniformidade Forte e propõe uma distinção de notação entre traços- ϕ , pessoa, número e gênero, e traços- δ , os traços configuracionais do discurso, como tópico e foco. A questão crucial da teoria de Miyagawa (2017) continua sendo a de que a concordância nas línguas é gerada por um sistema único e unificadas sob a perspectiva de um conjunto de traços gramaticais (os traços- ϕ e os traços- δ). Essa proposta teórica prediz duas configurações para a concordância nas línguas por meio da herança de traços de C por T, conforme ilustrado em (61) (MIYAGAWA, 2017, p. 4).



Em adição, a proposta de Miyagawa prevê quatro tipos distintos de línguas: (i) línguas com traços- ϕ em C e traços- δ em T; (ii) línguas com traços- δ em C e traços- ϕ em T; (iii) línguas tanto com traços- ϕ quanto com traços- δ em T; e (iv) línguas tanto com traços- ϕ quanto com traços- δ em C. A seguir, apresentamos a categorização tipológica dessas línguas conforme ela foi ilustrada pelo autor:

- (62) Línguas previstas pela hipótese da Uniformidade Forte de Miyagawa (2017, p. 4)
- Categoria I: C ϕ , T δ – Japonês
- Categoria II: C δ , T ϕ – Inglês
- Categoria III: C, T ϕ/δ – Espanhol
- Categoria IV: C ϕ/δ , T – Dinca (língua nilo-saariana)

A partir de (62), Miyagawa (2017) analisa a concordância envolvendo traços- ϕ e traços- δ por meio do estudo de vários fenômenos sintáticos, como a concordância alocutiva, a topicalização, o parâmetro *pro-drop*, perguntas “*por que*”, foco, alternância de caso etc. Em relação especificamente a *pro-drop* radical, o autor desenvolve uma sofisticada análise do licenciamento e da identificação de argumentos nulos de acordo com os traços ϕ e δ no chinês e no japonês, lançando um olhar mais acurado sobre a natureza de *pro* nessas línguas.

Miyagawa (2017) sugere que o sujeito nulo em chinês é um pronome fraco, e não o resultado da elipse de argumento, como é amplamente aceito na literatura. Ele argumenta que em chinês a topicalização do sujeito *pro* ocorre quando *pro* não recebe os traços- ϕ de T (destaca-se que, nessa perspectiva, o chinês possui traços- ϕ), enquanto *pro* em línguas românicas é topicalizado como parte das propriedades de concordância e de movimento dessas línguas. Para o autor, em japonês e nas línguas românicas, o sujeito *pro* comporta-se como um pronome lexical devido à sua capacidade de poder referir-se livremente a entidades no discurso dentro de um contexto razoável e também em remeter a um sujeito, a um objeto ou a outros sintagmas dentro de uma estrutura sentencial. Por sua vez, em chinês, o sujeito *pro* é limitado em seu potencial de referência: é capaz de se referir a uma entidade discursiva em contextos muito restritos, e internamente seu antecedente é limitado ao sujeito. O autor argumenta que o sujeito *pro* em chinês atua de acordo com a Categoria II de (62), dependendo dos traços- ϕ de concordância para referência interna na sentença, e, quando essa opção não é escolhida, muda para os traços- δ para se referir a uma entidade discursiva.

Miyagawa (2017) mostra que, quando *pro* em chinês se refere a uma entidade do discurso, o antecedente não precisa ser um sujeito. Assim, dado que *pro* em (63b), por exemplo, está em uma posição de tópico, é natural que ele procure por um tópico como seu referente, de modo que o antecedente para *pro* em (63b) deve ser *Mali* de (63a), gerando uma leitura *sloppy* de *pro* – leitura pouco esperada no chinês (cf. MIYAGAWA, 2017).⁴⁸

- (63) a. *Mali, Zhangsan hen xihuan ta.* (Chinês) (MIYAGAWA, 2017, p. 73)
 Maria Zhangsan muito gosta her
 ‘A Maria, Zhangsan gosta muito dela.’
- b. *Danshi, Lisi shuo [pro yijing jiehun le].*
 Mas Lisi disse já casar PRF
 ‘Mas, Lisi disse que [*pro* = Maria] está pronta para casar-se.’

⁴⁸ Na linguística gerativa, uma leitura *sloopy* consiste em um argumento nulo em um VP elidido poder ser interpretado como tendo um referente diferente daquele do VP antecedente, e uma leitura *strict* consiste em um argumento elidido em um VP elidido denotar o referente do VP antecedente.

Em relação ao sujeito *pro* em chinês ser um pronome fraco, Miyagawa (2018) fornece como evidência da presença de traços- ϕ de concordância em chinês o bloqueio de ligação de anáforas, ilustrado em (64). Miyagawa (2018) adota uma abordagem para o bloqueio de ligação de anáforas fundamentada na ideia de que a ligação de longa distância de anáforas é implementada pelo movimento coberto da anáfora para o núcleo cujo especificador contém o sujeito da oração, ou seja, um antecedente em potencial. Isso baseia-se na suposição de que a anáfora é subespecificada para algum traço, mais precisamente os traços- ϕ de pessoa, e se move para um núcleo cujo especificador pode fornecer o traço necessário para possibilitar que a anáfora encontre o antecedente apropriado (MIYAGAWA, 2018).

- (64) a. *Lisi_i juede [Zhangsan_j dui zij_i_{vj} mei xinxin]*
 Lisi achar Zhangsan ter REFL não confiança
 ‘Lisi acha que Zhangsan não tem confiança em si mesmo/nele.’
- b. *Lisi_i juede [wo/ni_j dui zij_i^{*}_{vj} mei xinxin]*
 Lisi achar eu/você ter REFL não confiança
 ‘Lisi acha que eu/você não tenho/tem confiança em mim/você mesmo.’
- c. *Wo_i juede [ni_j dui zij_i^{*}_{vj} mei xinxin].*
 Eu achar você ter REFL não confiança
 ‘Eu acho que você não tem confiança em você mesmo.’ *zij_i=você/*eu*
- d. *Wo_i juede [Zhangsan_j dui zij_i^(*)_{vj} mei xinxin]*
 Eu achar Zhangsan ter REFL não confiança
 ‘Eu penso que Zhangsan não tem confiança em si mesmo.’ *zij_i = Zhangsan/(*)eu*
- e. *Nashi wo_i juede Zhangsan_j dui zij_i_{vj} mei xinxin jiu fangqi le*
 naquele tempo eu achar Zhangsan ter REFL não confiança então desistir PERF
 ‘Naquele tempo, eu pensei que o Zhangsan não tinha confiança em si mesmo, então (eu/ele) desisti/desistiu.’ (MIYAGAWA, 2018, p. 286)

Consoante com Miyagawa (2018), (64a) mostra que a anáfora pode ser ligada localmente, ou pode assumir uma relação de ligação de longa distância com o sujeito da oração mais alta. Em (64b), verifica-se que a interpretação de longa distância é bloqueada se o sujeito local for um pronome com traço semântico de participante de 1^a/2^a pessoa. Em (64c), o bloqueio pelo sujeito local de 1^a/2^a pessoa se mantém mesmo se o sujeito mais alto é de 1^a/2^a pessoa. Em (64d), o autor evidencia que, enquanto o sujeito local de 1^a/2^a pessoa desencadeia

o bloqueio, como em (64b-c), um sujeito local de 3ª pessoa não o faz para muitos falantes (cf. MIYAGAWA, 2017, 2018). Em (64e), tem-se outro exemplo em que um sujeito local de 3ª pessoa não desencadeia concordância.

Miyagawa (2018) supõe que o efeito de bloqueio em (64) se aplica no chinês devido a algum tipo de sistema de concordância de pessoa, assim, a ausência de bloqueio por um sujeito local de 3ª pessoa em chinês equivale à dicotomia encontrada em muitas línguas entre concordância morfológica de participantes e não participantes. Para Miyagawa (2018), enquanto a concordância de participante tem os traços de concordância completos, a concordância de não participantes não o tem, de modo que, como efeito, é uma concordância não marcada.⁴⁹

Para finalizar essa subseção, cabe destacar que, a nosso ver, as propostas de Miyagawa (2010, 2017, 2018) e de Saito & Kim (2012), cada uma com suas particularidades, parecem se fundar no postulado de Huang (1984) de que a ausência total de traços de concordância conduz as LSN orientadas para o discurso de um modo bastante distinto das LSN com morfologia verbal rica. Nessa perspectiva, por meio dos dados empíricos fornecidos pelos autores, verificamos que a presença (ainda que abstrata) de traços- ϕ de concordância em T, como ocorre na 3ª pessoa do singular no inglês coloquial de Singapura e tal como foi proposto por Miyagawa (2017, 2018) para o chinês, altera a direção do licenciamento e da identificação dos sujeitos nulos nas línguas de orientação discursiva. Os traços formais presentes em T parecem, então, ter um papel crucial no licenciamento e na identificação de sujeitos nulos nas línguas de um modo geral.

3.2.3 *Línguas de sujeito nulo consistente*

Conforme vimos no início do capítulo, de acordo com Chomsky (1982, 1986) e Rizzi (1982, 1986), existem algumas propriedades comumente associadas às LSN consistentes: sujeito nulo referencial em orações finitas; inversão livre em orações simples; morfologia verbal rica, que geralmente distingue pessoa e número; pronomes resumptivos nulos em orações subordinadas; e aparente violação do filtro [*that-trace*]. Dentre as distintas categorias nominais vazias identificadas nas línguas naturais, foi postulado que o sujeito nulo é uma

⁴⁹ Cabe destacar que os trabalhos de Miyagawa (2010, 2017, 2018) estão voltados para a compreensão dos mecanismos de concordância nas línguas e seus efeitos na sintaxe e constituem uma proposta inovadora na tentativa de compreender a natureza de *pro*, especialmente por meio das LSN radical. Dessa forma, não nos estenderemos na proposta desse pesquisador, porque o escopo deste capítulo é mais precisamente tratar da tipologia das LSN. Sugerimos, portanto, a leitura dos trabalhos do autor para uma melhor compreensão de sua proposta sobre a natureza de *pro*, especialmente nas línguas orientadas para o discurso.

categoria nominal vazia regida por concordância forte e clíticos. Mais precisamente, esses estudos postularam que, se uma dada língua de sujeito nulo não tem um elemento manifesto no especificador da projeção flexional ou temporal, a exigência do EPP é satisfeita por *pro* sendo concatenado nessa posição.

Contudo, Holmberg (2005) alega que as análises iniciais sobre os sujeitos nulos desenvolvidas no âmbito da teoria de regência e ligação são conflitantes com as premissas teóricas do programa minimalista para a checagem de traços. Como na teoria de checagem de traços de Chomsky (2000, 2001) os traços- ϕ do verbo não são interpretáveis, eles não podem ter seu conteúdo identificado antes de serem valorados; uma análise nos termos da teoria de regência e ligação implicaria dizer que os traços de T teriam de ser valorados por *pro*, mas essa operação de cópia é incompatível com princípios que regulam o modo como as estruturas são construídas no minimalismo, em particular a condição de inclusividade (do inglês *Inclusiveness Condition*), que exige que as propriedades de um nó terminal sejam recuperáveis no léxico e que as propriedades de um nó não terminal sejam recuperáveis na estrutura que ele domina (cf. CHOMSKY, 1995; NUNES, 1999). Copiar informações de T para *pro*, portanto, viola a condição de inclusividade, pois, caso *pro* copiasse os traços- ϕ de T, *pro* adquiriria traços de um nó que ele não domina (HOLMBERG, 2005).

Holmberg (2005) afirma que existem correntemente duas análises para os sujeitos nulos, relacionadas à concordância, que aderem às premissas minimalistas. Primeiramente, considerando especialmente línguas europeias como o grego moderno e o espanhol, supõe-se que *pro* pode ser eliminado do aparato teórico das LSN e que, nessas línguas, a informação verbal contida em T é interpretada como o sujeito da oração, um pronome referencial definido, embora fonologicamente expresso como um afixo (cf. P. BARBOSA, 1995, 2009, 2018; ORDOÑEZ, 1997; ALEXIADOU & ANAGNOSTOPOULOU, 1998; KATO, 1999; MANZINI & SAVOIA, 2002; entre outros). A proposta alternativa a essa tem sido presumir que o sujeito nulo é especificado por um conjunto completo de traços- ϕ e, portanto, valora os traços- ϕ não interpretáveis de T e se move para [SPEC-TP]. Isso implica dizer que a omissão do sujeito é uma questão fonológica: o sujeito nulo é um pronome, ou um NP/DP, que não é pronunciado (cf. HOLMBERG, 2005, 2010a; ROBERTS, 2010; entre outros).

Antes de terem sido teoricamente contestadas por Holmberg (2005), as análises de *pro* desenvolvidas na teoria de regência e ligação e nos primeiros estudos minimalistas foram previamente desafiadas por Alexiadou & Anagnostopoulou (1998), que propuseram que o parâmetro do sujeito nulo necessitava de uma formulação teórica melhor haja vista a existência de línguas em que não há *pro* em [SPEC-TP]. Trata-se, especialmente, de um

estudo contrário àqueles que argumentam que a ordem VS resulta da inserção de um pronome nulo especial, *pro*, numa posição pré-verbal (cf. CHOMSKY, 1982, 1995; RIZZI, 1982; 1986). Alexiadou & Anagnostopoulou (1998) argumentam sistematicamente contra trabalhos que postulam a existência de *pro* expletivo em sentenças com ordem VS, como ilustrado nos contrastes em (65a-b) e (65c-d).

- (65) Espanhol (ALEXIADOU & ANAGNOSTOPOULOU, 1998, p. 492)
- a. *Juan leyó el libro.* (Espanhol).
‘João leu o livro.’
- b. *Leyó Juan el libro.*
Leu João o livro
Grego moderno
- c. *O Petros pandreftike tin Ilektra.*
‘O Pedro casou-se com Elektra.’
- d. *Pandreftike o Petros tin Ilektra.*
Casou-se o Pedro com Elektra

O estudo das autoras centra-se nas LSN que permitem ordem VS em todos os predicados eventivos (66) e não sofrem de restrição para interpretação da definitude em construções inacusativas (67).

- (66) Grego moderno (ALEXIADOU & ANAGNOSTOPOULOU, 1998, p. 495)
- a. *efige o Petros.* (VS com verbo inacusativo).
saiu o Pedro / ‘Pedro saiu.’
- b. *epekse o Petros.* (VS com verbo inergativo).
jogou o Pedro / ‘Pedro jogou.’
- c. *ektise i Maria to spiti.* (VS com verbo transitivo).
construiu Maria casa / ‘Maria construiu uma casa.’
- (67) Grego moderno (ALEXIADOU & ANAGNOSTOPOULOU, 1998, p. 512)
- a. *Eftase ena pedi.*
‘Chegou uma criança.’
- b. *Eftase o Jorgos.*
‘Chegou o Jorge.’

c. *Eftase kathe filos mu.*

‘Chegou cada amigo meu.’

Esse último caso constitui uma forte evidência contra a proposta de Chomsky (1995) de que o expletivo *pro* na ordem SV(O) nas LSN carrega unicamente o traço categorial [D]. Para Chomsky (1995), o expletivo carregando o traço categorial [D] associa-se a um complemento NP. A partir da observação de que essa proposta não abarca dados como em (67a-b), Alexiadou & Anagnostopoulou (1998) postulam que não há *pro* em posição pré-verbal na ordem VS nas LSN consistentes. Para Alexiadou & Anagnostopoulou (1998), os traços de concordância do verbo finito nessas línguas, como ilustrado no contraste em (68a), são nominais o suficiente para satisfazer a exigência do EPP.

(68)	a. Inglês	b. Grego moderno	
	<i>I love</i>	<i>Agapo</i>	‘(Eu) amo.’
	<i>You love</i>	<i>Agapas</i>	‘(Tu) amas.’
	<i>He loves</i>	<i>Agapa</i>	‘(Ele) ama.’
	<i>We love</i>	<i>Agapame</i>	‘(Nós) amamos.’
	<i>You love</i>	<i>Agapate</i>	‘(Vós) amais.’
	<i>They love</i>	<i>Agapane</i>	‘(Eles) amam.’

(ALEXIADOU & ANAGNOSTOPOULOU, 1998, p. 517)

A afirmação das autoras pressupõe que em línguas como o grego moderno: os afixos de concordância são pronomes [+definidos], têm traços- ϕ interpretáveis e atribuem Caso estrutural; o sujeito é interno ao VP; e o movimento do verbo para o núcleo funcional flexional/temporal satisfaz o EPP e estabelece a ordem VS. Em línguas que não permitem sujeito nulo, como o inglês, um elemento XP deve ocupar [SPEC-TP] para satisfazer o EPP.⁵⁰

Essa análise aparentemente simples explica a correlação entre riqueza morfológica e inversão livre. A inversão livre, nesses termos, é uma consequência natural do fato de que [SPEC-TP] não precisa ser preenchido nas LSN consistentes, e a concordância morfológica produz a propriedade do sujeito nulo, satisfazendo o EPP via movimento do verbo (ALEXIADOU & ANAGNOSTOPOULOU, 1998). Isto é, se uma dada LSN não tem um

⁵⁰ Para Alexiadou & Anagnostopoulou (1998), como o movimento A para [SPEC-TP] não é necessário nas LSN consistentes, todo movimento para uma posição pré-verbal, como em (65a) e (65c), é movimento A-barra.

elemento manifesto em [SPEC-TP], a exigência do EPP deve ser satisfeita por um traço-D no núcleo funcional T.⁵¹

Uma implementação importante da proposta de Alexiadou & Anagnostopoulou (1998) diz respeito à derivação das orações com sujeitos nulos, ou seja, das orações em que os afixos da morfologia do verbo finito têm valor pronominal. Defendendo uma distinção entre sistemas de concordância forte e concordância fraca, as autoras argumentam que no primeiro sistema os afixos de concordância têm entradas lexicais independentes, como se fossem elementos pronominais, e possuem um traço categorial [D]. Desse modo, o radical do verbo e o afixo [+D] são independentes um do outro e, portanto, estão separados na numeração. Considerando que as LSN consistentes possuem concordância forte, as autoras apontam que há duas opções para o afixo [+D] no percurso da derivação: primeiro, ser concatenado com a raiz verbal em um estágio inicial da derivação, ou seja, em seu domínio interno, de modo que o verbo será projetado e o elemento complexo resultante dessa concatenação será fundido em alguma estrutura maior – esse é o caso de quando há movimento de V-para-T, e esse movimento satisfaz o EPP. Noutra opção, o afixo [+D] pode ser concatenado diretamente no núcleo funcional flexional/temporal, assim o verbo se move para esse núcleo – esse seria o caso dos clíticos sujeitos nos dialetos italianos (ALEXIADOU & ANAGNOSTOPOULOU, 1998). Nos sistemas de concordância fraca, prevê-se que os afixos de concordância não são independentes um do outro na derivação e não têm um traço categorial independente. Consoante com essa análise, a posição do sujeito [SPEC-TP] sequer é projetada nas orações declarativas com verbos finitos e ordem VS(O) nas LSN como o grego moderno e o espanhol.

Se a concordância no verbo finito é suficiente para satisfazer o EPP na análise de Alexiadou & Anagnostopoulou (1998), isso não ocorre na abordagem minimalista para os sujeitos nulos proposta por Holmberg (2005). Nessa segunda abordagem, a posição de [SPEC-TP] é ocupada por um argumento nulo (pronominal ou nominal) que verifica o EPP e, portanto, não está disponível para outra categoria. Uma das análises no âmbito dessa abordagem é a do próprio Holmberg (2005), que defende uma revisão das hipóteses sobre as LSN, especialmente a proposta de Rizzi (1986), argumentando que não é a referencialidade, mas, sim, a definitude [D] a propriedade que deve ser crucialmente tomada para caracterizar

⁵¹ Alexiadou & Anagnostopoulou (1998) estendem a proposta dos afixos pronominais aos clíticos sujeitos, como os dos dialetos trentino e fiorentino. Elas afirmam, no entanto, que os clíticos sujeitos possivelmente satisfazem o traço-EPP via *Merge*, diferentemente das LSN consistentes, que o fazem via movimento (*Move*) (cf. exemplos (7)-(9) neste capítulo para dados com clíticos sujeitos no trentino). Os sujeitos clíticos do francês, por sua vez, seguem o padrão XP em [SPEC-TP] do inglês, segundo as autoras. Para elas, o que marca a diferença entre os clíticos sujeitos do francês e os dos dialetos italianos é o fato de os primeiros sofrerem restrição para interpretação da definitude em construções inacusativas – restrição que consiste no não licenciamento de um nome nu em construções inacusativas com ordem VS, como no exemplo: ‘*Il est arrivé Jean.’ (cf. *Ibid.*, p. 518).

as LSN consistentes. Holmberg (2005, 2010a) mostra que, nas LSN consistentes, quando um pronome nulo na 3ª pessoa do singular entra numa relação de concordância com T e T possui um traço de definitude não interpretável [μ D], é gerada uma sentença como em (69).

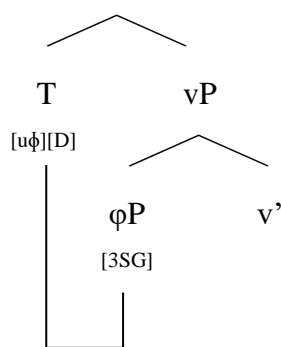
- (69) Italiano (HOLMBERG, 2010a, p. 88)
Verrà.
 vir-3SG-FUT
 ‘(Ele/ela) virá.’

Holmberg (2005, 2010a) analisa os sujeitos nulos em (69) como uma estrutura pronominal menor que os pronomes regulares, formalmente ϕ P. Seguindo essa visão, ele se baseia em duas premissas fundamentais para sua abordagem sobre os sujeitos nulos nas LSN:

- (70) a. Os pronomes são DPs, com a estrutura [$_{DP}$ D [$_{\phi P}$ ϕ [$_{NP}$ N]]], ou ϕ Ps;
 b. Pronomes nulos são ϕ Ps. (HOLMBERG, 2010a, p. 94)

Com base em (70), Holmberg (2010a) ilustra em (71) a relação de quando T sonda um ϕ P na 3ª pessoa do singular e tem seus traços- ϕ valorados por esse pronome, resultando dessa união um pronome nulo definido.⁵²

- (71) Relação sonda-alvo entre T e um ϕ P (HOLMBERG, 2010a, p. 95)



De acordo com Holmberg (2010a), na derivação de sujeitos nulos definidos de 3ª pessoa que são dependentes de um antecedente nas LSN consistentes, há um tópico de mudança de orientação (do inglês *Aboutness-shift topic*) localizado em C, nulo ou manifesto,

⁵² Na subseção 3.2.4 a seguir, explicaremos as implicações da relação sonda-alvo entre T e um ϕ P nas LSN em que T não possui um traço [μ D] para ser valorado.

mediando a relação de coindexação entre o sujeito nulo e seu antecedente, que corresponde ao tópico de uma oração anterior, como representado em (72):⁵³

- (72) Italiano (HOLMBERG, 2010a, p. 96)
- a. *Questa mattina, Gianni ha visitato la mostra. Più tardi __ ha visitato l'università.*
 ‘Nesta manhã, Gianni visitou a mostra. Mais tarde __ visitou a universidade.’
- b. [CP <Gianni₁> [*questa mattina Gianni₁ ha visitato la mostra*]]
 [CP <Ø₂> [*più tardi ha* φP₂ *visitato l'università*]] 1 = 2

O pronome nulo de 3ª pessoa na oração subordinada em (72) é dependente de um tópico. Holmberg (2010a) argumenta que o índice referencial do sujeito nulo (Ø₂/φP₂) vem do índice do DP expresso fonologicamente no discurso anterior (*Gianni₁*) por intermédio de uma cadeia de tópicos. Nessa direção, Holmberg (2010a) explica que a relação de compartilhamento de índice entre o tópico nulo e o sujeito nulo em (72b) envolve crucialmente T: o tópico nulo valora o traço [*u*D] de T, e a valoração consiste em [*u*D] copiar o índice referencial do tópico nulo. Esses procedimentos são capazes de satisfazer o EPP, de acordo com Holmberg (2010a), como ilustrado na derivação da oração declarativa em (73).

- (73) Italiano (HOLMBERG, 2010a, p. 105)
- Ha comprato una macchina nuova.*
 Tem comprado um carro novo (‘[Ele] comprou um carro novo.’).
 [CP <DP₁> [TP *ha* + T[*D*₁, 3SG, EPP] [*v*P <φP[3SG, NOM]> *comprato ...*]]]

Quanto aos sujeitos nulos de 1ª e 2ª pessoa, Holmberg (2010a) adota a hipótese de que cada oração possui traços semânticos representando o falante e o destinatário no domínio de C. Dessa forma, o falante e o ouvinte estão sempre disponíveis como antecedentes locais. Isso impede, nesses casos, a atribuição de um índice referencial a T por outro tópico frasal em [SPEC-CP].

⁵³ Traduzimos *Aboutness-shift topic* como um tópico de mudança de orientação, pois, para Holmberg (2010a), a noção de tópico consiste essencialmente em uma expressão que denota um indivíduo ou um grupo já estabelecido no discurso, sobre o qual o predicado diz alguma coisa. Esse tópico encontra-se, na perspectiva do autor, no domínio do núcleo mais alto da sentença.

Adotando a proposta de Holmberg (2005, 2010a) de que as LSN consistentes possuem um traço-D em T, Roberts (2010) formula a hipótese em (74) em relação ao traço-D e à especificação dos traços- ϕ nas LSN consistentes.

(74) Se uma categoria α tem um traço-D, então todos os traços- ϕ de α são especificados.

(ROBERTS, 2010, p. 82)

Com base em (74), Roberts (2010) defende que a supressão de um pronome na derivação somente é possível se ele identifica os traços do núcleo funcional T. Desse modo, estipula-se que qualquer língua que tenha um traço-D não interpretável em T e cinco ou seis morfemas distintos para recuperar as pessoas gramaticais permite sujeitos nulos (ROBERTS, 2010). Roberts (2010) propõe que os sujeitos nulos nessas línguas são pronomes DP, mais precisamente pronomes fracos, deletados em PF antes de irem para *Spell-out*. O pesquisador assume que sujeitos nulos são pronomes fracos que devem ocupar uma posição de especificador, isto é, [SPEC-TP], têm um traço-D valorado como definido e valoram o traço [uD] em T. Tendo em vista que T contém os traços- ϕ que correspondem aos do sujeito, na relação *probe-goal* do sujeito com T, o sujeito é um alvo defectivo (do inglês *defective goal*), no sentido de que seus traços são esgotados por aqueles presentes na sonda (*probe*) – ou seja, alvos defectivos nunca têm uma realização em PF, independentemente de suas sondas (ROBERTS, 2010). Por ser um alvo defectivo, o sujeito é deletado em PF pela operação *Chain Reduction* (ROBERTS, 2010).⁵⁴

Crucialmente, Roberts (2010) propõe que o apagamento do sujeito nulo implica a projeção de [SPEC-TP], de modo que a diferença entre as LSN consistentes e as demais LSN está na composição de T: T carrega traços nominais que, em combinação com sua especificação verbal, asseguram que um pronome fraco projetado em [SPEC-TP] constitua um conjunto de traços daqueles presentes em T, justificando assim o apagamento de pronomes de acordo com suposições mais gerais em termos de *Chain Reduction*, isto é, em termos de um processo de apagamento em PF, que consiste na exclusão de todas as cópias idênticas em uma dependência, exceto a da cadeia mais alta (ROBERTS, 2010). Como

⁵⁴ *Chain* (cadeia) é um conjunto de ocorrências; se as ocorrências estão em contextos completos, não é preciso dizer que uma cadeia é uma sequência, pois há uma relação de contenção entre os contextos que permite reconstruir o que for necessário (cf. CHOMSKY, 1995). *Chain Reduction* consiste no seguinte axioma: elimine o número mínimo de constituintes de uma cadeia não trivial CH que seja suficiente para que CH seja mapeado numa ordem linear de acordo com o axioma de correspondência linear (NUNES, 2004). O axioma de correspondência linear consiste em: sejam X, Y não terminais e x, y terminais tal que X domina x e Y domina y. Então, se X assimetricamente c-comanda Y, x precede y (KAYNE, 1994).

consequência do postulado em (74), T em uma língua sem um traço-D não tem seus demais traços- ϕ especificados. Assim, Roberts (2010) propõe que se T perde um traço-D, *pro*, sendo um pronome fraco, não é um alvo defeutivo e, portanto, não pode ser nulo, ou seja, não pode ser deletado (ROBERTS, 2010). Nesses termos, Roberts (2010) salienta que as LSN consistentes não têm traços- ϕ empobrecidos em T; antes, T tem um traço [*uD*] e *pro* é um alvo defeutivo e, portanto, deletado antes de chegar em PF (ROBERTS, 2010).⁵⁵ Já nas línguas de sujeito não nulo, como o inglês, os traços- ϕ em T são empobrecidos, e, portanto, essas línguas não apresentam traço-D em T nem *pro*.

Em sua argumentação de que se T tem um traço-D, então todos os traços- ϕ de T devem ser especificados, Roberts (2010) integra a generalização de *pro* de Müller (2006), ilustrada em (75).

(75) Generalização de *pro* de Müller (2006, p. 10)⁵⁶

Um *pro* DP argumental não pode sofrer concordância (*Agree*) com núcleo funcional α se α foi submetido (talvez de forma vaga) ao empobrecimento neutralizante dos traços- ϕ na numeração.

Sob o escopo da Morfologia Distribuída, adotada por Müller (2006), empobrecimento é definido como uma operação que elimina traços morfossintáticos em morfemas abstratos em certos contextos específicos (cf. BONET, 1991; MÜLLER, 2006, 2008; entre outros); ou seja, um morfema que sofre essa operação passa a ter um conjunto de traços menos especificado. Assim, em determinadas línguas, existem certas operações de empobrecimento que sistematicamente suprimem, sob valoração, traços- ϕ em T, levando as especificações de traços dos morfemas a se tornarem idênticas, tendo a mesma realização fonológica. Nessa direção, de acordo com (75), *pro*, como qualquer sujeito DP, estabelece uma relação de concordância com T, mas, diferentemente dos DPs, não pode entrar em tal relação se T tiver sido empobrecido. Em outras palavras, o T empobrecido não pode concordar com um sujeito nulo (cf. MÜLLER, 2006, 2008). Destaca-se que Müller (2006, 2008) propõe que a

⁵⁵ De acordo com Neeleman & Szendrői (2007), os sujeitos nulos nas línguas *pro-drop* radical podem ser vistos como pronomes comuns que são eliminados em PF. Na visão dos autores, a diferença entre LSN consistentes e orientadas para o discurso é que as primeiras necessariamente envolvem uma operação de apagamento que referencia as propriedades (pro)nominais de T, enquanto nas últimas o apagamento se aplica independentemente das propriedades de T. Esse tipo de postulação de *pro* para as LSN de orientação para o discurso é possível porque, devido à ausência de concordância nessas línguas, o problema com a operação de cópia não existe.

⁵⁶ Apesar de enfatizar em seus estudos a categoria pessoa, Müller (2006, 2009) reconhece que, de acordo com sua generalização, as categorias de número e modo, por exemplo, também podem levar ao empobrecimento de um paradigma flexional.

morfologia flexional é pré-sintática, mas, apesar disso, emprega subespecificação. Isto é, para o autor, o empobrecimento pode ser visto como uma operação pré-sintática desde que os traços afetados sejam invisíveis para a morfologia, mas não para operações sintáticas subsequentes.

A generalização proposta por Müller (2006) explicaria por que os sujeitos se tornam gradualmente manifestos em línguas nas quais os traços- ϕ em T sofrem empobrecimento. O postulado de Roberts (2010) em (74), por sua vez, prevê que o empobrecimento de T (necessário para a perda da morfologia rica que, de acordo com Roberts (2010), também caracteriza as LSN consistentes) se inicia com a ausência de um traço-D em T. Em outras palavras, Roberts (2010) sugere que a perda do traço-D de T em uma LSN consistente resulta na queda de sujeitos nulos e, possivelmente, no empobrecimento do paradigma verbal dessa língua. Essa é a abordagem que tomaremos na análise dos nossos dados no Capítulo 6.

Trabalhando no âmbito da Morfologia Distribuída, Saab (2016) desenvolve uma abordagem alternativa da ideia de que os sujeitos nulos são instâncias elididas de argumentos manifestos a partir de uma perspectiva teórica em que a elipse é compreendida como uma operação de bloqueio de inserção de léxico na derivação. Saab (2016) sustenta a ideia de que a diferença entre um pronome manifesto e um pronome nulo é que o primeiro não possui [+I] e obtém conteúdo por meio de regras de inserção lexical, enquanto o último possui [+I], o que bloqueia a aplicação dessas regras (cf. SAAB, 2016, para detalhes sobre as regras de inserção lexical e de bloqueio dessas regras). Vejamos (76).

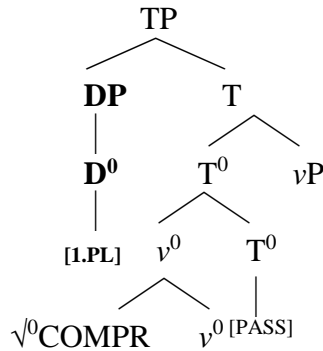
(76) Atribuição morfológica de I (do inglês *Morphological I-Assignment*) [elipse de núcleo] (SAAB, 2016, p. 54)

Dada uma palavra morfossintática Y^0 , um recurso [+I] é atribuído a Y^0 se e somente se houver um nó X^0 idêntico a Y^0 contido em uma palavra morfossintática adjacente ou imediatamente local a Y^0 (onde a noção de contenção é reflexiva).

Saab (2016) salienta que localidade imediata é a relação entre um núcleo e um núcleo de seu complemento. A operação em (76) sucintamente prediz um conjunto de interações entre sintaxe, elipse de núcleo e outras operações pós-sintáticas, como concordância (SAAB, 2016). Para Saab (2016), (76) explicita um processo de afixação de morfemas em PF que instancia os sujeitos nulos nas LSN consistentes, como ilustrado em (77).

(77) Espanhol: sintaxe de ‘Comparamos.’

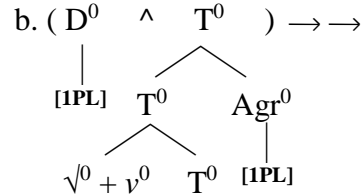
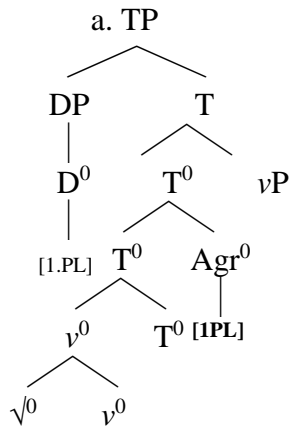
(SAAB, 2016, p. 56)



De acordo com (77), um sujeito nulo é uma projeção máxima D em [SPEC-TP] (SAAB, 2016). Assume-se, assim, que a concordância é implementada exclusivamente em PF por meio da introdução de um morfema dissociado, que é uma cópia dos traços formais do sujeito (SAAB, 2016). Formalmente, uma estrutura sintática como (78a) é linearizada como (78b), e D pode ser valorado por um traço [+I].⁵⁷ Esse traço é adicionado aos núcleos elípticos sob identidade formal e bloqueia as regras de inserção lexical, resultando em um sujeito nulo.

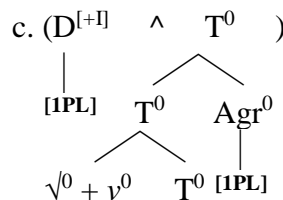
(78) *Concordância (Agreement)*

Linearização



→ → → →

Atribuição morfológica de I



(SAAB, 2016, p. 56)

Saab (2016) chama a atenção para o fato de que em (77) o nó D⁰ na posição de sujeito é uma palavra morfossintática, mas o morfema dissociado não o é; isso prediz que o sujeito

⁵⁷ Linearização consiste basicamente no modo como os falantes ordenam seus enunciados. Para uma abordagem formal desse conceito, tal como proposto por Saab (2016), remetemos o leitor a Embick (2007).

pode ser elidido, ou seja, não pronunciado, mas o morfema dissociado não o pode, como ilustrado em (79).

- (79) Espanhol (SAAB, 2016, p. 57)
- a. $\bar{D}(P)[1_{PH}]$ *compramos un libro.*
- b. **Nosotros compramos un libro.*
- ‘(Nós) compramos um livro.’

Saab (2016) conjectura que a introdução de morfemas dissociados, isto é, de concordância morfológica, e o EPP, ou seja, a exigência de que T deve estar associado a um traço nominal ou D na sintaxe ou em PF, estão em distribuição complementar. Para Saab (2016), a conjectura que ele esboça e a observação de Kato de que o surgimento de formas fracas se correlaciona com a perda de propriedades *pro-drop* fornecem uma explicação clara para lidar com o problema da variação linguística em algumas línguas *pro-drop*. Nessa perspectiva, o autor defende a proposta de que se pode pensar a mudança sofrida na sintaxe dos sujeitos no PB como uma mudança que vai da morfologia à sintaxe: por causa da perda da concordância morfológica, o traço-EPP no PB tem que ser satisfeito na sintaxe (cf. M. E. DUARTE, 1993 [2018], 1995, 2000).

Para P. Barbosa (2009), as LSN consistentes não são guiadas pelo requerimento de um elemento XP mover-se para [SPEC-TP] devido às propriedades de sua riqueza morfológica. Adotando a proposta de Pesetsky & Torrego (2007) para *Agree*, P. Barbosa (2009) defende que T nas LSN consistentes é semelhante a um clítico pronominal, isto é, possui um conjunto de traços- ϕ inerentemente valorados, e que *pro* é um nominal minimamente especificado, cujos traços valorados são fornecidos pela operação *Agree* com o conjunto valorado, embora não interpretável, de traços- ϕ de T. Em estudo mais recente, P. Barbosa (2018) propõe que os traços- ϕ e o traço-D em T são interpretáveis nas LSN consistentes. A autora indica como um dos corolários dessa abordagem o fato de que os sujeitos pré-verbais não quantificados devem ser considerados como clíticos deslocados à esquerda (do inglês *Clitic Left Dislocated*). Nessa perspectiva, o Princípio Evite Pronome, que induz à não realização fonológica do sujeito quando sua plena identificação for possível, reduz a opcionalidade de sujeitos manifestos por não concatenar um pronome como um clítico deslocado à esquerda, a não ser que seja necessário sinalizar mudança de tópico ou ênfase (P. BARBOSA, 2018). Ainda nessa perspectiva, a alegação de que T hospeda um traço-D e um conjunto interpretável de traços- ϕ torna concebível postular que a posição temática do sujeito seja preenchida por um NP

fonologicamente nulo – formalmente, um nP , no sentido de que esse item consiste minimamente em um núcleo de categorização n (cf. MARANTZ, 2001), isto é, corresponde a um n que não se funde com uma raiz.

3.2.4. Línguas de sujeito nulo parcial

Conforme vimos anteriormente, Holmberg (2005, 2010a) mostra que as análises sobre sujeitos nulos no âmbito da teoria de regência e ligação são incompatíveis com as versões atuais da teoria de valoração de traços de Chomsky (2000, 2001). Nessa direção, ele propõe uma distinção entre as LSN que licenciam sujeitos nulos definidos e aquelas que não o fazem. Para Holmberg (2010a), sujeitos nulos são formas pronominais – ϕP – e as LSN consistentes caracterizam-se primordialmente por possuírem um traço de definitude não valorado [uD] em T. O resultado da união entre um ϕP na 3ª pessoa do singular e os traços- ϕ e [uD] de T nas LSN consistentes é um pronome nulo definido. Por outro lado, se for considerada uma LSN sem [uD] em T, a relação sonda-alvo entre um ϕP na 3ª pessoa do singular e T não fornece um valor de definitude. O resultado é um pronome nulo indefinido cuja interpretação é a de um sujeito nulo com traço [+ humano], como em (80b). Contrariamente, nas LSN consistentes, os sujeitos indefinidos na 3ª pessoa do singular são marcados fonologicamente, como em (80a).

- (80) a. É assim que *se* faz o doce. (PE/PB) (HOLMBERG, 2010a, p. 92)
 b. É assim que ___ faz o doce. (PE*/PB)

Resumidamente, quando um traço-D está ausente do conjunto de traços de T numa LSN e um pronome nulo na 3ª pessoa do singular e T entram numa relação de *Agree*, o pronome nulo só pode ser interpretado indefinidamente, como em (80b).⁵⁸ Como resultado dessa análise, Holmberg (2010a) classifica as LSN que licenciam sujeitos nulos definidos na 3ª pessoa do singular como LSN consistentes e as LSN que licenciam sujeito nulo indefinido na 3ª pessoa do singular como LSN parciais. Nessa primeira classe, encontra-se o PE; na última, o PB.

Em relação aos sujeitos nulos definidos, Holmberg (2010a) propõe duas maneiras de derivá-los nas LSN: a primeira é por meio da incorporação de um pronome sujeito em T, na

⁵⁸ Conforme vimos na Introdução da tese, Holmberg (2010a) afirma que o pronome nulo indefinido no PB possui ambas as leituras genéricas, quais sejam, inclusiva e exclusiva.

qual o sujeito nulo é uma cópia apagada numa cadeia nucleada por T. Esse modelo deriva sujeitos nulos definidos nas LSN consistentes. A outra maneira de derivá-los é por meio do apagamento de um pronome em [SPEC-TP] controlado por uma oração mais alta. De acordo com Holmberg (2010a), essa é a maneira como os sujeitos nulos definidos podem ser gerados nas LSN parciais. Isso implica dizer que pronomes referenciais nulos numa LSN parcial, como o finlandês e o PB, são licenciados desde que localmente c-comandados por um antecedente (81a), diferentemente do que ocorre nas LSN consistentes, como o italiano (81b).

(81) a. Finlandês (HOLMBERG, 2010a, p. 92)

*Juha₁ ei ole sanonut mitään, mutta Pauli₂ sanoo että *Ø₁ haluaa ostaa uuden auto.*

‘O João₁ não disse nada, mas Paulo₂ disse que (ele)₁ deseja comprar um carro novo.’

b. Italiano

Gianni₁ non ha detto niente, ma Paolo₂ ha detto che Ø₁ vuole comprare una macchina nuova.

‘O João₁ não disse nada, mas o Paulo₂ disse que (ele)_{1/2} deseja comprar um carro novo.’

De acordo com Holmberg (2010a), o problema exibido em (81) pode ser expresso da seguinte forma: o sujeito nulo da coordenada não pode ser c-comandado pelo DP *Juha*, de modo que o único antecedente possível é o DP *Pauli*. Em italiano, c-comando não é um requerimento quando o antecedente é um tópico, o que possibilita a interpretação do DP *Gianni* como o referente do sujeito nulo da coordenada em (81b) – curiosamente, caso o sujeito da coordenada fosse manifesto em ambas as línguas, a interpretação do referente do sujeito seria ambígua entre os dois DPs da sentença.⁵⁹

Dentre as características das LSN parciais, Holmberg & Sheehan (2010) apontam o fato de que essas línguas apresentam sujeitos nulos quase argumentais em construções com verbos meteorológicos, como em “Está chovendo”. Eles afirmam também que a primeira e a segunda pessoas gramaticais podem ser recuperadas na morfologia verbal de determinadas orações finitas nessas línguas e que um sujeito nulo referencial de terceira pessoa na posição de sujeito só pode ser nulo quando é controlado por um argumento mais alto (cf. HOLMBERG, 2010a; HOLMBERG & SHEEHAN, 2010). Holmberg (2010a) e Holmberg &

⁵⁹ Holmberg, Nayudu & Sheehan (2009) propuseram que existe uma escala de tipos de controle que licencia sujeitos nulos nas LSN parciais. Em particular, eles argumentam que, além do controle obrigatório e não obrigatório, há casos intermediários a esses e que não são atestados em ambientes infinitivos. Remetemos o leitor à leitura dos autores (incluindo obras mais recentes) para detalhes sobre essa questão.

Sheehan (2010) pontuam que as LSN parciais diferem entre si quanto a se pronomes de 1ª e 2ª pessoa podem ser nulos quando usados como indexicais, isto é, quando expressos como elementos que dependem do contexto de fala para estabelecer seus referentes. Nesse contexto, o hebraico moderno, interessantemente, permite sujeitos nulos de 1ª e 2ª pessoa via movimento do verbo para T. Entretanto, o verbo nessa língua se move para T somente no passado e no futuro. Conseqüentemente, os sujeitos nulos de 1ª e 2ª pessoa em orações matrizes no hebraico moderno só são possíveis nos tempos passado e futuro (LANDAU, 2004), como ilustrado em (82).

- (82) Hebraico moderno (LANDAU, 2004, p. 815)
- a. *(ani) dibarti / (ata) dibarta / *(hu) diber.*
 (eu) falei-1SG.PASS / (você) falou-2SG.PASS / *(ele) falou-3SG.PASS
- b. *(ani) adaber / (ata) tedaber / *(hu) yedaber.*
 (eu) falarei-1SG.FUT / (você) falará-2SG.FUT / *(ele) falará-3SG.FUT
- c. **(ani/ata/hu) medaber.*
 *(eu) falo-1SG.PRES / *(você) fala-2SG.PRES / *(ele) fala-3SG.PRES

Uma previsão que essa descrição sobre o hebraico moderno faz é que nas LSN parciais, mesmo os pronomes de 1ª e 2ª pessoa devem ser ϕ Ps quando não são interpretados referencialmente, isto é, quando não são interpretados como índices, mas como pronomes ligados. Como foi exposto, para Holmberg (2010a), os sujeitos nulos referenciais de 1ª e 2ª pessoa nas LSN parciais são sempre pronomes definidos completos, e as LSN parciais diferem quanto a esses sujeitos serem, ou não, pronunciados em PF. Nas LSN consistentes, por sua vez, os indexicais nulos são pronomes deficientes, e a interpretação referencial é atribuída através de um traço-D em T. A proposta de Holmberg (2010a) para as LSN parciais em contraste com as LSN consistentes pode ser resumidamente descrita como em (83):

- (83) Línguas de sujeito nulo consistente e parcial (HOLMBERG, 2010, p. 94)
- a. *Línguas de sujeito nulo consistente:*
 Presença de sujeitos nulos definidos ('ele/ela' nulos);
 Ausência de sujeito nulo indefinido na 3ª pessoa do singular.
Ex.: italiano, espanhol, grego moderno, PE.

b. Línguas de sujeito nulo parcial:

Presença de sujeitos definidos nulos se localmente c-comandados por um antecedente;

Presença de sujeito nulo indefinido na 3ª pessoa do singular

Ex.: finlandês, marata, hebraico moderno, PB.

Holmberg (2010a) argumenta que um pronome definido nas LSN parciais não está adjungido a T e, portanto, é atraído para [SPEC-TP]. Dessa forma, uma vez que T nas LSN parciais, como o finlandês e o PB, perdeu o traço-D, um pronome definido sujeito deve ter seu próprio traço-D, isto é, seu próprio índice (HOLMBERG, 2010a). Ao ser sondado por T, esse pronome será forçado pelo EPP de T a se fundir mais uma vez, agora em TP. A cópia mais alta da cadeia deve ser pronunciada (HOLMBERG, 2010a). A derivação em (84) resume essa análise.

(84) Finlandês (HOLMBERG, 2010a, p. 105)

Hän on ostanut uuden auton.

ele tem comprado carro novo ('Ele comprou um carro novo.')

[TP *Hän*_{[D1, 3SG, NOM]] [T' *on*+T_{[3SG, EPP]] [VP <*hän*_[D1, 3SG, NOM]] > *ostanut* ...]]]}}

Uma consequência da proposta de derivação acima é que não há incorporação em T se um pronome nulo de 3ª pessoa possui um traço-D defectivo e não valorado (rotulado *uDP*), de modo que esse pronome será projetado para [SPEC-TP], onde satisfará o EPP e será interpretado somente se for controlado pelo argumento de uma oração mais alta, como ilustrado em (85a). (85b) é a derivação de (85a) antes da operação de controle, que subsequentemente atribuirá o índice 1 ao pronome nulo não valorado; ou seja, em (85b), o pronome é nulo não em virtude da incorporação em T, mas em virtude de ter um antecedente local, do qual resulta a interpretação definida do sujeito (HOLMBERG, 2010a).

(85) Finlandês (HOLMBERG, 2010a, p. 102 e 105)

a. Jari sanoo että _ istuu mukavasti tässä.

João disse-3SG que senta confortavelmente aqui

'O João disse que (ele) senta confortavelmente aqui.'

≠ 'O João disse que se senta confortavelmente aqui.'

b. Jari₁ ... [CP että [TP uDP [T' istuu+T_{3SG, EPP} [VP <uDP_{3SG}> <istuu> mukavasti tässä]]]]

Outra consequência da proposta de Holmberg (2010a) é que o pronome nulo indefinido de 3ª pessoa singular nas LSN parciais deve ocupar [SPEC-vP] e, portanto, não pode checar o EPP (HOLMBERG, 2010a), permitindo que outras categorias na sentença satisfaçam o EPP. Assim, uma vez que o sujeito nulo indefinido não tem um tópico antecedente, o EPP é checado por alguma outra categoria, como em (86a) e em sua derivação em (86b).

(86) Finlandês (HOLMBERG, 2010a, p. 102 e 105)

a. Jari sanoo että tässä istuu mukavasti.
 João disse-3SG que aqui senta confortavelmente
 ‘O João disse que aqui se senta confortavelmente.’
 ≠ ‘O João disse que ele senta confortavelmente aqui.’

b. [CP että [TP tässä [T' istuu+T_{3SG, EPP} [VP φP_{3SG} <istuu> mukavasti <tässä>]]]]

Em estudo sobre as propriedades sintáticas e semânticas do pronome nulo genérico no finlandês, Holmberg (2010b) observa que: (i) ele sempre possui leitura genérica inclusiva (isto é, inclui o falante o ouvinte) e traço [+humano], ou seja, é um argumento nulo com papel temático de agente; (ii) faz tudo o que pronomes manifestos fazem, contudo, não satisfaz o traço-EPP de T; desse modo, (iii) um advérbio locativo, como (87b), ou um pronome expletivo manifesto, como em (87c), satisfaz o traço-EPP em [SPEC-TP], desencadeando, por incorporação, concordância com a terceira pessoa do singular e atribuindo Caso nominativo ao constituinte que ocupa [SPEC-TP].

(87) Finlandês (HOLMBERG, 2010b, p. 210)

a. **Istuu mukavasti tässä.*
 Senta-3SG confortavelmente aqui
 b. *Tässä istuu mukavasti.*
 Aqui senta-3SG confortavelmente
 ‘Aqui (se) senta confortavelmente.’
 c. *Sitä istuu mukavasti tässä.*
 EXP senta-3SG confortavelmente aqui
 ‘Senta-se confortavelmente aqui.’

Comparando os dados do finlandês com o PB, outra LSN parcial que licencia sujeitos nulos genéricos, verifica-se que construções com sintagmas adverbiais em posição pré-verbal e pronomes nulos genéricos também são encontradas no PB, com a diferença de que no PB são possíveis tanto a leitura inclusiva (88) quanto a leitura exclusiva (ou seja, aquele que exclui o falante e o ouvinte) dos pronomes nulos genéricos. Ambas as construções também possuem interpretação [+ humana].

(88) Nos nossos dias não __ usa mais saia. (PB) (GALVES, 1987, p. 37)

(89) a. João me contou que na praia *(e_1) vende cachorro quente.

b. João me contou que e_1 vende cachorro quente na praia.

(C. RODRIGUES, 2004, p. 142)

C. Rodrigues (2004) propõe que o sujeito nulo genérico no PB seja um expletivo nulo. Ela argumenta que o advérbio em (89a) bloqueia a extração do sujeito da oração encaixada pelo sujeito da oração matriz, permitindo a interpretação genérica do sujeito nulo da subordinada. Em (89b), isto é, na ausência de um elemento intervindo entre a subordinada e a oração matriz, o sujeito da oração principal c-comanda o sujeito nulo da subordinada (C. RODRIGUES, 2004). Modesto (2008) reconhece que, caso um adjunto adverbial ocupe uma posição deslocada à esquerda da frase, a interpretação para uma sentença como em (90) é ambígua entre uma leitura definida e uma leitura genérica no PB.

(90) Na praia, o Feco me falou que vende cachorro quente.⁶⁰

Figueiredo Silva (2000) constatou que um sujeito nulo de 3ª pessoa com leitura genérica no PB também pode ser licenciado numa interrogativa, como em (91).

(91) Onde compra cartão postal aqui? (FIGUEIREDO SILVA, 2000, p. 131)

Figueiredo Silva (2000) argumenta que um dado como em (91) só seria interpretado referencialmente no PB caso um tópico sentencial ou discursivo ancorasse a interpretação do

⁶⁰ Chamamos a atenção para o fato de que a ambiguidade em estruturas como em (90) pode ser restrita dependendo do escopo do adjunto adverbial.

sujeito nulo; na ausência desse tópico, a autora afirma que as construções com sujeitos nulos na 3ª pessoa do singular no PB são interpretadas como se o sujeito nulo fosse o clítico *se-impessoal*, um sujeito nulo genérico. Com base na análise desses e outros fenômenos do PB, Figueiredo Silva (2000) sugere que o PB se tornou “uma língua de sujeito nulo parcial fazendo uso de estratégias especiais para identificar a categoria vazia sempre que os sujeitos nulos devem ser interpretados referencialmente” (FIGUEIREDO SILVA, 2000, p. 127).⁶¹

Figueiredo Silva (2017, p. 196) declara que os sujeitos nulos genéricos no PB parecem ser licenciados sob o “aspecto exibido pelo presente do indicativo ou pelo pretérito imperfeito, mas crucialmente não pelo pretérito perfeito”, tendo em vista que esse tempo verbal atuaria como um operador que deveria licenciar a interpretação referencial, não mais disponível na gramática do PB. Conforme apresentamos na Introdução desta tese, essa argumentação não serve para o PB oitocentista em Goiás: em (92), temos um sujeito nulo arbitrário, isto é, de leitura genérica exclusiva, licenciado pelo pretérito perfeito.

(92) Dia 19 __ **pintou** o estuque da sala. (NAVES & BORGES, 2015, p. 83)

Lunguinho & Medeiros Júnior (2009) já haviam descrito que no PB o pretérito imperfeito e o presente do indicativo atuam como operadores que licenciam a leitura genérica inclusiva das construções com sujeitos nulos indeterminados na 3ª pessoa do singular, como nos exemplos em (88) e (91), respectivamente, ao passo em que o pretérito perfeito atua como um operador que licencia a leitura arbitrária (episódica) nessas mesmas construções. Parece ainda ser uma questão em aberto se um adjunto adverbial no PB obrigatoriamente precisa ocupar a posição de especificador de TP nas construções com sujeitos nulos genéricos de leitura inclusiva e exclusiva, como ocorre nas construções com pronomes de leitura genérica inclusiva no finlandês. Como podemos verificar nos exemplos (93) e (94), pronomes nulos genéricos podem ser licenciados no PB sem a presença de um advérbio em posição pré-verbal.

(93) a. Matou um rapaz no show do Zezé di Camargo e Luciano ontem.
 b. Montou o armário lá em casa semana passada.
 c. Telefonou aí da CEB para você.
 d. Lava sofá. (LUNGUINHO & MEDEIROS JÚNIOR, 2009, p. 10)

⁶¹ No original “*a partial pro-drop language making use of special strategies to identify the empty category whenever null subjects must be interpreted referentially*” (FIGUEIREDO SILVA, 2000, p. 127).

(94) Capina, rastela, limpa quintais e bate veneno.⁶²

Ainda que não se identifique um elemento adverbial pré-verbal manifesto em (93) e (94), Pilati, Naves & Salles (2017), considerando parcialmente o trabalho de Holmberg (2010a), argumentam que existe uma âncora adverbial discursiva que licencia essas construções: em (93a-c), essa âncora é manifesta (“no show”, “lá em casa”, “aí da CEB”) e em (93d) e (94) é recuperada pragmaticamente, isto é, remete ao local ou ao agente a que a informação está associada por meio do contexto adequado de exposição. Crucialmente, o trabalho de Pilati, Naves & Salles (2017) aproxima o tratamento do PB ao tratamento proposto por Holmberg (2010a), para o finlandês, com a diferença de que no caso do finlandês o elemento adverbial ou expletivo (ou outro elemento) precisa ser realizado fonologicamente em [SPEC-TP], mas no PB essa realização fonológica em [SPEC-TP] não é obrigatória.

Ainda sobre os pronomes nulos genéricos no finlandês, Holmberg *et al.* (2009) mencionam que um DP objeto também pode satisfazer o traço-EPP de T nas construções em que eles são licenciados, como ilustrado em (95).

(95) Finlandês (HOLMBERG *et al.*, 2009, p. 63)
Nuorten mielipiteitä kuuluu arvostaa.
 jovens-GEN opiniões dever-3SG respeitar
 ‘Deve-se respeitar as opiniões dos jovens.’

Curiosamente, verificamos que um fenômeno similar ocorre no PB, com a diferença de que no PB o objeto frontado desencadeia a concordância com o verbo (o que não ocorre no finlandês), atuando, indiscutivelmente, como o sujeito da oração (cf. NEGRÃO & VIOTTI, 2010, 2015; para estudos sobre essas construções no PB, denominadas de *construções absolutas* pelas autoras).

(96) Objeto frontado no PB
 a. Apartamento vende bem em Goiânia.
 b. A cada um minuto quatro coisas vendem.

⁶² Dado registrado na caçamba de uma moto estacionada na feira municipal de uma cidade no interior de Goiás.

Holmberg (2010b) afirma haver um pronome genérico manifesto no finlandês (*sä*), cuja morfologia do pronome e do verbo é a da 2ª pessoa do singular, como visto em (97).

- (97) Finlandês (HOLMBERG, 2010, p. 203)
Sä saat töitä jos sä puhut saksa.
 você-SG conseguir-2SG trabalho se você-SG falar-2SG alemão
 ‘Você consegue um trabalho se você fala alemão.’

No PB, há também pronome genérico manifesto, mais precisamente dois pronomes: um correspondente à 2ª pessoa do singular (*você*) e outro correspondente à 1ª pessoa do plural (*a gente*). Interessantemente, a morfologia de concordância de ambos é idêntica à da 3ª pessoa do singular, e ambos foram introduzidos no paradigma pronominal do PB com essa interpretação genérica somente na segunda metade do século 20 (cf. M. E. DUARTE, 1993 [2018], 1995, 2000; KATO & DUARTE, 2014).

- (98) a. *Øgen* não pode fumar aqui. (KATO & DUARTE, 2014, p. 9)
 b. A gente não pode fumar aqui.
 c. Você não pode fumar aqui.

Os dados em (98), com ou sem pronome manifesto, são de leitura genérica inclusiva. Conforme constataram Naves & Borges (2015) por meio de dados diacrônicos, a leitura genérica exclusiva (arbitrária) no PB impede a inserção de pronomes lexicais na posição de sujeito, como ilustrado em (99):

- (99) a. Dia 15 depois da novena (*a gente/*você/*alguém/*o pessoal) **alevantou** mastro de Nossa Senhora do Carmo.
 b. Dia 29 eu Maria e Anica Macedo fomos no teatro, (*a gente/*você/*alguém/*o pessoal) **representava** *Direito por linhas torta*. (NAVES & BORGES, 2015, p. 85)

Os dados em (99) mostram, portanto, que a leitura arbitrária não permite a inserção de pronomes genéricos manifestos na posição de sujeito. Isso sugere que o sujeito genérico de leitura exclusiva licenciado pelo pretérito perfeito é semanticamente menos marcado para os traços de pessoa relacionados ao falante e ao ouvinte do que o sujeito nulo genérico de leitura

inclusiva, que possui uma leitura existencial e traços semânticos de pessoa que podem remeter genericamente tanto ao falante quanto ao ouvinte.

Em trabalho em que analisam a interface morfologia-sintaxe em línguas *pro-drop*, Holmberg & Roberts (2013) mostram que LSN consistentes, como o italiano, o árabe marroquino e o PE, não permitem sujeitos nulos indefinidos na 3ª pessoa do singular e que LSN parciais como o finlandês, o islandês e o PB permitem esse tipo de sujeito nulo. O primeiro argumento utilizado pelos autores para explicar essa observação é o de que as flexões dos verbos finitos nas LSN consistentes são semanticamente ricas e, por isso, não permitem a interpretação genérica do sujeito. Desse modo, os autores afirmam que, nessas línguas, sujeitos nulos definidos podem ser nulos e sujeitos genéricos devem ser manifestos, como ilustrado nos contrastes em (100a-b) e (101a-b).

