

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

JACQUELINE DE SOUSA BORGES DE ASSIS

**DE *SE* A (VO)CÊ: O PERCURSO DA INDETERMINAÇÃO
NO PORTUGUÊS BRASILEIRO - PB**

UBERLÂNDIA

2017

JACQUELINE DE SOUSA BORGES DE ASSIS

**De se a (vo)cê: o percurso da indeterminação
no Português brasileiro - PB**

Tese apresentada ao Programa de Pós-graduação em Estudos Linguísticos, do Instituto de Letras e Linguística da Universidade Federal de Uberlândia, como requisito parcial para a obtenção do título de Doutora em Estudos Linguísticos.

Área de concentração: Estudos em Linguística e Linguística Aplicada. **Linha de pesquisa:** Teoria, Descrição e Análise linguística.

Orientadora: Prof^ª. Dr^ª. Maura Alves de Freitas Rocha

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

2017

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

A848d Assis, Jacqueline de Sousa Borges de, 1965-
2017 De se a (vo)cê : o percurso da indeterminação no Português
brasileiro / Jacqueline de Sousa Borges de Assis. - 2017.
117 f. : il.

Orientadora: Maura Alves de Freitas Rocha.
Tese (doutorado) - Universidade Federal de Uberlândia, Programa
de Pós-Graduação em Estudos Linguísticos.
Inclui bibliografia.

1. Linguística - Teses. 2. Sociolinguística - Teses. 3. Gramática
comparada e geral - Sintagma nominal - Teses. 4. Língua Portuguesa -
Teses. I. Rocha, Maura Alves de Freitas. II. Universidade Federal de
Uberlândia. Programa de Pós-Graduação em Estudos Linguísticos. III.
Título.

CDU: 801

“Mire e veja: o mais importante e bonito, do mundo, é isto: que as pessoas não estão sempre iguais, ainda não foram terminadas – mas que elas vão sempre mudando. Afinam ou desafinam. Verdade maior. É o que a vida me ensinou”.

João Guimarães Rosa

Dedico este trabalho a todos aqueles que se empenham em desvendar os enigmas da mudança linguística.

E ao meu filho, Pedro Paulo.

AGRADECIMENTOS

Um dia direi: não foi fácil, mas com esforço eu venci. *Obrigada meu Deus! Sem Seu amor, eu nada seria!*

O mérito da conclusão deste trabalho deve ser compartilhado com algumas pessoas a quem gostaria de agradecer.

Primeiramente, sou grata a minha orientadora, Prof^a. Dr^a. Maura Alves de Freitas Rocha, que, acreditando em mim, incentivou-me a realizar outra pesquisa nesta área tão dinâmica e instigante da linguística: a variação e mudança. Meu reconhecimento por seu incentivo, apoio, dedicação e paciência em me orientar para a realização deste desafio.

Às professoras Dras. Joaquina Aparecida Nobre da Silva, Vânia Cristina Casseb Galvão, Elisete Maria de Carvalho Mesquita, Luísa Helena Borges Finotti, pela disponibilidade em participar da banca de defesa da tese, pela leitura cuidadosa do texto e pela colaboração com pertinentes considerações.

Aos professores Drs. Maria Eugênia Duarte, Juanito Avelar e José Magalhães, por terem aceitado participar das bancas de qualificação, quando contribuíram com valiosas sugestões.

Aos companheiros da pós-graduação, dentre os quais destaco Joaquina, Marília e minha prima-irmã e amiga, Valdete, por ter dividido comigo os melhores e piores momentos desta jornada, pela força, consideração, dicas e inesquecíveis viagens para congressos.

Ao diretor do CEFET-MG/ unidade de Araxá, professor Dr. Henrique José Avelar, pela sensibilidade em entender minha necessidade de afastamento do trabalho para dedicar-me a esta tese. E as minhas amigas cefetianas de longa jornada, agradeço pela torcida, apoio, palavras de carinho e momentos de descontração. Dentre os amigos, ganha destaque minha amiga de ‘todas as horas’, Elisa, a quem devo o encorajamento para esta batalha. Meu reconhecimento por sua importância nesta conquista!

Sou imensamente grata ao meu filho, meu ponto de apoio, pelo companheirismo e carinho de sempre, e ao meu marido por saber compreender minhas ausências e respeitar minhas angústias. A todos meus irmãos pela constante demonstração de amor, e a minhas queridas cunhadas, Cris e Clénize, tias e sobrinhos, pelo carinho em me acolher, minha eterna gratidão!

Aos colegas de jornada da biblioteca do Uniaraxá, pelas muitas horas compartilhadas e pela cordialidade. Vocês deixaram meus dias de trabalho mais prazerosos. Meu muito obrigada a todos, em especial ao Tales, por ser essa pessoa tão ímpar!

Por fim, deixo um agradecimento especial ao meu pai, que se orgulha de meu esforço. Obrigada, pai, por me dar uma grande razão para buscar sempre mais!

Trabalho parcialmente financiado pela **CAPES**.

RESUMO

Diante do quadro de alterações na gramática do PB como sintomas da remarcação do Parâmetro do Sujeito Nulo – PSN por que tem passado a língua, investigo a emergência do pronome de 2^a. pessoa com marcação genérica – *você_{gen}*, largamente empregado na língua falada como forma alternativa aos pronomes: *nulo* genérico e não referencial, *se*, *a gente* e *nós*. Esta tese parte da hipótese de que *você_{gen}* exerce no PB a função de pronome indeterminador à semelhança do clítico *se*, cujo apagamento ou deixa vazia essa posição ou é substituído por *você_{gen}*. Associo essa possibilidade às mudanças sintáticas processadas no PB: apagamento do clítico *se*; surgimento do paradigma de pronomes fracos nominativos no lugar do sistema pronominal *Agr* como compensação da perda do sujeito nulo referencial (Kato, 2000); perda da propriedade de genericidade do pronome *a gente*; emergência de construções existenciais com *ter*; tendência ao preenchimento da posição de sujeito que caminha para a perda do nulo não referencial e genérico, em uma clara evidência do encaixamento linguístico da variação. O modelo teórico-metodológico seguido é o da Sociolinguística Paramétrica proposto por Kato & Tarallo (1989). Assim, tendo em vista a variação interlinguística, tem-se como princípio norteador da discussão formal das variantes a teoria gramatical de Princípios & Parâmetros, de Chomsky (1981), e avanços em sua versão minimalista (Kato, 2000; Holmberg, 2010). O ponto relevante deste trabalho em relação à implementação de *você_{gen}* está na proposição de que o interior de INFL é o locus onde o pronome é inicialmente inserido antes de ser movido para a posição de sujeito gramatical das construções de *ter_{ex}* nas quais o pronome não estabelece relação temática com o verbo. Considerando a variação paramétrica, um estudo comparativo com o PE comprova que quanto à realização de sujeitos indefinidos, o PB e o PE também exemplificam contrastes relevantes, confirmando a hipótese de que as mudanças sintáticas que estão desencadeando alteração de valor do PSN é um fenômeno que acomete somente o PB.

Palavras-chave: Sociolinguística Paramétrica, Parâmetro do Sujeito Nulo, variação linguística, indeterminação do sujeito, LSNP.

ABSTRACT

*Considering the alterations in the grammar of PB as symptoms of the remarking of the Null Subject Parameter - PSN that the language has passed, I investigate the emergence of the 2nd person pronoun with generic mark-up, widely used in the spoken language as an alternative form to the pronouns: null non referential and generic, 'se', 'nós' and 'a gente'. This thesis is based on the hypothesis that você_{gen} is used on the function of indeterminating pronoun like the clitic *se*, whose deletion leaves this position empty or is substituted by você_{gen} . I associate this possibility with the syntactic changes processed in PB: deletion of clitic *se*; emergence of the nominative weak pronoun paradigm in place of the pronominal Agr system as compensation for the loss of the referential null subject (Kato, 2000); loss of the property of genericity of the pronoun 'a gente'; emergence of existential constructions with 'ter'; tendency to fill the position of subject that moves to the loss of the non referential and generic null, in clear evidence of the linguistic embedding of the variation. The theoretical-methodological model followed is that of Parametric Sociolinguistics proposed by Kato & Tarallo (1989). Thus, in view of the interlinguistic variation, this thesis adopt as guiding principle Chomsky's grammatical theory of Principles & Parameters (1981), and advances of this theory in its minimalist version (Kato, 2000; Holmberg, 2010). The relevant point of this work in relation to the implementation of você_{gen} is in the proposition that the interior of INFL is the locus where the pronoun is initially inserted before being moved to the position of grammatical subject of the constructions of ter_{ex} in which the pronoun does not establish thematic relation with the verb. Considering the parametric variation, a comparative study with the EP proves that in relation to the performance of indefinite subjects, BP and PE also exemplify relevant contrasts, confirming the hypothesis that the syntactic changes that are triggering alteration of the PNS value is a phenomenon that affects only PB.*

Key words: Parametric Sociolinguistics, Parameter of the Null Subject, linguistic variation, indetermination of the subject, LSNP.

SUMÁRIO

Lista de siglas e abreviaturas.....	8
Lista de gráficos, tabelas e quadro.....	9
INTRODUÇÃO	10
1. MODELO DE VARIAÇÃO E MUDANÇA	22
1.1 A Sociolinguística laboviana	24
1.1.1 A comunidade de fala: uma das questões polêmicas da sociolinguística laboviana	27
1.2 A Sociolinguística Paramétrica.....	33
2. O PARÂMETRO <i>PRO-DROP</i>	41
2.1 A proposta de Kato (2000).....	41
2.1.1 Considerações parciais.....	53
2.2 A proposta de Holmberg (2010)	54
2.2.1 Resumo.....	63
3. SINTOMAS DA REMARCAÇÃO DO PARÂMETRO <i>PRO-DROP</i> NA GRAMÁTICA DO PB	64
3.1 Reanálise e apagamento do clítico <i>se</i>	64
3.2 O surgimento do paradigma de pronomes fracos.....	66
3.3 Estratégias inovadoras de indeterminação	71
3.3.1 <i>A gente</i> cede lugar a <i>você</i> para referência genérica	71
3.3.2 A sintaxe das construções existenciais com <i>ter</i>	77
3.3.3 Propriedades do verbo <i>ter_{ex}</i> e sua predicação	80
3.4 Propostas de análise	89
3.4.1 Do caráter não expletivo de <i>você</i> genérico em existenciais	85
3.4.2 Do caráter indeterminado de <i>você</i> genérico.....	89
3.4.3 Sobre a derivação de construções com nulo genérico e nulo expletivo	94
3.4.4 Referência arbitrária x referência genérica (+) e (-) inclusiva.....	98
3.5 Considerações parciais.....	102
4. METODOLOGIA	104
4.1 Material selecionado... ..	104
4.2 Fatores.....	107
5. ANÁLISE E DISCUSSÃO DOS DADOS	115
5.1 Sujeitos nulos <i>versus</i> plenos	119
5.2 Sujeitos nulos em variação com sujeitos genéricos pronominais	119
5.3 Nulo não referencial x <i>você_{gen}</i> em construções existenciais com <i>ter</i>	124
5.4 <i>A gente</i> cede lugar a <i>você</i> : o que dizem os dados.....	131
5.5 Grau de indeterminação	136
5.6 O tempo verbal.....	141
5.7 O caráter hipotético da construção.....	144
5.8 Presença x ausência de SADVs e SPs locativos e temporais	145
5.9 Construções finitas x infinitivas.....	148
5.10 Intercambialidade	152
5.11 Paralelismo	153
5.12 As forças sociais exercidas sobre as formas em variação	154
5.13 O percurso da indeterminação no PB	158
6. CONCLUSÕES	162
7. REFERÊNCIAS BIBLIOGRÁFICAS	165
Apêndice	171

LISTA DE SIGLAS E ABREVIATURAS

Agr - núcleo do sintagma de concordância; afixo pronominal de concordância (Kato, 1999).

AGR - núcleo frasal (*clause head*)

AgrP - sintagma de concordância (*agreement phrase*)

Caso default – o Caso que por definição não é checado.

D - determinante

DP - sintagma determinante (*determiner phrase*)

D – traço de difinitude (Holmberg)

uD – traço de difinitude não interpretável (*Unvalued definiteness feature*) / traços não valorados.

EPP – Princípio de Projeção Estendido, o qual, segundo Chomsky (1998), requer que a posição de SpecTP seja preenchida por um XP.

Exp – expletivo

I - flexão (núcleo do sintagma flexional, de *inflection*)

INFL - sintagma flexional (*inflection*) - responsável por agregar traços de tempo e de concordância.

LSN – Língua de Sujeito Nulo

LSNN - Língua de Sujeito não Nulo

LSNP - Língua de Sujeito Nulo Parcial

NP – sintagma nominal

Numeração: conjunto de pares (LI, *i*), onde LI é um item do léxico e *i* um índice correspondente ao número de vezes que aquele item será utilizado na derivação.

PB - Português brasileiro

PE - Português europeu

PF - forma fonética (*phonetic form*)

PP - sintagma preposicional (*prepositional phrase*)

pro - pronome nulo

pro_{exp} - pronome nulo expletivo

PSN - Parâmetro do Sujeito Nulo

Pro-drop – é a classificação atribuída a uma língua que atende o parâmetro de sujeito nulo: realização do traço de *pessoa* com pronome nulo.

SN – sintagma nominal

SN – sujeito nulo (Kato)

Spec - especificador (*specifier*)

SpecTP – especificador do sintagma temporal.

T – *tense*

TP - sintagma temporal (*tense phrase*). No núcleo TP, há os traços-*phi* (*pessoa*, *número* e *gênero*) e um traço EPP, que atrai o sujeito que está em posição inferior.

t - vestígio de movimento (*trace*)

Traços- ϕ (traços-*phi*) - traços formais de *gênero*, *número* e *pessoa*.

V - verbo

VP - sintagma verbal (*verb phrase*)

VOS – ordem dos sintagmas (verbo-objeto-sujeito)

Σ P – sintagma da periferia esquerda

ϕ P - Pronome de terceira pessoa defectivo

LISTA DE GRÁFICOS, TABELAS E QUADROS

Gráfico 1: Ocorrência das variantes no PB em tempo real	116
Gráfico 2: Ocorrência das variantes no PE em tempo real.....	117
Gráfico 3: Distribuição das variantes nas construções existenciais no PB em tempo real.....	127
Gráfico 4: Distribuição das variantes nas construções pessoais no PB em tempo real.....	130
Gráfico 5: Distribuição das variantes no PB nos tipos de construção	134
Gráfico 6: Grau de indeterminação das variantes no PB.....	138
Gráfico 7: Distribuição das variantes em relação aos tempos verbais no PB - déc. 70.....	141
Gráfico 8: Distribuição das variantes em relação aos tempos verbais no PB - déc. atual.....	142
Gráfico 9: Distribuição das variantes em construções hipotéticas no PB em tempo real	144
Gráfico 10: Emprego das variantes na presença x ausência de SADVs e SPs no PB.....	146
Gráfico 11: Distribuição das variantes em existenciais na presença de locativos no PB.....	147
Gráfico 12: Frequência das variantes em relação ao tipo de sentença no PB	148
Gráfico 13: Distribuição das variantes nas construções de infinitivo no PB	149
Gráfico 14: Ocorrência das estratégias de indeterminação no PE por tipo de construção	151
Gráfico 15: Emprego das variantes em relação ao paralelismo no PB.....	153
Gráfico 16: Distribuição das variantes por faixa etária no PB	156
Tabela 1: Ocorrência das estratégias de indeterminação no PB por década	115
Tabela 2: Ocorrência das estratégias de indeterminação no PE por década.....	116
Tabela 3: Distribuição das variantes em construções existenciais no PB em tempo real	171
Tabela 4: Distribuição das variantes nas construções pessoais no PB em tempo real	171
Tabela 5: Grau de indeterminação das variantes por período no PB	137
Tabela 6: Distribuição das variantes em relação a tempos verbais no PB em tempo real	171
Tabela 7: Distribuição das variantes nas construções hipotéticas no PB em tempo real	172
Tabela 8: Emprego das variantes na presença x ausência de SADVs e SPs no PB	172
Tabela 9: Distribuição das variantes em existenciais na presença de locativos no PB	172
Tabela 10: Frequência das variantes em relação ao tipo de sentença no PB.....	152
Tabela 11: Intercambialidade das variantes na década atual no PB.....	173
Tabela 12: Distribuição das variantes em relação ao fator paralelismo no PB	173
Tabela 13: Estratificação das variantes em relação a gênero no PB	155
Tabela 14: Distribuição das variantes por faixa etária: tempo aparente no PB.....	173
Tabela 15: Ocorrência das estratégias de indeterminação no PE por tipo de construção	174
Tabela 16: Distribuição das variantes em relação a tempos verbais no PE.....	174
Tabela 17: Distribuição das variantes por faixa etária: tempo aparente no PE.....	174
Quadro 1: Frequência dos pronomes nulo genérico e expletivo no PB nas duas décadas.....	140

INTRODUÇÃO

O fato de as línguas mudarem com o passar do tempo tem sido analisado por vários teóricos sob diversas perspectivas, desde fins do século XVIII, quando teve início a Linguística Histórica e Comparada. De acordo com esta abordagem, a Linguística firmou-se como ciência com os estudos históricos, a partir do reconhecimento de Willian Jones (1786 apud Faraco, 2011) de que as semelhanças entre o sânscrito, o latim e o grego apontavam para uma origem comum entre essas três línguas. Esses estudos desencadearam o princípio de que as línguas mudam com o tempo.

Assim, a mudança linguística se tornou investigação científica com base mais sólida por meio do método comparativo, e da tese de que as correlações sistemáticas apontam para uma origem comum entre as línguas. Mas foi com o movimento neogramático (1878) que ela ganhou rigor metodológico, ao se questionarem os pressupostos tradicionais da prática histórico-comparativa, e se estabelecerem uma orientação metodológica diferente e um conjunto de postulados teóricos para a interpretação da mudança linguística.

Com a publicação do Manual do Pensamento Neogramático em 1880, *Hermann Paul* (1970) propôs uma diretriz para os estudos da mudança linguística, a de universalizar as condições de vida da língua, traçando de uma maneira geral as linhas fundamentais de uma teoria da evolução da mudança. Para Paul, os princípios fundamentais da mudança linguística deveriam ser buscados nos fatores psíquicos e físicos tomados como determinantes dos objetos culturais como a língua.

Ao entender que o fundamento da cultura era o elemento psíquico e que só havia uma psicologia individual, Paul propõe a tese de que a fonte de toda mudança linguística era o indivíduo falante e de que a propagação da mudança se dava por meio da ação recíproca dos indivíduos, perspectiva adotada pelos gerativistas, embora estes assumam não um psicologismo, mas um biologismo na base do processo. Paul propõe ainda que a mudança linguística é originada principalmente no processo de aquisição da língua.

Entretanto, o psicologismo na base da concepção dos neogramáticos também foi alvo de críticas, visto que sua constituição está na tensão entre o social e o individual.

Dentre os linguistas que participaram da crítica aos neogramáticos destaca-se o austríaco Hugo Schuchardt. Paradoxalmente, esse linguista tinha visão subjetivista da língua e, para se opor à lei fonética, chamou a atenção para as variedades de fala existentes em uma comunidade, as quais são condicionadas por fatores sociais, tema esse retomado pela

sociolinguística na década de 1960. O linguista busca nesse contexto compreender o processo de mudança linguística, introduzindo no século XX um tratamento da variação como condicionada pelo contexto social e cultural da língua, trilhando o caminho da sociolinguística.

É a partir desse ponto de vista que o século XX trouxe nova luz para os estudos linguísticos. O trabalho de William Labov vem ampliando, desde o início da década de 1960, nosso conhecimento sobre um dos pontos centrais da linguística, a mudança linguística, em que esta passa a ser analisada também em uma perspectiva sincrônica e vinculada a fatores sociais. A sociolinguística laboviana configurou uma abordagem diferente, de natureza dinâmica, em que a língua deixa de ser vista como uma estrutura estática e homogênea e passa a ser vista como um sistema em constante mutação e profundamente comprometido com a estrutura social em que se insere.

Esta pesquisa segue a abordagem da Sociolinguística Paramétrica, tendo em vista a sua preocupação em explicar a variabilidade linguística e sua inter-relação com fatores sociais e o papel que desempenham no processo de mudança. Nesta perspectiva, e considerando que a diversidade linguística no Português do Brasil (doravante PB) pode ser observada em todos os níveis, no presente trabalho pretende-se analisar um fenômeno sintático do PB em comparação ao Português europeu (PE).

O fenômeno em análise diz respeito à inovadora estratégia de indeterminação/generalização do sujeito no PB, o pronome genérico de 2ª. pessoa (doravante *você_{gen}*), cujo uso vem sendo observado tanto na língua falada como escrita, como uma forma alternativa às canônicas *se*, *a gente*, *nós*, *nulo genérico* e *nulo expletivo*. Entende-se a indeterminação do sujeito como situações em que não se consegue ou não se quer estabelecer uma referência precisa para o ser sobre o qual se faz algum tipo de declaração. A partir desse entendimento, este estudo pretende analisar as construções de/com: i) clítico *se*, tanto nominativo como passivo em construções em que a interpretação passiva passa a ter uma interpretação indeterminada (reanálise do *se*)¹; ii) pronomes pessoais (*você* e *cê* - *forma foneticamente reduzida*)² com referência genérica; (*nós*) com referência indeterminada e o indefinido (*a gente*) e; iii) nulo não referencial e nulo genérico, conforme os exemplos a seguir:

¹O termo ‘reanálise de *se*’, empregado por Nunes (1990), diz respeito à mudança de interpretação do clítico que passa de apassivador à indeterminador. Ressalto que não faço distinção entre as construções ditas de *se passivo* ou de *se-indeterminador* devido ao fato de ambas serem construções em que o argumento externo é interpretado como indeterminado ou genérico, pelos traços semânticos do pronome *se*.

² A redução fonética de *você* será abordada no item 1.4.2 sobre a teoria de pronomes fracos, de acordo com Kato (1999).

- (1) Para resolver os problemas do Brasil hoje, precisa-*se* de investimentos.
- (2) Para resolver os problemas do Brasil hoje, *você^{gen}* precisa de investimentos.³
- (3) Para resolver os problemas do Brasil hoje, *a gente* precisa de investimentos.
- (4) Para resolver os problemas do Brasil hoje, *nós* precisamos de investimentos.
- (5) Para resolver os problemas do Brasil hoje, *Ø_{gen}* precisa de investimentos.
- (6) Não *Ø_{exp}* tem investimento na economia hoje.

Constata-se, assim, que ao lado das possíveis formas de indeterminação do sujeito consagradas pela norma gramatical, como no exemplo (1), o PB tem apresentado outras que conduzem a uma interpretação indeterminada para o sujeito das sentenças, ora realizado, com o emprego de formas pronominais com generalização de sentido, como nos exemplos de (2) a (4), ora nulo, exemplo (5). Além do preenchimento da posição de sujeito nulo de 3ª pessoa de referência genérica, os pronomes plenos indeterminados estão co-ocorrendo na gramática do PB com o pronome nulo não referencial, exemplo (6).

O que se observa em processos de variação é que eles não ocorrem de forma isolada. Normalmente, ligam-se ou encaixam-se em outros processos em variação em outras partes do sistema. Assim, as mudanças processadas no PB como sintomas da remarcação do parâmetro *pro-drop* apontam para encaixamento linguístico.

Por um lado, como já foi amplamente discutido por diversos autores, como Galves (1993, 2001), Figueiredo Silva (1996), Duarte (2000), Cavalcante (2007), dentre outros, alterações do *status pro-drop* no PB associam-se à redução do paradigma flexional. Sem a possibilidade de se recuperar o sujeito por meio da morfologia flexional, uma vez que se evidencia no PB um enfraquecimento da concordância pela redução de seu paradigma verbal com a entrada no sistema pronominal das formas *a gente* e *você*, é preciso lexicalizá-lo.

Por outro lado, a emergência no PB de formas pronominais de referência indefinida na posição de sujeito pode estar associada à reanálise do *se apassivador* (cf. Nunes, 1990); ao apagamento do *se*, que, por sua vez, pode ser atribuído à perda em curso dos clíticos no PB (cf. Duarte, Kato e Barbosa, 2003), culminando em construções de sujeito nulo de 3ª pessoa de referência genérica, exemplo (5); e na sequência, ao surgimento de pronomes fracos no lugar do sistema pronominal *Agr*, como compensação da perda do sujeito nulo referencial no PB (Kato, 1999).

³ Exemplo extraído do Programa *Em Pauta* da Globo News, em 06/02/2016.

Nunes (1990), com base no quadro da Teoria de Regência e Vinculação – TRV desenvolvido por Chomsky (1981), propõe que houve no PB uma reanálise sintática em que o expletivo da posição de sujeito de construções com *se passivo* foi reinterpretado como um pronome nulo referencial, levando à interpretação indeterminada do *se passivo* e, conseqüentemente, ao seu apagamento em frases finitas.

Com relação ao apagamento da partícula *se*, Nunes (1990) aponta que estaria relacionado ao fato de que, enquanto o PE licencia uma referência determinada na identificação do pronome nulo de 3ª pessoa do singular em sentenças finitas, o PB licencia uma referência indeterminada. Duas causas poderiam estar, de acordo com Nunes, desencadeando esse apagamento: a discordância na 3ª pessoa do plural – *ele/eles fala*, ou a tendência de supressão de clíticos no PB.

É interessante notar que depois de implementada uma das propriedades das línguas de sujeito não nulo, o preenchimento dos sujeitos referenciais, o sistema começa a caminhar no sentido do preenchimento dos sujeitos não referenciais e genéricos, integrando um conjunto de propriedades características do PB que o distanciam do PE, razão por que o PE foi estabelecido para a análise interlinguística.

De acordo com Duarte (1997, 2000, 2003), seria natural esperar que os sujeitos não referenciais ou expletivos começassem também a se realizar foneticamente, apresentando nosso sistema um conjunto de estruturas nessa posição que vem preenchê-la.

Como consequência do preenchimento da posição do sujeito em construções de sujeito nulo com referência indeterminada e não referencial, substantivos como *gente* passaram por processo de pronominalização, incorporando gradualmente em seu significado a noção de pronome pessoal (Omena e Braga, 1996), (Omena, 2003), e pronomes pessoais, como *você*, passam a expressar sentido genérico.

A hipótese de que as estratégias em análise tenham sido desencadeadas por sucessivos processos de mudança efetivados na língua está fundamentada em Weinreich, Labov e Herzog, 1968 (doravante WLH) que postulam que nenhum traço novo se implementa no sistema sem produzir outras estruturas a ele relacionadas de forma "não acidental".

Este, portanto, trata-se de um típico processo de mudança linguística encaixada, em que o PB está perdendo uma de suas propriedades mais marcantes, a categoria vazia de sujeito, ou sujeito nulo, evidenciando um processo gramatical mais amplo: a tendência ao preenchimento da posição de sujeito, tanto referencial como não. É fato que a posição de sujeito vem sendo cada vez mais preenchida no PB.

Constitui evidência da alteração paramétrica que parece operar no sistema do PB o fato de que são cada vez mais frequentes estruturas com pronomes plenos ocupando posição não-argumental, como no caso das construções existenciais com *ter* que permitem a ocorrência de diferentes estratégias para o preenchimento dessa posição, conforme exemplos de (7) a (11) reestruturados a partir do exemplo (6):

(7) *Ø_{exp}* não *tem* investimento na economia hoje.

(8) Não *se* *tem* investimento na economia hoje.

(9) *Você/ê* não *tem* investimento na economia hoje⁴.

(10) *A gente* não *tem* investimento na economia hoje.

(11) *Nós* não *temos* investimento na economia hoje.

A literatura gerativista contempla diversas pesquisas como as de Lunguinho (2006), Duarte (2003), Avelar (2009), Avelar e Callou (2011), Marins (2013), dentre outras, que abordam a tendência do PB em inserir elementos na posição estrutural de sujeito como efeito da perda da propriedade de licenciar sentenças em que a posição de sujeito aparece vazia.

Este estudo parte, assim, da hipótese de que ao perder seu sujeito nulo referencial, e um paradigma de pronomes fracos nominativos surgir no lugar do sistema pronominal *Agr* (cf. Kato, 2000), o PB está caminhando para a perda do nulo não referencial ou expletivo e do nulo genérico.

O surgimento de formas fracas pode ser visto no PB em contextos nos quais a língua licenciaria sujeito nulo, como uma estratégia para compensar a perda da propriedade de licenciar sujeito nulo e para valoração do traço EPP da sentença. Assim, com base na proposta de Kato (1999), considero que com o surgimento do paradigma de pronomes fracos no PB emergiu na língua um pronome com referência genérica (2ª pessoa), o qual está sendo empregado como forma de evitar que o verbo ocorra em posição inicial, tanto para satisfação do EPP, em construções de nulo genérico, como em atendimento ao parâmetro ϕ -independente⁵ do PB, em construções de nulo expletivo.

Nessa perspectiva, este estudo se propõe a analisar o processo de variação de *se*, partícula de indeterminação na Língua Portuguesa padrão, até *ê* – a forma foneticamente

⁴ Exemplo extraído do Programa *Em Pauta* da Globo News, em 06/02/2016.

⁵ Este parâmetro será explicitado no item 1.4.5.

reduzida de *você*, perpassando pela baixa produtividade de clíticos no PB, que deixa vazia a posição do sujeito indeterminado, dando origem ao nulo indeterminado, o qual vem perdendo espaço para a forma genérica *você/cê*, que, por sua vez, pode estar exercendo a função de elemento indeterminador no PB. Deve-se considerar que a forma reduzida *cê* também tem se mostrado produtiva no PB para referência genérica em posição de sujeito, cujo comportamento sintático está sendo comparado no presente estudo ao clítico *se*, razão por que proponho que o percurso da indeterminação no PB vai de *se* a *cê*.

Considerando-se a Teoria de Princípios e Parâmetros de Chomsky (1981) – especificamente no tocante ao Parâmetro do Sujeito Nulo - PSN (ou parâmetro *pro-drop*), que é responsável por distinguir as línguas entre aquelas que permitem um sujeito foneticamente nulo em frases finitas e aquelas que não permitem – e a versão minimalista desta teoria (cf. Chomsky, 1995 e trabalhos subsequentes), e considerando-se, ainda, o nulo genérico como uma das formas de indeterminação do sujeito no PB, este estudo pretende averiguar se as estratégias de indeterminação com formas pronominais, mais especificamente a forma inovadora de 2ª pessoa - *você* com sentido genérico, constituem indícios de um processo de competição de formas variantes que levarão a uma mudança mais geral de remarcação do valor do PSN no PB, em contraste com o PE.

Por meio do levantamento bibliográfico acerca do objeto desta pesquisa, foi possível constatar que mais recentemente a área genérica tem recebido maior atenção pelo interesse que as construções genéricas com a 2ª. pessoa têm despertado, como resultado do incremento de sua frequência durante os últimos anos em diversas línguas.