(100) Italiano (HOLMBERG & ROBERTS, 2013, p. 120)

- a. __ Può controllare questo macchinario con una mano sola.
 Pode-3SG controlar-INF essa máquina com uma mão só
 ‘(Ele/ela) pode operar essa máquina com uma mão apenas.’
- b. Si può controllare questo macchinario con una mano sola.
 Se pode-3SG controlar-INF essa máquina com uma mão só
 ‘Pode-se operar essa máquina com uma mão apenas.’

(101) Árabe marroquino (HOLMBERG & ROBERTS, 2013, p. 120)

- a. ka-i-qder i-xeddem had l-makina b-yedd weHda.
 PRT-3M-pode 3M-operar essa máquina com-mão uma
 ‘Ele pode operar essa máquina com uma mão.’
- b. t-eqder t-xaddem had l-makina b-yedd weHda
 2-pode 2-operar essa máquina com-mão uma
 ‘Você pode operar essa máquina com uma mão.’

Holmberg & Roberts (2013) reconhecem, contudo, que a concordância rica não pode ser o único fator a ser considerado para a distinção (manifesto *versus* nulo) no licenciamento de sujeitos genéricos nas LSN consistentes e nas LSN parciais, afinal, o islandês distingue cinco formas no paradigma do presente do indicativo, duas no singular e três no plural (cf. exemplo em (29) deste capítulo, em que se constata um sincretismo entre a 2ª e a 3ª pessoa do singular no islandês). Já em relação ao finlandês, o paradigma verbal é ainda mais rico,

ênfatisam os autores, possuindo, como no italiano, seis formas verbais finitas para o presente do indicativo, três no singular (cantar: *laulan* (1SG), *laulat* (2SG), *laulaa* (3SG)) e três no plural (cantar: *laulamme* (1PL), *laulatte* (2PL), *laulavat* (3PL)). Diante dessas evidências, Holmberg & Roberts (2013), seguindo seus trabalhos anteriores (cf. HOLMBERG, 2005, 2010a, 2010b; ROBERTS, 2010), presumem que a diferença crucial entre essas línguas é que a morfologia dos verbos finitos em LSN consistentes como o italiano tem um componente de definitude, codificado pelos autores como um traço [*uD*], isto é, um traço de definitude (traço-D) não valorado. Esse traço está ausente no finlandês, no islandês e no PB. Holmberg & Roberts (2013) explicam que os traços- ϕ não especificados de T [*u- ϕ*] são valorados pelo sujeito em línguas com riqueza morfológica, porém, na ausência de um traço de definitude não especificado em T [*uD*], o valor de definitude do sujeito não pode ser copiado por T. Nas palavras desses autores, isso significa que esse sujeito não é uma cópia de T e por isso não pode ser deletado. Desse modo, os autores sugerem que um sujeito só é deletado em uma dada língua com riqueza morfológica se essa língua possui um traço [*uD*] em T. Por outro lado, se uma dada língua apresenta certa riqueza morfológica em seus verbos finitos, como o finlandês e o islandês, mas não possui um traço [*uD*] em T, seus sujeitos definidos devem ser manifestos. Com base nessas argumentações, Holmberg & Roberts (2013) formulam o seguinte postulado, relacionando sujeitos nulos, morfologia verbal rica e o licenciamento de sujeitos genéricos.

- (102) Se uma língua tem sujeito pessoal *pro-drop* em sentenças finitas e ativas, e tem concordância verbal rica, então essa língua não tem sujeito genérico *pro-drop* no mesmo contexto, e vice-versa.⁶³

(HOLMBERG & ROBERTS, 2013, p. 121, tradução nossa)

O postulado em (102) também prevê que, se uma língua tem sujeito genérico *pro-drop* em sentenças finitas e ativas, e tem concordância verbal rica, então essa língua não tem sujeito pessoal *pro-drop* no mesmo contexto, como é o caso do finlandês. Em resumo, a partir do que foi exposto, podemos pontuar que os pronomes nulos genéricos são uma das características definidoras das LSN parciais e diferem no PB e no finlandês no que concerne ao fato de que o PB não licencia pronomes expletivos nas construções em que esses sujeitos

⁶³ No original: “If a language has personal subject *pro-drop* in active, finite clauses, and has subject agreement, then it does not have generic subject *pro drop* in the same context, and vice versa.” (HOLMBERG & ROBERTS, 2013, p. 121).

são licenciados (mas o faz com pronome lexical de leitura genérica inclusiva, como em (98)) e possui sujeitos nulos tanto de leitura genérica inclusiva quanto de leitura genérica exclusiva, apresentando o tempo verbal como morfologia distintiva especial para distinguir essas interpretações (cf. LUNGUINHO & MEDEIROS JÚNIOR, 2009) e contrariando o argumento de Holmberg, Nayudu & Sheehan (2009) de que o PB apresenta ambas as leituras – inclusiva e exclusiva, ou existencial e arbitrária – sem qualquer morfologia verbal especial para essa distinção. Se as categorias especiais que satisfazem o EPP em [SPEC-TP] nas construções com sujeitos nulos genéricos no finlandês atuam da mesma forma no PB, é uma questão aparentemente em aberto, que retomaremos no próximo capítulo.

3.3 Síntese do capítulo

Neste capítulo, observamos que, desde Permluter (1971), a caracterização das LSN consistentes leva em conta o papel do paradigma verbal da língua em recuperar a referencialidade dos sujeitos nulos por meio de suas terminações flexionais. Nesse sentido, discutimos brevemente as primeiras abordagens teóricas para explicar o parâmetro do sujeito nulo, bem como apresentamos um resumo dos principais estudos mais atuais sobre a sintaxe dos sujeitos nulos. Pontuamos que, na teoria de regência e ligação, os sujeitos nulos foram tratados como uma variante morfofonológica de sujeitos manifestos. Assim, se uma dada língua teria ou não sujeitos nulos dependia de os núcleos flexionais relevantes serem capazes de satisfazer as condições de licenciamento e identificação desses sujeitos. A relação entre elementos nulos e fonologicamente realizados levou a se postular *pro* referencial e sua contrapartida expletiva. No decorrer do capítulo, todavia, elucidamos que o licenciamento de sujeitos nulos não implica o licenciamento obrigatório de expletivos nulos (finlandês) e que o licenciamento de expletivos nulos não implica o licenciamento de sujeitos nulos definidos na sintaxe (alemão). Mostramos que Huang (1984) contestou o papel da morfologia verbal rica no licenciamento e na identificação de argumentos nulos, tomando como evidência línguas asiáticas de tópico proeminente que permitem argumentos nulos, a exemplo do sujeito da sentença, desde que estejam vinculados à noção discursiva de tópico. Com a proposta de Huang (1984), uma profícua discussão em torno dos argumentos nulos em línguas sem qualquer marca de concordância verbal, como o chinês e o japonês, foi promovida na linguística gerativa (cf. TOMIOKA, 2003; NEELEMAN & SZENDRÖI, 2007; MIYAGAWA, 2010, 2017, 2018; SAITO & KIM, 2012; HOLMBERG & ROBERTS, 2013; entre outros).

Foi visto que, dentro do programa de pesquisa diacrônica gerativa, a relação entre perda morfológica e mudança sintática tem sido um ramo de investigação frutífero, especialmente sob o escrutínio da hipótese da concordância rica, que vincula a expressão fonológica das flexões dos verbos finitos a vários outros fenômenos sintáticos, como, por exemplo, movimento de V-para-T, sujeitos nulos e inversão livre. Tratar dessas questões, ainda que indiretamente, é condição *sine qua non* em estudos diacrônicos gerativos sobre a sintaxe dos sujeitos nulos. Vimos que, indiscutivelmente, as propriedades flexionais de um verbo frequentemente se correlacionam com a disponibilidade de sujeitos nulos. Porém, não existe uma noção universalmente válida de riqueza de concordância que sirva como condição suficiente para explicar o licenciamento de todos os tipos de sujeitos nulos. Conforme exemplificamos, em línguas orientadas para o discurso, como o chinês, os sujeitos nulos são identificados no discurso. Vimos também que, no geral, não há evidência convincente para a universalidade do EPP nem de *pro* como um requisito para preencher a posição [SPEC-TP] nas línguas naturais, principalmente naquelas que parecem licenciar sujeitos nulos mais livremente, como as LSN consistentes.

Verificamos que os objetivos essenciais dos estudos sobre a sintaxe dos sujeitos nulos na gramática gerativa são: verificar a hipótese do EPP e explicar as condições de licenciamento e de interpretação dos sujeitos nulos por meio da recuperabilidade de seus traços- ϕ , independentemente de o sujeito nulo ser *pro* referencial, PRO ou uma variável ligada, ou se as características referenciais são recuperadas por meio de um traço-D em um núcleo funcional temporal (cf. CHOMSKY, 1982, 1986; RIZZI, 1982, 1986; JAEGGLI & SAFIR, 1989; ALEXIADOU & ANAGNOSTOPOULOU, 1998; HOLMBERG, 2005, 2010a, 2010b; ROBERTS, 2010; SAAB, 2016; P. BARBOSA, 2018; entre outros). Resumidamente, esses estudos tentam explicar a distribuição de sujeitos nulos em termos de licenciamento e identificação e supõem que: *pro* em [SPEC-TP] ou um traço-D em T satisfazem o EPP.

Apresentamos quatro tipos de LSN dentro do escopo do programa minimalista, quais sejam: as LSN expletivo, as LSN orientadas para o discurso, as LSN consistentes e as LSN parciais. A partir dessa abordagem, mostramos que em relação às LSN consistentes, há duas análises predominantes: (a) pronomes são deletados sob adjacência com T nas línguas com morfologia verbal rica e um traço-D em T, isto é, nas LSN consistentes (cf. HOLMBERG, 2005, 2010a; ROBERTS, 2010); (b) a flexão de concordância tem valor pronominal, ou seja, é uma categoria D, a partir do que se pode postular que o movimento de V-para-T funde a categoria D ao verbo e satisfaz o traço EPP de T nas LSN consistentes, mas não exclusivamente nelas (cf. ALEXIADOU & ANAGNOSTOPOULOU, 1998). Ambas as

análises corroboram com a intuição presente no trabalho seminal de Permulter (1971). Em relação à configuração das LSN parciais, tratamos da proposta de Holmberg (2005, 2010a), que, em síntese, propõe que a operação que gera sujeitos nulos definidos na 3ª pessoa do singular nas LSN consistentes gera um sujeito nulo indefinido na 3ª pessoa do singular nas LSN parciais, pois, nessas línguas, o traço de definitude está ausente do conjunto de traços de T. A título de ilustração, vimos algumas semelhanças e divergências entre os sujeitos nulos genéricos no PB e no finlandês.

Com base na revisão teórica e descritiva dos principais estudos sobre a sintaxe dos sujeitos nulos no PM desenvolvida neste capítulo (cf. ALEXIADOU & ANAGNOSTOPOULOU, 1998; HOLMBERG, 2010a; NEELEMAN & SZENDRÖI, 2007; ROBERTS, 2010; SAITO & KIM, 2012; SAAB, 2016; P. BARBOSA, 2018; dentre outros), mostramos que há uma intrínseca relação entre morfologia, sintaxe e semântica no licenciamento e na identificação de sujeitos nulos. Os estudos de Alexiadou & Anagnostopoulou (1998), Holmberg (2005, 2010a), Roberts (2010) e Holmberg & Roberts (2013), indo além da dicotomia morfologia-sintaxe, por exemplo, relacionam um traço semântico de definitude e os traços formais de pessoa, número e gênero presentes nos verbos finitos ao licenciamento de sujeitos nulos definidos. Holmberg (2005, 2010a) e Holmberg & Roberts (2013) relacionam, ainda, a ausência do traço de definitude e traços- ϕ ao licenciamento de sujeitos indefinidos – manifestos ou nulos – de leitura genérica inclusiva ou exclusiva, nas LSN consistentes e nas LSN parciais. Alexiadou & Anagnostopoulou (1998) e Holmberg (2005, 2010a) mostram que as flexões de um verbo podem ser tomadas como categorias sintáticas das orações (viz. o sujeito), embora sejam realizadas como partes de um verbo. O estudo de Alexiadou & Anagnostopoulou (1998), mais especificadamente, mostra que certos aspectos da sintaxe de uma oração, como a ordem das palavras, podem estar correlacionados com a propriedade semântica de definitude e com as propriedades de flexão dos verbos finitos nas LSN canônicas.

Destacamos, por fim, que, em nossa análise no Capítulo 6, serão essenciais a proposta de Alexiadou & Anagnostopoulou (1998) para a ordem VS nas LSN consistentes e a proposta de Holmberg (2005, 2010a) em relação ao papel do traço-D em T nas LSN consistentes e nas LSN parciais. É a partir de uma combinação das implementações dessas propostas que delinearemos nossa análise para a mudança na sintaxe do sujeitos nulos entre os séculos XVIII e XIX no dialeto goiano. O postulado de Roberts (2010) em (74) também será essencial para a noção de empobrecimento de T que adotamos em nossa análise.

4 UM PANORAMA DA SINTAXE DO SUJEITO NO PORTUGUÊS BRASILEIRO

Nas últimas três décadas, houve um número extensivo de pesquisas gerativas sobre mudanças na gramática do PB (cf. ROBERTS & KATO (1993) [2018]; KATO & NEGRÃO, 2000; GALVES, 2001; CASTILHO *et al.*, 2007; NUNES, 2009; KATO & NASCIMENTO, 2015; M. DUARTE, 2012; KATO & ORDOÑEZ, 2016; PILATI, SALLES & NAVES, 2017; CYRINO & TORRES MORAIS, 2018; entre outros). Apresentamos, neste capítulo, os achados de alguns desses estudos sobre a gramática do PB na perspectiva da mudança linguística. A primeira seção do capítulo relata uma série de inovações na sintaxe do sujeito no PB, especialmente aquelas que estão vinculadas a alguma caracterização ou propriedade especial das línguas de sujeito nulo, conforme vimos no capítulo precedente. A segunda seção dedica-se especificamente às construções com *se*-passivo e *se*-impessoal no PB e no PE. Consoante com a afirmação de Galves (1986) de que sobre o *se* ainda correrá muita tinta, nossa intuição é a de que, de fato, o estudo dessas construções ainda tem muito a nos dizer sobre a gramática do PB.

4.1 Um mosaico de mudanças na sintaxe do sujeito no PB

4.1.1 *Sujeitos nulos versus sujeitos manifestos*

O surgimento de construções sintáticas inovadoras no PB, nas quais se constata um maior preenchimento da posição de sujeito por elementos de natureza (pro)nominal, a exemplo das construções de tópico-sujeito em (1), nas quais um verbo inergativo ou inacusativo aparece acompanhado de um NP em posição pré-verbal, aparentemente extraído de um PP, sem qualquer marca morfológica que indique alguma modificação na projeção de sua estrutura argumental, provocou uma série de estudos funcionalistas e gerativistas sobre a sintaxe do sujeito no PB (cf. PONTES, 1981, 1986, 1987; GALVES, 1984, 1998).

- (1) a. *O Carlos André* cresceu o nariz.
 b. *A Sarinha* tá nascendo dente.
 c. *Essa casa* bate bastante sol.
 d. *Esse caro* cabe 60l de gasolina. (PONTES, 1986, p. 18)

Os estudos seminais de Pontes (1981, 1986, 1987) identificaram que o PB apresentava, seguindo a terminologia de Li & Thompson (1976), propriedades de línguas de tópico proeminente. Pontes (1986) chama a atenção para o fato de que o primeiro NP no conjunto de dados em (1) está topicalizado. Pontes (1986) argumenta que esses NPs são o sujeito frásico das orações, tendo em vista: (i) a ausência da pausa prosódica usada para assinalar fonologicamente o tópico; (ii) a ordem SVO das sentenças; (iii) a agramaticalidade da anteposição do sujeito caso o NP pós-verbal fosse o sujeito (2); (iv) a agramaticalidade da sentença em caso de concordância entre o verbo e o NP pós-verbal (3); (v) a impossibilidade de apassivação dessas sentenças, considerando o primeiro NP como objeto (4).

- (2) a. *?A Sarinha dente tá nascendo. (PONTES, 1986, p. 18)
 b. *?Essa casa bastante sol bate.
 c. *?Esse carro 60l de gasolina cabe(m).
 d. ?O Carlos André o nariz cresceu.
- (3) a. *A Sarinha tão nascendo dentes. (PONTES, 1986, p. 18)
 b. *Os meninos cresceu o nariz.
 c. *O carro furaram os pneus.
 d. *A Belina cabem 60l de gasolina.
- (4) a. *O nariz foi crescido pelo Carlos André. (PONTES, 1986, p. 19)
 b. *A casa é batida pelo sol.
 c. *O dente é nascido pela Sarinha.
 d. *A gasolina é cabida pela Belina.

Para Pontes (1981, 1986), construções de tópico-sujeito não poderiam ser explicadas por meio de um modelo gerativo da linguagem. Em contraposição a essa suposição, Galves (1984, 1998) propôs que as inovações na gramática do PB, incluindo aquela exemplificada em (1), estavam relacionadas à caracterização dessa língua, nos termos gerativos de Huang (1984), como uma LSN orientada para o tópico/discurso – consoante Huang (1984), a estrutura básica das orações em línguas orientadas para o discurso não é sujeito-predicado, isto é, a estrutura básica da oração de uma língua orientada para o discurso favorece a função

informativa dos constituintes, como tópico e foco.⁶⁴ Galves (1998) argumenta que mudanças no paradigma flexional do PB resultaram no requerimento do preenchimento da posição de sujeito, mesmo com os chamados sujeitos não canônicos. Para Galves (1998), há na morfologia verbal do PB uma ausência de distinção morfológica entre 2ª e a 3ª pessoa que legitima não somente construções de tópico-sujeito, mas também construções com verbo transitivo e o DP argumento interno na posição de sujeito, como em (5).

(5) A balança está consertando. (GALVES, 1998, p. 19)

Galves (1998) sugere que o comportamento das estruturas de topicalização do PB está associado a uma categoria funcional denominada por ela de Pessoa e assume que essa categoria possui traços- ϕ que requerem ser checados por um movimento invisível de traços pronominais. Galves (1998, p. 27) exemplifica sua proposta em (6).

(6) [_{PersP} Essa competência [_{PersP} ϕ_i [_{TP} *ela*_i é de natureza mental]]]

Nessa estrutura, de acordo com Galves (1998), os traços do pronome *ela* se movem para checar os traços- ϕ de Pessoa e o DP *essa competência é* é adjungido à projeção de Pessoa. Galves (1998) enfatiza que sua teoria sobre a checagem dos traços de Pessoa nessas construções prediz que, mesmo diante da ausência de um pronome lexical, haverá um pronome nulo na posição de sujeito, como ilustrado em (7).

(7) [_{PersP} DP [_{PersP} ϕ_i [_{TP} *pro*_i ...]]] (GALVES, 1998, p. 27)

Galves (1998) propõe que a projeção de Pessoa também deve ser a projeção de Tempo, tendo em vista que os traços- ϕ nas estruturas de tópico-sujeito não podem ser checados numa relação de especificador-núcleo. Para que a derivação não fracasse, a única opção viável é a de que os traços- ϕ do verbo se movam para Pessoa, sendo checados pelos traços- ϕ dessa categoria. Tomando essa direção, a proposta de Galves (1998) legitima o movimento de um NP para a categoria Pessoa, sugerindo a estrutura subjacente das construções de tópico-sujeito em (8):

⁶⁴ Galves (1984) fez sua primeira afirmação de que o PB era uma LSN orientada para o discurso em um estudo cuja análise se centrava nas assimetrias no licenciamento de pronomes lexicais e de categorias vazias nas posições de sujeito e objeto. Galves (1998), por sua vez, fornece uma análise centrada em construções com tópico-sujeito.

(8) [PersP NP_j [Pers' φ_i [TP V_i t_j]]] (GALVES, 1998, p. 27)

Em (8), a autora demonstra que o NP pré-verbal está numa relação de especificador-núcleo com Pessoa, uma vez que ele checa o traço AGR que Pessoa, pelo movimento dos traços-φ do verbo, herdou de Tempo. Por conta dessa configuração especificador-núcleo dos traços-φ do NP e do verbo, de acordo com Galves (1998), é que é estabelecida a concordância morfológica entre o verbo e o NP, diferentemente do que acontece nas estruturas com pronome lembrete. Galves (1998) contrasta (8) com a estrutura em (9), que representa as construções de tópico-sujeito com pronome lembrete.⁶⁵

(9) [PersP esta casa [PersP φ_i [TP pro_{expl} bate sol nela_i]]] (GALVES, 1998, p. 30)

Adotando uma proposta diferente da de Galves (1998) para o parâmetro do sujeito nulo no PB, M. E. Duarte (1993 [2018], 2000) propôs que o PB estaria num processo de mudança de uma língua *pro-drop* para uma língua não *pro-drop*, devido especialmente à reestruturação dos paradigmas pronominal e verbal da língua. Com base em 1100 sentenças coletadas de sete peças de teatro brasileiras populares, uma de cada um dos períodos representados na Tabela 1, o *corpus* de M. E. Duarte (1993 [2018], 2000) mostra um aumento significativo entre os anos de 1937 e 1992 no número de sujeitos manifestos em contexto de retomada pronominal (forma plena *versus* forma nula de preenchimento do sujeito pronominal).

Tabela 1: Sujeitos pronominais manifestos na diacronia do PB.

Período	1845	1882	1918	1937	1955	1975	1992
Porcentagem	20%	23%	25%	46%	50%	67%	74%

Fonte: M. E. Duarte (1993 [2018], 2000).

De acordo com M. E. Duarte (1993 [2018], 2000), os resultados acima mostram a perda do princípio Evite Pronome, que, segundo Chomsky (1981), consiste na postulação de que se um pronome nulo é possível, ele deve ser preferido no lugar do pronome manifesto. M.

⁶⁵ Para estudos recentes sobre as construções com tópico-sujeito no PB, remetemos o leitor à leitura de Lunguinho (2006), Munhoz (2011), Munhoz & Naves (2012), Andrade & Galves (2014), Avelar (2015), Kato (2015) e Andrade (2018).

E. Duarte (1993 [2018]) menciona que nas peças de teatro populares que compõem seu *corpus* o pronome *tu* não aparece nos dados referentes aos períodos de 1937, 1955 e 1975, períodos em que a autora constata uma queda substancial no licenciamento de sujeitos nulos no contexto da retomada pronominal. A autora destaca, contudo, que, apesar de seu *corpus* não apresentar dados do uso de *tu* nesses três períodos analisados, foi constatado que, na peça de teatro *A vida em três andares*, escrita em 1938 por um autor não muito popular chamado Humberto Couto, os pronomes *tu* e *você* são utilizados de modo intercambiável. O percentual geral do uso de sujeitos nulos nessa peça foi de 68%, o de pronomes nulos da 2ª pessoa *tu* foi de 89% e o de pronomes nulos da 2ª pessoa *você* foi de 45% (M. E. DUARTE, 1993 [2018]). Para a autora, esses resultados evidenciam que o uso intercambiável dos pronomes *tu* e *você* estavam presentes na variedade falada do PB no Rio de Janeiro nos períodos em que o uso do *tu* não foi encontrado nas peças de teatro populares.

M. E. Duarte (2000) afirma que a categoria pessoa exerce um papel determinante no preenchimento ou não do sujeito ao longo do período analisado: a 3ª pessoa é um pouco mais resistente ao preenchimento da posição de sujeito em comparação com a 1ª e a 2ª. Considerando o período inicial e final das peças analisadas (1845 e 1992), o resultado obtido pela pesquisadora sobre esse tópico é o seguinte:

- (10) Forma plena do sujeito de acordo com a categoria pessoa
- a. 1ª pessoa: de 31 a 82%.
 - b. 2ª pessoa (desinência não distintiva): de 7 a 78%.
 - c. 3ª pessoa: de 17 a 45%. (M. E. DUARTE, 1993 [2018], p. 89-92)

M. E. Duarte (1993 [2018]) afirma, ainda, que a reestruturação do paradigma verbal do PB foi iniciada com a perda do pronome *vós* e pela inserção dos pronomes *você* (2ª pessoa do singular) e *a gente* (1ª pessoa do plural) no sistema pronominal do PB. A fusão subsequente dos pronomes *você* e *a gente* em termos de restrição de significado/referência com os pronomes *tu* e *nós*, respectivamente, teria tornado essas formas pronominais intercambiáveis no início do século XX.⁶⁶ Um padrão semelhante é observado para a segunda pessoa do plural: o pronome *vós* desapareceu e foi substituído pelo pronome *vocês*. O Quadro 1 ilustra as observações de M. E. Duarte (1993 [2018]).

⁶⁶ Estudos sobre a diacronia do pronome *você* no PB (cf. LOPES & CAVALCANTE, 2011; LOPES, RUMEU & CARNEIRO, 2013; RUMEU, 2013; entre outros) mostram que o uso inicial desse pronome como pronome pessoal se deu no fim do século XIX.

Quadro 1: Evolução nos paradigmas flexionais do PB por M. E. Duarte (1993 [2018]).⁶⁷

	PRONOMES NOMINATIVOS	PARADIGMA 1 SÉCULO XIX	PARADIGMA 2 SÉCULO XX/1	PARADIGMA 3 SÉCULO XX/2
1PS	eu	<i>canto</i>	<i>canto</i>	<i>canto</i>
1PP	nós <i>a gente</i>	<i>cantamos</i> -	<i>cantamos</i> <i>canta</i> ∅	<i>cantamos</i> <i>canta</i> ∅
2PS	tu <i>você</i>	<i>cantas</i> -	<i>cantas</i> <i>canta</i> ∅	<i>canta(s)</i> <i>canta</i> ∅
2PP	vós <i>vocês</i>	<i>cantais</i> -	- <i>cantam</i>	- <i>canta(m)</i>
3PS	ele, ela	<i>canta</i> ∅	<i>canta</i> ∅	<i>canta</i> ∅
3PP	eles/elas	<i>cantam</i>	<i>cantam</i>	<i>canta(m)</i>

Fonte: M. E. Duarte (1993 [2018]), p. 85).

A partir do Quadro 1, M. E. Duarte (1993 [2018]) relaciona o empobrecimento do paradigma verbal no PB à reconfiguração do quadro pronominal. Esses fenômenos teriam sido, portanto, o gatilho para a mudança em curso no PB na visão da autora. Galves (1993 [2018]) sugere que o empobrecimento do paradigma flexional do PB teria provocado não somente a perda do princípio Evite Pronome, mas também teria reestruturado o padrão sentencial do PB, que teria passado de uma língua voltada para a sentença a uma língua voltada para o discurso.

Contrárias às propostas de Duarte (1993 [2018], 2000) e Galves (1993 [2018]), Negrão (1999) e Negrão & Viotti (2000) argumentam que foram as propriedades discursivas do PB que minaram a função que outrora era desempenhada pelo paradigma flexional dos verbos finitos na orientação da língua para o sujeito; assim o empobrecimento da concordância não estaria diretamente relacionado com o preenchimento da posição de sujeito por meio da realização de pronomes plenos no PB. As autoras afirmam que, se o empobrecimento da concordância verbal fosse o gatilho para a perda do princípio Evite Pronome, era de se esperar que o aumento desses sujeitos ocorresse mais robustamente nos casos em que as formas verbais não se distinguem, isto é, nos casos de 2^a e 3^a pessoa. Desse modo, para as autoras, os achados de M. E. Duarte (1993) mostrados em (10), em que há um aumento mais substancial da frequência dos pronomes plenos com a 1^a e a 2^a pessoa, desafiam a hipótese de que o empobrecimento do paradigma verbal tenha desempenhado um papel tão

⁶⁷ Com o trabalho de M. E. Duarte (1993 [2018]) republicado, o Quadro 1 é a versão mais atual da proposta da autora a respeito da evolução dos paradigmas flexionais do PB.

decisivo nas estratégias de preenchimento (pro)nominal da posição de sujeito no PB.⁶⁸ Negrão & Viotti (2000) defendem que as mudanças no uso do sujeito pronominal indicam uma especialização de formas, e não a substituição do sujeito nulo pelo sujeito pronominal. Elas afirmam que há no PB contextos em que os sujeitos nulos são usados como pronomes ligados e os pronomes plenos como sujeitos referenciais; ademais, a gramática do PB apresenta assimetria na interpretação de sujeitos nulos e manifestos em orações encaixadas, como em (11).⁶⁹

- (11) a. Só o Maluf₁ acha que ₁ vai ganhar as eleições.
b. Só o Maluf₁ acha que ele₁ vai ganhar as eleições.

(NEGRÃO & VIOTTI, 2000, p. 110)

Em (11a), Maluf é a única pessoa que acredita que vai ganhar as eleições, de modo que (11a) seria falsa se outro candidato também acreditasse na própria vitória. Em (11b), Maluf é a única pessoa que acredita na vitória dele nas eleições, de modo que seria falsa se mais alguém acreditasse na vitória de Maluf. Parece, então, que algo mudou no PB em relação ao parâmetro do sujeito nulo, e não se trata apenas da substituição de formas pronominais nulas por formas pronominais fonologicamente representadas.

M. E. Duarte (2000) mostra que não somente pronomes expressos tornaram-se mais usuais, mas surgiu uma nova tendência na gramática do PB de preenchimento da posição de sujeitos não referenciais, como na comparação entre (12) e (13), onde temos: em (13a), a posição de sujeito de uma construção com verbo impessoal preenchida pelo pronome genérico *você*; em (13b), pronomes lexicais na posição de sujeito de uma construção com verbo de alçamento; em (13c), o movimento do complemento de uma infinitiva para a posição de sujeito; em (13d), a expansão da estrutura argumental de um verbo que, então, projeta um pronome lexical na posição de sujeito.

- (12) a. Não há/tem clientela no centro da cidade. (M. E. DUARTE, 2000, p. 31-32)
b. Parece que vocês não pensam a sério a vida.

⁶⁸ Cyrino, Duarte & Kato (2000) propuseram uma hierarquia referencial para descrever o uso das formas plenas dos sujeitos de 3ª pessoa no PB. Com base nessa proposta e em 1145 dados de peças de teatro, Duarte, Mourão & Santos (2012) constatam que quanto [+humano] e [+específico] forem os traços semânticos do antecedente do sujeito de 3ª pessoa maior será a probabilidade de o sujeito ser expresso. O traço [-específico] do DP antecedente é o fator mais resistente na implementação de sujeitos referenciais manifestos, de acordo com os autores.

⁶⁹ Modesto (2000) também declara que o PB é uma língua orientada para o discurso e argumenta que a identificação do sujeito nulo nessa língua ocorre por meio de uma ligação a uma posição não argumental.

- c. Leva tempo para aprender essas coisas.
- d. Vale a pena salvá-los.

- (13) a. Você não tem clientela no centro da cidade. (M. E. DUARTE, 2000, p. 31-32)
- b. Vocês₁ parecem que *pro*₁/vocês₁ não pensam a sério na vida.
 - c. Essas coisas levam tempo para aprender.
 - d. Eles valem a pena ser salvos.

Outro exemplo na direção da mudança no funcionamento dos padrões gramaticais do PB são as construções com duplicação de sujeitos deslocados à esquerda, que ocorrem tanto com sujeitos definidos (14a) quanto com sujeitos de leitura genérica (14b).

- (14) a. A Clarinha_i ela_i cozinha que é uma maravilha. (*PE/PB)
- b. Você, no Canadá, você pode ser o que você quiser. (*PE/PB)
- (M. E. DUARTE, 2000, p. 28)

Pires (2008) observa que o PB é crucialmente diferente de outras línguas que permitem sujeitos nulos, como o espanhol, o grego e o italiano, em relação às propriedades dos sujeitos pré-verbais. Seguindo parcialmente a análise de Ordóñez (1997) de que os sujeitos nas línguas românicas (com morfologia verbal rica) ocupam [SPEC-TP] e de que os sujeitos pré-verbais em espanhol são constituintes periféricos à esquerda, isto é, especificadores de um núcleo funcional topicalizado [SPEC-TopP], Pires (2008) argumenta que DPs argumentais ou interpretados como argumento do verbo deslocados à esquerda da sentença podem ser gerados por meio de dois mecanismos no PB: em posição A e em posição A-barra. No primeiro caso, um DP argumental se move para [SPEC-TP] para valorar seu Caso nominativo e, ao mesmo tempo, acionar a concordância de seus traços- ϕ sobre o núcleo temporal da sentença. Pires (2008) salienta que os DPs que se deslocam para [SPEC, TP] no PB podem ser o argumento externo de verbos transitivos e não-ergativos, ou o argumento interno de verbos inacusativos e passivos, considerando a análise padrão de movimento-A no modelo de princípios e parâmetros. No segundo caso, um DP interpretado como argumento do verbo é realizado dentro do domínio de CP da oração, sendo pronunciado como um elemento deslocado à esquerda. Em (15) e (16), Pires (2008) resume esses mecanismos.

(15) Sujeitos manifestos em [SPEC-TP]: sujeitos em posição A
 DPs de argumento que se movem para [SPEC-TP] para valorar seu Caso nominativo e, ao mesmo tempo, acionar a concordância de seus traços- ϕ (pessoa e número) sobre o núcleo flexional da oração. (PIRES, 2008, p. 115)

(16) DPs manifestos no domínio de CP: DPs em posição A-barra
 DPs interpretados como argumentos do verbo, mas que são realizados dentro do domínio de CP da oração. Eles são pronunciados como elementos deslocados à esquerda. (PIRES, 2008, p. 116)

Pires (2008) afirma que o PB é diferente de línguas românicas como o espanhol (cf. ORDÓÑEZ, 1997) no que diz respeito ao fato de que um pronome lembrete manifesto pode ocupar [SPEC-TP] ao mesmo tempo que um DP interpretado como argumento do verbo ocupa [SPEC-TopP], como ilustrado em (17).

(17) [TopP A Silvia_i, [TP ela_i [T^v saiu cedo]] (PIRES, 2008, p. 120)

As evidências de elipse do VP com leitura *strict* na forma afirmativa em (18a) e na forma negativa em (18b) também fornecem suporte à proposta adotada por Pires (2008) de que pronomes lembretes podem ocorrer em posição de sujeito ao mesmo tempo em que um DP interpretado como argumento do verbo ocupa uma posição de tópico à esquerda.

(18) a. A Silvia, ela saiu cedo e [TopP a Paula, [TP ela também]] (PIRES, 2008, p. 121)
 b. A Silvia, ela saiu cedo e [TopP a Paula, [TP ela não]]

Pires (2008) afirma que os dados acima não mostram se um DP completo pode ocorrer como parte dos fragmentos remanescentes da elipse de VP. O pesquisador argumenta, contudo, que é possível que um DP completo ocupe [SPEC-TP] nas estruturas com fragmentos remanescentes da elipse de VP no PB, embora esse sujeito manifesto enfrente restrições em relação à correferência nessa elipse, como ilustrado pelo autor em (19).

(19) [TopPVocês, [TP_{POS} policiais federais_i vão interrogar]], mas [TopP_A Carla, [TP_{POS} bastardos_{i/*k} /eles_{i/??k} não vão]] (PIRES, 2008, p. 122)

Na oração principal em (19), Pires (2008) mostra que um DP completo (*os policiais federais*) pode ocorrer como um sujeito manifesto em [SPEC-TP] logo após um objeto em [SPEC-TopP]. Adicionalmente, na mesma estrutura, um pronome restritivo ou um epíteto (*bastardos*) pode ocorrer em adição a um DP deslocado à esquerda como parte dos fragmentos remanescentes da elipse de VP. Pires (2008), por fim, enfatiza que a única restrição aparente em (19) é que tanto o pronome quanto o epíteto nos fragmentos remanescentes da elipse devem ocorrer em correferência com o sujeito da oração antecedente. Em suma, os dados apresentados por Pires (2008) mostram que tanto [SPEC-TP] quanto [SPEC-TopP] estão atualmente disponíveis como posições para DPs interpretados como argumento interno de um verbo no PB.

Figueiredo Silva (1996, 2000) observou que sujeitos nulos definidos de 3ª pessoa em contextos pragmaticamente neutros não são licenciados pela morfologia do verbo no PB. Isso quer dizer que sentenças como em (20b) só são licenciadas no PB caso haja algum tópico no discurso que possa ser identificado como o sujeito da oração.

- (20) a. Encontrei a Maria ontem. (FIGUEIREDO SILVA, 2000, p. 134)
 b. *Encontrou a Maria ontem

Figueiredo Silva (1996, 2000) observou, ainda, que sujeitos nulos no PB estão sujeitos a um requisito de localidade não observado no PE. A autora mostra que há sujeitos nulos referenciais no PB de dois tipos, cada um definido de acordo com a posição de seu antecedente: sujeito nulo anafórico, como em (21a), em que o antecedente está em uma posição A; e sujeito nulo variável, como em (21b), em que o antecedente, por ser um tópico, está em uma posição A-barra.

- (21) a. João₁ disse que e_1 comprou um carro (FIGUEIREDO SILVA, 2000, p. 135)
 b. A Maria₁, o João disse que e_1 comprou um carro.

Figueiredo Silva (2000) conclui, assim, que a correferência entre o sujeito nulo de uma oração encaixada e o sujeito da oração principal não é obrigatória no PB. Contrários a essa proposta, Ferreira (2000, 2009) e C. Rodrigues (2000, 2004) afirmam que sujeitos nulos de orações finitas incorporados imitam o comportamento de PRO, obrigatoriamente

controlado no PB.⁷⁰ Seguindo a teoria do controle do movimento de Hornstein (1999), os autores argumentam que os sujeitos nulos nessas orações são traços de movimento A, assim como o PRO controlado o *é*, e assumem, com Chomsky (2000, 2001), que o Caso estrutural congela DPs para propósitos de movimento A e que o sujeito de uma oração finita permanece sem Caso quando um T sem traços- ϕ é selecionado; assim, o sujeito de uma oração encaixada pode passar por um movimento A para a posição de sujeito da matriz, a fim de receber Caso. Nesses termos, sentenças como (22a) são derivadas como em (22b), onde o T encaixado é incompleto, ou seja, sem traços- ϕ , e o T da frase matriz é completo, isto é, com traços- ϕ . O sujeito encaixado, ou seja, o DP *João* em (22b) move-se da oração encaixada para o vP da oração matriz e entra em uma relação temática com seu predicado; posteriormente, move-se para o [SPEC-TP] da oração matriz, onde recebe Caso estrutural.

- (22) a. O João₁ disse que *e*₁ viu a Maria (C. RODRIGUES, 2004, p. 116)
 b. [TP [DP o João_{Caso}]₁ [T- ϕ] [T' [vP *t*₁ [v' [VP disse [CP que [C' [TP *t*₁ [T' [T- ϕ -def] [vP *t*₁] [v' [VP viu a Maria]]]]]]]]]]]]]

Analisando dados similares a (22), Modesto (2000, 2007, 2008), um dos defensores da abordagem de que o PB é uma língua orientada para o discurso, chega a conclusões diferentes das de Ferreira (2000, 2009) e C. Rodrigues (2000, 2004) sobre o PB. Modesto (2000, 2007, 2008) analisa o sujeito nulo de orações finitas encaixadas e argumenta que há uma interpretação ambígua do sujeito nulo dessas orações: na maioria dos casos, o sujeito nulo da encaixada deve ser interpretado como correferente ao sujeito da oração matriz, como em (23) e (25a); entretanto, o sujeito da oração matriz deixa de ser o antecedente quando um sujeito da oração subordinada é questionado, como ilustrado em (24) e (25b).

- (23) a. A Dani₁ convenceu o Feco₂ que *e*_{1/*2} pode se eleger. (MODESTO, 2007, p. 235)
 b. A Dani₁ avisou o Feco₂ que *e*_{1/*2} precisa trabalhar até mais tarde.

- (24) Quem₂ que a Dani₁ convenceu *t*₂ que *e*_{*1/2} pode se eleger? (MODESTO, 2007, p. 238)

- (25) a. [FP a Dani₁ [TP *t*₁ convenceu o Feco₂ [CP que [FP *e*₁ [TP *t*₁ pode se eleger]]]]]

⁷⁰ Apesar de chegarem a conclusões bastante similares, os trabalhos de Barra Ferreira (2000) e Rodrigues (2000) foram desenvolvidos independentemente. Rodrigues (2004), por sua vez, constitui o primeiro estudo sobre as propriedades do PB como uma LSN parcial.

b.[_{CP} Quem₂ que [_{FP} t₂ [_{TP} a Dani₁ convenceu t₂ [_{CP} que [_{FP} e₂ [_{TP} t₂ pode se eleger]]]]]]]
 (MODESTO, 2007, p. 243-244)

Modesto (2007) argumenta que o sujeito nulo no PB é um ϕ P (cf. HOLMBERG, 2005) e, por isso, ele não refere por si só e só pode ser interpretado quando ligado. Nessas condições, Modesto (2007) enfatiza que o PB é uma língua de tópico proeminente, de modo que os sujeitos nulos são movidos para uma posição A-barra, qual seja, [SPEC-FP], que permite ϕ Ps incorporados na posição de sujeito. Para o autor, sua proposta explicaria por que línguas do tipo do inglês não mostram sujeitos incorporados nulos como em (23) e (24), afinal, os sujeitos em línguas como o inglês devem permanecer em [SPEC-TP] e, portanto, não podem vincular uma variável a um sujeito numa oração encaixada. Figueiredo Silva (2017, p. 196) salienta que a implementação de movimento A ou de dependência A-barra para explicar o sujeito nulo em orações encaixadas no PB ainda está “*sujeita a debate, na dependência inclusive dos desdobramentos teóricos do PM.*”

De acordo com Duarte & Figueiredo Silva (2016), o PB e o PE comportam-se da mesma forma no que diz respeito ao estatuto referencial dos sujeitos nulos de 3ª pessoa em orações encaixadas: diferentemente dos pronomes manifestos de 3ª pessoa, eles devem procurar por um antecedente A ou A-barra para serem interpretados. No entanto, eles se comportam de maneira diferente em relação ao princípio Evite Pronome; e o PE proíbe a correferência entre um pronome manifesto encaixado e o sujeito de uma oração matriz (26).

(26) O João disse que ele comprou o carro.

PE: *o João_i/ele_i

PB: \surd o João_i/ele_i (DUARTE & FIGUEIREDO SILVA, 2016, p. 240)

4.1.2 Questões de ordem: inversão livre do sujeito e movimento do verbo

Embora a inversão livre do sujeito possa ser um tanto restrita no PE (cf. ZUBIZARRETA, 1982; BARBOSA, DUARTE & KATO, 2005; COSTA & FIGUEIREDO SILVA, 2006; DUARTE & FIGUEIREDO SILVA, 2016), são surpreendentes as mudanças relativas à ordenação dos elementos sintáticos no PB. Berlinck (2000), por exemplo, mostra uma perda quase total da inversão livre do sujeito na história do PB com verbos transitivos. Os dados diacrônicos de Berlinck (2000) mostram como a frequência da ordem SV no PB

aumentou de 42% no início do século XIX para 96% na segunda metade do século XX. O oposto tem sido o caso das ordens com sujeitos pós-verbais, que quase desapareceram. As ordens VSX caíram de 34% para 2% e as ordens VXS de 24% para 2%. A Tabela 3 apresenta os resultados da pesquisa de Berlinck (2000) divididos em seis períodos inseridos no decurso dos séculos XIX e XX. Ressalta-se que na Tabela 2 não constam os dados com verbos inacusativos.

Tabela 2: Frequência da ordenação dos elementos sintáticos em Berlinck (2000)

Período de tempo	I	II	III	IV	V	VI
Ordem das palavras:						
SV	42%	76%	90%	89%	94%	96%
VSX	34%	14%	6%	5%	2,5%	2%
VXS	24%	10%	4%	6%	3,5%	2%

Fonte: Berlinck (2000, p. 183)

De acordo com Berlinck (2000), a ordem VS no PB resiste especialmente com os verbos inacusativos. Nos dados da pesquisadora, a média de construções com verbos inacusativos e ordem VS mantém-se praticamente invariável no decorrer dos seis períodos analisados, apesar de a porcentagem de construções com a ordem SV com esses verbos ter crescido nesse período.⁷¹

Santos & Soares da Silva (2012) também apresentam um estudo diacrônico da ordem VS com dados de peças de teatro e discutem a restrição da monoargumentalidade para a ordem VS no PB, considerada uma das propriedades de sujeito nulo mais resistentes na língua, como exemplificado em (27).

(27) Holly: Pronto, ___ chegou o outro estudante.

(SANTOS & SOARES DA SILVA, 2012, p. 122)

Com 734 dados de verbos monoargumentais separados em sete períodos de tempo entre 1833 e 1992, os achados da pesquisa dos autores estão expressos na Tabela 3.

⁷¹ Nascimento (1984) declarou que não existe inversão livre do sujeito no PB e que a ordem VS no PB só era possível com verbos inergativos e inacusativos.

Tabela 3: Frequência da ordem V-DP (*versus* DP-V) em Santos & Soares Silva (2012)

Período de tempo	I	II	III	IV	V	VI	VII
Ordem V-DP	56%	53%	66%	53%	45%	52%	50%

Fonte: Santos & Soares Silva (2012, p. 132)

Santos & Soares Silva (2012) mostram que os verbos inacusativos, de estado e de permanência são mais resistentes à ordem SV, e que DPs com traço [+humano] são mais facilmente atraídos para a posição pré-verbal. Ademais, eles afirmam que elementos dêiticos favorecem a ordem VS. Os autores chegam à conclusão de que DPs [+definidos], introduzidos por artigos definidos, pronomes possessivos, demonstrativos e quantificadores, e modificados por oração relativa, são frequentemente usados na ordem SV, ao passo em que DPs [-definidos], marcados por artigos ou pronomes indefinidos e expressões partitivas, frequentemente são utilizados na ordem VS.⁷²

Em relação ao comportamento sincrônico da ordem das palavras no PB, Pilati (2006) declara que as construções com ordem VS são produzidas em contextos bastante restritos nessa língua, pois são licenciadas sob o requerimento de uma relação dêitica com o tempo da fala, como, por exemplo: simultaneamente ao término de uma ação, como em (28a) e (30a); ou no momento em que se dá/recebe uma notícia, como em (29a).

(28) a. Entrou Dante. (PILATI, 2006, p. 199)

Dante entrou (aqui/nesse lugar)

b. Dante entrou.

Dante entrou (em algum lugar)

(29) a. Morreu Fellini. (PILATI, 2006, p. 199)

Fellini acabou de morrer

(Eu acabei de ouvir que Fellini morreu)

b. Fellini morreu.

(Fellini morreu (há algum tempo))

⁷² Para mais um estudo sobre ordem VS em perspectiva diacrônica, sugerimos ao leitor a leitura de Gravina (2014), que trata das mudanças na sintaxe dos sujeitos nulos e na ordenação sintática em um *corpus* constituído por jornais brasileiros e portugueses nas duas metades do século XIX e na primeira metade do século XX.

- (30) a. Telefonou Beatriz. (PILATI, 2006, p. 199)
Beatriz ligou (aqui, para este lugar)
 b. Beatriz telefonou.
Beatriz ligou (para algum lugar, fez telefonemas)

Pilati (2006) propõe que exista um *pro* nulo de interpretação locativa em posição pré-verbal nas sentenças com ordem VS em dados como (28)-(30), o que explicaria as distintas interpretações para as sentenças com ordem VS e SV no PB. Dessa forma, o pronome nulo de natureza locativa forneceria a informação dêitica necessária para a interpretação das orações com ordem VS. Em estudo posterior, seguindo algumas das observações originais da proposta de cisão do sistema pronominal no PB de Rabelo (2010), em que os morfemas de 3ª pessoa no PB não são lexicalmente especificados para o traço de definitude, ao passo que os morfemas de 1ª e 2ª pessoa são lexicalmente especificados para esse traço, Pilati, Naves & Salles (2017) argumentam que constituintes não argumentais podem ser projetados em [SPEC-TP] ou movidos para essa posição para satisfazer o traço-EPP no PB em construções em que a posição de sujeito não é comumente preenchida por sujeitos lexicais: (i) construções com inversão da ordem VS; (ii) construções com tópico-sujeito; (iii) construções com verbos meteorológicos; e (iv) construções com sujeito nulo arbitrário na 3ª pessoa do singular. Nesses termos, em casos em que há um elemento pré-verbal manifesto, como (31), o elemento locativo/temporal deve ser gramaticalmente interpretado como um pronome adverbial que satisfaz o EPP; e em casos em que não há um elemento pré-verbal manifesto, como em (32), deve-se interpretar que na posição do sujeito há um expletivo nulo locativo, licenciado, conforme apontou Pilati (2006), sob uma relação dêitica com o tempo da fala.

- (31) a. Ali_i dormem as crianças porque *e_i* é mais quente.
 b. A localidade_i não ocorreu nenhum problema porque *e_i* é segura.
 c. As cidades da Amazônia chovem muito.
 d. Aqui vende frutas. (PILATI, NAVES & SALLES, 2017, p. 71-77)
- (32) a. Tomou posse o ministro.
 b. Faltou sorte para o Fluminense no segundo tempo.
 c. Não chove com frequência em Brasília.
 d. Vende frutas. (PILATI, NAVES & SALLES, 2017, p. 71-77)

Em estudo recente, Naves & Pilati (2018) argumentam que, na verdade, os sujeitos nulos de 3ª pessoa singular não carregam obrigatoriamente traço de pessoa do discurso no PB. Seguindo também a proposta tipológica de Bhat (2004) para o sistema pronominal nas línguas de um modo em geral e a proposta de Biberauer (2010) para a natureza dos sujeitos expletivos nas LSN expletivo (cf. subseção 3.2.1 do capítulo anterior), Naves & Pilati (2018) argumentam que esses sujeitos devem ser tratados como proformas locativas que atuam na sintaxe como expletivos quase argumentais e podem ser nulos ou manifestos. As autoras mostram que essa proforma locativa é licenciada em construções com verbos meteorológicos e existenciais.

- (33) a. *e* Chove muito em Brasília. (NAVES & PILATI, 2018, p. 205)
 b. *e* Tem prédios lindos em Brasília.

- (34) a. Aqui/ali/lá chove muito. / Aqui/ali/lá tem prédios lindos.
 b. Brasília chove muito. / Brasília tem prédios lindos.
 (NAVES & PILATI, 2018, p. 206)

Para as autoras, *aqui/ali/lá* são proformas expressas que estão em distribuição complementar com um expletivo de natureza locativa no PB [*proLOC*]. Seguindo a linha de análise em Pilati, Naves & Salles (2017), Naves & Pilati (2018) também associam esse pronome expletivo às construções com ordem VS (35) e com sujeito nulo arbitrário (36).

- (35) a. Entrou Dante. (NAVES & PILATI, 2018, p. 207)
 Dante entrou (aqui /nesse lugar).
 b. Dante entrou.
 Dante entrou (em algum lugar).

- (36) a. [*proLOC*] Conserta sapatos. (NAVES & PILATI, 2018, p. 207)
 b. Aqui conserta sapatos.

Naves & Pilati (2018) enfatizam três evidências intralinguísticas de que as proformas locativas *aqui/ali/lá* ocupam [SPEC-TP] no PB: em (37), a proforma pode controlar o sujeito nulo de uma oração encaixada; em (38), assim como o *there* do inglês, a proforma está sujeita

ao alçamento (CHOMSKY, 1995); em (39), a proforma impede o controle do sujeito encaixado pelo sujeito da matriz.⁷³

- (37) a. As crianças_i dormem ali porque *e*_i querem. (NAVES & PILATI, 2018, p. 207)
 b. *Ali dormem as crianças_i porque *e*_i querem.
 c. Ali_i dormem as crianças porque *e*_i é mais quente.
- (38) a. Parece que aqui conserta sapato. (NAVES & PILATI, 2018, p. 207)
 b. Aqui parece que conserta sapato.
- (39) a. O Pedro_i disse que *e*_{i/*j} conserta sapato. (NAVES & PILATI, 2018, p. 207)
 b. O Pedro_i disse que aqui*_{i/j} conserta sapato.

Conforme vimos no capítulo anterior, a ordem VS e a disponibilidade de sujeitos nulos também estão associadas ao movimento do verbo para um núcleo funcional flexional ou temporal em algumas línguas. Haja vista que a maioria dos estudos gerativistas assume que os traços de pessoa e número em T no PB foram empobrecidos (cf. DUARTE & FIGUEIREDO SILVA, 2016), muitos pesquisadores têm sugerido que o verbo deixa a posição em que é gerado no PB, porém não ocupa mais T. Galves (2001) declarou que a distinção morfológica entre singular e plural nos paradigmas dos verbos finitos seria suficiente para desencadear o movimento do verbo. Já Costa e Galves (2002) propõem que não há diferença em relação à posição de projeção do verbo nessas duas variedades do português, o PB e o PE. Contra Galves (2001), Costa & Figueiredo Silva (2006) pontuam que as diferenças em relação à extensão do movimento do verbo no PB coloquial e no PE não podem ser derivadas de diferenças morfológicas de número, pois o PB tem uma morfologia de número diferente no domínio verbal quando comparado com PE, mas, segundo os autores, ambas as línguas teriam movimento do verbo na mesma extensão. Eles argumentam que as diferenças observadas entre as línguas devem estar relacionadas ao fato de que o sujeito é mais frequentemente manifesto no PB do que no PE. Costa & Galves (2002) e Costa & Figueiredo (2006) assumem, portanto, que no PB e no PE há apenas um movimento curto do verbo: os autores

⁷³ Duarte & Figueiredo Silva (2016) afirmam que permanece uma questão de disputa considerar formas verbais de 3ª pessoa singular precedidas por advérbios locativos como instâncias de sujeitos nulos genéricos. Para as autoras, contudo, a alta frequência de construções com tópicos-sujeitos locativos e genitivos (cf. MUNHOZ, 2011; MUNHOZ & NAVES, 2012) sugere que, no PB, o advérbio locativo em sentenças como (32a) preenche a posição de [Spec-TP].

argumentam que V se moveria em ambas as línguas até T, mas não para AgrS, um nó flexional mais alto, como ilustrado em (40).

- (40) PE: [_{AgrSP} Subj [_{TP} t_{Subj} V [_{VP} t_{Subj} t_v]]] (COSTA & GALVES, 2002, p. 116)
 PB: [_{AgrSP} Subj [_{AgrSP} [_{TP} *pro/ele* V [_{VP} t_{pro} t_v]]]]

Numa perspectiva cartográfica para a sintaxe dos advérbios, Tescari Neto (2013) propõe que a variação na posição do verbo em relação ao AdvP no PB pode ser vista de duas maneiras: o movimento do verbo lexical seria obrigatório até um dado núcleo na parte inferior do núcleo funcional T, ou seja, o verbo lexical se moveria para uma projeção intermediária; e o verbo só pode passar por advérbios baixos (cf. TESCARI NETO, 2013).⁷⁴ Em trabalho posterior, Tescari Neto (2016) sugere que o enfraquecimento do sistema de concordância no PB não é suficiente para explicar a razão de o verbo não subir para T no PB em comparação com outras línguas românicas. O autor sugere que é o enfraquecimento de T que explicaria porque o verbo seria projetado para uma posição medial no PB. Conforme proposta de Cinque (1999) para as estruturas de hierarquia dos advérbios, Tescari Neto (2016) implementa um núcleo funcional abaixo de T, qual seja, T_{Anterior}P, como a posição medial para a qual o verbo se move no PB. Tescari Neto (2016) ilustra a sintaxe dessa projeção como em (41).

- (41)
-
- ```

graph TD
 T_AnteriorP --> AdvP
 T_AnteriorP --> T_Anterior_prime[T_Anterior']
 AdvP --> ja[já]
 T_Anterior_prime --> T_Anterior_0[T_Anterior^0]
 T_Anterior_prime --> VP
 T_Anterior_0 --> sei_i[seii]
 VP --> Eu[Eu]
 VP --> V_prime[V']
 V_prime --> V_0[V^0]
 V_prime --> DP
 V_0 --> sei[sei]
 DP --> portugues[português]

```

Ao propor que o enfraquecimento de T teria sido o gatilho para a perda do movimento do verbo para T no PB, a proposta de Tescari Neto (2016) parece concordar com a ideia, tanto da versão fraca quanto da versão forte da hipótese da concordância rica, de que concordância rica e movimento do verbo para T estão interligados.

<sup>74</sup> Sugerimos a leitura de Tescari Neto (2013 e trabalhos posteriores) para detalhes da proposta do autor para a sintaxe dos advérbios altos e baixos no PB sob a perspectiva da sintaxe cartográfica.

## 4.2 Panorama das construções com *se*-passivo e *se*-impessoal no português

Nesta seção, apresentamos as divergências e as convergências no comportamento sintático-semântico das construções com *se*-passivo e *se*-impessoal no PB e no PE, com o objetivo de justificar nossa escolha de analisar essas construções em nosso *corpus*, mostrando que essas construções ainda constituem uma produtiva agenda de pesquisa para estudar comparativamente a gramática dessas duas línguas.

Estudos descritivos da gramática do PE classificam construções como em (42), nas quais um verbo transitivo e sem argumento externo está conjugado na 3ª pessoa do singular, como construções com *se*-impessoal ou *se*-nominativo, pois nelas o clítico *se* é interpretado como o sujeito frásico (cf. BRITO, DUARTE & MATOS, 2003; I. DUARTE, 2013). Em favor dessa análise de que o *se*-impessoal possui conteúdo argumental e funcional, acrescentam-se dois fatos de natureza sintático-semântica: o primeiro é que o caráter referencial do *se*-impessoal permite a extração simultânea de clítico em orações coordenadas, como visto em (42); o segundo é que o *se*-impessoal é referencial e, por isso, não pode ocorrer associado a uma posição de pronome expletivo, como ilustrado em (43).

- (42) Informa-se que se aluga apartamentos e \_\_ vende moradias. (PE)  
(BRITO, DUARTE & MATOS, 2003, p. 837)

- (43) a. \*Há-se muitos livros nesta biblioteca. (PE)  
b. \*Parece-se a toda a gente que os professores comprem livros em excesso. (PE)  
(BRITO, DUARTE & MATOS, 2003, p. 837)

I. Duarte (2013, p. 444-447) restringe as construções com *se*-impessoal no PE àquelas realizadas com: verbos transitivos na 3ª pessoa do singular que expressam um traço de indeterminação do sujeito, parafrazeável por “alguém”, “as pessoas (em geral)” ou por “há pessoas que”, como em (44a); com verbos intransitivos que não selecionam argumento interno, como em (44b); e com verbos inacusativos, como em (44c) – esses dois últimos tipos de verbos, a autora ressalta, não são permitidos nas construções com o chamado *se*-passivo, exemplificadas em (45).

- (44) a. Diz-se que esse ministro vai ser demitido. (PE) (I. DUARTE, 2013, p. 445)  
b. Incorre-se frequentemente nesse erro. (PE)

c. Morria-se de sarampo no princípio do século passado. (PE)

(45) a. O raciocínio algébrico ensina-se *propositadamente* na escola. (PE)

b. As obras mais recentes publicaram-se numa editora espanhola *para garantir maiores tiragens*. (PE) (I. DUARTE, 2013, p. 444)

Quando o verbo das construções com *se* é transitivo ou ditransitivo e conjugado na 3ª pessoa do singular, I. Duarte (2013) afirma que existe nessas construções uma ambiguidade de interpretação entre uma leitura ativa e uma leitura passiva, conforme ilustrado no exemplo em (46). Para a autora, em (46) tem-se, na verdade, *se*-passivo.

(46) Descobriu-se um gene responsável pela obesidade. (PE) (I. DUARTE, 2013, p. 446)

(Interpretação 1, passiva pronominal: ‘Foi descoberto um gene responsável pela obesidade.’)

(Interpretação 2, *se*-impessoal: ‘Alguém descobriu um gene responsável pela obesidade.’)

De acordo com I. Duarte (2013), nas construções com *se*-passivo o DP argumento interno controla os traços de número do verbo, como em (46) – o *se*-passivo ocorre, portanto, apenas nos contextos em que o verbo da oração é transitivo ou ditransitivo. A autora afirma, então, que essas construções podem ser licenciadas tanto com DPs no plural, como em (45b), quanto com DPs no singular, como em (45a) e (46). A pesquisadora propõe que, nas construções com *se*-passivo com DPs no plural ou no singular, o clítico denota uma entidade arbitrária relacionada com o papel temático de agente da passiva, de modo que o argumento externo está implícito. Para a autora, extrai-se a interpretação do agente da passiva nessas construções pela possibilidade de sua ocorrência com um advérbio orientado para o agente, como em (45a), e com orações subordinadas finais formadas a partir de um adjunto cujo sujeito só pode ser controlado pelo argumento externo implícito da passiva, como em (45b).

Outra característica das construções com *se*-passivo é que elas não admitem o redobro do clítico, de modo que um argumento interpretado como agente não pode ser expresso por meio da preposição *por*, como em (47).

(47) \*Venderam-se hoje muitos livros *por alguém* na feira do livro. (PE)

(BRITO, DUARTE & MATOS, 2003, p. 839)

Diante das evidências apresentadas, a autora postula que nas construções com *se*-passivo o clítico atua como um morfema apassivador, expressando a diátese passiva da construção sem realizar qualquer alteração na morfologia do verbo e bloqueando a atribuição de papel temático à posição do argumento externo e de Caso acusativo ao argumento interno (cf. BRITO, DUARTE & MATOS, 2003; I. DUARTE, 2013). Nessa perspectiva, o *se*-passivo insere-se dentro do escopo dos clíticos quase argumentais (cf. BRITO, DUARTE & MATOS, 2003; I. DUARTE, 2013).<sup>75</sup>

Em estudo pioneiro, Naro (1976) observou que as construções com *se*-passivo, desde o português clássico (PC), podiam ser licenciadas sem que o verbo exibisse concordância com o DP argumento interno. Nesse mesmo período, houve o apagamento do sintagma preposicional com papel temático de agente dessas construções – em (48) exemplifica-se uma construção com *se*-passivo prototípica, um dos versos de *Os Lusíadas*, de Luís de Camões.

(48) Aqui se escreverão novas histórias, por gentes estrangeiras. (NARO, 1976, p. 781)

Para Naro (1976), já no período clássico da língua, especialmente a perda do sintagma preposicional com papel temático de agente e o enfraquecimento da concordância entre o DP argumento interno e o verbo teriam desencadeado uma reanálise de leitura passiva para uma leitura ativa na voz dos verbos de construções com *se*-passivo. Essa reanálise na leitura da voz dos verbos transitivos nas construções com *se* abriu, então, caminho para o surgimento das construções com *se*-impessoal com verbos intransitivos.

De acordo com Martins & Nunes (2016), construções com *se*-passivo não são licenciadas no PB, e as raras e inconsistentes construções remanescentes desse tipo são atualmente restritas à escrita e ao estilo formal proposto pela tradição gramatical prescritivista, que condena o uso de *se*-impessoal com verbos transitivos. Para propor uma descrição geral das construções com *se*-passivo e *se*-impessoal no português contemporâneo, Martins & Nunes (2016) contrastam algumas das propriedades sintático-semânticas dessas construções, bem como apresentam as convergências e as divergências relacionadas a elas nas gramáticas do PB e do PE. Elencamos a seguir alguns dos achados de Martins & Nunes

<sup>75</sup> Dobrovie-Sorin (1998) defende que o *se*-passivo e o *se*-impessoal não estão associados a uma posição argumental e, com base nisso, denonima-os de *se*-acusativo. De acordo com Dobrovie-Sorin (1998), a possibilidade de se ter uma interpretação com um argumento externo implícito nessas orações é decorrente das propriedades semânticas dos verbos em que o *se*-acusativo hospeda-se. Em contrapartida, Brito, Duarte & Matos (2003) apontam que essa análise não explica a correlação entre a presença do clítico e a suspensão de atribuição de relação temática ao argumento externo no caso, por exemplo, do *se*-passivo.

(2016), a fim de evidenciar o papel crucial que essas construções desempenham na descrição da diferença entre as gramáticas do PE e do PB:

A) o argumento interno das construções com *se*-passivo pode mover-se para uma posição pré-verbal (49), assim como o das passivas verbais (50):

(49) a. Os bolos comeram-se ontem. (PE) (MARTINS & NUNES, 2016, p. 323)

b. \*Os bolos comeu-se ontem. (PE)

(50) a. Estas maçãs todas apanharam-se sem estarem maduras. (PE)

b. Estas maçãs todas foram apanhadas sem estarem maduras. (PE)

(MARTINS & NUNES, 2016, p. 325)

Os autores salientam que a construção impessoal em (49b) pode ser julgada como aceitável com uma entonação marcada, graficamente representada por uma vírgula: “Os bolos, comeu-se ontem”. Por sua vez, o *se*-passivo na construção em (49a) não requer qualquer entonação especial para ser licenciada. Martins & Nunes (2016) tratam o contraste entre (49a) e (49b) como uma evidência de que o *se* das construções em que o DP plural não concorda com o verbo deve ser analisado como *se*-impessoal, e não como *se*-passivo. Em relação à posição pré-verbal do DP em (50a), mencionamos o trabalho clássico de Raposo & Uriagereka (1996), no qual os autores apresentam uma explicação, em termos minimalistas, para a concordância entre o verbo e o argumento interno das construções com *se*-passivo no PE. Eles argumentam que, nessas construções, o *se*-passivo ocupa a posição de sujeito e o DP argumento interno pode ser projetado para uma posição à esquerda de T, onde é amalgamada a concordância. Para averiguar o estatuto do clítico e o comportamento do DP nessas construções em comparação com o DP argumento externo em outras construções, Raposo & Uriagereka (1996) realizam alguns testes sintáticos. Um dos testes realizados mostra o comportamento do sujeito de sentenças infinitivas complementos de predicados adjetivos, o qual pode ocupar uma posição pré-verbal nas formas ativa e passiva (51a-b). Em comparação com as construções com *se* (51c-d), o DP dessas construções não é aceito nessa posição (51d).

(51) a. Vai ser difícil os tribunais aceitarem os documentos.

b. Vai ser difícil os documentos serem aceites.

c. Vai ser difícil aceitarem-se os documentos.

d. \*Vai ser difícil os documentos aceitarem-se.

(RAPOSO & URIAGEREKA, 1996, p. 754)

Outro teste que os autores realizam, em relação ao comportamento do DP das construções com *se*, constata que, em (52a), a leitura é ambígua entre uma interpretação passiva ou recíproca, e, em (52b), a leitura é obrigatoriamente reflexiva/recíproca.

(52) a. Os especialistas consultaram-se durante a operação.

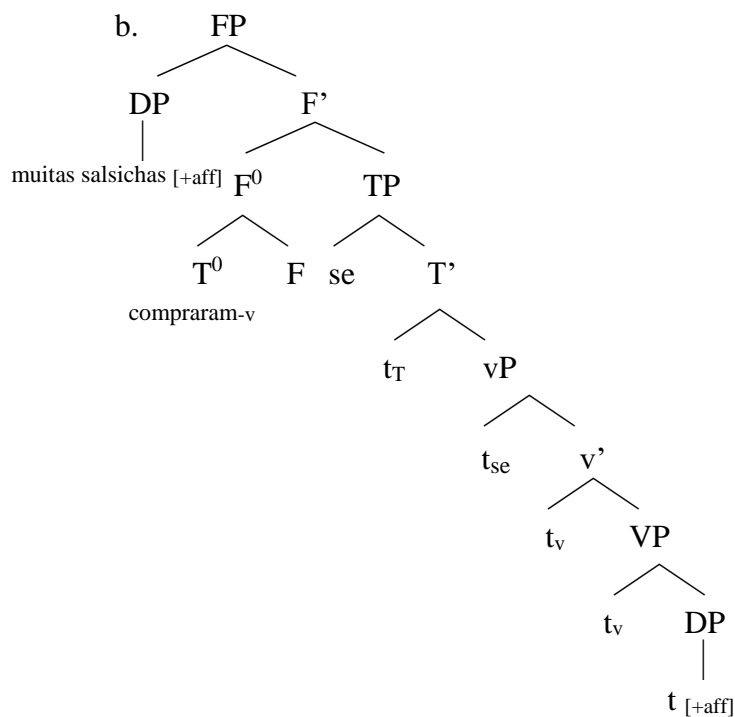
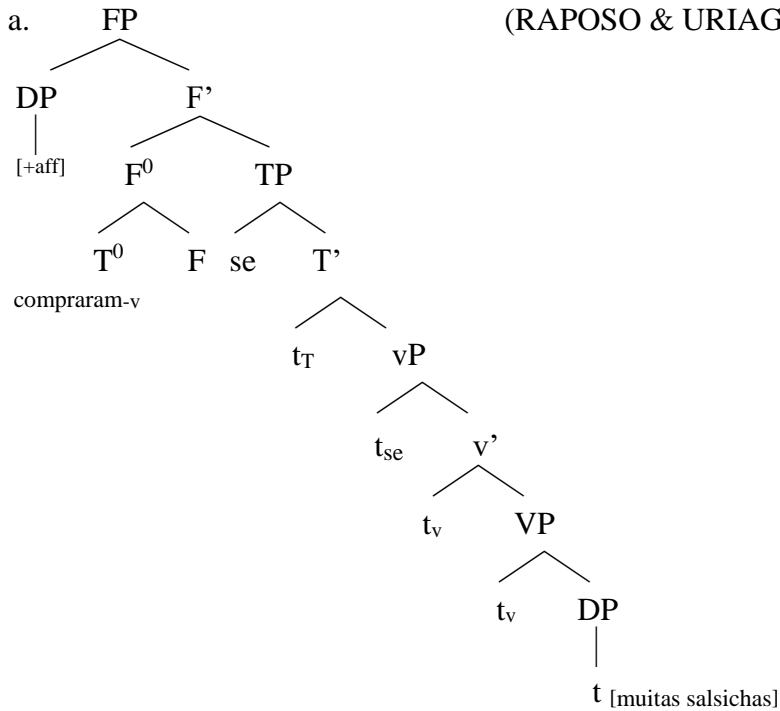
b. Em que momento da operação os especialistas se consultaram?

(RAPOSO & URIAGEREKA, 1996, p. 765-766)

Ao analisar as sentenças em (52), Raposo & Uriagereka (1996) constataam que a construção em (52b) possui apenas leitura recíproca, pois o adjunto é deslocado para uma posição de tópico, e, dessa forma, o DP assume a posição de sujeito. Para os autores, é justamente por causa de o DP ocupar a posição de sujeito que a construção (52b) recebe leitura recíproca. Já a leitura da construção em (52a) é ambígua entre uma interpretação passiva ou recíproca devido ao traço [+humano] do DP e à possibilidade, segundo os autores, de o DP ocupar duas posições. Os autores sugerem que, nas construções com *se*-passivo, o DP ocupa uma posição à esquerda de T, a saber: FP (do inglês *ForceP*). Dessa forma, a leitura passiva de (52a) ocorre com o DP na posição de especificador de F, enquanto a leitura recíproca ocorre com o DP na posição de especificador de T. As representações arbóreas em (53) resumem a proposta de Raposo & Uriagereka (1996) para a posição do DP argumento interno nas construções com *se*-passivo – que atua, nos termos usados pelos autores, como um sintagma afetivo (do inglês, *affective phrase*), isto é, um elemento sintático que no PE ocupa uma posição acima de T/IP, mas não é um sujeito nem um tópico.



(53) a. (RAPOSO &amp; URIAGEREKA, 1996, p. 786)



Para Raposo & Uriagereka (1996), a checagem de traços em F pode ser manifesta, de modo que há então o movimento do DP (53b) – destaca-se que o clítico *se* não bloqueia o movimento do DP em (53b), pois ele não possui traços que poderiam ser atraídos por F. Ou, ainda, pode haver apenas a checagem dos traços formais, de modo que o DP permanece *in situ*, como em (53a). Em outras palavras, o que é atraído não é a categoria completa, mas um traço formal particular. Os traços de uma categoria completa são checados em movimento

manifesto apenas por razões de convergência na PF, o que, por sua vez, implica dizer que o movimento implícito na LF é apenas para a atração de traços formais, nunca da própria categoria (RAPOSO & URIAGEREKA, 1996). Assim, quando um traço formal específico de uma determinada categoria é atraído na LF, o movimento carrega automaticamente todo o conjunto de traços formais da categoria. As operações de movimento em (53a) são: a subida do clítico *se* para [SPEC-T], a subida do complexo [V-v] para T e o subsequente movimento de T contendo o complexo [V-v] para F. O movimento do DP para [SPEC-FP] é visto pelos autores como uma operação de último recurso.

Para dar conta do Caso nominativo valorado entre o DP e o verbo na interpretação passiva da construção em (52a) e em (53), Raposo & Uriagereka (1996) propõem que no PE o núcleo F, assim como T, pode valorar o Caso nominativo. Dessa forma, para esses pesquisadores, F é a categoria em que, no PE, o DP argumento interno das chamadas passivas com *se* recebe Caso nominativo. Já nas construções em que o DP não estabelece concordância com o verbo, o DP permanece na posição em que foi gerado, interno ao VP, e, ali, recebe Caso acusativo (54).

(54) Compra-se sempre demasiadas salsichas no talho Sanzot.

(RAPOSO & URIAGEREKA, 1996, p. 750)

Em relação à posição do clítico *se*, Raposo & Uriagereka (1996) afirmam que o *se* é um DP mínimo que possui traços semânticos reduzidos, como PRO; assim: (a) nas construções com *se*-passivo, o clítico checa o traço-D de T, mas recebe Caso nulo, ou seja, nessas construções o clítico *se* não recebe Caso nominativo em T e também não impede a atribuição de Caso nominativo por F; (b) nas construções com *se*-impessoal o clítico checa o traço-D de T e ali recebe Caso nominativo, tendo em vista ser o único elemento disponível para essa operação. A explicação de Raposo & Uriagereka (1996) parece boa para as construções com *se*-passivo no PE, mas não serve para o PB atualmente, uma vez que o DP argumento interno plural de construções com *se* no PB não desencadeia concordância com o verbo e não pode ocupar uma posição à esquerda de T como F, tendo em vista que esse núcleo não está mais ativo na gramática do PB (cf. PIRES, 2006).

Pires (2006) salienta que a projeção FP deve ter estado ativa na gramática do PB até pelo menos o século XIX, o que torna plausível sugerir que a perda das propriedades desse mecanismo gramatical pode estar relacionada com várias outras mudanças desencadeadas concomitantemente na gramática do PB. O autor cita como exemplos dessas mudanças a

perda do movimento do verbo para uma posição acima de T, a perda de infinitivos flexionados, a mudança na colocação de clíticos, o empobrecimento do sistema flexional, entre outros. Nessa direção, sugerimos que a perda das construções com *se*-passivo na gramática do PB também pode ser colocada nesse rol de mudanças possivelmente provocadas pela perda do núcleo funcional F na gramática do PB.<sup>76</sup> Pires (2006) mostra que a perda do núcleo funcional F no PB eliminou o movimento aberto do verbo para a posição à esquerda de um clítico, provocando a preferência da próclise no lugar da ênclise, como ilustrado em (55).

- (55) a. Ela me/te viu ontem. (\*viu-me/\*te) (PB) (PIRES, 2006, p. 158)  
 b. (Ela) viu-me/te ontem. (\*me/\*te viu). (PE/século XVIII)

O autor incrementa seu argumento mencionando os resultados do trabalho de Cyrino (1993, p. 167-168) sobre a posição do clítico em peças de teatro, os quais indicam uma perda sistemática da ênclise com verbos no infinitivo impessoal e no gerúndio no PB. Pires (2006) afirma que esses resultados fornecem evidência consistente de que o movimento do verbo para F foi em geral perdido na segunda metade do século XX. Pires (2006) sistematizou os resultados do trabalho de Cyrino (1993) em relação à posição enclítica dos clíticos com verbos no infinitivo impessoal e no gerúndio conforme a Tabela 4 a seguir.

Tabela 4: Percentual de clíticos em posição enclítica em Cyrino (1993).

| Período | Infinitivos impessoais | Gerúndios |
|---------|------------------------|-----------|
| 1700    | 100%                   | 100%      |
| 1800    | 100%                   | 100%      |
| 1850    | 86%                    | 100%      |
| 1900    | 56%                    | 25%       |
| 1950b   | 25%                    | 66%       |
| 1950b   | 0%                     | 0%        |

Fonte: Pires (2006, p. 149).

<sup>76</sup> Pires (2016) observa que, contanto que o movimento de um núcleo ocorra localmente, no movimento de V-para-C no português, V teria que se mover primeiramente para F, e só então para C. Dessa forma, a perda de V-para-C pode ter sido um dos gatilhos cruciais para a perda de V-para-F no PB.

Feita essas considerações sobre os trabalhos de Raposo & Uriagereka (1996) e Pires (2006), retomamos nossa discussão a respeito do estudo de Martins & Nunes (2016) sobre os aspectos descritivos formais das construções com *se* no PE e no PB.

B) construções com *se*-passivo não atribuem Caso acusativo ao argumento interno:

- (56) a. \*Compraram-se-os ontem. (PE)  
 b. Comprou-se-os ontem. (PE) (MARTINS & NUNES, 2016, p. 324)
- (57) a. \*Podem-se comprá-los amanhã. (PE)  
 b. Pode-se comprá-los amanhã. (PE) (MARTINS & NUNES, 2016, p. 324)
- (58) a. \*Histórias de lobisomens, ouviam-se-lhas vezes sem conta. (PE)  
 b. Histórias de lobisomens, ouvia-se-lhas vezes sem conta. (PE)  
 (MARTINS & NUNES, 2016, p. 324)

Por meio dos dados em (56)-(58), em que o clítico *se* vem seguido por um clítico acusativo, Martins & Nunes (2016) mostram que as construções com *se*-impessoal atribuem Caso acusativo ao argumento interno em diferentes contextos: quando os clíticos estão encadeados em um mesmo hospedeiro (56b); quando ambos são colocados em hospedeiros diferentes (57b); e, inclusive, quando um clítico dativo intervém entre o *se* e o clítico acusativo (58b). Já os dados em (56a), (57a) e (58a) evidenciam que as construções com *se*-passivo não atribuem Caso acusativo ao argumento interno de seus verbos. Essa evidência ratifica a análise de Martins & Nunes (2016) de que, apesar de as construções com *se*-impessoal apresentarem propriedades semânticas semelhantes às construções com *se*-passivo, como aquelas apontadas por I. Duarte (2013) em (46), elas não são sintaticamente caracterizadas como construções passivas com *se* – ou seja: nas construções impessoais com *se*, o argumento interno (plural ou singular) não controla os traços de flexão do verbo, enquanto nas passivas com *se* o argumento interno deve estar no plural e controlar os traços de flexão do verbo.<sup>77</sup> Assim, a agramaticalidade de (56a), (57a) e (58a) pode ser tomada, consoante Martins & Nunes (2016), para mostrar que o *se*-passivo se comporta como um

<sup>77</sup> Castelar de Carvalho (1990) já havia afirmado que apenas a interpretação impessoal do *se* é disponível no PB.

particípio de morfologia passiva na medida em que desativa a propriedade de o verbo atribuir Caso acusativo.<sup>78</sup>

C) ausência de um sintagma preposicional para expressar o agente da passiva, isto é, ausência de *by-phrase*, nas construções com *se*-passivo, como exemplificado em (59) e (60):

- (59) a. Os bolos foram comidos (pelos meninos). (MARTINS & NUNES, 2016, p. 326)  
 b. Comeram-se os bolos (\*pelos meninos). (PE)  
 c. Os bolos comeram-se (\*pelos meninos). (PE)

- (60) a. As casas foram destruídas. (MARTINS & NUNES, 2016, p. 326)  
 b. Destruíram-se as casas. (PE)

Diferentemente das passivas verbais e das construções com *se*-passivo prototípicas do período clássico do português, as construções com *se*-passivo no PE contemporâneo não licenciam um PP com o agente da passiva exposto, conforme ilustrado nos dados em (59b-c). Martins & Nunes (2016) mostram que o argumento externo implícito das passivas verbais é compatível com uma interpretação [+humana] ou [-humana]. Isso quer dizer que em (56a) as casas podem ter sido destruídas tanto por seus proprietários quanto pela chuva: (i) “as casas foram destruídas *pelo dono*” ou (ii) “as casas foram destruídas *pela chuva*”. Em contraste, construções com *se*-passivo, como em (60), somente permitem uma interpretação [+humana] para seu argumento externo implícito (MARTINS & NUNES, 2016). Em outras palavras, (60b) não é empregada para descrever a destruição de casas pela chuva. Seguindo Naro (1976), Martins & Nunes (2016) afirmam que as construções com *se*-impessoal são resultantes de uma reanálise daquelas com *se*-passivo e mostram que as construções com *se*-impessoal herdaram das construções com *se*-passivo a restrição de uma interpretação [+humana] para seu argumento externo implícito, como ilustrado em (61).

- (61) Destruiu-se as casas. (MARTINS & NUNES, 2016, p. 326)

---

<sup>78</sup> Martins & Nunes (2016, p. 324) observam que existe, para alguns falantes, uma restrição de adjacência independente no PE que exclui o *se* seguido por um clítico acusativo, como mostrado a seguir com o *se*-reflexivo: (i) \*O João deu-se-o. Os autores informam que os dados (56)-(58) foram julgados considerando-se o dialeto da primeira autora.

D) com variações na gramaticalidade de acordo com a variedade da língua e o dialeto do falante, o *se*-impessoal é licenciado em construções com verbos não transitivos:

- (62) a. Precisa-*se* de funcionários. (Verbo transitivo com complemento preposicional).  
 b. Trabalha-*se* muito nesta cidade. (Verbo inergativo).  
 c. Chegava-*se* cedo ao trabalho. (Verbo inacusativo).  
 d. Quando *se* é promovido, as coisas ficam mais fáceis. (Verbo passivo).  
 e. Não *se* ficou contente com a nova situação. (Verbo de ligação).  
 f. Parecia-*se* ir ganhar o jogo. (Verbo de alçamento: PE<sup>ok</sup>; BP\*).

(MARTINS & NUNES, 2016, p. 328-329)

Dado que o *se*-impessoal (um sujeito frásico) surgiu de uma reanálise do *se*-passivo (um particípio de morfologia passiva) em construções com verbos transitivos, Martins & Nunes (2016) defendem que ele deixou de se restringir a esses verbos e passou a ser utilizado em outros tipos do verbo, conforme ilustrado em (62).

E) somente construções com *se*-impessoal licenciam predicado secundário orientado para o sujeito:

- (63) a. \*Criam-*se* avestruzes despreocupado. (PE) (MARTINS & NUNES, 2016, p. 327)  
 b. Cria-*se* avestruzes despreocupado. (PE)

Essa restrição de um predicado secundário orientado para o sujeito poder ocorrer somente em construções com *se*-impessoal leva à suposição de que o *se*-impessoal checa o Caso nominativo disponível na derivação e é, portanto, o sujeito da oração.

F) construções sem o clítico *se*-impessoal, como em (64), são licenciadas no PB, mas não no PE, e o uso do *se*-impessoal como sujeito de verbo no infinitivo, como no contraste em (65)-(66), são frequentes no PB, mas não no PE.<sup>79</sup>

---

<sup>79</sup> Cavalcante (2006) mostra que há uma grande diferença entre a porcentagem de uso do clítico *se* sujeito de infinitivo no PB e no PE. Os resultados de sua pesquisa, com dados dos séculos XIX e XX, mostram o seguinte: 8% de *se* junto a infinitivos na fala e na escrita no PE; e 20% na fala e 50% na escrita de *se* junto a infinitivo no PB – além de 20% de *a gente* e *você* (ambos genéricos) como sujeitos de infinitivo na fala do PB.

- (64) a. Não usa mais esse estilo de redação. (PB) (MARTINS & NUNES, 2016, p. 329)  
 b. Casava muito cedo no século passado. (PB)  
 c. No futuro vai descobrir remédio para tudo quanto é doença. (PB)
- (65) a. O mar está perigoso para nadar. (PB<sup>ok</sup>; PE<sup>ok</sup>) (MARTINS & NUNES, 2016, p. 330)  
 b. O mar está perigoso para *se* nadar. (PB<sup>ok</sup>; PE:??)
- (66) a. Nova Iorque e Barcelona são cidades fáceis de gostar. (PB ?\*; PE<sup>ok</sup>)  
 b. Nova Iorque e Barcelona soa cidades fáceis de *se* gostar. (PB<sup>ok</sup>; PE\*)  
 c. Nova Iorque e Barcelona soa cidades fáceis *da gente* gostar. (PB<sup>ok</sup>; PE\*)  
 (MARTINS & NUNES, 2016, p. 330)

Em relação aos dados em (64), Martins & Nunes (2016) mencionam que alguém pode se perguntar se essas sentenças, na verdade, resultariam: (i) da exclusão do *se*-impessoal; ou (ii) do enfraquecimento da morfologia de 3ª pessoa no plural, que também licencia sujeitos nulos de leitura arbitrária, como ilustrado em (67).

- (67) a. Telefonaram para você. (MARTINS & NUNES, 2016, p. 329)  
 b. Estão batendo na porta.

Martins & Nunes (2016) ressaltam, contudo, que uma diferença importante entre as sentenças em (64) e (67) é a de que o sujeito indefinido nulo necessariamente exclui o falante em (67), mas não em (64). Os autores salientam, ainda, que as construções com *se*-impessoal podem ou não incluir o falante e mostram que, em determinados dialetos do PE, isso pode até ser desambiguizado via duplicação/redobro do clítico, conforme ilustrado no exemplo (68) – que também é uma evidência para supor que o *se*-impessoal recebe Caso nominativo e ocupa a posição de sujeito da oração.

- (68) a. Chama-*se*-lhe *a gente* espigas. (%PE) (MARTINS & NUNES, 2016, p. 327)  
 b. Há varias qualidades que até ainda *nos* não *se* conhecemos. (%PE)  
 c. Sei é de real certeza que isto era com o que *se eles* batiam o centeio. (%PE)

Para Martins & Nunes (2016), como o PB tornou-se uma LSN parcial, sua gramática também começou a favorecer sujeitos abertamente expressos em contextos sintáticos onde uma LSN do tipo românico escolhe um pronominal nulo como a opção não marcada. Em decorrência disso, o *se*-impessoal passou a ser mais frequentemente expresso em orações com verbos infinitivos no PB quando comparado com o PE (MARTINS & NUNES, 2016). De acordo com essa visão, isso seria uma estratégia para apoiar uma interpretação arbitrária/genérica para o sujeito do infinitivo. Esse contraste entre PB e PE é especialmente evidente em orações infinitivas como as ilustradas em (65) e (66), nas quais o PE desfavorece visivelmente, ou mesmo exclui, a presença de *se*-impessoal, da mesma forma que exclui outros sujeitos, como o *a gente*, que são genéricos no PB (MARTINS & NUNES, 2016).

G) apesar de a gramática do PB ser proclítica, o clítico das construções com *se*-impessoal é sempre enclítico no PB, salvo quando há uma partícula atratora de próclise:

- (69) a. Me viram/\*viram-me no cinema. (MARTINS & NUNES, 2016, p. 331)  
 b. Eles não te criticaram/\*criticaram-te.  
 c. Você deve, sem sombra de dúvida, se inscrever na competição.  
 d. \*Você se deve/deve-se, sem sombra de dúvida, inscrever na competição.  
 e. Você não deve, sem conhecer as regras, se inscrever na competição.  
 f. \*Você não se deve/deve-se, sem conhecer as regras, inscrever na competição.
- (70) a. Contratou-se/\*Se contratou um novo professor.  
 b. Não se contratou/\*contratou-se um novo professor.  
 c. \*Deve, sem sombra de dúvida, se contratar um novo professor.  
 d. Deve-se, sem sombra de dúvida, contratar um novo professor.  
 e. \*Não deve, sem haver justa causa, se despedir um professor.  
 f. \*Não deve-se, sem haver justa causa, despedir um professor.  
 g. Não se deve, sem haver justa causa, despedir um professor.  
 (MARTINS & NUNES, 2016, p. 331)
- (71) \*Pode-se se sentar em qualquer lugar. (PB) (MARTINS & NUNES, 2016, p. 333)
- (72) a. Não se pode se divertir com um barulho desses. (PB)  
 b. Não se deve se levantar tarde. (MARTINS & NUNES, 2016, p. 333)



A restrição do *se*-impessoal à próclise em (70), exceto nos casos em que há partículas atratoras, quando comparada com os demais clíticos da língua em (69), leva os autores a concluir que o *se*-impessoal no PB possui especificações lexicais que o diferenciam dos outros clíticos da língua, incluindo o clítico reflexivo de terceira pessoa, que é homófono do *se*-impessoal. Martins & Nunes (2016) argumentam também que o que resulta na gramaticalidade, ou não, de dados como em (71) e (72) é adjacência. Isto é, as sentenças em (72) são bem formadas porque os pronomes não estão adjacentes um ao outro, enquanto em (71) o *se*-impessoal e o *se*-reflexivo não são permitidos num mesmo domínio. Assim, os autores presumem que sentenças agramaticais como (71) podem tornar-se gramaticais se um material parentético romper a adjacência entre os dois clíticos, como ilustrado em (73):

(73) Pode-se, salvo engano, se sentar em qualquer lugar.

(MARTINS & NUNES, 2016, p. 333)

No Quadro 2 a seguir, sintetizamos as propriedades das construções com *se*-passivo e *se*-impessoal de acordo com o que apresentamos até aqui (cf. BRITO, DUARTE & MATOS, 2003; I. DUARTE, 2013; MARTINS & NUNES, 2016; NARO, 1976; RAPOSO & URIAGEREKA, 1996; entre outros), a fim de explicitar nosso especial interesse nessas construções, que têm sido amplamente utilizadas como um dos argumentos cruciais para marcar a distinção entre a gramática do PB e do PE e, mais especificamente, classificar o PB como uma LSN parcial e o PE como uma LSN consistente.

Quadro 2: Quadro panorâmico das construções com *se* passivo e impessoal no português.

| CONSTRUÇÕES                                                                                                                                                           | CARACTERÍSTICAS                                                                                                                                                                                                                                                                                                                                                             |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>Construção com <i>se</i>-passivo prototípica:</b></p> <p>(i) Aqui se escreverão novas histórias por gentes estrangeiras que virão.<br/>(NARO, 1976, p. 781)</p> | <p>Foram registradas com sintagma preposicional expressando o agente da passiva até o século XVII; nesse período, inicia-se a reanálise do <i>se</i>-passivo para o <i>se</i>-impessoal devido à perda do agente da passiva e da concordância entre o verbo transitivo e o argumento interno. Surgimento de construções com <i>se</i>-impessoal e verbos intransitivos.</p> |

|                                                                                                                                                                                                                                  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>Passivas pronominais / construções com <i>se</i>-passivo (restritas ao PE):</b></p> <p>(ii) As obras mais recentes publicaram-se numa editora espanhola para garantir maiores tiragens. (I. DUARTE, 2013, p. 444)</p>      | <p>Nessas construções, o <i>se</i> atua como um morfema passivo, sem realizar alteração na morfologia do verbo, e bloqueia a atribuição de papel temático à posição do argumento externo e de Caso acusativo ao argumento interno. O <i>se</i>-passivo insere-se dentro do escopo dos clíticos quase argumentais e o argumento interno, que controla os traços de pessoa e número do verbo, está ligado a uma posição de tópico, pois pode ser alçado a uma posição pré-verbal, onde recebe Caso nominativo. Extraí-se a interpretação do argumento interno implícito das construções com <i>se</i>-passivo por meio de testes com um advérbio orientado para o agente e com orações subordinadas finais formadas a partir de um adjunto cujo sujeito só pode ser controlado por argumento externo implícito. Nessas construções, não se admite o redobro do clítico, mas se admite a extração simultânea de clítico.</p> |
| <p><b>Construções com <i>se</i>-impessoal (PE/PB):</b></p> <p>(iii) a. Precisa-se de funcionários.</p> <p>b. Trabalha-se muito nesta cidade.</p> <p>c. Chegava-se cedo ao trabalho. (MARTINS &amp; NUNES, 2016, p. 328-329).</p> | <p>O <i>se</i>-impessoal tem valor referencial, é analisado como o sujeito frásico e recebe Caso nominativo. Nessas construções, permite-se a extração simultânea de clítico e, em alguns dialetos do PE, o redobro do clítico. Além dos verbos transitivos e intransitivos, o <i>se</i>-impessoal também é licenciado com verbos inergativos, inacusativos, passivos, de ligação e, no caso do PE, de alçamento. Na ausência de uma partícula atratora de próclise, o <i>se</i>-impessoal é enclítico no PB, ao passo que, inversamente, os demais clíticos no PB são proclíticos. No PE, o <i>se</i>-passivo e o <i>se</i>-impessoal são enclíticos.</p>                                                                                                                                                                                                                                                                |
| <p><b>Construções SEM o clítico <i>se</i>-impessoal (restritas ao PB):</b></p> <p>(iv) Casava muito cedo no século passado.</p> <p>(MARTINS &amp; NUNES, 2016, p. 329).</p>                                                      | <p>Por ser o PB uma LSN parcial, quando um sujeito nulo indefinido na 3ª pessoa do singular entra em uma relação de <i>Agree</i> com T, ele só pode ser interpretado indeterminadamente. Assim, sujeitos dessa natureza denotam pessoas em geral e podem incluir o falante e o ouvinte (leitura inclusiva), bem como denotam pessoas em geral, mas excluem o falante e o ouvinte (leitura exclusiva). São tidos como resultantes da reanálise das construções com <i>se</i>-impessoal.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                |

|                                                                                                                                                                                          |                                                                                                                                                                                                                                                                                                                                                              |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>SE-sujeito de verbo no infinitivo (mais produtivas no PB):</b></p> <p>(v) Nova Iorque e Barcelona são cidades fáceis de se gostar.</p> <p>(MARTINS &amp; NUNES, 2016, p. 330).</p> | <p>Como o PB tornou-se uma LSN parcial, passou a favorecer sujeitos manifestos em contextos onde uma LSN do tipo românico escolhe um sujeito nulo como a opção não marcada; assim, o <i>se</i>-impessoal como sujeito de infinitivos pode ser interpretado como uma estratégia para apoiar uma leitura arbitrária/genérica para o sujeito do infinitivo.</p> |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Fonte: o autor da tese.

O fato que mais chama a atenção no Quadro 1 é o de que as construções sem o clítico *se*-impessoal são restritas ao PB e, conforme propôs Holmberg (2010a), são uma das propriedades definitórias do PB como uma LSN parcial, tendo em vista que a operação de concordância entre um pronome nulo de 3ª pessoa e T numa LSN parcial resulta num sujeito nulo indeterminado como aquele das construções sem o clítico *se*-impessoal.

### 4.3 Síntese do capítulo

Neste capítulo, relatamos algumas das principais mudanças atestadas na sintaxe do sujeito no PB, tanto em perspectiva diacrônica quanto sincrônica. Verificamos que o debate em torno da classificação do PB dentre as línguas de sujeito nulo é uma constante desde estudos seminais como os de Galves (1984), M. E. Duarte (1993) e outros. Apresentamos dados diacrônicos que confirmavam mudanças na expressão do sujeito pronominal e na livre inversão do sujeito no PB. Vimos que a ordem VS com verbos monoargumentais constitui, talvez, a mais resistente característica das LSN consistentes presente no PB. Vimos certas assimetrias na interpretação dos sujeitos nulos e manifestos no PB e no PE, especialmente em orações encaixadas. Exemplificamos algumas das inovações na gramática do PB, como a interpretação do sujeito nulo de uma oração subordinada com correferência restrita ao sujeito da oração matriz, bem como o licenciamento de sujeitos nulos genéricos/arbitrários, propriedades que, conforme vimos no Capítulo 3, definem as LSN parciais (cf. HOLMBERG, 2005, 2010a). Discutimos análises em que o verbo no PB parece não ser mais movido para T, e observamos que as construções com *se*-impessoal como sujeito de infinitivo podem ser tomadas como exemplos da gramática de sujeito nulo parcial do PB em comparação com o

PE, uma LSN consistente. No Capítulo 6, à luz de alguns dos fenômenos presentes nesse capítulo, como a inversão livre do sujeito e as construções com *se*-passivo e *se*-impessoal, apresentaremos os resultados estatísticos de nossa pesquisa sobre a sintaxe do sujeito no português goiano nos séculos XVIII e XIX. Antes, porém, no Capítulo 5 a seguir, discorreremos sucintamente sobre a metodologia empregada na criação, compilação, transcrição e codificação das duas amostras que constituem nosso *corpus*.

## 5 QUESTÕES DE MÉTODO: *CORPUS* HISTÓRICO E TRATAMENTO ESTATÍSTICO DE DADOS LINGÜÍSTICOS

Neste capítulo, apresentamos as metodologias empregadas na criação, na compilação e na codificação das amostras dos manuscritos históricos utilizadas na composição do *corpus* desta tese, criado com o intuito de testar nossas hipóteses de trabalho e munir a tese com evidências empíricas que possam prover e sustentar as respostas para as questões de pesquisa propostas na Introdução. O capítulo está organizado da seguinte forma: a seção 5.1 fornece uma definição de *corpus* linguístico e discute sua importância no estudo da mudança sintática, assim como apresenta: (a) os critérios de coleta das fontes escritas e (b) os critérios e os procedimentos utilizados na compilação das amostras das fontes. A seção 5.2 fornece explicações sobre os modelos estatísticos de análise quantitativa de dados que adotamos para analisar os dados do *corpus* da tese. A seção 5.3 sintetiza o capítulo.

### 5.1 A criação de um *corpus* para um estudo histórico do português em Goiás

O desenvolvimento de aparatos tecnológicos como o gravador, capaz de registrar a produção da fala como realmente ela é, e o computador, equipado com softwares avançados de edição de textos, revolucionou a linguística em seu fazer científico a partir da segunda metade do século XX. Impulsionados por esses avanços tecnológicos na área da computação, os estudos com *corpora* linguísticos estabeleceram-se, nas últimas décadas, como um dos meios mais profícuos para estudar aquisição de primeira e de segunda língua, processamento prosódico e sintático, assim como variação e mudança linguística (GRIES & NEWMAN, 2014). Gries & Newman (2014) definem um *corpus* linguístico como um conjunto de textos escritos ou orais legíveis por máquina, tendo sido produzidos em um ambiente natural de comunicação e compilados com a intenção de serem representativos e balanceados em relação a uma variedade, registro ou gênero particular, além de serem analisados linguisticamente. Gries & Newman (2014) apontam que, seguindo essa tendência, os *corpora* em geral costumam ser:

(a) *legíveis por máquina*: atualmente, a maioria dos grandes *corpora* linguísticos está compilada e armazenada em arquivos de programas de computadores – tendo geralmente como padrão os arquivos com editor XML;

(b) *produzidos em um ambiente natural de comunicação*: isso quer dizer que os textos foram falados ou escritos para algum propósito comunicativo autêntico, não com a finalidade de serem colocados em um *corpus*;

(c) *representativos* (em relação a uma variedade): isso significa que as diferentes partes da variedade linguística que interessam ao pesquisador estão todas manifestadas no *corpus*;

(d) *balanceados*: o que significa que todas as partes de uma variedade não só devem ser amostradas no *corpus*, mas também devem refletir a proporção da parte que compõem nesta variedade, afinal, de modo geral, um *corpus* é criado com a intenção de representar, com o máximo de fidelidade possível, algo maior do que ele.

Os parâmetros acima estabelecidos, dentre outros, nem sempre são ou podem ser estritamente seguidos na criação de um *corpus* linguístico (GRIES & NEWMAN, 2014). Por exemplo, na linguística histórica, área de estudo em que não são aplicados os métodos experimentais e indutivos de coleta e obtenção de dados tipicamente utilizados na criação de *corpora* com dados de falantes nativos, como a introspecção, a eliciação, questionários ou experimentos, o pesquisador precisa tratar do problema da transição entre gramáticas exclusivamente por meio da análise documental de textos que sobrevivem às vicissitudes do tempo (cf. KROCH, 2001). Em geral, os *corpora* modernos são construídos essencialmente com a intenção de fornecer uma amostra representativa de uma comunidade de fala. Na pesquisa histórica, isso raramente é possível, pois registros vernaculares, incluindo a fala, simplesmente não estão mais disponíveis (PINTZUK, TAYLOR & WARNER, 2017). Isso impõe ao pesquisador obstáculos para acomodar a criação de seu *corpus* às condições de *representatividade* e *balanceamento*. Diante disso, os dados da pesquisa em linguística histórica são limitados aos vestígios que os documentos oferecem, de modo que a reconstituição de etapas das gramáticas das línguas dá-se a partir dos sinais deixados nos textos datados de épocas anteriores, dos quais não se podem extrair evidências negativas, a não ser pela ausência de evidência, e os dados devem ser tomados como o objeto empírico para a formulação de hipóteses a respeito dos processos envolvidos no objeto teórico da pesquisa diacrônica, qual seja: a mudança gramatical (cf. PAIXÃO DE SOUSA, 2006).

Van Kemenade & Los (2014) elucidam que nem sempre o estudo em linguística histórica pode ser feito com o uso e o auxílio exclusivo de máquina, especialmente nas

situações em que um pesquisador dispõe de *corpora* anotados. Para as autoras, um acadêmico não pode confiar exclusivamente nos números encontrados nas buscas aos *corpora* eletrônicos. As autoras sugerem que o pesquisador deve verificar nos arquivos dos *corpora* de sua pesquisa se há dados incluídos nas buscas que devam ser excluídos, ou se há dados excluídos nas buscas que precisam ser incluídos no estudo (VAN KEMENADE & LOS, 2014). Ressalta-se que a realização dessa tarefa sugerida pelas autoras não pretende desqualificar o uso de *corpora* eletrônicos em linguística histórica, cujas maiores vantagens são o acesso a um conjunto vasto de dados e a facilidade de se gerarem resultados estatísticos confiáveis, afinal, um *corpus* sintaticamente anotado torna rápidos e relativamente fáceis os aspectos técnicos de coleta de dados. Se o pesquisador sabe o que quer, e o *corpus* inclui categorias relevantes, então ele pode extrair todos os exemplos possíveis em instantes (cf. VAN KEMENADE & LOS, 2014; PINTZUK, TAYLOR & WARNER, 2017). Saber quais dados coletar e como classificá-los, por sua vez, não é algo fácil, é uma tarefa que deve ser guiada por uma compreensão da sintaxe do(s) fenômeno(s) em estudo (PINTZUK, TAYLOR & WARNER, 2017).

De acordo com Pintzuk, Taylor & Warner (2017), a pesquisa em linguística histórica fundamentou-se tradicionalmente no conhecimento dos acadêmicos sobre as línguas estudadas ou em impressões sobre padrões gerais dessas línguas, às vezes com o uso de exemplos individuais de fontes secundárias. Paralelamente, os pesquisadores também alicerçavam seus estudos no uso de um conjunto de exemplos recolhidos de uma amostra. Atualmente, no entanto, a disponibilidade de *corpora* eletrônicos, sintaticamente anotados ou não, transforma completamente o escopo, a precisão e a velocidade das investigações em linguística histórica, instando os acadêmicos a utilizarem *corpora* eletrônicos para estabelecer a base de suas pesquisas (PINTZUK, TAYLOR & WARNER, 2017). Nesse sentido, temos alguns exemplos de *corpora* históricos criados no Brasil sobre a língua portuguesa: o *Corpus Histórico do Português Tycho Brahe* (GALVES, ANDRADE & FARIA, 2017), criado por pesquisadores da Universidade Estadual de Campinas, e o *Corpus do Projeto de História do Português Paulista*, criado por pesquisadores da Universidade de São Paulo.<sup>80</sup>

Cabe-nos destacar que, embora textos escritos sejam resultado de uma atividade que denota o desempenho linguístico do falante, eles podem ser tomados pelo pesquisador gerativista como produtos da linguagem humana que mostram “evidências sobre os

---

<sup>80</sup> O *Corpus Histórico do Português Tycho Brahe* está disponibilizado em: <http://www.tycho.iel.unicamp.br/corpus/>. Acesso em: 29 set 2018. O *Corpus do Projeto de História do Português Paulista* está disponível em: <http://phpp.fflch.usp.br/corpus>. Acesso em: 03 nov 2018.

mecanismos internos da mente e sobre os modos como esses mecanismos operam ao executar as ações e ao interpretar a experiência” (CHOMSKY, 2005, p. 33). Pintzuk, Taylor & Warner (2017) ressaltam que os *corpora* de textos históricos são particularmente importantes no estudo diacrônico-formal, pois não há outra fonte de dados para analisar a língua-I em perspectiva histórica. Para Pintzuk, Taylor & Warner (2017), a utilização de dados quantitativos no estudo diacrônico-formal da sintaxe deve, sobretudo, permitir a análise da variação e da mudança linguística. Desse modo, a fim de coletar, analisar e interpretar adequadamente os dados de um *corpus*, é necessário o pesquisador adotar uma teoria da gramática que lide adequadamente com os conceitos de variação e mudança linguística – conforme fizemos no Capítulo 2 desta tese. Sem essa base teórica, é impossível saber quais dados coletar e como analisá-los e interpretá-los apropriadamente. É nessa perspectiva que os *corpora* linguísticos são ferramentas essenciais para os estudos que descrevem e analisam a sintaxe das línguas naturais a partir da perspectiva dinâmica da mudança sintática.

Como investigadores, os linguistas precisam estar cientes da natureza e das limitações de quaisquer *corpora* utilizados numa pesquisa, afinal, diferentes *corpora* podem apresentar resultados diferentes mesmo quando o objeto de estudo é o mesmo (PINTZUK, TAYLOR & WARNER, 2017). Conforme salientam Pintzuk, Taylor & Warner (2017), a presença ou a ausência de uma construção em um *corpus* não equivale diretamente à sua gramaticalidade ou agramaticalidade. A ausência de uma construção, em particular, pode ser difícil de se interpretar. O tipo de construção crucial para a pesquisa, por sua vez, pode ser raro, ou improvável de ocorrer nos tipos de texto disponíveis, que são necessariamente escritos e restritos ao gênero. A possibilidade de uma lacuna acidental é difícil de excluir. A interpretação da ausência é sempre um julgamento por parte do pesquisador. Evidências positivas também devem ser tratadas com cuidado. Um único exemplo pode indicar a gramaticalidade, mas, dada a natureza do processo de produção do texto, pode ser igualmente um erro ou ser mal interpretado (PINTZUK, TAYLOR & WARNER, 2017).<sup>81</sup>

Diante desse cenário teórico-metodológico, descrevemos, nas subseções seguintes, os critérios e os procedimentos adotados para seleção e compilação das fontes escritas, bem como para transcrição, codificação e comparabilidade das amostras das fontes manuscritas.

---

<sup>81</sup> Pintzuk, Taylor & Warner (2017) mostram, por exemplo, que Biberauer e Roberts (2005) confiam no dado em (i), extraído de Fischer *et al.* (2000, p. 203), para apoiar uma análise particular da mudança envolvendo a ordem VO no inglês médio; Pintzuk, Taylor & Warner (2017) afirmam, entretanto, que uma pesquisa da ordem objeto-verbo-partícula em *corpora* eletrônicos mais robustos não revela mais exemplos dessa ordem.

(i) þe swuch fulðe speteð ut in eni encre eare  
 that.one that such filth spews out in any anchoress's ear  
 'Aquele que vomita tal imundície no ouvido de qualquer âncora.'



### 5.1.1 As fontes escritas

Gries & Newman (2014) enfatizam que a primeira decisão a ser tomada por um pesquisador na criação de um *corpus* linguístico concerne às categorias que irão formar a base de sua amostra. Os autores elencam três possibilidades que os pesquisadores têm na escolha das categorias fundamentais a serem consideradas na composição de um *corpus*, a saber: (i) a categoria de usuários da língua (sexo/gênero, raça, idade, classe socioeconômica, localização geográfica etc.); (ii) a categoria de produtos da língua (língua falada, língua escrita, registros de uso etc.); ou (iii) uma combinação das categorias em (i) e (ii). No caso desta pesquisa, trabalhamos especialmente com a categoria de produtos da língua, mais especificamente com textos escritos em Goiás nos séculos XVIII e XIX na forma de diários. Por outro lado, esbarramos na possível influência que a categoria de usuários da língua possa ter nos resultados obtidos por meio das amostras de nossas fontes escritas. Em sociedades amplamente analfabetas como aquelas que foram colonizadas pela Coroa Portuguesa a partir das Grandes Navegações, grande parte do material escrito era produzido por uma elite e geralmente restrito às suas funções, e isso – bem como alguma variação desse arquétipo – tem algum reflexo nos resultados de qualquer pesquisa de natureza linguístico-diacrônica.

Nessa direção, cabe destacar que, no caso dos dois diários que usamos para o *corpus* desta tese, não só os contextos sócio-históricos de produção de ambos são bastante distintos, como também o sexo/gênero e a origem geográfica dos autores. No Quadro 1, listamos os títulos e a datação dos manuscritos selecionados para a composição do *corpus*, bem como informamos os períodos dos quais foram retiradas as amostras para transcrição e codificação.

Quadro 1: Fontes manuscritas do *corpus* da tese.

| <b>TRADIÇÃO DISCURSIVA</b>        | <b>MANUSCRITO</b>                                                                                  | <b>PERÍODO CODIFICADO</b> |
|-----------------------------------|----------------------------------------------------------------------------------------------------|---------------------------|
| <b>DIÁRIOS<br/>E<br/>MEMÓRIAS</b> | 1. Coleção de notícias da Capitania de Goyaz: diários de viagem do Barão de Mossâmedes (1772-1773) | 1772-1773                 |
|                                   | 2. Memorial de lembrança de Anna Joaquina da Silva Marques (1881-1930)                             | 1882-1884                 |

Fonte: o autor da tese.

Em relação ao contexto sócio-histórico, o primeiro diário, redigido entre 1772 e 1773, foi concebido com um objetivo duplo: (a) registrar para a Coroa Portuguesa e para o futuro governador da região informações sobre os aspectos geográficos, sociais, culturais, políticos e econômicos da capitania de Goiás naquele período; e (b) ser utilizado como ferramenta para a elaboração de um mapa da capitania de Goiás (um dos escribas do texto era cartógrafo). O segundo diário, escrito entre 1881 e 1930, registrou os eventos cotidianos e os fatos históricos que marcaram a vida cotidiana de sua autora e de seus familiares na cidade de Goiás.

O manuscrito *Coleção de notícias da Capitania de Goyaz: diários de viagem do Barão de Mossâmedes (1772-1773)* possui uma cópia (manuscrito 1596) arquivada na Biblioteca Geral da Universidade de Coimbra, Portugal, e foi editado em versão fac-símile por Pinheiro & Coelho (2006) – obra consultada para a edição dos dados desta pesquisa. O manuscrito é dividido em dois diários. Foi compilada e transcrita quase a totalidade dos dois diários para alcançarmos o número total de 1250 sentenças – por essa razão, estabelecemos esse mesmo número de sentenças para a amostra de dados do documento do século XIX.<sup>82</sup>

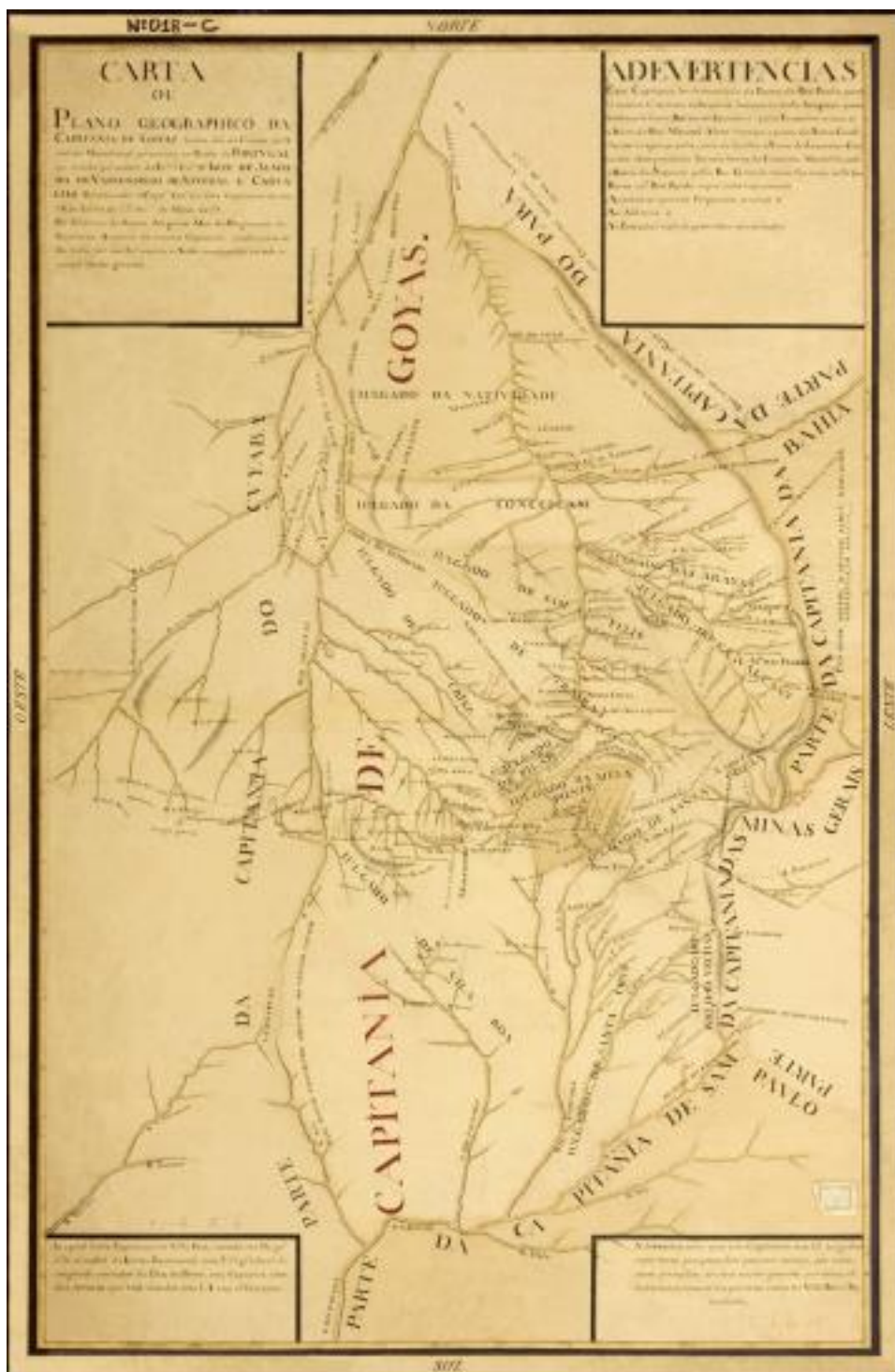
Os diários relatam os caminhos percorridos na capitania de Goiás pelo então governador José de Almeida Vasconcelos Soveral e Carvalho, administrador colonial português nomeado governador da capitania de Goiás por Sebastião José de Carvalho e Melo, o marquês de Pombal. Os diários foram escritos por dois escribas. Um deles é o sargento-mor e cartógrafo português Tomás de Souza Vila Real, autor de um dos mapas da capitania de Goiás (ver Figura 1 na página seguinte). O outro escriba só é identificado no texto por possuir uma grafia distinta da grafia do primeiro. Independentemente da origem dos escribas (provavelmente ambos eram homens portugueses que saíram de Lisboa para a capitania de Goiás junto à comitiva de José de Almeida Vasconcelos Soveral e Carvalho, futuro barão de Mossâmedes), destacamos que o principal local de produção dos manuscritos foi a capitania de Goiás, nos anos de 1772 e 1773, quase um quarto de século depois de oficialmente instaurado o processo colonizador naquela região.<sup>83</sup>

---

<sup>82</sup> Catalogação do manuscrito *Coleção de notícias da Capitania de Goyaz* no sítio da Biblioteca Geral da Universidade de Coimbra: <http://web.bg.uc.pt/cman/show.asp?i=1596&p=122>.

<sup>83</sup> A capitania de Goiás foi oficialmente instituída como parte do processo colonizador em 1748 e integrava em seu domínio político-administrativo o atual estado de Tocantins, que se tornou uma unidade federativa autônoma em 1988, e outras partes de territórios vizinhos. Desse modo, essas regiões acabam por fazer parte da capitania de Goiás como instituição histórica, política e geográfica. Antes da institucionalização colonial, as primeiras povoações na região ocorreram a partir de 1727, depois de Bartolomeu Bueno ter fundado às margens do Rio Vermelho o arraial de Sant’Ana, doze anos depois chamado de Vila Boa, a atual cidade de Goiás.

Figura 1: carta ou plano geográfico da capitania de Goiás.



Autoria: Tomás de Souza Vila Real, 1788.<sup>84</sup>

<sup>84</sup> De acordo com Vieira Júnior (2015, p. 120), há dois documentos originais aquarelados e duas cópias do mapa da Figura 1 catalogados no Arquivo Histórico do Exército, no Rio de Janeiro.

O *Memorial de lembrança de Anna Joaquina da Silva Marques (1881-1930)* está arquivado no Instituto de Pesquisas e Estudos Históricos do Brasil Central (IPEHBC), em Goiânia (Goiás), e é composto por 28 volumes registrados em distintos suportes de escrita, como cadernos, livros e outros materiais avulsos, como folhas de papéis de diversos tipos. O documento foi escrito no decorrer de cinco décadas por Anna Joaquina da Silva Marques, mulher goiana que nasceu em 29 de janeiro de 1855, na cidade de Goiás, e faleceu em 1932, aos 77 anos, vítima de um acidente de carro quando estava a caminho de Itaberaí (cf. CARVALHO, 2008; REIS, 2013). Fotografias do manuscrito original foram usadas na compilação das amostras para o *corpus*. Para totalizar 1250 sentenças, foram compilados e transcritos 28 meses de escrita do documento original: de janeiro de 1882 a abril de 1884.

A memorialista Anna Joaquina era natural da cidade de Goiás e filha de Luísa Joaquina da Silva, mulher que nunca se casou e foi mãe de outros cinco filhos: Esmira, Antônio (Totó), Pacífica, Maria (Mariquinha) e Luísa (CARVALHO, 2008). Duas de suas filhas tornaram-se ilustres professoras da cidade de Goiás: Pacífica Josefina de Castro, mestra Nhola (ou Inhola), cujo apelido deu nome à Escola Estadual Mestra Nhola, localizada na cidade de Goiás; e Luísa Joaquina da Silva Marques, mestra Lili, uma das mestras lembradas por Cora Coralina em seus poemas (cf. CARVALHO, 2008, 2012). Historiadores alegam que a mãe de Anna Joaquina da Silva Marques possivelmente tinha ascendência ameríndia devido ao fato de um testamento escrito em 1841 pelo pai de sua filha mais velha, Esmira, mencionar essa condição social (CARVALHO, 2008). Luísa Joaquina da Silva, a mãe da memorialista Anna Joaquina, nasceu por volta de 1818 e faleceu em 1891, aos 73 anos (CARVALHO, 2008).

O pai de Anna Joaquina era o cônego Pio Joaquim Marques, cujo pai era o também sacerdote Emydio Joaquim Marques, que também era filho de um sacerdote (Gonçalo Marques) e foi o primeiro diretor do Liceu de Goiás (1847-1854), assim como professor de latim, a mais disputada disciplina do Liceu (CARVALHO, 2008). Com o pedido de exoneração do cargo de direção do Liceu por parte do padre Emydio em 1854, quem assumiu como diretor interino foi o cirurgião Vicenti Morette Foggia, padrinho de batismo da memorialista Anna Joaquina da Silva Marques (CARVALHO, 2008). Posteriormente, já na segunda metade da década de 1860, o próprio cônego Pio foi nomeado como inspetor do Liceu, trabalhando ali por quase 9 anos (CARVALHO, 2008). Há registros cartoriais que indicam que a mãe de Pio se chamava Prisca Epifânia da Silva e era natural do Arraial do Carmo (CARVALHO, 2008), na região norte da capitania de Goiás. Ao que tudo indica, a ascendência paterna de Anna Joaquina da Silva Marques já havia se estabelecido há algum

tempo em Goiás e era composta, em três gerações seguidas, por homens de fé e de instrução. A despeito das questões e tensões sociais envolvendo ser filha de um padre naquela época, Anna Joaquina e sua irmã Luísa, também filha do cônego, mantiveram contato constante com o pai até a morte dele, registrada no memorial em 14 de novembro de 1883. Ambas, inclusive, constaram como herdeiras no testamento do cônego Pio (CARVALHO, 2008).