O emprego do pronome de 2ª. pessoa para referência genérica e indeterminada é uma estratégia comum em línguas como o Inglês, Francês e Espanhol. No PB, essa referência já fora realizada pelo pronome *a gente*, que, ao passar a ser empregado com referência mais determinada de primeira pessoa do plural, portanto quase-inclusiva nos termos de Holmberg (2010), perde a propriedade de genericidade. Um estudo comparativo com essas línguas que apresentam comportamento semelhante no tocante à substituição de pronomes de referência indeterminada/ genérica, privilegiando o pronome de 2ª. pessoa, mostra que elas têm em comum o fato de seus sistemas lançarem mão de outro pronome para referência genérica sempre que o pronome empregado com esse fim perde esta característica, e que, portanto, esta variação é inerente ao sistema dessas línguas. Diante dessa possibilidade de substituição de pronomes para referência genérica no PB, *a gente* por *você*, um estudo específico a respeito integrará esta tese.

Uma possível explicação para esse comportamento similar entre as línguas em relação ao emprego do pronome de 2^a. pessoa para referência genérica tem respaldo na teoria das propriedades paramétricas, segundo a qual as línguas podem convergir em determinadas partes de sua gramática, revelando movimentos “sincronizados e espelhados” (Kato e Tarallo, 1989).

Tendo em vista a variação interlinguística, ou paramétrica, o estudo comparativo entre as duas variedades do Português – PB e PE quanto à variação em análise visa comprovar a hipótese de que as mudanças sintáticas que estão desencadeando alteração de valor do PSN são fenômenos que acometem somente o PB, uma vez que as formas de indeterminação nulo genérico e *você_{gen}*, não constam da gramática do PE.

Estudos comparativos sobre sujeitos, como os de Kato & Tarallo (1986), Nunes (1990) e Duarte, M.E; Kato, M e Barbosa, P. (2003) mostram que o PE segue em direção oposta ao PB: enquanto o PB usa formas pronominais nominativas (expressas ou nulas) para a representação de sujeitos indefinidos (ou indeterminados), o PE privilegia as construções com *se* para expressar a indeterminação do sujeito. Já nas sentenças infinitivas, de acordo com Duarte, M. E; Kato, M. e Barbosa, P. (2003), ambas as variedades preferem o sujeito nulo, com o *se* aparecendo em cerca de 10% das sentenças. A diferença está no fato de que o PB já começa a usar *você* em substituição ao *se*, como nos exemplos (12) e (13) mostrados pelas autoras:

(12) Você esquece como é legal [**você** ligar uma rádio e ter alternativas]⁶ (PB)

(13) É até um clichê no *rock'n roll* [**você** não precisar ler (partituras musicais)]
(PB)

Dessa forma, pelas possibilidades de emprego de *você_{gen}*, dentre as quais, como sujeito de infinitivo e de existenciais, contextos que não selecionam argumento externo, é que o emprego de *você* como estratégia de indeterminação do sujeito se mostra altamente produtivo na língua falada espontânea no PB, suplantando o emprego da forma pronominal *a gente*, conforme tem sido mostrado por pesquisas empíricas como as de Duarte (1995) e de Almeida (2010).

Os resultados das análises aqui propostas – quantitativos e qualitativos, subsidiarão a caracterização das repercussões da mudança por que está passando o PB, como consequência de

⁶ Os exemplos (12) e (13) foram extraídos de Duarte, M.E; Kato, M e Barbosa, P (2003).

um epifenômeno que pode estar conduzindo à alteração de *status* de língua de sujeito nulo parcial (doravante LSNP) ou *pro-drop* parcial, para língua de sujeito não nulo (LSNN).

Assim, um alcance destas análises será a verificação da perda da propriedade de licenciamento do sujeito nulo genérico no PB, uma característica que o distingue das línguas de sujeito nulo prototípicas (LSN) e o aproxima das LSNP, conforme Holmberg (2010). Sob essa perspectiva, é possível que estejam ocorrendo no PB fenômenos linguísticos que estariam refletindo uns nos outros e provocando instabilidade no sistema: se por um lado se observa o esvaziamento da posição do sujeito indeterminado pela supressão do *se*, originando o nulo de referência genérica, por outro, se observa a tendência do sistema de preenchimento dos sujeitos referenciais tanto definidos como indefinidos com formas pronominais plenas, além dos não referenciais em sentenças existenciais com *ter*, o que parece estar perfeitamente de acordo com as demais alterações sofridas na gramática do PB.

A relevância da análise das variantes elencadas neste trabalho está em detectar se se trata de mudança em progresso, a quais fatores sociais e linguísticos estão relacionadas, de que forma elas se encaixam nos concomitantes linguísticos das formas em questão, e se podem ser consideradas sintomas da alteração paramétrica na língua. Além disso, um estudo comparativo com o PE será revelador para confirmar ou refutar a hipótese de Tarallo (1993) de que o PB e o PE seguem por caminhos divergentes, ao mesmo tempo em que subsidiará a análise interlinguística, conforme o modelo metodológico que me propus a seguir: a Sociolinguística Paramétrica, modernamente denominada variação paramétrica.

Por meio da observação preliminar do vernáculo em relação ao objeto de estudo desta pesquisa, delimitaram-se as perguntas a serem respondidas. Busco, assim, com esta pesquisa resposta para as seguintes questões:

1. Considerando-se a proposta de Kato (1999) de que a perda do sujeito nulo referencial no PB correlaciona-se com o surgimento do paradigma de pronomes fracos, possíveis portadores de traços-*phi*, como compensação para a perda do licenciamento de sujeito nulo, pode-se afirmar que a perda do nulo genérico e do não referencial também está correlacionada a esse paradigma?
2. O preenchimento da posição vazia do sujeito nulo genérico e não referencial com pronomes genéricos plenos deve-se à satisfação de EPP, uma vez que, de acordo com Holmberg (2010), o sujeito nulo genérico não checa o EPP?

3. A presença de pronomes pessoais na posição de sujeito gramatical⁷ de sentenças existenciais com *ter* está em consonância com as das demais alterações sofridas pela gramática do PB?
4. O pronome *você_{gen}* tem a função de elemento indeterminador no PB?
5. O emprego do pronome de 2^a. pessoa para marcação genérica no PB é uma estratégia da qual a língua lança mão para compensar a perda da propriedade de referência genérica do pronome *a gente*?
6. O sujeito nulo indeterminado que permanece no PB é para estabelecer referência arbitrária?

Fundamentada no construto teórico da Sociolinguística Paramétrica (Kato e Tarallo, 1989), com o objetivo de esclarecer aspectos de variação que apontam para a mudança linguística, por meio não só de um modelo de análise que busca a interação entre fatores internos e externos, mas também que aponta para a variação interlinguística, esta pesquisa foi norteada pelas seguintes hipóteses:

- i) Após a perda do nulo referencial, o PB tende ao preenchimento da posição do sujeito nulo não referencial e genérico, perdendo uma de suas características mais marcantes: a categoria vazia de sujeito.
- ii) No PB, a preferência pela estratégia inovadora de indeterminar o sujeito (forma pronominal plena *você_{gen}*) está afastando-o do PE, onde o *se* passivo é estratégia preferida.
- iii) O pronome *você_{gen}* e sua forma reduzida “*cê*” estão atuando no PB como elementos indeterminadores com função similar ao clítico *se*.
- iv) O emprego de *você_{gen}* como estratégia de indeterminação do sujeito é mais produtivo na fala espontânea do que as demais formas elencadas como suas variantes, suplantando inclusive a forma *a gente*, cuja substituição se deve à perda da sua capacidade de referenciação genérica.

⁷ A distinção entre sujeito gramatical e semântico referenda-se em Avelar e Callou (2011, p.254): “sujeito semântico é o constituinte que recebe papel temático como argumento externo do verbo (em termos minimalistas, na posição de Spec-v); sujeito gramatical é o constituinte que concorda com o verbo, podendo ou não coincidir com o sujeito semântico da oração (em termos minimalistas, Spec-T/Infl é a posição naturalmente destinada ao sujeito gramatical em línguas como o Português) ”.

v) Constituem fatores linguísticos favorecedores da forma pronominal genérica *você*: construções finitas, construções existenciais, contextos hipotéticos, SADVs e SPs locativos e temporais, paralelismo, operador temporal (presente e futuro do subjuntivo).

vi) O pronome nulo que permanece na língua é para referência arbitrária que *você_{gen}* não atribui.

A partir dessas hipóteses, foram elencados os objetivos gerais e específicos que direcionaram a pesquisa e a análise dos dados.

Objetivo Geral:

i) Analisar as construções com: i) *se* nominativo e *se* passivo com interpretação *indeterminada*; ii) formas pronominais (*nós, a gente, você e cê*) com generalização de sentido; iii) construções de sujeito nulo não referencial, genérico e arbitrário, que se encontram em estado de variação no Português falado do Brasil como formas alternativas de indeterminação/ generalização do sujeito, em uma perspectiva sincrônica, tomando por base o modelo teórico-metodológico da Sociolinguística Paramétrica.

Objetivos Específicos:

i) Averiguar se o emprego genérico da forma pronominal *você*, na posição de sujeito, é relevante para confirmar a hipótese de que o PB tende ao preenchimento da posição do nulo não referencial e genérico, alterando as propriedades paramétricas do PSN.

ii) Comparar o emprego, nas duas variedades do Português – PB e PE, das variantes elencadas como estratégias de indeterminação do sujeito no PB, tendo em vista que no PE a passiva com concordância ainda é a forma preferida de indeterminação, conforme estudos comparativos entre essas línguas. Comprovar, assim, a proposta de Tarallo (1993) sobre aspectos da sintaxe do PB de que o fazem divergir do PE.

iii) Correlacionar a tendência ao preenchimento do sujeito nulo genérico e não referencial com formas pronominais plenas aos processos de mudança em curso no PB, tais como a supressão do clítico *se*, a perda da propriedade de licenciar sujeito

nulo e o surgimento do paradigma de formas fracas, como evidências do encaixamento linguístico da mudança paramétrica na língua.

iv) Comparar a frequência dos empregos de *você (cê)* e *a gente* como estratégias de indeterminação do sujeito no PB nas duas décadas em tempo real de curta duração (décadas de 70 e atual) para avaliar a evolução e difusão desses empregos.

v) Investigar os contextos linguísticos que favorecem e restringem o emprego da forma pronominal *você/cê* em oposição aos pronomes *se* e nulo genérico como estratégias de indeterminação do sujeito.

vi) Analisar se com o preenchimento da categoria vazia do sujeito nulo não referencial e genérico o PB está caminhando para se tornar uma LSN.

vii) Verificar se o pronome nulo que permanece na língua é para marcar a referência arbitrária, considerando-se que o pronome *você_{gen}* não atribui essa referência.

Diante dessas considerações iniciais, esta tese está assim dividida. No primeiro capítulo apresento o modelo de Variação e Mudança adotado, a Sociolinguística laboviana, bem como a proposta de compatibilização com a Teoria Sintática Gerativa, a Sociolinguística Paramétrica.

Na sequência, os avanços do Parâmetro *pro-drop* em sua abordagem Minimalista são apresentados no capítulo 2, nas perspectivas de Kato (2000) e de Holmberg (2010), que trazem contribuições teóricas sobre características do PB que permitem sua classificação como uma LSN.

No capítulo 3, abordo alguns sintomas da remarcação do PSN que podem ter repercutido na implementação da inovadora estratégia de indeterminação no PB, o pronome de 2ª. pessoa – *você* - para marcação genérica, a partir da proposição de um percurso que vai do pronome *se* ao *cê* (forma reduzida de *você*). São eles: reanálise e apagamento de clíticos no PB; o surgimento do paradigma de pronomes fracos; *a gente* cede lugar a *você*; a sintaxe das construções existenciais com *ter*; a derivação de construções com nulo genérico e nulo expletivo; referência genérica x referência arbitrária das formas pronominais. Na sequência, apresento propostas de análise sobre: o caráter indeterminado de *você_{gen}*; a derivação de construções com nulo genérico e nulo expletivo; e a referência arbitrária x genérica (+) e (-) inclusiva.

No capítulo 4 é delineada a metodologia da pesquisa, nos moldes do modelo laboviano, com apresentação do material analisado e dos grupos de fatores elencados como possíveis favorecedores das variantes em análise e o embasamento teórico para esse levantamento.

No capítulo 5 estão apresentados os dados processados pelo programa estatístico *GoldVarb X*, distribuídos em tabelas e gráficos, com valores percentuais, seguidos da discussão e análise quantitativa e qualitativa dos dados.

Por fim, nas conclusões são retomadas as discussões iniciais sobre como o surgimento de um pronome para marcação genérica no PB constitui um sintoma da remarcação do PSN encaixado em outras alterações em processo na língua que passam a caracterizar o PB atual como uma LSNP em face do PE.

1. MODELO DE VARIAÇÃO E MUDANÇA

Os estudos da sintaxe do Português falado vêm se implementando nos últimos anos tanto entre os sociolinguistas como entre os gerativistas brasileiros, em diferentes perspectivas. Por serem os aspectos da sintaxe também objeto de confronto constante entre o que prediz a norma padrão e o que se encontra na fala, pesquisas nesta área são imperativas.

A investigação sociolinguística tem comprovado que o estudo da variação linguística deve-se fazer tomando em conta o contexto social, posto que a língua está a serviço da comunidade de fala. Portanto, um dos princípios básicos da sociolinguística laboviana é que não é possível compreender o desenrolar de uma mudança fora da estrutura social da comunidade em que ocorre. Nesta perspectiva, as diferenças na fala dos membros de uma comunidade correlacionam-se com certos fatores sociais, comprovando que muitas das variações têm uma motivação social.

Desta forma, partindo do pressuposto de que a inovadora estratégia de indeterminação do sujeito no PB constitui caso de variação sintática, a investigação da área de atuação das variantes no meio social possibilitará organizar os seus possíveis grupos de fatores extralinguísticos condicionadores.

Se entendermos que os fatores linguísticos motivadores das variantes provêm de hipóteses sobre o fenômeno em variação, um critério para especificá-los seria ter como ponto de partida uma hipótese mais geral sobre a língua, formulada por meio de uma teoria gramatical. Nesta perspectiva, é necessário um conjunto de informações definidas a partir de uma teoria gramatical, tanto na descrição das formas consideradas variantes como na descrição do que conta como fator linguístico.

O modelo metodológico seguido, portanto, é o da Sociolinguística Paramétrica proposto por Kato & Tarallo (1989). Adoto o mesmo caminho por eles empreendido, que “resgata a compatibilidade entre as propriedades paramétricas do modelo gerativo e as probabilidades do modelo variacionista, seja para provar seu espelhamento e reflexo, seja para realinhar um modelo em função do outro” (KATO E TARALLO, 1989, p. 5).

Assim, considerando esta proposta de compatibilização, e tendo em vista um direcionamento mútuo entre a variação intra e interlinguística, este estudo terá como princípio norteador para a análise formal das variantes a teoria gramatical de Princípios e Parâmetros, de Chomsky (1981), e seus avanços.

A teoria de Princípios e Parâmetros parte do entendimento de que a gramática universal se constitui em um conjunto de princípios rígidos, leis gerais a que todas as línguas humanas devem obedecer, e parâmetros, cujo valor final é atingido por meio da fixação de uma opção entre duas possíveis. É nos parâmetros que se encontra o lócus da variação entre as línguas. Deve-se a esta teoria, a proposta da existência do parâmetro denominado de Sujeito Nulo (doravante PSN ou *pro-drop*), que é responsável por distinguir as línguas entre aquelas que permitem um sujeito foneticamente nulo em frases finitas e aquelas que não permitem. Nessa concepção de gramática, uma mudança linguística é descrita como um conjunto de manifestações superficiais decorrentes da alteração do valor de parâmetro.

Um alcance dessa proposta de compatibilização de teorias é o de que, tendo em vista os pressupostos teóricos de Princípios e Parâmetros, não seria previsível que o PB seguisse por caminhos diversos do europeu, conforme postulado por Tarallo (1993). Se ambos possuem a mesma estrutura, portanto, as mesmas propriedades paramétricas, o PE deveria necessariamente caminhar para as mesmas mudanças por que está passando o PB.

Dentre outros motivos, o comportamento diferenciado do Português brasileiro em relação ao europeu pode ser explicado por fenômenos linguísticos que se processaram no PB desencadeando mudanças em teias, que refletiram umas nas outras, denominando-se, em WLH (2009) de encaixamento linguístico. Esse eco sintático de um processo de mudança a outro é teoricamente previsível a partir de um paradigma sintático forte para a análise linguística: o modelo chomskyano (cf. Kato e Tarallo, 1989).

No caso específico da variação em análise, este fenômeno parece sugerir que as estratégias de indeterminação do sujeito no PB, tanto com formas pronominais plenas, como com sujeito nulo genérico, representam um conjunto de mudanças em andamento no Português do Brasil desencadeado por outros processos de mudança efetivados na língua.

1.1 A SOCIOLINGUÍSTICA LABOVIANA

Na perspectiva teórica laboviana, a associação língua e sociedade é inquestionável: a língua é um fato social. As necessidades coletivas moldam a língua e suas alterações são reflexo de alterações na sociedade. A sociedade atua, assim como os fatores linguísticos, para determinar seu curso, suas variações e prováveis mudanças. Segundo Labov, “a língua é uma forma de comportamento social” (LABOV, 2008, p. 215). Nessa perspectiva de linguagem como um fenômeno de natureza essencialmente social, o termo sociolinguística torna-se redundante. Para Labov, sociolinguística é linguística.

O modelo teórico-metodológico da Sociolinguística quantitativa proposto por Labov na década de 1960 aparece como uma reação à ausência do componente social no modelo gerativo. De acordo com o autor,

A Linguística tem sido definida de forma a excluir o estudo do comportamento social ou o estudo da fala. Esta definição tem sido conveniente para seus formuladores, os quais, por inclinação pessoal, preferiram trabalhar com seu próprio conhecimento, com informantes individuais ou com materiais secundários (LABOV, 2008, p. 219).

Labov postula a possibilidade de se sistematizar a variação existente e própria da língua falada e tem no fato linguístico seu ponto de partida. E é exatamente no método de pesquisa que recai sua crítica mais severa ao modelo gerativo, mais especificamente, na limitação dos dados, que conduziu Chomsky à convicção de que a teoria não é determinada pelos dados.

O linguista adota uma visão oposta à gerativista, de acordo com a qual, com o estudo da língua diretamente em seu contexto social, a quantidade de dados disponíveis aumenta enormemente e oferece-nos maneiras e meios para decidir qual das muitas prováveis análises é a correta. Assim, para Labov, o estudo da estrutura linguística deve se dar por meio do exame dos fatos linguísticos, conforme manifestados no dia a dia das pessoas, no contexto social da comunidade de fala.

No entendimento de Labov, o estudo dos fenômenos da fala vale por aquilo que eles podem nos dizer sobre a estrutura linguística. E é exatamente neste contexto que seu trabalho procura, por meio da variação linguística, possibilitar predições sobre a questão da mudança linguística, usando o presente para clarificar o passado. De acordo com Labov:

*The close examination of the present shows that much of the past is still with us. The study of history benefits from the continuity of the past as well as from analogies with the present*⁸ (LABOV, 1994, p.27).

Embora o referencial teórico adotado neste estudo seja o da Sociolinguística laboviana e paramétrica, os pressupostos teóricos subjacentes a esse quadro, que também serão importantes para a discussão proposta, baseiam-se em WLH, os quais passo a apresentar.

Neste seu trabalho, WLH propõem a superação do ideal de homogeneidade linguística identificada com estrutura, com a defesa da heterogeneidade sistemática, partindo da indagação de que “se uma língua tem de ser estruturada para funcionar eficientemente, como é que as pessoas continuam a falar enquanto a língua muda, isto é, enquanto passa por períodos de menor sistematicidade?” (WLH, 2009, p. 35). Ou seja, se a noção de estrutura implica efetivo funcionamento do sistema, a identificação entre homogeneidade e estrutura somente poderia ser aceita se a variação na fala fosse admitida como devidamente estruturada e com funcionamento efetivo dentro da comunidade em questão; em caso contrário, não haveria comunicação eficaz durante períodos de mudança, em que a língua passaria por perda de sistematicidade.

Os autores consideraram, assim, que um modelo de linguagem que acomodasse os fatos de uso variável e de seus denominadores sociais propiciaria uma teoria de mudança linguística que pudesse superar os paradoxos não resolvidos da linguística histórica, cuja solução estaria no rompimento da identificação entre estrutura e homogeneidade.

Neste enfoque, o sistema deixa de ser visto como uma sucessão de etapas estáticas, e a história do sistema linguístico passa a ser focalizada em uma perspectiva dinâmica, em que seu funcionamento e suas características estruturais são consideradas nos diversos pontos no tempo. Afinal, para que os sistemas mudem, eles necessariamente sofrem algum tipo de variação. E o vínculo entre variação e mudança só poderá ser constatado se aceitarmos a história e o passado como refletidos no presente.

Desta forma, um dado momento de nosso sistema, em um recorte sincrônico, revela, a partir de sua estrutura e funcionamento, suas relações com o passado e possíveis projeções para o futuro; o estado sincrônico é, assim, resultado de um desenvolvimento passado que continua no presente.

⁸ Um rigoroso exame do presente mostra que muito do passado continua entre nós. O estudo histórico se beneficia pela continuidade do passado, bem como pelas analogias com o presente.

Embasando toda a teoria está o argumento segundo o qual há heterogeneidade linguística nas comunidades de fala, e é dentro delas que se deve buscar estrutura, sistema e funcionamento. Segundo WLH, “a chave para uma concepção racional da mudança linguística – e mais, da própria língua – é a possibilidade da descrição da diferença ordenada em uma língua que serve a uma comunidade” (WLH, 2009, p.36).

Uma vez que as línguas humanas consistem em sistemas organizados de forma e conteúdo, seria estranho que a diversidade não fosse uma de suas propriedades. Na realidade, a diversidade é uma propriedade inerente aos sistemas linguísticos e o papel da Sociolinguística é exatamente tratá-la como objeto de estudo. Assim, a natureza variável da língua é, para a Sociolinguística, um pressuposto fundamental, que orienta a observação e a descrição do comportamento linguístico.

A Sociolinguística vê a diversidade não como um problema, mas como constitutiva da língua. Um dos corolários desta abordagem é que “numa língua que serve a uma comunidade complexa (*i.e.*, real), a falta de heterogeneidade estruturada é que pode ser disfuncional” (WLH, 2009, p.36).

Sendo um dos princípios básicos da teoria não ser possível compreender o desenrolar de uma mudança fora da estrutura social da comunidade em que ocorre, as diferenças na fala dos membros de uma comunidade normalmente correlacionam-se com fatores sociais e estilísticos, razão por que muitas das variações têm motivação social. Nas palavras de Labov (2008, p. 21), “as pressões sociais estão operando continuamente sobre a língua, não de algum ponto remoto no passado, mas como uma força social imanente agindo no presente vivo”. É necessário, portanto, recorrer às variações derivadas do contexto social para se buscar explicações para as variações inerentes ao sistema linguístico.

Assim, na pesquisa sociolinguística, o que se quer analisar é o grupo de indivíduos e não o indivíduo tomado isoladamente, pois, conforme argumenta Labov (2008), a língua, como os demais fatos sociais, não é propriedade do indivíduo, mas propriedade da comunidade. Para Labov, a *langue* é um fato social que determina o comportamento linguístico e o locus da linguagem está na comunidade.

Portanto, por ser o locus de investigação da teoria laboviana, a comunidade de fala constitui um dos construtos fundamentais de sua proposta. Razão por que é preciso pautar-se em critérios para a definição do que delimita o pertencimento do indivíduo a uma comunidade de fala.

1.1.1 A comunidade de fala: uma das questões polêmicas da sociolinguística laboviana

Ao se apresentar o quadro teórico da sociolinguística laboviana, mais especificamente em relação ao conceito de comunidade de fala, é importante ressaltar alguns questionamentos que incidem sobre a teoria laboviana, principalmente sobre a não incorporação das dimensões sociais, tendo em vista que Labov, de acordo com Figueroa (1994), fornece escassa teoria social e vaga definição dos seus construtos fundamentais. Nesse contexto, embora constitua um dos pontos centrais da sociolinguística laboviana, o lócus de investigação, a definição de comunidade de fala apresenta aspectos controversos e polêmicos. Nos trabalhos de Labov, essa noção aparece fundamentada em dois aspectos: em termos de atitudes normativas (no nível consciente) ou em termos de propriedades formais (no nível inconsciente).

O próprio Labov assume que esta não é uma noção fácil de definir: “Deve-se admitir que não há acordo quanto à forma de definir uma comunidade de fala, e pode ser indagado se a comunidade de fala existe mesmo como um objeto definível”⁹ (LABOV, 1980 apud FIGUEROA, 1994, p. 85).

Diante disso, analiso o conceito laboviano de comunidade de fala e qual critério Labov prioriza em seus estudos para delimitá-la. A partir daí, espero chegar a um consenso para subsidiar a delimitação da comunidade de fala a ser investigada nesta pesquisa.

As incoerências e inconsistências em relação à noção de comunidade de fala são tantas que, de acordo com Figueroa (1994), Labov chega a admitir que não há acordo quanto a esta definição. Em uma das tentativas de conceituá-la, Labov o faz nos seguintes termos: “Uma comunidade de fala não pode ser concebida como um grupo de falantes que usam todas as mesmas formas; ela é mais bem definida como um grupo que compartilha as mesmas normas a respeito da língua” (LABOV, 2008, p. 188).

A princípio, a fala da comunidade é definida por Labov em termos de normatividade, o que implica que os membros de uma comunidade de fala não necessariamente utilizam as mesmas formas ou compartilham as mesmas normas de uso, mas partilham o mesmo sistema normativo de valores. Entretanto, no trabalho de Labov (1989) parece haver uma tendência ao afastamento da normatividade para definições mais formalistas de comunidade, como por exemplo, no seu estudo sobre o uso ou não uso do ‘*short*’ a em Filadélfia.

⁹ Original: “It must be admitted that there is no agreement on how to define a speech community, and can be asked whether the speech community even exists as a definable object.”

Nesse trabalho, Labov (1989) define os falantes não negros de Filadélfia como comunidade em termos de padrões de uso do ‘*short*’ a. Ou seja, é o uso real de uma forma linguística que é o fator distintivo e não a interpretação normativa. Segundo o linguista, “o ‘*short*’ a de Filadélfia é preciso o suficiente para servir de guia para quem quiser falar como um falante de Filadélfia”¹⁰ (LABOV, 1989, p. 51). Labov afirma ainda que, na aprendizagem da língua, os membros do grupo são socializados para o ‘contrato social’, que é vinculativo, assim, “se o ‘*short*’ a em Filadélfia é parte da *langue* para filadelfianos brancos, então ele é vinculativo sobre o indivíduo” (LABOV, 1989, p. 53).

Nessa direção, Labov assume que, além de valores conscientes em relação à língua (atitudes normativas), os falantes de uma mesma comunidade de fala compartilham, inconscientemente, aspectos essenciais do sistema linguístico – as regras gramaticais, sendo que os indivíduos adquirem tal sistema sem que possam escolher falar de uma ou de outra forma.

Ainda em relação aos critérios formais linguísticos como base para a adesão em uma comunidade, Labov (1989) distingue o sistema linguístico que se adquire simplesmente por crescer em uma comunidade, do sistema sócio-simbólico, que se pode adquirir de outras maneiras, ou que se pode manipular por razões simbólicas. “A estrutura uniforme gramatical sobre a qual a utilização baseia-se não é algo que o indivíduo tenha qualquer controle, portanto, uma falante branca de Filadélfia não opta por usar o ‘*short*’ a; ela não pode usar, mas o usa” (LABOV, 1989, p. 58).

Figuerola (1994) sugere ainda que, na realidade, este problema pode estar relacionado com o desejo aparente de Labov de manter a sociolinguística dentro da linguística padrão e longe da etnografia da comunicação e da sociologia da linguagem. O que se percebe é que Labov, ao adotar uma ou outra definição, defende o ponto de vista subjacente à delimitação de sua pesquisa.

Apesar de Labov enfatizar o aspecto social da língua na formulação do conceito de comunidade de fala, Figuerola (1994) não deixa claro se o aspecto de consciência das atitudes dos falantes é priorizado por Labov em detrimento das normas gramaticais compartilhadas pelo grupo para delimitar a comunidade de fala, da forma como é afirmado por Severo (2008):

Considerando-se os dois níveis (consciente e inconsciente) envolvidos na delimitação da comunidade de fala, fica claro que Labov prioriza o caráter de consciência das atitudes dos falantes em relação às normas gramaticais

¹⁰ Original: “The ‘*short*’ a of Philadelphia is accurate enough to serve as a guide for those who want to talk as a speaker of Philadelphia.”

compartilhadas pelo grupo para caracterizar a comunidade de fala: “uma comunidade de fala não pode ser concebida como um grupo de falantes que utiliza as mesmas formas; ela é mais bem definida como um grupo que compartilha as mesmas normas em relação à língua (LABOV, 1972, p. 158 apud SEVERO, 2008).¹¹

Por outro lado, Figueroa (1994) sugere que talvez Labov tenha preferido as atitudes dos falantes para determinar a comunidade de fala devido à busca da homogeneidade da comunidade de fala, evitando, também, um certo tipo de variação:

Esperava-se que, ao nos concentrarmos sobre os julgamentos dos falantes nativos em vez de sua fala real, muito desta variação poderia ser desviada. De certa forma, esta esperança justifica-se: membros de uma comunidade de fala compartilham um conjunto comum de padrões normativos, mesmo quando nós encontramos variação altamente estratificada na fala real (LABOV, 1972, p. 192 apud FIGUEROA, 1994).¹²

Portanto, se para Labov os membros de uma comunidade de fala não têm, necessariamente, de falar da mesma forma, eles simplesmente compartilham uma série de avaliações sobre a fala, é um indício de que Labov não propõe que a comunidade de fala seja homogênea, como inferem vários linguistas, como Figueroa (1994), Romaine (1982) e Severo (2008). Essa questão gera a seguinte dúvida: será que o compartilhamento de atitudes e valores em relação às variantes da língua por um grupo de falantes implica uniformidade a tal ponto de se pressupor uma comunidade de fala homogênea em termos linguísticos?

O questionamento de Severo (2008) sobre o posicionamento de comunidade homogênea de Labov, parte do fato de que, diante da pluralidade de contextos de uso da língua, seus membros circulam por diversos grupos sociais, o que faz com que a comunidade seja heterogênea. Por isso, segundo a autora, se a realidade dinâmica e conflitiva da fala dos indivíduos de uma mesma comunidade for considerada, a noção de comunidade de fala é uma abstração teórica tomada *a priori*.

Acredito, todavia, que no trabalho de Labov não haja evidências deste entendimento, uma vez que, ao situar o estudo da língua no contexto social, o autor rompe com as correntes anteriores que analisavam a língua como estrutura homogênea, e considera que a existência de heterogeneidade é que permite ao sistema linguístico se manter em funcionamento mesmo nos períodos de variação linguística.

¹¹ Tradução do original.

¹² Tradução do original.