Em tese, a família de Anna Joaquina da Silva Marques configurou-se num modelo à margem do modelo tradicional de família patriarcal e escravocrata nos idos de oitocentos, especialmente por conta da independência financeira e dos casos amorosos de sua mãe e de sua ascendência possivelmente mestiça. Entretanto, os membros dessa família nitidamente gozaram de certo prestígio social para os padrões da época. No que concerne à vida profissional de Anna Joaquina, sabe-se por intermédio de seus manuscritos que ela auxiliava suas irmãs no professorado, porém, tinha possivelmente como principal fonte de renda o corte, a costura e a engomagem de roupas (cf. CARVALHO, 2008, 2012; REIS, 2013).

Na Figura 2, temos um retrato raro de uma das irmãs da memorialista, a mestra Nhola, reunida com uma de suas turmas escolares. A mãe da memorialista pode ser a mulher sentada no centro da última fileira, logo atrás de mestra Nhola, que está em pé, no centro da imagem, rodeada por seus alunos. Carvalho (2012) questiona se nessa fotografia não estaria também a memorialista Anna Joaquina, a saber: a primeira mulher em pé da esquerda para a direita, provavelmente acompanhada de sua irmã Luísa, a mestra Lili.

Figura 2: Turma de primeiras letras de mestra Nhola.



Fonte: Bretas (1991).

Foi a própria Anna Joaquina quem chamou seus escritos de *Memorial de lembrança*. Nas palavras dela, tudo o que era escrito no memorial era para não deixar que nada ao seu redor ficasse no esquecimento; dessa forma, há no texto relatos simples da vida cotidiana, como a visita de um conhecido, um batismo, um casamento, um falecimento, as idas frequentes à igreja para beijar o Senhor dos Passos, assim como relatos que contam a história da cidade de Goiás sob a visão da memorialista, como o incêndio da Igreja da Boa Morte e a recepção dos vilaboenses às notícias da queda do regime monárquico e da proclamação da República, entre outros fatos de natureza histórica.<sup>85</sup> A prática de escrita de Anna Joaquina era acompanhada por seus familiares, que, por poucas vezes, fizeram pequenas interferências no texto. Talvez por isso as notas no memorial são tão lacônicas e nem um pouco intimistas.

Cabe destacar que nossa proposta de examinar e comparar *diários de viagem* e *memoriais/memórias*, cujas principais características são a descrição e a narratividade (cf. SIMÕES, 2007; SIMÕES & KEWITZ, 2017), não inicia ou estabelece uma nova tradição nos estudos linguístico-históricos sobre a língua portuguesa no Brasil. Estudos no âmbito do *Projeto de História do Português Paulista* enquadram diários de viagem e memoriais em um mesmo conjunto de textos dadas as tradições discursivas ligadas ao campo da história no período colonial que os entrelaçam (cf. SIMÕES, 2007; SIMÕES & KEWITZ, 2017).<sup>86</sup>

Embora em momentos anteriores tenhamos adotado o termo *diário(s)* como referência genérica aos manuscritos do Quadro 1, destacamos que essencialmente levamos em consideração para a escolha desses manuscritos como fontes para as amostras do *corpus* da tese o fato de que ambos pertencem a uma mesma tradição discursiva, e não especificamente a um dado gênero. A partir de agora, passamos a chamar esses documentos de *Diários de viagem* e *Memorial de lembrança*.

Em relação às tradições discursivas que circunscrevem a produção de textos ligados ao campo da história na América colonial, Simões (2007) pontua que:

---

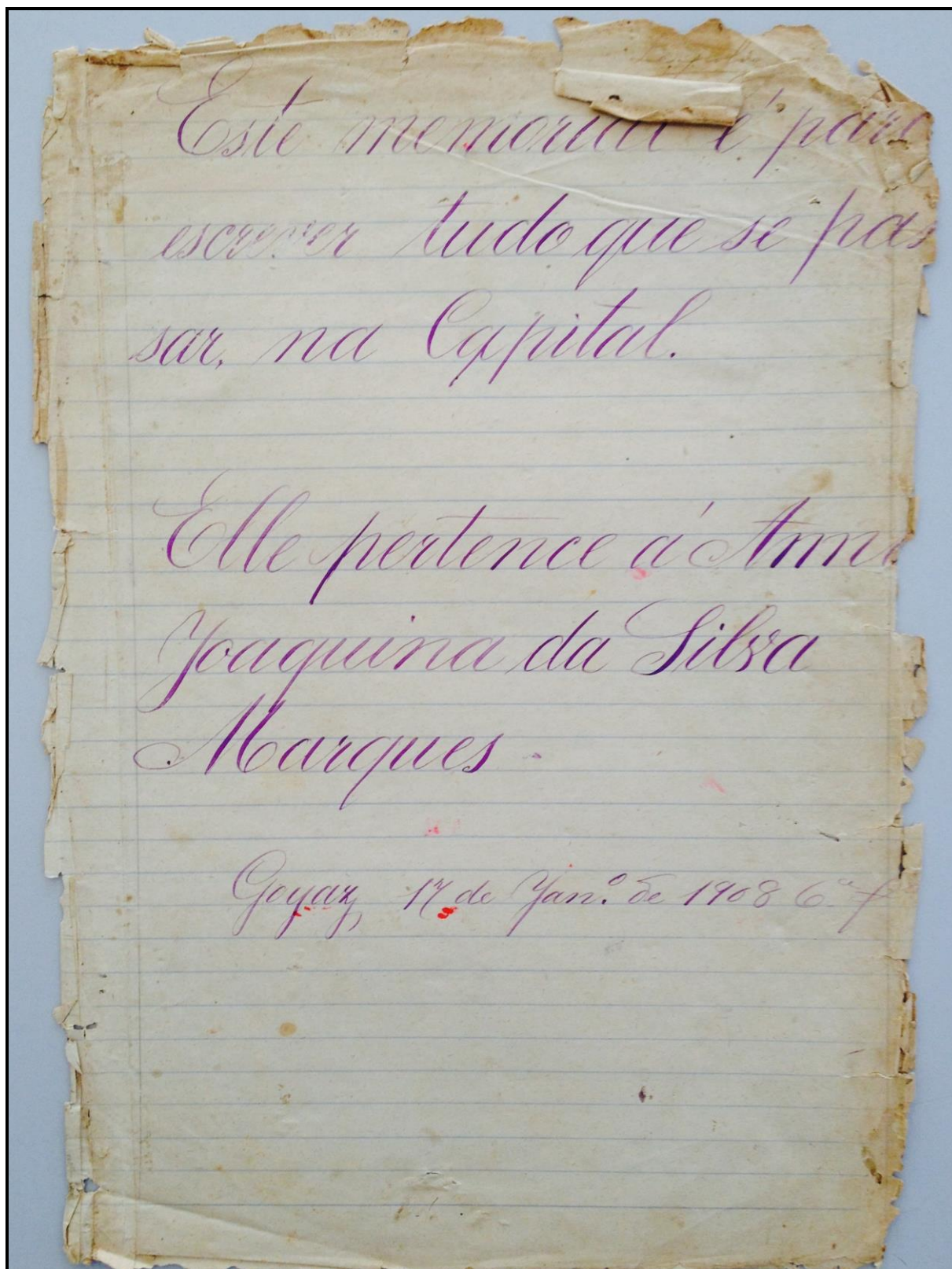
<sup>85</sup> Foi o emocionado relato do incêndio da Igreja da Boa Morte que inicialmente nos chamou a atenção para a ocorrência de sujeitos nulos indeterminados na 3ª pessoa do singular no *Memorial de lembrança*: “Às 3 horas da madrugada pegou fogo na Igreja da Boa Morte. Quando os sinos dobraram o povo todo na cidade despertaram e foram em socorro, mas foi um pouco tarde: as imagens já estavam reduzidas a cinza, o consistório estava todo incendiado, só as Nossas Senhoras. das Dores e do Parto, São Joaquim escaparam de serem queimados, mas o povo forão tão deligentes que quando foi 8 horas já estava apagado o incêndio. Foi um clamor geral que mais nunca se esquecerá desta cena!!! **Tinha enfeitado** a Igreja com melhor gosto possível, para ser exposto o Santíssimo Sacramento, mas Deus não quiz! **Fez as cerimonia no Seminário, e a tarde o Lava-pés também.**”

<sup>86</sup> De acordo com A. Barbosa (2012, p. 594, grifos do autor): “Uma **Tradição Discursiva** é uma **abstração da cultura escrita** em uma dada sociedade **que identifica uma projeção histórico-discursiva** instituída a partir de normas textuais materializadas em um modo recorrente e continuado de uso de elementos linguísticos ou epilinguísticos.”

Na tradição portuguesa, das *cartas-notícia* e dos *diários*, como é o caso da *Carta de Pero Vaz de Caminha* (1500) e do *Diário de Viagem de Pero Lopes* (1530), chegamos às *notícias práticas*, relatos escritos por bandeirantes e outros desbravadores, que registravam a progressão de suas viagens e seus feitos, numa mescla entre *carta* e *diário*. Paralelamente, à medida que se constituía uma nova identidade luso-brasileira, foi preciso registrar a diacronia dos fatos históricos que demarcavam eventos e feitos significativos da história do Brasil. Nesse momento surgem as *memórias históricas*, algumas ufanistas, como a de Frei Gaspar da Madre de Deus (1797), outras mais críticas, como a *Dissertação* de Marcelino Pereira Cleto (1781). (SIMÕES, 2007, p. 202).

Pode-se inferir a partir de Simões (2007) que as memórias históricas no período colonial brasileiro foram feitas por homens eminentes e de vida pública, cujos escritos tinham grande potencial de preservação e circulação. Com isso e considerando também o fato de ter sido escrito no período imperial, se pode argumentar que o *Memorial de lembrança* não se enquadra nessa tradição discursiva, pois sua autora, Anna Joaquina, não era uma eminente mulher de vida pública e seus escritos, restritos à sua forma original até os dias de hoje, parecem ter circulado inicialmente apenas entre os membros de sua família. Ademais, as memórias de Anna Joaquina, escritas no período imperial, indiscutivelmente, não possuem o vigor literário presente na maioria das memórias escritas na tradição discursiva ligada ao discurso da história do período colonial brasileiro. Apesar disso, vamos defender que o *Memorial de lembrança* pode ser visto como uma memória histórica cujas formas e normas linguísticas são herdadas daquela tradição discursiva do século XVIII descrita na citação de Simões (2007). Conforme verificamos em frontispícios e em páginas introdutórias dos cadernos que compõem o documento original, a autora do *Memorial de lembrança* tinha a intenção de que seus escritos fossem tratados como um relato singular de suas memórias sobre a cidade de Goiás: “Este livro é para escrever tudo o que se passa na Capital. Ele pertence a Anna Joaquina da Silva Marques. Goiás, 17 de janeiro de 1908 – 6<sup>a</sup>-f.” (cf. Figura 3 na página seguinte). Desse modo, as diferenças cruciais entre o *Memorial de lembrança* e as memórias históricas do século XVIII são as condições de autoria e do local de produção e circulação desses textos: na condição de mulher sem participação na vida pública de sua cidade, restou a Anna Joaquina registrar suas memórias sobre a cidade de Goiás a partir do âmbito privado.

Figura 3: Memorial de lembrança, 1908.



Fonte: o autor da tese.



Simões & Kewitz (2017, p. 234) contam que, dentre aqueles que registraram suas viagens em diários no período colonial, há alguns que se dedicaram “à tarefa de relatar viagens, sobretudo encomendadas por algum superior em Portugal durante o período colonial”. Conforme relatamos anteriormente, esse é, inquestionavelmente, o caso dos *Diários de viagem*.

### 5.1.2 Estabelecendo os critérios de transcrição, codificação e comparabilidade dos dados

Amostras dos manuscritos mencionados foram transcritas em edição semidiplomática no processador de textos Word e no editor de planilhas Excel, a fim de serem codificadas e analisadas estatisticamente. A edição semidiplomática das amostras realizada no Word caracteriza-se por uma tentativa de prover uma melhor interpretação do texto por meio da eliminação de dificuldades paleográficas e representa uma tentativa de melhoramento do acesso ao texto por meio da divisão de palavras indevidamente grafadas juntas, do desdobramento de abreviaturas e da organização da pontuação. Cambraia (2005) explica que a edição semidiplomática, ou paleográfica, é um grau mediano de decodificação e interpretação de um manuscrito. Nesse tipo de edição, por exemplo, o autor afirma que pode haver a inserção ou a supressão de elementos por conjectura desde que explicitamente assinalada na edição. Cambraia (2005) ressalta que um dos objetivos centrais da edição paleográfica, ou semidiplomática, é tornar o texto editado o mais acessível possível a um público mais amplo.

Pintzuk, Taylor & Warner (2017) asseveram que uma coisa aprendida recentemente pelos pesquisadores que trabalham com *corpora* é que o tamanho do *corpus* é importante. Conjuntos de dados pequenos podem não ser balanceados e representativos. Os autores exemplificam que Pintzuk (1999), com base em um *corpus* coletado manualmente com cerca de 2.300 sentenças, calculou a taxa de movimento do verbo no inglês antigo aumentando em 11,2%, enquanto a taxa quando todo o *York–Toronto–Helsinki Parsed Corpus of Old English Prose* (YCOE)<sup>87</sup> é de 30,4% (cf. HAEBERLI & PINTZUK, 2011).

Em nosso estudo, foram transcritas 1250 sentenças de cada um dos dois manuscritos descritos na Tabela 1, totalizando uma amostra de 2500 sentenças ( $n = 2500$ ), ou seja, de 1250 sentenças por século. Reconhecemos que a quantidade de sentenças analisadas em nossa tese não se compara àquelas possíveis de serem coletadas nos grandes *corpora* anotados,

<sup>87</sup> Disponível em: <http://www-users.york.ac.uk/~lang22/YCOE/YcoeHome.htm>. Acesso em: 17 fev 2019.

como o YCOE, que possui 1.75 milhões de palavras, das quais 1.5 milhões estão anotadas. Ressaltamos, todavia, que nosso *corpus* é o primeiro *corpus* linguístico-histórico sistematizado, balanceado e representativo da sintaxe do sujeito no português em Goiás nos séculos XVIII e XIX, o que pode constituir um primeiro e importante passo no estudo diacrônico-formal da sintaxe desse dialeto.<sup>88</sup> Ademais, é quantitativamente maior do que o conjunto de dados de outros estudos diacrônico-formais sobre a gramática do PB que lograram resultados satisfatórios e sólidos para descrever e analisar a evolução da gramática do PB (cf. DUARTE, 1993 [2018]; CAVALCANTE, 2006; entre outros).

Das 2500 sentenças codificadas, teve-se como unidade de análise dos dados o núcleo oracional representado por um verbo lexical. Os verbos de cada sentença foram classificados de acordo com aspectos de sua natureza morfossintática: transitivos, bitransitivos, intransitivos, inacusativos e impessoais. Ocorrências de perífrases verbais com verbos aspectuais, modais e auxiliares, locuções verbais, verbos elípticos e no infinitivo também foram considerados unidades de análise. Deste modo, de acordo com a quantidade de verbos lexicais – unidades de análise – presentes na oração, orações coordenadas e subordinadas constituem ao menos duas unidades de análise. A codificação das sentenças (unidades de análise) no editor de planilhas consistiu em averiguar o comportamento de propriedades gramaticais relacionadas à sintaxe dos sujeitos nulos nas amostras dos textos. Essas propriedades constituem as *variáveis* do nosso estudo. Dentre essas variáveis, destacamos: a ausência ou a presença fonológica do argumento externo do verbo; o valor de referencialidade do argumento externo; a pessoa e o número gramatical do argumento externo; a pessoa e o número gramatical das flexões verbais; a natureza morfossintática do argumento externo (DP, NP, pronome definido etc.); a posição pré ou pós-verbal do argumento externo manifesto; a presença e a posição pré-verbal de adjuntos adverbiais; e, por fim, a natureza da oração, isto é, se simples, coordenada, subordinada, passiva verbal, construção com *se* etc. (cf. Apêndice, para uma compreensão mais acurada dos procedimentos de codificação dos dados das amostras).

Em relação à temporalidade, a distância temporal entre as amostras transcritas é de pouco mais de um século: 110 anos. Do século XVIII, foram transcritos textos escritos de

---

<sup>88</sup> Um dos trabalhos pioneiros no âmbito do grupo de pesquisa “O Centro-Oeste na história do português brasileiro”, D. Borges (2008) sistematizou um *corpus* histórico sobre o português em Goiás. Em seu estudo, D. Borges (2008) analisa a diacronia de construções causativas no português em Goiás, um dos documentos estudados pelo autor é o *Diário de viagem do Barão de Mossâmedes (1771-1773)*, de Pinheiro & Coelho (2006), também consultado nesta tese. A pesquisa de D. Borges (2008) parte da hipótese de que o isolamento linguístico vivenciado pela população goiana no período colonial deve ser tomado como ponto de partida para o desenvolvimento de análises linguísticas sobre a constituição da gramática do PB em Goiás.

1772 a 1773; e do século XIX, textos escritos de 1882 a 1884. Cabe destacar que os diários listados na Tabela 1 foram os únicos textos desse gênero disponibilizados pelas instituições arquivísticas consultadas ou visitadas em busca de material para esta pesquisa.<sup>89</sup> O padrão de comparabilidade estabelecido para esta pesquisa é a análise comparada da sintaxe dos sujeitos nulos nos textos em tela no interstício de um século para o outro, mais precisamente, de 110 anos. Os achados da pesquisa são sustentados por essas análises.

## 5.2 O tratamento estatístico de dados linguísticos

Nesta seção, apresentamos crucialmente as razões para adotarmos um teste estatístico específico na análise dos dados do nosso *corpus*. Cabe destacar que estatística pode ser definida como um conjunto de procedimentos e princípios para a coleta de dados e análise de informações com o propósito de ajudar um pesquisador, ou pessoas em geral, a tomar decisões quando confrontados com a incerteza frente a um vasto número de dados e informações (cf. UTTS & HECKARD, 2012). Esse conjunto de procedimentos e princípios de coleta e análise de dados requer certa sofisticação e divide a estatística em duas vertentes: descritiva e inferencial.

A estatística descritiva é usada para descrever as características básicas dos dados em um estudo. Fornece resumos simples sobre a amostra, isto é, resume as complexidades dos dados a um pequeno conjunto gerenciável de números, abstraindo-se dos detalhes e dos ruídos para descrever as propriedades gerais básicas dos dados (cf. JOHNSON, 2014). Juntamente com a análise de gráficos simples, como os histogramas, por exemplo, forma a base de praticamente todas as análises quantitativas de dados. A estatística inferencial, por sua vez, caracteriza-se pela utilização de métodos que ajudam o pesquisador a fazer inferências ou previsões baseadas na amostra de dados. A partir dela, portanto, o pesquisador tenta chegar a conclusões que se estendem além dos resultados da amostra obtidos por meio de estatística descritiva. Dos testes empregados em estatística inferencial, a opção por um teste de significância em particular é feita com base nas questões que cobrem os vários aspectos do estudo conduzido pelo pesquisador, o número e o tipo de variáveis envolvidos e o tamanho e a distribuição do conjunto de dados (cf. GRIES, 2014).

Na aplicação dos testes de significância, a noção de hipótese é fundamental. Hipótese em estatística inferencial é uma afirmação que faz uma previsão sobre a distribuição de uma

---

<sup>89</sup> Foram consultadas as seguintes instituições para aquisição dos manuscritos: o Instituto de Pesquisas e Estudos Históricos do Brasil Central (Goiânia-GO) e a Biblioteca Geral da Universidade de Coimbra (Portugal).

variável ou sobre a relação entre duas ou mais variáveis em uma amostra e que tem a estrutura de uma oração condicional *se... então...* (cf. GRIES, 2014). Nesse sentido, o pesquisador lida com uma hipótese alternativa ( $H_1$ ) *versus* uma hipótese nula ( $H_0$ ). A hipótese alternativa é a afirmação sobre um efeito, uma diferença ou uma correlação entre uma ou mais variáveis. A hipótese nula é a contraparte lógica da  $H_1$ : é uma afirmação que prediz a ausência de um efeito, de uma diferença ou de uma correlação entre uma ou mais variáveis (cf. GRIES, 2014). Os testes de significância, por sua vez, são baseados em algumas etapas específicas: primeiro, computa-se o efeito observado nos dados (por exemplo, distribuição de frequência, uma correlação etc.); depois, computa-se a chamada *probabilidade de erro p* para obter a probabilidade de o efeito observado e todos os outros resultados se desviarem de  $H_0$ ; por fim, compara-se  $p$  ao nível de significância (usualmente 5%, isto é, 0.05), de modo que, se  $p$  é menor do que o nível de significância, rejeita-se a hipótese  $H_0$  e aceita-se a hipótese  $H_1$ . Cabe salientar que não se prova  $H_1$ : há apenas uma probabilidade  $p$  de que a  $H_1$  se sobreponha a  $H_0$ .

Há, sem dúvida, um grande pacote de ferramentas estatísticas disponível para o estudo de sintaxe histórica. Dentre os testes de significância para amostras independentes, como é o caso das amostras desta pesquisa (1250 dados do século XVIII e 1250 dados do século XIX), o teste  $t$  de Student é o mais recomendado por possibilitar a análise de tendências em estudos envolvendo a comparação de médias de dois conjuntos de dados independentes, usando amostras representativas (cf. GRIES, 2014). Desse modo, presume-se que as amostras são independentes e a variável de interesse é normalmente distribuída nos dados dos quais as amostras foram colhidas. Testa-se se as médias das variáveis de interesse nas duas amostras independentes diferem significativamente entre si. Se o resultado for significativo, então  $p\text{-value} < 0,05$  (ou  $\text{valor-p} < 0,05$ ).

Além do teste  $t$  de Student, uma das técnicas estatísticas recorrentes nos estudos em sintaxe histórica é a regressão logística (cf. PINTZUK, TAYLOR & WARNER, 2017). A equação usada na regressão logística mostra uma curva S de aumento, ou de diminuição, no decorrer do tempo. Existem dois coeficientes: um define a inclinação da curva S (a taxa de variação) e o outro localiza a curva no eixo do tempo. Defendendo que o reflexo da mudança sintática em *corpora* eletrônicos envolve um longo período de competição entre variantes inovadoras, Pintzuk, Taylor & Warner (2017) argumentam que a trajetória de tais mudanças ao longo do tempo tem sido bem documentada e quase invariavelmente segue uma ‘curva S’, com uma introdução lenta da nova variante, um rápido aumento no período intermediário e, em seguida, uma rejeição à medida que a variante conservadora é perdida. Apesar de a regressão logística ser uma excelente ferramenta para modelar estatisticamente a variação e a

mudança dentro de um eixo diacrônico (cf. KROCK, 1989), destacamos que não adotaremos essa técnica neste estudo, pois o escopo da nossa pesquisa é especificamente a mudança (e não a variação) no eixo diacrônico, que pode ser facilmente capturada sob o escrutínio de testes de significância, como o teste *t* de Student. É possível, no entanto, que, em pesquisas futuras, possamos aplicar a regressão logística para análises que considerem aspectos mais específicos da variação linguística nos dados das amostras de nosso *corpus*.

### 5.3 Síntese do capítulo

De um modo geral, tratamos, neste capítulo, dos impasses vivenciados pelo pesquisador que converge seus interesses de estudo em gramática gerativa e em linguística histórica, ou seja: falamos sobre a necessidade de organização e sistematização de um *corpus* para estudos nessa área e sobre o posicionamento desse pesquisador diante dos dados e resultados obtidos pelo *corpus* de sua pesquisa. De modo mais específico, falamos da seleção das fontes manuscritas utilizadas na composição do *corpus* linguístico-diacrônico da tese (os *Diários de viagem* e o *Memorial de lembrança*), bem como fornecemos informações sobre a natureza sócio-histórica de produção desses manuscritos. Informamos o leitor de que as fontes manuscritas selecionadas para nosso *corpus* linguístico-diacrônico pertencem às tradições discursivas ligadas ao campo da história produzidas na capitania de Goiás nos períodos colonial e imperial, respectivamente. Dadas as distinções que circunscrevem a categoria de usuários da língua nesses manuscritos, vimos que lidar com as informações e os dados relativos aos autores e ao contexto de produção dos manuscritos pode ser relevante para a pesquisa. Sucintamente, apresentamos os procedimentos técnicos e os critérios linguísticos para a compilação, a transcrição e a codificação dos dados do *corpus* (cf. Apêndice, para mais detalhes). Discutimos as vantagens de utilizar testes de significância em dados linguísticos de amostras independentes, como é o caso do nosso conjunto de dados, e, por fim, informamos o leitor de que a significância dos nossos dados será computada pelo teste *t* de Student. No capítulo a seguir, apresentamos os resultados estatísticos dos dados de nossa pesquisa e os analisamos à luz das teorias linguísticas que apresentamos nos capítulos anteriores.

## 6 ASPECTOS DA SINTAXE DO SUJEITO NA HISTÓRIA DO PORTUGUÊS EM GOIÁS: SÉCULOS XVIII E XIX<sup>90</sup>

Este capítulo contém os resultados quantitativos da distribuição de propriedades gramaticais relacionadas ao parâmetro do sujeito nulo codificadas nas amostras do nosso *corpus*, assim como nossa proposta de análise linguística para esses resultados. Foram codificadas 1250 orações de cada um dos dois manuscritos no formato diário/memórias que compõem nosso *corpus*, totalizando 2500 sentenças. Ressalta-se que os documentos foram escritos em Goiás nos séculos XVIII e XIX, respectivamente. O primeiro documento foi escrito majoritariamente por Tomás de Souza Vila Real, escriba português, e por um homem de origem não identificada durante viagens pela capitania de Goiás entre os anos de 1772 e 1773. O segundo documento foi escrito por Anna Joaquina da Silva Marques, mulher nascida na cidade de Goiás em 1855. Suas memórias registradas em diários começaram a ser escritas em 1882 e terminaram em 1930. À luz dos desenvolvimentos teóricos descritos nos capítulos anteriores, argumentamos ao longo deste capítulo que esse documento é – até o momento de realização desta tese – a evidência empírica mais antiga do desenvolvimento do PB como uma LSN parcial. O Quadro 1 fornece informações sobre os documentos codificados.

Quadro 1: *Corpus* analisado.

| <b>Manuscrito:</b>                                                                     | <b>Período codificado:</b> | <b>Número de orações:</b> |
|----------------------------------------------------------------------------------------|----------------------------|---------------------------|
| 1. Coleção de notícias da Capitania de Goyaz: diários de viagem do Barão de Mossâmedes | 1772-1773                  | 1250                      |
| 2. Memorial de lembrança de Anna Joaquina da Silva Marques                             | 1882-1883                  | 1250                      |

Fonte: o autor da tese.

Separamos os dados da nossa amostra, constituídos de orações finitas e não finitas, em seis eixos de análise: em primeiro e em segundo, respectivamente, o eixo dos sujeitos nulos definidos e dos sujeitos manifestos definidos – isto é, sujeitos que valoram os traços- $\phi$  do verbo em T, especialmente o traço-D, e aos quais o verbo faz referência. Crucialmente, uma das características dos textos no formato diário/memórias que analisamos nesta tese é a

<sup>90</sup> Alguns dos resultados reportados neste capítulo foram apresentados anteriormente no formato pôster nos seguintes eventos: *The 91st Annual Meeting of the Linguistic Society of America* (Texas, EUA, 2017) e *Workshop Emerging grammars in language contact situations* (Creta, Grécia, 2017) e aparecem em um artigo publicado – em coautoria com Acrísio Pires – no *Proceedings of the Linguistic Society of America* (2017), o qual foi, posteriormente republicado, a convite, na *Revista Linguística* (2017). Alguns dados foram corrigidos.

ausência total de dados referentes à 2ª pessoa gramatical. Desse modo, todos os dados de sujeitos nulos definidos e de sujeitos manifestos definidos do nosso *corpus* referem-se somente à 1ª ou à 3ª pessoa gramatical. A partir dos dois primeiros eixos, estabelecemos outros dois eixos de análise, a saber: o eixo da inversão livre do sujeito e o da concordância verbal, propriedades sintáticas definidoras do parâmetro do sujeito nulo na maioria das línguas românicas. Em quinto, tem-se o eixo das construções com *se*-passivo e *se*-impessoal. Separamos essas construções das demais porque, seguindo a proposta de Raposo & Uriagereka (1996) apresentada no Capítulo 4, consideramos que o clítico *se*, mesmo ocupando a posição de especificador de T em ambas: (a) atua como um morfema apassivador nas construções com *se*-passivo; e (b) possui leitura indefinida nas construções com *se*-impessoal. Destaca-se que a noção de definitude é essencial para a análise aqui proposta: conforme mencionamos em capítulos anteriores, a hipótese primordial desta pesquisa é a de que ocorreu uma série de mudanças na gramática oitocentista do PB em Goiás devido à perda do traço-D em T nessa gramática, possibilitando a emergência de uma LSN parcial.<sup>91</sup> Em sexto e último, temos o eixo das construções impessoais com verbos meteorológicos e existenciais do tipo *ter/haver*, pois, *a priori*, essas construções são constituídas de sujeitos nulos que não apresentam uma contraparte manifesta em distribuição complementar na língua, isto é, não apresentam um expletivo manifesto.<sup>92</sup>

Na Tabela 1 e na Figura 1, a seguir, apresentamos a distribuição geral dos dados das amostras do nosso *corpus* de pesquisa.

Tabela 1: Distribuição geral dos dados.

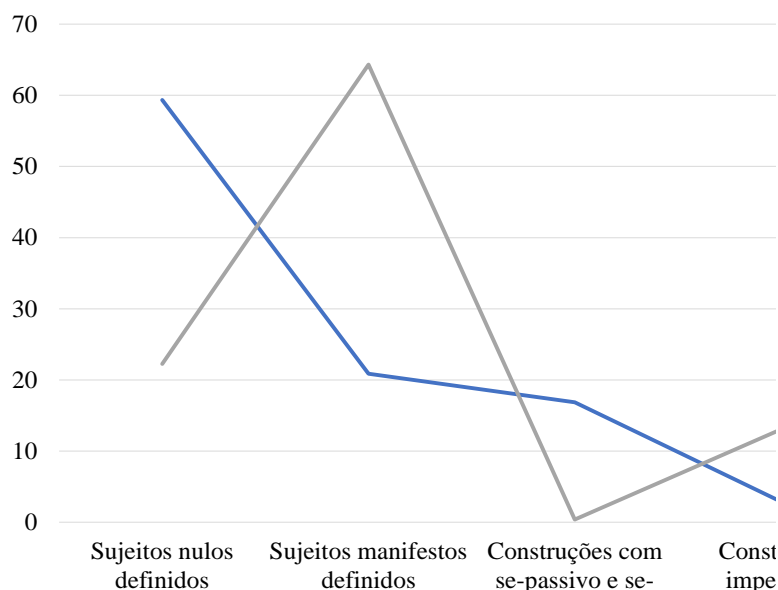
|               | Construções com sujeitos definidos |                     | Construções com <i>se</i><br>(passivo e/ou<br>impessoal) | Construções<br>impessoais |
|---------------|------------------------------------|---------------------|----------------------------------------------------------|---------------------------|
|               | Sujeitos nulos                     | Sujeitos manifestos |                                                          |                           |
| Século XVIII: | 59,36%<br>742(1250)                | 20,88%<br>261(1250) | 16,88%<br>211(1250)                                      | 2,88%<br>36(1250)         |
| Século XIX:   | 22,24%<br>278(1250)                | 64,32%<br>804(1250) | 0,4%<br>5(1250)                                          | 13,04%<br>163(1250)       |

Fonte: o autor da tese.

<sup>91</sup> Conforme mencionamos anteriormente, essa hipótese está fundamentada na proposta de Holmberg (2005, 2010a) de que as LSN parcial não carregam um traço-D em T, e uma consequência natural dessa configuração é que sujeitos nulos podem ser interpretados como indefinidos/impessoais quando entram em relação de concordância com T nessas línguas.

<sup>92</sup> Conferir Naves & Pilati (2018) para uma análise em que certos advérbios (*aqui/ali/lá*) podem atuar como expletivos quase argumentais no PB. A análise das autoras também se baseia parcialmente na proposta de Holmberg (2005, 2010a) para as LSN parciais.

Figura 1: Distribuição geral dos dados.



Fonte: o autor da tese.

O detalhamento dos resultados quantitativos presentes na tabela e na figura acima e a análise linguística desses resultados são apresentados nas seções seguintes deste capítulo, organizadas da seguinte maneira: a seção 6.1 apresenta a distribuição de sujeitos nulos e manifestos definidos, fornecendo evidência empírica robusta da perda de sujeitos nulos na gramática do PB em Goiás no século XIX em comparação com os resultados do século XVIII. Nessa seção, apresentamos nossa proposta de derivação para a checagem do traço-EPP nas gramáticas da língua portuguesa em Goiás nos séculos XVIII e XIX. A seção 6.2 exhibe os resultados relativos à inversão livre do sujeito nos dados do *corpus*, os quais corroboram os estudos que mostram a perda da inversão livre do sujeito no PB com verbos transitivos e sua ocorrência em construções inacusativas e existenciais. A seção 6.3 traz resultados que questionam o papel da concordância verbal na perda de sujeitos nulos nos dados históricos do PB em Goiás, minorando argumentos em favor de uma relação unidirecional em que a perda de morfologia verbal teria sido o fator primário que levou à queda de sujeitos nulos na gramática do PB. A seção 6.4 mostra os resultados relacionados às construções com *se*-passivo em comparação com as construções com *se*-impessoal, evidenciando a perda total das construções com *se*-passivo na gramática de sujeito nulo parcial desenvolvida no século XIX. A seção 6.5 trata das construções impessoais com verbos meteorológicos e existenciais do tipo *ter/haver* e analisa a posição sintática dos adjuntos adverbiais que as acompanham. A seção 6.6, por fim, apresenta uma discussão geral da análise linguística proposta para os



dados históricos do português goiano sob a perspectiva de algumas das implementações teóricas adotadas nas seções anteriores.

### 6.1 Sujeitos manifestos definidos *versus* sujeitos nulos definidos

A análise quantitativa dos dados dos séculos XVIII e XIX mostra um aumento acentuado de sujeitos manifestos definidos de uma média de 20,88% para 64,32% (valor- $p < 0,01$ ). Excluindo as construções impessoais (i.e., construções com sujeitos nulos não argumentais e com verbos meteorológicos, que, em português, não possuem pronome expletivo manifesto), a porcentagem de sujeitos nulos definidos no *corpus* diminuiu de 59,36% para 22,24% entre os séculos XVIII e XIX (valor- $p < 0,01$ ). Em (1)-(2) e (3)-(4), respectivamente, exemplificamos ocorrências de sujeitos manifestos e nulos definidos extraídos das amostras do nosso *corpus* de pesquisa, a fim de que possamos analisar os mecanismos de licenciamento desses sujeitos nessas amostras sob o escrutínio de algumas das teorias linguísticas que apresentamos nos capítulos anteriores, especialmente as propostas de Alexiadou & Anagnostopoulou (1998) e Holmberg (2010a).

#### (1) Sujeitos manifestos definidos no século XVIII

a. Este arraial teve princípio no ano de 1736, por seu fundador Carlos Marinho, o qual vindo do Rio de São Francisco fugiu da Justiça, foi casualmente ali ter, e achando muito ouro naquele lugar se estabeleceu. Hoje estão as suas minas muito decandentes, e o que mais o faz avultar é a casa de fundição, que a criou o Senhor Conde dos Arcos no ano de 1752.

b. Pouco dantes de chegar a este pouso se apartou o Dr. Ouvidor de Sua Excelência o qual o acompanhava do Arraial de Pilar e seguiu a estrada do Julgado das Arraias (...) O Dr. Ouvidor sempre cheio de liberalidades, mandou à mulher do capitão Roque um grande presente de saias, camisas, e caponas, e desta forma satifez Sua Excelência aqueles povos.

c. Na madrugada do dia 8 de agosto voltou Sua Excelência para Natividade pelo mesmo caminho, cuja jornada fez em 5 léguas digo fez em cinco marchas entrando segunda vez naquele Julgado no dia 17 de agosto. (...) Recolheu na casa da sua aposentadoria sem mais outra cerimônia. Residiu Sua Excelência neste arraial até o dia 6 de agosto.

- (2) Sujeitos manifestos definidos no século XIX
- a. Dia 21 eu passei dia em casa de Mariquinha. De tarde eu fui com ela visitar a mulher do tenente Pedro Ribeiro. De noite quando vim fui em casa de Cavalcante.
- b. Dia 7 minha mãe e Esmira forão em casa de Titia. Dia 9 Lulú Pitaluga esteve aqui de noite. Dia 10 eu e Lili fomos nos Cavalinhos. Dia 11 eu e Lili fomos em casa do Imperador (Manuel Antônio). Dia 12 às 5 horas da manhã cazou-se Nhanhã do Alexandrino com Salvador.
- c. Dia 25 Mariquinha minha irmã passou dia aqui. De noite Prescilianna e a Professora Maria Ciriaca estiverão aqui. Nesse dia o Dr. Engenheiro veio ver a casa. Dia 27 a senhora Jozefa esteve aqui e Prescillianna e Sanchinha estiverão aqui, depois fomos beijar Senhor dos Passos. Dia 28 eu e Lili fomos de tarde na Cambaúba, depois fomos na chácara do senhor João Godinho.
- (3) Sujeitos nulos definidos no século XVIII
- a. No dia 18 \_\_ marchou 7 léguas, \_\_ pousou na Taboca. No dia 19 \_\_ marchou 7 léguas e \_\_ tomou quartel no Engenho de Manoel Dias
- b. Na madrugada do dia 24 de agosto saiu Sua Excelência das Arraias para o Regimento de São Domingos, e \_\_ marchando 4 léguas \_\_ pousou no ribeirão da água fria. No dia 25 \_\_ marchou três léguas, \_\_ tomou quartel no morro do chapéu. No dia 26 \_\_ marchou 6 léguas \_\_ fez Rancho no córrego do engenho dos macacos. No dia 27 \_\_ marchou 7 léguas \_\_ pousou no ribeirão do Galheiro.
- c. Mas é certo que alguns paulistas sabiam que neste rio chamado hoje Vermelho e então Goyá, havia muito ouro, principalmente Bartolomeu Bueno<sup>1</sup>, o qual \_\_<sub>1</sub> ofereceu este Descoberto ao senhor Rodrigo Cézar governador e capitão general da Capitania de São Paulo<sup>2</sup>, e com efeito \_\_<sub>2</sub> veio a este continente no ano de 1735, e \_\_<sub>2</sub> deu ao manifesto muitas riquezas.<sup>93</sup>
- (4) Sujeitos nulos definidos no século XIX
- a. Dia 7 foi domingo Totó Ludovico veio passar o dia aqui. depois \_\_ fomos em casa de Mariquinha depois \_\_ fomos no Matadouro. De noite \_\_ fomos ao Mez de Maria. Dia 8 \_\_ fomos todos em casa de Tiasenhora.

<sup>93</sup> Rodrigo Cézar Menezes, além de governador da capitania de São Paulo entre 1721 e 1728, foi governador da capitania-geral do Reino de Angola de 1732 a 1738.

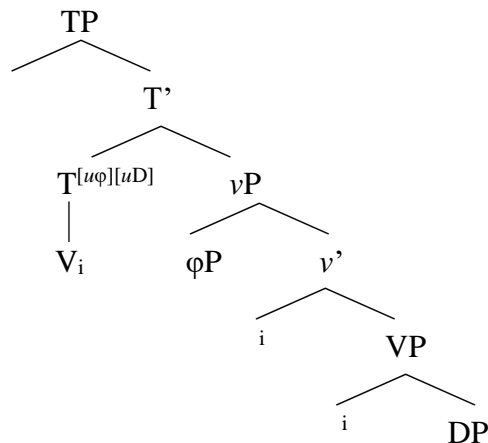
b. Dia 3 eu e minha mãe fomos em casa de Mariquinha depois \_\_ fomos dar pêsames para Filizbina da morte do marido. De tarde \_\_ voltamos \_\_ beijamos Senhor dos Passos quando \_\_ chegamos aqui, estava Dr. Azeredo<sub>2</sub> que \_\_<sub>2</sub> veio ver a casa por ordem de Dr. Theodorico.

c. Dia 23 \_\_ passei dia em casa de Mariquinha. Dia 24 \_\_ fui lá passar dia. De noite veio notícia que Augusto<sub>1</sub>, escravo que foi de Antônio Manoel, matou o novo Senhor<sub>2</sub> que \_\_<sub>2</sub> [o] comprou poucos meses.

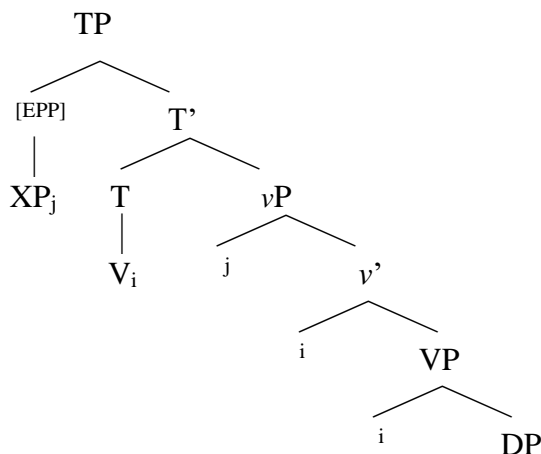
Com base na observação empírica dos dados em (1)-(2) e (3)-(4) e nos resultados estatísticos obtidos, propomos que a gramática oitocentista do PB em Goiás perdeu o traço-D de T em comparação com a gramática setecentista do português em Goiás – esta última provavelmente representa uma gramática com propriedades muito similares com o PE, uma LSN consistente.<sup>94</sup> A perda do traço-D de T teria sido, então, responsável pelo aumento acentuado de sujeitos manifestos em posições temáticas entre os dois períodos, especialmente na 3ª pessoa, como exemplificado em (2b-c). Seguindo Holmberg (2010a), essa afirmação prediz que a gramática oitocentista do PB em Goiás passou a preencher a posição de sujeito (i.e., SPEC-TP) com sujeitos lexicais nas orações declarativas devido à ausência de um traço-D não interpretável em T [*uD*] que pudesse ser valorado por um sujeito nulo em uma relação *probre-goal* – a presença desse traço licenciaria a interpretação definida dos sujeitos. Essa nova gramática, então, passou a se comportar sintaticamente como uma LSN parcial. Considerando essa hipótese e os resultados da distribuição de construções com sujeitos definidos, ilustramos a seguir duas propostas de derivação para a checagem do traço-EPP para os dados do *corpus*: (5a) representa o mecanismo predominante de checagem do traço-EPP para a 3ª pessoa no PB setecentista, hipoteticamente uma LSN consistente; (5b) representa o mecanismo predominante de checagem do traço-EPP para a 3ª pessoa no PB oitocentista, hipoteticamente uma LSN parcial. Destacamos que estamos propondo que o mecanismo predominante de checagem do traço-EPP para sujeitos nulos em (5a) é completamente ausente da gramática do século XIX. O mecanismo predominante de checagem do traço-EPP em (5b) também existia para os sujeitos manifestos no século XVIII, mas com a presença de um traço-D em T – no capítulo seguinte, discutiremos mais especificamente sobre esse mecanismo.

<sup>94</sup> Conforme mencionamos no Capítulo 5, um dos escribas dos *Diários de viagem* é um militar e geógrafo português. Não temos informação, contudo, sobre a identidade do outro escriba, que é identificado no texto somente por conta de sua grafia distinta. É por conta desse escriba desconhecido que tomamos a cautela de não afirmar que a gramática exibida nos *Diários de viagem* é um exemplo da vertente histórica da gramática do PE.

- (5) a. Mecanismo de checagem do EPP para a 3<sup>a</sup> p. no PB setecentista em Goiás



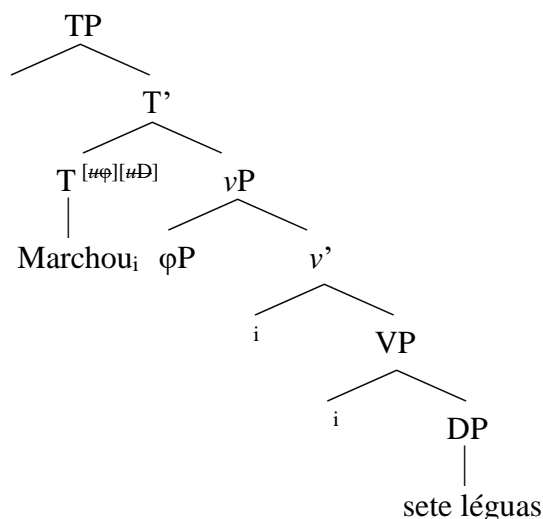
- b. Mecanismo de checagem do EPP para a 3<sup>a</sup> p. no PB oitocentista em Goiás



Em (5a),  $[uD]$  é a rotulação do traço-D de T a ser valorado na checagem de traços na derivação (cf. HOLMBERG, 2010a). Seguindo a proposta de Holmberg (2010a), a valoração de  $[uD]$  e dos demais traços não especificados de T por meio do movimento do verbo para esse núcleo funcional é suficiente para satisfazer o EPP em (5a). Em outras palavras, em (5a), o EPP é um traço-D em T (cf. CHOMSKY, 1995, 2015; ALEXIADOU & ANAGNOSTOPOULOU, 1998; HOLMBERG, 2005, 2010a). Consideramos que os sujeitos nulos de 3<sup>a</sup> pessoa gerados por um mecanismo como em (5a) são especialmente dependentes de um tópico – nulo ou manifesto – de mudança de orientação localizado em C, mediando a relação de coindexação entre o sujeito nulo e seu antecedente, conforme pontuam Holmberg (2005, 2010a) e Roberts (2010) para as LSN consistentes (cf. Capítulo 3). Também adotamos em (5a) a premissa de Holmberg (2005, 2010a) de que sujeitos nulos são  $\phi$ P. Assim, em (5a), seguindo a proposta de Holmberg (2010a), o sujeito nulo  $\phi$ P permanece em sua primeira posição de concatenação e não se move para [SPEC-TP]. A partir da proposta em (5a),

derivamos em (6) e em (7), respectivamente, orações com sujeitos nulos definidos de 3ª pessoa presentes na amostra de dados setecentistas do PB em Goiás.<sup>95</sup> Na estrutura de colchetes em (7), é possível captar uma cadeia de tópicos frasais que licencia o sujeito nulo definido de 3ª pessoa, tal como prevemos para a gramática desse dialeto no século XVIII.

- (6) Derivação de oração declarativa com sujeito nulo de 3ª p. no século XVIII (cf. (3a))



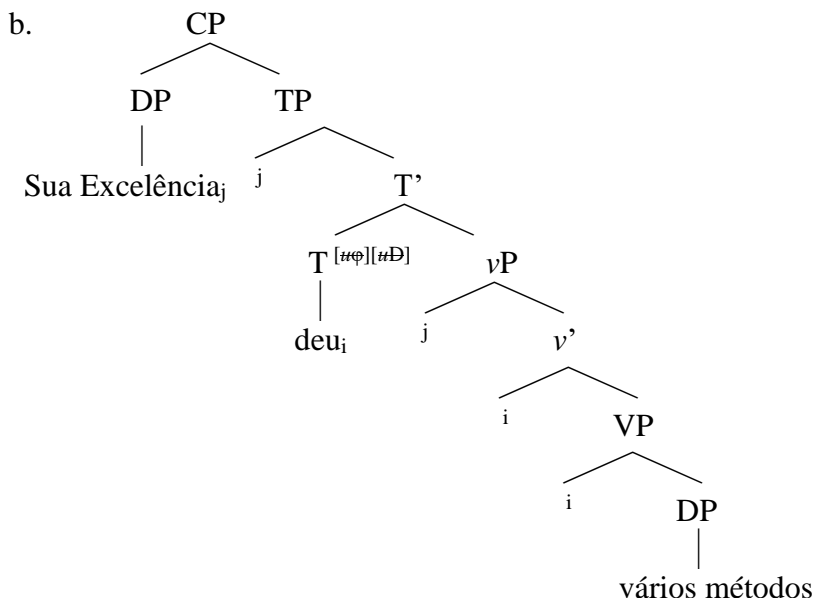
- (7) a. (...) No que Sua Excelência tanto se tem empenhado. No dia 30 último de setembro marchou 4 léguas pousou nos Montes Claros. Pela manhã, marchou sete léguas.  
 b. [<sub>CP</sub> < No que Sua Excelência<sub>1</sub> tanto tem se empenhado > [No dia 30 último de setembro \_\_<sub>1</sub> marchou 4 léguas, \_\_<sub>1</sub> pousou nos Montes Claros]]  
 [<sub>CP</sub> < Ø<sub>2</sub> > [No dia 30 último de setembro φP<sub>2</sub> marchou 4 léguas, φP<sub>2</sub> pousou nos Montes Claros Pela manhã, φP<sub>2</sub> marchou sete léguas]] (1 = 2)

Como mencionado, o traço-EPP em (5a) é satisfeito pelo movimento do verbo para T, movimento este que também é responsável por valorar os traços-φ não especificados de T. Nessa direção, seguindo as intuições de Alexiadou e Anagnostopoulou (1998), propomos que sujeitos manifestos definidos, tanto na 1ª quanto na 3ª pessoa, devem ocupar uma posição A-barra na gramática setecentista, como ilustrado em (8).

<sup>95</sup> Para os casos de pronomes nulos de 1ª pessoa em nossos dados, propomos, seguindo Holmberg (2010a), que cada oração dessa natureza teoricamente possui traços semânticos representando o falante no domínio de C, como ilustramos em (i):

- (i)
-

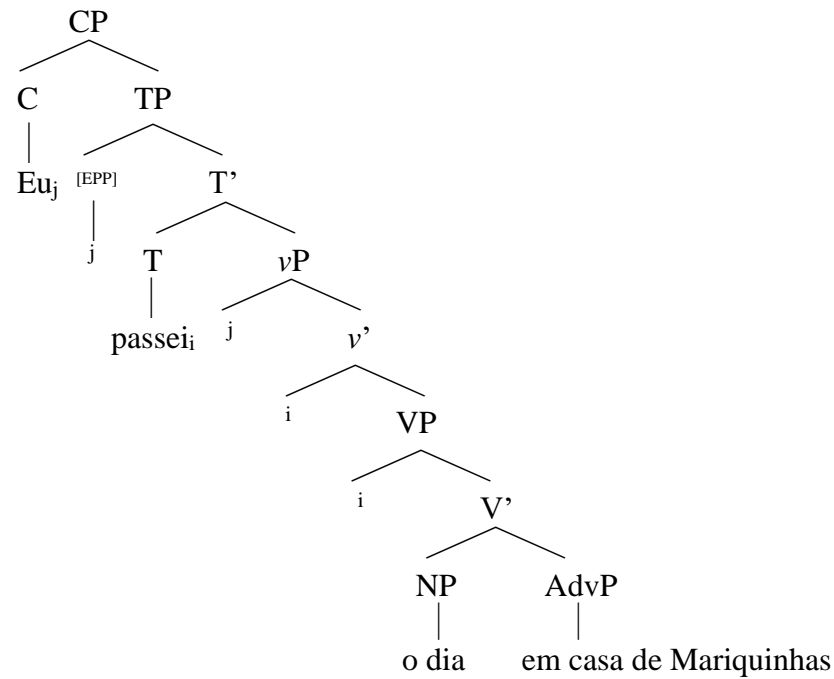
- (8) Oração declarativa com sujeito de 3ª pessoa manifesto no século XVIII
- a. Sua Excelência deu vários métodos para procurarem novos haveres na sua vizinhança.



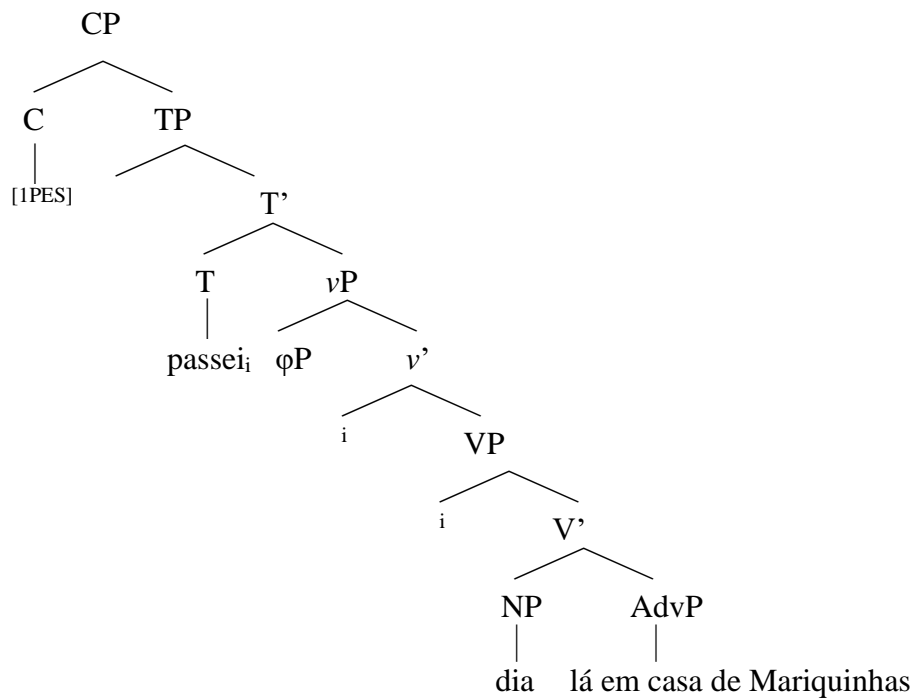
Vejamos agora algumas das implicações da nossa proposta de checagem do traço-EPP na gramática oitocentista. Como ilustrado em (5b), o traço [ $uD$ ] não está presente no conjunto de traços formais de T, de modo que propomos que o EPP passa a ser satisfeito por meio do movimento de um elemento lexical de natureza nominal – um NP ou um DP – para a posição de especificador de T na gramática do século XIX, especialmente na 3ª pessoa. Esse movimento de um XP satisfaz o traço-EPP de [SPEC-TP] nos moldes propostos por Chomsky (1995, 2001, 2015) e Alexiadou & Anagnostopoulou (1998) para línguas como o inglês, onde o EPP é satisfeito pela concatenação (*internal Merge* ou *external Merge*) de um XP – um traço nominal forte, nos termos de Chomsky (2001) – em [SPEC-TP]. Isso, contudo, não significa que a variedade que estamos analisando tenha se tornado uma língua de sujeito não nulo, como o inglês.

Conforme ilustramos em (9) a seguir, a posição de [SPEC-CP] está disponível na gramática do século XIX para receber sujeitos manifestos de 1ª pessoa – assim como estava na gramática do século XVIII para os sujeitos manifestos de 1ª e 3ª pessoa. (cf. HOLMBERG, 2010a). Já nos casos em que o pronome de 1ª pessoa é nulo, propomos que ele está em [SPEC-vP] e é controlado pelos traços semânticos de 1ª pessoa presentes no domínio de C.

- (9) Oração declarativa com pronome manifesto de 1ª pessoa no século XIX



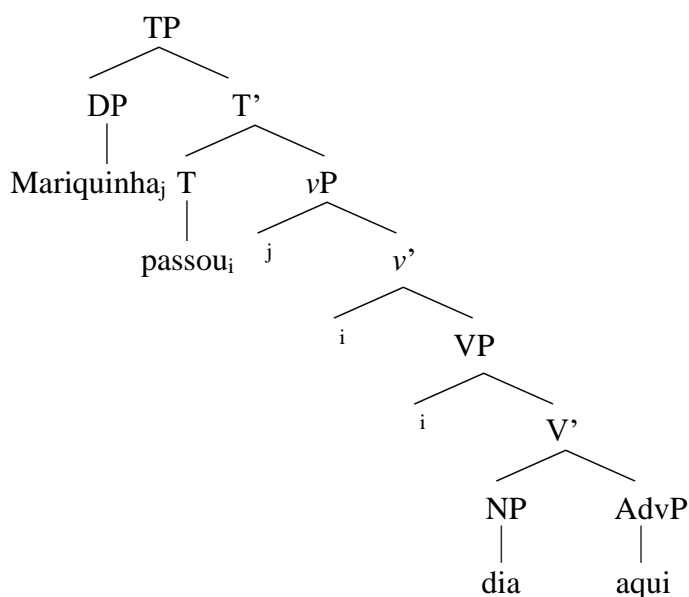
- (10) Oração declarativa com pronome nulo de 1ª pessoa no século XIX (cf. (4c))



Em (5b), adotamos também a proposta de Holmberg (2010a) de que sujeitos definidos de 3ª pessoa não são adjungidos a T por meio de operações em cadeias, razão pela qual devem ser expressos fonologicamente em [SPEC-TP] nas LSN parciais. Destaca-se que um DP lexical não pode ser adjungido a T nas LSN parciais, pois ele possui traços formais cujos valores não podem ser copiados por T, já que T não possui os traços não interpretáveis

necessários (cf. HOLMBERG, 2010a; ROBERTS, 2010). Assim, uma vez que T na gramática do século XIX não possui um traço-D a ser valorado pelo movimento do verbo, um sujeito definido de 3ª pessoa deve conter seu próprio índice de referencialidade e mover-se para [SPEC-TP] nessa gramática (cf. HOLMBERG, 2010a). Com isso, os sujeitos de 3ª pessoa, ao serem sondados por T, são forçados pelo EPP a se concatenar em [SPEC-TP], onde serão expressos como em (11). Isso explica o fato de que a maioria dos sujeitos definidos de 3ª pessoa na amostra de dados oitocentista são itens lexicais manifestos de natureza nominal.<sup>96</sup>

(11) Oração declarativa com sujeito definido de 3ª p. manifesto no século XIX (cf. (2c))



Interessantemente, os escassos dados de sujeitos nulos definidos de 3ª pessoa em orações declarativas na amostra do século XIX têm sua interpretação ancorada a um tópico frasal/discursivo, como ilustrado em (12).

- (12) a. Dia 7 de tarde eu, Nhola e Lili fomos no casamento<sub>i</sub> do senhor João Athanzio, \_\_<sub>i</sub> foi na igreja do Rosário, de noite houve baile.
- b. Dia 18 às 3 horas da tarde fomos todos na casa da rua do Carmo ver o casamento<sub>i</sub> do Senhor João Gustavo, com Ernestina filha do Torquato Barros. \_\_<sub>i</sub> Teve lugar em casa de D. Leonor mer. de S. Pero e \_\_<sub>i</sub> foi celebrado no Carmo.

<sup>96</sup> Muitos dos sujeitos manifestos definidos na 3ª pessoa na amostra oitocentista são NPs nus, isto é, não possuem um determinante aparente. Isso é condizente com o que Holmberg (2010a) prevê para as LSN parcial: maior ocorrência de NPs nus como argumento externo nas orações declarativas. Aliás, ressalta-se que o PB contemporâneo licencia certas estruturas com NPs nus de leitura genérica que são agramaticais no PE, como o NP plural nu a seguir (cf. FERREIRA & CORREA, 2016):

(i) Cachorros morderam o Pedro. (\*PE/√PB)



c. Dia 1º de Julho de 1883 fomos a festa do Imperador de noite houve Theatro (Manda quem pode) depois houve baile<sub>i</sub> em casa do Imperador.   <sub>i</sub> acabou às 3 horas da manhã.

Um corolário da proposta de que o PB oitocentista em Goiás é uma LSN parcial seria o de que os pronomes nulos definidos de 3ª pessoa nas orações encaixadas devem ser comandados por um argumento de uma oração mais alta, e não pela adjunção do pronome a T (cf. HOLMBERG, 2010a). Como dados dessa natureza na amostra do século XIX são em sua maioria casos de orações relativas, ou seja, orações encaixadas em um DP/NP (presentes em línguas de sujeito não nulo, como o inglês), não podemos argumentar que os dados em (13) e (14) a seguir são relevantes para a hipótese central que lançamos sobre a gramática do PB em Goiás no século XIX.

- (13) a. Dia 20 nós todos fomos em casa de Mariquinhas ver procissão de São Sebastião<sub>i</sub> que   <sub>i</sub> passeava na pinguela.  
 c. Depois a mulher do tenente Cavalcante esteve aqui, vimos procissão de Nossa Senhora da Luz<sub>i</sub> que   <sub>i</sub> passou aqui.  
 d. Dia 10 Nhola e eu fomos em casa de Manuel Antônio ver a menina dele<sub>i</sub> que   <sub>i</sub> nasceu que   <sub>i</sub> estava doente.
- (14) a. Antes disso esteve aqui de tarde o senhor Carlos Leitão<sub>i</sub> que   <sub>i</sub> veio visitar o Totó.  
 b. Dia 18 o senhor Joaquim Fernandez esteve aqui de noite, quando ele saiu entrou a Anna Joaquina<sub>i</sub> daqui de frente que   <sub>i</sub> veio despedir para ir para Barra.  
 c. De tarde voltamos beijamos Senhor dos Passos quando chegamos aqui, estava Dr. Azeredo<sub>i</sub> que   <sub>i</sub> veio ver a casa por ordem de Dr. Theodorico.