Além disso, Labov enfatiza que as atitudes compartilhadas são sobre formas estigmatizadas, de forma que, se não houver marcas sociais vinculadas às variáveis, as normas podem não ser tão uniformes. Nesses casos, a delimitação da comunidade de fala não poderia se restringir unicamente aos valores compartilhados pelos falantes, pois há variáveis que não são, necessariamente, conhecidas por eles.

Se o conceito de comunidade tivesse caráter homogêneo, a pressuposição de uniformidade quanto à avaliação das variáveis linguísticas pelos membros da comunidade inviabilizaria a correlação de fatores sociais como motivadores da variação linguística. Inviabilizaria, grosso modo, o objetivo da sua proposta teórico-metodológica, que é o de correlacionar aspectos da língua e da sociedade na identificação de grupos de falantes que possuem características linguísticas em comum.

Portanto, parece-me um pouco inconsistente a crítica feita por Figueroa (1994) a Labov quanto a haver uma tensão permanente no trabalho de Labov, entre a afirmação de que o objeto de investigação linguística é ordenado, sistemático e não particularista (se estudado corretamente), e o desacordo de Labov com Saussure, Chomsky e outros que insistem na necessária homogeneidade do objeto linguístico. Figueroa (1994) afirma que se, por um lado, Labov está em acordo com Saussure sobre a *langue* ser fundamentalmente homogênea no nível da comunidade, por outro, está em desacordo com Saussure quanto ao estudo científico da língua precisar ignorar heterogeneidade real, bem como quanto ao entendimento de que a *parole* é caótica e desmotivada.

Ora, o que a teoria da mudança linguística esboçada por WLH (2009) defende, ao propor ordem e sistematicidade, é a heterogeneidade estruturada, ou seja, a motivação e a natureza ordenada da variação, mantendo a ideia de que “os desvios de um sistema homogêneo não são todos eles erros aleatórios de desempenho, mas são em um alto grau codificados e partem de uma descrição realista da competência de membros de uma comunidade de fala” (op cit, p. 60).

A noção de comunidade de fala vem sendo discutida desde a década de 60 por autores como Gumpers (1996 apud Figueroa, 2004), para quem o início da sociolinguística moderna é marcado pelo reconhecimento de que a correlação entre aspectos linguísticos e sociais deve considerar a comunidade de fala. Portanto, para a sociolinguística, a comunidade de fala, e não o indivíduo ou a língua, é a unidade de estudo. Segundo Gumpers, uma comunidade de fala pode ser constituída por grupos pequenos por meio do contato face a face ou mesmo pode cobrir grandes regiões, dependendo do nível de abstração que se deseja alcançar. Outro autor que

discute a definição de comunidade de fala, Patrick (2004 apud Figueroa, 2004), ressalta que esta noção não deve ser tomada como “entidades predefinidas à espera de serem pesquisadas” (ibid, p. 593).

Nessa perspectiva, se o fenômeno a ser investigado for observado como específico de um grupo de falantes que possuem características linguísticas em comum, mais bem caracterizado como rede social ou comunidade de prática, a delimitação da comunidade de fala deve se restringir a este grupo, seguindo a direção micro de análise.

Por outro lado, quando se trata de um fenômeno em variação mais amplo na língua, que esteja afetando o sistema como um todo, e se observa que essa variação não é restrita a grupos específicos, tais como os processos em variação que apontam para uma possível alteração de parâmetro na língua, entendo que a delimitação da comunidade de fala pode ser baseada tanto em limites geográficos como em critérios formais ou normativos, desde que a comunidade seja constituída por falantes nativos.

Além disso, deve-se notar ainda que um indivíduo pode pertencer a várias comunidades de fala, identificando-se com uma ou com outra de acordo com a situação. Portanto, considerando-se as diversas possibilidades de identidade assumidas pelos indivíduos, torna-se complexo restringir essa noção *a priori*.

Dessa forma, considerando ser este conceito uma elaboração metodológica, entendo que a comunidade de fala deve ser definida em conformidade com as necessidades e o momento da pesquisa, cabendo ao pesquisador delimitar o lócus do seu objeto de estudo que melhor se adéque ao fenômeno em investigação diante das questões do pesquisador. Ou seja, a comunidade de fala também precisa ser investigada, e não tomada como dada, e deve ser estabelecida no decorrer da pesquisa.

Ainda concernente aos pressupostos teóricos labovianos, destacam-se as linhas mestras referenciadas neste trabalho, quais sejam, o encaixamento e o uniformitarismo. A abordagem do princípio do uniformitarismo, que prevê que tendências hoje em curso devem ter atuado em estágios anteriores e possivelmente continuarão a atuar, será relevante diante da constatação de que a emergência de pronomes inovadores de 2ª pessoa para referência genérica tem sido demonstrada por línguas românicas em seu processo de mudança, como o Francês e o Espanhol, em substituição ao nome/ pronome que perdeu essa propriedade por estar sendo mais usado com referência específica para se referir a 1ª pessoa plural. Isso aconteceu no PB quando *nós* genérico passou a ser intercambiável com *a gente*.

Por encaixamento linguístico deve-se entender a forma como as mudanças observadas estão encaixadas na matriz de concomitantes linguísticos das formas em questão, e a quais outras mudanças se acham associadas. No caso da variação em estudo, parto do pressuposto de que a inovadora estratégia de indeterminação do sujeito *você_{gen}* está sendo empregada no PB como forma de preenchimento da categoria vazia da posição de sujeito das construções de sujeito nulo não referencial ou com referente genérico, como consequência, entre outras, do apagamento do clítico *se*, e do surgimento do paradigma de pronomes fracos na língua para compensar a perda do licenciamento do sujeito nulo. Esse pode ser um indício de provável encaixamento linguístico.

Caso este estudo constate que a variação em questão indica uma mudança em curso, o fato de estar ligada a um conjunto de mudanças mais radicais em andamento no PB também evidenciará encaixamento linguístico.

1.2 A SOCIOLINGUÍSTICA PARAMÉTRICA

Conforme demonstrado, na perspectiva da Sociolinguística laboviana, para se explicar a variação linguística é necessário correlacioná-la às diferenças existentes na estrutura social da comunidade de fala, por entender cada domínio, o social e o linguístico, como fenômenos estruturados.

A partir da constatação de que “muitos dos fatores considerados em análises variacionistas não são explicativos, mas distribucionais e organizadores no universo da gramática”, Kato e Tarallo, 1989 (p.11) propõem um elo entre a variação intra e interlinguística: a Sociolinguística Paramétrica; propõem, assim, o aparato de uma teoria gramatical para uma análise formal das variantes. No caso específico de variação sintática, uma teoria gramatical contribuirá para a explicação do fenômeno em análise.

Como este trabalho não se caracteriza por um trabalho sociolinguístico no sentido clássico, mas se propõe a seguir o modelo teórico da Sociolinguística Paramétrica, ele adota o mesmo caminho proposto por Kato e Tarallo (1989), aquele que “resgata a compatibilidade entre as propriedades paramétricas do modelo gerativo e as probabilidades do modelo variacionista, seja para provar seu espelhamento e reflexo, seja para realinhar um modelo em função do outro” (KATO E TARALLO, 1989, p. 5).

Ao proporem a Sociolinguística Paramétrica, Kato e Tarallo consideraram que os pressupostos do modelo paramétrico da Teoria Gerativa, que atua à base de princípios e não mais de regras e procura resgatar a variação interlinguística, se aproximam dos pressupostos da teoria da variação, posto que “também o variacionista está interessado em projetar, antecipar e aprofundar resultados cujo valor exceda os limites do intralinguístico para o universo do interlinguístico” (KATO E TARALLO, 1989, p. 7).

Este é o ponto onde as duas teorias convergem: a linguística de probabilidades prevê como um dialeto de uma determinada língua pode começar a realinhar as propriedades de seus parâmetros sintáticos. Ou seja, os resultados intralinguísticos podem ser úteis ao realinhamento das propriedades paramétricas previstas no modelo interlinguístico.

Na realidade, o que a sintaxe gerativa busca com a parametrização é dar conta da diferença entre as línguas. Como as diferenças entre as línguas se manifestam também em estudos realizados sobre a variação nas línguas, a Sociolinguística Paramétrica vale-se da variação dentro da língua para mostrar a variação entre as línguas.

Tanto na Gramática Gerativa quanto na Sociolinguística é possível evidenciar o reconhecimento de que “as línguas podem convergir em determinadas partes de suas gramáticas, revelando movimentos sincronizados e espelhados” (KATO E TARALLO, 1989, p. 19), ou seja, variações sintáticas observadas entre estágios diferentes de uma mesma língua são da mesma natureza de variações observadas entre línguas diferentes em um mesmo período dado. A esses momentos de generalização translinguística os gerativistas denominam de propriedades paramétricas.

É nesta perspectiva que Kato e Tarallo propõem a compatibilização da linguística de probabilidades com a de propriedades paramétricas por entenderem que os resultados de uma podem ser úteis ao realinhamento de outra, ou que um modelo pode ser realinhado em função do outro. A Sociolinguística Paramétrica, portanto, indica que o alcance dos resultados e o poder explanatório das análises via probabilidades e/ou propriedades são compatíveis entre si.

O alcance desta proposta é explicitado pelos autores a partir de três momentos. Em um primeiro momento, os autores consideram que são esses parâmetros que permitem reconhecer, por exemplo, que muitos dos fatores que condicionam a inversão do sujeito em línguas como o Francês canadense, o Espanhol mexicano e o Português carioca, atuam na mesma direção, a saber, a propriedade de haver um sistema produtivo de clíticos acusativos e o parâmetro *pro-drop*, ou +/- clítico acusativo e +/- suj0. Assim, tem-se um fenômeno de variação que interessa igualmente à Gramática Gerativa e à Sociolinguística, sendo que para esta o instrumento de análise é a avaliação da produtividade de duas formas sintáticas definidas previamente.

Um segundo alcance desta compatibilização diz respeito ao realinhamento de uma propriedade de um componente da gramática, a partir dos resultados probabilísticos sobre outro fenômeno variável, presente em outra parte da mesma gramática. Como exemplo, Kato e Tarallo citam, dentre outros, o trabalho de Naro (1981) sobre as restrições morfológicas ao apagamento do sujeito em Português, que apontam para a tendência de o PB perder as propriedades do Parâmetro do Sujeito Nulo.

De forma análoga, em estudos recentes sobre variação paramétrica, informações relativas à frequência de uma forma aparecem como um dos argumentos a favor de ter havido alteração do valor de parâmetro. A alta produtividade de uma forma em um determinado período de tempo e sua baixa produtividade em períodos subsequentes ou a sua não ocorrência são tomadas como evidência de alteração gramatical.

Um terceiro momento em que as duas teorias se encontram relaciona-se à possibilidade de um estudo variacionista fazer previsões de que o dialeto de uma língua em situações de contato pode começar a realinhar as propriedades de seus parâmetros sintáticos. Já uma

linguística típica de propriedades anteciparia que, no caso de contato entre duas línguas com as mesmas propriedades paramétricas, não haveria interferência nos moldes previstos por Weinreich (1953). Como exemplo, Tarallo e Kato citam a interferência sintática por contato entre o Português brasileiro e o Espanhol americano quanto ao parâmetro *pro-drop*. Um estudo variacionista mostra que o Português de fronteira obedece à mesma organização sistêmica do Espanhol, permitindo inclusive a ordem OVS, inexistente no Português da costa do Brasil.

Diante dos estudos apresentados por Kato e Tarallo (1989), em que as análises intra e interlinguística se complementam no sentido de realinhar os parâmetros sintáticos para um aprimoramento da análise linguística, constata-se a coerência do caminho por eles empreendido de atenuar o embate entre empiristas e racionalistas que dificultou que a linguística tivesse maiores alcances.

Na prática, o modelo da Sociolinguística Paramétrica propõe uma reavaliação dos estudos linguísticos de propriedades paramétricas que não levam em conta a heterogeneidade dentro de uma língua, com o objetivo de realinhar as propriedades paramétricas.

Indiscutivelmente, são muitas as contribuições da sintaxe gerativa para os estudos gramaticais das línguas em geral, os quais contribuem para o delineamento das propriedades paramétricas relativas, por exemplo, ao parâmetro do sujeito nulo – PSN. Por meio desses estudos, derrubaram-se pressupostos e formularam-se outros desde o estabelecimento do PSN, em 1981, por Chomsky.

Em meio a esses desdobramentos da proposta de Chomsky (1981), há o de que o Português do Brasil, dentre outras línguas, não constitui uma língua *pro-drop* prototípica, mas semi *pro-drop* ou *pro-drop* parcial. Essa caracterização baseia-se no fato de que o PB vem apresentando propriedades que o distanciam de línguas *pro-drop* e o aproximam de línguas não *pro-drop*.

De acordo com Holmberg (2010), as LSN consistentes diferem das caracterizadas como parciais, em que enquanto naquelas os pronomes sujeitos definidos são obrigatoriamente nulos em alguns contextos, nas parciais, além de não serem obrigatórios, eles só podem ser nulos se comandados localmente por um antecedente. Por outro lado, as LSN parciais permitem um pronome nulo indefinido na posição de sujeito, o que não é possível nas consistentes. Ambas LSN têm em comum o fato de que terem pronomes nulos não temáticos em expressões de tempo, sentenças de extraposição e sentenças existenciais e apresentacionais.

Deve-se ponderar, entretanto, que tais estudos não partem de uma concepção de língua em movimento, analisada em recortes sincrônicos, em correlação com outros estágios da mesma língua. Cabe avaliar, portanto, se tais línguas não estariam perdendo as propriedades e se

distanciando do parâmetro ao qual pertenciam, ou seja, passando por estágios de mudança de um parâmetro a outro, + *pro-drop* a - *pro-drop*.

Nesse contexto, confirma-se a importância dos estudos labovianos para dar uma nova luz a essas discussões. O modelo da Sociolinguística Paramétrica dá conta de explicar o fenômeno por que passam línguas que, face ao não atendimento a todas as propriedades do parâmetro *pro-drop*, têm sido caracterizadas como de sujeito nulo parcial por linguistas, como Holmberg (2010), Biberawer (2010), Roberts (2010), Kato (2000), dentre outros.

Se esses estudos levassem em conta a variação intralinguística, seria possível perceber que essas línguas estão passando por estágios de mudança, em vez de se postular que nem todas as línguas são “bem-comportadas” no que toca ao PSN, apresentando características mistas, ou ainda, que nem todas as propriedades observadas por Rizzi (1982 apud Holmberg, 2010), integram o PSN, tais como, se INFL pode ser especificado, é + pronominal, e o INFL que é + pronominal pode ser referencial.

Pesquisas empíricas que atestam a remarcação do PSN sob a perspectiva interlinguística têm sido empreendidas por diversos linguistas. O trabalho de Duarte (1995), sobre representação do sujeito pronominal de referência definida no PB, exemplifica a aplicação da Sociolinguística Paramétrica. A análise empírica mostrou que de um total de 1.424 dados, 71% apresentaram o sujeito pronominal expreso contra 29% de sujeitos nulos.

Diante desses resultados, a autora considera que se está diante de evidências empíricas que não mais permitem incluir o PB entre as línguas (+Sujeito Nulo). Dentre os fatores levantados, o estudo mostrou como relevantes para a realização da variável os grupos: a pessoa gramatical, a estrutura com o pronome de retomada com a mesma referência do sujeito da principal e a faixa etária dos entrevistados. O levantamento dos fatores revela ainda que os contextos em que o sujeito nulo ainda mostra alguma resistência são a 3ª. pessoa (em relação à 1ª. e 2ª. pessoas, em que a mudança se mostra quase completa), a existência de correferência entre o sujeito da subordinada e da principal, o traço [-animado] do referente e o grupo etário mais velho.

É dessa forma que o modelo propicia um estudo empírico mais significativo das línguas, que serve de subsídio para uma análise interlinguística, a partir do fenômeno estudado, além de prover dados a respeito da produtividade do fenômeno.

Assim, um estudo comparativo entre o PB e o PE, que objetiva detectar em que aspecto as duas variantes se diferenciam no que diz respeito às inovadoras formas de indeterminação do sujeito e, conseqüentemente, ao preenchimento do sujeito nulo, poderá apresentar contribuições para as discussões sobre a proposta de Tarallo (1993) de que o PB e o PE seguem

por caminhos divergentes, ao mesmo tempo em que subsidiará a análise interlinguística, conforme o modelo teórico a que me propus seguir.

1.2.1 A compatibilização de teorias em questão

Ressalvas à proposta de conciliação desses dois modelos, a Sociolinguística laboviana e a Teoria Gerativa, são apresentadas por alguns linguistas, dentre os quais, destaca-se o argumento da incomensurabilidade da compatibilização, apresentado por Borges Neto (2004). A empreitada de Borges Neto de combater veementemente a proposta de Tarallo e Kato de compatibilização das teorias de probabilidades e de propriedades paramétricas apresenta tanto argumentos fundamentados como não fundamentados.

A crítica do autor é construída a partir da noção de incomensurabilidade de teorias, tendo em vista a motivação ideológica que subjaz a elas, passando por um trabalho prévio de exegese da proposta da sociolinguística paramétrica para, no final, apresentar o construto teórico dessa proposta.

Nesse percurso, o autor antecipa algumas inferências a partir de expressões empregadas no texto de Tarallo e Kato, tomadas de forma descontextualizada, para depois refutá-las. Como exemplo, as relativas à pressuposição de que dizer que há compatibilidade entre teorias significa dizer que as motivações ideológicas e os procedimentos metodológicos são os mesmos, que é possível reduzir uma teoria à outra e que seus objetivos são complementares. Nessa perspectiva, o autor sugere que a afirmação de Tarallo de que a proposta se trata de “ênfatar a complementaridade entre os modelos” induz ao entendimento de que “o pretendido no texto é demonstrar que as duas teorias são complementares” (Borges Neto, 2004, p. 199).

Pode-se perceber um equívoco no emprego do termo ‘complementar’ para significar compatibilização, uma vez que o termo não condiz com o que Tarallo e Kato propõem. A partir de uma citação de Oliveira (1986), Borges Neto faz uma descrição coerente da proposta, mas não sem deixar de apresentar graves críticas ao modelo laboviano, ao apontar que há falhas na sociolinguística, as quais, para serem eliminadas, de acordo com o autor, devem receber descrições estruturais, e que “é mais fácil e econômico usar as descrições já existentes” (Borges Neto, 2004, p. 202). Ora, cada modelo teórico tem seus pressupostos e propósitos; não se pode exigir de um modelo teórico-metodológico ter um aparato descritivo, assim como não se espera de um modelo formal, como o gerativismo, que considere em sua teoria aspectos sociais.

Não é escopo da Teoria Gerativa investigar a variação intralinguística, mas entre as línguas, daí os parâmetros. Como a variação na língua e os fatores condicionadores da variação indicam que pode estar havendo alteração em um parâmetro, esses resultados podem ser úteis para o realinhamento do parâmetro, ou melhor dizendo, para se entender por que uma ou mais propriedades de determinado parâmetro em uma língua não condizem com as de outra língua que é caracterizada pelo mesmo parâmetro, valendo-se, para tanto, do poder explanatório do quadro teórico da gramática gerativa. É nesse sentido que se deve entender a compatibilização, entre resultados, sem tocar em conceitos e ideologias inerentes a cada vertente teórica. Não é o caso de, conforme sugerido por Borges Neto, “preencher lacunas”, até porque cada proposta tem o seu referencial, o seu modelo de abordagem.

Outra inferência de Borges Neto que parece equivocada diz respeito à sua afirmação de que para Almeida (1989) “a sociolinguística tem condições de “completar” a gramática gerativa” (Borges Neto, 2004, p. 51), a partir da afirmação de Almeida (1989) de que “a gramática gerativa também apresenta lacunas (ou falhas): falta-lhe um tratamento para a variação e a mudança e falta-lhe o encaixamento linguístico no social”.

Entretanto, o que se observa nessa passagem de Almeida (1989) é a defesa da contribuição da ciência linguística social para o acervo do conhecimento sobre a linguagem. Não se percebe exatamente uma crítica do autor aos trabalhos chomskianos, mas a observação de que o estudo dos fenômenos linguísticos não poderia se resumir à descrição da competência. Como se pode constatar em sua fala, Almeida (1989) enfatiza a importância de cada uma das ciências para os estudos linguísticos.

Na constatação e explicação da variação, de sua estabilização e/ou de seu processo de desenvolvimento rumo à mudança, levando-se em conta tanto as motivações linguísticas quanto as extralinguísticas — ou sociais —, reside, sem dúvida, a grande contribuição da ciência linguística social — a Sociolinguística — para o acervo do conhecimento que o estudo do fenômeno linguagem não se poderia resumir à descrição e explanação da competência linguística — tal como a define Chomsky —, por mais completos, exaustivos e convincentes que sejam tais procedimentos. (ALMEIDA, 1989, p. 86).

A proposta de compatibilização de resultados não tem o objetivo de “complementar” as teorias nem “aproximá-las”, mas de utilizar os resultados de uma para o realinhamento dos resultados de outra. Daí ser inconcebível articular sobre a incomensurabilidade das teorias, posto que Tarallo e Kato (1989) não sugeriram que elas seriam comensuráveis. Ao propor a compatibilização entre as propriedades paramétricas do modelo gerativo e as probabilidades do

modelo variacionista, eles acreditaram em um direcionamento mútuo entre a variação intra e inter-linguística, enfim em um modelo que permitisse compatibilizar os resultados de pesquisas empreendidas em ambas as teorias, em relação às questões no âmbito da teoria gerativa com questões que interessavam à teoria da variação.

Em trabalho realizado sobre os avanços no estudo da mudança sintática, Duarte (2015) aborda a importância do componente gramatical associado ao componente não-linguístico, ou seja, a associação da Teoria da Variação e Mudança a uma teoria linguística, seja funcionalista, fonológica, morfológica, semântica ou sintática, o que, de acordo com a autora, constitui condição *sine qua non* para levar a efeito a aplicação do modelo teórico para estudar a mudança.

Dentre os autores que defendem essa perspectiva, Duarte cita Pagotto (2006), para quem “qualquer abordagem que se diga sociolinguística e não se proponha a dar uma resposta satisfatória à questão de como a estrutura linguística se engendra na estrutura social está, de saída, claudicante com relação às tarefas que propõe investigar” (op. cit: 52).

De acordo com Duarte (2015), quando essa compatibilização foi proposta por Kato e Tarallo, na década de 1980, a noção de parâmetros era ainda incipiente, e a teoria Gerativa, então focada na busca de princípios universais, não se interessava pela mudança linguística, seria natural o espanto diante da tentativa de utilizar uma teoria baseada na “competência” linguística para analisar os dados da língua-E, a “performance” dos indivíduos de uma comunidade de fala, na busca de evidências para a mudança paramétrica em curso.

Hoje, diante dos vários trabalhos realizados na perspectiva dessa proposta, pode-se constatar que muitas questões e correlações atestadas na observação da sintaxe do PB teriam passado despercebidas. Assim, é possível inferir que se Borges Neto tivesse produzido este texto nos dias atuais, talvez sua crítica tivesse sido menos intensa.

Outra questão que está em causa na polêmica gerada pela compatibilização do modelo da Teoria da Variação e Mudança com a Teoria Linguística Gerativa diz respeito à interpretação da variação como inerente ao sistema ou como formas de diferentes gramáticas em competição, ou seja, qual perspectiva teórica sobre a mudança linguística deverá ser adotada para a análise do fenômeno em variação.

Conforme mostrado por Duarte (2015), quando se observa uma mudança em curso, a proposta de gramáticas em competição se depara com alguns problemas. Henry (2002 apud Duarte, 2015) questiona o modelo de Kroch (1994) de competição de gramáticas. Mais especificamente em relação à análise do aparecimento do *do* perifrástico em Inglês, o próprio Kroch enfatiza que a mudança não é instantânea, envolvendo a alternância de formas durante um longo período de tempo, o que sugere que gerações sucessivas adquiririam regras variáveis

que a cada geração evoluíam a favor da variante nova até a completude da mudança. A inconsistência de Kroch, para Henry (2002), parece estar em apresentar dados em variação e concluir que não se trata de verdadeira opcionalidade mas de gramáticas em competição.

Outra questão colocada por Henry (2002) em relação à proposta de gramáticas em competição como perspectiva teórica para a análise da mudança linguística, diz respeito à variação nos estágios iniciais da aquisição. Se as gramáticas fossem invariáveis, então era de se esperar que as crianças adquirissem uma única forma e apenas mais tarde uma outra variante. Entretanto, em estudos sobre a aquisição do inglês em Belfast, com crianças entre 2 e 4 anos, Henry (2002) encontra evidências de que as crianças não somente adquirem regras variáveis mas são sensíveis ao *input* a que estão expostas.

Diante desses fatos, o posicionamento de Duarte (2015, p. 94) de que embora as discussões envolvendo a interpretação da variação encontrem argumentos em defesa de uma e outra interpretação, eles não invalidam os resultados de estudos realizados à luz de um e outro modelo de mudança, será endossado nesta tese.

Tendo em vista que um dos objetivos desta tese é analisar a difusão e evolução da variação na comunidade de fala, em tempo aparente e real de curta duração, a teoria da mudança adotada é na perspectiva da variação como uma propriedade inerente ao sistema. Assim, este estudo irá mostrar que descrições linguísticas deveriam levar em conta o fato de que a mudança paramétrica implica um longo período de variação antes que uma nova opção seja estabelecida.

2. O PARÂMETRO *PRO-DROP*

À medida que os estudos gerativistas se implementaram no sentido de se identificarem as propriedades relacionadas ao licenciamento do sujeito nulo, ou parâmetro *pro-drop*, a noção de parâmetro passou por refinamentos. Na teoria gerativa que prevê princípios universais comuns a todas as línguas e parâmetros como possibilidades binárias em aberto associadas com a variação linguística, a escolha de um valor em um parâmetro representa a diferença entre sistemas gramaticais. No entanto, estudos recentes mostram que outros parâmetros podem estar afetando a realização de pronomes nominativos, com a proposição de mais categorizações para as línguas no tocante ao parâmetro *pro-drop*, como as semi *pro-drop* ou *pro-drop* parciais e as consistentes. Dentre os vários autores que estão discutindo o comportamento dessas línguas, apresento as propostas de Kato (2000) e de Holmberg (2010), nas quais as discussões e hipóteses desta tese se referendam.

2.1 A proposta de Kato (2000)

Na introdução do livro *Brasilian portuguese and the null subject parameter*, organizado por Kato e Negrão (2000), Kato (2000) mostra que a variação paramétrica foi recentemente reconceitualizada como uma função do léxico, a partir da concepção de que ela reside nas propriedades lexicais das categorias funcionais. O parâmetro do sujeito nulo não é mais conceitualizado como uma diferença sintática particular entre as línguas, no sentido de elas permitirem ou não a omissão do sujeito pronominal. O parâmetro passa a ser entendido como uma propriedade morfológica abstrata do núcleo funcional. É postulado, assim, que a variação interlinguística é devida à morfologia, sendo a estrutura gramatical fundamentalmente a mesma para todas as línguas.

O desafio passou a ser o de interpretar a natureza de uma propriedade morfológica abstrata e caracterizar o conjunto de propriedades relacionadas com as línguas de sujeito nulo, levando a um avanço teórico.

Kato (2000) objetiva apresentar uma teoria do parâmetro do sujeito nulo, de acordo com Kato (1999), como sendo um feixe de propriedades que inclui a possibilidade de sujeitos nulos e a inversão livre, as quais podem ser derivadas da mesma propriedade do sistema de concordância, ou do mesmo parâmetro. A proposta teórica de Kato (1999) tem por base os estudos de Rizzi (1982) e de Burzio (1986), que apontam que INFL em LSNs é uma espécie de

pronome ou clítico, e de Everett (1996) que propõe que afixos de concordância, clíticos e pronomes são realizações alomórficas dos traços-*phi*. A proposta de Kato (1999) estende este ponto de vista para a variação interlinguística: para a mesma função as línguas escolhem uma dessas formas para o pronome nominativo. Esta análise elimina *pro* referencial como uma categoria descritiva, tendo em vista que *pro* era projetado para valorar os traços de *Agr*.

Diante das afirmações de que “o PB está exibindo mudanças em progresso em ambas as propriedades”¹³, quais sejam, a perda do sujeito nulo e a perda da inversão livre, e de que “o PB está no caminho para se tornar uma língua não SN)”¹⁴, é possível inferir que Kato (2000) admita implicitamente que o PB está passando por processo de mudança paramétrica, embora a perspectiva do texto seja teórico-sincrônica com o aparato da gerativa. Há de se considerar ainda que a teoria de pronomes fracos, adotada e refinada por Kato (1996), tem como pressuposto que o surgimento do paradigma de pronomes fracos constitui uma das propriedades que identificam LSN, e que em LSN típicas ou consistentes não há formas fracas nem no paradigma pronominal nem como pronome locativo.

Para dar sustentação à sua proposta, Kato (2000) apresenta alguns estudos desenvolvidos sobre o tema, como os de Duarte (1993) e Galves (1993). Conforme mostrado por Duarte (1993), o decréscimo de sujeitos nulos no PB tem sido impulsionado pela substituição da 2ª pessoa “tu” pela forma de tratamento em sua forma reduzida “você”, a qual seleciona concordância de 3ª pessoa. Kato (2000) nota que, embora o desenvolvimento de um paradigma de concordância pobre dê sustentação à hipótese de morfologia rica como determinante do Parâmetro *pro-drop*, tal relação é invalidada pelo fato de que a 1ª pessoa, a única ainda marcada com morfologia distinta, ter sido o primeiro sujeito a se tornar mais frequentemente expresso.

A partir dos resultados quantitativos de Duarte (1993), Kato (2000) afirma que o fato de o PB contemporâneo favorecer o preenchimento do sujeito, particularmente nos contextos em que há morfemas de 1ª. pessoa, serve de sustentação adicional a favor da hipótese de que, de fato, *Agr* não é um fator preponderante no acionamento do sujeito nulo. Além disso, ainda que a 2ª pessoa do singular indireta *você* e a 3ª pessoa do singular tenham a mesma flexão, a 2ª pessoa tem o pronome realizado mais frequentemente do que a 3ª.

¹³Original: “*Brasiliian Portuguese (BP) has been exhibiting a change in progress in both properties.*” (p.223)

¹⁴Original: “*PB is on its way to becoming a non-NSL language.*” (p. 232).

Galves (1993 *apud* Kato, 2000) interpreta os fatos empíricos de Duarte (1993) como uma mudança no sistema de concordância. A autora usa AGR como um núcleo frasal e o princípio geral de economia, alegando que AGR é projetado somente quando a língua tem uma morfologia de concordância forte (rica).

A análise de Kato (1999) adota o ponto de vista de Galves (1993) de que a concordância em LSNs não forma uma categoria sincrética com tempo, entretanto, discorda do fato de que a flexão de concordância seja o núcleo de uma projeção oracional, mas do DP, o qual se funde com o verbo como seu argumento.