Curiosamente, nas relativas apositivas em (14), o referente do pronome relativo é o sujeito de uma oração com ordem VS – as quais, como veremos na próxima seção, são restritas aos verbos intransitivos de natureza inacusativa e existencial na amostra de dados do século XIX. Vejamos em (15) uma proposta de derivação para o dado em (14c).

- (15) [CP quando chegamos aqui [TP [T estava<sub>i</sub> [VP t<sub>i</sub> [DP Dr. Azeredo<sub>j</sub> [CP <que> [TP Ø<sub>j</sub> [T veio ver ...]]]]]]]]]]

Em (15), chamamos a atenção para a posição do sujeito da oração matriz, mais precisamente questionamos que elemento ocupa a posição de [SPEC-TP] nas estruturas com ordem VS, com ou sem uma subordinada relativa, na amostra dos dados do século XIX (e, paralelamente, nos dados do século XVIII). Trataremos dessa questão na seção seguinte. Cabe destacar que outros questionamentos também emergem da análise que propomos nesta seção. Por exemplo, com a perda do traço-D de T na gramática oitocentista do PB em Goiás, que levou o EPP a ser predominantemente satisfeito por um XP em [SPEC-TP] nas construções com sujeitos definidos e verbos transitivos, questiona-se se realmente o núcleo flexional T continua a valorar os traços- $\phi$  via movimento do verbo. Outra questão que emerge é se a perda do traço-D de T também teria provocado a perda dos traço- $\phi$  de T na gramática oitocentista do PB em Goiás. No intuito de fornecer uma explicação para esses questionamentos, veremos em uma das seções seguintes que a implementação da proposta que adotamos prevê o empobrecimento do paradigma verbal como consequência da perda do traço-D de T.

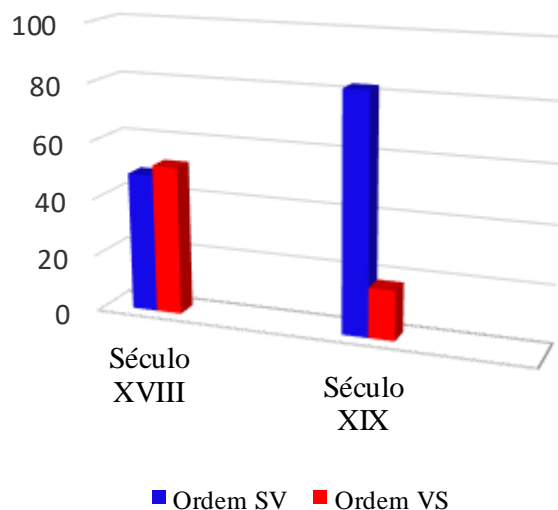
## 6.2 Inversão livre do sujeito

A frequência de dados com inversão livre do sujeito em nosso *corpus* caiu de 51,72% no século XVIII para 17,41% no século XIX, ao passo em que a ordem SV subiu de 48,28% para 82,59%. Na Tabela 2 e no histograma a seguir são apresentados os resultados obtidos.

Tabela 2: Ordenação SV/VS.

|              | Ordem SV          | Ordem VS          |
|--------------|-------------------|-------------------|
| Século XVIII | 48,28%<br>126/261 | 51,72%<br>135/261 |
| Século XIX   | 82,59%<br>664/804 | 17,41%<br>140/804 |

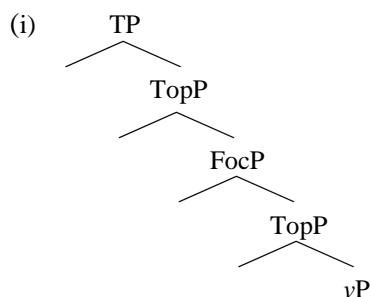
Figura 2: Ordenação SV/VS.



Fonte: o autor da tese.

Destacamos que as ocorrências de sentenças com ordem VS nos dados oitocentistas estão restritas a verbos intransitivos de natureza inacusativa (*ir, entrar, chegar, partir* etc.) e existencial (*estar*). Esses resultados, exemplificados em (15)-(16), vão ao encontro da análise de Berlinck (1988, 1989) de que a ordem VS resiste no PB especialmente com verbos inacusativos (cf. também PILATI, 2006; PILATI, NAVES & SALLES 2017; NAVES & PILATI, 2018).<sup>97</sup>

<sup>97</sup> Utilizamos a expressão ordem VS para nos referirmos a todas as seguintes ordenações: VSO, VOS, OVS, VSX, VXS, XVS, em que O é objeto e X é um outro termo da oração. Apesar de uma abordagem minimalista permitir a realização de análises microscópicas sobre o comportamento dos constituintes de uma sentença, não iremos detalhar nossa análise tratando individualmente de todos esses tipos de ordenações. Cabe destacar também que não estamos considerando em nossa análise as interpretações de tópico e foco das construções com ordem VS. Com base em dados do italiano, Beletti (2004) propôs uma configuração como em (i) para as orações com ordem VS, considerando que tópico e foco nessas estruturas são interpretados nos sistemas de interface:

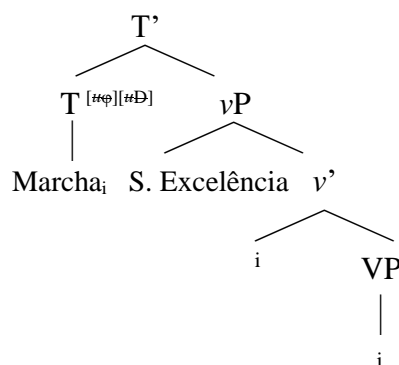


Em relação à configuração em (i), Beletti (2004) pontua que o constituinte que ocupa FocP tem uma entoação específica e, por isso, apresenta interpretação distinta da interpretação em que esse mesmo constituinte ocupa uma das posições de tópico disponíveis (TopP). Sobre a questão de foco nas estruturas com ordem VS no PB, remetemos nosso leitor à tese de Pilati (2006), que afirma que as estruturas VS no PB atual apresentam foco identificacional como resultado da presença de um elemento locativo ou temporal em [SPEC-TP] com escopo

- (16) Ordem SV no século XVIII
- a. O Dr. Ouvidor sempre cheio de liberalidades mandou à mulher do capitão Roque um grande presente de saias, camisas e caponas.
- b. Em todo o tempo que Sua Excelência residiu neste Arraial, que foram nove dias, assistiu aos festejos que lhe fizeram os moradores; (...) O doutor José dos Santos elogiou a Sua Excelência com uma outra e discreta oração latina.
- (17) Ordem SV no século XIX
- a. Dia 17 de noite eu Nhola e Lili fomos no Palácio. Dia 18 eu e Lili fomos a Cambaúba. Dia 19 eu passei dia em casa de Mariquinha. Dia 20 de abril de 1883 sexta-feira eu e Lili fomos beijar Senhor dos Passos.
- b. Dia 27 Nhola de Anna Joaquina veio para conosco acompanhar Senhor dos Passos. Dia 28 eu e Lili fomos em casa do capitão Florambel. Dia 30 eu passei de tarde a cavalo e de noite fui ao Centenário. Dia 31 foi 6ª feira das Dores a procissão passou aqui.
- (18) Ordem VS no século XVIII
- a. Marcha Sua Excelência, do Julgado de Traíras para o de São Félix. Saiu Sua Excelência do Arraial de Traíras no dia trinta de junho pelas sete horas da manhã acompanhado de todos os moradores daquele Arraial, e na retaguarda a Companhia da Cavalaria Auxiliar do Arraial de São José, onde tomou sua Excelência quartel tendo marchado duas léguas.
- b. Ali foi Sua Excelência recebido do vigário com muita decência que ainda que não há Julgado neste Arraial sempre é freguesia, e foi fundado no mesmo tempo que o de Traíras, e com as mesmas circunstâncias. Neste Arraial reformou Sua Excelência a Companhia da Cavalaria Auxiliar.
- (19) Ordem VS no século XIX
- a. Dia 29 foi domingo, fiz anos, ganhei muito, de noite estive aqui Lulu Pitaluga e o L. de Castro. (...) de tarde estive aqui os meninos do Luiz Jardim.
- b. Dia 20 de junho de 1883 4ª-feira partiu para corte o senhor Joaquim Fernandes de Carvalho em companhia do senhor Simãozinho.

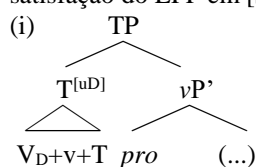
Com base nas implementações teóricas da seção anterior e nos dados com estrutura VS do século XVIII em (18), nossa proposta para a inversão livre no PB setecentista em Goiás é a de que, como o movimento do verbo valorava o traço-D em T, não havia necessidade de o EPP ser checado em [SPEC-TP], de modo que o DP se mantém na primeira posição em que foi concatenado. Em outras palavras, o traço-D em T na gramática setecentista equivale ao EPP em [SPEC-TP] nas línguas de sujeito não nulo, como o inglês. Seguindo Alexiadou & Anagnostopoulou (1998), a ordem VS na gramática setecentista do PB em Goiás é, portanto, uma consequência natural da valoração do traço de definitude em T por meio do movimento do verbo, como ilustrado em (20).

(20) Derivação da inversão livre em oração declarativa na gramática setecentista



Conforme mencionamos no Capítulo 4, Alexiadou & Anagnostopoulou (1998) argumentam contrariamente à postulação teórica de que exista um *pro* em posição pré-verbal nas construções com ordem VS em línguas com morfologia verbal rica. Como evidência para esse ponto de vista, elas mostram que, em grego e em espanhol, todos os sujeitos pré-verbais ocupam uma posição A-barra, a não ser se houver algum gatilho para o movimento A de um DP para [SPEC-TP]. Para analisar os dados com ordem VS no português setecentista em Goiás, seguimos a proposta das autoras para a ordem VS nas LSN canônicas: i.e., argumentamos que não existe *pro*, ou qualquer outro elemento fonologicamente não realizado, em posição pré-verbal nas estruturas VS setecentistas. Ou seja, o movimento do verbo para T, valorando o traço-D de T, é suficiente para satisfazer o EPP nessas estruturas.<sup>98</sup>

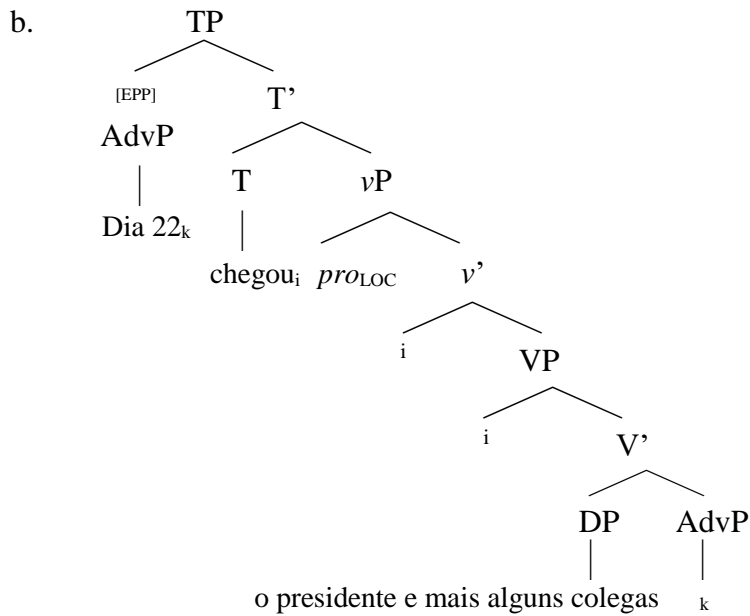
<sup>98</sup> Interessantemente, para dar conta do EPP nas línguas que manifestam o comportamento ilustrado por Alexiadou & Anagnostopoulou (1998) para as estruturas com ordem VS no grego sem deixar de postular a satisfação do EPP em [SPEC-TP] e *pro*, Sheehan (2017, p. 109) propõe a seguinte estrutura.



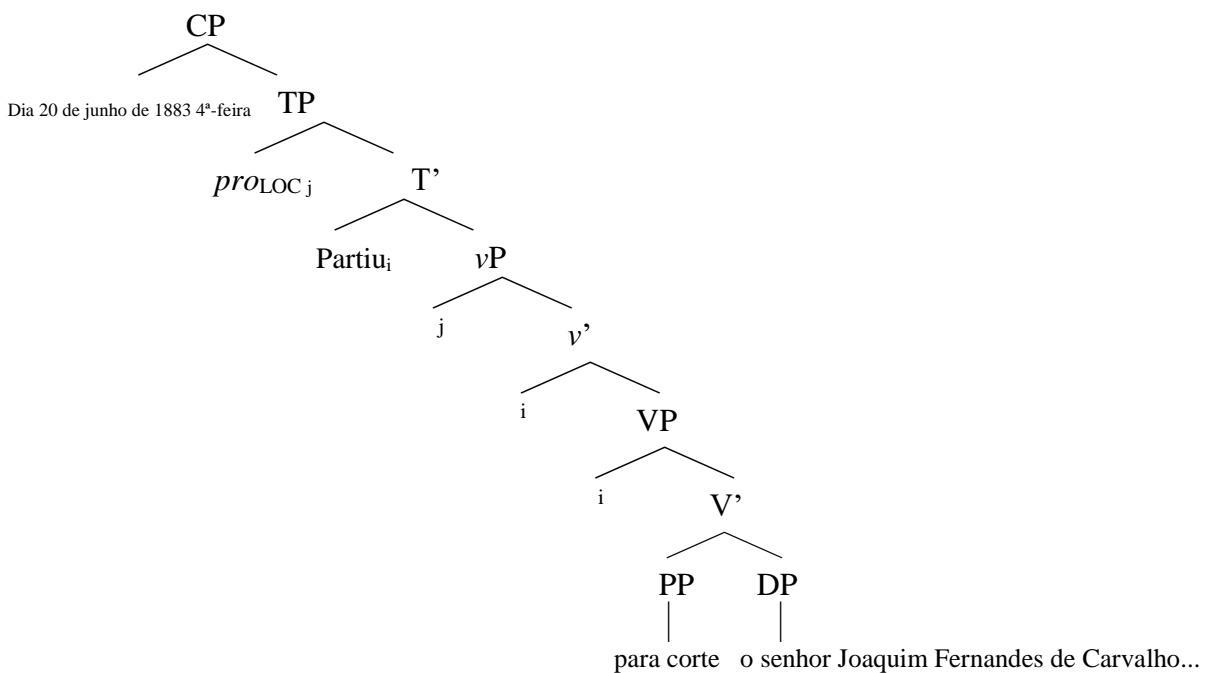
Se na gramática do século XVIII apenas o movimento do verbo para T era suficiente para a valoração do traço-D em T, gerando a ordem VS (inclusive nos casos em que os sujeitos eram nulos: [V [ $\phi$ P]]), uma das consequências naturais da perda do traço-D em T na gramática do século XIX seria o estabelecimento da ordem SV devido à exigência de que algum elemento lexical com índice de definitude ocupe a posição de [SPEC-TP] para satisfazer o traço-EPP, como em (5b) e (11). Essa predição parece certa para os verbos transitivos, que perderam a ordem VS no século XIX. Contudo, a ordem VS persiste nas construções inacusativas e existenciais. Nesses casos, argumentamos que o DP argumento interno desses verbos têm a opção de se manter numa posição abaixo de T. Inicialmente, isso implica ter que assumir que o estabelecimento da concordância do verbo com o DP argumento interno nas construções inacusativas com ordem VS ocorre por conta do movimento do verbo para T, o que sugere que o movimento do verbo para T na gramática oitocentista ainda continua a valorar os traços- $\phi$  de T, embora essa operação não seja mais capaz de expressar/valorar a definitude do sujeito da sentença. Ademais, caso os traços- $\phi$  de alguns verbos finitos na ordem VS ainda sejam valorados pela subida do verbo para T na gramática oitocentista, isso reforça nossa hipótese de que foi a perda do traço-D de T, e não o empobrecimento dos traços- $\phi$  dos verbos finitos, a principal razão para a perda de sujeitos nulos na gramática oitocentista do PB em Goiás. Todavia, caso essa proposta esteja correta, resta uma questão em aberto: como se dá a satisfação do traço-EPP nas construções inacusativas do século XIX, tendo em vista que o movimento de V-para-T não valorava nenhum traço de definitude na gramática desse período?

Considerando a proposta seminal de Pilati (2006) e seus trabalhos posteriores de que sentenças com verbos inacusativos na ordem VS no PB contemporâneo são casos de inversão locativa em que um pronome expletivo locativo [*pro*<sub>LOC</sub>] ou um elemento dêitico locativo ou temporal satisfaz o traço EPP de [SPEC-TP], propomos que os casos de ordem VS registrados no documento do século XIX são casos de inversão locativa com [*pro*<sub>LOC</sub>] ou com um elemento de natureza dêitica satisfazendo o EPP em [SPEC-TP]. Assim, há duas formas coexistindo para derivar as estruturas com ordem VS na gramática oitocentista do PB em Goiás. Na primeira forma, *pro*<sub>LOC</sub> é um expletivo gerado em [SPEC-vP] e permanece nessa posição, possibilitando que um item lexical de natureza dêitica se mova para [SPEC-TP], satisfazendo o EPP, conforme ilustrado em (21). A outra forma para derivar uma estrutura VS na gramática oitocentista consiste basicamente em *pro*<sub>LOC</sub> ser projetado em [SPEC-TP], como ilustrado em (22). Mais do que alternativas teóricas em competição, sugerimos que ambas as estruturas ((21) e (22)) coexistiam na gramática do PB em Goiás no século XIX.

(21) a. Dia 22 chegou o presidente e mais alguns colegas



(22) Ordem VS na gramática oitocentista (cf. (19b))



Ressaltamos que *pro<sub>LOC</sub>* em (21) e (22) é um dado novo no sistema gramatical oitocentista do PB em Goiás. Além disso, enfatizamos que não estamos propondo que a satisfação do traço-EPP por um expletivo nulo ou um elemento dêitico seja uma exigência da gramática oitocentista. Ao contrário, essa inversão é um recurso para a satisfação do traço-EPP apenas quando o DP argumento interno não se move para [SPEC-TP]. Por fim, destacamos que em relação a estruturas como em (21) no PB contemporâneo, Teixeira (2015) declara que a posição de especificador de TP no PB provavelmente é, nos casos de inversão

locativa, preenchida por advérbios locativos/temporais porque eles, assim como os itens lexicais da categoria N, possuem um índice de referencialidade em sua composição interna que possibilita que eles cheguem o traço-EPP no PB. Isso reforça a ideia de Holmberg (2010a) de que o sujeito na posição de [SPEC-TP] em uma LSN parcial, como o PB, deve possuir algum índice de definitude/referencialidade.

### 6.3 A morfologia de concordância

Em capítulos precedentes, mostramos que, desde Permulter (1971), a caracterização das LSN consistentes leva em conta o papel das terminações flexionais do paradigma verbal da língua em recuperar a interpretação definida dos sujeitos nulos. A maioria dos trabalhos desenvolvidos no âmbito da teoria gerativa – quer seja no modelo de princípios e parâmetros, quer seja no programa minimalista – tem como norte a ideia de que as terminações flexionais do verbo possuem um papel de destaque no licenciamento e na identificação de sujeitos nulos nas LSN consistentes, especialmente nas LSN românicas, como o espanhol e o português peninsulares e o italiano. Diante disso, dentro do escopo de nossas amostras, buscamos analisar o possível papel da morfologia verbal na queda de sujeitos nulos definidos na gramática oitocentista do PB em Goiás em comparação com a gramática setecentista.

Em primeiro lugar, destacamos que não houve registro significativo de ausência de concordância de flexão entre verbo e sujeito nos dados do século XVIII no que concerne à primeira e à terceira pessoa – reiteramos que nosso *corpus* não possui dados da 2ª pessoa. Foram registrados quatro casos de ausência de concordância entre um sujeito de 3ª pessoa e o verbo no manuscrito setecentista. Três desses casos envolvem DPs coordenados em estrutura VS, tendo ocorrido apenas um caso de DP coordenado em estrutura SV. Pode-se dizer que nas estruturas VS a relação de concordância é parcial, isto é, ocorre apenas entre o verbo e o primeiro elemento da coordenação, como no exemplo em (23a) a seguir. No único caso de ausência de concordância em estrutura SV (23b), a causa para isso pode ser o fato de o DP coordenado ser pesado, isto é, apresentar três palavras.

- (23) a. Sua Excelência intentou ir ao regimento do Duro; porém, como **veio** ao Arraial de Natividade o vigário e o capitão Roque, regente dos índios, que ali ainda se conservam aldeados, desistiu desta ideia.
- b. O seu esmoler, ânimo, e a sua liberalidade, [o] **fez** despender nesta digressão mais de 4 mil cruzados.



Esses casos escassos de ausência de concordância verbal sugerem que a gramática setecentista tinha morfologia verbal capaz de referir e recuperar as pessoas do discurso por meio das terminações flexionais de seus verbos finitos. Vejamos em (24) exemplos de concordância verbal no plural na gramática setecentista.

- (24) a. No dia sete de julho marchou três léguas, e tomou quartel no Arraial de São Félix. Foi Sua Excelência recebido nesse Arraial com grande aplauso; os moradores com o Intendente<sub>1</sub> o vieram buscar<sub>1</sub> ao caminho; o senado o esperou na grande ponte, que acaba na entrada do Arraial, e debaixo do pátio o conduziram<sub>1</sub> à Igreja Matriz, onde recebeu o Vigário com as decentes cerimônias do costume.
- b. O Intendente e os mais oficiais da Casa da Fundação<sub>1</sub> fizerão<sub>1</sub> fazer no tempo que Sua Excelência ali residia quatro óperas em um excelente Teatro, agradeceu-lhe muito o obséquio; mas sempre lhe disse que era descontente com as despesas que tinham<sub>1</sub> feito.
- c. Hoje estão<sub>1</sub> as suas minas muito decandentes<sub>1</sub>, e o que mais o faz avultar é a Casa de Fundação, que a criou o senhor Conde dos Arcos no ano de 1752.

Em relação aos dados oitocentistas, menos de 12% das orações com um sujeito pluralizado manifesto ou nulo, isto é, 54 dados de um total de 464, não apresentaram concordância explícita entre o verbo e o sujeito, conforme ilustramos na Tabela 3 a seguir.

Tabela 3: Concordância com verbo no plural no século XIX.

|            | Concordância | Ausência de concordância |
|------------|--------------|--------------------------|
| Século XIX | 88.36%       | 11.64%                   |
|            | 410/464      | 54/464                   |

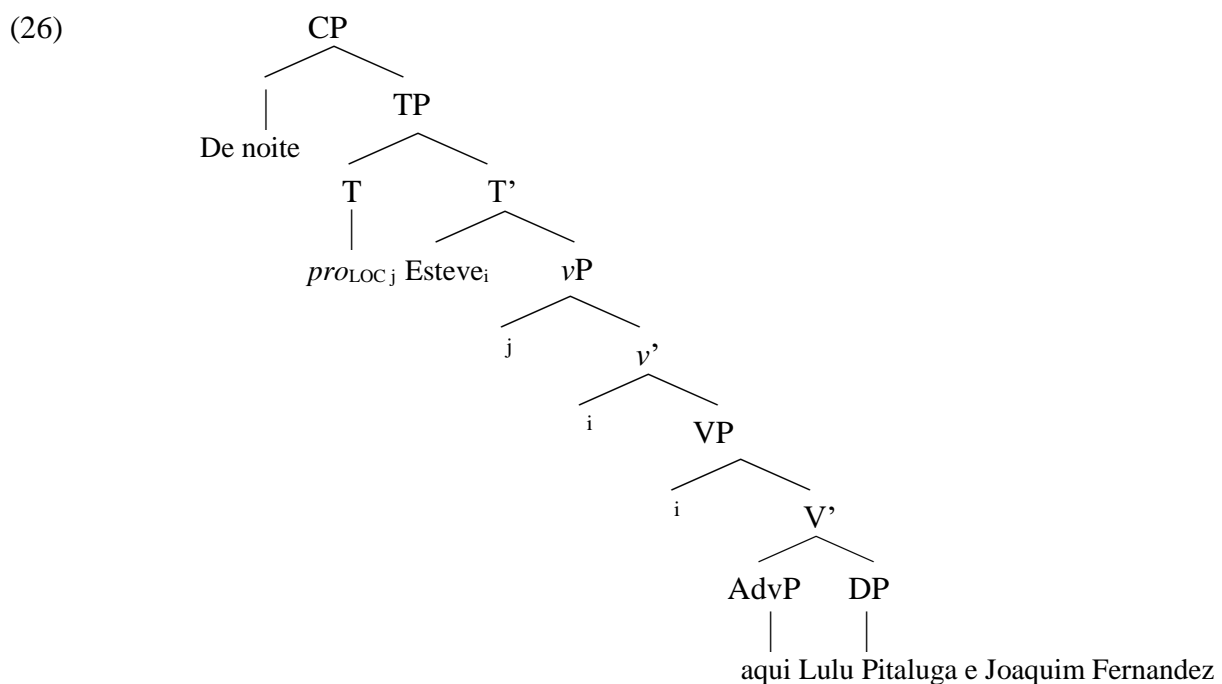
Fonte: o autor da tese.

Destacamos que na codificação das 464 sentenças amostradas na Tabela 3 foram considerados todos os tipos de verbos não impessoais (transitivos, intransitivos, inergativos etc.), com sujeito manifesto ou nulo. Curiosamente, todos os casos de sujeitos nulos referenciais no plural apresentaram concordância manifesta com o verbo, independentemente da natureza semântica do verbo. Ressaltamos, por outro lado, que todos os casos de incompatibilidade de concordância na gramática oitocentista envolvem crucialmente a 3ª pessoa do plural e verbos de natureza existencial, como em (25a-b), ou de natureza

inacusativa, como em (25c-e). A maioria desses dados são estruturas SV (57,4%) com DPs pesados e coordenados, como no exemplo em (25c). No caso da gramática setecentista, tivemos um único caso dessa natureza dentre os quatro casos encontrados de ausência de concordância. Essa predominância de construções com DPs pesados e coordenados em posição pré-verbal deve estar crucialmente ligada ao estabelecimento da ordem SVO – que mencionamos na seção anterior – na gramática oitocentista.

- (25) Ausência de concordância com verbo existencial e inacusativo no PB oitocentista
- a. Dia 22 foi 4<sup>a</sup>-feira de cinza. De noite esteve aqui Lulu Pitaluga e Joaquim Fernandez.
  - b. Dia 14 esteve aqui o Luiz de Castro e Lulu Pitaluga de noite.
  - c. Dia 17 Lulu Pitaluga e o Cavalcante veio aqui.
  - d. De noite veio aqui as filhas do Carneiro.
  - e. Dia 22 chegou aqui o Ilustríssimo Excelentíssimo Senhor Dr. Antônio Gomes Pereira Júnior, Presidente para Goyaz com sua digna consorte; e mais alguns seus colegas chefe de polícia.

Em (26), derivamos sintaticamente a ausência de concordância numa construção existencial com ordem VS (estrutura que representa 42,6% dos dados sem concordância verbal do manuscrito oitocentista).



Em relação à ausência de concordância verbal em exemplos como os da derivação em (26), é possível afirmar que a concordância é parcial, ou seja, é estabelecida entre o verbo e o primeiro elemento nominal coordenado. Haveria aí, então, a coexistência da concordância parcial e da concordância plena em estruturas VS com DPs coordenados, a qual também é atestada em três dados escassos na amostra setecentista. A diferença crucial na ausência de concordância dos dados do século XIX é verificada, então, em ocorrências como em (25d), na qual se verifica a ausência de concordância entre o DP plural (não coordenado) e o verbo em uma oração com estrutura VS. Ao todo, há cinco construções desse tipo no manuscrito oitocentista (cf. mais exemplos em (27)).

- (27) a. Dia 20 faleceu duas crianças com coqueluche.  
b. De tarde esteve aqui os meninos do Luiz Jardim.

Consideradas as propriedades sintático-semânticas dos verbos inacusativos e existenciais e de seus argumentos, uma explicação possível para a ausência de concordância em casos de ordem VS com DPs plurais com a estrutura DET + SN pode ser que a valoração dos traços- $\phi$  de T pelo DP argumento interno pluralizado começa a deixar de ser uma exigência nas construções com esses verbos na gramática oitocentista (apesar de ser produtiva).

Nossos dados do século XIX sugerem, por meio da 1ª e da 3ª pessoa do plural, que a morfologia verbal não teria tido um papel significativo na mudança na gramática do PB no que concerne à sintaxe do sujeito, diferentemente da proposta de M. E. Duarte (1993 [2018]) ilustrada na Tabela 4.

Tabela 4: Evolução nos paradigmas flexionais do PB por M. E. Duarte (1993 [2018]).

|            | <b>PRONOMES<br/>NOMINATIVOS</b> | <b>PARADIGMA 1<br/>SÉCULO XIX</b> | <b>PARADIGMA 2<br/>SÉCULO XX/1</b> | <b>PARADIGMA 3<br/>SÉCULO XX/2</b> |
|------------|---------------------------------|-----------------------------------|------------------------------------|------------------------------------|
| <b>1PS</b> | eu                              | <i>canto</i>                      | <i>canto</i>                       | <i>canto</i>                       |
| <b>1PP</b> | nós<br><i>a gente</i>           | <i>cantamos</i><br>-              | <i>cantamos</i><br><i>canta</i> ∅  | <i>cantamos</i><br><i>canta</i> ∅  |
| <b>2PS</b> | tu<br><i>você</i>               | <i>cantas</i><br>-                | <i>cantas</i><br><i>canta</i> ∅    | <i>canta(s)</i><br><i>canta</i> ∅  |
| <b>2PP</b> | vós<br><i>vocês</i>             | <i>cantais</i><br>-               | -<br><i>cantam</i>                 | -<br><i>canta(m)</i>               |
| <b>3PS</b> | ele, ela                        | <i>canta</i> ∅                    | <i>canta</i> ∅                     | <i>canta</i> ∅                     |
| <b>3PP</b> | eles/elas                       | <i>cantam</i>                     | <i>cantam</i>                      | <i>canta(m)</i>                    |

Fonte: M. E. Duarte (1993 [2018]), p. 85).

Tendo em vista os achados de nossos dados diacrônicos, não tomamos o Paradigma 2 e o Paradigma 3 de Duarte (1993 [2018]) como argumento para a queda de sujeitos nulos na variedade oitocentista do PB em Goiás, pois: (i) não temos evidência empírica de que o pronome *a gente* (presente nos paradigmas 2 e 3) possa ter sido integrado no sistema pronominal daquela variedade já no século XIX, assim como (ii) não temos evidência empírica significativa da perda da morfologia de número da 3ª pessoa do plural indicada no Paradigma 3. Com base nos paradigmas 2 e 3, M. E. Duarte (1993 [2018]) sugere que a inserção do pronome *a gente* (concorrendo com a 1ª pessoa do plural ‘*nós*’) e a perda da morfologia de número da 3ª pessoa do plural teriam, de algum modo, contribuído para o aumento de sujeitos manifestos no contexto da retomada pronominal em dados de peças de teatro escritas no Rio de Janeiro no século XX.<sup>99</sup> Cabe destacar que, como nosso *corpus* não possui dados da 2ª pessoa, a evolução nos paradigmas flexionais do PB proposta por M. E. Duarte (1993 [2018]) na Tabela 4 só poderia explicar o uso substancial de sujeitos manifestos no século XIX em Goiás se tomássemos uma possível perda da morfologia de 2ª pessoa – *tu/vós* – como suficientemente capaz de desencadear a mudança que levou à perda do estatuto do PB de LSN consistente. Isto é, a evolução nos paradigmas flexionais do PB detectada por M. E. Duarte (1993 [2018]) para a variedade carioca do PB no século XX só seria adequada para sugerir uma mudança paramétrica na gramática do português oitocentista em Goiás se a inserção dos pronomes *você* e *vocês* referentes, respectivamente, à 2ª pessoa do singular e do plural no paradigma pronominal do PB tivesse ocorrido no século XIX, um século antes do que é constatado para a maioria das variedades do PB, e fosse crucialmente responsável pela perda no PB do estatuto de LSN consistente.

Conforme mencionamos no Capítulo 4, estudos sobre a diacronia do pronome *você* constata que seus primeiros usos como pronome pessoal em variedades do PB surgiram nas décadas finais do século XIX (cf. LOPES & CAVALCANTE, 2011; LOPES, RUMEU & CARNEIRO, 2013; RUMEU, 2013; entre outros).<sup>100</sup> Lopes & Duarte (2003, p. 8) alegam que “enquanto no Brasil a forma gramaticalizada *você* invade paulatinamente o sistema pronominal no fim do século XIX, anunciando uma concorrência com o pronome *tu*, em Portugal, o pronome *tu* passa a suplantar os outros usos nesse mesmo período”. Esses estudos sugerem, portanto, que é possível que o pronome *você* também fosse utilizado como pronome pessoal na variedade oitocentista do PB em Goiás. Para M. E. Duarte (1993 [2018]), contudo,

<sup>99</sup> Destaca-se que não há evidência da pronominalização de *a gente* em nossos dados do século XIX.

<sup>100</sup> Lopes, Rumeu & Carneiro (2013, p. 200) afirmam que em cartas cariocas escritas entre 1880 e 1900: “as duas estratégias (*tu* e *você*) apareceram como formas variantes no mesmo contexto de uso”.

nenhuma mudança substancial em relação à expressão do sujeito na gramática da língua parece ter ocorrido devido a esses primeiros usos de *você* como pronome pessoal no século XIX. A autora afirma que uma mudança significativa no licenciamento de sujeitos nulos no PB só é detectada na década de 30 do século XX (de 25% a 46%, comparando, respectivamente, os anos de 1918 e 1937), quando a coexistência das duas formas pronominais (*tu* e *você*) desaparece nas peças analisadas pela autora, sendo praticamente exclusivo o uso de *você*:

(...) uma forma verbal com desinência zero ou de 3ª pessoa do plural – que constituem sincretismos – não afeta a identificação do sujeito nulo de segunda pessoa. A súbita queda, de 69% de sujeitos nulos de 2ª pessoa em 1918 para 25% em 1937, coincide com a opção pela forma já pronominalizada *você(s)*. (DUARTE, 1993 [2018], p. 89).<sup>101</sup>

Destaca-se que, apesar de os dados datados da década de 30 do século XX em diante no *corpus* de Duarte (1993 [2018]) apresentarem a opção pela forma pronominalizada *você*, a autora pondera, com base na peça de um autor não muito popular (Humberto Couto), que na fala carioca a forma *tu* e *você* provavelmente continuaram intercambiáveis no português carioca da década de 30 e após. Em relação à morfologia da 2ª pessoa do plural, cabe destacar que a substituição do pronome *vós* pelo pronome *vocês* também é constatada em muitas variedades do PE (cf. CINTRA, 1986), que permanece inalteradamente uma LSN consistente. A nosso ver, os estudos que indicam que os primeiros usos de *você* no século XIX como pronome pessoal não afetaram a sintaxe do sujeito da língua naquele período e o fato de que o pronome *vós* também foi substituído em variedades do PE minoram a possibilidade de que a morfologia de 2ª pessoa poderia ter sido o fator primário para a perda do estatuto de LSN consistente do PB.

Com base nos resultados de nossos dados e nas perspectivas teóricas que adotamos nesta tese, propomos que, na verdade, um mecanismo de natureza sintático-semântica, mais especificamente um traço categorial de definitude [i.e, um traço-D], teria tido um papel decisivo no desenvolvimento do PB em Goiás como uma LSN de sujeito nulo parcial no século XIX (e na conseqüente perda de morfologia verbal constatada na gramática da língua no século XX). Para sustentar nossa hipótese, retomamos o postulado de Roberts (2010),

<sup>101</sup> Caso a perda de sujeitos nulos no PB realmente tenha como uma das suas principais causas a perda da morfologia de 2ª pessoa, a citação de M. E. Duarte (1993 [2018]) gera um questionamento interessante a respeito da possibilidade de variedades do PB que ainda registram o uso do pronome *tu* e da desinência de 2ª pessoa singular -s, como variedades do Nordeste e do Sul, produzirem mais dados de sujeitos nulos do que as demais variedades do PB.

reproduzido em (28), sobre o papel do traço-D na especificação dos traços- $\phi$  nas LSN consistentes.

- (28) Se uma categoria  $\alpha$  tem um traço-D, então todos os traços- $\phi$  de  $\alpha$  são especificados.  
(ROBERTS, 2010, p. 82)

Um corolário do postulado em (28) é o de que a riqueza morfológica e a valoração dos traços- $\phi$  são dependentes da presença de um traço-D no núcleo funcional T numa dada língua de sujeito nulo. Diante dessas colocações, sugerimos que o empobrecimento atestado na morfologia verbal na gramática do PB, especialmente se ele se tornou mais proeminente principalmente no século XX, como observado por M. E. Duarte (1993 [2018], 2000), pode ser, na verdade, o resultado da perda histórica do traço-D em T na gramática do PB, ao invés de um gatilho primário para ela. Desse modo, o empobrecimento na concordância verbal no PB seria uma consequência da mudança na configuração paramétrica da língua, provocada pela perda do traço-D de T.

#### 6.4 Construções com *se*

Um dos argumentos mais recorrentes na literatura gerativista sobre a ausência das construções com *se*-passivo no PB propõe que o empobrecimento da concordância nos paradigmas flexionais é responsável pela perda dessas construções no PB coloquial. Com base nos resultados estatísticos desta pesquisa e sob a luz de aspectos das análises de Raposo & Uriagereka (1996), propomos nesta seção que a perda das construções com *se*-passivo no PB oitocentista é um resultado da perda do traço-D de T da gramática desse dialeto. Argumentamos, ainda, que, concomitantemente à perda das construções com *se*-passivo, o PB em Goiás passou a licenciar sujeitos nulos arbitrários, constituindo não somente a mais antiga evidência de pronomes nulos impessoais na 3ª pessoa do singular no PB, mas também uma forte evidência para a hipótese da perda do traço D em T na gramática oitocentista do PB, tendo em vista que para Holmberg (2010a) essa é uma das propriedades definitórias das LSN parcial.

A amostra do manuscrito setecentista apresenta um uso produtivo das construções com *se* (16,88%; 211/1250), principalmente da construção com *se*-impessoal (137 de 211 construções com *se*), enquanto o manuscrito do século XIX traz pouquíssimos usos de

construções com *se* impessoal (apenas 0,4% dos dados; 5/1250) e carece completamente de construções com *se* com interpretação passiva.<sup>102</sup>

Contrariando vários estudos sobre as construções com *se*-passivo na diacronia do PB (cf. GALVES, 1985; NUNES 1990; entre outros), descobrimos que os dados do século XIX de Goiás careciam de construções com *se*-passivo (mostrando uma mudança em relação aos dados do século XVIII) sem que isso estivesse ligado a um empobrecimento da morfologia de número da 3ª pessoa do plural dos verbos transitivos. Baseando-nos em algumas das intuições de Holmberg (2010a), propomos que a perda do traço-D em T na gramática do PB também está envolvida na perda de construções com *se*-passivo, na diminuição de construções *se*-impessoal e na ascensão de um tipo inovador de construção em comparação com PE, qual seja, as construções com sujeito nulo impessoal na 3ª pessoa do singular, como em (29), as quais aparentemente estão restritas às LSN parciais (cf. HOLMBERG, 2010a).<sup>103</sup>

(29) Pronome nulo indeterminado na 3ª pessoa do singular no século XIX

a. Dia 18 o senhor Joaquim Fernandez esteve aqui de noite, quando ele saiu entrou a Anna Joaquina daqui de frente que veio despedir para ir para Barra. Dia 17 faleceu o cadete Cândido Gonzaga, e \_\_ **enterrou** dia 18.

b. Dia 14 de novembro de 1883 3ª-feira às 5 horas da manhã faleceu o cônego Pio à tarde foi sepultado nessa mesma tarde faleceu Pedro cunhado do Antônio Manoel \_\_ **sepultou** no dia seguinte.

c. Dia 28 foi domingo do Espírito Santo a procissão do Divino passou aqui quando \_\_ **leu** a sorte foi Manoel Antônio o Imperador do ano seguinte.

De acordo com Holmberg (2010a), nas LSN parciais (línguas sem um traço-D em T), a relação sonda-alvo entre T e um  $\varphi$ P de 3ª pessoa não supre o valor de definitude, e o resultado dessa operação é um pronome nulo indeterminado na 3ª pessoa do singular atuando como o sujeito da oração. Seguindo Holmberg (2010a), propomos que essa segunda operação gera as construções com sujeitos nulos indeterminados no PB oitocentista, como em (29) – destacamos que não encontramos essas construções na amostra de dados do século XVIII. Com esses dados, temos uma evidência empírica da perda do traço-D em T na gramática

<sup>102</sup> Embora tenhamos utilizado uma amostra de 1.250 sentenças do manuscrito do século XIX neste estudo, destacamos que não encontramos registro de construções passivas com *se* em todo o manuscrito, que foi escrito ao longo de 50 anos: de 1881 a 1930.

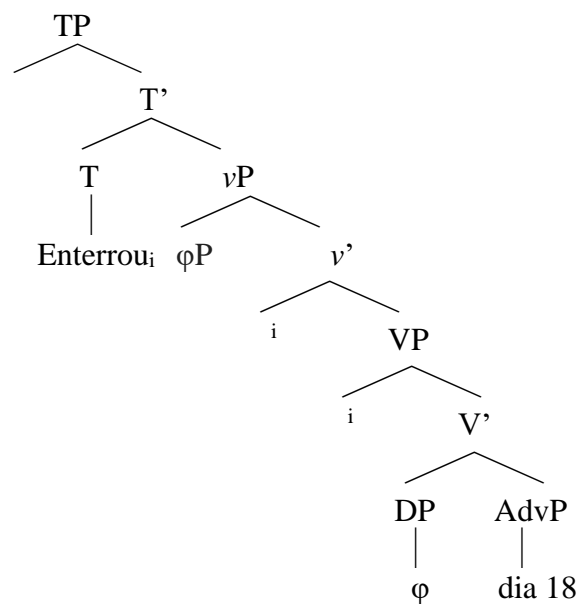
<sup>103</sup> Neste estudo, foram identificadas 05 construções com sujeitos nulos arbitrários na 3ª pessoa do singular no período de escrita do documento que vai de 1883 a 1884. H. Borges (2014), por sua vez, identificou 38 sentenças dessa natureza no mesmo manuscrito considerando o período de 1881 e 1899.

inovadora do PB oitocentista em Goiás, a qual consiste na possibilidade de licenciar construções impessoais na 3ª pessoa do singular sem um sujeito manifesto, isto é, sem o clítico *se*-impessoal – em (30) exemplificamos uma construção oitocentista com *se*-impessoal.

(30) Dia 23 às 3 horas da manhã Manoel daqui morreu, às 2 horas enterrou-se.

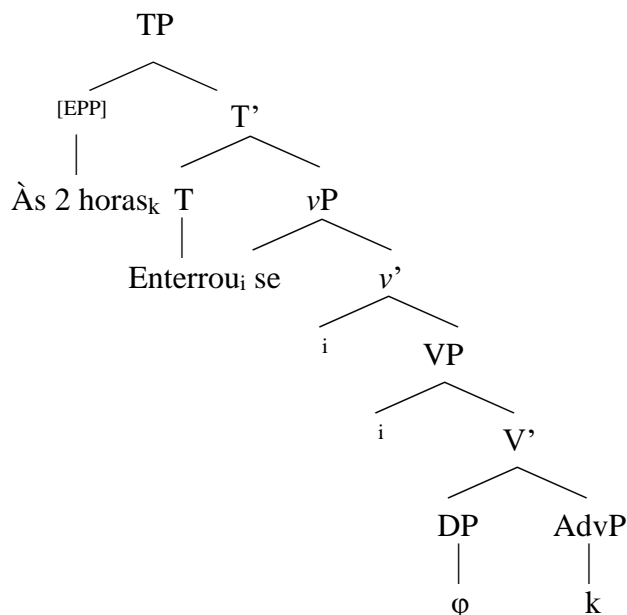
Assumimos aqui que o pronome de 3ª pessoa no singular de leitura arbitrária no PB gerado em construções como em (29) é um  $\phi$ P, consoante Holmberg (2005, 2010a). Desse modo, assim como no caso do pronome nulo genérico no finlandês (cf. HOLMBERG, 2010b), sugerimos que o pronome nulo impessoal de leitura arbitrária gerado em construções como (29) no PB permanece onde é gerado, em [SPEC-vP]. Igualmente, sua contraparte manifesta, o *se*-impessoal em (30), permanece em sua posição de base, em [SPEC-vP]. Esse último caso se relaciona com o fato de que o *se*-impessoal no PB atual ocorre exclusivamente em posição enclítica na ausência de alguma partícula atratora (cf. MARTINS & NUNES, 2016). Em (31) e (32), propomos, respectivamente, uma derivação para o sujeito nulo indeterminado de 3ª pessoa no singular e outra para sua contraparte manifesta, o *se*-impessoal, considerando os dados da nossa amostra oitocentista.

(31) Pronome nulo indeterminado na 3ª p. do singular ((cf. (29a))<sup>104</sup>)



<sup>104</sup> Como mencionado no Capítulo 4, a diferença entre a leitura genérica inclusiva e a leitura genérica exclusiva (arbitrária) das construções impessoais com pronome nulo na 3ª pessoa do singular no PB atual parece estar ligada aos traços de tempo do verbo, que atuam como operadores que licenciam uma ou outra interpretação. No caso dos dados do manuscrito oitocentista, o único operador atuando nesse sentido é o pretérito perfeito, licenciando a leitura arbitrária.



(32) Construções com *se*-impessoal no século XIX (cf. (30))

Mostramos em (33) alguns dados de construções com *se*-passivo e *se*-impessoal na amostra de dados setecentista.

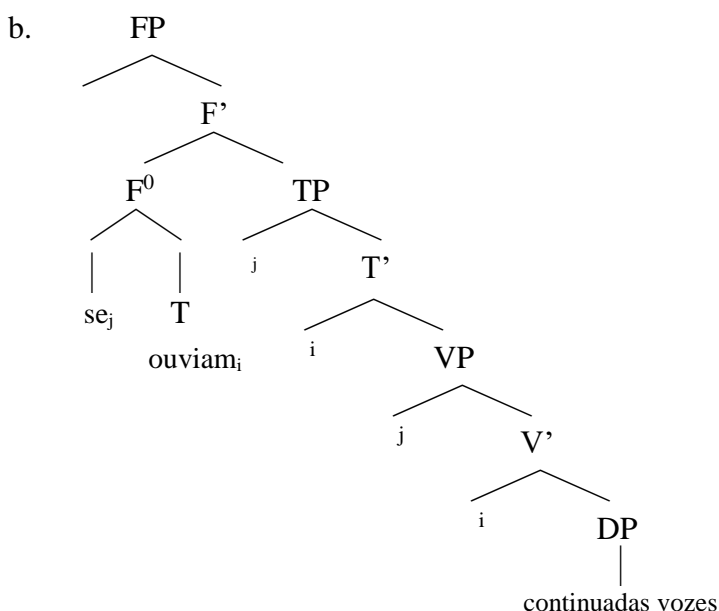
- (33) Construções com *se*-impessoal (33a) e *se*-passivo (33b) na amostra do século XVIII
- a. Pelas cinco e meia da manhã se partiu do rancho antecedente, e às onze se chegou ao de Manoel Correia, passando-se a grande serra chamada dos Órgãos ou da Boavista, sumamente levantada sobre o nível do Mar.
- b. nesta serra se encontrarão dois maus Ranchos, e mais adiante um chamado Itamaraty que não é mau, e umas três ribeiras com suas pontes de madeira: (...) andaram-se cinco léguas até o Rancho sobredito de Manoel Correia. Aqui se contarão as pessoas de que toda a comitiva se compunha.

No caso das construções com *se*-passivo registradas somente na amostra setecentista, seguindo a análise de Raposo e Uriagereka (1996) para essas construções no PE, propomos que o *se*-passivo possui um traço de definitude capaz de satisfazer o traço-D em [SPEC-TP] – é por isso, aliás, que a derivação não gera um pronome nulo indeterminado em [SPEC-TP], o que não é permitido numa LSN consistente. Adicionalmente, V move-se para T, valorando seus traços-φ em T. Observa-se que o DP argumento interno das construções com *se*-passivo deve concordar em seus traços-φ com T, como suposto pela correspondência manifesta de concordância entre eles ilustrada em (34a) a seguir. Segundo Raposo & Uriagereka (1996), essa concordância se estabelece no PE por conta de um núcleo funcional

periférico especificado para os traços- $\phi$  e o Caso nominativo – viz., FP. Em outras palavras, a concordância entre o verbo e o DP argumento interno em (34) é estabelecida por meio da atração dos traços formais de FP e do DP argumento interno, que, como resultado, por exemplo, permite que, em alguns casos, o DP argumento interno ocupe uma posição à esquerda do verbo (cf. Capítulo 4, para exemplo de quando essa concordância se manifesta fonologicamente, isto é, com o DP em posição pré-verbal).

(34) Construções com *se*-passivo no PB setecentista

a. Pelas árvores da vizinhança do povo, se ouviam continuadas vozes dos meninos  
*Viva sua Excelência.*



Aparentemente, a perda de construções com *se* na gramática oitocentista poderia ter duas explicações: na primeira delas, o núcleo FP não estaria mais ativo nessa gramática e, por isso, teria sido responsável pela perda do estabelecimento da concordância do verbo com o DP argumento interno, impossibilitando, portanto, que essa gramática gerasse passivas com *se*. Na segunda explicação para a perda das passivas com *se* na gramática oitocentista, sugerimos que a perda do traço-D de T em verbos transitivos bloquearia a valoração de traços semânticos de número no plural por T nos casos em que não houvesse em [SPEC-TP] um elemento com traços formais correspondentes – nota-se que é o clítico *se* o elemento que ocupa [SPEC-TP] nessas construções. Uma vez que T não mais carrega um traço-D, a única forma de um verbo transitivo desencadear a concordância de número plural passa a ser, então, quando um XP que contém esse traço formal se move para [SPEC-TP]. O *se*-passivo da gramática setecentista, todavia, assim como o *se*-impessoal, já não era capaz de satisfazer um

traço de definitude em [SPEC-TP] na gramática oitocentista. Destaca-se que, na proposta de Raposo & Uriagereka (1996), as operações de movimento explícito nas construções com *se*-passivo são, respectivamente: (i) a subida do *se* para [SPEC-TP], (ii) a subida do complexo [V-v] para T e (iii) o movimento de T contendo o complexo [V-v] para F. Com a impossibilidade de o clítico *se* subir para [SPEC-TP] e satisfazer o requerimento sentencial por um traço-D, o núcleo funcional F perde sua função de estabelecer a correspondência entre os traços- $\phi$  do DP argumento interno e os traços não interpretáveis do verbo, o que, por fim, extingue as construções com *se*-passivo da gramática de sujeito nulo parcial do PB oitocentista em Goiás. Cabe destacar que, tanto na primeira quanto na segunda explicação, atribuir a perda das construções com *se*-passivo à impossibilidade de o DP argumento interno estabelecer a concordância com o verbo por meio do núcleo funcional F vincula-se à proposta de Pires (2006), mencionada no Capítulo 4, a respeito de a perda do movimento do verbo de T para F no PB ter desencadeado uma série de alterações na gramática da língua, como a perda da ênclise, por exemplo.

Tomamos a segunda explicação como a mais provável para a perda das construções com *se*-passivo na gramática do PB oitocentista em Goiás, porque ela nos explica, ao mesmo tempo, a emergência das construções com sujeitos nulos indeterminados na 3ª pessoa do singular, uma das propriedades definidoras das LSN parcial. Em (35), exemplificamos uma construção – presente no manuscrito do século XIX, mas não na amostra desta tese – que poderia ser uma construção com *se*-passivo, mas é uma construção com sujeito nulo indefinido na 3ª pessoa do singular.

- (35) Dia 4 Faleceu a Dona Lin[...], mulher do capitão João Berquó para amanhecer. Faleceu a senhora Ninica de repente, nesse dia \_\_ enterrou ambas 1 de manhã a outra de tarde. (H. BORGES, 2014, p. 65)

## 6.5 Construções impessoais

As construções impessoais nas amostras do nosso *corpus* incluem tanto construções com verbos impessoais (*haver/ter* etc.) quanto com verbos meteorológicos (*chover/amanhecer* etc.), que no português possuem um sujeito expletivo nulo e não argumental. Representando 13,04% dos dados oitocentistas (contra apenas 2,88% dos dados setecentistas), todas as construções impessoais no século XIX estão acompanhadas por um elemento dêitico em posição pré-verbal, como exemplificado em (36) a seguir.

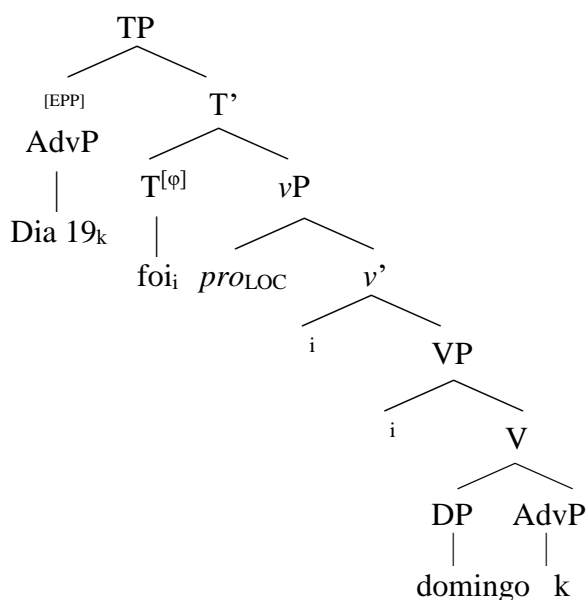
- (36) a. Dia 2 de tarde choveu a 1ª chuva e muito grande.  
 b. Nessa noite choveu muito.  
 c. Dia 19 foi domingo.  
 d. Dia 15 houve espetáculo no Teatro: A Sogra e o entremez Lucrecia Borges.

Posicionamentos teóricos recentes têm admitido que alguns constituintes no PB contemporâneo, como *aqui*, *ali*, e *lá*, podem atuar como se fossem expletivos quase argumentais em construções similares àquela em (36b) (cf. NAVES & PILATI, 2018).<sup>105</sup>

- (37) *Aqui/ali/lá chove muito.* (NAVES & PILATI, 2018, p. 206)

Contudo, a proposta de Pilati, Naves & Salles (2017) de que um elemento dêitico satisfaz o traço-EPP no PB é mais condizente com as construções impessoais nos dados desse tipo na amostra oitocentista. Essa foi também a intuição fornecida por H. Borges (2014) para construções dessa natureza no manuscrito do século XIX. A derivação das construções impessoais na gramática oitocentista pode, então, ser representada como em (38) a seguir.

- (38) Derivação de uma construção impessoal no século XIX (cf. (36c))



<sup>105</sup> O finlandês, uma LSN parcial, possui um pronome expletivo *sitü* que, concorrendo com elementos adverbiais, pode ocupar a posição de sujeito de construções com sujeitos nulos indefinidos.

Essa constatação nos parece interessante quando estabelecemos um paralelo desses resultados com a gramática histórica do francês, que devido ao contato linguístico com línguas germânicas se desenvolveu de uma língua de sujeito nulo para uma língua de sujeito não nulo (cf. BIBERAUER & ROBERTS, 2017). Para atender ao requerimento de realizar foneticamente o sujeito em sentenças impessoais, o francês especializou o uso do pronome *il*. No PB (não apenas na variedade goiana), cuja gramática se desenvolveu de uma LSN consistente para uma LSN parcial, não houve a especialização ou o desenvolvimento de um pronome expletivo para as construções impessoais, mas constituintes não argumentais passaram a ser movidos para [SPEC-TP] para satisfazer o EPP nessas construções.

## 6.6 Síntese do capítulo

Há algumas considerações a serem feitas sobre a perda de sujeitos nulos no PB de Goiás constatada neste capítulo por meio de dados empíricos e das análises estatística e linguística. Primeiro, a perda de sujeitos nulos nesse dialeto parece estar crucialmente ligada à perda do traço-D em T em sua gramática do século XIX. Portanto, o aumento acentuado de sujeitos manifestos nos dados do século XIX foi explicado como uma exigência de um elemento sintático com um índice de definitude se mover para [SPEC-TP] para satisfazer o EPP naquela gramática. Recapitulamos a proposta de Holmberg (2005, 2010a) de que, se uma língua não tem um traço-D em T (como nas LSN parcial e nas línguas de sujeito não nulo), o traço-EPP é checado em [SPEC-TP] por um DP lexical ou por outra projeção máxima que possa satisfazer esse requisito. Assim, em nossa proposta, a gramática oitocentista do PB – uma LSN parcial – exige a realização fonológica e o movimento de um elemento de natureza nominal para [SPEC-TP] para satisfazer o EPP, especialmente na 3ª pessoa.

Seguindo Alexiadou & Agnostopolou (199), nossa explicação para a inversão livre do sujeito (ordem verbo-sujeito) na gramática do século XVIII é a de que o traço EPP pode ser satisfeito pelo movimento de V-para-T nessa gramática. Assim, não há exigência de que os sujeitos DPs manifestos em estruturas VS ocupem a posição de [SPEC-TP], ou que um pronome expletivo o faça. Paralelamente, argumentamos que a perda acentuada da inversão livre do sujeito do século XVIII para o século XIX também é uma consequência da perda do traço-D em T na gramática oitocentista do PB em Goiás.

Dada a falta de um traço-D em T na gramática oitocentista, inferimos que os sujeitos nulos de 3ª pessoa passaram a ser licenciados por um antecedente correferente (para obter uma leitura referencial), comportamento esperado para as LSN parciais na proposta de

Holmberg (2010a), apesar de nossos dados serem pouco significativos a esse respeito. Curiosamente, constatamos que a gramática de sujeito nulo parcial no PB em Goiás no século XIX provavelmente surgiu independentemente do empobrecimento da concordância verbal. Inferimos que a perda da morfologia verbal constatada na gramática do PB atual seria, na verdade, uma consequência da perda do traço-D de T. Adicionalmente, mostramos que os dados do século XIX já mostram uso de construções impessoais com sujeito nulos arbitrários na 3ª pessoa do singular devido à perda do traço-D de T – de acordo com Holmberg & Roberts (2013), essas construções são encontradas nas LSN parciais, mas não nas LSN consistentes. Destacamos também que a gramática oitocentista do PB em Goiás apresentou perda do movimento do verbo de T para F. Nesse caso, seguindo a proposta de Pires (2006), verificamos que a inativação do núcleo funcional F impossibilitou completamente a produção de construções com *se*-passivo na gramática oitocentista do PB em Goiás – a ausência de passivas com *se* no PB é atestada por Martins & Nunes (2016) para o PB contemporâneo. Ao todo, esses resultados fornecem evidências empíricas suficientes para postularmos o surgimento de uma variedade do PB com propriedades de uma LSN parcial já na segunda metade do século XIX.

Por fim, devemos destacar que, apesar da perda da inversão livre do sujeito, não pudemos fornecer evidências significativas da perda do movimento de V-para-T na gramática oitocentista. Isso sugere que precisamos refinar os critérios de codificação das amostras de nosso *corpus* ou até mesmo aumentar nosso conjunto de amostras de dados diacrônicos para pesquisas futuras sobre o movimento do verbo na gramática oitocentista do PB em Goiás.

## **7 UMA HIPÓTESE DE LINGUÍSTICA DE CONTATO PARA A EMERGÊNCIA DA GRAMÁTICA DO PORTUGUÊS BRASILEIRO EM GOIÁS**

Seguindo uma profícua tradição nos estudos sobre a constituição do PB, este capítulo tem por objetivo discutir as implicações e as conjeturas de um possível papel do contato linguístico na emergência da gramática do PB em Goiás no século XIX. Mais precisamente, dissertamos sobre como a mudança caracterizadora da emergência da gramática do PB no século XIX em Goiás pode ser analisada como resultado do componente sócio-histórico do contato linguístico no período colonial brasileiro. O capítulo está organizado da seguinte maneira: na seção 7.1, discutimos algumas questões-chaves dos estudos de linguística de contato, mais especificamente o conceito de mudança induzida por contato (do inglês, *contact-induced language change*), bem como discorremos sobre alguns dos principais estudos atuais que tratam do papel do contato linguístico na formação do PB; na seção 7.2, descrevemos o contexto sociolinguístico em que emergiu a gramática do PB em Goiás no século XIX, com o objetivo de vincular a evidência da emergência da gramática oitocentista do PB em Goiás, uma das regiões mais isoladas do Brasil no período colonial, ao contexto sócio-histórico do contato linguístico no país, talvez o único fato de natureza sociolinguística a marcar uniformemente todo o território brasileiro no período colonial.