Esta é a base da proposta de Kato (1999), sendo o *Agr* pronominal sintaticamente definido como uma flexão de concordância que aparece na Numeração¹⁵ como um item independente do verbo. Pode ser considerado um *Agr* pronominal qualquer determinador portador de traços-*phi*¹⁶, que pode ser um pronome livre (fraco), clítico ou afixo pronominal, começando a derivação de uma maneira similar, sendo fundido na posição D.

Kato (1999) propõe que, da mesma forma que pronomes fracos e clíticos, esses afixos têm caso e traços-*phi*. A diferença entre pronomes fracos, clíticos ou afixos pronominais está em que pronomes ficam em Spec de T e clíticos e afixos se movem como núcleos. Spec de T é projetado somente no primeiro caso.

Nessa perspectiva, em línguas como o Inglês, de sujeito não nulo, o pronome livre sobe para SpecTP. Já em línguas como o Espanhol, prototipicamente de sujeito nulo, o sujeito clítico e afixos pronominais são atraídos e adjungidos a T. T tem seu caso nominativo eliminado depois da checagem, mas os traços-*phi* do *Agr* pronominal são retidos.

Esta proposta de Kato (2000) assenta-se na teoria do surgimento de formas pronominais fracas¹⁷, inicialmente proposta por Kato em 1996. Nessa sua proposta de pronomes fracos e fortes no PSN, Kato (1996a) postula que, como um item lexical independente na Numeração, *Agr* pronominal pode checar *uD* e traços de Caso, e Spec de T não é projetado. Se *Agr* forte pode desempenhar o papel de checagem de EPP em I, esta análise exclui *pro* como categoria descritiva uma vez que *pro* era projetado para valorar os traços de *Agr*. Assim, sem *pro* na posição Spec de T, esta posição pode ser ocupada por um determinante portador de

¹⁵ Numeração: conjunto de pares (LI, *i*), onde LI é um item do léxico e *i* um índice correspondente ao número de vezes que aquele item será utilizado na derivação. Assim, os elementos de léxico selecionados da Numeração são combinados recursivamente. Uma derivação linguística parte do conjunto de itens lexicais que constitui a Numeração. (Chomsky, 1995, p.225).

¹⁶ Os traços formais são de dois tipos, os com motivação semântica - traços-*phi* e categoriais, e os estritamente formais – Caso e EPP. São traços *phi* (gênero, número e pessoa), traços categoriais (N e V, por exemplo), e traços estritamente formais (traço de Caso e traço EPP - do inglês *Extended Projection Principle*).

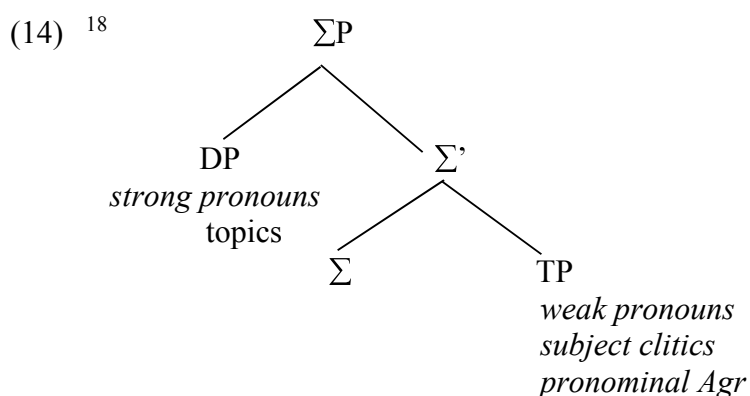
¹⁷ A teoria de pronomes fracos será detalhada na seção 3.2.

traços-*phi*. Entretanto, Kato (2000) afirma que em línguas como o PB que retêm o sujeito nulo não referencial, *pro* é mantido para sujeitos nulos arbitrários e anafóricos.

Em relação aos traços dos pronomes fortes e fracos, Kato (1999) assume que assim como DPs deslocados, pronomes fortes têm o caso nominativo “default” em LSNs. Uma vez que eles estão na relação núcleo de SpecTP com Σ P, pode-se considerar que o caso é checado nesta posição mais elevada, fora dos domínios T ou AGR.

Portanto, para Kato (1999), da mesma forma que em Francês, no PB o pronome forte bem como qualquer tópico DP estaria em Spec de Σ , enquanto o fraco estaria em SpecTP. Assim, quanto à derivação, em LSNs os pronomes livres fracos se movem para SpecTP para verificar D e traços-*phi* de T. Em LSNs, o afixo pronominal também move para posição acima, adjacente ao T da mesma forma que clíticos passam de DP objetos para T, onde verificam Caso e traços-*phi*. A diferença entre clíticos e afixos encontra-se somente na direção de adjunção.

A representação da derivação para o sujeito assumida por Kato (1999) é a que se segue:



A postulação de posições idênticas para o tópico e pronomes fortes baseia-se não somente na visão de que a variação é devida à morfologia e de que a estrutura gramatical é fundamentalmente a mesma para todas as línguas, mas também em contrastes empíricos, conforme os exemplos abaixo, extraídos de Kato (1999):

- (15) a. ME, I love wine.
 b. MOI, j’aime de vin.
 c. YO amo vino.

¹⁸ Exemplo (14) foi extraído de Kato (1999).

Esta proposta sustenta ainda a de que sujeitos duplos em LSNs não é um fenômeno que envolve um *pro* silencioso, mas o próprio afixo de concordância. Assim, enquanto o Inglês dobra o pronome fraco (16)a. e o Francês dobra o clítico (16)b., o Espanhol dobra a própria concordância (16)c., conforme mostrado nos exemplos abaixo:

- (16) a. Me, I ¹⁹
 b. Moi, jê
 c. Yo, V+agr

Para explicar a realização de pronomes sujeitos em LSNs, Kato (1999) propõe que essas línguas têm dois tipos de nominativo: o nominativo ‘*default*’ do DP lexical, o qual não precisa ser checado, e o nominativo com traços de *Agr* pronominal, o qual tem de ser checado contra os de T. Assim, sem *pro* ocupando a posição de SpecTP em LSN, é possível que esta seja ocupada por um pronome, um caso de duplicação do sujeito.

Comparando a), b) e c) em (16), Kato (1999) confirma a proposta de que o pronome forte tem um caso *default*, sendo o nominativo o *default* em línguas de sujeito nulo românicas, e que (16)c. é na realidade a configuração do caso nominativo pela concordância em termos da relação de núcleo de SpecTP.

Assim, em Espanhol, LSN prototípica, o pronome fraco é a própria concordância, e na estrutura (17) ele aparece como um pronome resumptivo do elemento lexical em ΣP . Esse DP lexical tem as propriedades de um nome deslocado: ele é definido ou específico, tem o Caso nominativo *default* e é interpretado como o sujeito da assertiva contida em TP. A explicação para esse DP ocupar a posição ΣP está em que, uma vez que a concordância pronominal desempenha todas as operações com TP, os elementos lexicais (no caso *yo*) são adjungidos no núcleo acima de TP da mesma forma que os casos normais de deslocamento à esquerda.

- (17) (ΣP *yo*_i (TP *ló com-i*_i (VP...))

Em relação às categorias que podem checar Caso e traços-*phi* do V+T no PB, Kato (2000) esclarece que são o recém-formado pronome nominativo fraco ou um DP cheio. Assim, o surgimento de formas fracas, que para Kato (2000) constitui uma das propriedades que identificam LSNN, pode ser visto no PB, em contextos nos quais a língua licenciaria sujeito

¹⁹ Os exemplos (16) e (17) foram extraídos de Kato (2000, p. 230 e 231).

nulo, como uma estratégia para compensar a perda da propriedade de licenciar sujeito nulo e para satisfação do EPP da sentença. Nos termos de Kato (1999, p. 6), “a perda de *Agr* pronominal correlaciona-se com o aparecimento de pronomes sujeitos fracos”. Ou seja, o surgimento de formas pronominais reduzidas no PB, quase homófonas às formas pronominais plenas, pode ser interpretado como sendo o reflexo da maneira como o PB contemporâneo permite a satisfação do EPP da sentença.

Assim, a teoria do surgimento de formas pronominais fracas, ao pressupor que um dos estágios de perda das propriedades de sujeito nulo correlaciona-se com o surgimento de um paradigma pronominal fraco em competição ou substituição ao paradigma forte, por si só, estabelece uma distinção entre LSN plenas e parciais.

Nesse trabalho, Kato (2000) busca explicar como esta teoria de pronomes fracos correlaciona com a possibilidade da inversão livre em LSN de uma forma natural. A inversão livre é definida como a possibilidade de o sujeito pós-verbal aparecer imediatamente após o verbo e antes do objeto ou após todo o predicado. Kato (2000) mostra que, ainda que Adams (1987) relacione a perda de sujeito nulo no Francês Antigo (doravante FA) à perda de sua ordem VSO, o padrão de inversão que aparece mais generalizado em línguas de sujeito nulo é VOS.

Conforme demonstrado por Kato (2000), as mudanças que ocorreram no PB são vistas contra a teoria do sujeito nulo e inversão propostas até então. Diferentemente do ocorrido em FA (correlação entre a perda do sujeito nulo e a ordem VSX), o PB, conforme atestado pelo estudo de Duarte e Kato (1998), manteve o sujeito nulo após a perda do movimento longo do verbo em interrogativas, sendo que a perda do sujeito nulo se correlaciona com a perda da ordem VOS, ou inversão livre.

De acordo com a análise de Kato (1999), o PB perdeu seu sujeito nulo referencial, e um pronome fraco nominativo apareceu no lugar do sistema pronominal *Agr*, o qual costumava ser identificado como sujeito gramatical. E esse pronome fraco está se cliticizando para compensar *Agr*. O nulo que resta no PB, de acordo com Kato (2000) é o expletivo, analisado como neutro \emptyset_2 . O \emptyset_2 não-argumental é o único afixo que ainda pode aparecer como um item independente na Numeração e pode ser fundido com V+T para checar seus traços-D. Todos os outros afixos de concordância são agora parte de entrada de V e, em seu lugar, o pronome fraco aparece. Isso se relaciona com a perda da inversão livre da seguinte forma.

Kato (2000) advoga que a perda das propriedades da inversão é meramente uma consequência do surgimento do paradigma flexional fraco. Como a concordância não é mais analisada como (+ pronominal), ela não pode aparecer como um item independente do verbo na numeração. As categorias que podem checar caso e traços-*phi* do V+T são o recém-formado

pronome nominativo fraco ou um DP cheio. Os pronomes sujeitos fracos, similarmente ao pronome nominativo em Inglês, requerem não somente a projeção do SpecTP, mas um traço D-forte em T, produzindo a ordem SVO. O PB apresenta assim derivação semelhante ao Inglês.

Assim, a inversão livre com sujeitos definidos parece ser impossível em línguas como o Inglês porque Spec de T é projetado. A fim de deixar o sujeito na posição nuclear canônica de ênfase no final da sentença, T', uma projeção não máxima, tem de se mover. O movimento do predicado, ou o TP, poderia constituir uma operação ilegal, uma vez que ele moveria uma projeção intermediária. O Inglês tem uma estratégia diferente de focalizar o sujeito, a qual é interna ao TP: focalização *in situ*. Pode ser este o caso do PB.

Comparando a inversão livre com concordância no PE e PB, Kato (2000) mostra que no PE a ordem VS tem uma representação na qual a projeção máxima TP é adjacente a ΣP. Uma vez que a concordância é pronominal, é o argumento afixal que se adjunge a T para checar seu Caso nominativo e traços-*phi*. Como o verbo que alça para T tem somente traços de tempo, os traços de concordância ficam retidos em *Agr*, uma vez que eles são traços interpretáveis. Spec de T não é projetado e o movimento do TP é possível.

Já no PB, a concordância não é pronominal e é parte da flexão verbal na numeração. O que é incorporado como argumento do verbo é o DP lexical, o qual sobe para Spec de T para checar seu Caso nominativo e traços-*phi*. O Caso nominativo do DP lexical no PB não é um Caso 'default' como no PE e precisa ser checado. O verbo flexionado sobe para T para checar os traços fortes de V. O constituinte que precisa de movimento para fornecer a ordem VS é uma projeção intermediária no caso, e, portanto, a ordem VS é descartada.

Resumindo, a perda da inversão livre com DPs definidos deriva do fato de que pronomes nominativos fracos foram criados, e eles têm de ocupar Spec de T. Uma vez que SpecTP foi projetado, a focalização do sujeito não pode ser obtida pela inversão do sujeito. Em vez disso, o que temos no PB é o processo encontrado nas línguas inglesa e alemã: focalização *in-situ*, conforme exemplo (18) do Inglês:

- (18) a. *JOHN ate an apple.*
b. O JOÃO comeu a maçã.

Os exemplos em (19) e (20) mostram que o PB atual tem um padrão para duplicar o sujeito similar ao Francês em seu estágio entre Francês Antigo e Francês Médio (*Jou, je agr* [-pronominal]), ou seja, o pronome forte é o nominativo. Já em relação ao Inglês, a diferença está

no fato de que, enquanto no PB os pronomes fracos e fortes são quase-homófonos porque o Caso *default* é o nominativo, em Inglês o Caso *default* é o acusativo.

- (19) a. Ele, ele é meu amigo.²⁰
b. *Him, he is my friend.*

- (20) a. Você, cê é meu amigo.
b. *You, you are my friend.*

O que ocorre com a inversão livre com DP indefinidos no PB é diferente. Eles são reanalisados como construções expletivas, sendo sua derivação similar aos acusativos e existenciais em Francês. Kato (2000) mostra que, apesar da crescente perda da inversão livre no PB, a língua falada retém uma ordem VS produtiva em construções existenciais e inacusativas, conforme exemplos (21) e (22). Essas construções exibem comportamento unificado quando a concordância está em causa, como pode ser visto em (22), na qual o verbo não concorda com o associado.

- (21) Tinha gatos de todos os tipos.

- (22) Chegou uns ovos.

Kato (2000) dá a seguinte explicação para estas construções no PB: elas eram mantidas com o verbo existencial invariavelmente no singular, e o verbo do tipo *ter* foi proposto para checar o Caso acusativo. Se o único argumento é o acusativo, ele não pode checar o traço nominativo de V+T. O que está imerso com V+T no ponto de checagem é o afixo neutro \emptyset_2 . Então, a construção existencial VS foi mantida porque o afixo neutro ainda é pronominal, devido ao fato de que nenhum expletivo lexical fraco apareceu. Como no PB é suposto que o associado tenha o Caso nominativo *default* (o qual por definição é o caso não checado), o afixo neutro \emptyset_2 é necessário para eliminar traços-*phi* de T, e o resultado é a falta de concordância.

Embora a ordem VOS possa ser naturalmente derivada em LSNs, mas não em LSNNs, com relação às construções inacusativas e existenciais, Kato (2000) mostra que elas apresentam o mesmo padrão em muitas línguas. Em Inglês, uma LSNN, construções inacusativas com verbos apresentativos (*arrive, appear*) e construções existenciais com *be* exibem

²⁰ Exemplos de (18) a (22) foram extraídos de Kato (2000, p. 252).

comportamento similar com relação à concordância: o verbo flexionado concorda com o nome pós-verbal, referido na literatura gerativa como expletivo associado.

(23) *There **are** cats under the table.*²¹

(Tem gatos embaixo da mesa.)

Ainda que o Inglês e o PB tenham em comum o traço-D forte em T, o que requer projeção de SpecTP, diferentemente do PB, em Inglês o locativo satisfaz o EPP, mas Caso e traços-*phi* não podem ser checados. Por isso, o Inglês recorre ao alçamento do FF do associado desencadeando concordância nominal.

Entretanto, esta não é característica apenas de LSNN: o Italiano, uma LSN, apresenta comportamento semelhante ao do Inglês em construções existenciais e inacusativas: o verbo flexionado concorda com o associado em ambos os casos.

(24) *Ci sono dei gatti sotto il tavolo.*

Já em Francês, inacusativos e existenciais também exibem comportamento similar onde a concordância está em causa, mas ao contrário do Inglês e do Italiano, não há nenhuma relação de concordância entre o verbo e o nome pós-verbal. O verbo concorda é com o expletivo *il*, conforme se vê em (25):

(25) *Il y a des chats sous la table.*

Kato (2000) aponta que em LSNs, como o Espanhol e o PE, as construções existenciais são mais similares às inacusativas e existenciais em Francês, no sentido de que o verbo perde concordância com o associado nas primeiras e o auxiliar é do tipo *have* nas existenciais. Mesmo quando nessas construções aparece um pronome fraco na posição de sujeito, não se estabelece concordância com o associado. Em Português, em vez de postular um *pro* nas construções existenciais, Kato (2000) alega que é o afixo de 3ª. pessoa que faz o papel de *il*. Mas ainda que faça parte da numeração, o afixo de 3ª. pessoa não tem papel de argumento, portanto ele adjunge diretamente a *ter* (V+T) para checar Caso e traços-*phi*.

²¹ Exemplos de (23) a (29) foram extraídos de Kato (2000, p. 243).

Antes de o PB perder a concordância na inversão, ele assemelhava-se ao Italiano. Conforme mostrado em (26)a., o associado tem propriedades vinculadas e controladas ligadas à sentença encaixada. Com a perda da concordância, essas propriedades são perdidas também, e o PB agora se assemelha com o Francês, como pode ser visto em (26)b.

- (26) a. Entraram três homens sem se identificar. (Italiano)
b. Entrou três homens sem se identificar. (Francês)

Com relação ao papel do locativo em existenciais, Kato postula que o locativo é uma expressão significativa e que requer checagem. Construções existenciais em geral exibem um elemento locativo fraco, mesmo se a língua não tem o traço D-forte em T, como o *ci* em Italiano, e afixo (*ha-y*) em Espanhol. Kato mostra que as formas locativas fracas se distribuem entre as línguas como os pronomes pessoais fracos, e que essas formas podem ser dobradas pelo PP, podendo tanto alçar para SpecTP (como o *there* em Inglês) ou adjungir a T como clíticos ou afixos (*y* em Francês). Assim, similarmente a estruturas envolvendo pronomes pessoais, Kato sugere que o pronome locativo forma um DP com o complemento PP, sendo que a forma fraca que move para TP, conforme mostrado em (27).

- (27) [_{DP} *There* [_{PP} *under the table*]].

Entretanto, Kato observa que o Português não tem nenhum clítico ou afixo com a função de locativo, e que, portanto, não ocorre alçamento de qualquer tipo.

Após mostrar que o PB se tornou similar ao Francês também em relação à construção inacusativa em que ambos perderam concordância, Kato (2000) aponta que, entretanto, há notáveis diferenças entre as duas. O Francês tem um expletivo claro, *il*, e o PB tem um afixo nulo. O Francês tem um locativo claro, *y*, em existenciais, mas não tem locativos em inacusativos. O PB não tem locativo pronominal fraco em nenhuma dessas construções. O Francês tem o auxiliar *être* com o verbo inacusativo e o verbo *avoir* com o existencial. O PB usa o *ter* como verbo existencial e como auxiliar em formas perifrásticas de verbos inacusativos, conforme exemplos (21) e (22).

Ao concluir sobre as semelhanças entre as mudanças em curso no PB e as ocorridas em FA, Kato (2000) sintetiza os três aspectos de suas gramáticas que as mantêm separadas: o abundante uso de clíticos em Francês, a lexicalização do expletivo, e o uso de *avoir* somente para existenciais.

Quanto ao expletivo lexical, Kato (2000) infere que o PB ainda pode adquiri-lo, mas entende que a preferência para lexicalizar a posição do sujeito tem escolhido um caminho diferente, como mostram Duarte (2000) e Viotti (1999). Viotti (1999) considera o transitivo *ter* e o existencial *haver* como o mesmo item lexical e mostra que a preferência pelo verbo *ter* em vez de *haver* autoriza uma construção pessoal como em (28), com o alçamento do locativo para a posição de sujeito, uma operação não disponível para *haver* (29).

(28) Esta cidade **tem** gatos abandonados.

(29) ? Esta cidade **há** gatos abandonados.

Para Duarte (2000), uma vez que o PB perdeu a omissão obrigatória do sujeito referencial, é de se esperar sinais de presença dessa mudança na representação também em sujeitos não referenciais. Algumas das estratégias de preenchimento do nulo não referencial que têm sido observadas no PB, citadas por Duarte, são: i) sentenças existenciais têm sido substituídas por estruturas com *você_{gen}* acompanhando *ter*; ii) verbos de alçamento²² muito frequentemente têm aparecido em estruturas com dois verbos flexionados, como em (30)a.; iii) complementos de sentenças infinitivas são usualmente movidos para a posição vazia do sujeito da sentença principal, como em (31)a. e; iv) verbos de um argumento podem aparecer com dois argumentos, como em (32)a. ; v) sentenças de extraposição²³ como em (33)a. são sempre retomadas pelo demonstrativo *isso* em uma típica construção de deslocamento à direita.

(30) *pro_{exp}* Parece que vocês não pensam a sério na vida.²⁴

(30)a. *Vocês_i* parecem *pro_i* não pensam a sério na vida.

(31) *pro_{exp}* Vale a pena salvá-los.

(31)a. Eles valem a pena ser salvos.

(32) *pro_{exp}* Era em torno de dez pessoas.

(32)a. Isso era em torno de dez pessoas.

²² As construções a que a Teoria Gerativa se refere como de alçamento envolvem o alçamento de um DP, sujeito de uma predicação, em que não pode receber Caso do seu predicador (um adjetivo, um infinitivo não flexionado, por exemplo), para uma posição em que a atribuição de caso seja possível. Os verbos *parecer* e *acabar*, nas sentenças em 1 a) *As professoras_i acabaram t_i estressadas*, e b) *Os mestrandos_i parecem t_i assustados*, tomam como argumento interno uma estrutura que contém uma predicação (as mini-orações ou *small clauses*) "os mestrandos assustados" e "as professoras estressadas", não selecionando um argumento externo, como ilustram os exemplos em 2 a) *Parecer* [_{sc} *os mestrandos assustados*] e b) *Acabar* [_{sc} *as professoras estressadas*] (HENRIQUES, 2008).

²³ 'Extraposição' é um termo utilizado na literatura gerativista referindo a não adjacência entre duas partes de uma construção, ou a um tipo específico de movimento sintático.

²⁴ Exemplos de (30) a (33) foram extraídos de Duarte (2000, p. 32, 33).

- (33) Que o partido tomou a medida errada, *pro_{exp}* é evidente.
(33)a. Que o partido tomou a medida errada, isso é evidente.

Essas constituem, assim, no entendimento de Duarte (2000), as estratégias usadas para evitar a posição vazia do expletivo pré-verbal. Entretanto, a autora destaca o uso do demonstrativo *isso* e a topicalização do argumento interno ou do complemento, exemplo (34) a seguir, como as mais importantes para esse propósito. É interessante notar que a topicalização do sujeito é uma das características do Francês falado, levando-se em conta que Duarte (2000) assume que as mudanças paramétricas em progresso no PB, ao mesmo tempo que o separam de LSNs românicas, como Italiano, Espanhol e PE, tornam-no mais similar ao Francês, uma LSNN.

- (34) Eu acho que o povo brasileiro_i ele_i tem uma grave doença.

2.1.1 Considerações parciais

Kato (2000) finaliza sua proposta com a afirmação de que a principal conclusão teórica do trabalho é que a projeção obrigatória de SpecT, ou EPP, é um epifenômeno, o qual resulta do surgimento de formas pronominais fracas. Esta é uma diferença básica em relação às LSN típicas, nas quais não existem formas livres fracas nem no paradigma pronominal nem como pronome locativo, e como consequência, elas nunca projetam SpecTP. Em suma, o SpecTP é projetado em línguas como o Francês, o Inglês e o PB atual porque não há outro lugar para onde as formas fracas podem ir, uma vez que a morfologia delas não tolera sua adjunção ao núcleo (cf. Nunes 1998 *apud* Kato, 2000).

Adoto esta proposta teórica de Kato (2000) para nortear minhas discussões acerca do emergente preenchimento da posição do sujeito nulo não referencial e nulo genérico no PB, com pronomes genéricos, como consequência do surgimento de formas pronominais fracas que levou à projeção obrigatória de SpecTP. Entretanto, este trabalho se propõe a um avanço em relação à proposição de Kato (2000) de que no PB a preferência para lexicalizar a posição do sujeito - Spec TP é via alçamento do locativo para a posição de sujeito.

Neste ponto considero que o PB está avançando no sentido de se aproximar cada vez mais da gramática do Francês, uma vez que se observa na língua não só um aumento no preenchimento do sujeito nulo genérico com formas pronominais expressas, como também uma

tendência para se evitar até mesmo o expletivo nulo, revelando um pronome foneticamente realizado nesta posição em construções existenciais.

Ainda, partindo do pressuposto de que este estudo assume com Kato (2000) que o PB constitui uma língua *pro-drop* parcial, entendo ser pertinente lançar mão de outros aportes teóricos que tratam da diferenciação entre línguas *pro-drop* parciais e consistentes, nos termos de Holmberg (2010), para consubstanciar esta classificação e algumas das hipóteses propostas nesta tese, o que passo a fazer com a apresentação da proposta de Holmberg (2010).

Entretanto, embora Holmberg (2010) e Kato (2000) categorizem o PB como língua *pro-drop* parcial, deve-se observar que as perspectivas adotadas são distintas. Enquanto para Kato (2000), o PB está passando por processos de mudança que o distinguem das línguas *pro-drop* típicas, portanto, *pro-drop* parcial é um estágio da língua portuguesa do Brasil, para Holmberg, este é um estado da língua (Kato e Duarte, 2014, p. 17).

2.2 A proposta de Holmberg (2010)

Embora o parâmetro do sujeito nulo (PSN) seja sempre referido no singular, para Holmberg (2010), mais de um parâmetro está envolvido na determinação de o pronome sujeito poder ou não ser nulo, posto que os sujeitos nulos podem ser derivados de mais de uma forma. O autor propõe uma reformulação da proposta paramétrica de Rizzi (1982) que reconhece dois parâmetros:

- (35) a. *INFL can be specified [+ pronoun]* ²⁵
b. *INFL which is [+ pronoun] can be referential*

Nos termos de Rizzi (1982), enquanto o parâmetro (35)a. faz distinção entre LSNs e não LSNs, o (35)b. distingue entre aquelas LSNs que permitem todos os tipos de sujeito nulo, referenciais e não referenciais, e aquelas que somente licenciam os não referenciais.

Entretanto, Holmberg (2010) alega que esses parâmetros não são suficientes para dar conta da variação encontrada nas línguas, uma vez que há outros parâmetros afetando a realização de pronomes sujeitos que complicam o quadro. Uma das questões levantadas pelo linguista diz respeito ao fato de que LSN consistentes não podem permitir todos os tipos de sujeitos nulos permitidos em LSN parciais. Há um tipo de sujeito nulo, o qual é inexistente em LSNs consistentes em sentenças finitas, nomeadamente, pronomes nulos genéricos. Isso significa que a relação entre os parâmetros de sujeito nulo não é como proposto por Rizzi (1982).

Holmberg (2010) propõe uma versão de (35)b. para distinguir o que ele categoriza de LSNs consistentes de outras línguas. Entretanto, a noção crucial para ele não é referencialidade, mas definitude: “há um parâmetro, uma versão de (35)b., o qual distingue entre línguas que permitem sujeitos nulos definidos (*he, she*), como em (36), de todas as outras línguas”²⁶ (Holmberg, 2010, p. 88).

- (36) *Verrà.* ²⁷ (Italiano)
(*He will come*)

²⁵ a. *INFL pode ser especificado [+ pronominal]*
b. *INFL que é [+ pronominal] pode ser referencial.*

²⁶ Original: “There is a parameter, a version of (1b), which distinguishes languages that allow definite null subjects (null ‘he/she’) from all other languages”.

²⁷ Exemplo extraído de Holmberg (2010, p.88)

É dessa forma que Holmberg (2010) propõe o traço de definitude (D), como uma propriedade que as LSNs consistentes têm que as parciais - LSNPs não têm, ou seja:

(37) Traço D (definitude) como parte de traços-*phi* que compõem T finito.²⁸

Dentre as outras línguas, há aquelas que permitem sujeitos nulos indefinidos (*one* nulo) e sujeitos nulos expletivos (*it* nulo), por exemplo o PB, e aquelas que não permitem sujeitos nulos de forma alguma, por exemplo, o Inglês. O primeiro compõe a classe das LSNPs, e o último, a classe das LSNs. Dentre as línguas categorizadas como LSNPs, Holmberg cita o PB, Filandês e Marati.

Esta proposta dá conta de explicar algumas particularidades, como a de que LSN não podem permitir todos os tipos de sujeitos nulos permitidos em LSNPs, especificamente pronomes nulos indefinidos que são inexistentes em LSNs consistentes, e o fato de que sujeitos nulos definidos somente podem ser derivados em LSNPs pela eliminação do pronome em Spec TP, sujeito a controle por uma sentença mais alta, como no exemplo (38):

(38) A Maria_i admite que *t_i* não fala bem inglês.²⁹

A explicação parte do fato de que em uma língua sem D³⁰ em T, a relação entre sonda³¹ (T) e alvo local (sujeito nulo ϕ P³²) não supre o valor de definitude. O resultado é um D-less, portanto pronome subjetivo indefinido. Se os traços-*phi* são de 3ª. pessoa, a interpretação é de um pronome genérico inclusivo, correspondendo a *one* em Inglês³³, em construções como *É assim que Ø faz o doce*, a ser analisada a seguir.

Isso explica a ausência de pronomes indefinidos nulos em sentenças em LSNs consistentes, uma vez que nessas línguas o traço de definitude em T bloqueia a interpretação indefinida do pronome de 3ª pessoa, ou ϕ P, e porque LSNs parciais têm sujeitos nulos com interpretação definida somente quando eles são controlados por um DP definido, exemplo (38).

²⁸ D(e)finite-feature as part of the ϕ -feature make-up of finite T.

²⁹ Exemplo extraído de Holmberg (2010, p. 131).

³⁰ D – traço de difinitude.

³¹ Em Chomsky (1999), assume-se que para que se estabeleça a operação *Agree* (operação sintática de concordância) um traço não interpretável, *sonda*, procura um traço do mesmo tipo, *alvo*. Os traços não-interpretáveis são traços sem valor especificado que devem ser valorados no decorrer da derivação linguística.

³² ϕ P (Pronome de terceira pessoa defectivo).

³³ Deve-se notar que no PB, além dos pronomes genéricos, os arbitrários integram a categoria dos nulos indefinidos.

A partir das propriedades propostas, Holmberg (2010) assume que podem ser caracterizados dois tipos de LSNs: as consistentes, que apresentam pronome subjetivo definido nulo (*he, she*) e não apresentam pronomes sujeitos indefinidos nulos (*one*), e as parciais, que apresentam pronomes sujeitos definidos nulos somente se c-comandados³⁴ pelo antecedente, mas apresentam pronomes subjetivos indefinidos nulos.

Outras distinções entre LSNs consistentes e parciais, apontadas por Holmberg (2010), dizem respeito aos contextos em que os sujeitos nulos são opcionais em LSNPs, mas obrigatórios em LSNs, e contextos nos quais eles são excluídos em LSNPs, mas permitidos em LSNs.