### **7.1 Algumas questões sobre linguística de contato e a formação do português brasileiro**

Thomason (2001) define contato linguístico como o uso de mais de uma língua em um mesmo lugar e em um mesmo tempo. Nessa direção, para a autora, situações não triviais de contato linguístico se caracterizam pelo uso de mais de uma língua por pelo menos um determinado grupo de pessoas em um mesmo lugar e em um mesmo tempo. Winford (2003) ressalta que a mistura de traços linguísticos resultante do contato linguístico é um processo criativo e regulado por regras que afeta, de um modo ou de outro, todas as línguas em contato. Os resultados do contato linguístico são, portanto, variados (cf. WINFORD, 2003).

Um dos resultados do contato linguístico é a mudança induzida por contato. Thomason (2001) define mudança induzida por contato como qualquer mudança linguística que teria sido menos provável de acontecer em determinada língua sem uma situação particular de contato e interação com outra(s) língua(s). Thomason (2001) argumenta que o uso da expressão “menos provável” (do inglês, *less likely*) em sua definição de mudança induzida por contato linguístico refere-se ao fato de que muitas das mudanças induzidas por

contato também possuem motivações estruturais internas e, portanto, podem ser resultantes de múltiplos fatores. Thomason (2010) enfatiza seu posicionamento de que tanto motivações internas quanto motivações externas devem ser consideradas em qualquer relato sistemático da história de uma língua: “a melhor explicação para qualquer mudança linguística levará em conta todos os fatores causais detectáveis, tanto internos quanto externos” (THOMASON, 2010, p. 34).<sup>106</sup> A autora explica, ainda, que sua definição de mudança induzida por contato não deve ser tomada como um critério/um procedimento para estabelecer a mudança, afinal, não há meios de se medir precisamente o que é “menos provável” para qualquer mudança linguística que ocorreu no passado (THOMASON, 2010). A definição de Thomason (2001) para mudança induzida por contato inclui propriedades de interferência transferidas para uma determinada língua A partir de uma dada língua B, como, por exemplo, transferência de morfemas, de morfemas e estrutura ao mesmo tempo, ou de estrutura apenas. De acordo com a autora, o processo de transferência ocorre com ou sem modificação estrutural das propriedades da língua-fonte. No Quadro 1, reproduzimos a proposta de Thomason (2001) para a mudança induzida por contato, indicando seus processos e resultados.

Quadro 1: Mudança induzida por contato: processos e resultados linguísticos.

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>Fatores sociais:</b><br/> Intensidade do contato<br/> Presença <i>versus</i> ausência de aquisição imperfeita<br/> Atitudes dos falantes</p> <p><b>Fatores linguísticos:</b><br/> Marcadores universais<br/> Grau com que traços são integrados ao sistema linguístico<br/> Distância tipológica entre a língua fonte e a língua alvo</p> <p><b>Uma tipologia dos efeitos sobre a estrutura da língua-alvo:</b><br/> Perda de traços<br/> Adição de traços<br/> Substituição de traços</p> <p><b>Uma tipologia de mecanismos de mudança induzida por contato:</b><br/> Troca de código<br/> Alternância de código<br/> Familiaridade passiva<br/> Negociação<br/> Estratégias de aquisição de segunda língua<br/> Aquisição bilíngue de primeira língua)<br/> Decisão deliberada (<i>deliberate decision</i>)</p> |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Fonte: Thomason (2001, p. 60).

<sup>106</sup> No original: “The best explanation for any linguistic change will take all discoverable causal factors into account, both internal and external” (THOMASON, 2010, p. 34).



A integração de estruturas linguísticas de duas línguas ou mais é o resultado de vários mecanismos que podem operar individualmente ou em combinação. De acordo com o Quadro 1, Thomason (2001) aponta sete mecanismos que atuam na integração de estruturas linguísticas em situações de contato linguístico, a saber: mudança de código (*code-switching*), alternância de código (*code alternation*), familiaridade passiva (*passive familiarity*), negociação (*negotiation*), estratégias de aquisição de segunda língua (*second-language acquisition strategies*), aquisição bilíngue de primeira língua (*bilingual first-language acquisition*) e decisão deliberada (*deliberate decision*).

Thomason & Kaufman (1988) e Thomason (2001) distinguem dois processos fundamentais de mudança envolvendo contato linguístico: manutenção (do inglês, *maintenance*) e substituição (do inglês, *shift*), que são desencadeados por fatores sociais, como a duração do contato, o número de falantes bilíngues, o tamanho da população e o domínio socioeconômico de determinados falantes. Ressalta-se que, independentemente do processo responsável pela mudança induzida por contato, Thomason (2001, p. 21) defende que “fatores sociais são os únicos que precisam ser considerados na avaliação da estabilidade: fatores linguísticos (como a similaridade estrutural geral das línguas em contato) parecem ser totalmente irrelevantes.”<sup>107</sup>

Nos casos de manutenção (*maintenance*), falantes nativos de uma língua A mantêm a sua língua, mas integram elementos de outra língua, a língua-alvo B. Sob baixa intensidade de contato e se o número de falantes bilíngues em uma determinada população é pequeno, o empréstimo é limitado ao vocabulário não básico. Com o aumento do contato e com um número crescente de falantes bilíngues, há mais empréstimo lexical e até mesmo empréstimos estruturais moderados ou intensos (cf. THOMASON & KAUFMAN, 1988; THOMASON, 2001).<sup>108</sup> Em outras palavras, a manutenção (*maintenance*) provoca primeiramente mudança lexical – ao passo que a substituição (que, doravante, chamaremos de *language shift*) desencadeia primeiramente mudanças fonológicas e sintáticas.

---

<sup>107</sup> No original: “social factors are the only ones that need to be considered in assessing stability: linguistic factors (such as overall structural similarity of the languages in contact) seem to be totally irrelevant” (THOMASON, 2001, P. 21).

<sup>108</sup> Quando se fala do termo empréstimo, geralmente ele é utilizado com um sentido bastante restrito, qual seja: a adoção de um item vocabular de uma língua B por uma língua A. Cabe destacar, contudo, que empréstimo na linguística de contato abarca também a adoção de sons ou formas, como, por exemplo, morfemas, palavras ou unidades maiores, por uma língua A em contato com uma dada língua B. Nessa perspectiva, o empréstimo diz respeito, essencialmente, à substância fonética, isto é, a todas as transferências de sons e regras fonológicas, de palavras com sua forma e significado e de estruturas morfológicas e morfossintáticas envolvendo tanto forma quanto significado (cf. KUTEVA, 2017). Por outro lado, há também os empréstimos que dizem respeito ao padrão estrutural, sem relação com o material fonológico, como a ordem das palavras, que envolve as exigências de ordenação dos elementos significativos de uma oração (cf. KUTEVA, 2017).

Nos casos de *language shift*, os falantes da língua A abandonam o uso de sua língua em favor da língua-alvo B. Quando falantes de uma língua A estão mudando para uma língua B, eles transferem recursos de sua língua nativa A para B. Os fatores sociais que determinam o resultado estrutural do contato linguístico são as relações socioeconômicas e políticas de dominação, o tamanho da população e o grau em que a língua-alvo B é acessível para os falantes da língua A (cf. THOMASON, 2001). O efeito do primeiro fator nos casos de *language shift* é direto e a cultura dominante de falantes da língua B promove o desaparecimento da língua A em favor de B (cf. THOMASON, 2001). Thomason (2001) afirma que, se o número de falantes da língua A é significativamente menor do que o número de falantes da língua B, a língua-alvo B é adotada sem deixar vestígios da língua A. Consoante Thomason (2001), o alto percentual de bilíngues em uma população em casos de *language shift* é relevante porque determina o papel daquilo que ela chama de aprendizagem imperfeita.<sup>109</sup> Na visão de Thomason (2001), os aprendizes adultos de segunda língua (L2) em situação de contato linguístico utilizam uma série de estratégias específicas que podem ter algum impacto na estrutura da língua-alvo sob certas condições sociais, como as que subjazem *language shift*.<sup>110</sup>

Dentre as estratégias delineadas pela autora, duas delas parecem especialmente relevantes: a primeira consiste em manter distinções e outros padrões da língua nativa dos aprendizes (a L1) na construção de sua versão da gramática da língua-alvo, projetando a estrutura de sua L1 (língua-fonte) na forma da língua-alvo. Thomason (2001) afirma que essa estratégia consiste basicamente em os aprendizes de L2 inserirem material da língua-fonte (língua materna) na língua-alvo. Outra estratégia usada pelos aprendizes de L2 relevante para Thomason (2001) é a de ignorar distinções, especialmente distinções marcadas, que são presentes na língua-alvo, mas opacas para os aprendizes em um estágio inicial de aquisição de L2. Essas “falhas” em aprender certos traços da língua-alvo, estão, de acordo com Thomason (2001), entre os casos mais comuns de interferência induzida por substituição (*shift-induced interference*). Ainda de acordo com a visão de Thomason (2010), os aprendizes de L2 em situação de contato linguístico certamente devem decidir, às vezes, consciente ou inconscientemente, empregar recursos que não são usados por falantes nativos da língua-alvo. Nota-se que, em sua proposta teórica, Thomason (2001) argumenta em favor da ideia de que o

---

<sup>109</sup> Thomason (2001) explica que é importante ter em mente que ela não usa o termo aprendizagem imperfeita para propor uma possível inabilidade dos indivíduos em aprender a língua-alvo em situações de contato linguístico ou mesmo a falta de acesso suficiente à língua-alvo para permitir a aquisição completa.

<sup>110</sup> Thomason (2001) acredita que, nos casos de manutenção com contato linguístico intenso e nos casos de *language shift* com um grande grupo de falantes de A e aprendizagem imperfeita, as consequências estruturais resultantes do contato linguístico parecem ser basicamente as mesmas.

falante da língua-fonte tem alguma consciência/percepção sobre a mudança que está provocando na língua-alvo. Essa ideia de que o falante desempenha um papel consciente na mudança induzida por contato não condiz com os pressupostos gerativistas sobre aquisição de língua e mudança sintática adotados nesta tese, especialmente com o conceito de descontinuidade apresentado no Capítulo 2.<sup>111</sup>

Tendo em vista que o conceito de descontinuidade é capaz de abarcar a mudança induzida por contato, consideramos a hipótese de que a mudança ocorrida na gramática do PB em Goiás nos idos de oitocentos é resultante de contato linguístico, mais especificamente um caso de interferência induzida por substituição (*shift-induced interference*). Propomos essa hipótese considerando o conjunto de dados da língua-E que servia de *input* para a aquisição e a estabilização do português como língua materna na América portuguesa. Esse conjunto de dados era bastante diversificado e provinha: (i) de falantes do português como primeira língua e (ii) de inúmeros falantes de línguas indígenas e africanas como língua materna. Nesse contexto, os dados do português produzidos como língua-E pelos falantes de línguas indígenas e africanas como língua materna eram: (a) influenciados por suas primeiras línguas e (b) distintos dos dados de língua-E produzidos por falantes de português como língua materna. Desse conjunto de múltiplas línguas-E é que propomos ter emergido a gramática caracterizadora das variedades da língua portuguesa na América. Voltaremos a essa questão mais detidamente na próxima seção; antes, porém, apresentaremos os procedimentos metodológicos propostos por Thomason (2001) para testar se, de fato, uma mudança numa dada língua foi induzida por contato. Esses procedimentos serão essenciais para sustentar nossa hipótese de que a gramática do PB oitocentista em Goiás seria resultado de contato linguístico.

Thomason (2001) afirma que procedimentos explícitos para estabelecer as propriedades linguísticas resultantes de mudança induzida por contato dificilmente são necessários quando um morfema de uma língua-alvo (língua de empréstimo) concorda em forma e função com um morfema da língua-fonte, ou quando um morfema de uma língua crioula concorda em forma e função com um morfema da língua lexificadora – embora as

---

<sup>111</sup> Lima-Salles (2006, p. 129) chama a atenção para o fato de que “o gerativismo diverge crucialmente de concepções que atribuem a variação e a mudança linguística a (...) mecanismos cognitivos desencadeados por pressões discursivas e por fatores de natureza perceptual e psicológica”. Lima-Salles (2006, p. 130-131) ressalta, contudo, que: “a abordagem da gramática gerativa não exclui o papel dos fatores externos, de natureza sócio-histórica, em cujas vicissitudes se assentam as condições singulares a que são expostas as sucessivas gerações de novos falantes no processo de aquisição de língua materna, tornando-se inevitável a mudança e, por consequência, a diversidade linguística. Nesse cenário, há ainda espaço para investigar o papel do contato de línguas no processo de mudança linguística, em termos de princípios (internos) da faculdade da linguagem (...).”

fontes de certos morfemas gramaticais sejam frequentemente debatidas por crioulistas e alguns desses morfemas possam ter múltiplas fontes. Para a autora, casos dessa natureza constituem exemplos não controversos de mudança induzida por contato. Há, contudo, segundo Thomason (2001), um segundo tipo de mudança induzida por contato linguístico que é indubitavelmente controverso. A autora aponta, por exemplo, para o fato de que, em casos de interferência induzida por substituição (*shift-induced interference*), a transferência de morfemas é menos comum do que a transferência de padrões estruturais. Nesses casos, a interferência estrutural ocorre especialmente por conta do contato cada vez mais intenso envolvendo bilinguismo generalizado (cf. THOMASON, 2001). Para casos assim, a autora afirma ser necessário estabelecer critérios que possam testar se, de fato, a mudança foi induzida por contato. Thomason (2001) propõe, então, cinco procedimentos a serem adotados para testar se os casos controversos realmente constituem um caso de mudança induzida por contato. Esses critérios são descritos a seguir.

Primeiramente, Thomason (2001) afirma que o pesquisador não pode atentar para um fato singular da língua, como, por exemplo, uma possível influência no sistema fonológico, mas para a língua como um todo (fonologia, morfologia e sintaxe). Thomason (2001) afirma que, se um determinado traço estrutural foi transferido de uma língua-fonte para uma língua-alvo, é altamente provável que outros traços estruturais também tenham sido transferidos da mesma língua-fonte. Para a autora, um argumento baseado em apenas um traço singular é pouco convincente. Em segundo lugar, uma língua-fonte precisa ser identificada para se estabelecer que essa língua-fonte tem precisamente a propriedade estrutural ou uma propriedade que razoavelmente pode ser argumentada como derivada de outra propriedade dessa língua. Em terceiro, o pesquisador precisa encontrar traços estruturais compartilhados entre a língua-fonte e a língua-alvo, os quais não precisam ser, e dificilmente serão, idênticos em todos os aspectos. Em quarto, o pesquisador precisa provar que os traços compartilhados não existiam na língua-alvo antes do intenso contato com a possível língua-fonte. Por último, o pesquisador precisa provar que o traço compartilhado realmente existia na língua-fonte antes de ela entrar em contato com a língua-alvo, quer seja por meio da comparação entre as duas línguas (língua-alvo e língua-fonte) diretamente, quer seja por meio da comparação da língua-alvo com línguas próximas da língua-fonte nos casos em que a língua-fonte já não mais existe porque todos os seus falantes a substituíram pela língua-alvo. Se um ou mais dos procedimentos acima não puder ser atendido, Thomason (2001, 2010) salienta que qualquer alegação de motivação externa para a mudança linguística em uma dada língua deve ser, na melhor das hipóteses, provisória. Para a autora, a adoção desses critérios no estudo de um

potencial caso de mudança induzida por contato pode colocar as hipóteses de origem de uma língua em um terreno mais firme do que às vezes elas têm estado.

Winford (2003) reconhece *language shift* como uma situação bastante diversificada de contato linguístico. O autor declara que a caracterização sociolinguística do processo de *language shift* tem impacto nos resultados dessa situação de contato linguístico. Por essa razão, ele delinea uma proposta interessante para as situações de contato linguístico que envolvem *language shift*, quais sejam: (i) substituição rápida e completa por grupo minoritário, com pouca ou nenhuma interferência de substrato da língua-fonte sobre a língua-alvo, como ocorre, por exemplo, com os grupos de imigrantes urbanos que trocam suas línguas maternas pelo inglês nos Estados Unidos; (ii) substituição rápida por grupos maiores ou por minorias de prestígio, com interferência de substrato leve a moderada, como no caso da substituição do francês da Normandia pelo inglês na Inglaterra; (iii) substituição por comunidades indígenas, com interferência moderada a intensa, como ocorreu com o uso de dialetos do inglês no século XVII em Barbados. Esse último tipo de situação reflete o caso da predominância no uso da língua portuguesa em detrimento das línguas indígenas autóctones e das línguas africanas trasladadas na América portuguesa. A partir da caracterização sociolinguística dos três tipos de *language shift* propostos por Winford (2003), podemos observar que haverá pouca influência da língua-fonte sobre a língua-alvo se o grupo que desencadeia a mudança abandonando a língua-fonte for menor do que o grupo da língua-alvo. A questão demográfica é, portanto, crucial para analisar os casos de contato linguístico resultantes de *language shift*. Além disso, Winford (2003) ressalta que o *input* linguístico proveniente das línguas e das situações de contato linguístico é extremamente relevante para estabelecer a natureza e a extensão das mudanças possíveis na língua-alvo.

Winford (2003) também menciona as situações de contato linguístico em que há como resultado a criação de novas línguas, viz., os pidgins, as línguas crioulas e as línguas mistas – Thomason (2001) classifica o contexto de emergência dessas línguas como uma situação extrema de contato linguístico. De acordo com Mufwene (2001, 2008), os pidgins e os crioulos comumente se caracterizam por serem novas variedades de línguas que se desenvolveram historicamente a partir dos contatos entre variedades coloniais não padrão de uma língua européia e várias línguas não européias ao redor do Atlântico e também nos oceanos Índico e Pacífico durante os séculos XVII e XIX. Os pidgins normalmente emergiam em colônias comerciais, que se desenvolviam em torno de fortes comerciais, como na costa da África Ocidental. Mufwene (2001, 2008) ressalta que os pidgins têm estruturas reduzidas e funções comunicativas restritas, tipicamente comércio. Inicialmente, eles também serviram

como língua franca não nativa para usuários que mantiveram seus vernáculos nativos em suas interações intra-étnicas (MUFWENE, 2001, 2008). Os crioulos, por sua vez, emergiram tipicamente em colônias de assentamento cuja indústria primária consistia no cultivo de cana de açúcar ou arroz em grandes *plantations*, para as quais foram empregados escravos não indígenas e não europeus, ou seja, escravos africanos, constituindo a esmagadora maioria das populações nessas plantações. Mufwene (2015) declara que os primeiros vernáculos crioulos no Novo Mundo não parecem ter surgido antes do século XVIII. O autor afirma que as primeiras referências às peculiaridades na fala dos escravos datam de uma mesma época, embora não se possa negar a existência de interlínguas, isto é, variedades em transição, ao longo dos períodos colonial e pós-colonial até o fim do tráfico de escravos.

Para Mufwene (2001), a história socioeconômica dos assentamentos no Novo Mundo sugere que a integração foi um fator crítico no desaparecimento geral das línguas africanas. Nesse contexto, o autor pontua que no contexto de escravidão em que viviam os africanos passaram a falar as línguas coloniais locais como seus vernáculos. Consequentemente, seus filhos adquiriram essas línguas coloniais locais como suas línguas nativas e vernaculares. Assim, cada nova geração de escravos visava ao vernáculo falado pelos membros locais das populações com quem interagem. Mufwene (2001) pontua, portanto, que a apropriação dos crioulos vernaculares como principal meio de comunicação levou ao desaparecimento das línguas africanas no Novo Mundo, cujos traços residem em alguma influência de substrato que possa ser identificada. A explicação fornecida por Mufwene (2001) para a perda das línguas africanas nas regiões no Novo Mundo onde emergiram línguas crioulas, como no Haiti (crioulo haitiano) e na Colômbia (palanquero), reside, portanto, no esforço por parte dos cativos africanos para sobreviver na nova ecologia em que viviam, usando vernáculos que lhes permitiam se comunicar da forma mais adequada possível.

Mufwene (2001, 2008) é um dos estudiosos que compreende a linguagem como uma população heterogênea de idioletos. Assim, cada falante de uma língua tem à sua disposição um conjunto de variantes que podem servir ao mesmo propósito comunicativo – palavras alternativas, pronúncias alternativas da mesma palavra, estruturas gramaticais alternativas etc. O idioleto de um indivíduo se desenvolverá de acordo com as características particulares da ecologia social na qual a aquisição de língua ocorre. Essa ecologia determinará qual variante será dominante em um idioleto (MUFWENE, 2008). Para Mufwene (2008), em qualquer interação linguística, os participantes acomodam-se constantemente aos idioletos uns dos outros por meio de inovações criativas. Essas inovações resultam de processos de competição e seleção das variantes disponíveis, o que pode provocar mudanças nos idioletos envolvidos

(MUFWENE, 2001). Em outros termos, o peso originalmente atribuído a cada variante pode mudar dependendo das pressões da situação comunicativa. Com base nesses argumentos, Mufwene (2001) adota uma posição uniformista em argumentar que os crioulos evoluíram pelos mesmos processos envolvidos na evolução das variedades de línguas não crioulas, sujeitos a peculiaridades apenas em relação às ecologias de sua emergência.<sup>112</sup>

Mufwene (2008) propõe que a aquisição de língua pode ser caracterizada como responsável pelo surgimento de padrões comunicativos particulares em uma comunidade linguística, assim como fenômenos emergentes na física ou na economia. O pesquisador afirma que as pesquisas que relacionam a aquisição de L2 ao estudo das línguas crioulas envolvem noções da aquisição de L2 focadas no indivíduo, e não nas comunidades de falantes, que geralmente produzem normas separadas dentro de uma mesma ecologia. Para o pesquisador, os estudos crioulos precisam lidar precisamente com esse aspecto particular da mudança. Mufwene (2008) argumenta que não é suficiente apenas provar que há transferência da primeira para a segunda língua, caracterizando uma influência de substrato; além disso, é preciso explicar como e por que determinados traços linguísticos de alguns falantes se espalham para uma população inteira, ou para partes dela, enquanto outras não. Adotando uma proposta de mudança linguística condizente com a de Mufwene (2001, 2008), Couto (2009) resume a perspectiva ecológica da mudança linguística assim:

(...) toda mudança linguística tem a ver com o contato. Para tanto é necessário reconhecer os diversos tipos de contato (...). São eles (i) o contato interlinguístico, ou contato de línguas propriamente dito, e o contato intralinguístico. Este último compreende (ii) o contato de dialetos e (iii) o contato de idioletos. (...) há também (iv) o contato entre gerações (contato intergeracional) e (v) o contato do aprendiz com a língua a ser adquirida. Até certo tipo de (vi) ausência de contato pode levar à mudança, como quando falantes de determinada língua se desgarram da comunidade-mãe de falantes. (...) Os pretensos casos de causalidade interna na mudança linguística, na verdade, têm a ver com contato de dialetos ou de idioletos, que leva a uma “replicação imperfeita”, como diz Salikoko Mufwene. No contato intergeracional, que abrange a aquisição de língua pela criança, a língua nunca é transmitida exatamente como existia na geração transmissora. (COUTO, 2009, p. 66).<sup>113</sup>

<sup>112</sup> Mufwene (2001) utiliza os termos ecologia em referência ao cenário sócio-histórico em que uma dada língua é falada. Para o autor, esse cenário é fortemente determinado por fatores socioeconômicos, especialmente quando se trata da ecologia de formação das línguas crioulas. Quanto ao termo evolução, o pesquisador utiliza-o principalmente para se referir à mudança na estrutura e às restrições pragmáticas associadas a uma dada língua.

<sup>113</sup> A partir dessa perspectiva, Couto (2009) realiza um extensivo estudo de diversas mudanças detectadas na gramática do português – do latim à sua evolução em solo americano – e declara que o PB é pura e simplesmente “uma nova modalidade do português, um dialeto, que evoluiu mais que o europeu devido ao contato com outras línguas no novo meio ambiente em que se inseriu” (COUTO, 2009, p. 78). Couto (2009) ressalta que esse é o entendimento de Mufwene (2001) para outras línguas europeias trasladadas, como o francês de Québec.

Como mencionado, Mufwene (2001, 2008) rejeita a ideia de que exista alguma diferença nos processos evolutivos que levam ao surgimento de línguas crioulas e línguas não crioulas. O autor pondera ainda que se de fato houver diferenças entre esses processos, tais diferenças equivalem a diferenças nos resultados determinados pela variação nas condições ecológicas de determinada língua. Winford (2017), contudo, desafia a visão de Mufwene (2001, 2008) e alega que a mudança com motivação interna tende a envolver mecanismos que exigem um longo processo antes que a inovação se estabeleça na gramática, ao passo que, em contraste, a mudança induzida por contato é um processo instantâneo, impulsionado pela transferência direta da função gramatical de um elemento numa dada língua de origem para um elemento similar na língua receptora (WINFORD, 2017).<sup>114</sup> Na visão de Winford (2017), portanto, um dos diagnósticos mais precisos para distinguir a mudança externa da mudança interna é o rápido surgimento de novas estruturas na língua-alvo – esse é o posicionamento de Thomason (2001) e da maior parte dos estudiosos de linguística de contato.

Winford (2017) reconhece que o quadro teórico de competição e seleção proposto por Mufwene (2001, 2008) é bem sucedido em descrever o contexto social e ecológico da mudança linguística e as características linguísticas que emergem na gênese crioula. Nessa direção, o autor pontua três grandes níveis de análise da ecologia social nos casos de contato linguístico, a saber: (i) as diferenças no sistema econômico, que influenciam a estrutura da população; (ii) as estruturas populacionais, que determinam quem interage com quem; e (iii) as formas demográficas das populações.

A partir de exemplos do inglês coloquial de Singapura, língua que discutimos no Capítulo 2 desta tese, Winford (2017) indica que há uma nítida evidência de que as novas estruturas no padrão sentencial dessa língua resultam da transferência de substrato do chinês e do malaio (línguas orientadas para o discurso), ao mesmo passo que mostra que esses tipos de estruturas não têm equivalentes em outras variedades do inglês que também foram trasladadas, como o inglês americano ou o inglês australiano (línguas de sujeito não nulo). Com isso, o autor pretende mostrar que diferentes tipos de reestruturação ocorrem com diferentes resultados em diferentes configurações ecológicas. A crítica de Winford (2017) à proposta de Mufwene (2001, 2008) consiste basicamente em defender uma teoria do contato linguístico que seja virtualmente composta por três componentes: o sociolinguístico, o

---

<sup>114</sup> Registramos um agradecimento especial à professora Marlyse Baptista por, em 2016, ter-nos apresentado o manuscrito de Winford (2017). Remetemos o leitor a Baptista (2017) para uma compreensão mais geral da crítica de Winford (2017) a Mufwene (2001, 2008).



linguístico e o psicolinguístico.<sup>115</sup> Para Winford (2017), Mufwene (2001, 2008) não aborda o processo psicolinguístico que gera as inovações observáveis em línguas crioulas. Desse modo, Winford (2017) defende uma visão tradicional de que todos os novos ingleses (*New Englishes*), e também as línguas crioulas em geral, surgem da aquisição de segunda língua envolvendo processos de reestruturação no sentido pretendido pelos pesquisadores de aquisição em primeira e segunda língua. Entre outras coisas, para Winford (2017), o processo de reestruturação inclui a replicação de padrões gramaticais da L1 em sistemas de interlíngua dos aprendizes da L2, o que envolve, segundo o autor, o mecanismo psicolinguístico de imposição, que consiste em aplicar os procedimentos de produção da L1 na produção de estruturas de uma L2 emergente, como no caso dos novos ingleses.

Desse modo, Winford (2017) defende a visão tradicional de que todos os novos ingleses, incluindo os crioulos, surgiram via aquisição de segunda língua envolvendo processos de reestruturação no sentido pretendido pelos pesquisadores de aquisição em primeira e segunda língua, não no sentido de Mufwene (2001, 2008). Entre outras coisas, o processo de reestruturação, para Winford (2017), inclui a replicação de padrões gramaticais da L1 em sistemas de interlíngua dos aprendizes da L2, o que envolve, segundo o autor, mecanismo psicolinguístico de imposição, que consiste em aplicar os procedimentos de produção da L1 na produção de estruturas de uma L2 emergente.

Feita essa sucinta discussão sobre algumas situações especiais de contato linguístico, sua ecologia e seus resultados, especialmente as situações que envolvem *language shift*, vamos discutir agora o contexto sócio-histórico de formação do PB, fortemente marcado por situações de contato linguístico envolvendo bilinguismo e multilinguismo durante e após o período de colonização da América portuguesa, permanecendo multilíngue em determinadas regiões do país (cf. MATTOS E SILVA, 2004). Para Mattos e Silva (2004), o multilinguismo generalizado na América portuguesa não impossibilitou que a língua portuguesa fosse instaurada como a única língua oficial em 1757 e que um esforço político fosse engendrado para que a língua portuguesa fosse difundida e instaurada em toda a América portuguesa, minorando e extinguindo as línguas nativas e as línguas africanas trasladadas para o Brasil. Como Mattos e Silva (2004) defende, mais do que os esforços políticos empreendidos pela coroa portuguesa para que a língua portuguesa se instaurasse por todo o território colonizado por Portugal na América, foram os africanos e seus descendentes os principais responsáveis pela difusão e conseqüentemente pela emergência da gramática inovadora do português em

---

<sup>115</sup> Na seção seguinte deste capítulo, buscamos adotar esses três níveis de análise para avaliar a ecologia social em que emergiu a gramática do PB em Goiás no século XIX.

solo americano. Em seus estudos, Mattos e Silva (2004) parece relegar um papel secundário aos ameríndios na formação do PB. Acreditamos, entretanto, que, para fornecer uma argumentação consistente sobre a mudança gramatical no PB, é preciso considerar tanto o aporte ameríndio quanto a matriz africana, sem, evidentemente, deixar de levar em conta as idiossincrasias sociolinguísticas, linguísticas e demográficas de cada uma dessas matrizes fundantes do PB. Lobato (2006) propõe que possíveis inovações gramaticais provenientes de matrizes indígenas e africanas na constituição do PB devem ser tratadas levando em consideração a hipótese de que os indivíduos daquelas línguas, ao aprenderem o português como L2, interpretaram as formas linguísticas não em suas propriedades intencionais, mas, sim, em suas propriedades extensionais. Ressalta-se que as propriedades extensionais a que a autora se refere dizem respeito ao desempenho linguístico dos falantes nas situações reais de fala, isto é, referem-se ao conjunto de enunciados da língua-E. Por seu turno, as propriedades intencionais são conceitos linguísticos construídos mentalmente, inerentes à competência linguística e se referem à capacidade linguística bioprogramada, à língua-I. Ao enfatizar a matriz indígena na formação do PB, Lobato (2006) reconhece a dificuldade em identificar a relação entre as estruturas da língua-fonte, da língua-alvo e da(s) interlíngua(s). Desse modo, a autora discute essa questão em termos de mecanismos cognitivos mais gerais e universais do processo de aquisição de L2. Isso sugere que mecanismos similares de aquisição de L2 podem ter sido utilizados por ameríndios e africanos no território colonial.

Estima-se que havia 1 milhão de povos tupis e 1.5 milhões de povos guaranis constituindo o grupo tupi-guarani quando os colonizadores portugueses chegaram à América portuguesa (SCHWARCZ & STARLING, 2015). Os chamados tupis-guaranis se subdividiam em diversos povos de línguas aparentadas. De acordo com A. Rodrigues (2005), as línguas ameríndias naquele período eram cerca de 1200, mas atualmente são cerca de 180. As línguas nativas que tiveram maior contato com o português foram o tupi e o tupinambá. A interação dessas línguas com o português resultou, inclusive, na criação de novas línguas: a língua geral paulista, oriunda do tupi de Piratininga, e a língua geral amazônica, ou *nheengatu*, oriunda do tupinambá do Pará (A. RODRIGUES, 2010). Em geral, o termo língua geral designa o conjunto de línguas que emergiram do contato entre colonos e povos ameríndios na América do Sul (A. RODRIGUES, 2010). Segundo A. Rodrigues (2010), a língua geral paulista teve grande abrangência no território e foi utilizada por parte expressiva daqueles que participaram da empreitada de desbravar o interior do país desde a segunda metade do século XVI até a segunda metade do século XIX. Diferentemente da extinta língua geral paulista, a língua geral amazônica, ou *nheengatu*, é falada na região amazônica até os dias de hoje.

Thomason (2001) menciona a prática de importar força de trabalho como responsável por complexas situações de contato linguístico em todo o mundo, tendo como o caso mais notório o comércio atlântico de escravos, responsável pela diáspora de cerca de 10 milhões de africanos escravizados para trabalharem nas plantações do chamado Novo Mundo. Nesse cenário, pelo menos 4 milhões de vidas africanas tiveram a América portuguesa como destino (cf. MATTOSO, 1982; FLORENTINO, 1999). De acordo com Bonvini (2009), as línguas concernentes aos africanos escravizados no Brasil provinham das regiões oeste-africana e austral, como exemplificado no Quadro 2. Apesar de os povos africanos terem sido distribuídos de modo desigual pelo território brasileiro, para que pessoas de uma mesma língua não pudessem se aglomerar numa única região, Bonvini (2009) crê que o quimbundo foi uma língua veicular desde portos de Angola até certas regiões da costa brasileira.

Quadro 2: Quadro das línguas africanas transplantadas para o Brasil.

|                                         | <b>Região oeste-africana ou sudanesa</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | <b>Região austral</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|-----------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Origem:</b>                          | Gana, Togo, Benin e Nigéria                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | Congo, República Democrática do Congo, Angola e Moçambique                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| <b>Características:</b>                 | As línguas da região sudanesa apresentavam uma grande diversidade tipológica e um número maior de famílias linguísticas.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | Havia um número reduzido de línguas na região austral, as quais eram essencialmente de base banto e faladas pela maioria dos cativos. São tradicionalmente divididas em grupos linguísticos.                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| <b>Famílias ou grupos linguísticos:</b> | <p><i>Atlântica:</i> (fula (fulfulde), uolofe, manjaco, balanta).</p> <p><i>Mandê:</i> (mandiga, bambara, maninca, diúla).</p> <p><i>Gur:</i> (subfamília gurúnsi).</p> <p><i>Cuá:</i> (subgrupo gbe: eve, fon, gen e aja, na maioria das vezes designadas pelo termo jeje).</p> <p><i>Ijóide:</i> (ijó).</p> <p><i>Benuê-congolesa</i>, que se subdividia em: <i>defóide</i> (falares iorubás designados pelo termo nagô-queto), <i>edóide</i> (edo); <i>nupóide</i> (nupe ou tapa); <i>ibóide</i> (ibo); <i>cross-Riber</i> (efique, ibíbio).</p> <p><i>Afro-asiático:</i> chádica (hauçá).</p> <p><i>Nilo-saariano:</i> saariana (canúri).</p> | <p><i>Congo (quicongo):</i> quissilongo, quissicongo (quissangala), quizombo, quissundi, quivili, iuoio e quiombe.</p> <p><i>Quimbundo</i> (falada pelos ambundos da região central de Angola), quissama e quindongo.</p> <p><i>Iaca-holo</i> (iaca, imbangala, chinji).</p> <p><i>Chôcue:</i> uchôcue, ochinganguela, chilucazi e luena (luvale).</p> <p><i>Luba:</i> chilula-cassai (lulua).</p> <p><i>Lunda:</i> chilunda, urunda.</p> <p><i>Macua:</i> omacua.</p> <p><i>Umbundo</i> (falado na região do Benguela, em Angola): umbundo, olucianeca.</p> <p><i>Cuaniana, indonga:</i> ochicuaniana, cumabi.</p> <p><i>Herero:</i> ochiherero.</p> |

Fonte: Bonvini (2009, p. 30-31).

Mattoso (1982) distingue o tráfico de africanos para o Brasil, de acordo com suas principais regiões de procedência, em quatro ciclos: no século XVI, houve o ciclo da Guiné, que trouxe para o Brasil povos sudaneses para trabalharem na produção de cana-de-açúcar; no século XVII, houve o ciclo do Congo e de Angola, que trouxe para a América Portuguesa povos bantos considerados excelentes agricultores para trabalharem na produção de cana-de-açúcar; no século XVIII, houve o ciclo da Costa da Mina, que retomou o tráfico de sudaneses para trabalhar nas minas de ouro – nesse ciclo, outros povos africanos teriam sido escolhidos para a escravização na agricultura e em serviços domésticos; no século XIX, houve o ciclo de Angola e Moçambique, que retomou o tráfico de povos bantos para trabalhar nas plantações de café.

Florentino (1997) diverge de Mattoso (1982) sobre ter sido a Costa da Mina a maior região provedora de povos africanos para a empresa do tráfico entre as costas do Atlântico Sul no século XVIII. Florentino (1997) afirma que, nesse período, teria sido dominante o sequestro de povos nativos de Angola. O autor apresenta um conjunto de dados e registros históricos que mostra que, mormente a partir da década de 30 do século XVIII, a maioria dos navios negreiros ancorados nos portos do Rio de Janeiro trazia cativos de Angola, de maneira a suplantarem os números do tráfico de viventes da Costa da Mina e a prover a região de Minas Gerais com a maioria de seus cativos. Por intermédio de uma análise criteriosa de fontes escritas, Florentino (1997) supõe que pelo menos metade dos africanos traficados para o Brasil no século XVIII aportou nos portos fluminenses, o que equivale a 850 mil negros.

De 1500 a 1800, há apenas poucos registros de censo populacional na América portuguesa. Os primeiros registros censitários mostram que a população brasileira cresceu de 1.5 milhões para cerca de 2 milhões de pessoas nas últimas três décadas do século 19. Schwartz (2001) atribui ao mercado de escravos, que traficava anualmente 20 a 40 mil africanos, o crescimento registrado da população no fim do século XVIII. Parte do incremento da população também se deu devido ao aumento do número de pessoas livres, notadamente os pardos, que apresentavam índices mais baixos de mortalidade do que os africanos escravizados (cf. SCHWARTZ, 2001). Segundo o autor, a população parda deveria se aproximar dos 28% em toda a Colônia, alcançando 30% em algumas regiões, como Bahia, Goiás, Pernambuco e Minas Gerais. O Recenseamento Geral de 1872, o primeiro censo oficial brasileiro, teve certa preocupação em constatar a demografia etnicorracial que se forjava no Império escravista. Em um estudo que reajustou os dados do censo imperial de 1872, Paiva *et al.* (2012) ratificam que a população brasileira chegara a quase dez milhões:

9.930.478. Desse total, 8.419.617 (84,8%) eram livres, enquanto 1.510.806 (15,2%) eram escravos. Ademais, a população brasileira em 1872 era composta por 3.801.692 (38,28%) pardos (afrodescendentes), 3.781.110 (38,07%) brancos, 387.234 (3,9%) caboclos, 176.057 (1,78%) africanos e 125.876 (1,27%) portugueses. Não há dados sobre os povos ameríndios.

Fundamentadas na perspectiva teórica da ecologia linguística de Mufwene (2001, 2008), Negrão & Viotti (2011, 2012, 2014a, 2014b) defendem que certas estratégias de impessoalização do sujeito no PB surgiram como um vernáculo colonial cuja gramática resulta de um processo de seleção de traços linguísticos oriundos de um banco de traços constituído pelo português vernáculo do século XVI e línguas africanas, principalmente o quimbundo, a língua africana de maior abrangência na rota do comércio transatlântico de escravos. Para Negrão & Viotti (2014a), a explicação para o estabelecimento da diferença sobre a gênese de línguas crioulas e de qualquer outro vernáculo colonial não é de natureza linguística; seguindo Mufwene (2008), as autoras argumentam que o surgimento das línguas crioulas ocorreu devido a um processo social de exclusão que caracterizou a maior parte das *plantations* da empresa colonial. Nesse contexto de exclusão, os africanos tinham pouco acesso à língua dos europeus, ao mesmo tempo em que iam perdendo o contato com suas línguas. De acordo com as autoras, ao lidar social e cognitivamente com esse impasse, os africanos tornaram-se agentes da emergência das línguas crioulas e, em contextos menos extremos, de variedades vernáculas como o PB. Negrão & Viotti (2011, 2012, 2014a, 2014b) defendem que os agentes da mudança no processo de formação do PB foram principalmente os falantes de quimbundo. Curiosamente, a primeira gramática do quimbundo foi escrita no Brasil no fim do século XVII pelo padre Pedro Dias, da Companhia de Jesus, chama-se *Arte da língua de Angola* (cf. ROSA, 2013). Além disso, Heywood (2002) declara que no século XVIII o quimbundo era uma língua franca falada majoritariamente por portugueses, afro-portugueses e africanos no Reino de Angola, um dos principais fornecedores de mão de obra escrava para o Brasil (cf. ALENCASTRO, 2000; MATTOSO, 1982; FLORENTINO, 1997).

A língua também era uma área que estava aberta à transformação sob a agência africana. Nos primeiros anos do século, o quimbundo, a língua dos Mbundu (Ambundu no português do século XVIII), era a principal língua falada pelos portugueses, afro-portugueses e africanos, livres e escravizados, no Reino de Angola. Mesmo antes do século XVIII, os portugueses reconheciam a importância do quimbundo como a língua franca da população, e os jesuítas já haviam publicado um catecismo em quimbundo em 1642. A maneira como o processo desenvolveu-se é reveladora. Embora os missionários usassem o quimbundo quando lidavam com a população africana, foi o seu uso em ambientes legais por oficiais portugueses

que permitiram que a população o usasse como língua franca. Por exemplo, a diretriz de 1737 do governador Rodrigo César de Menezes, criando uma posição de intérprete e juiz para lidar com multas e petições de liberdade em favor de escravos, estipulava que ela fosse mantida por uma pessoa de alto caráter moral e que fosse “conhecedora” da língua do povo Ambundu. (HEYWOOD, 2002, p. 103-104, tradução nossa).<sup>116</sup>

Desse contexto sócio-histórico, pode-se inferir que os usuários do quimbundo como língua franca no Reino de Angola a transplantaram para o Brasil pelas rotas do comércio transatlântico de escravos africanos; isso justificaria a elaboração de uma gramática do quimbundo em solo brasileiro quase 50 anos após padres jesuítas terem publicado um catecismo em quimbundo no Reino de Angola. Estabelecer quão significativo o quimbundo – ou qualquer outra língua africana – foi dentro do contexto histórico do Brasil colonial não é uma tarefa trivial. Além da coleta de registros e dados linguísticos provenientes de documentos históricos, como a gramática *Arte da língua de Angola*, os linguistas que lidam com a ecologia social da formação das variedades brasileiras da língua portuguesa precisam indiscutivelmente recorrer a registros censitários que evidenciam o papel do comércio transatlântico de escravos na formação étnica, linguística e populacional do brasileiro.<sup>117</sup>

Os estudos sociolinguísticos sobre as origens do PB são geralmente direcionados por duas vertentes diametralmente opostas: na primeira vertente, alega-se que mudanças em curso em dialetos do português arcaico gradualmente evoluíram para o PB quando esses dialetos foram trazidos para o Brasil; na segunda vertente, com fundamentos de uma abordagem crioulística da mudança linguística, alega-se que as mudanças no PB foram causadas por por aprendizagem parcial da gramática do português por ameríndios e africanos ou por interferência/empréstimo. Ambas as vertentes atestam que estiveram envolvidas no processo de formação do PB as línguas indígenas sul-americanas e as línguas africanas trasladadas à

---

<sup>116</sup> No original: “Language was also an area that was open to transformation under African agency. In the early years of the century, Kimbundu, the language of the Mbundu (Ambundu in eighteenth-century Portuguese), was the major language spoken by free and enslaved Portuguese, Afro-Portuguese, and Africans in the Reino de Angola. Even before the eighteenth century, the Portuguese recognized the importance of Kimbundu as the lingua franca of the population, and Jesuits had already published a Kimbundu catechism in 1642. The way the process developed is revealing. Although missionaries used Kimbundu when dealing with the African population, it was its use in legal settings by Portuguese officials that allowed the population to use it as the lingua franca. For example, Governor Rodrigo César de Menezes’s 1737 directive, creating a position of interpreter and judicial enquirer to deal with fines and petitions of liberty on behalf of slaves, stipulated that it be held by a person of high moral character and one who was “knowledgeable” in the Ambundu language.” (HEYWOOD, 2002, p. 103-104).

<sup>117</sup> Além do quimbundo, há também um registro de uma língua veicular africana oriunda da Costa da Mina falada na América portuguesa no século XVIII. Trata-se da língua mina-jeje, de base eve-fon, falada em Vila Rica, onde atualmente é Ouro Preto, Minas Gerais, e documentada na *Obra nova da lingoa geral de mina, traduzida, ao nosso idioma*, de autoria de Antônio da Costa Peixoto (cf. PESSOA DE CASTRO, 2002).

porção portuguesa do continente americano. Contudo, cada vertente de estudo atribui um valor específico a essa interação.

Baxter & Lucchesi (1997) sugeriram que o PB deve ser considerado como um caso de criouliização leve ou de semicriouliização devido à formação dessa língua estar, na visão dos autores, ligada a um processo de transmissão linguística irregular. Em trabalho mais recente, Lucchesi & Baxter (2009) afirmam que a transmissão linguística irregular no Brasil se deu inicialmente pela aquisição do português por africanos adultos, que, por serem adultos, não tinham acesso aos dispositivos necessários da faculdade da linguagem durante a aquisição do português. Como o contato linguístico se delongou, teria ocorrido a transmissão de estruturas gramaticais do português conforme aprendidas pelos adultos africanos para seus filhos em processo de aquisição de língua, do que teria resultado a mudança linguística (cf. LUCCHESI & BAXTER, 2009). A principal evidência de um processo de aquisição imperfeita do português pelos africanos fornecida pelos autores é a simplificação da morfologia verbal do PB.<sup>118</sup>

Lucchesi (2015) defende que situações de transmissão linguística irregular do tipo leve predominaram na formação das variedades populares do PB. Para o autor, essas situações foram inicialmente estabelecidas pela aquisição do português pelos cativos africanos que adquiriam o português como primeira língua e pela aquisição do português por africanos adultos. Como o contato linguístico e a polarização étnica se delongaram, houve a transmissão de estruturas gramaticais do português conforme aprendidas pelos africanos para as gerações em processo de aquisição de língua, daí teria resultado a difusão das primeiras mudanças na gramática do PB (LUCCHESI, 2015). Sob a concepção de que a realidade sociolinguística brasileira se fundamenta numa polarização entre a norma linguística da elite letrada e a norma da população marginalizada, Lucchesi (2015) descreve nas seguintes linhas o movimento histórico que teria levado a essa polarização na sociedade brasileira:

Com a consolidação da sociedade açucareira no Nordeste brasileiro no século XVII, a polarização passou a opor a língua da casa-grande, o português do senhor de engenho e dos seus, às línguas das senzalas, geralmente línguas da família banto, com proeminência para o quimbundo, o quicongo e o umbundo. No século seguinte, nas *Minas Geraes*, juntaram-se aos falantes dessas línguas os escravos do oeste-africano – falantes principalmente do iorubá, do fon, do ewe, e do hauçá – para se esfalfarem todos, nos rios e minas, na busca insana de ouro e de pedras preciosas. Por fim, já no período do Império,

<sup>118</sup> Lucchesi & Baxter (2009) distinguem entre transmissão linguística irregular do tipo que gera línguas crioulas e do tipo leve, que resulta em variedades como o PB. O primeiro tipo seria responsável pelo antigo crioulo de Helvécia no extremo sul da Bahia (cf. LUCCHESI, BAXTER & RIBEIRO, 2009).

voltaram a predominar nas grandes fazendas de café do Sudeste os escravos bantos; enquanto, na primeira capital do Brasil, predominavam os oeste-africanos, em geral os falantes do iorubá, aqui chamado nagô, língua que era usada pela população pobre de Salvador até o início do século XX. (LUCCHESI, 2015, p. 91).

Lucchesi (2015) também apresenta, entre muitos outros exemplos, a reestruturação do paradigma verbal do PB como resultado de uma erosão provocada pelo processo de transmissão linguística irregular leve envolvendo os povos de línguas africanas em contato com a língua portuguesa no Brasil. Naro & Scherre (2003, 2007) argumentam que o processo de redução morfológica do PB, na verdade, faz parte de uma deriva interna da língua, que teria tomado rumos distintos do PE por uma confluência de motivos e preponderado no PB pela “atração de forças de diversas origens – algumas oriundas da Europa; outras da América; outras, ainda, da África, que, juntas, se reforçaram para produzir o português popular do Brasil” (NARO & SCHERRE, 2007, p. 25). Com evidências de que no português arcaico havia casos de dados escritos com ausência de concordância verbal, como em (1), os autores afirmam que “não se verifica a existência de estruturas novas ou próprias do português brasileiro que teriam resultado de processos de ‘transmissão linguística irregular’, mas sim de estruturas previamente existentes cuja frequência e distribuição foram aumentadas durante tal processo” (NARO & SCHERRE, 2003, p. 295).

- (1) Português arcaico (NARO & SCHERRE, 2003, p. 293)
- a. mas DAVA lhe gram torva as portas carradas.
  - b. a todos aqueles que se fazem afora da carreira do pecado e TORNA a direita carreira.

Em suma, Naro & Scherre (2003, 2007) defendem um ponto de vista em que um certo grau de variação na concordância verbal já existia no português trazido para a América e que apenas se intensificou em uma direção, qual seja, a da não concordância. Com essa proposta, os autores rejeitam, portanto, a possibilidade de a simplificação morfológica do PB ter sido resultado, ou mesmo ampliada, devido a processos psicolinguísticos de aquisição de L2 em um contexto de *language shift*, como o vivenciado pelos povos indígenas e africanos na América portuguesa.

A fim de evidenciar que o contato linguístico também gerou estruturas inovadoras na gramática do PB, Lucchesi & Mello (2009) trazem como evidência da influência africana em variedades rurais do PB dados de alternância dativa em que se encontram duas opções



estruturais para a expressão do dativo: a construção de dativo preposicionado (2a) e a construção de objeto duplo (2b).

- (2) a. Dei os remédios aos meninos. (LUCCHESI & MELLO, 2009, p. 155)  
 b. Dei os meninos o remédio.

(2b) não é encontrada nas línguas românicas, que só permitem a construção de dativo preposicionado em (2a). Com isso, os autores afirmam que a presença de estruturas como (2b) em variedades rurais do PB podem ser tomadas como um exemplo da reestruturação original da gramática provocada pelo contato linguístico. Os autores reforçam sua hipótese mencionando o argumento de estudos em linguística de contato de que construções de objeto duplo são aparentemente um fenômeno geral nas línguas crioulas.

No âmbito dos estudos gerativistas, Avelar & Galves (2016), por exemplo, apresentam um estudo sobre um possível caso de mudança induzida por contato no PB. Os autores argumentam que, no PB, [SPEC-T] é criado assim que T é projetado, independentemente da valoração dos traços- $\phi$  de T, que são herdados de C (segundo Chomsky (2008)). No PE, por sua vez, [SPEC-T] é criado somente após C ser concatenado à estrutura oracional. A análise dos autores busca dar conta de uma variedade de fenômenos do PB, como o caso dos tópicos-sujeitos exemplificado em (3). Para os autores, essas construções são um resultado do contato linguístico do PB com línguas do grupo banto que licenciam estruturas com inversão locativa nas quais o verbo concorda com o locativo, como ilustrado em (4).

- (3) Esses carros tão parecendo que o pneu não foi trocado.  
 (AVELAR & GALVES, 2016, p. 240)

- (4) a. Kinande (AVELAR & GALVES, 2016, p. 244)

*Omo-mulongo mw-a-hik-a (?o-)mu-kali*

**LOC.19**-vila **18S**-T-chegar-FV (AUG)-CL1-mulher.1 ‘

‘Chegou uma mulher na vila’.

b. Otjiherero

*mò-ngàndá mw-á-hití òvá-ndú*

**18**-9.casa **SC18**-PASS-entrar 2-pessoas

‘Os convidados entraram na casa.’

## c. Quimbundo

*Mu njibela muala ni kitadi?*  
 LOC.18 bolso LOC18.estar com dinheiro  
 ‘Há dinheiro no bolso?’

Constratando a construção inovadora em (3) com os dados de línguas bantas em (4), Avelar & Galves (2016) questionam se não teria sido o contato linguístico o responsável por mudanças ligadas ao estabelecimento do estatuto do PB como uma língua de tópico proeminente. De acordo com os autores, há fortes razões para acreditar que os padrões oracionais com ênfase para o tópico no PB resultam de mudanças desencadeadas pelo contato linguístico envolvendo falantes de português e falantes de línguas nigero-congolesas. A alegação dos autores é a de que, no Brasil, sob a influência de línguas nigero-congolesas, especialmente línguas do grupo banto, os falantes dessas línguas em seu processo de aquisição do português transferiram propriedades de suas próprias línguas para o PB, ao passo em que também perderam propriedades da língua alvo, principalmente propriedades morfológicas. Na perspectiva dos autores, a difusão e a estabilização desse processo teriam sido possíveis graças ao enorme contingente populacional de povos africanos e afro-brasileiros no território colonial.<sup>119</sup>

Como mencionado anteriormente, Negrão & Viotti (2011, 2012, 2014b) alegam que uma série de traços linguísticos estariam depositados num banco de traços resultante do contato entre o português clássico e as línguas africanas no período colonial, especialmente o quimbundo; da competição e da seleção de traços dessas línguas teriam emergido certas inovações nas estratégias de impessoalização do sujeito na gramática do PB. Nesse sentido, as autoras argumentam que estratégias de impessoalização do sujeito como em (5) seriam um resultado da competição e da seleção de traços linguísticos presentes em (6) e (7).

(5) A casa do lado alugou ou vendeu? (NEGRÃO & VIOTTI, 2014b, p. 315)

<sup>119</sup> Apesar de a proposta de Avelar & Galves (2016) (e outras propostas, como a de Negrão & Viotti (2014b)) obedecer(em) a certos procedimentos metodológicos para o estudo da mudança induzida por contato, rejeitamo-la(s) como um exemplo autêntico de mudança induzida por contato pela seguinte razão: os dados atestados pelos autores como inovadores são descritos por ampla literatura como inovações registradas – até o momento – somente a partir da segunda metade do século XX. Consoante Thomason (2001), a mudança induzida por contato se caracteriza por apresentar um rápido surgimento de novas estruturas na língua receptora. Assim, era de se esperar que dados como em (3) (e (5)) tivesse(m) surgido na língua pelo menos no século XIX.

- (6) Português clássico (NEGRÃO & VIOTTI, 2014b, p. 318)  
 El Rei uma chamada Dona Urraqua casou com o Conde Dom Reymão de Tolosa.
- (7) Passiva/construção impessoal no quimbundo (NEGRÃO & VIOTTI, 2014b, p. 315)  
*Nzua a-mu-mono kwa meme.*  
 João eles-ele-viram por mim  
 ‘João foi visto por mim.’ (lit.: ‘João, eles o viram por mim’)

Negrão & Viotti (2011, 2012, 2014b) buscam explicar o surgimento da construção em (5) a partir da possibilidade de uma competição entre as sentenças em (6) e (7) no período colonial. Para as autoras, as construções (6) e (7) possuem uma estrutura subjacente bastante similar que as colocaria em competição, qual seja: em todas elas, o objeto é deslocado para uma posição pré-verbal e o sujeito vem logo em seguida acompanhando de um verbo que está na forma ativa, conforme representado em (8a) e (8b), respectivamente.<sup>120</sup>

- (8) a. [DP tema deslocado para posição pré-verbal ][Ø sujeito nulo anafórico ][V forma ativa/concordância com antecedente do sujeito nulo]  
 b. [DP tema deslocado para posição pré-verbal] [prefixo sujeito anafórico OU com leitura impessoal + prefixo objeto + V forma ativa]

Para as autoras, a inovação na gramática do PB constatada em (5) seria um possível resultado da competição e seleção dos traços ilustrados em (8). As autoras destacam que em (5), ao ser deslocado para uma posição pré-verbal, o objeto é reinterpretado como o sujeito da sentença no PB. Elas mencionam também que construções como (5) são amplamente tratadas na literatura como um resultado da perda do clítico *se*-impessoal no PB e da exigência dessa língua em preencher a posição de sujeito, ainda que com sujeitos não canônicos. Negrão & Viotti (2011, 2012, 2014b), todavia, não propõem uma relação direta entre a perda do clítico *se*-impessoal e a influência do contato linguístico – o que nos parece estranho. Para as autoras, a perda do clítico *se* em construções como (5) está relacionada a questões de natureza semântica, ao passo que a reanálise do objeto como sujeito nessa construção está relacionada

<sup>120</sup> A nosso ver, para Mufwene (2001, 2008), não é a congruência entre estruturas que possibilita a competição e a seleção de traços linguísticos, mas a ausência de correspondência morfossintática entre estruturas que possuem o mesmo significado, como no contraste que fornecemos a seguir.

- (i) Eu quero comprar o mesmo livro que [a Maria comprou].  
 (ii) Eu quero comprar o mesmo livro que [comprou a Maria].

com o contato linguístico. Negrão & Viotti (2014b) destacam que a evolução da passiva do quimbundo em (7) tem sua origem diacrônica na confluência de uma construção de deslocamento à esquerda, como ilustrado em (9a), e de uma construção com sujeito impessoal com morfologia de terceira pessoa do plural no verbo, como ilustrado em (9b):

- (9) a. Deslocamento à esquerda com NP-sujeito pleno

*Nzua, aana a-mu-mono*

João, filhos eles-ele-viram

‘João, os filhos o viram.’

- b. Deslocamento à esquerda com sujeito pronominal:

*Nzua, a-mu-mono*

João, eles-ele-viram

1. Ativa anafórica: ‘João, eles o viram.’ (anafórica/ativa)

2. Passiva impessoal: ‘João, ele foi visto.’ (impessoal/passiva)

(NEGRÃO & VIOTTI, 2014b, p. 308-309)

Em suas pesquisas, as autoras contataram um falante de quimbundo em Angola que forneceu evidências de que as construções em (9) são, de fato, produtivas na língua. Dados originais e respectivas traduções fornecidas pelo falante são exemplificados em (10).

- (10) a. Quimbundo: deslocamento à esquerda com NP-sujeito pleno

*Nzua, ana a-mu-mono*

João os filhos lhe viram

- b. Quimbundo: deslocamento à esquerda com sujeito pronominal

*Nzua, a-mu-mono*

João lhe viram

c. *Quimbundo: passiva*

*Nzua a-mu-mono kwa’na*

João lhe viram nos filhos

(NEGRÃO & VIOTTI, 2014b, p. 311)

As autoras constatarem que a estrutura da passiva no quimbundo teve um forte impacto na gramática do português angolano, possibilitando o surgimento das seguintes estruturas:

- (11) a. O João lhe bateram na mãe dele. (NEGRÃO & VIOTTI, 2014b, p. 313)  
 ‘O João foi batido pela mãe.’  
 b. O papá estão a lhe bater num doente.  
 ‘O papai está sendo batido por um doente.’  
 c. (...) quando tinham-lhe tirado o cabasso num tio dela.  
 ‘(...) quando o cabaço lhe foi tirado por um tio dela.’

A comparação entre (possíveis) resultados do contato linguístico do português com o quimbundo no continente americano com os resultados desse mesmo contato em Angola remete à proposta de Petter (2009) a respeito de serem os africanos os principais agentes responsáveis pelas variedades do português no Brasil e na África (Angola e Moçambique). Para Petter (2009), essas variedades constituem o que ela denomina de *continuum* afro-brasileiro do português. Apesar das idiosincrasias presentes na ecologia que propiciou a emergência de cada uma dessas variedades do português, Petter (2009) sugere que as semelhanças entre elas se devem ao contato do português com línguas do grupo banto com características linguísticas muito próximas.<sup>121</sup>

São tantas as semelhanças compartilhadas pelas três variedades de português [brasileira, angolana e moçambicana] nos três níveis de organização linguística selecionados (fonológico, lexical e morfossintático) que fica difícil defender que tais fatos sejam casuais, resultantes de uma deriva natural do português ou decorrentes da manutenção de formas antigas do PE. Por que as mesmas áreas da gramática do português foram “perturbadas”? A hipótese de que essas mudanças tenham sido introduzidas por falantes de línguas africanas, tanto na África quanto no Brasil, impõe-se de forma contundente, mesmo que se considere que no Brasil falantes de línguas indígenas e de outras línguas europeias tenham participado da constituição do PB. É provável que o contato com as LB (línguas do grupo banto) não seja a única explicação para as mudanças observadas no PB, mas esse contato é com certeza bastante relevante. (PETTER, 2010, p. 171-172).