Dada a sentença (39),

(39) *John said that **he** wanted to buy a car.*³⁵

em LSNs, como o Inglês e o Francês, o pronome precisa ser expresso. Em LSNs, como o Grego e o Italiano, o pronome deve ser nulo, sob o argumento de que não há contraste ou mudança de tópico envolvido. Em LSNPs, como o PB, o pronome é opcionalmente nulo.

Considerando um contexto em que outra pessoa está em discussão, em uma LSN consistente, o pronome ainda poderá ser nulo se não se assume nenhum contraste. Já em uma LSNP, como o PB, neste caso, o pronome precisa ser expresso.

Para uma sentença como (40),

(40) (Filandês)

a. *Juha₁ ei ole sanoonut mitäÄn. Mutta Pauli₂ sanoo että *Ø₁ haluaa ostaa uuden auton.*³⁶

*Joha₁ hasn't said anything, but Pauli₂ says that **he**₁ wants to buy a new car.*

(Italiano)

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

Gianni₁ hasn't said anything, but Paul₀₂ says that he_{1/2} wants to buy a new car.

(Gianni₁ não disse nada, mas Paul disse que ele₁ quer comprar um carro novo.)

³⁴ Relação de *c-comando* (comando de constituinte) trata-se, originalmente, de uma relação recíproca entre dois elementos com um nó estrutural comum, isto é, um nó A c-comanda B e B c-comanda A se ambos são imediatamente dominados por um nó C, e A e B não dominam um ao outro.

³⁵ Exemplo extraído de Holmberg (2010, p.91)

³⁶ Exemplo extraído de Holmberg (2010, p.92).

em uma LSNP, como o Finlandês por exemplo, o antecedente pretendido *Joha* não c-comanda o sujeito nulo, sendo que o único possível antecedente do sujeito nulo é o DP *Pauli*. Já em Italiano, uma LSN consistente, o pronome é nulo. A explicação dada por Holmberg (2010) é que em Italiano, c-comando não é uma exigência, visto que o antecedente é um tópico, constituindo este outro fator determinante para que as LSNs sejam categorizadas em consistentes e parciais.

Contudo, Holmberg (2010) esclarece que o favorecimento de sujeitos nulos em LSNs consistentes se dá somente no caso de pronomes definidos. Para sujeitos nulos indefinidos, a situação é contrária. Os nulos indefinidos são comuns em LSNPs, mas encontrados sob condições mais restritas em LSNs consistentes. Em particular, o que não se encontra em qualquer LSN consistente é o sujeito nulo indefinido de 3ª pessoa, o nulo *one*, em sentenças ativas finitas.

Comparando-se as sentenças em PB e PE,

(41) É assim que __ faz o doce. (PB)

(42) É assim que *se* faz o doce. (PE)

observa-se que, enquanto no PB, o sujeito correspondente ao pronome genérico inclusivo em Inglês *one* é nulo, em PE, esse pronome, *se*, precisa estar presente. Conforme demonstrado anteriormente, a razão por que LSNs consistentes não podem ter um *one* nulo é que esse pronome é um ϕ P vazio, o qual não pode valorar uD ³⁷, e consequentemente, permanece invalidado, com falha na derivação, pois a valoração de uD é uma propriedade de línguas desse tipo. Por outro lado, LSNs consistentes podem expressar leitura genérica exclusiva³⁸ com pronome sujeito de 3ª pessoa do plural, mas parece sempre recorrer a outras estratégias que não o pronome nulo de 3ª pessoa do singular para expressar leitura genérica inclusiva para o sujeito de uma sentença finita, tipicamente um nulo de 2ª pessoa ou alguma forma de passiva.

Para abordar a incorporação de ϕ P em T e a projeção de SpecTP, Holmberg (2010) aplica a teoria de pronomes clíticos de Roberts (no prelo) à de sujeitos nulos, e assume a incorporação do pronome nulo ϕ P como sendo o resultado direto de *Agree*, não envolvendo

³⁷ uD – traços de definitude não interpretáveis.

³⁸ A distinção entre marcação genérica inclusiva e exclusiva será abordada no item 1.4.6 Referência arbitrária x referência genérica (+) e (-) inclusiva.

qualquer movimento. A incorporação se dá da seguinte forma: T finito tem um feixe de traços não valorados e, portanto, sonda por uma categoria com traços valorados. O pronome sujeito defectivo tem os traços-*phi* requeridos valorados, e portanto, elimina os traços não valorados de T. Como resultado T compartilha todos os traços de ϕ P. Efetivamente o resultado é o mesmo que se ϕ P tivesse movido, incorporando em T. Consequentemente, o sujeito ϕ P não é pronunciado. Se a cadeia inclui traços D, pela virtude de traços D de T, e uma vez que D é valorado por um A-tópico em Spec CP, o resultado é uma construção de sujeito nulo definido.

Um fato a ser considerado é que sujeitos nulos de 3ª. pessoa são dependentes de um antecedente em LSNs consistentes também. Entretanto, há algumas condições, sendo a mais importante, que o antecedente precisa ser um tópico, na realidade, um tipo específico de tópico que atua como antecedente de um item nulo (Frascarelli, 2007 apud Holmberg, 2010): A-tópico (*aboutness shift topic*). Frascarelli (2007) postula que este A-tópico tem propriedades que atendem à hipótese D em T, como estar em uma posição designada na articulação com o domínio C³⁹. Assim, o antecedente de um sujeito nulo é um A-tópico do domínio C da oração imediatamente anterior.

Assim, segundo Holmberg (2010), o vestígio de referencialidade do sujeito nulo de 3ª. pessoa, em última análise, vem ou é identificado com o vestígio do DP explicitado no discurso precedente, via cadeia de A-tópicos. O resultado é que os únicos pronomes que permanecem nulos são aqueles indexados a um A-tópico nulo em Spec CP. Portanto, a generalização de que sujeitos nulos de 3ª. pessoa em LSNs são sempre definidos e sempre referem a uma pessoa ou objeto já introduzido como tópico é assim explicada.

Para elucidar o A-tópico, Holmberg apresenta o seguinte exemplo em italiano:

(43) *Questa mattina, la mostra è stata visitata di Gianni. Più tarde *0/egli ha visitato l'università.*⁴⁰

(Esta manhã, a exposição foi visitada por Gianni. Mais tarde, ele visitou a universidade.)

A impossibilidade neste exemplo (43) de um pronome nulo na posição de sujeito, apesar da não ambiguidade na sentença precedente, deve-se ao fato de que o pronome sujeito

³⁹ O domínio C corresponde a uma das categorias funcionais que são: D (determinante), T (tempo) e C (complementizador), que delimitam os domínios nominal, verbal e oracional, respectivamente. Os núcleos funcionais são projetados em camadas mais altas. Desta forma, DP, TP e CP constituem projeções máximas dos núcleos D, T e C, respectivamente (Cf. Corrêa, 2006).

⁴⁰ Exemplo extraído de Holmberg (2010, p. 96)

‘egli’ introduz o *Aboutness-shift topic*, em termos de Frascarelli (2007), uma vez que o tópico da sentença precedente é “*la mostra*”.

De acordo com esta proposta, nos casos de LSNs consistentes, em que o sujeito é incorporado em T e SpecTP não é projetado, o EPP é checado pelo A-tópico nulo, o qual atribui o valor *uD* em T. Dessa forma, uma estrutura sem SpecTP é permitida somente no caso de sujeitos nulos definidos, que é quando há um antecedente. Os outros casos, quando SpecTP precisa ser preenchido para checar o EPP, incluem:

- “ a) Quando o sujeito é um pronome nulo indefinido;⁴¹
- b) Quando o sujeito é um DP lexical ou um pronome pleno não incorporado e
- c) Existenciais e outras sentenças nas quais falta um A-tópico.”

(Holmberg, 2010, p.100)

Conforme mostrado por Holmberg, em outras línguas, LSNs não consistentes, ou parciais, T finito não tem o traço *uD*. No subconjunto dessas línguas, como o PB e Finlandês, por exemplo, o sujeito referencial ainda pode ser nulo, essencialmente pela mesma derivação como nas LSNs consistentes: T valoriza os traços-*phi* do sujeito. Esses valores são copiados por T, e, em troca, o sujeito tem seu traço de Caso avaliado. No caso em que o sujeito é um ϕ P vazio, T irá copiar todos os valores de traços do sujeito. Como resultado, T e ϕ P formam uma cadeia, e o sujeito permanece nulo pela redução da cadeia. Esses sujeitos nulos precisam ter um antecedente c-comandando localmente; eles precisam ser controlados. De outra forma, na ausência de D em T, valorado por um A-tópico, a interpretação do sujeito não pode ser de um pronome definido.

Dessa forma, em LSNPs, enquanto os sujeitos nulos de 3ª pessoa são ϕ P que podem receber uma interpretação definida por meio de uma relação de ligação ou controle com o antecedente em uma sentença mais alta, sujeitos nulos de 1ª e 2ª p. são DPs totalmente especificados, os quais são deletados. Na ausência de controlador, uma 3ª pessoa só pode ser interpretada como impessoal, tanto genérica como não temática. Esse é o caso em LSNs consistentes também: sujeitos nulos são ϕ Ps. A diferença crucial é que LSNs consistentes têm o traço-D em T, o qual assegura, e força, uma interpretação definida do ϕ P, ainda que sem um antecedente.

⁴¹ Original: a) *when the subject is a null indefinite pronoun;*

b) *when the subject is a lexical DP or a spelled-out, non-incorporated pronoun; and*

c) *existential and other thematic sentences, with lack na A-topic.*

Quanto à checagem de EPP, Holmberg (2010) argumenta que o sujeito que não é incorporado é atraído pelo EPP para SpecTP. Portanto, o sujeito nulo definido em LSNPs move-se para SpecTP e checa o EPP, enquanto o indefinido/ genérico é incorporado em T, está em SpecvP e não checa o EPP. As LSNPs têm em comum o fato de permitirem incorporação do pronome indefinido em T. Holmberg defende, assim, a hipótese de que em LSNPs o EPP, em sentenças com sujeito nulo definido, é checado pelo A-tópico nulo. Já na ausência de um controlador, como a 3ª pessoa é interpretada como indefinida, o EPP não é checado, e precisa ser checado por alguma outra categoria, normalmente expressões adverbiais.

O quadro de LSNPs pode ser assim resumido: quando o sujeito é um DP (lexical ou pronominal) ele não pode ser incorporado, e quando o sujeito (pronome nulo de 3ª. pessoa) é incorporado, ele só pode ser interpretado como indefinido, uma vez que T não tem o requisito de traço *uD*.

Holmberg (2010) mostra que o Filandês tem uma condição EPP, a qual é na maior parte das vezes satisfeita pelo sujeito, mas pode ser satisfeita por outras categorias também, incluindo advérbios circunstanciais. No exemplo a seguir (44)a., o sujeito de 3ª pessoa não realizou movimento para SpecTP, ao invés, o advérbio de lugar o fez. Em (44)b. o sujeito moveu-se para SpecTP, satisfazendo o EPP. Holmberg (2010) lembra que nessa posição o pronome sujeito pode ser nulo não em virtude da incorporação em T, mas em virtude de ter um antecedente local, daí a interpretação.

- (44) a. *Juri says that here sits comfortably.*
b. *Juri says that ϕ sits comfortably here.*

Situação similar é encontrada no PB:

- (45) a. João me contou que na praia vende cachorro-quente.
b. João me contou que \emptyset vende cachorro-quente na praia.

No PB, o pronome nulo só tem interpretação definida se se refere ao antecedente local; portanto em construções como essa (45)a., para ser definido, o pronome não pode ser nulo para não provocar ambiguidade.

Para Holmberg (2010), o movimento do advérbio em construções como (44) e (45) é condizente com as observações: quando um pronome sujeito é incorporado em T, em uma língua que não tem traços *uD* em T, os quais possam ser avaliados por um A-tópico, então,

somente então, o EPP precisa ser satisfeito por outras maneiras, especificamente, movimento de uma categoria para SpecTP. Holmberg (2010) sugere, então, que enquanto sujeitos nulos definidos em LSNs consistentes são ϕ Ps incorporados, sujeitos nulos definidos no PB e outras LSNPs são DPs que têm segunda imersão com SpecTP.

Holmberg (2010) assume, então, que se esta proposta estiver correta, não há nenhum *pro*, de nenhum tipo, em SpecTP em construções de sujeito nulo em LSNs consistentes ou em construções de sujeito nulo genéricas em LSNPs, mas há um em construções de sujeito nulo controlado em LSNPs.

Deve-se notar aqui a equivalência desta proposição com a de Kato (1996a) sobre a não projeção de Spec de T para sujeitos nulos referenciais, em função de *pro* não ser uma categoria descritiva. Entretanto, para Kato, como o PB retém o sujeito nulo não-referencial, *pro* é mantido para sujeitos nulos arbitrários e anafóricos, enquanto a proposta de Holmberg admite a existência de *pro* somente em construções de sujeito nulo controlado em LSNPs.

Ao considerar os efeitos da teoria proposta de reformulação do parâmetro do sujeito nulo, Holmberg (2010, p. 124) enfatiza que: “i) ter sujeito nulo definido de 3ª pessoa exclui ter sujeito nulo indefinido em sentenças finitas; ii) o sujeito nulo indefinido não satisfaz o EPP, enquanto o definido o satisfaz; iii) sujeitos nulos indefinidos podem ser genéricos mas não existenciais ou universais”.

Além do parâmetro “D em T”, outro parâmetro postulado por Holmberg (2010) que lança luz para a variação em análise neste estudo diz respeito ao da ϕ -(in)dependência⁴² de EPP, cujas propriedades permitem distinguir LSNPs das LSNs.

De acordo com este parâmetro, em línguas LSN consistentes, o EPP é conectado com a checagem dos traços- ϕ não-interpretáveis de tal forma que a única categoria que pode satisfazer o EPP é a sondagem de T, isto é, um DP ou pronome capaz de valorar seus traços- ϕ não-interpretáveis, ao qual é atribuído Caso nominativo. A essas línguas Holmberg se refere como ϕ -dependentes.

Por outro lado, se a checagem de traços- ϕ não-interpretáveis de T não está acessível para movimento para SpecTP, seja porque está muito distante, ou porque não tem nenhum alvo para valorá-los, então o EPP irá atrair outra categoria mais local, como no exemplo (45)a., que teve o locativo ‘na praia’ alçado para a posição SpecTP da sentença encixada. Nesses casos, o requerimento de EPP em T é ϕ -independente, uma vez que licencia categorias não argumentais

⁴² A ϕ -(in)dependência tem relação com a checagem dos traços- ϕ não interpretáveis de T – nesse caso, traços formais que codificam as informações de número e pessoa – contra os traços- ϕ interpretáveis presentes no DP que ocupará a posição de SpecTP.

na posição de Spec-TP. Dessa forma, uma vez que no PB, SpecTP pode receber tanto DPs argumentais como não argumentais, o PB é caracterizado como uma língua ϕ -independente, por autores como Avelar & Galves (2011), ao contrário do que ocorre no PE, língua ϕ -dependente, em que EPP atrai para SpecTP apenas DPs portadores dos traços- ϕ sondados por T.

2.2.1 Resumo

Sintetizando a proposta teórica de Holmberg, para o autor há mais de um parâmetro envolvido na questão do sujeito nulo, além dos postulados por Rizzi (2000). Um deles diz respeito à noção de definitude (D), a qual, segundo o autor, é fundamental para a explicação do sujeito nulo nas línguas. Definitude, um dos traços-*phi* que compõem T finito, é uma propriedade de LSN consistentes, e é o que as distingue das demais. É esse traço que atribui uma interpretação definida a um DP de 3ª pessoa, o qual, em línguas desprovidas de *uD* em T, tem interpretação indefinida.

Esta proposta é interessante por dar conta de explicar por que LSNPs licenciam sujeitos nulos indefinidos, enquanto as consistentes não os licenciam, ou seja, por que a interpretação de um nulo de 3ª pessoa é genérica em línguas como o PB, enquanto no PE esse nulo não pode ter interpretação indefinida. Por outro lado, as línguas consistentes apresentam pronomes sujeitos definidos nulos (*he, she*), enquanto as parciais os apresentam somente se comandados por um antecedente, visto que em uma língua sem D em T, a relação entre sonda (T) e alvo (sujeito nulo ϕ P) não supre o valor de definitude e o resultado é um D-less, portanto pronome sujeito indefinido. Evidencia-se assim uma propriedade das LSNPs que as distingue das LSNs consistentes.

Deve-se ressaltar ainda que esta proposta de Holmberg dá conta também da distinção entre *você* como pronome de 2ª. pessoa e *você* genérico, a qual pode ser explicada recorrendo-se ao traço de definitude exigido nos traços-*phi* das LSNs consistentes, não necessário nas LSNs parciais. É a presença/ausência do traço [+ definido] em T, que pode coocorrer ou não com o traço de pessoa, que possibilita a distinção de um T com *você* segunda pessoa definida e outro com *você* genérico.

O parâmetro da ϕ -(in)dependência de EPP também sinaliza uma motivação para a emergência de um tipo específico de estrutura no PB – mas não no PE, a estratégia alternativa a estruturas existenciais impessoais e ao nulo genérico, o pronome *você_{gen}*, a qual surge em atendimento às demandas do sistema que caminha no sentido do preenchimento da posição

sujeito, uma vez que o PB, como uma língua ϕ -independente, exige a presença de um DP em SpecTP para checar EPP.

Em comparação à proposta de Kato (2000), considero que elas são equivalentes em alguns aspectos e divergentes em outros. Quanto ao status de *pro*, tanto Kato (2000) como Holmberg (2010) são unânimes em postularem a inexistência de *pro* como categoria descritiva em SpecTP em LSNs consistentes, e admitirem sua existência em LSNPs em construções, denominadas por Holmberg (2010) de sujeito nulo controlado, e por Kato (2000), de pronomes anafóricos. Kato (2000) postula ainda a existência de *pro* em construções de pronomes que considera arbitrários, a partir do entendimento que o PB mantém o nulo não referencial.

Entretanto, as propostas divergem no diz respeito à checagem de traços. Enquanto Kato (1996a), em seu estudo sobre pronomes fracos e fortes postula que, como um item lexical independente na Numeração, *Agr* pronominal pode checar *uD* e traços de Caso, para Holmberg (2010) somente os nulos definidos em LSN e indefinidos em LSNPs têm os traços checados pela incorporação do pronome em T; no caso de nulos definidos em LSNPs, essa checagem se dá por um A-tópico nulo (no caso da 3ª pessoa), com vestígio referencial do A-tópico.

Outra divergência que merece destaque entre as duas propostas refere-se à postulação de *nulo* expletivo em construções com V inicial. Holmberg (2010) assume que, no caso de sentenças com V em posição inicial em línguas Românicas, como o Português, as quais podem ter sujeito pós-verbal, e aparentemente não projetar SpecTP, que elas têm um expletivo nulo em SpecTP checando EPP. Esse é o caso das construções existenciais. Já Kato (2000), em vez de postular um *pro* nas construções existenciais em Português, alega que é o afixo de 3ª. pessoa que faz o papel de *il*, mas, de acordo com a autora, ainda que faça parte da Numeração, o afixo de 3ª. pessoa não tem papel de argumento. Nos termos de Kato (2000), “o nulo não argumental (\emptyset_2) é o único afixo que ainda pode aparecer como um item independente na Numeração e pode ser fundido com V+T para checar seus traços-D”.

Finalmente, ressalto que, por dar conta de explicar por que a interpretação de um pronome nulo de 3ª pessoa é indefinida em línguas como o PB, o qual, em virtude da não checagem do EPP, encontra-se em estágio de variação com um pronome pleno, *você_{gen}*, é que a proposta de Holmberg (2010) se mostra adequada para nortear e referendar as discussões desta pesquisa. Por se complementarem, lançarei mão de alguns dos pressupostos teóricos de Kato (2000) para balizar a caracterização de LSNPs, relativamente ao surgimento do paradigma de pronomes fracos, como um sintoma da remarcação paramétrica por que está passando o sistema do PB.

3. SINTOMAS DA REMARCAÇÃO DO PARÂMETRO *PRO-DROP* NA GRAMÁTICA DO PB

Este capítulo dedica-se à discussão de alguns dos sintomas da remarcação do parâmetro *pro-drop* no PB, os quais podem ter desencadeado a implementação das emergentes estratégias de indeterminação do sujeito na língua, a partir da proposição de um percurso que vai do pronome *se* ao *(vo)cê*. Além da apresentação dos fenômenos sintáticos relacionados à remarcação paramétrica, serão discutidos conceitos que embasarão a análise dos dados.

3.1 Reanálise e apagamento do clítico *se*

Pesquisas na perspectiva da Sociolinguística Paramétrica, como as de Nunes (1990) e Duarte (2000), apontam que uma das maiores divergências entre o Português brasileiro e o europeu diz respeito à passiva pronominal sem concordância em oposição à com concordância que ainda é a estratégia de indeterminação preferida pelos lusitanos.

De acordo com Naro (1976, p.788 apud Nunes, 1990, p.76), a construção sem concordância em Português ganhou aceitação aproximadamente entre meados do séc. XV e meados do séc. XVI. Naro (1976) aponta que o surgimento da passiva pronominal sem concordância desencadeou um processo de mudança linguística no PB em que a forma com concordância foi perdendo espaço para sua concorrente, sem concordância, até que seu emprego se restringisse à modalidade escrita.

Para Nunes (1990), a reanálise⁴³ da passiva pronominal envolve um processo de reanálise sintática em que o expletivo da posição de sujeito de construções com *se* apassivador foi tomado como um pronome nulo referencial (*pro*). Segundo Nunes, esse processo é assim explicado: considerando-se as estruturas (46) e (47), com e sem concordância,

(46) (*expl*) alugam-se_{pe} casas_{pi}

(47) *pro*_{pe} aluga-se_{p0} casas_{pi}

⁴³ Para a definição de reanálise, Nunes (1990, p.92) cita Langacker (1976, p.58): “*I will define ‘reanalysis’ as a change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface manifestation*”.

na estrutura (46), o clítico recebe papel temático do argumento externo e caso acusativo e o SN *casas* é detentor do papel temático do argumento interno e recebe caso nominativo em cadeia com o expletivo. Em (47), há na posição de sujeito um pronome nulo referencial - portador do papel temático do argumento externo e de caso nominativo - que é indeterminado pelo clítico, e o SN *casas* continua detendo o papel temático do argumento interno, recebendo agora caso acusativo.

De acordo com essa perspectiva, o fato de a categoria vazia de (46) passar a ser interpretada como um pronome referencial e não como um expletivo, como em (47), forçou a interpretação do SN posposto como objeto direto, portador de caso acusativo, e não mais como sujeito, portador de caso nominativo e responsável pela concordância verbal.

Nesses termos, o PB passou a admitir construções em que um pronome referencial nulo *-pro* na posição de sujeito é indeterminado pelo clítico. Isso é condizente com o fato de que uma vez que a categoria vazia da posição de sujeito passou a ser interpretada como referencial, e, portanto, temática, era previsível que o alçamento do argumento interno passasse a ser bloqueado para a posição de sujeito.

Outra inovação que se seguiu a esta, de acordo com Nunes (1990), diz respeito ao apagamento do clítico *se* em construções de indeterminação como no exemplo (48) de Galves (1987, p. 37):

(48) Nos dias atuais não *Ø* usa mais saia.

Nunes (1990) mostra que no século XIX, momento em que as ocorrências das estruturas contendo o *pro* referencial na posição de sujeito superam as de estruturas com nulo expletivo, é que começa a haver a supressão do clítico em estruturas não coordenadas "passivas pronominais sem pronome a partir de construções com *se* indeterminador". O estudo mostra que o fenômeno de supressão do *se* indeterminador nas passivas pronominais sem concordância atinge, à época, um percentual de 79% em entrevistas (107 ocorrências em 135 construções).

Essa questão das passivas pronominais sem pronome salienta a distância entre o PB e o PE. Enquanto o apagamento de *se* atinge 79% em entrevistas do PB, conforme Nunes (1990), em entrevistas do dialeto europeu essa marca é de apenas 2% (2 ocorrências em 97 construções).

Aliada a essas duas inovações em relação ao clítico *se*, quais sejam, a reanálise do pronome nulo, que de expletivo passa a ser tomado como referencial (*pro*), e sua subsequente supressão em construções finitas, o PB registra ainda outra mudança mais recente: a inserção do pronome *se* em construções infinitivas.

Em relação às sentenças infinitivas, Galves (1987, p. 35) observa que há "estranha diferença entre o PE e o PB: enquanto que, neste, o *se* tem nítida tendência a desaparecer em todos os seus usos nas frases com tempo, ele reaparece maciçamente - distanciando-se nisso do uso do português europeu – nas infinitivas para expressar a indeterminação”.

Considerando os exemplos (49) e (50),

(49) É impossível *se* achar lugar aqui.⁴⁴

(50) É impossível *Ø* achar lugar aqui.

Galves (1987) atribui essa divergência a diferenças na interpretação do sujeito nulo das infinitivas. Enquanto no PB, toma-se como referência para esse sujeito nulo um antecedente na sentença ou no discurso (isto é, o tópico), em PE, toma-se um antecedente na sentença ou tem-se uma interpretação indeterminada. Galves (1987) conclui, assim, que o clítico indeterminador surge nas sentenças infinitivas do PB para impedir que seu sujeito seja referencialmente vinculado ao tópico.

Essas discussões sobre a supressão do *se* em sentenças finitas e sua inserção em infinitivas serão retomadas na análise dos dados.

3.2 O surgimento do paradigma de pronomes fracos

Nesta seção, Kato (1999) é retomada para um detalhamento do que a autora considera um dos sintomas da perda das propriedades de sujeito nulo em uma língua: o surgimento de um paradigma de formas pronominais fracas em competição ou substituição às formas fortes (Kato, 1999).

Kato (1999) propõe que todas as línguas têm pronomes fortes, que são dêiticos por natureza, enquanto as formas fracas, que são ou dêiticos ou referencialmente dependentes, aparecem apenas em línguas sem uma concordância [+ pronominal]. Dito de outra forma, se a língua não exhibe pronomes subjetivos fracos ou sujeitos clíticos, então ela será considerada uma língua de sujeito nulo. Dessa forma, pode-se afirmar que a teoria do surgimento de formas pronominais fracas pressupõe um dos estágios de perda das propriedades de sujeito nulo.

⁴⁴ Exemplo extraído de Galves (1987 apud Nunes, p.107).

Embora a morfologia rica seja uma condição para o licenciamento do sujeito nulo, seu enfraquecimento não é suficiente para dar conta da perda das propriedades desse parâmetro, conforme atestam vários estudos diacrônicos, dentre eles, o de Duarte (1993) para o PB, e de Adams (1987) e de Roberts (1993) para o Francês Antigo - FA.

Outro importante estudo que desafia a hipótese entre morfologia rica e licenciamento do sujeito nulo é o de Sigurðsson (1994 apud Kato, 1999), sobre o Islandês Antigo (IA), que mostra que, embora a morfologia *Agr* não tenha se alterado, o IA perdeu o sujeito nulo referencial.

Fato curioso no PB que de certa forma também sobrepuja essa hipótese da morfologia rica é que o sujeito nulo teve restrições mais fortes em relação à 1ª. e à 2ª. pessoas e menos fortes em relação à 3ª. pessoa, a qual no singular exibe uma forma não marcada. Um ponto interessante nesse estudo de Sigurðsson (1994), de acordo com Kato (1999), é a revelação de que uma língua que faz amplo uso de pronomes de terceira pessoa nulos não pode licenciar pronomes de 1ª. e 2ª. pessoa nulos. Essa constatação mostra-se importante para a compreensão do que aconteceu no PB, que apresenta comportamento do sujeito nulo parcialmente semelhante ao do IA, posto que ambos PB e IA não têm sujeitos nulos nem de 1ª. nem de 2ª. pessoas, mas têm de 3ª. pessoa anafórica. Ainda deve ser salientado desse estudo que, assim como o PB, tanto o IA quanto o Islandês moderno apresentam a mesma forma para a 2ª. e a 3ª. pessoas do singular. Esses fatos sinalizam momentos de generalização interlinguística, conforme previsto pela teoria de propriedades paramétricas.

Na versão minimalista de Chomsky (1995), fortemente motivada pelo princípio da economia de derivação, AGRP deixa de ser uma projeção independente e todos os verbos nascem na Numeração totalmente flexionados, com *head* capaz de verificar mais de um traço. Para Chomsky (1995), o que conta como categorias funcionais legítimas são apenas elementos interpretáveis em LF, que não é o caso de AGRP.

Diante de fatos como os ocorridos no PB e IA, e da proposta de Chomsky (1995) de que AGRP deixa de ser uma projeção independente, Kato (1999) afirma que AGRP tem que ser estudado em conjunto com outros portadores de traços-*phi*, tais como pronomes fracos e clíticos, que são considerados em distribuição complementar com a flexão *Agr* de natureza pronominal ⁴⁵.

⁴⁵ Esta proposta é baseada na proposta inicial de Everett (1993) de que os três tipos de manifestação de formas fracas, *Agr* pronominal, pronomes livres e clíticos, estão em distribuição complementar interlinguisticamente, definindo diferentes línguas, ou eles podem aparecer em distribuição complementar dentro da mesma língua, como em Hebreu. Nas conhecidas línguas *pro-drop* prototípicas, como Italiano e Espanhol, os pronominais fracos são os afixos de concordância.

Dessa forma, Kato (1999) propõe que a perda do licenciamento do sujeito nulo no PB correlaciona-se com o surgimento de pronomes subjetivos fracos, possíveis portadores de traços-*phi*, como compensação para essa perda. Esta hipótese do surgimento de pronomes fracos referenda-se em análises de processos ocorridos no Francês Antigo (FA). O FA tem sido analisado como uma LSN, a qual perdeu as propriedades relacionadas ao licenciamento do SN juntamente com a ordem VS.

Kato (1999) dá a seguinte interpretação para os fatos em Francês: quando o pronome nominativo coexistia com *Agr* pronominal, o sujeito nulo era possível. Quando apareceu o paradigma fraco nominativo quase homófono, *Agr* deixou de ser pronominal, o sujeito nulo desapareceu e a forma dativa assumiu a função de pronomes fortes. O paradigma de pronomes nominativos fortes foi substituído pelo paradigma do caso objetivo, o qual parece estar ainda em competição com as formas nominativas fortes em FA. O quadro a seguir é ilustrativo desses estágios:

- (51) a. JOU, Agr [+pronominal] {FA}⁴⁶
 b. JOU, je Agr [-pronominal] {entre Francês Antigo e Francês Médio}
 c. MOI, je Agr [-pronominal] {FM}

Esses fatos confirmam a hipótese de Kato (1999) de que *Agr* pronominal está em distribuição complementar com pronomes subjetivos fracos e que ambos podem co-ocorrer com pronomes fortes, ou seja, a perda de *Agr* pronominal correlaciona-se com o aparecimento de pronomes sujeitos fracos.