Utilizando dados similares aos dados de (5) a (11), Petter, Negrão & Viotti (2018) realizam um estudo comparativo entre o PB, o português angolano, o quimbundo e também outras línguas bantas da família de línguas nígero-congolesas na perspectiva de um *continuum*

<sup>121</sup> Holm (2009), um defensor do papel de processos de crioulização na formação do PB, também faz um estudo comparativo de certas propriedades das variedades brasileira e angolana da língua portuguesa. O autor defende que o contato linguístico induziu essas línguas à redução de suas propriedades morfossintáticas. Adicionalmente, o autor argumenta que, conforme esperado em linguística de contato, as mudanças ocorridas em cada uma dessas línguas tendem a tornar as variedades parcialmente reestruturadas mais semelhantes às suas línguas de substrato.

afro-brasileiro do português. As autoras afirmam que as construções brasileiras ou angolanas não derivam direta ou linearmente de uma estrutura do português arcaico ou de qualquer língua banta angolana. Elas argumentam que, na verdade, as similaridades entre esses dados reforçam a hipótese de um *continuum* linguístico África-Brasil, o qual é explicado pelas autoras a partir da proposta de Mufwene (2008) de que em situações de contato linguístico é constituído um conjunto de traços gramaticais do qual traços de diferentes línguas são selecionados pelos falantes em suas interações sócio-comunicativas: do uso recorrente dos traços selecionados é formada a base das línguas que emergem dessas ecologias de contato linguístico.

## **7.2 Algumas considerações sobre a origem sócio-histórica da gramática do português em Goiás**

Nesta seção, descrevemos sucintamente o quadro populacional da capitania e da província de Goiás (séculos XVIII e XIX), considerando especialmente o sistema escravista e a economia mineradora e rural da época. Nosso objetivo é fornecer evidência para a hipótese de que a variedade da língua portuguesa que emergiu nessa região tem certa influência do contato linguístico nos séculos em tela.

Inicialmente, destacamos que Goiás era um território multilíngue antes da chegada dos portugueses. Os nativos que habitavam a região são classificados, segundo os troncos linguísticos a que pertenciam(em), em dois grandes grupos: (i) os jês, grupo de maior número populacional, compostos pelos akuên (akroás, xabriacás, xavantes e xerentes), os caiapós (caiapós do sul e caiapós setentrionais), os timbiras (apinajés, krahós, gavião, canela, afotogés, corretis, otogés, porecramecrãs, macamecrãs e temembus) e os carajás (carajás, javaés e xambioás); e (ii) os tupis, mais precisamente os avá-canoeiros, os tapirapés, e os guajajaras (teneteara) (cf. CHAIM, 1974; ROCHA, 1998; KARASCH, 2016).

Entre 1772 e 1778, os registros censitários informavam uma população de cerca de 55.000 habitantes na capitania de Goiás. Em 1804, houve o primeiro censo oficial da capitania, que registrou uma queda no número de habitantes para 50.764, dos quais: 7.131 eram brancos; 16.531 eram pardos; 7.943 eram negros; e 19.159 eram escravos – isto é, 14.05% da população eram de pessoas brancas e 85.95% eram compostas por escravos, negros ou pardos (cf. PALACÍN & MORAIS, 2008, p. 49-64). A maioria dos censos realizados naquele período ignorava a existência dos povos indígenas em Goiás. Karasch (2016) chama atenção para o fato de que as narrativas oficiais setecentistas sobre Goiás não

especificam se os termos cativos, negros ou escravos faziam referência, de fato, a africanos ou cativos indígenas. Até mesmo o uso do termo *negro*, de acordo com a autora, pode ser enganador porque, naquele período, os chamados *negros da terra* não eram negros da África, mas povos ameríndios. Karasch (2016) afirma que os cativos que acompanhavam as bandeiras até Goiás eram provavelmente homens e mulheres indígenas, uma vez que os africanos traficados para São Paulo eram raros e caros. A autora afirma ainda que, com a exploração das minas, uma maior quantidade de homens africanos foram levados para buscar ou banhar ouro no rio Vermelho, próximo a ele, ou até a sua junção com o rio Araguaia.

Karasch (2018) explica que, em 1832, os registros censitários indicaram pela primeira vez o quantitativo de africanos escravizados em Goiás: da população total de 13.261 negros escravizados, 1.923 (16.6%) eram africanos. 9.652 (83.4%) eram brasileiros. E 1.686 eram chamados *pardos*. A autora afirma que a quantidade de africanos que continuaram a viver em Goiás após esse período é incerta, especialmente porque muitos deles escaparam para os quilombos ou compraram sua liberdade.<sup>122</sup> Em 1832, além de rastrear o declínio geral dos afro-brasileiros escravizados e a criouliização da população escrava, os censos registraram uma população escrava dominada por homens, com uma proporção de três homens para cada mulher (KARASCH, 2018).

Karasch (2018) afirma que o censo de 1804 revela a decadência geral na economia, evidenciada pelo declínio da população escrava e pelo aumento na porcentagem de mulheres na população escrava, isto é, 40% (nos períodos anteriores os homens eram majoritários devido ao trabalho nas minas de ouro). Os censos de 1825 e 1832, por fim, aponta Karasch (2018), registraram 45,2% e 45,6% de mulheres, indicando a tendência para o quadro populacional de negros escravizados no século XIX, quando a proporção de mulheres escravas sobe constantemente e supera a dos escravos do sexo masculino.

Na Tabela 1, temos informações a respeito dos povos africanos e afro-brasileiros escravizados na capitania de Goiás entre os anos de 1783 e 1832. Trata-se de um minucioso levantamento feito por Karasch (2018) em seus estudos sobre a presença de centro-africanos na região Centro-Oeste do Brasil. A autora utilizou diversas fontes, como registros de batismos e morte, de impostos, alforrias e irmandades, para chegar aos resultados e às principais conclusões de seu estudo.

---

<sup>122</sup> Segundo Palacín & Moraes (2008), todos os arraiais no século XVIII em Goiás viveram à sombra de quilombos. A maior comunidade brasileira de remanescentes de quilombo, o Kalunga, com aproximadamente 4 mil almas, está localizada em Goiás, em regiões montanhosas nos arredores do município de Cavalcante.

Tabela 1: Escravos em Goiás: 1783-1832

| <b>LOCALIZAÇÃO</b>      | <b>1783</b>   | <b>1789</b>   | <b>1792</b>   | <b>1804</b>   | <b>1825</b>   | <b>1832</b>   |
|-------------------------|---------------|---------------|---------------|---------------|---------------|---------------|
| <b>COMARCA DO SUL</b>   |               |               |               |               |               |               |
| Vila Boa                | 4,689         | 9,200         | 8,568         | 4,432         | 3,274         | 3,073         |
| Meia Ponte              | 1,682         | 4,777         | 4,855         | 2,282         | 1,842         | 1,800         |
| Santa Luzia             | 899           | 2,960         | 2,491         | 1,264         | 741           | 741           |
| Santa Cruz              | 723           | 1,223         | 1,153         | 997           | 887           | 1,094         |
| Pilar                   | 1,567         | 1,967         | 3,839         | 1,575         | 969           | 1,033         |
| Crixas                  | 1,207         | 2,444         | 2,045         | 634           | 699           | 384           |
| Rio das Velhas          | 299           | 277           | 2,261         | -             | -             | -             |
| Desemboque              | -             | -             | -             | 660           | -             | -             |
| Carretão                | -             | -             | -             | -             | -             | 5             |
| Traíras                 | -             | -             | -             | 2,807         | -             | -             |
| <b>Total</b>            | <b>11,066</b> | <b>22,848</b> | <b>25,212</b> | <b>14,651</b> | <b>8,412</b>  | <b>8,130</b>  |
| <b>COMARCA DO NORTE</b> |               |               |               |               |               |               |
| Vila de São João        | -             | -             | -             | -             | 78            | 228           |
| Traíras                 | 3,790         | 6,245         | 5,328         | -             | 1,493         | 1,441         |
| Cavalcante              | 923           | 993           | 950           | 1,209         | 456           | 474           |
| Flores                  | -             | -             | -             | -             | 478           | 561           |
| São Félix               | 648           | 2,707         | 2,599         | 641           | 142           | 231           |
| Arraias                 | 363           | 1,198         | 1,198         | 469           | 765           | 792           |
| Conceição               | -             | 986           | 908           | 684           | 271           | 156           |
| Natividade              | 923           | 2,332         | 2,338         | 1,529         | 904           | 879           |
| Porto Real/Imperial     | -             | -             | -             | -             | 376           | 325           |
| Carmo                   | -             | -             | -             | 844           | -             | -             |
| Carolina                | -             | -             | -             | -             | -             | 39            |
| Duro                    | -             | -             | -             | -             | -             | 5             |
| <b>Total</b>            | <b>6,647</b>  | <b>14,461</b> | <b>13,321</b> | <b>5,376</b>  | <b>4,963</b>  | <b>5,131</b>  |
| <b>TOTAL GERAL</b>      | <b>17,713</b> | <b>37,309</b> | <b>38,533</b> | <b>20,027</b> | <b>13,375</b> | <b>13,261</b> |

Fonte: Karasch (2018, p. 131).

Karasch (2018) assevera que os registros fiscais do início do século XIX consubstanciam os padrões de comércio entre os portos da costa e a capitania de Goiás. O resultado é o seguinte: os escravos enviados pelo porto de Salvador tendiam a ser do oeste da África e os enviados pelo Rio de Janeiro eram principalmente da região central. A autora relata, ainda, que os centro-africanos eram mais significativos em números na comarca do sul, enquanto os africanos ocidentais eram predominantes na comarca do norte, sendo que, de acordo com a autora, a maioria dos centro-africanos presentes na comarca do norte eram angolanos. Verifica-se, por meio de suas origens geográficas, que os escravos de Goiás eram extraordinariamente diversos em suas identidades étnicas.

Por meio de um inventário detalhado dos escravos pertencentes a José Francisco Hutim, cujas propriedades eram perto de Vila Boa (atual cidade de Goiás) e foram



confiscadas pelos portugueses sob a alegação de que ele e seus escravos estavam roubando ouro, Karasch (2018) identifica o que ela considera um padrão de mistura étnica entre escravos angolanos e africanos ocidentais em Goiás. Para a autora, os documentos sobre os escravos de Hutim são como um registro mais geral daquele tempo: 84 escravos descritos nos seus diversos locais de trabalho (serviço doméstico, engenho de açúcar etc.) no ano de 1805. Em duas das propriedades de Hutim, Karasch (2018) informa que homens angolanos mais velhos ocupavam posições de estatura: um como feitor e o outro como um especialista em açúcar. O inventário parece revelar que havia algum senso de hierarquia social dentro da força de trabalho escrava, e os angolanos podem ter tido alguma vantagem por causa da antiguidade, isto é, devido ao tempo de serviço. O inventário também sugere que as escravas e os escravos mais jovens se dedicavam aos serviços domésticos.

Em 1872, os dados do Recenseamento Geral do Império constataram que a população da província de Goiás ultrapassara os 160 mil habitantes, dos quais cerca de 140 mil (87,5%) eram de nascidos na província. Havia 132.027 pessoas livres, das quais 56.361 mil eram profissionais agrícolas: 52.361 eram lavradores e 3.829 eram criadores de animais, o que equivale a aproximadamente 35% da população livre. Os escravos somavam 10.652 almas, das quais 8.629 (ou seja, 81%) tinham nascido ali. Conforme podemos verificar na Tabela 2, pardos e negros excediam 70% da população goiana no último quarto do século 19.

Tabela 2: Quadro étnico na província de Goiás.

| CATEGORIA | TOTAL  | PORCENTAGEM |
|-----------|--------|-------------|
| Etnia:    |        |             |
| Branco    | 41.938 | 26,2%,      |
| Pardo     | 90.490 | 56,4%       |
| Preto     | 23.710 | 14,8%,      |
| Caboclos  | 4257   | 2,6%        |

Fonte: Paiva *et. al.* (2012).

Em relação aos dados relacionados às tabelas 1 e 2, ressaltamos que as categorias que definem as identidades étnicas, tanto da capitania quanto da província de Goiás, são imprecisas (a mesma imprecisão se estende a todo o território brasileiro), pois, além de os censos omitirem dados sobre os indígenas, não houve a utilização de critérios objetivos para distinguir quem era *negro*, *escravo* ou *pardo* no censo de 1804. Esses impasses em relação à atribuição de cor e de identidade étnica nos remetem a Soares (2009, p. p. 46), que afirma que tanto a nação de origem quanto a cor foram atributos da Colônia aos indígenas e aos negros

escravizados e estavam sujeitas a variações regionais e investidas de crenças muito pouco convincentes: “A documentação reflete o que se escreve dos negros (...) e sobre os indígenas. É a palavra daquele que domina dando cor e signos àqueles de quem se escreve”.

Dito isso, tomamos agora a noção de *descontinuidade* de Lightfoot (2017b), que prevê a existência de descontinuidades entre as estruturas sintáticas usadas de uma geração para a próxima, para pensar a mudança na gramática do PB em Goiás, considerando os aspectos populacionais dos séculos XVIII e XIX daquela região apresentados neste capítulo. Dado o alto contingente populacional de africanos e afro-brasileiros (descendentes de africanos escravizados) inseridos nas economias mineira e agrícola por meio do sistema escravista nos séculos XVIII e XIX em Goiás, é razoável supor que o maior número de dados de língua-E fornecidos como *input* para a aquisição do português como primeira língua tenha sido produzido pelo maior contingente de povos unidos étnica, social e linguisticamente naquele território: em 1804, mesmo com o declínio de escravos devido ao declínio do ciclo aurífero em Goiás, africanos e afro-brasileiros correspondiam a 85.95% do total da população; em 1872, excediam pouco mais de 70% da população. Parece não haver registros históricos de que os povos africanos e os afro-brasileiros em Goiás tenham desenvolvido algum sistema linguístico de comunicação, como um pidgin, por exemplo, ou tenham utilizado alguma língua franca na ecologia social em que viviam. O português foi provavelmente a língua predominantemente falada por esses povos em Goiás.

Desse modo, nos parece mais provável supor que a mudança na gramática do século XIX em Goiás tenha se originado, sobretudo, em dados da língua-E de africanos e afro-brasileiros (que adquiriram sua língua materna de dados linguísticos primários fornecidos por seus pais africanos), os quais difundiram esses dados por serem maiores em termos populacionais e devido à maior mobilidade que possuíam pelo território. Karasch (2018) relata que boa parte dos escravos africanos e afro-brasileiros em Goiás possivelmente viveu em outros territórios, como Salvador e Minas Gerais, antes de serem enviados para Goiás. Mattos e Silva (2004) apresenta a demografia e a mobilidade dos povos africanos como os dois fatores sociolinguísticos definidores da difusão do PB por todo o território brasileiro.<sup>123</sup>

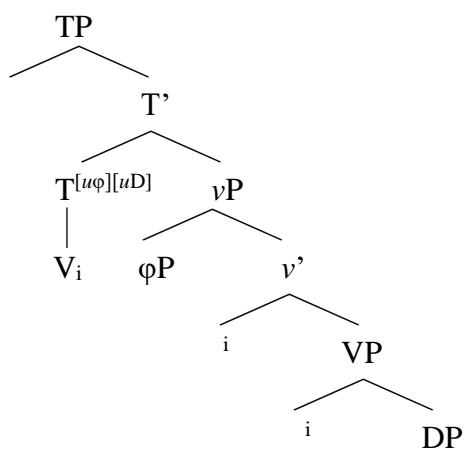
---

<sup>123</sup> Em estudo sobre formas de indeterminação do sujeito em quatro comunidades rurais afro-brasileiras da Bahia, Lucchesi (2014) mostra a ocorrência de construções com sujeito nulo indeterminado com verbo na 3ª pessoa do singular nessas comunidades – esse tipo de construção foi uma das evidências que apresentamos no Capítulo 6 como uma forte evidência da mudança na configuração da gramática do PB oitocentista em Goiás como uma LSN parcial. Dessa forma, se, assim como Mattos e Silva (2004), pensarmos que os quilombos e as comunidades rurais afro-brasileiros são como laboratórios da constituição do PB, por causa de a maioria deles ter tido e ainda ter o português como língua veicular e por ter admitido povos indígenas e também portugueses como membros de seu coletivo, a hipótese de que a matriz africana é mais robusta na formação do PB prevalece.

Consoante com Negrão & Viotti (2014a), sugerimos também que é bastante razoável supor que, devido à constante interação com africanos que falavam diferentes variedades de português como segunda língua, os falantes nativos de português possam ter sido influenciados por essas diferentes variedades, fazendo com que seus próprios idioletos se adaptassem às novas situações linguísticas que eles estavam enfrentando. Essa situação, aliás, é uma premissa de que em situações de contato linguístico, os interlocutores se interessam por projetos conjuntos de construção de significados, não por preservar a estrutura de suas próprias línguas nativas (cf. NEGRÃO & VIOTTI, 2014a).<sup>124</sup>

Como propomos no capítulo anterior, a perda do traço-D teria ocasionado uma série de mudanças na gramática do PB em Goiás no século XIX. Nosso principal candidato social para a introdução da perda desse traço na história linguística de Goiás são os escravos africanos e afro-brasileiros. Já o principal candidato linguístico, isto é, o gatilho inicial para formalizar o processo de mudança por meio de dados linguísticos primários é a perda da inversão VS.<sup>125</sup> No capítulo anterior, também ilustramos alguns mecanismos de checagem do traço-EPP na gramática do PB setecentista. Em um desses mecanismos, o mais predominante nos resultados estatísticos, o verbo se move para T e um sujeito nulo de 3ª pessoa ( $\varphi$ P) é projetado em [SPEC-TP], como ilustrado novamente em (12); no segundo mecanismo mais produtivo, o movimento do verbo para T era suficiente para satisfazer o traço-D de T (o EPP da oração), ao passo em que o sujeito permanecia em [SPEC-vP], gerando a ordem VS, como em (13).

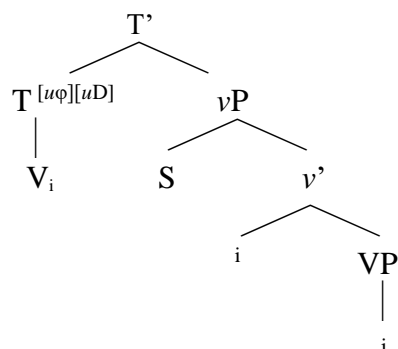
(12) Mecanismo de checagem do EPP para a 3ª p. na gramática setecentista: sujeito nulo



<sup>124</sup> Conforme Karasch (2018) constatou, escravos angolanos lideravam a segurança e o engenho em uma fazenda nas proximidades da cidade de Goiás. É sobre esse tipo de situação que estamos falando nesse parágrafo.

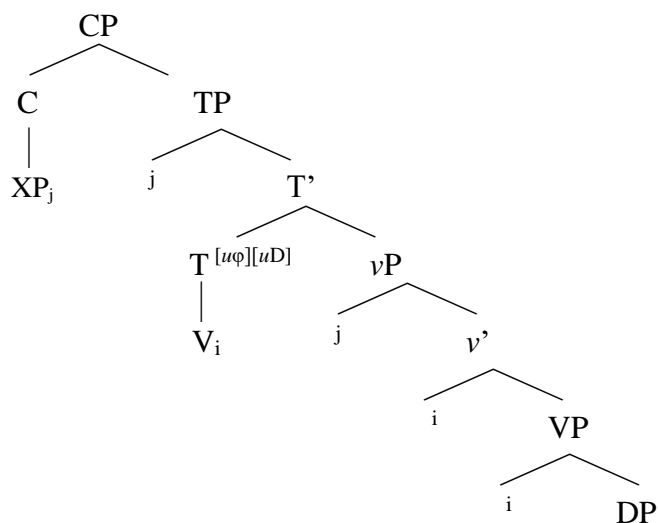
<sup>125</sup> Para uma leitura sobre o papel da ordem VS na emergência da gramática do PB sem qualquer correlação com o contato linguístico, remetemos nosso leitor ao trabalho de Pires (2006), pioneiro em relacionar a perda da inversão livre à emergência da gramática do PB.

- (13) Mecanismo de checagem do EPP na gramática setecentista: V-para-T e VS

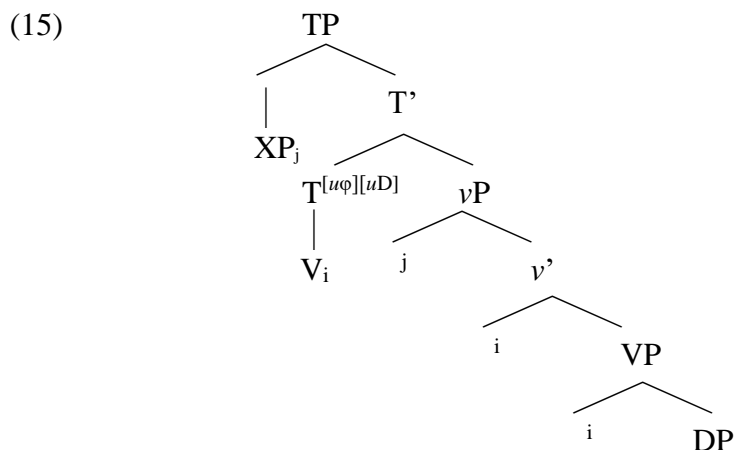


Adicionalmente, um outro mecanismo, ilustrado em (14) abaixo, também era utilizado para a checagem do traço-EPP na gramática setecentista. Assumimos que os sujeitos em posição pré-verbal numa estrutura como (14) devem ocupar uma posição A-barra (cf. ALEXIADOU & ANAGNOSTOPOULOU, 1998).

- (14) Mecanismo de checagem do EPP no século XVIII: movimento A-barra



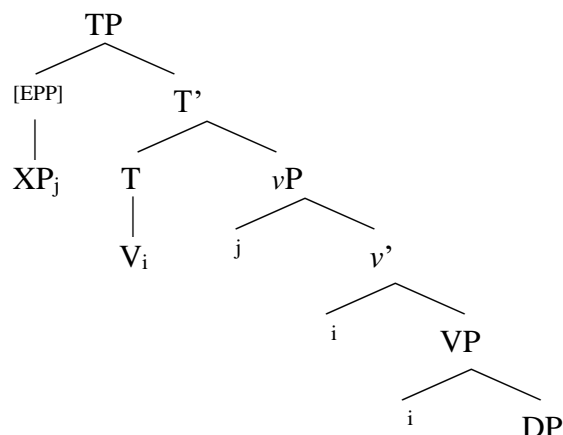
Fizemos, ainda, menção a uma quarta opção de satisfação do traço-EPP na gramática setecentista. Indicamos que nessa alternativa um sujeito manifesto, especialmente de 3ª pessoa, se moveria para [SPEC-TP], apesar de seu traço-D (isto é, seu traço-EPP) ter sido satisfeito pelo movimento do verbo para T. Representamos a seguir a derivação dessa estrutura.



Em (15), o traço de definitude do sujeito, rotulado [ $\mu$ D], é satisfeito pelo movimento do verbo para T; com o movimento de um XP para [SPEC-TP], um novo índice de referencialidade entra na derivação. A estrutura em (15) era extremamente rara nos dados setecentistas do PB em Goiás. Considerando (15), nossa proposta para a perda do traço-D do conjunto de traços de T na gramática do PB em Goiás (e muito possivelmente de muitas outras variedades do PB) é a seguinte: falantes não nativos do português inseridos nas principais atividades econômicas daquele período (mineração, agronegócio, serviços domésticos) produziam, como dados de suas línguas-E, mais estruturas como (15) do que aquelas em (12) e (13).

O uso massivo dessa estrutura por esses falantes não nativos do português teria levado, em um primeiro momento, a uma perda da ordem VSX nos dados de língua-E; assim, especialmente no processo de aquisição de língua das novas gerações, a estrutura em (15) era a mais robusta nos dados linguísticos primários a que as crianças eram expostas para a aquisição do português como língua materna. Nesse processo de aquisição de língua, a gramática das crianças interpretou o índice de referencialidade da estrutura do XP em [SPEC-TP] como o único necessário para satisfazer o EPP da sentença, descartando o traço-D de T de suas línguas-I. Isso teria gerado, portanto, o novo e predominante mecanismo de checagem do traço-EPP na gramática oitocentista do PB em Goiás ilustrado em (16) – nota-se que o mecanismo de checagem do traço-EPP em (16) é idêntico ao mecanismo em (15), mas não possui um traço-D em T.

## (16) Mecanismo de checagem do EPP na gramática oitocentista do PB em Goiás



O traço-D em T é um elemento semântico sem um expoente morfofonológico, por isso, é razoável pensar que se trata de uma noção cognitiva mais difícil de ser apreendida em situações de contato linguístico marcadas por multilinguismo generalizado e relações sociais tão estruturalmente desiguais como era a situação do contato linguístico no período colonial brasileiro. Nesse contexto, portanto, satisfazer o traço de definitude da oração por meio do movimento de um XP com índice referencial para [SPEC-TP] era a opção mais viável para aqueles que adquiriram o português como segunda língua, principalmente para o contingente de falantes de línguas nigero-congolesas, línguas que em sua maioria tinham como estrutura básica a ordem SVO (cf. PETTER, 2015; PAL & ARAÚJO, 2015).

Em sua gramática seiscentista sobre o quimbundo, língua banta veicular na rota do tráfico transatlântico, o padre jesuíta Pedro Dias (1697) declara, por exemplo, que a ordem SVO, ilustrada em (17b), era preferível em quimbundo (cf. ROSA, 2013).

(17) a. *O mutu ou eme ngamugibile.*

a pessoa esta eu a matei

‘Esta pessoa, eu a matei.’

b. *Eme ngagibile o mutu ou.*

eu matei a pessoa esta

‘Eu matei esta pessoa.’

(DIAS, 1697, *apud* ROSA, 2013, p. 191)

Pedro Dias (1697) explica que o uso de um nominativo em posição pós-verbal, menos recorrente, também era gramatical em quimbundo e fornece como um exemplo o dado em (18), em que um locativo aparece antes do verbo.

- (18) *Mo maca aa amba atu ayari*  
 nesta prática falam duas pessoas  
 ‘Nesta conversa falam duas pessoas.’ (DIAS, 1697, *apud* ROSA, 2013, p. 197)

A ordem SVO também era encontrada nas sentenças interrogativas, como exemplificado em (19).

- (19) *Nzambi üazola atu osso?*  
 Deus ama a todos  
 ‘Deus ama a todos?’ (DIAS, 1697, *apud* ROSA, 2013, p. 207)

Comparando o uso de artigos entre o português e o quimbundo, Pedro Dias (1697) descreve alguns sintagmas nominais no quimbundo seiscentista nucleados pela partícula *o*, pronunciada de modo quase “gutural e sem ápices” nas palavras do autor. O autor exemplifica esse uso no quimbundo com os dados em ordem SVO em (20).

- (20) a. *Nzambi ubana o mala o pembela yâyê.*  
 Deus dá a os homens a sua graça  
 ‘Deus dá aos homens a sua graça.’  
 b. *O ginganga jaiüaba.*  
 os padres são bons  
 ‘Os padres são bons.’ (DIAS, 1697, *apud* ROSA, 2013, p. 195)

Pedro Dias (1967) exhibe alguns dados do quimbundo em que o sujeito pronominal pode ser omitido tendo em vista sua retomada pelas flexões dos verbos (sempre prefixais em quimbundo), como em (21b).

- (21) a. *Eyè üacolo, eme pê ngacolo.* (DIAS, 1697, *apud* ROSA, 2013, p. 195 e 197)  
 vós estais bem, eu também estou bem  
 b. *Eyè üacolo, ngacolo pé.*  
 vós estais bem, estou bem também  
 ‘Vocês estão bem, (eu) também estou.’

Os pronomes nominativos no quimbundo seiscentistas são (DIAS, 1697, *apud* ROSA, 2013, p. 145):

- |      |                 |                |
|------|-----------------|----------------|
| (22) | a. No singular: | b. No plural:  |
|      | 1P: <i>Eme</i>  | 1P: <i>Etu</i> |
|      | 2P: <i>Eyè</i>  | 2P: <i>Enu</i> |
|      | 3P: <i>Ae</i>   | 3P: <i>Ao</i>  |

Já os morfemas distintivos de pessoa nos verbos são (DIAS, 1697, *apud* ROSA, 2013, p. 151):

- |      |                  |                |
|------|------------------|----------------|
| (23) | a. No singular:  | b. No plural:  |
|      | 1P: <i>Ngui-</i> | 1P: <i>Tu-</i> |
|      | 2P: <i>Ũ-</i>    | 2P: <i>Mu-</i> |
|      | 3P: <i>Û-</i>    | 3P: <i>A-</i>  |

Curiosamente, Pedro Dias (1697) descreve que os sujeitos pronominais de 1ª e 2ª pessoa (*eme* e *eyè* no exemplo a seguir), além de antecederem o verbo, eram manifestos quando nos predicados havia uma relação contrastiva entre esses sujeitos.

- (24) *Eme ngandala culunda o milonga yâ nzambi, eyè cuandala cuilunda.*  
*eu guardo os preceitos de Deus, tu não queres guardá-los.*  
 ‘Eu quero guardar as palavras de Deus, tu não queres guardá-las.’

(DIAS, 1697, *apud* ROSA, 2013, p. 197)

Tendo em vista a constatada perda da ordem VS na gramática oitocentista do PB em Goiás (cf. Capítulo 6) e o fato de que a língua portuguesa esteve em contato com um grande contingente populacional de povos africanos falantes de línguas nigero-congolesas durante o período colonial, especialmente os falantes de línguas bantas como o quimbundo, consideramos que o estabelecimento da ordem VSO> SVO no século XIX, em comparação com a predominante ordem VS no século XVIII, pode ser visto como uma influência de substrato de línguas nigero-congolesas, especialmente o quimbundo, na gramática oitocentista do PB em Goiás.



Em comunicação pessoal, Márcia Oliveira nos questiona se, ao estabelecermos uma relação entre a ordem SVO na gramática oitocentista do PB com as línguas nigero-congolesas faladas na América portuguesa, não estaríamos recorrendo ao princípio de cafeteria (*Cafeteria Principle*) proposto por Dillard (1970, *apud* PARKVALL, 2012). O princípio de cafeteria se caracteriza, na crioulística, em o pesquisador examinar um número razoável de línguas africanas até encontrar o traço desejado e, com o dado em mãos, apontá-lo como a origem do traço de um crioulo transatlântico. Discordamos da pesquisadora nesse sentido, pois, nossa proposta, ainda que necessite de estudos posteriores mais robustos, observa cuidadosamente aquilo que na tradição crioulística é chamado de o *Edito de Bickerton*, o contra-argumento do princípio de cafeteria: “para qualquer variedade linguística inovadora como a fonte de uma dada característica, é preciso que tenha havido falantes presentes no lugar certo e no momento certo” (PARKVALL, 2012, p. 46). Assim, a despeito das imprecisões sobre os conceitos de raça e cor nos censos coloniais, nossa proposta está fundamentada nas evidências demográficas de que povos africanos, especialmente os falantes de línguas nigero-congolesas, estavam “no lugar certo e no momento certo”: durante o comércio do tráfico atlântico pelo menos 4 milhões de africanos foram traficados para o Brasil, a maioria deles oriundos das regiões em que predominavam as línguas nigero-congolesas. Ademais, como vimos no início desta seção, no século XVIII, após instauradas as primeiras minas de exploração em Goiás, a maior parcela da população era composta por escravos africanos e seus descendentes<sup>126</sup>, estes últimos, também escravizados, provavelmente já haviam adquirido o português L2 de seus pais como primeira língua, o que sugere, ainda, que a mudança nos dados da fala pode ser anterior a mudança que constatamos na escrita.

A questão do papel do contato linguístico na gênese da gramática do PB em Goiás no século XIX não termina, todavia, com a formalização dessa proposta interpretativa que consideramos aqui. É preciso, ainda, auferirmos algum valor à contribuição indígena na emergência dessa gramática. Acreditamos que a influência ameríndia na gramática do PB em Goiás provém principalmente: (i) dos indígenas que saíram da região sudeste do Brasil junto às bandeiras para desbravar o Brasil Central e (ii) dos brasileiros de ascendência ameríndia falantes de português como primeira ou como segunda língua que levaram para a região Centro-Oeste do Brasil uma variedade do português marcada por influências indígenas,

---

<sup>126</sup>Como mencionamos anteriormente, os povos africanos eram possivelmente mais raros no período das bandeiras, isto é, das primeiras expedições colonizadoras ao interior de Goiás (cf. KARASCH, 2016). Assim, estamos considerando nessa afirmação o período após a elaboração dos dados censitários, que registram um número massivo de povos africanos e de seus descendentes após o início da exploração de minas na região.

inclusive influências da língua geral paulista<sup>127</sup>. Esse é o ponto de vista, por exemplo, de Borges, Lima-Salles e Pilati (2014): para esses autores, falantes da língua geral paulista como primeira língua foram para o território goiano e ali permaneceram até possivelmente o início do século XIX, quando a economia mineradora entrou em colapso e os bandeirantes falantes da língua geral paulista como primeira língua abandonaram a região em busca de outras fontes de renda. A argumentação dos autores, especialmente o papel atribuído aos povos indígenas na formação do PB, é fundamentada no trabalho de Lobato (2006).

Reconhecer os aportes indígenas e africanos ao olhar para os dados da gramática do PB oitocentista em Goiás é pensar também que, de certo modo, eles expressam um resultado de interações afro-indígenas vivenciadas no período colonial, tal como aquelas constatadas atualmente em comunidades rurais brasileiras, como a comunidade quilombola de Jurussaca, no Pará. Essas comunidades têm chamado a atenção de pesquisadores como Figueiredo & Oliveira (2013) e Oliveira *et al.* (2015), que argumentam que o que se denomina português afro-indígena é uma variedade falada da língua em comunidades rurais que deve ser vista como parte do conjunto das variedades que formam o contínuo dialetal do PB, como o português afro-brasileiro (cf. LUCCHESI, BAXTER & RIBEIRO, 2009) e o português indígena.<sup>128</sup> Essa perspectiva de análise para a história linguística de Goiás encontra respaldo, por exemplo, na formação dos Tapuios do Carretão, uma tribo indígena remanescente de um aldeamento criado no século XVIII, na região nordeste do estado de Goiás, e composta étnica e historicamente por ameríndios, negros (africanos e afro-brasileiros) e brancos.

### 7.3 Síntese do capítulo

Neste capítulo, foram apresentadas algumas das perspectivas teóricas sobre o contato linguístico e seus resultados. Ressaltamos que a mudança linguística em situações de contato linguístico, como a mudança induzida por contato e as situações extremas de contato, atua de

<sup>127</sup> Em comunicação pessoal, Márcia Oliveira (2019) nos informou que a ordem predominante das palavras na língua geral paulista era SVO, o mesmo padrão da maioria dos falantes de línguas nigero-congolesas que viviam no Brasil naquele período.

<sup>128</sup> Oliveira *et al.* (2015) apresenta a seguinte definição para português afro-indígena: “Uma variedade vernacular do português brasileiro falada por comunidades de fala ‘aquilombadas’. Destacamos que tais comunidades não são apenas as “terras de preto”, mas ainda as “terras indígenas” e as “terras mistas” como se dão no norte do Brasil. Nestas sociedades, verificam-se, como traços marcantes: (i) nas “terras de preto”: uma descendência de africanos; (ii) nas “terras indígenas”: etnias que perderam ou estão por perder por completo suas línguas maternas, mas mantêm seus laços identitários; (iii) nas “terras mistas”: comunidades aquilombadas no norte do Brasil cuja formação étnica é negra e indígena. Nestes três tipos de “terras” o português é traço de identidade e nossa hipótese é que cada comunidade de fala traga suas marcas específicas devendo cada uma delas, na medida do possível, ser inventariada.” (OLIVEIRA *ET AL.* 2015, P. 155).

modo bem diferente na gramática da língua em comparação com a chamada mudança interna da língua: quando as línguas se desenvolvem em contato, geralmente saltos significativos de uma construção a outra podem ocorrer na gramática da língua como resultado da mudança. Esses saltos também podem provocar a instabilidade em outras partes da gramática. Por outro lado, quando as línguas se desenvolvem sozinhas (mudança interna), elas percorrem caminhos muito específicos que levam de um tipo de construção gramatical a outro.

Mostramos uma série de fenômenos linguísticos da gramática do PB que têm sido associados ao contato do português com as línguas bantas no período colonial. Apresentamos informações sobre as línguas africanas trasladadas para o Brasil e os dados demográficos sobre a população africana e afro-brasileira no território colonial brasileiro e em Goiás, principalmente nos séculos XVIII e XIX. Propusemos a hipótese de que dados de línguas-E produzidos principalmente por africanos e afro-brasileiros no período colonial serviram de base para a emergência, difusão e estabilização das estruturas gramaticais inovadoras do PB, mais especificamente a ordem SVO. Para sustentar nossa hipótese, consideramos os cinco procedimentos propostos por Thomason (2001) para estabelecer se uma mudança numa dada língua foi realmente induzida por contato linguístico, isto é: (i) não nos detivemos em um dado singular da língua-alvo, mas na estrutura gramatical da língua como um todo; (ii) identificamos, dentre as línguas nigero-congolesas, o quimbundo como a língua-fonte que possui as condições históricas mais favoráveis para ter influenciado a gramática do PB, assim como possui precisamente a propriedade estrutural inovadora verificada na gramática do PB no século XIX; (iii) encontramos como traços compartilhados entre a língua-fonte a língua-alvo a ordem SVO; (iv) mostramos por meio de dados empíricos que a ordem SVO não era predominante na gramática do português setecentista, ao mesmo tempo em que (v) mostramos por meio de registros históricos que a ordem SVO era a ordem preferível no quimbundo seiscentista. Assim, sugerimos que a perda da ordem VS do século XVIII para o século XIX nos dados históricos do PB em Goiás seria o resultado de uma influência de substrato de línguas nigero-congolesas, majoritariamente línguas com a estrutura SVO, mais precisamente o quimbundo. Argumentamos que o estabelecimento da ordem SVO na gramática do PB que emergiu no século XIX teria sido responsável pela perda do traço-D de T e, conseqüentemente, por todos os demais fenômenos relatados no Capítulo 6 como resultados da perda da inversão livre e do traço-D em T na gramática oitocentista do PB em Goiás, a exemplo principalmente da perda substancial de sujeitos nulos.

## 8 CONSIDERAÇÕES FINAIS

Nesta tese, analisamos dados originais de manuscritos goianos escritos no formato diário/memórias nos períodos colonial e imperial, respectivamente, e evidenciamos a perda de sujeitos nulos na gramática do PB em Goiás do século XVIII para o XIX, sem que isso tivesse como gatilho um aparente empobrecimento do paradigma verbal, contrariando os achados seminais de Duarte (1993 [2018], 2000). Com base em observações empíricas e em análises teóricas, argumentamos que o manuscrito *Memorial de lembrança*, escrito entre os séculos XIX e XX em Goiás por uma mulher goiana nascida em 1855, constitui a evidência mais antiga – encontrada até a finalização deste tese – do desenvolvimento do PB como uma LSN parcial.

Duas questões de pesquisa, apresentadas na Introdução, delinearam os caminhos percorridos no desenvolvimento desta tese. A primeira questão indagou quais seriam os mecanismos sintáticos envolvidos na distribuição de sujeitos nulos nas gramáticas setecentistas e oitocentistas que explicariam a mudança paramétrica atestada nas amostras do *corpus* da tese. A esse respeito, tecemos os seguintes argumentos: o aumento de sujeitos manifestos na amostra do século XIX em comparação com a amostra do século XVIII, cujos dados foram produzidos por um escriba de origem portuguesa e outro de origem desconhecida, foi explicado como uma exigência da gramática oitocentista de que um item lexical com um índice de definitude, mais precisamente um NP/DP lexical, ocupasse [SPEC-TP] para satisfazer o traço-EPP em uma língua que já não possuía um traço-D em T. Esse mecanismo de satisfação do traço-EPP era predominante na gramática da língua no século XIX. Adicionalmente, o estabelecimento da ordem SVO do século XVIII para o XIX foi explicado pela perda da inversão livre do sujeito. Na gramática setecentista, o traço-D em T (um EPP) era satisfeito pelo movimento de V-para-T; com esse movimento do verbo, o traço-D de T era valorado e o sujeito ficava em uma posição abaixo de T, produzindo a ordem VS. Com a perda do traço-D de T, no entanto, o sujeito das estruturas VS passou a se mover para [SPEC-TP] a fim de satisfazer o EPP nessa posição.

A análise linguística que desenvolvemos para a primeira questão de pesquisa da tese uniu principalmente as propostas teóricas de Holmberg (2005, 2010a) com as de Alexiadou & Anagnostopoulou (1998) para lidar com a sintaxe dos sujeitos nas amostras setecentista e oitocentista do PB em Goiás. Nesse sentido, argumentamos que, na gramática setecentista, os traços não especificados de T ( $[u\phi]$  e  $[uD]$ ) eram valorados por meio do movimento de V-para-T. Esse movimento licenciava a ordem VS e checava o EPP por meio da valoração do

traço-D de T. Essas operações, associadas ao movimento do verbo para F, permitiam também que a gramática setecentista do PB em Goiás gerasse construções com *se*-passivo. Vimos que a perda do traço-D em T na gramática oitocentista do PB em Goiás, além de passar a exigir que um item lexical com índice de definitude ocupasse [SPEC-TP], eliminou as chamadas construções com *se*-passivo da gramática da língua e permitiu o licenciamento de sujeitos nulos indefinidos na 3ª pessoa do singular.

Como mencionamos, uma diferença crucial dos achados desta tese com outros trabalhos desenvolvidos sobre a mudança paramétrica na sintaxe do sujeito no PB consiste no fato de que não atribuímos um papel determinante na morfologia de concordância à perda da possibilidade de licenciar sujeitos nulos no PB. Com base em resultados estatísticos e seguindo o postulado de Roberts (2010, p. 82) de que “se uma categoria  $\alpha$  tem um traço-D, então todos os traços- $\phi$  de  $\alpha$  são especificados”, sugerimos que a mudança no paradigma flexional do PB – principalmente se foi desencadeada no século 20 – seria, na verdade, resultado de uma mudança relacionada à perda do traço-D em T, tendo em vista que um corolário do postulado de Roberts (2010) é o de que a riqueza morfológica e a valoração dos traços- $\phi$  numa dada LSN consistente são dependentes da presença de um traço-D no núcleo funcional T dessa língua. Assim, o empobrecimento atestado na morfologia de concordância no PB seria um resultado da perda do traço-D de T, ao invés de um gatilho primário para ela. Negrão & Viotti (2000), mesmo adotando uma perspectiva diferente, têm um posicionamento similar ao nosso: as autoras afirmam que propriedades discursivas minaram a função que era desempenhada pelo paradigma flexional dos verbos finitos no PB; assim, na proposta dessas autoras, o empobrecimento da concordância também não estaria diretamente relacionado com o preenchimento mais sistemático da posição de sujeito no PB.

Nossa segunda questão de pesquisa indaga sobre a possibilidade de propormos um modelo teórico de análise linguística que pudesse indicar o papel do contato linguístico no desenvolvimento do PB em Goiás como uma LSN parcial. Nesse sentido, apresentamos um panorama sociolinguístico de Goiás nos séculos XVIII e XIX para avaliar os possíveis efeitos do contato linguístico no declínio de sujeitos nulos na gramática oitocentista do português goiano. Fundamentados no conceito de descontinuidade de Lightfoot (2017), que prevê descontinuidades na aquisição da gramática de uma língua de uma geração para a outra, e nos procedimentos analíticos de Thomason (2001) para estabelecer se uma determinada mudança foi induzida por contato, tomamos como hipótese que múltiplos dados de línguas-E, produzidos especialmente por africanos e afro-brasileiros no período colonial, serviram de base para a emergência, a difusão e a estabilização das principais estruturas gramaticais

caracterizadoras do PB. Sugerimos assim que a perda da ordem VS do século XVIII para o século XIX seria o resultado de uma influência de substrato de línguas nigero-congolesas, majoritariamente línguas com a estrutura SVO, como o quimbundo. Nessa perspectiva, o estabelecimento da ordem SVO na gramática oitocentista do PB em Goiás, em comparação com a ordem VS na gramática setecentista do PB em Goiás, seria o resultado de uma influência de substrato na gramática do PB.

Destacamos que tanto o estabelecimento da ordem SVO quanto as restrições no licenciamento de sujeitos nulos na gramática do PB oitocentista em Goiás dizem respeito a mesoparâmetros. Os mesoparâmetros afetam categorias sintáticas individuais (cf. BIBERAUER & ROBERTS, 2017). Assim, a ordem SVO é um mesoparâmetro porque envolve propriedades formais dos núcleos funcionais, especialmente [SPEC-TP], que determinam a posição sintática dos constituintes de uma sentença. O parâmetro do sujeito nulo, por sua vez, envolve as propriedades formais de T finito e de Ds pronominais. Para Biberauer & Roberts (2017), os mesoparâmetros são mais suscetíveis a mudanças em situações de contato linguístico. Os autores citam como exemplo nesse sentido as variedades românicas do norte-europeu, como o francês e os dialetos do norte da Itália, que perderam diacronicamente a possibilidade de licenciar sujeitos nulos por conta do intenso contato com línguas germânicas. Eles mencionam o PB (coloquial) como uma variedade de língua românica que perdeu diacronicamente propriedades de uma LSN consistente.

Nesta tese, atribuímos o estabelecimento da ordem SVO na gramática oitocentista do PB ao contato linguístico do português com línguas nigero-congolesas que se instaurou na América portuguesa. Por outro lado, não estabelecemos uma correlação forte entre morfologia manifesta e sintaxe no estabelecimento da gramática de sujeito nulo parcial no PB oitocentista de Goiás. Em linhas gerais, propusemos que as mudanças encontradas no século XIX são o resultado de uma combinação de fatores sociolinguísticos e psicolinguísticos: primeiramente, uma tendência por parte principalmente de falantes de línguas nigero-congolesas, como o quimbundo, a adquirir e transferir estruturas similares às de suas línguas maternas para o português que adquiriam como segunda língua (L2) no período colonial, como, por exemplo, orações com ordem SVO. Conseqüentemente, houve um declínio da inversão livre do sujeito nos dados de língua-E produzidos por falantes não nativos de português. A predominância da ordem SVO e o declínio da inversão livre do sujeito no século XIX teriam impulsionado novos adquirentes do português como língua materna a não adquirir o elemento categorial D presente no núcleo funcional T da gramática setecentista do PB, favorecendo a mudança na expressão da sintaxe do sujeito que constatamos na gramática oitocentista do PB em Goiás.

## REFERÊNCIAS

- ALENCASTRO, L. F. *O trato dos viventes: formação do Brasil no Atlântico Sul*. São Paulo: Companhia das Letras, 2000.
- ALEXIADOU, A.; ANAGNOSTOPOULOU, E. *Natural Language & Linguistic Theory*, 16: 491, 1998. <https://doi.org/10.1023/A:1006090432389>.
- ANDRADE, A. Tópicos-sujeito locativos no português brasileiro: análise de hipóteses sintáticas e semântico-pragmáticas. *Letras Escreve*, v. 7, p. 11, 2018.
- ANDRADE, A; GALVES, C. A Unified Analysis for Subject Topics in Brazilian Portuguese. *Journal of Portuguese Linguistics*, v. 13, p. 117-147, 2014.
- AVELAR, J. O. Sobre a emergência das construções de tópico-sujeito no português brasileiro: mudança desencadeada por contato? In: J. Avelar; L. Álvarez-López (eds). *Dinâmicas Afro-latinas: Língua(s) e História(s)*. Frankfurt am Main: Peter Lang, 127-148, 2015.
- AVELAR, J. O.; GALVES, C. From European to Brazilian Portuguese: A parameter tree approach. *Cadernos de Estudos Linguísticos*, Campinas, SP, v. 58, n. 2, p. 237-256, set. 2016. ISSN 2447-0686.
- AVELAR, J. O.; GALVES, C. From European to Brazilian Portuguese: a parameter tree approach. *Cadernos de Estudos Linguísticos*, 58(2), 237-256, 2016.
- BAKER, M. 2008. The macroparameter in a microparametric world. In: BIBERAUER, T. (Org.). *The limits of syntactic variation*. Amsterdam: John Benjamins, 2008, p. 351–74.
- BAPTISTA, M. *Pidgins and Creoles: Syntax* (Oxford Handbooks Online). Oxford University Press, 2017. DOI: 10.1093/oxfordhb/9780199935345.013.13.
- BAPTISTA, M. *The Syntax of Cape Verdean Creole, the Sotavento varieties*. Amsterdam/Philadelphia: John Benjamins, 2002.
- BARBOSA, A. Tradição discursiva e tratamento de *corpora* históricos: desafios metodológicos para o estudo da formação do português brasileiro. In: LOBO, T.; CARNEIRO, Z.; SOLEDADE, J.; ALMEIDA, A.; RIBEIRO, S. (Org.). *ROSAE: Linguística histórica, história da língua e outras histórias*. Salvador: EDUFBA, p. 589-606, 2012.
- BARBOSA, P. *Null subjects*. PhD dissertation: MIT, 1995.
- BARBOSA, P. *pro* as a Minimal nP: Towards a Unified Approach to *pro*-drop. *Linguistic Inquiry*, manuscrito (just accepted), 1-78, 2018.
- BARBOSA, P. Two kinds of subject *pro*. *Studia Linguistica* 63: 2–58, 2009.
- BAXTER, A.; LUCCHESI, D. A relevância dos processos de pidginização e criouliização na formação da língua portuguesa no Brasil. *Estudos Linguísticos e Literários*, Salvador - Bahia, v. 19, p. 65-83, 1997.
- BERLINCK, R. A. Brazilian Portuguese VS Order: a diachronic analysis. In: Kato, M.; Negrão, E. (Eds.). *Brazilian Portuguese and the Null Subject Parameter*. Madrid, Iberoamericana, pp. 125-194, 2000.

BERTRAN, P. Notícia geral da capitania de Goiás em 1783. Goiânia: Universidade Católica de Goiás, Universidade Federal de Goiás; Solo Editores, 1997.

BERWICK, R. C.; CHOMSKY, N. The biolinguistic programa. The current state of its development. In: A. M. Di Sciullo and C. Boeckx (Eds.). *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*. Oxford University Press, 2011, p. 19-41.

BERWICK, R. C.; CHOMSKY, N. *Why only us*. Cambridge, MA: The MIT Press, 2016.

BHAT, D. N. S. *Pronouns*. Oxford: Oxford University Press, 2004.

BIBERAUER, T. Semi null-subject languages, expletives and expletive *pro* reconsidered. In: *Parametric Variation: Null Subjects in Minimalist Theory*, 153–99. Cambridge: Cambridge University Press, 2009. DOI:10.1017/CBO9780511770784.005.

BIBERAUER, T.; ROBERTS, I. Parameter Setting. In: *The Cambridge Handbook of Historical Syntax*. Ed. por Adam Ledgeway e Ian Roberts. Cambridge Handbooks in Language and Linguistics. Cambridge: Cambridge University Press, 2017, p. 134–62. DOI: 10.1017/9781107279070.008.

BOBALJIK, J. D. Realizing Germanic inflection: Why morphology does not drive syntax. *Journal of Comparative Germanic Linguistics*, v. 6, p. 129–167, 2002.

BOBALJIK, J. D.; THRÁINSSON, H. Two heads aren't always better than one. *Syntax*, v.1, 1998, p. 37-71.

BOECKX, C. Approaching parameters from below. In: A. M. Di Sciullo and C. Boeckx (Eds.). *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*. Oxford University Press, 2011, p. 205–221.

BONET, E. *Morphology after syntax: Pronominal clitics in Romance*. Tese (Doutorado). Cambridge MA, MIT, 1991.

BONVINI, E. Línguas africanas e português falado no Brasil. FIORIN, J. L.; PETTER, M. T. (Orgs.). *África no Brasil – a formação da língua portuguesa*. 1ª ed. [2ª reimpressão]. São Paulo: Contexto, 2009, p. 15-62.

BORER, H. *Parametric syntax: Case studies in Semitic and Romance languages*. Dordrecht: Foris, 1984.

BORGES, D. V. C. *Construções causativas no português do Centro-Oeste nos séculos XVIII-XIX e no português atual*. 2008. 122 f. Dissertação (mestrado) - Universidade de Brasília, Instituto de Letras, Departamento de Linguística, Português e Línguas Clássicas, 2008.

BORGES, D. V. C.; SALLES, H. M. M. L.; PILATI, E. O projeto O Centro-Oeste na história do Português Brasileiro: subsídios da sócio-história ao estudo da mudança linguística e hipóteses de trabalho. CARDOSO, C.; SCHERRE, M. M.; SALLES, H. M. M. M. (Org.). *Variação linguística: contato de línguas e educação*. Belo Horizonte: Ed. Pontes, 2012, v. 1, p. 219-237.

BORGES, H. *Sujeitos nulos indeterminados no português brasileiro: uma investigação diacrônica em Goiás*. (Dissertação de mestrado). Programa de Pós-Graduação em Linguística da Universidade de Brasília, 2014.



- BORGES, H.; PIRES, A. The emergence of Brazilian Portuguese: Earlier evidence for the development of a partial null subject grammar. *Proceedings of the Linguistic Society of America*, v. 2, p. 31:1-15, June 2017.
- BRETAS, G. F. *História da instrução pública em Goiás*. Goiânia: CEGRAF/UFG, 1991.
- BRITO, A. M.; DUARTE, I.; MATOS, G. Tipologia e distribuição das expressões nominais. In: Mateus, M. H., A. M. Brito, I. Duarte, H. Faria, S. Frota, G. Matos, F. Oliveira, M. Vigário & A. Villalva (eds.). *Gramática da Língua Portuguesa*. Lisboa: Caminho, p. 826 - 848., 2003.
- BURZIO, L. *Italian Syntax. A government-binding approach*. Dordrecht: Reidel Publish Company, 1986.
- CAMACHO, J. *Null subjects*. Cambridge University Press, 2013.
- CAMBRAIA, C. N. *Introdução à crítica textual*. São Paulo: Martins Fontes, 2005.
- CARDINALETTI, A. 1990. *Impersonal constructions and sentential arguments in German*. Padua: Unipress.
- CARDINALETTI, A.; STARKE, M. 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In Henk van Riemsdijk (ed.). *Clitics in the languages of Europe*. New York: Mouton de Gruyter, 145-233.
- CARVALHO, E. F. *O rosário de Aninha: os sentidos da devoção rosarina na escritura de Anna Joaquina Marques (Cidade de Goiás, 1881-1930)*. 2008. 285 f. Dissertação (Mestrado) – Universidade Federal de Goiás, Faculdade de Ciências Humanas e Filosofia, 2008.
- CARVALHO, E. F. Retalhos de uma experiência feminina: mestra Lili, professora pública da capital goiana (1858-1945). *Temporis(ação)*, v. 12, n. 1, p. 18 - 39, jan./dez., 2012
- CASTILHO, A.; TORRES MORAIS, M. A.; LOPES, R. E. V.; CYRINO, S. M. L. (Org.). *Descrição, história e aquisição do português brasileiro*. São Paulo: FAPESP, Campinas: Pontes Editores, 2007.
- CAVALCANTE, S. R. Mudança e estabilidade em construções “passivas” na história do português. *Linguística*, Montevideo, v.32, n. 1, p. 41-60, 2016. <http://dx.doi.org/10.5935/2079-312X.20160003>.
- CAVALCANTE, S. R. *O uso de se com infinitivo na História do Português: Do Português Clássico ao Português Europeu e Brasileiro Moderno*. 2006. 206f. Instituto de Estudos da Linguagem. Universidade Estadual de Campinas, Campinas.
- CAVALCANTE, S. R.; MARCOTULIO, L. L. Novo olhar sobre as construções com SE: para além da questão da concordância. In: M. E. L. Duarte. (Org.). *O sujeito em peças de teatro (1833-1992): estudos diacrônicos*. São Paulo: Parábola, p. 143-160, 2012.
- CHAIM, M. M. *Os aldeamentos indígenas na capitania de Goiás: sua importância na política do povoamento (1749-1811)*. Goiânia: Oriente, 1974.
- CHOMSKY, N. *Lectures on government and binding theory*. Dordrecht: Foris, 1981.
- CHOMSKY, N. *Barriers*. Cambridge, MA: MIT Press, 1986b.

- CHOMSKY, N. Beyond explanatory adequacy. In: A. Belletti (Ed.). *Structures and beyond: The cartography of syntactic structures*. Oxford University Press, 2004, p. 104–113.
- CHOMSKY, N. Derivation by phase. In: *Ken Hale: A Life in Language*. Ed.: Michael Kenstowicz. Cambridge MA: The MIT Press, 2001, p. 1–52.
- CHOMSKY, N. *Knowledge of language: its nature, origin, and use*. New York: Praeger, 1986a.
- CHOMSKY, N. Minimalist inquires: the framework. In: MARTIN, R.; MICHAEL, D. URIAGEREKA, J. (Eds.). *Step-by-step: essays in minimalist syntax in honor of Howard Lasnik*. Cambridge, MA: The Mit Press, 2000. p. 89-155.
- CHOMSKY, N. *Novos horizontes no estudo da linguagem e da mente*. São Paulo: Editora Unesp, 2005b.
- CHOMSKY, N. On phases. In: R. Freidin, C. Otero and M. L. Zubizarreta (eds.) *Foundational issues in linguistic theory*. Cambridge (MA): MIT Press, pp. 133-166, 2008.
- CHOMSKY, N. Problems of projection extensions. In: *Structures, Strategies and Beyond: Studies in Honour of Adriana Belletti*. Ed.: Elisa Di Domenico, Cornelia Hamann e Simona Matteini. Amsterdam: John Benjamins, 2015, p. 1-16.
- CHOMSKY, N. Problems of projection. *Lingua*, v. 130, 2013, p. 33-49.
- CHOMSKY, N. *Some concepts and consequences of the theory of government and binding*. Cambridge, Mass.: The MIT Press, 1982.
- CHOMSKY, N. *The Minimalist Program*. Cambridge, Mass.: MIT Press, 1995.
- CHOMSKY, N. Three factors in language design. *Linguistic Inquiry*, v. 36:1, p. 1-22, 2005a.
- CHOMSKY, N; LASNIK, H. Filters and control. *Linguistic Inquiry*, v. 8, n. 3, 425-504, 1977.
- CINQUE, G. *Adverbs and functional heads: A cross-linguistic perspective*. New York: Oxford University Press, 1999.
- COSTA, J.; FIGUEIREDO SILVA, M. C. 2006. Notas sobre a concordância verbal e nominal em português. *Estudos Linguísticos*, XXXV, p. 95-109, 2006.
- COSTA, J.; GALVES, C. External Subjects in Two Varieties of Portuguese: Evidence for a Non-Unified Analysis. In: Beyssade, C. et al. (Eds.). *Proceedings of Going Romance 2000*. Amsterdam/Philadelphia, John Benjamins, 2002.
- COUTO, H. H. A concordância e a função comunicativa da linguagem: uma visão ecolinguística. *Confluência* (Rio de Janeiro), n. 46, 2014.
- COUTO, H. H. Hipótese da relexificação na gênese dos crioulos e pidgins. *Revista da ABRALIN*, vol. 1, no 1, p. 221-250, julho 2002.
- COUTO, H. H. *Linguística, ecologia e ecolinguística*. São Paulo: Editora Contexto, 2009.
- CYRINO, S. M. L.; DUARTE, M. E. L. & KATO, M. A. Visible subjects and invisible clitics in Brazilian Portuguese. In: Kato, M. A. & Negrão, E. V. (Eds.) *Brazilian Portuguese and the Null Subject Parameter*. Frankfurt: Vervuert-Iberoamericana, 2000, p. 55-104.

CYRINO, S.; TORRES MORAIS, M. A. *História do português brasileiro: mudança sintática do português brasileiro: perspectiva gerativista*. (Coordenador geral: Ataliba Castilho). São Paulo: Contexto, 2018.

DOBROVIE-SORIN, C. Impersonal se Constructions in Romance and the Passivization of Unergatives. *Linguistic Inquiry*, v. 29, n. 3, p. 399-437, 1998. DOI: 10.1162/002438998553806.

DUARTE, I. Construções Ativas, Passivas, Incoativas e Médias. In: Paiva Raposo et al. (eds.). *Gramática do Português*, vol I: p. 429-458. Lisboa: Fundação Calouste Gulbenkian, 2013.

DUARTE, I.; FIGUEIREDO SILVA, M. C. The Null Subject Parameter and the Structure of the Sentence in European and Brazilian Portuguese. In: Wetzels, W.L.; J. Costa; S. Menuzzi (eds.) *The Handbook of Portuguese Linguistics*. Sussex: John Wiley & Sons, Inc., p. 234-53, 2016.

DUARTE, M. E. L. (Org.). *O sujeito em peças de teatro (1983-1992): estudo diacrônicos*. São Paulo: Parábola, 2012.

DUARTE, M. E. L. *A perda do princípio “evite pronome” no português brasileiro*. 141 f. Tese (Doutorado). Instituto de Estudos da Linguagem. Universidade Estadual de Campinas, 1995.

DUARTE, M. E. L. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. ROBERTS, I.; KATO, M. (Org.). *Português brasileiro: uma viagem diacrônica*. Campinas: Editora da Unicamp, 1993 [2018].

DUARTE, M. E. L. The Loss of the ‘Avoid Pronoun’ Principle in Brazilian Portuguese. In: Kato, M.; Negrão, E. (Eds.) *Brazilian Portuguese and the Null Subject Parameter*. Frankfurt am Main, Vervuert, pp. 17-36, 2000.

DUARTE, M. E. L.; MOURÃO, G.; SANTOS, H. (2012). Os sujeitos de terceira pessoa: revisitando Duarte 1993. In: DUARTE, M. E. L. (Org.). *O sujeito em peças de teatro (1983-1992): estudo diacrônicos*. São Paulo: Parábola, 2012, p. 21-44.

EMBICK, D. Linearization and local dislocation: Derivational mechanics and interactions. *Linguistic Analysis*, v. 33, p. 2–35, 2007.

EMONDS, J. The Verbal Complex V'-V in French. *Linguistic Inquiry*, 9 (2), pp. 151-175, 1978.

EPSTEIN, S.; GROAT, E. M.; KAWASHIMA, R.; KITAHARA, H. *A derivational approach to syntactic relations*. Oxford: Oxford University Press, 1998.

EPSTEIN, S.; SEELY, D. *Derivations in Minimalism*. Cambridge Studies in Linguistics. Cambridge: Cambridge University Press, 2006. DOI: 10.1017/CBO9780511550607.

FAUSTO, B. *História concisa do Brasil*. São Paulo: EdUSP, 2011.

FERREIRA, M. B. *Argumentos nulos em português brasileiro*. Dissertação de Mestrado. Campinas, Unicamp, 2000.

FERREIRA, M. B.; CORREA, C. N. The semantic of DPs. In W. L. Wetzels, J. Costa and S. Menuzzi (eds.). *The Handbook of Portuguese Linguistics*, Wiley-Blackwell, p. 356-373, 2016.

- FIGUEIREDO SILVA, M. C. *A posição de sujeito em português brasileiro: frases finitas e infinitivas*. Campinas: Editora da Unicamp, 1996.
- FIGUEIREDO SILVA, M. C. Main and embedded null subjects in Brazilian Portuguese. In: NEGRÃO, E. & KATO, M. *Brazilian Portuguese and the Null Subject Parameter*. Editora: VervuetIberoamericana. 2000. p.127-145.
- FIGUEIREDO SILVA, M. C. O estado da arte dos estudos sobre sujeitos nulos, posições de sujeito e marcas flexionais. *Revista Linguística*, v. 13, p. 1-21, 2017.
- FIGUEIREDO, C. G.; OLIVEIRA, M. S. D. Português do município do Libolo, Angola, e português afro-indígena da comunidade de Jussara, Brasil: cotejando os sistemas de pronominalização. *Papia*, v. 23 (2), p. 1-87, 2013.
- FIORIN, J. L.; PETTER, M. (Org.). *África no Brasil: a formação da língua portuguesa*. São Paulo: Contexto, 1ª ed. [2ª reimpressão]. São Paulo: Contexto, 2009.
- FISCHER, O.; VAN KEMENADE, A.; KOOPMAN, W.; VAN DER WURFF, W. *The syntax of early English*. Cambridge University Press. 2000.
- FLORENTINO, M. *Em costas negras: uma história do tráfico de escravos entre a África e o Rio de Janeiro (séculos XVIII e XIX)*. São Paulo: Companhia das Letras, 1997.
- FUß, E. Language Change. In: *The Oxford Handbook of Universal Grammar*. Oxford University Press, 2016, p. 12-22. DOI: 10.1093/oxfordhb/9780199573776.013.19.
- GALVES, C. A sintaxe do português brasileiro. *Cadernos de Linguística e Teoria da Literatura*, n. 13, p. 33-52, 1985.
- GALVES, C. *Ensaio sobre as gramáticas do português*. Campinas: Editora da Unicamp, 2001.
- GALVES, C. M. C. Pronomes e categorias vazias em Português do Brasil. *Cadernos de Estudos Linguísticos (UNICAMP)*, Campinas, v. 7, p. 107-136, 1984.
- GALVES, C. Tópicos, sujeitos, pronomes e concordância no português brasileiro. *Caderno de Estudos da Linguagem*, v. 34, p. 19-32, Campinhas, jan./jun., 1998.
- GALVES, C.; CYRINO, S. C.; LOPES, R.; SANDALO, F.; AVELAR, J. (Eds.). *Parameter Theory and Linguistic Change*. Oxford: Oxford University Press, 2012. Oxford Scholarship Online, 2013. DOI: 10.1093/acprof:oso/9780199659203.001.0001.
- GALVES, C.; ANDRADE, A; FARIA, P. *Corpus Histórico do Português Tycho Brahe (fase IV)*. 2017.
- GILLIGAN, G. M. *A cross linguistic approach to the pro-drop parameter*. Tese (Doutorado em Linguística). University of Southern California: 1987.
- GRAVINA, A. P. *Sujeito nulo e ordem VS na história do português brasileiro: um estudo diacrônico-comparativo baseado em Corpus*. Tese (Doutorado). 2014.
- GRIES, S. Basic significant testing. In: *Research methods in Linguistics*. Ed. Robert Podesva e Devyani Sharma. Cambridge: Cambridge University Press, 2014, p. 316-336. DOI: 10.1017/CBO9781139013734.017.

- GRIES, S.; NEWMAN, J. Creating and using corpora. In: *Research methods in Linguistics*. Ed. Robert Podesva e Devyani Sharma. Cambridge: Cambridge University Press, 2014, p. 257-287. DOI:10.1017/CBO9781139013734.015.
- GUY, G. *Linguistic variation in Brazilian Portuguese: aspects of phonology, syntax and language history*. Unpublished PhD Dissertation University of Pennsylvania. 1981.
- HAEBERLI, E.; PINTZUK, S. Verb (projection) raising in Old English. In: D. Jonas, J. Whitman and A. Garrett (eds.). *Grammatical change: Origins, nature, outcomes*. Oxford University Press, pp. 219–38, 2011.
- HAEGEMAN, L. Understood subjects in English diaries, *Multilingua*, 9, p. 157-159, 1990.
- HAERBELI, E. Syntactic effects of inflectional morphology and competing grammars. In: E. Fuß and C. Trips (Eds.). *Diachronic Clues to Synchronic Grammar*. Amsterdam: John Benjamins, p. 101-130, 2004.
- HALLE, M., & MARANTZ, A. Distributed morphology and the pieces of inflection. In K. Hale, & S. J. Keyser (Eds.). *The view from building 20*. Cambridge, MA: The MIT Press, p. 111-176, 1993.
- HEYCOCK, C.; SUNDQUIST, J. D. Don't rush to rehabilitate: A remark on Koenenman & Zeijlstra 2014. *Linguistic Inquiry*, 48: 173-179, 2017.
- HEYWOOD, L. M. Portuguese into African: The Eighteen-Century Central African Background to Atlantic Creole Cultures. In: *Central Africans and cultural transformations in the American diaspora* [ed.Linda M. Heywood]. Cambridge University Press: p. 91-113, 2002.
- HOLM, J. The genesis of the Brazilian Vernacular: insights from the indigenization of Portuguese in Angola. *Papia* 19, p. 93-122, 2009.
- HOLMBERG, A. & NIKANNE, U. Expletives, subjects, and topics in Finnish. In: P. Svenonius (Ed.) *Subjects, expletives, and the EPP*. Oxford: Oxford University Press, p. 71-106, 2002.
- HOLMBERG, A. 2005. Is there a little pro? Evidence from Finnish. *Linguistic Inquiry* 36: 533–64.
- HOLMBERG, A. Null Subject Parameters. In: BIBERAUER, T. et al. *Parametric Variation: null subjects in Minimalist Theory*. Cambridge: Cambridge University Press, p. 88-124, 2010a. DOI:10.1017/CBO9780511770784.003.
- HOLMBERG, A. The null generic subject pronoun in Finnish: a case of incorporation in T. In: BIBERAUER, T. et al. *Parametric Variation: null subjects in Minimalist Theory*. Cambridge: Cambridge University Press, p. 200-231, 2010b. DOI:10.1017/CBO9780511770784.006.
- HOLMBERG, A. Universal Grammar. In: *The Cambridge Handbook of Historical Syntax*. Ed.: Adam Ledgeway e Ian Roberts. Cambridge: Cambridge University Press, 2017, p. 275-300. DOI: 10.1017/9781107279070.014.
- HOLMBERG, A.; M. SHEEHAN. Control into finite clauses in partial null-subject languages. BIBERAUER, T. et al. *Parametric Variation: null subjects in Minimalist Theory*. Cambridge: Cambridge University Press, 2010.

- HOLMBERG, A.; NAYUDU, A.; SHEEHAN, M. Three Partial null subject languages: a comparison of Brazilian Portuguese, Finnish and Marathi. *Studia Linguistica*, 63: p. 59-97, 2009.
- HOLMBERG, A.; ROBERTS, I. The syntax–morphology relation. *Lingua*, v. 130, p. 111-131, 2013.
- HORNSTEIN, N. Movement and Control. *Linguistic Inquiry*, Volume 30, Number 1, p. 69–96, 1999.
- HORNSTEIN, N.; NUNES, J.; GROHMANN, K. *Understanding Minimalism*. Cambridge: Cambridge University Press, 2005. DOI: 10.1017/CBO9780511840678.
- HUANG, C.-T. J. *Anaphora: A cross-linguistic study*. Oxford: Oxford University Press. 2000.
- HUANG, C.-T. J. On the distribution and reference of empty pronouns. *Linguistic Theory*, v.15, p. 531-573, 1984.
- HUANG, C.-T. J.; ROBERTS, IAN. Principles and Parameters of Universal Grammar. In: *The Oxford Handbook of Universal Grammar*. Oxford University Press, 2016. DOI: 10.1093/oxfordhb/9780199573776.013.14.
- JAEGGLI, O.; SAFIR, K. The null subject parameter and parametric theory. In. O. Jaeggli & K. J. Safir (Eds.). *The null subject parameter*. Dordrecht: Kluwer, 1989.
- JOHNSON, D. Descriptive statistics. In: *Research methods in Linguistics* (ed. By Robert Podesva & Devyani Sharma). Cambridge University Press. 2014.
- KARASCH, M. Centro-africanos no Brasil Central, de 1780 a 1835. *Diáspora negra no Brasil*. 2ª ed. [3ª reimpressão]. São Paulo: Contexto, 2018.
- KARASH, M. C. *Before Brasília: Frontier Life in Central Brazil*. Albuquerque: University of New Mexico Press, 2016
- KATO, M. A. Strong and weak pronouns in the null subject parameter. *PROBUS*, v. 11(1), p. 1- 38. 1999.
- KATO, M. A.; NASCIMENTO, M. (Org.). *Gramática do Português Culto Falado no Brasil*. A construção da sentença. São Paulo: Contexto, 2015.
- KATO, M. A.; NEGRÃO, E. V. (Org.) *Brazilian Portuguese and the Null Subject Parameter*. Frankfurt: Vervuert-Ibero Americana, 2000.
- KATO, M.; DUARTE, M. E. L. Restrições na distribuição de sujeitos nulos no Português Brasileiro. *Veredas (UFJF. Online)*, v. 18, p. 1-21, 2014.
- KATO, M.; ORDÓÑEZ, F. (Org.). *The morphosyntax of Spanish and Portuguese in Latin America*, p. 49-77, 2016.
- KAYNE, R. Extensions of binding and Case marking. *Linguistic Inquiry*, 11, 75-96, p. 1980.
- KAYNE, R. *French Syntax: the transformational cycle*. Cambridge, Mass, MIT Press., 1975.
- KAYNE, R. S. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press, 1994.

- KEENAN, E. Explaining the creation of reflexive pronouns in English. In: D. Minkova and R. Stockwell (Eds.). *Studies in the history of English: A millennial perspective*. Berlin: Mouton de Gruyter, pp. 325–55, 2002.
- KIPARSKY, P. “Elsewhere” in phonology. In: *A festschrift for Morris Halle*. Ed.: Stephen Anderson e Paul Kiparsky. New York: Holt, Rinehart and Winston, p. 93–106, 1973.
- KOENEMAN, O. *The flexible nature of verb movement*. Tese (Doutorado em Linguística), Utrecht University. Utrecht: LOT Publications, 2000.
- KOENEMAN, O; ZEIJLSTRA, Hedde. The Rich Agreement Hypothesis rehabilitated. *Linguistic Inquiry*, 45:571–615, 2014.
- KOOPMAN, H.; SPORTICHE, D. The Position of Subjects. *Lingua* 85, 1991, pp. 211-258.
- KROCH, A. Modeling language change and language acquisition. Manuscrito. 2005.
- KROCH, A. Reflexes of grammar in patterns of change. *Language Variation and Change*, 1989 p. 199-244.
- KROCH, A. Syntactic change. BALTIN, M.; COLLINS, C. (Orgs.). *The handbook of contemporary syntactic theory*. Oxford: Blackwell, 2001.
- KUTEVA, T. Contact and Borrowing. In: *The Cambridge Handbook of Historical Syntax*, ed. Adam Ledgeway e Ian Roberts, 163–86. Cambridge: Cambridge University Press, 2017. DOI:10.1017/9781107279070.009.
- LANDAU, I. The scale of finiteness and the calculus of control. *Natural Language and Linguistic Theory*, v. 22, p. 811–77, 2004.
- LEDGEWAY, A. *Grammatica diacronica del napoletano*. Tübingen: Max Niemeyer Verlag, 2009.
- LEDGEWAY, A.; ROBERTS, I. (Eds.) *The Cambridge Handbook of Historical Syntax*. Cambridge: Cambridge University Press, 2017. DOI: [10.1017/9781107279070](https://doi.org/10.1017/9781107279070).
- LI, C.; THOMPSON, S. Subject and topic: a new typology of language. In C. N. Li. (Org). *Subject and Topic*. New York: Academic Press Inc, 1976.
- LIGHTFOOT, D. W. Acquisition and Learnability. In: *The Cambridge Handbook of Historical Syntax*. Ed.: Adam Ledgeway e Ian Roberts. Cambridge: Cambridge University Press, 2017b, p. 381-400. DOI: [10.1017/9781107279070.019](https://doi.org/10.1017/9781107279070.019).
- LIGHTFOOT, D. W. Imperfect Transmission and Discontinuity. In: *The Cambridge Handbook of Historical Syntax*. Ed.: Adam Ledgeway e Ian Roberts. Cambridge: Cambridge University Press, 2017c, p 515-533. DOI: [10.1017/9781107279070.025](https://doi.org/10.1017/9781107279070.025).
- LIGHTFOOT, D. W. Transparency. In: *The Cambridge Handbook of Historical Syntax*, Ed.: Adam Ledgeway e Ian Roberts. Cambridge: Cambridge University Press, 2017a, p. 322-337. DOI: [10.1017/9781107279070.016](https://doi.org/10.1017/9781107279070.016).
- LIGHTFOOT, D. *How New Languages Emerge*. Cambridge: Cambridge University Press, 2006. DOI: [10.1017/CBO9780511616204](https://doi.org/10.1017/CBO9780511616204).
- LIGHTFOOT, D. *How to set parameters: Arguments from language change*. Cambridge, MA: MIT Press, 1991.

- LIGHTFOOT, D. *Principles of diachronic syntax*. Cambridge, UK: Cambridge University Press, 1979.
- LIGHTFOOT, D. *The development of language: Acquisition, change and evolution*. Oxford: Blackwell, 1999.
- LIMA-SALLES, H. M. M. L. A. Diversidade e mudança linguística na perspectiva da gramática gerativa. In: Denize Elena Garcia da Silva. (Org.). *Língua, Gramática e Discurso*. Goiânia: Editorial Cãnone, p. 124-132, 2006.
- LOBATO, L. Sobre a questão da influência ameríndia na formação do português do Brasil. SILVA, D. E. G. (Org.). *Língua, gramática e discurso*. Goiânia: Cãnone Editorial, 2006, p. 54-86.
- LOBATO, L. Sobre a questão da influência ameríndia na formação do português do Brasil. SILVA, D. E. G. (Org.). *Língua, gramática e discurso*. Goiânia: Cãnone Editorial, 2006, p. 54-86.
- LONGOBARDI, G. Formal syntax, diachronic minimalism, and etymology: The history of French *chez*. *Linguistic Inquiry*, 32(2), 2001, p. 275–302.
- LOPES, C. R. S.; CAVALCANTE, S. R. O. A cronologia do voçamento no português brasileiro: expansão de você-sujeito e retenção do clítico-*te*. *Revista Linguística*, Santiago de Chile, v.25, p.30-65, jun. 2011.
- LOPES, C. R. S.; DUARTE, M. E. L. “De “Vossa Mercê” a “Você”: análise da pronominalização de nominais em peças brasileiras e portuguesas setecentistas e oitocentistas”. In: BRANDÃO, S. F. & MOTA, M. A. (orgs.). *Análise contrastiva de variedades do português: primeiros estudos*. Rio de Janeiro: In-Fólio, 2003.
- LOPES, C. R. S.; RUMEU; M. C. de B.; CARNEIRO, Z de O. N. A configuração diatópico-diacrônica do sistema de tratamento do português brasileiro. *Revista do GELNE*, Natal, v. 15, n. 1/2, p. 163-186, 2013.
- LUCCHESI, D. A simplificação morfológica na expressão do sujeito indeterminado no português afro-brasileiro. *Revista Linguística*, v. 10, n. 1, junho de 2014, p. 277-298.
- LUCCHESI, D. *Língua e sociedade partidas: a polarização sociolinguística do Brasil*. São Paulo: Contexto, 2015.
- LUCCHESI, D. O conceito de “transmissão linguística irregular” e o processo de formação do português do Brasil. RONCARATI, C.; ABRAÇADO, J. (orgs.). *Português brasileiro: contato linguístico, heterogeneidade e história*. Rio de Janeiro: 7 Letras, p. 272-284, 2003.
- LUCCHESI, D.; BAXTER, A. A transmissão linguística irregular. LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. (Orgs.). *O português afro-brasileiro*. Salvador: EDUFBA, 2009, p. 101-124.
- LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. (Orgs.). *O português afro-brasileiro*. Salvador: Edufba, 2009.
- LUCCHESI, D.; MELLO, C. F. A alternância dativa no português afro-brasileiro: um processo de reestruturação original da gramática. *Papia* (Brasília), v. 19, p. 153-184, 2009.



LUNGUINHO, M. V. S. Partição de constituintes no português brasileiro: características sintáticas. In: SILVA, D. E. (ed.) *Língua, Gramática e Discurso*. Goiânia: Cãnone; Grupo de Estudos Linguísticos do Centro-Oeste, p. 133-147, 2006.

LUNGUINHO, M. V. S.; MEDEIROS JÚNIOR, P. Inventou um novo sujeito: características sintáticas e semânticas de uma estratégia de indeterminação do sujeito no português brasileiro. *A Interdisciplinar: Revista de Estudos em Língua e Literatura*. Sergipe, v. 9, p. 7-21, 2009.

MADARIAGA, N. Reanalysis. In: *The Cambridge Handbook of Historical Syntax* Ed.: Adam Ledgeway e Ian Roberts. Cambridge: Cambridge University Press, 2017, p. 70-91. DOI: 10.1017/9781107279070.005.

MAGALHÃES, T. M. V. A valoração de traços de concordância dentro do DP. *DELTA. Documentação de Estudos em Linguística Teórica e Aplicada*, São Paulo, v. 20, n. 1, p. 149-170, 2004.

MANZINI, R.; SAVOIA, L. Null subjects without pro. *UCL Working Papers in Linguistics*, 9, p. 1-12, 1997.

MARTINS, A. M.; NUNES, J. Passives and se Constructions. In: *The Handbook of Portuguese Linguistics*. Ed.: Leo Wetzels, Sérgio Menuzzi e João Costa. (Org.). Hoboken: Wiley-Blackwell, 2016, p. 318-337.

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

MATTOSO, K. *Ser escravo no Brasil*. São Paulo: Brasiliense, 1982.

MIOTO, C.; FIGUEIREDO SILVA, M. C. LOPES, R. E. V. *Novo manual de sintaxe*. São Paulo: Contexto, 2013.

MIYAGAWA, S. Strong and weak pronouns in the covert system of pronouns. *Journal of Japanese Linguistics*. 34(2): 281-301, 2018. DOI:10.1515/jjl-2018-0017.

MIYAGAWA, S. *Why agree? Why move? Unifying agreement-based and discourse configurational languages*. Linguistic Inquiry Monograph 54, MIT Press, 2010.

MIYAGAWA, S. *Agreement beyond phi*. Linguistic Inquiry Monograph 75. Cambridge, MA: MIT Press, 2017.

MODESTO, M. Null subjects in Brazilian Portuguese and Finnish: they are not derived by movement. In: *New Horizons in the Analysis of Control and Raising*. Ed.: Davies, W. & Dubinsky, S. Dordrecht: Springer, 2007.

MODESTO, M. *On the identification of null arguments*. Tese (Doutorado), Los Angeles, University of Southern California, 2000.

MODESTO, M. Topic prominence and null subjects. In: BIBERAUER, T. (Org.). *The limits of syntactic variation*. Amsterdam: John Benjamins, 2008. p. 375-409.

MUFWENE, S. *Language evolution: Contact, competition, and change*. New York: International Publishing Group, 2008.

MUFWENE, S. Race, racialism, and the study of language evolution in America. *Language variety in the South: Historical and contemporary approaches*, ed. by Michael Picone & Katherine Davis. Tuscaloosa: University of Alabama Press, p. 449-474, 2015.

MUFWENE, S. *The ecology of language evolution*. Cambridge: Cambridge University Press, 2001. DOI: 10.1017/CBO9780511612862.

MÜLLER, G. Pro-drop and impoverishment. In: P. Brandt and E. Fuss (Eds.). *Form, structure and grammar*. A Festschrift presented to Günther Grewendorf on the occasion of his 60th birthday. Tübingen: Narr, p. 93–115, 2006.

MÜLLER, G. *Some consequences of an impoverishment-based approach to morphological richness and pro-drop*. Manuscrito. Universidade de Leipzig, 2008.

MUNHOZ, A. T. M. *A estrutura argumental das construções de tópico-sujeito: o caso dos sujeitos locativos*. Dissertação (Mestrado) – Universidade de Brasília, 2011.

MUNHOZ, A. T. M.; NAVES, R. R. Construções de tópico-sujeito: uma proposta em termos de estrutura argumental e de transferência de traços de *C. Signum*. *Estudos de Linguagem*, 15:245–265, 2012.

NARO, A. The genesis of the reflexive impersonal in Portuguese: a study in syntactic change as a surface phenomenon. *Language* (Baltimore), New York, v. 52, p. 779-810, 1976.

NARO, A.; SCHERRE, M. M. P. O conceito de transmissão linguística irregular e as origens estruturais do português brasileiro: um tema em debate. In: RONCARATI, C.; ABRAÇADO, J. (Org.). *Português brasileiro: contacto linguístico, heterogeneidade e história*. Rio de Janeiro: 7 Letras, 2003.

NARO, A.; SCHERRE, M. M. P. *Origens do português brasileiro*, São Paulo, Parábola Editorial, 2007.

NASCIMENTO, M. *Sur la Postposition du Sujet dans le Portugais du Brésil*. Tese (Doutorado em Linguística), Universidade de Paris VIII, 1984.

NAVES, R. R.; BORGES, H. A emergência da gramática do português brasileiro em Goiás. *Confluência* (Rio de Janeiro), v. 47, p. 68-103, 2015.

NAVES, R. R.; PILATI, E. Aqui/ali/lá em posição pré-verbal de construções com inversão locativa e com sujeito arbitrário em português brasileiro. In: D. Carvalho; D. Brito. (Org.). *Pronomes: morfossintaxe e semântica*. Salvador: EDUFBA, p. 179-194, 2018.

NEELEMAN, A.; SZENDRÖI, K. Radical pro drop and the morphology of pronouns. *Linguistic Inquiry*, v. 38, 2007.

NEGRÃO, E. V. *O português brasileiro: uma língua voltada para o discurso*. 1999. 214 F. Tese (Livre-Docência) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, 1999.

NEGRÃO, E. V.; EVANI, V. Elementos para a investigação da semântica do crítico SE no português brasileiro. *Cadernos de Estudos Linguísticos* (UNICAMP), v. 57, p. 41-59, 2015.

NEGRÃO, E. V.; VIOTTI, E. A estrutura sintática das sentenças absolutas no português brasileiro. *Linguistica* (Madrid), v. 23, p. 37-58, 2010.

- NEGRÃO, E. V.; VIOTTI, E. Brazilian Portuguese as a discourse-oriented language. In: Kato, M., Negrão, E. (Eds.) *Brazilian Portuguese and the null subject parameter*. Madrid: Iberoamericana, 2000, p.105-125.
- NEGRÃO, E. V.; VIOTTI, E. Brazilian Portuguese as a transatlantic language: Agencies of linguistic contact. *InterDISCIPLINARY Journal of Portuguese Diaspora Studies* Vol. 3.1, 2014a.
- NEGRÃO, E. V.; VIOTTI, E. Contato entre quimbundo e português clássico: impactos na gramática da impessoalização do português brasileiro e angolano. *Revista Linguística* (Online), v. 30, p. 289-330, 2014b.
- NEGRÃO, E. V.; VIOTTI, E. Em busca de uma história linguística. *Revista de Estudos da Linguagem*, 20-2: 309-342, 2012.
- NEGRÃO, E. V.; VIOTTI, E. Epistemological aspects of the study of the participation of African languages in Brazilian Portuguese. In: M. M. T. Petter e M. Vanhove (Org.), *Portugais et langues africaines. Études afro-brésiliennes*, Paris, Karthala:13-44, 2011.
- NEWMeyer, F. Goals and Methods of Generative Syntax. In: *The Cambridge Handbook of Generative Syntax*. Ed.: Marcel den Dikken. Cambridge: Cambridge University Press, 2013, p. 61-92. DOI:10.1017/CBO9780511804571.005.
- NOLL, V. & DIETRICH, W. (Ed.). *O português e o tupi no Brasil*. São Paulo, Contexto, 2010.
- NUNES, J. (Ed.). *Minimalist Essays on Brazilian Portuguese Syntax*. Amsterdam: John Benjamins, 2009.
- NUNES, J. *Linearization of chains and sideward movement*. Cambridge, MA: MIT Press, 2004.
- NUNES, J. *O famigerado se: uma análise sincrônica e diacrônica das construções com se apassivador e indeterminador*. Dissertação (Mestrado). Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, 1990.
- NUNES, J. *The copy theory of movement and linearization of chains in the minimalist program*. Tese (Doutorado em Linguística), University of Maryland, College Park, 1999.
- O'NEIL, W. The evolution of the Germanic inflectional systems: A case study in the causes of language change. *Orbis*, 27: 248–85, 1978.
- OLIVEIRA, M. S. D. CAMPOS, E. A. ; CECIM, J. F. ; LOPES, F. J. ; SILVA, R. A. . O português afro-indígena e a comunidade de Jurussaca. In: Ornelas de Avelar, Juanito; López, Laura Álvares. (Org.). *Dinâmicas Afro-Latinas - Língua(s) e História(s)*. 1ed.Berlin: Peter Lang, 2015, v. 1, p. 149-178.
- OLIVEIRA, M. S. D.; ARAÚJO, G. A. *O português na África Atlântica*. São Paulo: Humanitas, 2018.
- ORDÓÑEZ, F. *Word Order and Clause Structure in Spanish and Other Romance Languages*. Tese (Doutorado), CUNY, New York 1997.
- PAIVA, C. A. et al. *Publicação crítica do Recenseamento Geral do Império do Brasil de 1872*. Belo Horizonte, Cedeplar/UFMG, 2012.

- PAIXÃO DE SOUSA, M. C. Linguística Histórica. PFEIFFER, C.; NUNES, J. H. (Org.). *Introdução às Ciências da Linguagem: Língua, Sociedade e Conhecimento*. Campinas: Pontes, 2006, v. 3, p. 11-48.
- PAL, D. C.; ARAÚJO, P. P. Sintaxe e semântica. In: PETTER, M. *Introdução à linguística africana*. São Paulo: Contexto, 2015.
- PALACÍN, L.; MORAES, M. A. de S. *História de Goiás: (1722-1972)*. 7ª Ed. Revisada. Goiânia: Editora Vieira e Editora da UCG, 2008.
- PARKVALL, M. *Da África para o Atlântico*. Campinas, São Paulo: Editora da Unicamp, 2012.
- PERLMUTTER, D. *Deep and surface constraints in syntax*. New York: Holt, Rinehart and Winston, 1971.
- PESETSKY, D. TORREGO, E. The syntax of valuation and the interpretability of features. *Phrasal and clausal architecture: Syntactic derivation and interpretation*. In honor of Joseph E. Emonds. Eds.: Simin Karimi, Vida Samiian e Wendy K. Wilkins, 2007.
- PETTER, M. *Introdução à linguística africana*. São Paulo: Contexto, 2015.
- PETTER, M. O continuum afro-brasileiro do português. In: C. Galves, H. Garmes & F. Ribeiro (eds.). *África-Brasil: caminhos da língua portuguesa*. Campinas: Editora da Unicamp, p. 158-173, 2009.
- PETTER, M. O continuum afro-brasileiro do português. In: C. Galves, H. Garmes & F. Ribeiro (eds.). *África-Brasil: caminhos da língua portuguesa*. Campinas: Editora da Unicamp, p. 158-173, 2009.
- PILATI, E. *Aspectos sintáticos e semânticos das orações com ordem verbo-sujeito no português do Brasil*. Tese (Doutorado) – Universidade de Brasília, Instituto de Letras, Departamento de Linguística, Português e Línguas Clássicas, 2006.
- PILATI, E.; NAVES, R. R.; SALLES, H. (Org.). *Novos olhares para a gramática do português brasileiro*. Campinas: Pontes Editores, 2017.
- PINHEIRO, A. C. C.; COELHO, G. N. *Diário de viagem do Barão de Mossâmedes (1771-1773)*. Goiânia: Trilhas Urbanas, 2006.
- PINTZUK, S. *Phrase structures in competition: Variation and change in Old English word order*. New York: Garland, 1999.
- PINTZUK, S.; TAYLOR, A.; WARNER, A. Corpora and quantitative methods. In: *The Cambridge Handbook of Historical Syntax*. Ed.: Adam Ledgeway e Ian Roberts. Cambridge: Cambridge University Press, 2017, p. 218-240. DOI: 10.1017/9781107279070.012.
- PIRES, A. The subject, it is here! The varying structural positions of preverbal subjects. *DELTA*, São Paulo, v. 23, n. spe. p. 113-146, 2007.
- PIRES, A. The minimalist syntax of defective domains: gerunds and infinitives. Amsterdam: John Benjamins, 2006.
- PLATZACK, C. The Scandinavian languages and the null-subject parameter. *Natural Language and Linguistic Theory*, v. 5, p. 377-401, 1987.

- POLLOCK, J.-Y. Verb Movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry*, 20(3): 365-474, 1989.
- PONTES, E. Da importância do tópico em português. In: Encontro Nacional De Linguística, 1980, Rio de Janeiro. *Anais...* Rio de Janeiro: PUC, 1981.
- PONTES, E. *O tópico no português do Brasil*. Campinas: Pontes, 1987.
- PONTES, E. *Sujeito: da sintaxe ao discurso*. São Paulo: Ática; (Brasília): INL, Fundação Nacional Pró-Memória, 1986.
- RABELO, P. C. *Argumentos (EPP) nulos no português do Brasil em contextos oracionais finitos e infinitos*. 216 f. Tese (Doutorado em Linguística). Universidade de Brasília, Brasília, 2010.
- RAPOSO, E.; URIAGEREKA, J. Indefinite SE. *Natural language and linguistic theory*, Dordrecht, v. 14, n. 2, 1996, p. 749-810.
- REIS, D. S. B. *Práticas de escrita e de arquivamento: o Memorial de Anna Joaquina Marques (Cidade de Goiás, 1881-1930)*. 279 f. Dissertação (Mestrado em Historia) - Universidade Federal de Goiás, Goiânia, 2013.
- RIZZI, L. *Issues in Italian syntax*. Dordrecht: Foris, 1982.
- RIZZI, L. Null subjects in Italian and the theory of *pro*. *Linguistic Inquiry*, v. 17, p. 501-557, 1986.
- ROBERTS, I. 2010. A deletion analysis of null subjects. In: BIBERAUER, T. *et al. Parametric variation: null subjects in Minimalist Theory*. Cambridge: Cambridge University Press, 2010. DOI:10.1017/CBO9780511770784.002.
- ROBERTS, I. *Diachronic syntax*. Oxford: Oxford University Press, 2007.
- ROBERTS, I.; HOLMBERG, A. Introduction: parameters in minimalist theory. BIBERAUER, T. *et al. Parametric variation: null subjects in Minimalist Theory*. Cambridge: Cambridge University Press, 2010. DOI:10.1017/CBO9780511770784.001.
- ROBERTS, I.; KATO, M. (Org.) *Português brasileiro: uma viagem diacrônica*. Campinas: Editora da Unicamp, 1993 [2018].
- ROBERTS, I.; ROUSSOU, A. *Syntactic change. A minimalist approach to grammaticalization*. Cambridge: Cambridge University Press, 2003.
- ROCHA, L. M. *O Estado e os índios: Goiás 1850-1889*. Goiânia: Ed. UFG, 1998.
- RODRIGUES, A. Sobre as línguas indígenas e sua pesquisa no Brasil. *Ciência e Cultura* (SBPC), São Paulo, v. 57, n.2, p. 35-38, 2005.
- RODRIGUES, A. Tupi, tupinambá, línguas gerais e português do Brasil. In: V. Noll & W. Diatrich (Ed.). *O português e o tupi no Brasil*. São Paulo: Contexto, 2010.
- RODRIGUES, C. *Deriving Referential Null Subjects From Movement*. General's paper, University of Maryland, College Park, 2000.
- RODRIGUES, C. *Impoverished Morphology and A-Movement out of Case Domains*. Tese (Doutorado). Graduate School of University of Maryland (Department of Linguistics), 2004.

- ROEPER, T. Universal Bilingualism. *Bilingualism: Language and Cognition* 2(3). 169-186, 1999. DOI: 10.1017/S1366728999000310.
- ROHRBACHER, B. *The Germanic languages and the full paradigm*. Tese (Doutorado em Linguística), University of Massachusetts, Amherst, 1994.
- ROHRBACHER, B. W. 1999. *Morphology-Driven Syntax: A theory of V to I raising and pro-drop*. Amsterdam, John Benjamins, 1999.
- ROSA, M. C. *Uma língua africana no Brasil Colônia de Seiscentos: o quimbundo ou língua de Angola na Arte de Pedro Dias*. Rio de Janeiro: 7Letras, 2013.
- RUMEU, M. C. B. A variação ‘tu’ e ‘você’ no português brasileiro oitocentista e novecentista: reflexões sobre a categoria social gênero. *Alfa: Revista de Linguística* (UNESP. Online), v. 57, p. 545-576, 2013.
- SAAB, A. On the notion of partial (non-) *pro*-drop in Romance. Ed.: Mary Kato & Francisco Ordóñez. *The morphosyntax of Spanish and Portuguese in Latin America*, p. 49-77, 2016.
- SAFIR, K. *Syntactic chains and the definiteness effect*. Tese (Doutorado em Linguística), Cambridge, MA, MIT, 1982.
- SAITO, M. Notes on East Asian argument ellipsis. *Language Research*, 443, p. 203–227, 2007.
- SANTORINI, B. Variation and change in Yiddish subordinate clause word order. *Natural Language and Linguistic Theory*, 10:595–640, p. 1992.
- SANTOS, D. R.; SOARES DA SILVA, H. A ordem V-DP/DP-V com verbos inacusativos. In: DUARTE, Maria Eugênia Lammoglia (org.). *O sujeito em peças de teatro (1833-1992): estudos diacrônicos*. São Paulo: Parábola, p. 121-42, 2012.
- SATO, Y.; KIM, C. Radical *pro drop* and the role of syntactic agreement in Colloquial Singapore English. *Lingua*, 122 (8). p. 858-873, 2012.
- SCHWARCZ, L. M.; STARLING, H. M. *Brasil: Uma Biografia*. São Paulo: Companhia das Letras, 2015.
- SCHWARTZ, S. B. *Escravos, roceiros e rebeldes*. Bauru, SP: EDUSC, 2001.
- SHEEHAN, M. Basic word order in Brazilian Portuguese: A hybrid extended projection principle (EPP). *Estudos Linguísticos e Literários*, 58. pp. 104-129, 2018.
- SHEEHAN, M. The EPP and null subjects in Romance. Tese (Doutorado em Linguística). Newcastle upon Tyne: Newcastle University, 2006.
- SIMÕES, J. S. *Sintaticização, discursivização e semanticização das orações de gerúndio no português brasileiro*. Tese (Doutorado em Filologia e Língua Portuguesa) - Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2007. DOI: 10.11606/T.8.2007.tde-04102007-140928.
- SIMÕES, J. S.; KEWITZ, V. O corpus histórico do português paulista: características e potencialidades. *Revista da Abralin*, v. 16, p. 227-243-243, 2017.
- SOARES, M. C. *Devotos da cor: identidade étnica, religiosidade e escravidão no Rio de Janeiro, século XVIII*. Rio de Janeiro: Civilização Brasileira, 2009.

- SPEAS, M. Economy, agreement and the representation of null arguments. In: Ackema, P., Brandt, P., Schoorlemmer, M., Weerman, F. (Eds.). *Arguments and Agreement*. Oxford University Press, Oxford, p. 35–75, 2006.
- SPEAS, M. Null arguments in a theory of economy of projection. In: Benedicto, E., Runner, J. (Eds.). *University of Massachusetts Occasional Papers in Linguistics*. GLSA, University of Massachusetts, Amherst, pp. 179–208, 1994.
- SUNDQUIST, J. D. *Morphosyntactic change in the history of the mainland Scandinavian languages*. Tese (Doutorado em Linguística), Indiana University, Bloomington, 2002.
- SUNDQUIST, J. D. The Rich Agreement Hypothesis and Early Modern Danish embedded-clause word order. *Nordic Journal of Linguistics*, v. 26, p.233–258, 2003.
- TARALDSEN, K. T. *On the nominative island constraint, vacuous application and the that-trace filter*. Indiana University Linguistics Circle. Bloomington: Indiana, 1980.
- 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. (Org.). *Português Brasileiro: uma viagem diacrônica*. Campinas: Editora da Unicamp, 69-106, 1993 [2018].
- TAYLOR, A. The change from SOV to SVO in Ancient Greek. *Language Variation and Change*, 6:1–37, 1994.
- TEIXEIRA, Z. D. Propriedades sintáticas e semânticas dos advérbios no português brasileiro. Tese (Doutorado). Programa de Pós-Graduação em Linguística: Universidade de Brasília, 2015.
- TESCARI NETO, A. *On Verb Movement in Brazilian Portuguese: a cartographic study*. Tese (Doutorado em Ciências da Linguagem) – Università Ca' Foscari di Venezia, Italia, 2013.
- TESCARI NETO, A. Verb raising, the impoverishment of the verbal paradigm and the weakening of Tense in BP. *Revista do GEL*, v. 13, p. 75, 2016.
- THOMASON, S. G. Contact explanations in linguistics. In: Raymond Hickey (Ed.). *The handbook of language contact*. Oxford: Blackwell, p. 31-47, 2010.
- THOMASON, S. G. *Language contact: an introduction*. Edinburgh: Edinburgh University Press; Georgetown University Press, 2001.
- THOMASON, S. G.; KAUFMAN, T. *Language Contact, Creolization and Genetic Linguistics*. University of California Press, 1988.
- THRÁINSSON, H. Syntactic variation, historical development, and minimalism. In: *Minimalist syntax*. R. Hendrick (Ed.). Malden: Blackwell, p. 152-191, 2003.
- TOMIOKA, S. The semantics of Japanese null pronouns and its crosslinguistic implications. In: SCHWABE, K.; WINKLER, S. (Eds.). *The interfaces: Deriving and interpreting omitted structures*. Amsterdam: John Benjamins, 2003, p. 321-340.
- TSAO, F. *A Functional Study of Topic in Chinese: The First Step toward Discourse Analysis*. PhD dissertation, USC, Los Angeles, California, 1977.
- UTTS, J; HECKARD, R. F. *Mind on statistics*. Boston, MA: Brooks/Cole Cengage Learning, 2012.

VAN KEMENADE, A.; LOS, B. Using historical texts. In: *Research methods in Linguistics*. Ed. Robert Podesva e Devyani Sharma. Cambridge: Cambridge University Press, 2014, p. 216-232. DOI:10.1017/CBO9781139013734.012.

VIEIRA JÚNIOR, W. C. J. *Cartografia da Capitania de Goyaz no século XVIII: intenção e representação*. 193 f. Tese (Doutorado em Arquitetura e Urbanismo) – Universidade de Brasília, Brasília, 2015.

VIKNER, S. Vo-to-Io movement and inflection for person in all tenses. Haegeman, Liliane (Ed.). *Elements of grammar: handbook in generative syntax*. Dordrecht: Kluwer Academic Publishers, p. 189-213, 1997.

WALKDEN, G. The Actuation Problem. In: *The Cambridge Handbook of Historical Syntax*. Ed.: Adam L. & Ian R. Cambridge Handbooks in Language and Linguistics. Cambridge: Cambridge University Press, 2017, p. 403–24. DOI:10.1017/9781107279070.020.

WEINREICH, U.; LABOV, W.; HERZOG, M. Empirical foundations for a theory of language change. In: W. Lehmann and Y. Malkiel (Eds.). *Directions for historical linguistics*. Austin: University of Texas Press, p. 95–189, 1968.

WINFORD, D. The Ecology of Language and the New Englishes: toward an integrative framework. In: *Changing English*. Berlin, Boston: De Gruyter, p. 25-56, 2017.

WINFORD, D. *An introduction to contact linguistics*. New Jersey: Blackwell Publishing, 2003.

YANG, C. *Knowledge and learning in natural language*. Oxford: Oxford University Press, 2002.



## APÊNDICE

Este documento contém as instruções utilizadas para a codificação – em duas planilhas de um arquivo no Excel – das amostras de dados do *corpus* da tese *A sintaxe dos sujeitos na história do português em Goiás: evidências oitocentistas de uma língua de sujeito nulo parcial*. Cada uma das duas planilhas possui um conjunto de dados composto por 1250 sentenças retiradas, respectivamente, de cada um dos manuscritos que compõem o *corpus* da tese, conforme descritos na Tabela 1.

Tabela 1: Manuscritos com amostras transcritas e codificadas.

| <b>TÍTULO DA FONTE ESCRITA:</b>                                                                       | <b>PERÍODO<br/>CODIFICADO:</b> | <b>NOME DA<br/>PLANILHA:</b> |
|-------------------------------------------------------------------------------------------------------|--------------------------------|------------------------------|
| 1. Coleção de notícias da Capitania de Goyaz: os diários de viagem do Barão de Mossâmedes (1772-1773) | 1772-1773                      | DV18                         |
| 2. Memorial de lembrança de Anna Joaquina da Silva Marques (1881-1930)                                | 1882-1883                      | ML19                         |

### **Procedimentos de transcrição e codificação dos dados nos arquivos do Excel**

No arquivo Excel, as duas primeiras colunas nas planilhas são, respectivamente, as colunas “número da sentença/sentence number” e “fonte/source”, que contêm o número da sentença (de 1 a 1250) seguido do código do manuscrito (isto é, o nome da planilha). A terceira coluna é a coluna “dados/data”, que contém as sentenças transcritas dos manuscritos originais. As colunas seguintes nas planilhas correspondem às propriedades da sintaxe do sujeito que estão no escopo desta investigação – essas propriedades estão listadas na Tabela 2 a seguir. Nas planilhas, os códigos das propriedades relacionadas à sintaxe do sujeito estão em inglês porque se considerou a terminologia gerativa em inglês como referência na elaboração dos critérios de codificação dos dados da pesquisa. Ademais, o trabalho de codificação dos dados integrou um conjunto de atividades desenvolvidas nos Estados Unidos durante estágio de doutorando na Universidade Michigan, Ann Arbor, com apoio financeiro da CAPES (processo: 007549/2015-06), sob a supervisão do Prof. Dr. Acrísio Pires.

Eis algumas observações relevantes para a transcrição e codificação dos dados:

(i) a **unidade de análise** é o **núcleo oracional** representado por um **verbo lexical**;

(ii) verbos aspectuais, modais e auxiliares são considerados uma única unidade de análise;

(iii) orações coordenadas e subordinadas incluem duas unidades de análise, isto é, duas linhas na planilha do Excel separadas por uma barra vertical ( | ), como em (1):

- (1)      a. Ela saiu | para ir ao mercado.  
           b. A Maria comprou uma torta de maçã | e a comeu.

(iv) se numa linha da coluna de dados há sentenças transcritas entre colchetes – [ ] –, isso significa que essas sentenças não são consideradas na codificação das colunas com propriedades da sintaxe do sujeito referentes àquela linha;

(v) se reticências são colocadas entre dois parêntesis em uma sentença na coluna de dados – (...) –, isso significa que as palavras do manuscrito original não puderam ser transcritas;

(vi) propriedades sintáticas relacionadas ao argumento interno do verbo são codificadas somente nos contextos em que a sentença possui verbo inacusativo ou é uma construção com *se* (passivo, impessoal, reflexivo etc.).

Tabela 2: Propriedades sintáticas relacionadas aos sujeitos nulos.

| PROPRIEDADES          | CÓDIGO | DESCRIÇÃO                                                                                                                                                                                                                                    |
|-----------------------|--------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Verbo intransitivo | Int_V  | <p>Inserir 1 se há um verbo intransitivo que atribui papel temático à estrutura argumental da sentença.</p> <p>(1) A Maria <u>esteve</u> aqui.</p> <p>Verbos auxiliares, aspectuais, inacusativos e de ligação não são codificados aqui.</p> |

|                        |         |                                                                                                                                                                                                                                                                                                                                                         |
|------------------------|---------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                        |         | <p>(2) Ela <u>era</u> uma boa professora.</p> <p>Contudo, sentenças em que um verbo de ligação ocorre com um predicado verbal são classificadas considerando as propriedades do predicado verbal:</p> <p>(3) Ela foi <u>convidada</u>.</p>                                                                                                              |
| 2. Verbo inacusativo   | Unacc_V | <p>Inserir 1 se o verbo é inacusativo:</p> <p>(4) <u>Caiu</u> o pano.</p>                                                                                                                                                                                                                                                                               |
| 3. Verbo impessoal     | Imp_V   | <p>Inserir 1 se o verbo é existencial ou meteorológico:</p> <p>(5) a. <u>Havia</u> muitos padres ali.<br/> b. <u>Tinha</u> muita gente assistindo à peça.<br/> c. <u>Chovia</u> a cântaros.</p>                                                                                                                                                         |
| 4. Verbo elíptico      | Ellip_V | <p>Inserir 1 se um verbo elíptico atribui papel temático, o que pode ser indicado pela presença de um verbo de ligação, modal, auxiliar ou aspectual:</p> <p>(6) A: Ela vai casar com o João?<br/> B: Vai [<u>casar</u>].</p> <p>Nota: não considerar fragmentos de sentença como um verbo elíptico, como em: “Ela chegou, <u>e o João também.</u>”</p> |
| 5. Verbo no infinitivo | Inf_V   | <p>Inserir 1 se o verbo está no infinitivo:</p> <p>(7) A Maria vai sair / para __ comprar pão.</p> <p>(8) A Maria forçou o João <u>sair</u>.</p>                                                                                                                                                                                                        |

|                                  |        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|----------------------------------|--------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                  |        | Nota: não inserir 1 para dados com verbos elípticos como em (6B).                                                                                                                                                                                                                                                                                                                                                                                                   |
| 6. Locução verbal                | Cx_V   | <p>Inserir 1 se a sentença é composta por uma locução verbal:</p> <p>(9) a. Ele <u>vai sair</u>.<br/> b. Ele <u>vai começar a trabalhar</u>.<br/> c. Ele <u>vai poder começar a trabalhar</u>.<br/> d. <u>Começou a chegar</u> cartas de azeite.<br/> e. Maria pediu à sua mãe   para <u>poder ir</u> ao shopping.</p> <p>(10) A: A Mariana <u>foi falar</u> com o prefeito?<br/> B: <u>Foi</u>.</p> <p>Nota: em (10B), insere-se 1 porque o verbo é aspectual.</p> |
| 7. Pessoa do verbo:<br>1ª ou 2ª? | Pers_V | <p>Inserir 1 se o verbo se refere à primeira ou segunda pessoa:</p> <p>(11) a. (Eu) <u>cheguei</u> cedo.<br/> b. (Nós) <u>chegamos</u> cedo.</p> <p>(12) a. (Tu) <u>chegaste</u> cedo.<br/> b. (Você) <u>chegou</u> cedo.<br/> c. (Vós) <u>chegastes</u> cedo.<br/> d. (Vocês) <u>chegaram</u> cedo.</p>                                                                                                                                                            |
| 8. Verbo no plural               | Num_V  | <p>Inserir 1 se o verbo está conjugado no plural:</p> <p>(13) a. <u>Chegamos</u>.<br/> b. <u>Chegastes</u>.<br/> c. <u>Chegaram</u>.</p>                                                                                                                                                                                                                                                                                                                            |

|                                |           |                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|--------------------------------|-----------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                |           | <p>Também inserir 1 se o verbo no infinitivo está no plural:</p> <p>(14) a. Quando vós chegard<u>es</u> lá,   descansai.<br/> b. Quando vocês chegare<u>m</u> lá,   descansem.<br/> c. Quando chegarm<u>os</u> lá,   descansaremos.</p>                                                                                                                                                                                                              |
| 9. Argumento externo manifesto | Overt_EA  | <p>Inserir 1 para argumentos externos manifestos de verbos transitivos, bitransitivos, no infinitivo ou no gerúndio.</p> <p>(15) a. <u>Minha irmã</u> comprou um vestido novo.<br/> b. <u>Eu</u> dei um vestido novo para minha irmã.<br/> c. <u>João</u> esteve aqui.<br/> d. <u>A Maria</u> viu / <u>o João</u> sair.<br/> e. Vi   <u>o João</u> limpando a casa toda.</p>                                                                         |
| 10. Argumento externo ausente  | Absent_EA | <p>Inserir 1 nos casos de verbos meteorológicos:</p> <p>(16) ___ Chovia<u>a</u> cântaros.</p> <p>Inserir 1 se o verbo da sentença é inacusativo:</p> <p>(17) ___ Chegaram as cartas de aceite.</p> <p>Inserir 1 para o argumento externo nulo de orações coordenadas:</p> <p>(18) A Maria comprou uma torta de maçã   e a ___ comeu.</p> <p>Nota: o conceito argumento externo ausente não é usual na literatura gerativista. Empregamo-lo nessa</p> |

|                            |         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|----------------------------|---------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                            |         | <p>codificação apenas para distinguir 3 casos em que o argumento externo no PB é sempre nulo ou inexistente, quais sejam: (i) o caso dos verbos meteorológicos, que possuem um pronome expletivo nulo na posição de sujeito e não possuem argumentos externo e interno; (ii) o caso das sentenças com verbos inacusativos, que não possuem argumento externo em sua estrutura argumental; e (iii) o caso das estruturas em que um TP é coordenado a outro (cf. FERNÁNDEZ-SALGUEIRO, 2008), isto é, no caso das orações coordenadas com o mesmo sujeito sintático. Essa nomenclatura utilizada na codificação não afeta a análise linguística dos dados, pois é possível recuperar: os pronomes nulos expletivos de verbos meteorológicos; as ocorrências de verbos inacusativos; e os casos em que as estruturas coordenadas possuem sujeitos sintáticos distintos.<sup>129</sup></p> |
| 11. Argumento externo nulo | Null_EA | <p>Inserir 1 para o argumento externo nulo de verbos existenciais:</p> <p>(19) __ <u>Havia</u> muita gente lá.</p> <p>Inserir 1 para o argumento externo nulo de verbos no infinitivo e no gerúndio:</p> <p>(20) a. A Maria forçou o João __ <u>sair</u>.<br/> b. João passou o dia   __ limpando a casa toda.</p> <p>Inserir 1 para argumento externo nulo e referencial:</p> <p>(21) __ Ofereceu o livro à Maria.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                               |

FERNÁNDEZ-SALGUEIRO, G. (2008). *Aspects of the syntax of (TP-)Coordination, across-the-board extraction, and parasitic gaps* (Doctoral dissertation). Tese (Doutorado em Linguística) University of Michigan.

|                                                   |           |                                                                                                                                                                                                                                                                                                                                                              |
|---------------------------------------------------|-----------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                   |           | <p>Inserir 1 para argumentos externos nulos e não referenciais:</p> <p>(22) __ Alugam apartamentos bonitos naquele prédio.</p> <p>Inserir 1 para o argumento externo nulo de verbos inergativos:</p> <p>(23) __ Ali trabalha muito.</p>                                                                                                                      |
| 12. Argumento externo referencial                 | Ref_EA    | <p>Inserir 1 para argumentos externos referenciais (nulos ou manifestos):</p> <p>(24) <u>Pedro e João</u> foram à casa de Maria. Começou a chover. __ Ficaram na casa de Maria   até a <u>chuva</u> passar.</p> <p>Inserir 1 para argumentos externos que se referem à primeira pessoa:</p> <p>(25) a. __ Cheguei cedo.<br/>b. <u>Nós</u> chegamos cedo.</p> |
| 13. Argumento externo diferente de um DP ou um NP | Non_DP/NP | <p>Inserir 1 se o argumento externo é um CP ou um VP (verbo no infinitivo):</p> <p>(26) a. <u>O homem que trabalha</u>   vai prosperar.<br/>b. <u>Nadar</u> é bom para a saúde.</p>                                                                                                                                                                          |
| 14. Argumento externo pronominal                  | Pron_EA   | <p>Inserir 1 se o argumento externo é um pronome nulo ou manifesto:</p> <p>(27) a. Quando __ saímos,   a chuva começou.<br/>b. Marchou <u>Sua Excelência</u> duas léguas.<br/>c. O João<sub>1</sub> disse que <u>ele</u><sub>1/2</sub> comprou a casa.</p>                                                                                                   |

|                                                                        |           |                                                                                                                                                                                                                                                                                                                 |
|------------------------------------------------------------------------|-----------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 15. <i>Topic-shift</i>                                                 | Top_Shift | <p>Inserir 1 se o argumento externo é um <i>topic-shift</i>, como o argumento externo da oração subordinada a seguir:</p> <p>(28) O João<sub>1</sub> disse que   <u>ele</u><sub>2</sub> comprou a casa.</p>                                                                                                     |
| 16. Argumento externo pós-verbal                                       | PostV_EA  | <p>Inserir 1 se o argumento externo é pós-verbal:</p> <p>(29) Estiveram aqui <u>a Maria e o João</u>.</p>                                                                                                                                                                                                       |
| 17. Pessoa do argumento externo:<br>1 <sup>a</sup> ou 2 <sup>a</sup> ? | Pers_EA   | <p>Inserir 1 se o argumento externo se refere à primeira ou segunda pessoa:</p> <p>(30) a. (<u>Eu</u>) cheguei cedo.<br/>b. (<u>Nós</u>) chegamos cedo.</p> <p>(31) a. (<u>Tu</u>) chegaste cedo<br/>b. (<u>Você</u>) chegou cedo.<br/>c. (<u>Vós</u>) chegastes cedo.<br/>d. (<u>Vocês</u>) chegaram cedo.</p> |
| 18. Argumento externo pluralizado                                      | Num_EA    | <p>Se o argumento externo está no plural, inserir 1:</p> <p>(32) a. <u>A Maria e o João</u> estiveram aqui.<br/>b. <u>Eles</u> estiveram aqui.</p>                                                                                                                                                              |
| 19. Argumento interno ausente                                          | Absent_IA | <p>Inserir 1, se o argumento interno é nulo ou ausente, como no caso dos verbos inergativos:</p> <p>(33) A: O João chegou?<br/>B: Chegou ___.</p> <p>(34) Trabalha muito aqui.</p>                                                                                                                              |
| 20. Argumento interno pronominal                                       | Pron_IA   | <p>Inserir 1 se o argumento interno é um pronome manifesto.</p> <p>(35) <u>Eles</u> chegaram.</p>                                                                                                                                                                                                               |



|                                     |         |                                                                                                                                                                                                    |
|-------------------------------------|---------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 21. Argumento interno pré-verbal    | PreV_IA | Inserir 1 se o argumento interno é pré-verbal:<br><br>(36) <u>As cartas de aceite</u> chegaram.                                                                                                    |
| 22. Argumento interno na 3ª pessoa  | Pers_IA | Inserir 1 se o argumento interno se refere à terceira pessoa:<br><br>(37) Enterrou-se <u>o João</u> .                                                                                              |
| 23. Argumento interno pluralizado   | Num_IA  | Inserir 1 se o argumento interno está no plural:<br><br>(38) Alugam-se <u>casas</u> .                                                                                                              |
| 24. Verbo bitransitivo              | Sec_IA  | Inserir 1 se o verbo é bitransitivo:<br><br>(39) Minha irmã <u>me</u> comprou <u>um vestido novo</u> .                                                                                             |
| 25. Sentença com adjunto adverbial? | Pres_AA | Inserir 1 se há um adjunto adverbial na sentença:<br><br>(40) <u>No verão</u> , chegaram as cartas de aceite.                                                                                      |
| 26. Adjunto adverbial pré-verbal?   | PreV_AA | Inserir 1 se o adjunto adverbial está em posição pré-verbal:<br><br>(41) <u>Ontem</u> choveu muito.                                                                                                |
| 27. Subordinada                     | SC      | Inserir 1 se a sentença é uma oração subordinada.<br><br>(42) Não sei   <u>se ele concorda comigo</u> .                                                                                            |
| 28. Coordenada                      | CC      | Inserir 1 para as sentenças com conjunções coordenativas:<br><br>(43) a. A Maria comprou uma torta de maçã   <u>e</u> a comeu.<br>b. A Maria comprou uma torta de maçã, <u>mas</u> o João a comeu. |

|                              |          |                                                                                                                                                                                                                                                                                                           |
|------------------------------|----------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 29. Construção com <i>se</i> | SE_Const | <p>Inserir 1 se o pronome clítico <i>se</i> relacionasse com o verbo da sentença:</p> <p>(44) a. Ele <u>se</u> matou. (reflexivo)<br/> b. Alugam-<u>se</u> casas. (médio-passivo)<br/> c. Não <u>se</u> usa mais sai. (indefinido)<br/> d. Isso é fácil   de <u>se</u> fazer. (sujeito de infinitivo)</p> |
| 30. Passiva verbal           | V_Pass   | <p>Inserir 1 se a sentença é uma passiva verbal:</p> <p>(45) <u>A Maria foi convidada para o baile</u>, mas não foi.</p>                                                                                                                                                                                  |

*Nota final:*

Após analisadas todas as linhas de dados, as células das colunas com propriedades sintáticas dos sujeitos nulos que não foram codificadas com 1 foram preenchidas com 0.