A proposta de Kato (1999) mostra que a emergência de pronomes fracos no PB se reflete na duplicação produtiva de pronomes subjetivos, com pronomes fortes e fracos que exibem formas quase-homófonas. Em quadro apresentado por Nunes (1990, p. 2) em Kato (1999), percebe-se que a redução fonética afetou todo o paradigma de pronomes nominativos:

1 ^a . sing.	[ew] -> [o]
2 ^a . sing.	[vosé] -> [cê]
3 ^a . sing. masc	[Eli] -> [e ^h]
3 ^a . sing. fem.	[Ela] -> [E ^h]
3 ^a . pl. masc.	[Elis] -> [ez]

⁴⁶ Exemplo (51) foi extraído de Kato (1999).

O surgimento do pronome fraco de 2ª. pessoa *cê* é atribuído por Kato (1999) à redução fonética do pronome forte *você*, por meio de gramaticalização. Nos dizeres de Kato (1999, p. 25),

A gramaticalização de *você*, um pronome forte, em um quase-clítico *cê* parece ser a chave para a mudança de *Agr* de [+pronominal] a [-pronominal]. O pronome *você* foi apenas um pronome forte, mas sua redução fonética criou um clone fraco, que usurpou a função de argumento de seu *Agr*, que agora pode aparecer apenas como um afixo ao verbo.”⁴⁷.

Para ilustrar a duplicação do sujeito e a redução fonética do pronome fraco no PB, Kato (1999, p. 10) cita os exemplos:

(52) VOCÊ, 'cê não me pega! (Kato, 1996)

you you not me catch+S3rd

(53) EU, eu (foneticamente [o]) sinto demais isso, né? (Duarte, 1995, p.110)

I feel+S1st too much this right

Em relação ao PE, segundo Kato, o pronome fraco é o afixo verbal de concordância, responsável pela duplicação em LSN. O afixo tem propriedades de um pronome: é um argumento do V em VP e checa caso em T. A diferença para um pronome fraco livre é que não exige uma projeção do Especificador de T.

Um importante questionamento levantado por Kato (1999) acerca das discussões apresentadas diz respeito à lógica dos fenômenos, qual seja, uma língua de sujeito nulo se torna de sujeito não nulo porque sua morfologia perde alguma propriedade, ou essa perda é consequência da introdução de outras formas fracas? A motivação para essa questão baseou-se em que em alguns dialetos do sul do Brasil não houve mudança no paradigma pronominal, uma vez que ‘*tu*’ foi mantido, e, mesmo assim, o sistema *Agr* tornou-se não pronominal. Isso reforça a tese de que a mudança no sistema flexional por si só não representa a perda das propriedades de sujeito nulo.

A proposta de Kato (1999) para este tipo de dialeto é que o sistema realmente tem um paradigma de pronomes fracos que é homófono a um forte, mas pode ser a perda da flexão distintiva de 2ª. pessoa que realmente desencadeou a criação de pronomes fracos para duplicar

⁴⁷ Original: *The grammaticalization of você, a strong pronoun, into a quasi-clitic cê seems to be the key to the change of Agr from [+pronominal] to [-pronominal]. The pronoun você was only a strong pronoun, but its phonetic reduction created a weak clone, which usurped the argument role of its Agr, which can now appear only as an affix to the verb, except for the third person, which is still [+pronominal].*

os fortes. Kato (1999) mostra que no centro-sul do Brasil houve um processo oposto. O surgimento de pronomes fracos mudou *Agr* para [-pronominal].

Outro questionamento apontado nesta proposta de Kato (1999) que interessa especialmente a este estudo diz respeito ao *status* da forma reduzida *cê*: se seria um pronome fraco como os pronomes sujeitos franceses ou um clítico subjetivo como os pronominais em Trentino e Fiorentino. Kato (1999) cita alguns linguistas brasileiros como Ramos (1996) e Vitral (1996) que têm considerado a mudança desse pronome em um possível clítico. Considerando esses estudos, Kato (1999) argumenta que se o *cê* brasileiro pode ser considerado como os clíticos sujeitos franceses, os quais são analisados como pronomes fracos sintaticamente, por outro lado, eles se mostram diferentes dos clíticos sujeitos do italiano do norte, que apresentam posições variáveis no que diz respeito à negação, levando à conclusão de que estes últimos é que apresentam comportamento típico dos grupos dos clíticos. De qualquer forma, para Kato (1999) esse pronome fraco (você) está cliticizando para compensar *Agr*.

No sentido de ampliar a discussão de Kato (1999) quanto à forma foneticamente reduzida – *cê* – como um pronome fraco, deve-se considerar que esta forma também tem se mostrado produtiva no PB para referência genérica em posição de sujeito, cujo comportamento sintático está sendo comparado no presente estudo com o clítico *se*, os quais podem ser usados em contextos semelhantes com similar função indeterminadora. Deve-se ressaltar, entretanto, que tendo em vista as dimensões do país, qualquer generalização sobre este fenômeno na variedade brasileira do Português só pode ser tomada em termos relativos.

Assim, a hipótese do surgimento do paradigma de formas pronominais fracas é relevante para o presente estudo por consistir em um dos estágios da perda das propriedades de sujeito nulo, o que confirma a mudança paramétrica por que passa o PB, além de prover indícios para o surgimento da forma pronominal reduzida *cê*, a qual está concorrendo, ao lado de *você_{gen}*, com as canônicas estratégias de indeterminação. Por esta hipótese é possível ainda perceber a distinção entre *você* 2ª pessoa definida e *você* pronome genérico, como pertencentes a paradigmas distintos, respectivamente de pronomes fortes e fracos.

3.3 Estratégias inovadoras de indeterminação

3.3.1 *A gente cede lugar a você para referência genérica*

Breve percurso interlinguístico

Estudos como os de Fernández (2013), Rubenstein (2010), Sankoff (1980) e Coveney (2003) mostram que o emprego do pronome de 2ª pessoa para referência genérica e indeterminada é uma estratégia comum em línguas como o Inglês, o Francês e o Espanhol, dentre outras, conforme exemplos (54), (55) e (56) abaixo, apresentando-se como uma variante não-padrão que coexiste com as demais estratégias de indeterminação dessas línguas, o que sugere que esta pode ser uma propriedade sintática relacionada à mudança do parâmetro *pro-drop*.

(54) ***You** never know when it's goin to rain.*⁴⁸

(55) *Parfois **tu** tombes des gens qui ont le même âge et qui ont une plus grande maturité.*
(Sometimes you come across people who are the same age but are more mature)

(56) *De los alguaciles dijo que no era mucho que tuviesen algunos enemigos, o **prenderte** o **sacarte** la hacienda de casa o **tenerte** en la **suya** en guarda y comer a **tu** costa...*⁴⁹
(About some sheriffs he said it was not much they had some enemies, as their job was to catch **you** or remove **you** from **your** home or take **you** into custody in theirs and eat at **your** expense...)

Consolida o fato de esse fenômeno ser atestado em várias línguas românicas a declaração de Coveney (2003):

*In several other Romance languages, the indefinite use of the 2nd personal singular pronoun is certainly attested, even though it is apparently less frequent than other strategies, such as equivalents of 'on' or pronominal verbs.[...] According to Benveniste, 2nd person forms in fact serve as equivalents of 'on' in many languages.*⁵⁰

⁴⁸ Exemplos (54) e (55) foram extraídos de Coveney (2003, p. 165).

⁴⁹ Exemplo (56) foi extraído de Fernández (2013, p. 93).

⁵⁰ Tradução: “Em várias outras línguas românicas, o uso indefinido do pronome de 2ª p. singular é, certamente, atestado, ainda que aparentemente é menos frequente do que outras estratégias, tais como equivalentes de 'on' ou verbos pronominais. [...] De acordo com Benveniste, formas de 2ª pessoa de fato servem como equivalentes de 'on' em muitas línguas”.

No PB esta estratégia de indeterminação com a 2ª pessoa (indireta) genérica – *você genérico* – igualmente tem se mostrado bastante produtiva em discursos tanto informais como formais, embora, diferentemente de línguas como o Inglês, Francês e Espanhol, não haja no PB um pronome equivalente a ‘*on*’ para referência genérica. Estudos como os de Duarte (1995), Callou e Martins (2003), Avelar & Callou (2000), Avelar (2004), Silva e Almeida (2010), entre outros, analisam esse emprego.

Em relação aos contextos de emprego, percebe-se que ‘você genérico’ no PB é empregado nos mesmos contextos que a 2ª pessoa indefinida ‘*you*’ o é em Inglês: sentenças com ‘você genérico’ justificam ações, não excluindo o falante, mas incluindo outras pessoas. Pela observação da fala espontânea tanto na comunidade de fala que integra a pesquisa empírica utilizada nesta tese, como em outros meios de comunicação, foi possível constatar maior incidência do emprego de ‘você genérico’ em argumentos, na apresentação de experiências comuns, semelhantemente aos contextos identificados por Rubenstein (2010) para a língua inglesa:

*A speaker can employ ‘you’ in arguments, to present an experience as common or universal. He can use it to justify taking an extra long turn of speech in order to tell a story, by presenting the story not as personally indulgent but as universally relevant. But these functions can be filled by other indefinite constructions as well.*⁵¹

Em um estudo do Português numa perspectiva diacrônica translinguística, Kato (2005) chama a atenção para essa similaridade do Português da América em alguns aspectos com o Inglês, língua que mais se afastou das germânicas e, em outros aspectos, com o Francês, língua que mais se afastou das irmãs românicas.

Em Francês, o uso indefinido de *tú/vous* é favorecido por informantes que descrevem em detalhes vários processos em geral. O surgimento desse uso pode ter sido motivado pela sobrecarga de funções de *on*, uma vez que no Francês cotidiano *on* tem sido usado como pronome sujeito de 1ª pessoa do plural (COVENEY, 2003).

Conforme identificado por Laberge & Sankoff (1979), em Francês, os pronomes de 2ª pessoa *tú/ vous* só mais recentemente assumiram as funções indefinidas de *on* e, por isso, aparecem como menos adequados. Os dados apresentados pelos autores em várias pesquisas sugerem que a forma de 2ª pessoa é mais usada por falantes mais jovens e do sexo masculino,

⁵¹ Tradução: “Um falante pode empregar “você” em argumentos, para apresentar uma experiência tão comum ou universal. Ele pode usá-lo para justificar a tomada de um longo turno extra de uma conversa, a fim de contar uma história, apresentando a história não como pessoalmente indulgente mas como universalmente relevante. Mas essas funções podem ser preenchidas por outras construções indefinidas também.”

menos preocupados com a adequação de seu discurso (p. 432- 433). Laberge & Sankoff (1979) atribuem essa tendência em Francês não tanto a uma preferência para a 2ª pessoa pelos homens, mas sim a uma preferência por falantes mulheres para a forma mais apropriada de 3ª pessoa. Os autores argumentam que, uma vez que a 2ª pessoa é recém-chegada à função indeterminada em Francês, é desfavorecida por falantes mais cuidadosos ou para quem a fala é bem compensada.

Em Espanhol, a 2ª pessoa *tú* também é empregada com referência indefinida em alternância com *uno*, assim como ocorre a variação *one/ you* em Inglês. Porém, o estudo de Rubenstein (2010), que compara os empregos e as funções da 2ª pessoa indeterminada nas duas línguas, mostra que a construção com a 2ª pessoa foi produzida em percentuais mais baixos em Espanhol do que em Inglês.

Isso é bastante compreensível quando se observa que o Espanhol tem maior variedade de construções indefinidas ou impessoais do que o Inglês. O Espanhol conta com as funções impessoais para as construções de 1ª pessoa do singular e do plural (*nosotros, tu / vos / usted*), o *uno* para tempo indeterminado, o nulo ambíguo de 3ª pessoa do singular e plural, assim como uma construção que não tem equivalente em Inglês: a partícula impessoal ‘se’ (Haverkate, 1985, Morales, 1995 apud Rubenstein, 2010). Esta construção é extremamente produtiva em Espanhol e, de acordo com Morales (1995), é o pronome indefinido original da língua, e só recentemente o aumento do uso de *tú e uno* inespecíficos contribuíram para seu declínio gradual em algumas comunidades.

Uma possível explicação para esse comportamento similar entre as línguas em relação ao pronome de 2ª pessoa como estratégia de indeterminação/ generalização do sujeito pode ser buscada na teoria das propriedades paramétricas, segundo a qual, de acordo com Tarallo e Kato (1989) as línguas podem convergir em determinadas partes de sua gramática, revelando movimentos sincronizados. É exatamente o que observo nesse processo, e como as propriedades estão relacionadas ao parâmetro, é possível prever como uma língua, no caso o PB, pode realinhar as propriedades de seus parâmetros sintáticos a partir dos resultados intralinguísticos de outra língua, ou seja, é possível projetar resultados que ultrapassam os limites do intralinguístico para o universo do interlinguístico.

Nessa perspectiva, vale destacar que, ao que parece, o fenômeno está ocorrendo simultaneamente nas línguas analisadas: Espanhol, Francês e Português do Brasil, apresentando-se como recente em todas elas.

O estudo de Blondeau (2008) sobre a dinâmica dos pronomes em línguas em contato em Quebec constatou a ausência da 2ª pessoa com referência indefinida em Francês no século

XIX, o que confirma que a mudança em progresso identificada no século XX é recente. De acordo com o autor, a forma *tú* com referência indefinida em Francês cresceu drasticamente entre os anos 1971 e 1984, assim como o uso do marcador discursivo “*tu sai*”.

Fernandez (2013), ao tratar das construções pronominais impessoais em Espanhol, chama a atenção para o incremento de frequência das construções genéricas com a 2ª pessoa durante os últimos anos. Elas foram mencionadas pela primeira vez em gramática espanhola⁵² em 1969, e o primeiro exemplo autocriado foi fornecido vinte anos depois pela gramática de Seco (1989):

En el uso coloquial de hoy, el valor impersonal de uno es asumido frecuentemente por el pronombre personal tú y la forma tú del verbo: Vas por la calle tan tranquilo y te cortan el paso. (Seco, 1989:163 apud Fernandez, 2013: 93)

Processos de substituição de formas variantes, em se tratando de variação/ mudança sintática, acontecem quando “uma mudança sintática torna-se estilisticamente marcada e cede lugar a uma outra, enquanto permanece em uso em contextos que são pesadamente regulados por fatores estilísticos. Similarmente, outras formas variantes desenvolvem-se a partir de outras estilisticamente marcadas” (RISSANEM, 1986 apud Ramos, 1991).

O Francês do Canadá, por exemplo, passa por um processo de mudança no uso do pronome *on*, de referência genérica para específica (1ª pessoa do plural). Paralelamente, o pronome *tú*, de referência específica (2ª pessoa do singular) está sendo usado para referência genérica, suprimindo a lacuna deixada pelo avanço de *on* em direção à referência específica (LABERGE & SANKOFF, 1979).

No PB parece estar ocorrendo um processo similar. De acordo com Zilles (2005), a forma *a gente*, predominantemente genérica até o século XX, tem sido mais usada com referência específica (1ª pessoa do plural). O estudo de Zilles (2005) verifica o uso de ‘a gente’ em novos contextos, em análise qualitativa e quantitativa:

Quantitativamente, o uso de a gente na posição de sujeito da oração aumenta significativamente dos anos 1970 para os anos 1990. Qualitativamente, expande-se para novos contextos, onde antes não era possível. Assim, inicialmente usado apenas com sentido genérico, como no exemplo (1), estende-se para os contextos de referência específica, como no exemplo (2). Outra face da extensão é a de que a gente passa a ocorrer como pronome anafórico dentro da oração, como mostra a comparação entre (3) e (4) abaixo: (1) O estado, a gente tem que conhecer, né?

⁵² De acordo com Blondeau (2008 apud Fernandez, 2013), “a aparição tardia da expressão 2ª pessoa indefinida em gramáticas deve-se ao fato de a construção pertencer quase exclusivamente à língua oral, porque ela está longe de ser nova, uma vez que foi atestada no latim vulgar”.

- (2) Nós não tínhamos muito contato com outras crianças, a vó nunca deixou a gente sair assim [do] – do portão pra fora pra brincar com outras crianças.
- (3) A gente_i olha pra gente_i e vê que... (uso anafórico)
- (4) *A pessoa_i olha pra pessoa_i e vê que... (anáfora impossível com substantivo)

O emprego do pronome *a gente* com referência mais específica é também atestado por Omena (2003). De acordo com a autora:

... à medida que a forma *a gente* vai se estabilizando como pronome, substitui mais e mais a forma antiga. Nesse caso passa a ser usada mais frequentemente nos contextos de determinação como aconteceu com o *on* francês que, na alternância com o *nous* é caracterizado por Laberge (1977) como definido. (Omena, 2003, p. 68).

Em análise em tempo real de curta duração sobre o uso de *a gente* vs. *nós* segundo o traço [+/- determinado] do referente, Omena (2003) observa que no período entre 1980 e 2000, *a gente* passa a ser usado mais frequentemente nos contextos de determinação. A proximidade dos pesos relativos confirma essa falta de especialização da forma *a gente* já na década de 80: .53 com referência indeterminada *versus* .44 com referência determinada. Já a amostra de 2000 indica preferência por *a gente* para referência determinada: .61. *versus* .43.

Portanto, a proposição aventada neste estudo de que o PB está passando por processo semelhante ao Francês de substituição de pronomes, referenda-se no fato de que *você* pode estar suprimindo a carência de um pronome para referência genérica, causada pelo avanço de *a gente* em direção à referência específica. Ressalto que neste fenômeno de substituição pronominal é possível evidenciar dois princípios básicos da teoria sociolinguística: o encaixamento e o uniformitarismo.

Considera-se que a variação sintática é um tipo particular de mudança linguística em virtude de certas características que lhe são inerentes. Uma dessas características é que se trata de um processo linguisticamente motivado e altamente encaixado no sistema linguístico. Dessa forma, os fatores que a favorecem não estão exclusivamente em um determinado fenômeno em variação, mas em mudanças simultâneas, que estão ocorrendo ou já ocorreram em subsistemas linguísticos relacionados. Pode-se falar, então, em feixe de mudanças inter-relacionadas.

Assim, quando se analisa *você* como um pronome que ao passar a ser empregado como pessoal, contribuiu para a redução do paradigma verbal que, por sua vez, desencadeou o processo de preenchimento do sujeito na língua, pode-se perceber aí um feixe de alterações interligadas ao sistema pronominal. Por seu turno, o pronome até então usado para fazer

referência genérica no PB, *a gente*, passa a ser usado para referência específica, cedendo lugar a outro pronome, *você* genérico, assim como *nós* há décadas atrás foi substituído, em muitos contextos, por *a gente*, conforme nos mostram os estudos de Omena e Braga (1996), Lopes (1999) e Maia (2009).

De acordo com Omena e Braga (1996),

A ideia de coletividade do substantivo “gente” contribuiu para uma referência “indeterminadora”. Assim, a origem embasada na referência genérica, aliado a um espaço para referência genérica presente na forma ‘nós’ é que propiciaram que a forma lexical ‘a gente’ se tornasse forma gramatical, pronome de 1ª pessoa plural, entendido como não “eu + tu/você”, mas como “pessoa que fala + pessoa com que se fala”, podendo apresentar um valor “indeterminado, abrangente, genérico e até difuso.

De forma semelhante, Maia (2009) aponta que *a gente* suplantou *nós* para referência genérica porque havia um espaço para essa referência na forma pronominal *nós* que permitiu a entrada da forma lexical *a gente* em seu domínio discursivo. Traço esse coincidente com a história da variante *a gente*, conforme apresentado em Lopes (1999), que favoreceu que a forma *a gente* passasse a ocupar o espaço pronominal de 1ª pessoa do plural dominado pela forma *nós*.

Essas evidências de mudanças que ocorreram no passado nos ajudam a interpretar os resultados das investigações atuais. Esta tese atua, assim, sob o pressuposto do princípio da uniformidade segundo o qual “os mesmos mecanismos que operaram para produzir as mudanças em larga escala do passado podem ser observados em ação nas mudanças que presentemente ocorrerem à nossa volta” (LABOV, 2008, 192).

Outra questão interessante concernente à marcação genérica dos pronomes de 2ª pessoa é abordada por Laberge & Sankoff (1979) em relação aos mecanismos que permitem esta leitura. Para os autores, uma razão por que o falante indetermina o agente é para justificar ações tomadas ou emoções experienciadas por um indivíduo referido. Os pronomes indefinidos podem fazer isso por minimizar a ligação pessoal da experiência descrita ou por apresentá-la como compartilhada por uma comunidade em geral. Ao usá-los, o falante efetivamente se remove de uma possível censura por parte do interlocutor. Nas palavras de Laberge & Sankoff (1979, p.430):

There is a great difference between the discursive effect of a sentence like —When I get drunk I wake up with a headache and that of —When one gets drunk one wakes up with a headache[...] The first can call forth a

*reaction of disapproval from one's interlocutor, but he or she can do little more than disagree with the second.*⁵³

Assim, a 2ª pessoa é um entre os muitos recursos disponíveis para se fazer generalizações, e compartilha algumas propriedades com as outras construções indefinidas. Em qualquer referência indefinida, o falante evita aludir a pessoas específicas.

Embora sejam intercambiáveis *you/ one*, Haverkate (1985:1 apud Fernández, 2013) nota que, por se referir a um participante no discurso, a 2ª pessoa tem um efeito defocalizante menor, mas Kitagawa & Lehrer (1990 apud Fernández, 2013) consideram que o envolvimento do destinatário no discurso por meio do uso do pronome de 2ª pessoa tem um efeito mais profundo. Para esses autores, o uso da 2ª pessoa em uma declaração generalizada convida o ouvinte a aprovar algum tipo de papel ou de alguma forma participar, representacionalmente, na situação descrita: o falante está solicitando que o interlocutor aceite um papel hipotético, que ele/ela coloque a si mesmo no lugar do falante. Segundo Kitagawa & Lehrer (1990 apud Fernández, 2013), esta troca de papéis influencia ativamente a relação criada entre falante e ouvinte.

Os autores argumentam que a construção com a 2ª pessoa indefinida é usada para expressar um conhecimento estrutural, incluindo “situações de inserção”, “formulações de moral ou truísmo” e “dramas da vida”. Essas três categorias podem ser explicadas como retratando conhecimento sobre “como as coisas funcionam”, “como as pessoas são” e “como as coisas acontecem”.

3.3.2 A sintaxe das construções existenciais com *ter*

O emprego de *ter* como verbo existencial tem sido apontado como uma das diferenças mais relevantes entre o PB e o PE, razão por que muitos estudos têm se dedicado à análise desse fenômeno no PB. Estudos como os de Avelar (2009 a, b) e Avelar & Callou (2007, 2011) apontam que mudanças no paradigma flexional no PB podem ter desencadeado a emergência de *ter* como o verbo preferencial das construções existenciais e, por extensão, as condições para licenciamento de sujeito nulo: visto que o PB se afastou do comportamento tipicamente *pro-*

⁵³ Tradução: Há uma grande diferença entre o efeito discursivo de uma frase como: Quando eu fico bêbado, eu acordo com uma dor de cabeça e de: Quando se fica bêbado se acorda com uma dor de cabeça [...] O primeiro pode evocar uma reação de desaprovação de um interlocutor, mas ele ou ela pode fazer pouco mais do que discordar do segundo.

drop do PE, sujeitos nulos deixaram de ser naturalmente interpretados como referenciais, tendendo a receber uma interpretação indefinida.

Enquanto uma construção como (57),

(57) Tinha muito dinheiro dentro da bolsa.⁵⁴

no PE é interpretada como uma sentença possessiva com sujeito referencial nulo, falantes do PB tendem a interpretá-la como uma sentença existencial pela impossibilidade de associar um termo interpretado como possuidor em construções em que *ter* não apresenta um sujeito foneticamente realizado. Ou seja, a redução do paradigma flexional no PB gerou a perda do *pro* referencial, o que impede a atribuição de uma interpretação possessiva ao sujeito nulo das sentenças com *ter*, levando a uma reanálise das construções possessivas em construções existenciais, posto que estas dispensam a realização de um sujeito pleno, e o nulo na posição de sujeito tende a ter interpretação indefinida.

Considerando-se o fato de o conceito de existência relativamente às construções existenciais com *ter* não ser bem delineado, em concordância com Avelar e Callou (2011), para este trabalho será considerada a sintaxe das orações e não os valores atribuídos aos verbos que se prestam à expressão de existência. Portanto, quanto à configuração sintática das construções com *ter*, PE e PB diferem-se em que neste, o verbo *ter* é largamente empregado em orações consideradas impessoais, por dispensarem a presença de um argumento na posição de sujeito (Avelar e Callou, 2011).

Entretanto, o que se observa no PB, é que, além do pronome nulo (expletivo) na posição de sujeito de existenciais com *ter*, esta posição tem sido frequentemente preenchida por pronomes indeterminadores como *você_{gen}*, *se*, *a gente* e *nós*, conforme os exemplos de (7) a (11) apresentados na introdução e aqui retomados:

(7) *Ø* não tem investimento em cultura no Brasil.

(8) Não *se* tem investimento em cultura no Brasil.

(9) *Você/ cê* não tem investimento em cultura no Brasil.

(10) *A gente* não tem investimento em cultura no Brasil.

(11) *Nós* não temos investimento em cultura no Brasil.

⁵⁴ Exemplo extraído de Avelar & Callou (2011).

O PB passa, assim, a contar com estratégias alternativas à estrutura existencial impessoal, o que atende às demandas do sistema, que está mudando em direção ao preenchimento do sujeito.

Uma alternativa para estabelecer coerentemente a emergência de existenciais com *você_{gen}* é apontada por Avelar & Callou (2011), em relação ao Kato & Duarte (2003) chamam de “condição anti-V1”, da seguinte forma. Kato & Duarte (2003) em análise da fala observam que falantes do PB tendem a evitar a ocorrência do verbo em primeira posição e inserem, em posição pré-verbal, tanto constituintes argumentais quanto não-argumentais. Os exemplos em (a) a seguir, apresentados pelas autoras, ilustram essa condição.

(58) a. Isso *não tem nem dúvida*.⁵⁵

b. *Não tem nem dúvida sobre isso*.

(59) a. Eles *ainda faltavam receber o dinheiro do patrão*.

b. *Ainda faltava eles receberem o dinheiro do patrão*.

Para Kato & Duarte (2003), a condição anti-V1 não deve ser tratada como um requerimento do componente sintático, mas do componente fonológico, que impõe restrições para a realização do verbo em primeira posição no PB por razões que ainda precisam ser compreendidas. Por isso, construções com verbo em primeira posição são bem formadas, embora cada vez menos frequentes no PB.

De qualquer forma, de acordo com as autoras, os dados refletem uma inovação em relação ao PE, que é a tendência a inserir elementos argumentais ou não argumentais na posição de sujeito gramatical, evitando a ocorrência do verbo em posição inicial.

Nesse contexto, a emergência de construções existenciais com *ter* em que a posição de sujeito é preenchida por formas pronominais será analisada como uma tendência do PB ao preenchimento da posição de sujeito, a qual sinaliza uma inovação encaixada em um conjunto mais amplo de mudanças na gramática do PB, em contraste com o PE: a redução do paradigma flexional restringiu os contextos de licenciamento do sujeito nulo (cf. Duarte, 1995), o que está ampliando o preenchimento do sujeito para contextos de nulo genérico e nulo expletivo.

Sob outra perspectiva, Franchi et al (1998) observam que a detematização da posição de argumento externo não exclui, no caso das existenciais do PB, outras estratégias de

⁵⁵ Exemplos (58) e (59) foram extraídos de Avelar & Callou (2011, p. 258) ex. 10 e 11.

indeterminação do "sujeito". Há construções “cuja interpretação corresponde às existenciais, mas em que um relativo "esvaziamento" da posição de "sujeito" se faz mediante outro mecanismo sintático: a presença de um "sujeito" indeterminado - como *você, a gente* - em contextos bem claros que excluem uma interpretação predicativa”.

Em vista dessa possibilidade de construções existenciais formadas com o verbo *ter* apresentarem DPs pronominais ocupando a posição estrutural de sujeito, as quais não são interpretadas como possessivas, mas como um tipo específico de construções existenciais, pois que se prestam à expressão da existência, tais construções passam a ser consideradas como *ter existencial pessoal*, tendo em vista um tratamento unificado para os dados desta pesquisa que considere todas as ocorrências de *ter_{ex}* em que os pronomes *você_{gen}, se, a gente* e *nós*, variantes deste estudo, sejam realizados na posição de sujeito dessas construções. A classificação de *ter* existencial como verbo pessoal sustenta-se no estudo de Berlinck, Duarte e Oliveira (2009) que destacam dois aspectos relevantes sobre as construções existenciais com *ter/haver* no PB: a substituição de *haver* por *ter* e a implementação de sentenças pessoais com *ter*.

É lícito afirmar, portanto, que a presença de pronomes pessoais como sujeito gramatical de sentenças existenciais com *ter* conforma-se com as demais alterações na gramática do PB, como o aumento de sujeitos referenciais plenos e a emergência de estratégias de indeterminação do sujeito com pronomes nominativos em substituição ao uso de *se* indeterminador, em uma clara evidência do encaixamento linguístico da variação.

3.3.3 Propriedades do verbo *ter_{ex}* e sua predicação

Uma explicação para a emergência de *ter* como verbo existencial no PB, proposta por Avelar (2009a), parte da constatação de que muitos verbos no PB podem dispensar um sujeito nominal na presença de um locativo, como em (60):

(60) *(Ali) naquela loja vende bastante roupa.*

Para esclarecer a razão por que a presença de um sintagma locativo em sentenças existenciais com *ter* produz uma interpretação existencial no PB, Avelar (2009a) baseia-se na perspectiva de Freeze (1992), que postula que a seguinte condição está presente na gramática universal: estruturas copulares com um sintagma locativo em sua posição de sujeito recebem uma interpretação existencial. Essa propriedade é ilustrada por Freeze (1992) em sentenças

como ‘*There is a man in the room*’, em diversas línguas, tais como Hindi, Chamorro, Tagalog e Filandês.

Para Avelar (2009a), o fato de que sentenças com *ter* na presença de um locativo PP na posição de sujeito recebem interpretação existencial é coerente com o fato de que a sentença possessiva tem estrutura copular. Para explicar a implementação de sentença existencial com *ter* no PB, Avelar (2009a), em uma abordagem não lexicalista (HALLEY; NOYER, 2003 apud Avelar, 2009), argumenta que as construções possessivas, copulativas e existenciais apresentam a mesma estrutura de base: *ter* resulta de características correspondentes à cópula ‘estar’ e a preposição ‘com’. Desta forma, para Avelar (2009a, p. 10), “a inovação no PB não envolve exatamente uma reanálise de *ter* como um verbo existencial, mas o uso da estrutura possessiva (com características copulares e preposicionais) para expressar existência”.⁵⁶

Para sustentar esta proposta, Avelar (2009a) mostra o paralelismo existente em sentenças com ‘estar com’ e *ter* existencial, em que o único contraste é aspectual, sendo que a estrutura copulativa, mas não a existencial, apresenta episódio temporário, como se vê em (61):

- (61) a. No centro da cidade *tava com* um engarrafamento enorme⁵⁷.
b. No centro da cidade *tinha* um engarrafamento enorme.

Em relação à estrutura das sentenças, tem-se que em ambas o SN pós-verbal é complemento da preposição e o sintagma locativo está na posição de sujeito. Para Avelar (2009), esse paralelismo sugere que *ter_{ex}* também é obtido a partir das propriedades de ‘estar com’ em contextos existenciais. A ocorrência do locativo em posição final nas existenciais não constitui problema para a proposta de Avelar, que postula que a posição pré-verbal pode ser ocupada por um pronome dêitico co-indexado com o sintagma locativo, assim como em construções de tópico com sujeito em posição final na sentença.

Quanto às restrições para sentenças com *ter_{ex}*, Avelar (2009a) sugere que há uma conexão entre a presença do sintagma locativo e a interpretação existencial, de forma tal que o autor propõe que “a presença de um sintagma locativo livra sentenças com *ter* sem um sujeito de serem mal formadas no PB”.⁵⁸

⁵⁶ Original: “The innovation in BP does not exactly involve a reanalysis of *ter* as an existential verb, but the use of the possessive structure (with copular and prepositional features) to express existence”. (Avelar, 2009a, p.10)

⁵⁷ Exemplo extraído de Avelar (2009a).

⁵⁸ Original: “The presence of a locative phrase rescues *ter* sentences without a subject from being ill-formed”. (Avelar, 2009a, p. 6)

Avelar (2009a) mostra que os fatores que levaram à interpretação existencial do verbo *ter* no PB, como a ausência de sujeitos nulos referenciais e a possibilidade de inserção de PPs locativos na posição de sujeito, não ocorrem em PE. O autor associa esta oposição entre o PB e o PE com as propriedades dos traços-*phi* de T: elementos sem traços-*phi* interpretáveis (como PPs e advérbios) podem aparecer na posição de sujeito de sentenças transitivas no PB, porque o PB conta com uma versão de T sem traços-*phi* não-interpretáveis para serem avaliados.

O mesmo não ocorre no PE onde T sempre apresenta traços-*phi* completos, uma situação que requer a presença de um elemento com traços-*phi* interpretáveis (DP/NP) acessível à sonda de T. Portanto, em contraste com o PB, falantes do PE não utilizam a mesma estratégia de *ter* com sentido existencial porque o PE conta com sujeitos nulos referenciais, que atribuem a leitura possessiva a *ter*, e não licenciam PPs locativos na posição de sujeito de sentenças com *ter*. Avelar confirma, assim, que no PE somente o verbo *haver* pode ser usado em contextos existenciais.

Outro fator que evidencia a estrutura possessiva na base de *ter* existencial no PB, de acordo com Avelar (2009a), tem a ver com o uso do pronome *você_{gen}*, o qual é licenciado em sentenças existenciais com *ter* (62), mas não com *haver* (63).

(62) (Você) tem prédios altíssimos em Nova York.

(63) *(Você) há prédios altíssimos em Nova York.

O que condiciona este licenciamento, para Avelar (2009) é que o pronome *você_{gen}* não pode ser diretamente inserido em [Spec TP], mas em uma posição temática para ser licenciado. Nessa perspectiva, esta posição deve ser o especificador da preposição [Spec PP], reservado, de acordo com o autor, para elementos destinados para a posição de sujeito de sentenças existenciais, para onde ele se move. O mesmo não ocorre com o verbo *haver*, porque não tem preposição subjacente em estruturas com *haver*, visto que este verbo não expressa posse em Português. Esses fatos confirmam que o emprego de *ter* como verbo existencial no PB envolve não somente a substituição de *haver* por *ter*, mas a reanálise de uma estrutura possessiva capaz de expressar existência.

Por dar conta de explicar o licenciamento de um SN foneticamente realizado na posição de sujeito de construções existenciais com *ter* é que adoto neste estudo a proposta de Avelar (2009) de que *ter* existencial e possessivo possuem a mesma estrutura de base. Em outros termos, considero que como verbo existencial, *ter* guarda a memória da estrutura sintática de que deriva, e uma posição de sujeito se manifesta em todas as construções.

Entretanto, na contramão desta proposta de Avelar (2009a) de que o sintagma locativo está na posição de sujeito nas construções existenciais, no PB há ocorrências de construções existenciais que não exibem qualquer elemento de natureza locativa em sua *coda*, em que nem mesmo é necessário pressupor um constituinte locativo apagado para sua ancoragem contextual, como no exemplo (64) a seguir.

Esse argumento baseia-se ainda em Avelar (2004), quando o autor afirma que um dos contrastes entre as existenciais com *ter* e os casos de ergativização que invalidam uma possível correlação entre essas construções tem a ver com o *status* do locativo. De acordo com Avelar (2004), se houver outro constituinte com potencial temático, sintagmas com expressões locativas não são necessários para a boa formação das existenciais, como em (64), o mesmo não se verificando nas transitivas ergativizadas, com em (65).

(64) Tem muito carro baratinho (em Campinas)⁵⁹.

(65) Rouba muito carro baratinho?? (em Campinas).

Em consonância ao fato de que há construções existenciais que não exibem qualquer elemento de natureza locativa em sua *coda*, bem como pela divergência apontada entre as construções existenciais com *ter* e as ergativizadas, é que não defendo a proposta de Avelar (2009a) sobre a conexão entre a presença do sintagma locativo e a interpretação existencial.

Em relação à predicação de *ter* existencial é importante abordar que, devido à leveza semântica do verbo existencial, ele é incapaz de atribuir papel temático aos seus argumentos. Portanto, para sua boa formação, construções existenciais precisam de um outro constituinte, além do DP pós-verbal, o que se convencionou chamar de *coda*, a partir da proposta de Milsark (1974, 1977 apud Avelar, 2004). Assim, de acordo com Avelar (2004), no caso do PB, em função de *ter* existencial ser defectivo quanto à atribuição temática, é preciso uma categoria com potencial predicativo para atribuir papel semântico ao seu complemento. Os contrastes a seguir evidenciam essa questão:

(66) Tem muitas universidades* (que não têm curso de linguística).⁶⁰

(67) Tem vários livros da biblioteca* (emprestados com a Ana).

⁵⁹ Exemplo extraído de Avelar (2004, ex. 61)

⁶⁰ Exemplo extraído de Avelar (2004, p.63)

Por meio dos exemplos acima pode-se atestar que a predicação do verbo *ter_{ex}* é realizada por um outro elemento predicator dentro da *coda*⁶¹. Assim, em função de *ter* com sentido existencial não estabelecer nenhuma relação temática na sentença, a predicação se estabelece entre os dois elementos da *coda* das orações existenciais (cf. Viotti, 1998; Avelar, 2004). Nessa perspectiva, tanto o constituinte locativo quanto os outros constituintes formadores da *coda* consistem no centro temático da sentença.

É nesse sentido que se explica a detematização⁶² do agente de *ter_{ex}*: como esse verbo não seleciona argumentos, “as relações semânticas estabelecidas não estão inscritas como propriedade temática do verbo *ter_{ex}* (...): a interpretação depende do sentido dos sintagmas nominais e preposicionados que formam as expressões entre as quais o verbo *ter* expressa uma relação muito abstrata e inespecífica”⁶³ (FRANCHI ET AL, 1998, p. 21). A detematização de *ter_{ex}* pode ser evidenciada em (68):

(68) Me preocupo com o humano se embananando ele sozinho com as coisas que ele cria, sabe? Porque **você tinha** civilizações antigas, mas o que ela criava, o que ela produzia era *muito* menos do que uma de hoje em dia cria, certo? (FRANCHI ET AL, 1998: ex. 90 -SP, 343)

O sentido genérico desta construção (68) evidencia o não estabelecimento de relação temática entre o sujeito gramatical e o verbo, uma vez que *tinha* não pode ser interpretado como "possessivo", nem *você* como "possuidor", já que não se possuem civilizações antigas que eram concedidas a outrem.

Assim, o licenciamento de um pronome genérico/ indeterminado na posição de sujeito gramatical evidencia que *ter_{ex}* pode ser considerado como um verbo detematizado, mas não necessariamente impessoalizado, por razões tais como a capacidade de licenciar uma forma pronominal foneticamente realizada na posição de sujeito e com ela estabelecer concordância, mas sem atribuir-lhe papel temático, uma vez que *ter_{ex}* não seleciona argumento externo.

⁶¹ "Coda" é o termo utilizado por Milsark (1974, 1977) e Reulan e Meulen (1987) para denominar a parte da sentença existencial que segue o verbo.

⁶² Detematização diz respeito à perda sofrida pelo verbo da capacidade de atribuir papéis temáticos aos argumentos com os quais se combina.

⁶³ O artigo consultado é uma versão *online* publicado em DELTA, portanto não tem páginas numeradas.

3.4 PROPOSTAS DE ANÁLISE

3.4.1 Do caráter não expletivo de *você_{gen}* em existenciais

A análise de empregos de *você_{gen}* em sentenças existenciais, à luz dos pressupostos teóricos de Freeze (1992), atestou que esse pronome não constitui uma forma expletiva, contrariando a hipótese inicial motivadora desta pesquisa, de que nesses contextos *você* poderia ser assim considerado. Essa possibilidade foi aventada a partir da suposição de que se *você_{gen}* não fosse considerado expletivo, sentenças existenciais com a ocorrência desse pronome não seriam aceitáveis, posto que como impessoais, tais verbos não selecionam argumento externo.

A verificação do caráter expletivo de *você_{gen}* teve como ponto de partida a hipótese preliminar da teoria proposta por Freeze (1992) sobre paradigmas locativos universais, de que em línguas V1, em construções existenciais com expletivo, denominados por Freeze (1992) de proformas existenciais, tais proformas são lexicalmente locativas.

Em línguas como o Francês, a proforma *y* é cognata com *hi* Catalão e *y* Espanhol, e provavelmente cognata com *ci* Italiano. Similarmente, sua morfologia é explicitamente locativa, uma vez que as proformas consistem em uma locução preposicional com a 3ª pessoa do pronome acusativo ou de uma forma historicamente derivada de tal locução. Ainda como característica da proforma, o fato de ela não poder se mover (*move alfa*) sugere que não é argumento ou não está numa posição argumental.

De acordo com Freeze (1992), em poucas línguas, as sentenças existenciais têm uma proforma co-ocorrendo com um constituinte em outro lugar na sentença; a proforma distingue as sentenças existenciais dos predicados locativos. São exemplos de proforma: *there* em Inglês, *y* em Francês, *ci* em Italiano.

Constitui evidência sintática para a locatividade da proforma o fato de que a distribuição complementar intralinguística do predicado locativo e do existencial permanece constante com a proforma existencial. Outra evidência, nas línguas estudadas por Freeze (1992), é que o tema é sujeito do predicado locativo (69)a. e o locativo é o sujeito da construção existencial (69)b. A presença ou ausência da proforma é sintaticamente irrelevante. Entretanto, a proforma ocorre somente com sujeitos locativos, e nunca com sujeitos temáticos.

(69) a. Predicado locativo: *The book is on the bench.*

b. Existencial: *There is a book on the bench.*

Esse fato sugere que a proforma desempenha um papel especial no paradigma locativo. Ou seja, a proforma tem uma estrita relação de precedência com a locução locativa. A posição da proforma existencial provê maiores evidências para sua locatividade. Freeze (1992) argumenta que a proforma é o *spell out* de um traço em INFL⁶⁴ que é lexicalizado na PF (*phonological form*). Essa afirmação se sustenta com as seguintes constatações: a proforma é sempre adjacente a elementos em INFL e nunca à locução locativa; ela precisa preceder a locução locativa. Em línguas SVO, a proforma é adjungida à esquerda da cópula. As restrições positivas para que a proforma seja contígua a elementos de INFL (cópula e *tense*) e outras propriedades posicionais da proforma existencial em línguas V1 sugerem que ela é considerada parte de INFL.

Em Espanhol, que não tem proforma lexical, pode-se observar evidência da posição da proforma em INFL. A proforma *y* contígua com *ha* (tempo presente da cópula existencial em Espanhol) produz uma forma lexical inalisável *hay*, testemunhando íntima relação entre INFL e a proforma existencial.

Ainda, conforme mostrado por Freeze (1992), proforma não é sujeito ou qualquer outro argumento; não há, por consequência, nenhuma posição adicional na qual ela possa ser gerada ou da qual ela possa ter sido movida. O fato de ela não poder se mover sugere que ela não é argumento ou não está em posição de argumento.

Em sentenças sem pronome sujeito pleno, os pronomes *ci* italiano e *hi* do catalão, em função da posição, podem erroneamente ser tomados como estando na função de sujeito, mas na realidade não estão. Isso é confirmado pelo fato de que, como em francês, eles podem ser diretamente precedidos pela negativa, conforme se pode observar no exemplo em italiano: “*Non ci sono uomini in casa*”. O locativo em posição final também é consistente com a análise na qual o sujeito locativo é posposto.

Finalmente, a proforma é inseparável de AGR (quando não há cópula lexical). A mais simples constatação desses fatos é que a proforma é a pronúncia de um traço em INFL; a diferença entre línguas com ou sem proforma existencial, então, reside em PF. Em línguas com proforma existencial, o traço INFL [+LOC] é pronunciado como proforma existencial. Em línguas sem proforma existencial esse traço [+LOC] falha em ser lexicalizado.

Em trabalho sobre o *status* do suposto expletivo no PB, *você_{gen}*, Avelar (2009b) argumenta contra a ideia de que o pronome seja um expletivo quando é realizado em sentenças

⁶⁴ INFL - sintagma flexional (*inflection*) - responsável por agregar traços de tempo e de concordância (*cf.* Chomsky, 1981).

com *ter* existencial. Uma vez que o pronome mantém caráter indefinido, ele carrega um valor semântico, o que não é possível no caso dos expletivos. Outro ponto que invalida a caracterização de *você_{gen}* como expletivo, de acordo com o autor, diz respeito à impossibilidade de inserção do pronome em sentenças com o verbo *haver*. Posto que *haver* também é existencial, seria de esperar que sentenças com esse verbo licenciassem o pronome expletivo.

A proposta de Avelar (2009b) é a de que uma versão relevante de *você_{gen}* corresponde a um pronome indeterminado idêntico aos casos apresentados em contextos não existenciais, como em (70) abaixo, intercambiável com existenciais como (71):

(70) ***Você pode encontrar*** roupas bem baratinhas no centro. (Avelar, 2009b, p. 2)

(71) ***Você tem*** roupas bem baratinhas no centro.

Por outro lado, a proposta de Avelar (2009b) se conforma com a teoria de Freeze (1992) no tocante ao *status* locativo do pronome em construções existenciais. Para lidar com o possível conflito entre a presença de *você* e o caráter impessoal da sentença existencial, e com a questão de como esse pronome é licenciado como sujeito de sentenças impessoais, Avelar (2009b) argumenta que não há relação temática entre o verbo existencial e o pronome, mas entre o pronome e a categoria predicativa encabeçada pela locução locativa. Portanto, Avelar (2009b) propõe que os passos derivacionais para se gerar uma sentença existencial com *você* são: *você* é inicialmente conectado em [Spec LocP] onde recebe papel temático (sujeito indeterminado do sintagma locativo)⁶⁵; considerando que [Spec PP] é uma posição sem Caso, *você* move para [Spec TP] para satisfazer requerimentos formais relacionados com Caso e traços de EPP.

Assim, de acordo com o autor, uma sentença como: *Você tem muitos castelos na Europa* tem a seguinte representação derivacional:

(72) [TP *você_i* [T' *tem* [XP [DP *muitos castelos*] [X' [LocP *ti na Europa*]]]]]

⁶⁵ De acordo com Avelar (2009b, p.10): “If existential, the pronoun is inserted in LocP, where it is interpreted as the indeterminate subject of the locative predication” (Avelar, 2009).

⁴⁶ De acordo com Avelar (2009b, p.6): “Sentences with *ter* are subjected to restrictions that reveal to us a connection between the presence of a locative phrase and the existential interpretation. As shown in (8), the presence of a locative phrase rescues *ter* sentences without a subject from being ill-formed in BP, providing it with an existential meaning”.

A principal razão para Avelar (2009b) propor um *link* entre o pronome e a locução locativa está no fato de que o suposto expletivo, de acordo com o autor, é licenciado somente se a sentença existencial apresentar um aparato locativo. Deve-se considerar, contudo, que no PB *você_{gen}* é empregado em sentenças com *ter* com sentido existencial sem a presença de um locativo, como no exemplo (73):

(73) **Você tem** provas de linguística fáceis.

Além disso, é possível perceber um problema nesta proposta: se existe um *link* entre o pronome e a locução locativa, o pronome necessariamente se caracteriza como proforma ou pronome pleonástico, visto que estes são instâncias de elementos dêiticos originários ou modificadores de lugar, no caso. A coindexação do pronome *você* ao locativo pode atribuir-lhe implicitamente um caráter expletivo. Conforme mostra a comparação com uma proforma típica ‘*there*’ no exemplo (74) a seguir, o pronome indeterminado *você* não parece compatível com essa adjunção.

(74) a. *there (europe)*
b.* *você (europa)*

Esta proposta de não coindexação do pronome *você_{gen}* com a locução locativa tem respaldo em Franchi *et al* (1998) que postulam a presença de um expletivo nulo nas construções existenciais como parte da representação de uma estrutura sintática subsistente, mas que, para os autores, “não parece adequado nem confundí-lo com um locativo (Freeze, 1992), nem identificá-lo, mediante coindexação a qualquer título, com o SN argumento”. Como *você_{gen}* ocupa a posição do expletivo nulo e não altera o valor existencial da construção com *ter*, de forma semelhante ao nulo, não parece que esse pronome seja compatível com uma correlação com o locativo. O ponto de vista de Kato (2000) de que no PB não há nenhum um pronome locativo fraco também corrobora este entendimento.

Diante dessas questões de natureza sintática e da posição ocupada pelo pronome *você_{gen}* em sentenças com *ter* existencial, reconheço que, diferentemente da proforma existencial em línguas como o Francês, Inglês e Italiano, o pronome *você_{gen}* é tipicamente não locativo, morfológica ou sintaticamente. Ainda que pudesse ser considerado um elemento dêitico, *você_{gen}* não poderia estar coindexado ao locativo, visto que não remete a um lugar, no

sentido de ser parte ou representante dele, dado seu caráter genérico. Por essa razão, de expressar genericidade, o pronome *você_{gen}* precisa ser legível no componente semântico da gramática, ao contrário do que se espera de um autêntico expletivo.

Assim, pelas razões de não apresentar uma estrita relação com a locução locativa, e de não ter outros traços de expletivo, como não possuir carga semântica nem se referir a nada no contexto, o pronome genérico *você* não pode ser considerado uma proforma ou expletivo.

Por outro lado, conforme se tem discutido, esse pronome não é sujeito semântico das construções existenciais. Uma vez que *ter_{ex}*, como um verbo detematizado, não pode atribuir papel temático ao pronome em posição pré-verbal, este não pode ocupar posição argumental. Portanto, as construções com *ter_{ex}* no PB com pronome foneticamente realizado na posição de sujeito constituem um caso à parte em relação às construções existenciais interlinguisticamente.

Cabe analisar a função sintática e a localização desse pronome nas sentenças existenciais no PB.

3.4.2 Do caráter indeterminado de *você* genérico

Apesar da função apresentativa dos verbos existenciais, de expressar a existência do argumento do verbo, conforme discutido, o pronome *você_{gen}* se realiza em posição de sujeito das construções existenciais com *ter*, sem, contudo, estabelecer relação temática com o verbo. Isso explica por que a inserção desse pronome não implica, em construções como (75) a seguir, uma leitura com possuidor referencialmente indeterminado: como não é um sujeito semântico de *ter*, o pronome não recebe qualquer papel temático relacionado à interpretação possessiva atribuído pelo verbo, o que inviabiliza a expressão de posse⁶⁶.

(75) *Cê* não **tem** mais burca na Síria.

⁶⁶ Deve-se ressaltar que, conforme mostrado por Avelar e Callou (2011), é possível que construções existenciais com *você_{gen}* recebam uma interpretação possessiva se realizadas em contextos como o apresentado em: ‘*Se você_{gen} tem muitos castelos na Europa, você_{gen} deve ser rico*’, em que uma oração condicional trazendo *você_{gen}* antecede outra ocorrência do mesmo pronome em um predicado nominal. A única interpretação possível para o pronome, nesse caso, é a de um possuidor referencialmente indeterminado, dado que a oração com *ter* receberá obrigatoriamente a interpretação possessiva. Essa construção recebe a seguinte leitura: *se uma pessoa possui muitos castelos na Europa, essa pessoa deve ser rica*. Nesse caso, *você_{gen}* é sujeito semântico de *ter*, pois recebe papel temático do verbo, com sentido equivalente a de um pronome indeterminador como “uma pessoa”, “alguém”, e não sentido indefinido.

Diante do obstáculo gerado pela presença do pronome *você/cê* na posição de sujeito e pelo caráter existencial do verbo *ter*, o licenciamento de um pronome nessa posição sugere que *ter* não perde o caráter pessoal ao ser empregado com sentido existencial, ou seja, ele “guarda a memória” da estrutura sintática de que deriva (*ter_{poss}*), conforme discutido em 3.3.3.

Partindo então do entendimento de que construções de *ter_{ex}* não são estruturas impessoais, e de que não há qualquer relação temática entre *você_{gen}* e o complemento de *ter*, e descartada as hipóteses de este pronome constituir um expletivo e, ainda, de poder ser interpretado como sujeito indeterminado da locução locativa (cf. Avelar, 2009a), é possível considerar que essas construções com o pronome genérico *você* na posição de sujeito constituem construções de indeterminação do “sujeito”, com sentido genérico.

Note-se que uma sentença como (75) equivale a (76):

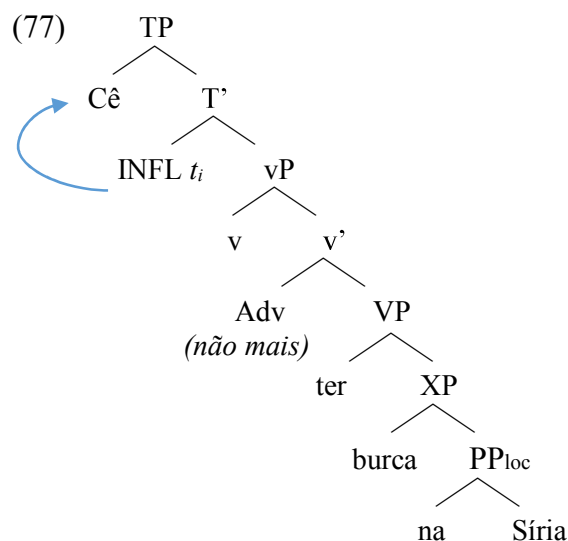
(76) Não **se tem** mais burca na Síria.

Assim, a relação não temática do pronome *você_{gen}* com o verbo *ter* em (75) permite considerá-lo um elemento indeterminador à semelhança de *se* indeterminador em (76). Como não recebe papel temático do verbo, pode-se supor que *você_{gen}* não absorve papel temático de outro constituinte na sentença.

Por não constituir argumento do verbo, o pronome *você_{gen}* não pode surgir em posição argumental, o que pressupõe outra posição na sentença na qual ele possa ser gerado ou da qual ele possa ser movido. Diante da suposição de que *você_{gen}* não absorve papel temático, proponho que este pronome é gerado em INFL, lugar onde é inicialmente inserido e posteriormente movido para SpecTP, para satisfação do requerimento EPP, e não na locução locativa, conforme proposto por Avelar (2009b) e Avelar e Callou (2011), mostrado na seção 3.4.1.

A derivação representada a seguir (77) ilustra a situação aqui proposta para a interpretação de *você_{gen}* em uma sentença como (75) “*Cê não tem mais burca na Síria*”: o pronome genérico é inserido no interior de INFL, e desse locus é movido para a posição de sujeito gramatical (SpecTP). Se esta proposta estiver no caminho correto, é possível que o DP *você_{gen}* tenha seus traços de Caso valorados/ checados pela relação de concordância com Infl/T⁶⁷.

⁶⁷ De acordo com (Avelar, 2009b): ‘I am assuming that a DP in [Spec,PP] can have its Case feature checked/valued by an agreement relation with Infl/T’.



Esta proposta tem respaldo na estrutura de indeterminação da língua portuguesa, que tem o clítico *se* adjungido a INFL (T + *se* indeterminado), além de que, em termos minimalistas, Spec-T/Infl é a posição naturalmente destinada ao sujeito gramatical em línguas como o Português (Avelar e Callou, 2011).

Igualmente lança luz sobre esta proposta o argumento de Kato (2000) de que nas construções existenciais em Português é o afixo neutro \emptyset_2 que faz o papel de expletivo, o qual adjunge diretamente a *ter* (V+T) para checar Caso e traços-*phi*, uma vez que ele não é argumento. Se o pronome *você_{gen}* está ocupando a posição do afixo neutro \emptyset_2 nessas construções, pode-se supor que a posição de imersão é a mesma para ambos pronomes.

Corroborar ainda esta hipótese de *você_{gen}* como elemento indeterminador a teoria de Kato (1999) do surgimento de pronomes sujeitos fracos no PB para compensar a perda de *Agr* pronominal, especialmente em relação à redução fonética do pronome *você*, que de acordo com a autora, criou um clone fraco - *cê*, o qual eliminou a função de argumento de seu *Agr*, e agora pode aparecer apenas como um afixo do verbo.

Pode referendar também esta proposta de equivalência entre as construções com *você_{gen}* + *ter_{ex}* e *v* + *se*, o exemplo (78) de Franchi *et al* (1998) o qual, comparado ao (79), mostra o paralelismo sintático e interpretativo entre existenciais com um clítico *se*, que marca a supressão do argumento externo de *ter* com indeterminação do sujeito, e existenciais com *você_{gen}*:

(78) Eu acho que qualquer lugar é diferente daqui do Rio, do ponto de vista do clima. Qualquer ponto onde *você* andar por aí é diferente. — Aqui não *se tem* definição de coisa nenhuma.⁶⁸

(79) Eu acho que qualquer lugar é diferente daqui do Rio, do ponto de vista do clima. Qualquer ponto onde *você* andar por aí é diferente. — Aqui *você* não *tem* definição de coisa nenhuma.⁶⁹

Com base na explicação de Franchi *et al* para construções como (78), de que o *se* é o traço morfológico da supressão de um argumento externo, e considerando que a inserção de *você_{gen}* não altera o sentido existencial do verbo *ter*, explica-se o paralelismo sintático e interpretativo de *se* com *você* em sentenças de *ter_{ex}*: esse pronome marca a supressão do argumento externo, atribuindo sentido indeterminado ao “sujeito”.

Semelhantemente, em construções pessoais indeterminadas, o pronome *você_{gen}*, tanto na forma plena como na reduzida (*cê*), marca a supressão do sujeito, com manutenção de sentido indeterminado, como no exemplo (80) da fala espontânea, intercambiável com (81):

(80) *Cê* percebe claramente um movimento para enfraquecer a Lava Jato.

(81) Percebe-*se* claramente um movimento para enfraquecer a Lava Jato.

As construções predicativas indeterminadas diferem das existenciais em que naquelas a indeterminação é do sujeito semântico, que é argumento do verbo, enquanto nestas *você_{gen}* constitui sujeito gramatical, o qual não interage semanticamente com o verbo.

Sob a perspectiva de Holmberg (2010), discutida em 2.2, pode-se inferir que o sentido indefinido de *você_{gen}* também é decorrente da incorporação em T dos traços de ϕP , o que o torna equivalente ao *se* indeterminado, ou seja, uma vez que o pronome genérico inclui todo mundo, sua referência é irrestrita, o que significa ausência de traço definido em T.

Outra possível correlação entre o pronome *você* e o clítico *se* é que eles compartilham os mesmos traços de localização, uma vez que ambos predicados contêm essencialmente os mesmos constituintes, mas em diferentes ordens, como se pode constatar nos exemplos (78) e (79). Nessa direção, Vitral (1996) compara o comportamento análogo da forma reduzida do pronome *você* (*cê*) e o do clítico *se*, mostrando que ambos podem ser utilizados em contextos

⁶⁸ Exemplo extraído de Franchi et al (1998: Ex. 94 - RJ, p. 168)

⁶⁹ Exemplo criado a partir do (78).

semelhantes, e também aponta como diferenças as posições estruturais ocupadas pelos pronomes: o primeiro sempre acima de negação e o último sempre abaixo dela.

Assim, proponho que o pronome *você_{gen}* exerce a função de elemento indeterminador no PB, concorrendo com o clítico *se*, e ensejando que predicados existenciais mantenham traços semânticos de verbos apresentacionais. Ou seja, a interpretação para *você_{gen}* como elemento indeterminador não deve ser confundida com a de um possuidor indeterminado em contextos de *ter_{ex}*, nos quais há restrições pragmáticas para esta leitura, conforme se constata em exemplos como o (75). Além desse aspecto, explica a não leitura possessiva dessas construções de *ter_{ex}* com *você_{gen}* na posição de sujeito gramatical o fato de que como o pronome não é inserido na posição de argumento externo do verbo, o papel temático de possuidor não lhe é atribuído e, conseqüentemente, a marcação de posse não é instanciada.

A esse respeito, deve-se ressaltar que a possibilidade de ocorrência de *ter* + *se* indeterminado no PE em construções possessivas, onde o *se* é largamente empregado em contraste com o PB, poderia invalidar esta análise, se se considerar que sentenças com *ter* + *você_{gen}* no PB correspondem a sentenças com *ter* + *se* indeterminado em PE. Como no PE *ter* não é usado com sentido existencial, esse fato poderia levar a considerar os casos de *você* + *ter* também como possessivos e não existenciais.

Para elucidar essa questão, Avelar (2009b) argumenta que o *ter* + *se* indeterminado no PE não é utilizado nos mesmos contextos que *você* + *ter* o é no PB. Os contextos exemplificados em (82) a seguir, por exemplo, rejeitam o *se* indeterminado em PE, mas licenciam *você* no PB.

- (82) a. * Tem-se oito planetas no sistema solar. (PE)⁷⁰
Ter-SE oito planetas no sistema solar.
Lit .: Um tem oito planetas no sistema solar.
- b. Você tem oito planetas no sistema solar. (PB)
Você existe oito planetas no sistema solar.
Há oito planetas no sistema solar.

A inaceitabilidade de construções como (81)a. no PE deve-se ao fato de que elas não podem ser interpretadas como possessivas por restrições pragmáticas, que proíbem a inserção de um possuidor indeterminado indicado por *se*.

Já no PB a realização de *você* nestes contextos, como em (81)b., é aceitável porque os falantes não interpretam tais sentenças como possessivas, mas como existenciais. A explicação

⁷⁰ Exemplo extraído de Avelar (2009).

está em que, como esse pronome funciona como um sujeito gramatical que não interage tematicamente com o verbo, o pronome não recebe qualquer papel temático relacionado à interpretação possessiva atribuído pelo verbo, o que inviabiliza a expressão de posse.

Neste sentido, a não relação possessiva entre o pronome *você_{gen}* e o complemento de *ter* é assumida nesta tese. Esse contraste entre as duas variedades do Português é suficiente para evidenciar que construções como (82)b. no PB são existenciais e não possessivas.

3.4.3 Sobre a derivação de construções com nulo genérico e nulo expletivo

A despeito do paralelismo sintático e interpretativo das construções existenciais com *ter* e de predicados indeterminados quanto à possibilidade de inserção de um pronome genérico na posição de sujeito em competição com o nulo não referencial e o genérico, deve-se ressaltar que, enquanto em predicados indeterminados com nulo genérico, de acordo com Holmberg (2010), há incorporação dos traços de ϕ P na categoria funcional T, e portanto, SpecTP não é projetado, o que licencia a inserção de um pronome nesta posição, nas construções existenciais, o nulo expletivo ocupa SpecTP e checa o EPP.

Considere-se a equivalência de sentido entre as construções existencial em (83)a. com a indeterminada⁷¹ em (84)a. com pronome nulo:

(83) a. *Ø_{expl}* Não *tem* mais burca na Síria.

b. *Cê* não *tem* mais burca na Síria.

(84) a. Não *Ø_{gen}* *usa* mais burca na Síria.

b. *Cê* não *usa* mais burca na Síria.

Uma explicação provável para a possibilidade de inserção de um pronome genérico em SpecTP, tanto em construções em que esta posição não é projetada (84)b, como nas em que há projeção de SpecTP ocupado por *pro_{exp}*, como em (83)b, sem alçamento de constituinte, respalda-se no parâmetro da ϕ -(in)dependência de EPP de T (Holmberg, 2010) e em representações distintas que entram na Numeração destas construções (*cf.* Lunguinho, 2006).

⁷¹ Neste estudo consideram-se como indeterminadas também as construções passivas, com base na proposta Nunes (1990) de reanálise do *se*.

Com base no parâmetro da ϕ -(in)dependência, Avelar & Galves (2011) propõem que PB e PE são línguas diferentes quanto a esse parâmetro. Enquanto o PE tem uma condição EPP de T ϕ -dependente, o que impõe que somente DPs argumentais portadores dos traços-*phi* sondados por T possam ser movidos para a posição de SpecTP, o PB, que exige a presença de um DP em SpecTP, permite o alçamento de DPs não argumentais para SpecTP, tendo, portanto uma condição EPP de T ϕ -independente. Dada essa propriedade ϕ -independente do PB, Avelar & Galves (2011) e Marins (2013) advogam que, no caso de sentenças existenciais com *ter* sem qualquer constituinte na posição de sujeito, SpecTP é projetado e é ocupado por um expletivo nulo. O posicionamento de que a posição de SpecTP sempre será projetada é assumida nesta tese.

A presença de um expletivo nulo nas construções de *ter_{ex}* na posição SpecTP também é defendida por Duarte e Kato (2008). A partir da proposta de que em línguas de proeminência de tópico o EPP deve ser satisfeito em SpecTP, as autoras defendem que para as sentenças existenciais com *ter*, se forem impessoais, SpecTP apresenta um expletivo nulo, como se pode ver em (85):

(85) [_{TP} tem_v [_{VP} t_v [prédios lindos em Londres]]

Partindo, assim, da consideração de que construções existenciais com *ter* projetam SpecTP, e que o expletivo nulo que ocupa essa posição satisfaz EPP, sugiro que a presença de um elemento não argumental em SpecTP dessas construções, como *você_{gen}*, não se deve ao atendimento a esse princípio. Portanto, a presença desse pronome em sentenças de *ter_{ex}* em substituição ao nulo expletivo, pode estar relacionado ao fato de que no PB há uma tendência a inserir elementos tanto argumentais quanto não na posição de sujeito gramatical (parâmetro da ϕ -(in)dependência), evitando, assim, que o verbo ocorra em posição inicial absoluta.

Sob esse ponto de vista, é possível afirmar que o PB passou a contar com mais uma estratégia alternativa a estruturas existenciais impessoais com *ter* em substituição ao expletivo nulo na posição de sujeito gramatical, além do alçamento do locativo, em atendimento à demanda do sistema, que muda em direção ao preenchimento do sujeito.

Com relação à derivação dessas sentenças existenciais com *pro_{exp}* ou um pronome na posição SpecTP, considero que a Numeração através da qual elas são derivadas tem entradas distintas. Esta proposição tem respaldado em Lunguinho (2006) que propõe Numerações distintas para derivação de construções de tópico-sujeito, como a que se vê em (86)b:

- (86) a. Furou o pneu do carro.
b. O carro furou o pneu.

O autor tenta estabelecer uma representação formal que dê conta das sentenças como (86)b. Para tanto, Lunguinho propõe que, como o fenômeno em análise diz respeito a verbos com só um argumento, há necessidade de se postular um expletivo nulo na posição de sujeito quando não há partição de constituintes e o argumento interno está depois do verbo. Assim, para uma sentença como em (86)a há um *pro* na Numeração, a partir da qual seria derivada a seguinte representação:

- (87) [TP *pro_{exp}* [T furou [VP ~~furou~~ [DP [o pneu do carro]]]]]
(LUNGUINHO, 2006, p.141)

Por sua vez, a Numeração relativa à sentença (86)b. com partição de constituinte não apresenta nem o expletivo nulo nem a preposição *de*, responsável pela semântica possessiva. A derivação para esta sentença é a representada em (88):

- (88) [TP [DP o carro] [T furou [VP ~~furou~~ [DP [DP ~~o carro~~] [D' o [NP pneu]]]]]]]

Para explicar que o fato de o DP o carro não ocupar SpecTP em (86)a, Lunguinho alega que para validar os seus traços-*phi* não interpretáveis, T sonda o DP *o pneu do carro* que, por não ter seu traço de Caso ainda valorado, pode ser tomado como alvo da operação de Agree. Uma vez que a Numeração tem um expletivo nulo, este ocupará a posição de SpecTP, e a subida desse DP é impedida, fazendo com que ele permaneça na posição onde foi gerado. Como discutido acima, o expletivo nulo na posição de SpecTP satisfaz o EPP de T.

Por não interessar especificamente a este trabalho o alçamento de um elemento da sentença para SpecTP, no caso de construções de tópico-sujeito, esse tipo de derivação não será detalhado.

Na direção desta proposta de Lunguinho, considero que a Numeração a partir da qual são geradas sentenças existenciais impessoais apresenta um *pro_{exp}*. Já as sentenças existenciais com um pronome lexicalizado na posição de sujeito gramatical seriam derivadas através de uma Numeração sem a presença de *pro_{exp}*, o que licencia a inserção do pronome. Assim, *pro_{exp}* é um

item que compõe a Numeração relativa à sentença em (83)a. “Não tem mais burca na Síria”, derivando a representação (89) a seguir:

(89) [SpecTP *pro_{exp}* T [v/VP ter v' [Adv não mais [XP [NP burca] [X' X [[LocP na Síria]]]]]]]

Dessa forma, dado que o PB é uma língua ϕ -independente, a posição de SpecTP é projetada em sentenças existenciais com *ter*. Na ausência de um elemento para preencher SpecTP, o expletivo precisa ser manifesto, como em (88).

Na derivação em (90), a inserção de um pronome fraco aparece como uma alternativa em substituição ao nulo expletivo. Conforme proposto no item anterior, 3.4.2, por não estabelecer relação temática com o verbo, o pronome *você_{gen}* não surge em posição argumental, mas é inicialmente inserido na posição INFL para depois ser alçado para SpecTP, derivando a representação em (90):

(90) [SpecTP *cê_i* T/ Infl *t_i* [v/VP ter v' [Adv não mais [XP [NP burca] [X' X [[LocP na Síria]]]]]]]]]

Se por um lado em línguas ϕ -independentes como o PB, postula-se um *pro_{exp}* em SpecTP em construções existenciais, o qual satisfaz EPP, por outro, em construções pessoais indeterminadas com pronome nulo genérico, como em (84)a “Não *Ø_{gen}* usa mais burca na Síria”, SpecTP não é projetado e, portanto, EPP não é checado.

Portanto, proponho que essas construções de nulo genérico são derivadas a partir de Numerações que não contêm *pro*, o que faculta a inserção de um pronome genérico lexicalizado, em atendimento à satisfação do requerimento EPP.

Em vista desses fatos, assumo que a ocorrência de um pronome genérico pleno em construções existenciais com *ter* não se deve ao atendimento do EPP, ao contrário do que ocorre em construções de indeterminação em que o pronome é inserido para satisfação deste princípio.

Deve-se ressaltar, considerando a proposta de Holmberg (2010) de categorização de dois tipos de LSNs, as consistentes, que não apresentam pronomes sujeitos nulos indefinidos, e as parciais, que apresentam sujeitos nulos indefinidos, que a implementação de um sujeito nulo de referência genérica num sistema que tende ao preenchimento dos sujeitos referenciais definidos com formas pronominais plenas, em vez de se mostrar na direção contrária em relação à remarcação do parâmetro *pro-drop*, revela a característica parcialmente *pro-drop* do PB.

3.4.4 Referência arbitrária x referência genérica (+) e (-) inclusiva

Para fechar a discussão dos fenômenos imbricados no processo em variação analisado, deve-se abordar a questão das possíveis referências atribuídas pelas variantes em estudo. O estabelecimento da distinção entre referência genérica e arbitrária do sujeito nulo é relevante para o presente estudo em vista da hipótese de que o pronome de 2ª pessoa *você_{gen}* ocupa no PB a posição dos nulos não referenciais e de referência genérica, mas não dos de referência arbitrária, os quais se pressupõe que tendem a permanecer no PB. Além disso, a hipótese de substituição no PB de pronomes para referência genérica, *a gente* por *você*, parte do pressuposto de que *a gente* esteja perdendo a propriedade (+) genérica e caminhando para uma referência (+) específica.

Holmberg (2015)⁷² faz a distinção da marcação genérica entre inclusivos, quase-inclusivos e exclusivos. É característica dos pronomes genéricos de marcação inclusiva que eles denotam pessoas em geral, incluindo o falante e o ouvinte. Como incluem todo mundo, sua referência é irrestrita, isso significa que os genéricos inclusivos não têm traços-*phi*.

Em LSNPs, o pronome genérico tem de ter alguns traços valorados de pessoa e número porque os traços de concordância têm de ter um valor atribuído. Os traços favoritos são os de 3ª pessoa, porque eles são os traços mínimos valorados. Isso significa que os traços são devido à exigência morfológica. Além disso, a referência genérica inclusiva pode ser obtida também pela semântica do predicado e outros fatores externos ao pronome, como se verá na análise dos dados.

É a leitura do pronome impessoal inglês *one*, e também do típico *you* genérico, pronome de 2ª pessoa usado como genérico em muitas línguas.

(91) *You speak different languages in India.*⁷³

Já o pronome genérico exclusivo tem como característica denotar pessoas em geral em domínios que não incluem o falante nem o destinatário:

(92) *They speak different languages in India.*⁷⁴

⁷² O paper *Generic Pronouns and Phi-features: evidence from Thai* (Anders Holmberg) foi apresentado no Encontro Intermediário do GT em Teoria da Gramática da ANPOL 2015, na UFMG.

⁷³ Exemplo extraído de Holmberg (2015).

⁷⁴ Exemplo meu criado a partir do ex. (90).

Duarte *et al* (2015)⁷⁵ estabelecem essa distinção de marcação de referência em termos de pronomes genéricos e arbitrários. Nessa concepção, sujeitos arbitrários referem a uma entidade específica e particular (ou um grupo homogêneo finito), a qual não pode ser estabelecida pelo contexto, como em (93):

(93) Uma vez me *definiram* Austrália como a Inglaterra de bermuda.⁷⁶

Já a referência de pronomes genéricos aplica-se a qualquer entidade, ou qualquer elemento de um grupo, o qual não é finito, restrito ou homogêneo, como em (94) e (95), mas não em (96):

(94) Mas agora não *se* tem mais inverno. (referência genérica)

(95) *Você* come bem em São Paulo. (referência genérica)

(96) *Eles* comem bem em São Paulo. (referência arbitrária)

Duarte *et al* (2015) propõem a seguinte escala de gradação de arbitrariedade:

(97) (arbitrário/ menos genérico) eles < eu < nós/ a gente < se/ zero/ você (mais genérico)

Correlacionando as duas propostas, de Holmberg (2105) e de Duarte *et al* (2015), pode-se pressupor que quanto mais inclusiva for a marcação do pronome, mais genérica será sua referência e, em contrapartida, quanto mais exclusiva a marcação, mais arbitrária a referência.

O ponto intermediário na escala diz respeito aos pronomes que referem a qualquer elemento de um grupo mais específico, estabelecido pelo contexto, não incluindo o destinatário, pronomes de (+1ª pessoa plural). Para o tratamento desses pronomes nos *corpora* analisados, que incluem *a gente* e *nós*, adoto para efeito deste estudo as denominações ‘menos inclusivo’ em oposição a ‘mais inclusivo’, conforme sua leitura inclua ou não o ouvinte. Nos termos de Holmberg, os pronomes que referem a pessoas em geral incluindo o falante mas não o ouvinte denominam-se quase-inclusivos.

⁷⁵ DUARTE, Maria Eugênia Lammoglia; MARINS, Juliana Esposito; SILVA, Humberto Soares da. *Revisitando Duarte (1995): por uma análise gradiente dos sujeitos indeterminados em PB*. Comunicação apresentada no Encontro Intermediário do GT em Teoria da Gramática da ANPOL 2015, na UFMG.

⁷⁶ Exemplo extraído de Duarte *et al* (2015).

Tomando-se por base a escala proposta por Duarte *et al* (2015), proponho que os pronomes elencados como variantes neste estudo, *nulo* \emptyset , *se* e *você*, situam-se na referência genérica (+) inclusiva, e os pronomes *a gente* e *nós* podem ter marcação (+) ou (–) inclusiva. Com os traços de pronome (+) inclusivo, *você_{gen}* é o pronome prototípico de referência genérica.

Além dessa categorização aqui proposta para os pronomes indeterminados em (+) ou (–) inclusivos, é importante ressaltar a diferença de marcação genérica e arbitrária do pronome nulo: é considerado arbitrário o pronome nulo passivo, e genérico, o nulo nominativo. Essa distinção pode ser exemplificada pelas sentenças (98) e (99) a seguir:

(98) \emptyset_{arb} aceita cheque na padaria.

(99) \emptyset_{gen} Tem que ser bem claro e simples na televisão e curto também; já no jornal *online* \emptyset_{gen} tem que ser rápido. (AAX – M1 -14)⁷⁷

Aplicando-se o diagnóstico de substituição do pronome *nulo* (\emptyset) por *você_{gen}*, tipicamente genérico, para comprovar o caráter arbitrário ou genérico dos pronomes nulos, é possível constatar que, enquanto no exemplo (98), \emptyset não é intercambiável com *você_{gen}*, no (99) o é, como mostram os exemplos (98)a. e (99)a. a seguir:

(98) a. **Você_{gen}* aceita cheque na padaria.⁷⁸

(99) a. *Você_{gen}* tem que ser bem claro e simples na televisão e curto também; já no jornal *online* *você_{gen}* tem que ser rápido.

Por essa razão, uma das hipóteses desta tese é de que o pronome nulo que tende a permanecer no PB é para atribuição de referência arbitrária que *você_{gen}* ou não consegue, ou não é licenciado⁷⁹, conforme discutido em 3.1 acerca da reanálise do clítico *se*. O que se tem observado no PB em construções como (98) de nulo arbitrário é o alçamento do locativo para a posição de especificador de TP para satisfazer o EPP de T, que se dá via movimento, e não via inserção de pronome, conforme exemplo (98)b.:

(98) b. **A padaria** aceita cheque.

⁷⁷ Os exemplos citados neste trabalho referentes ao PB obedecem a seguinte convenção: as letras iniciais são siglas da instituição onde o informante trabalha ou de sua cidade de origem (AAX = Araxá); na sequência as letras F e M indicam o gênero e os números de 1 a 3, a idade; a seguir tem-se o número da entrevista.

⁷⁸ No exemplo (97)a. o pronome *você* não atribui referência genérica.

⁷⁹ Deve-se considerar a proposta de *pro* na posição de sujeito dessas construções, conforme discutido em 3.1

A emergência desse tipo de estrutura em que o constituinte locativo ou parte dele (mais especificamente o PP_{loc}) aparece realizado na posição de sujeito de sentenças existenciais com *ter* é investigada por Marins (2013). A autora mostra que o PB teria passado a permitir, ao lado das existenciais e possessivas com *ter*, sentenças com o mesmo verbo em que há um DP na posição de sujeito que, entretanto, não pode ser interpretado imediatamente como possuidor, como se vê em “A geladeira *tem* leite na porta”. De acordo com Marins (2013), uma possibilidade é interpretar essas sentenças como existenciais, propondo para ela o mesmo tratamento dado às existenciais impessoais, em que não se verifica qualquer elemento na posição de sujeito. Para isso, o DP na posição de sujeito seria movido de outra posição da sentença – do interior do constituinte locativo –, tendo seu papel temático atribuído no domínio onde é gerado. Assim, uma vez realizado como sujeito, o DP manteria sua semântica original, sem entrar em interação semântica com o restante da sentença. A autora denomina essas sentenças de *existenciais de tópico-sujeito*.

O alçamento do locativo para a posição de sujeito também foi abordado por Viotti (1999) ao relacionar a preferência pelo verbo *ter* em vez de *haver* ao fato de *ter* autorizar uma construção pessoal com o locativo na posição de sujeito, uma operação não disponível para *haver*.

Outra possível constatação a partir de diagnóstico de substituição do pronome sujeito nulo por *você_{gen}* é que construções como (98) de nulo arbitrário podem aceitar a inserção de *você_{gen}* mediante o acréscimo da expressão “*não + v + mais*”, como em (100). Esta expressão atribui um valor habitual, usual, ao predicado, tornando-o genérico, no sentido de se aplicar a qualquer elemento de um grupo, possibilitando assim a inserção do pronome genérico *você* na posição de sujeito, conforme se pode constatar nos exemplos a seguir:

(100) *Você_{gen}* **não** aceita **mais** cheque (nos dias de hoje).

(101) *Você_{gen}* **não** usa **mais** saia (no Brasil).

(102) *Você_{gen}* **não** amola **mais** faca (nos dias de hoje).

Exemplos da amostra desta pesquisa a serem apresentados e discutidos na análise dos dados serão esclarecedores quanto à categorização aqui proposta acerca da marcação (+) ou (-) inclusiva dos pronomes genéricos/indeterminados considerados neste estudo.

3.5 Considerações parciais

Nesta seção busquei, por meio da apresentação de estudos que evidenciam alguns sintomas da mudança linguística em processo no PB, delinear o percurso da indeterminação no PB, de *se* a *cê*, tendo em vista a consideração do *status* do pronome *você* como elemento indeterminador, em função semelhante à do clítico *se*. Para embasar a proposta de que os pronomes *se* e *você* apresentam características semelhantes, foram discutidas ainda questões relativas ao caráter não expletivo do pronome *você*, além de considerações acerca da referência genérica e arbitrária do sujeito nulo, e (+) e (-) inclusiva dos pronomes indeterminados. Toda a trajetória aqui delineada foi embasada pelas seguintes considerações e hipóteses:

a) A baixa produtividade de clíticos no PB (Nunes, 1990), uma das formas de indeterminação na língua portuguesa padrão, é responsável pelo surgimento da estratégia de nulo genérico, uma vez que ao não mais serem empregados, os clíticos deixaram vazia esta posição.

b) Após perder seu sujeito nulo referencial, e um paradigma de pronomes fracos nominativos aparecer no lugar do sistema pronominal *Agr* (cf. Kato, 2000), o PB está perdendo o nulo não referencial e o genérico. O surgimento de formas pronominais fracas pode ser visto no PB, em contextos nos quais a língua licenciaria sujeito nulo (expletivo e genérico), como uma estratégia para compensar a perda da propriedade de licenciar sujeito nulo e para valoração do traço EPP.

c) A referência genérica inclusiva já foi realizada no PB pelo pronome *a gente*, que ao passar a ser empregado com referência (+) determinada de 1ª pessoa do plural, portanto (-) inclusiva, perde a propriedade de genericidade. É preciso que o sistema lance mão de outro pronome para essa referência. O pronome de 2ª. pessoa *você* passa a ser empregado no PB com marcação genérica.

d) Por ser licenciado na posição de sujeito do verbo *ter* existencial, presume-se, *a priori*, que *você* genérico tenha propriedades de pronome expletivo. Contudo, o próprio caráter indefinido do pronome, que revela um valor semântico, invalida esta interpretação.

e) Considerando a propriedade ϕ -independente do PB, em sentenças existenciais com *ter* sem qualquer constituinte na posição de sujeito, SpecTP é projetado e ocupado por um expletivo nulo, portanto, a realização de um pronome pleno genérico nessas construções se deve à tendência do PB de preenchimento da posição de sujeito com formas foneticamente realizadas. Já em construções predicativas em que não há projeção de SpecTP, a realização do pronome genérico é para satisfação do EPP.

f) A inserção de *você_{gen}* na posição de SpecTP, tanto de construções existenciais com *ter* como de pessoais, torna essas construções equivalentes a estruturas de indeterminação pela correlação que guardam com estruturas de indeterminação formalmente semelhantes.

g) Na base da proposta anterior, proponho que *você_{gen}* atua como elemento indeterminador com características semelhantes à partícula indeterminadora *se*. Justifica-se, assim, a equivalência dos elementos de indeterminação no PB: *se* e *cê*.

h) A proposta do caráter indeterminador do pronome *você_{gen}* é referendada pelos pressupostos teóricos de Holmberg (2010) sobre os traços de indefinidade do nulo de 3ª pessoa no PB. A referência irrestrita de *você_{gen}* pressupõe ausência de traço definido em T, o que o torna equivalente ao *se* indeterminado.

4. METODOLOGIA

Os procedimentos metodológicos desta pesquisa foram norteados pelas perguntas, hipóteses e objetivos elencados, pela definição da amostra e das variantes, bem como pelo estabelecimento dos possíveis grupos de fatores condicionadores das variantes utilizados na codificação e análise dos dados. De acordo com Tarallo (1997), o envelope de variação é a descrição das variantes por meio da qual se conhecem os contextos que favorecem o emprego de uma variante em detrimento de outra.

A análise proposta toma por base pesquisas empíricas disponíveis para consulta, consubstanciada pelas discussões, hipóteses e teorias propostas sobre os fenômenos que desencadearam a emergente forma de indeterminação no PB. Tendo em vista a comprovação da hipótese de que os fenômenos analisados neste estudo são sintomas do processo de remarcação do PSN no PB, a análise interlinguística foi realizada entre o PB e o PE, língua considerada prototipicamente *pro-drop*.

4.1 Material selecionado

Para subsidiar a análise da frequência das variantes em estudo, os pronomes indeterminados *se*, *nulo*, *você*, *a gente* e *nós* em variação com *você_{gen}*, e sua evolução em tempo real de curta duração, os *corpora* do PB e do PE são compostos por amostras coletadas em dois momentos distintos, décadas de 1970 e atual (2013 a 2015), com um interstício de aproximadamente 50 anos entre as coletas.

Justifica a seleção de *corpora* de dois períodos distintos o fato de que a articulação entre presente e passado permite evidenciar estágios variáveis e mudanças que aconteceram (tempo real) e que estão em curso (tempo aparente), ou seja, permite identificar a transição entre dois estágios da mudança (Labov, 2008). Permite ainda fazer o estudo a que Labov se refere como estudo de tendência⁸⁰, com a possibilidade de investigar possíveis mudanças em curso, no seio de uma comunidade, em tempo real de curta duração.

Considerando com Labov (2008, p. 291) que “não se pode fazer nenhum avanço importante rumo ao entendimento do mecanismo da mudança linguística sem o estudo sério dos fatores sociais que motivam a evolução linguística”, para testar a hipótese de que as

⁸⁰ Há duas abordagens para obtenção de dados em tempo real, de acordo com Labov: recontato (estudo do tipo painel) e e constituição de nova amostra (estudo de tendência), o qual compara amostras aleatórias da mesma comunidade de fala, estratificadas com base nos mesmos parâmetros sociais, em dois momentos do tempo.

estratégias de indeterminação em análise constituem processo de mudança em andamento, utilizaram-se na quantificação dos dados os grupos de fatores gênero e faixa etária, que constitui condição *sine qua non* para que se diferencie uma regra variável que se caracteriza em mudança em andamento de uma que se caracteriza como variável estável.

A distribuição dos fatores sociais teve como referência os critérios do Projeto NURC. Assim, para manutenção de regularidade dos fatores sociais, todos os informantes selecionados para comporem as amostras, tanto do PB como do PE, estão estratificados por gênero e idade, com igual distribuição de faixas etárias: de 25 a 35 anos (faixa 1), de 36 a 55 anos (faixa 2) e de 56 anos em diante (faixa 3).

Além disso, como o Projeto NURC procede à descrição dos padrões de uso na comunicação oral adotados pelo estrato social composto por indivíduos de escolaridade superior no Brasil, este critério foi adotado para seleção dos informantes desta pesquisa. Assim, os informantes dos *corpora* do PB e do PE preenchem os requisitos de possuírem curso universitário completo e serem filhos de falantes nativos de língua portuguesa, preferencialmente nascidos na cidade em análise.

O fator social ‘escolaridade superior’ faculta a abordagem nas entrevistas de temas da atualidade e relacionados à formação e/ou atividade profissional dos informantes, motivando-os a emitirem opinião, discutirem, enunciarem, o que pressupõe a predominância do tipo de discursivo argumentativo em suas falas. Tendo em vista que a indeterminação do sujeito, objeto deste estudo, é empregada mais frequentemente em argumentações, a presença desse tipo argumentativo nas entrevistas utilizadas favoreceu o emprego das variantes em análise.

Portanto, os fatores escolaridade, por estar definido *a priori*, e tipo de discurso, pressuposto, não foram selecionados como condicionadores. Assim, possuir formação universitária constituiu pré-requisito para a seleção dos informantes de toda a amostra.

Considerando-se os acervos disponíveis para consulta da década de 70 do Projeto NURC dos estados do Rio de Janeiro e de São Paulo, e um banco de dados do estado de Minas Gerais da década atual, a amostra do PB é composta por falantes de três estados da região Sudeste: Minas, Rio e São Paulo.

Para o *corpus* do tempo presente foram utilizadas entrevistas semiestruturadas da fala espontânea, gravadas com informantes do estado de Minas Gerais, que fazem parte de um banco de dados ainda não institucionalizado, coordenado pela profa. Dra. Maura Alves de Freitas

Rocha, da UFU, bem como de entrevistas orais realizadas por um pesquisador⁸¹ do Centro Universitário - Uniaraxá. A amostra do PB totaliza 4.895 dados, sendo 2.137 da década de 70, e 2.758 da década atual.

Deve-se ressaltar que, como em relação ao PB esta pesquisa se limita à investigação do comportamento linguístico em cidades de três estados da região sudeste, tendo em vista as dimensões do país, qualquer generalização sobre este fenômeno na variedade brasileira do Português só pode ser tomada em termos relativos.

Para a composição da amostra do PE foram observados os mesmos critérios de seleção de entrevistas da amostra do PB, entretanto, as décadas analisadas foram 1970 e 2010, em função da disponibilização de entrevistas orais com informantes do PE para consulta em *sites*.

Os dados da década de 70 do PE foram extraídos do CRPC - *Corpus* de Referência do Português Contemporâneo do Centro de Linguística da Universidade de Lisboa, e os da década atual, do Projeto de Cooperação Internacional Brasil-Portugal vinculado ao Programa de pós-graduação em Letras Vernáculas da Faculdade de Letras da Universidade Federal do Rio de Janeiro (UFRJ). Os informantes do PE são das regiões de Lisboa, Porto e Cacém. A amostra totaliza 931 dados, sendo 303 da década de 70, e 628 da década atual.

Diferentemente da maioria das variações sintáticas, por se tratar de muitas variantes (cinco), a frequência desta variação é relativamente alta, embora não se possa desconsiderar que qualquer tentativa de especificar estruturas sintáticas, implica inevitavelmente formas que podemos não ouvir em uma investigação limitada (cf. Labov, 2008).

O programa de análise estatística adotado para o processamento dos dados foi o *GoldVarb X*, sendo que os dados do PB e do PE foram submetidos a rodadas estatísticas separadamente.

O levantamento dos fatores linguísticos potencialmente favorecedores das variantes seguiu a própria intuição do pesquisador, tomando por base a observação do vernáculo e outras pesquisas já realizadas sobre os temas: preenchimento do sujeito, estratégias de indeterminação e verbos existenciais.

Os dados foram primeiramente submetidos a uma análise quantitativa e em seguida a uma análise qualitativa dos resultados obtidos, a exemplo de Tarallo (1989). A análise quantitativa baseou-se em metodologia variacionista, segundo Labov (2008), seguida da descrição formal dos resultados obtidos a partir da sintaxe gerativa à luz da Teoria de Princípios

⁸¹ Parte das entrevistas utilizadas foram realizadas pelo professor Dr. Valter Gomes para elaboração da tese intitulada: A construção de estratégias: estudo de caso em instituições de ensino superior. Universidade de Trás-os-Montes e Alto Douro. Portugal, 2017.

e Parâmetros (Chomsky, 1981), tendo em vista a proposta de Kato e Tarallo (1989) da Sociolinguística Paramétrica.

A análise qualitativa teve como ponto de partida o pressuposto de que o PB está caminhando para a perda das propriedades de sujeito nulo, com base na Teoria de Princípios e Parâmetros, de Chomsky, em sua versão Minimalista, bem como na análise de outros processos em variação em outras partes do sistema que se encaixam no fenômeno em questão, e que estariam repercutindo no sistema.

4.2 Fatores

Consideram-se como relevantes para esta análise os fatores:

Variantes dependentes:

- 0) nulo de 3ª pessoa não referencial e de referência genérica e arbitrária (\emptyset_{exp} , \emptyset_{gen} , \emptyset_{arb})
- 1) *você (genérico)*
- 2) *a gente (indeterminado/ genérico)*
- 3) *nós (indeterminado)*
- 4) clítico *se*

Devem-se tecer alguns esclarecimentos sobre a definição das variantes dependentes, visto que, de acordo com Labov, enunciados que se referem ao mesmo estado de coisas com o mesmo valor de verdade constituem-se como variantes de uma mesma variável. De forma semelhante, para Tarallo, “variantes linguísticas são diversas maneiras de se dizer alguma coisa em um mesmo contexto e com o mesmo valor de verdade” (TARALLO, 1997, p.8).

Portanto, para certificar que as variantes propostas são intercambiáveis, com o mesmo valor de verdade, só foram selecionadas as ocorrências em que o sujeito (formas pronominais em análise) não pudesse ser claramente delimitado, identificado, ou não houvessem opções para o referente. Outro aspecto a ser observado, é que *a priori* só deveriam ser considerados os empregos em que as cinco variantes selecionadas fossem equivalentes por meio de diagnósticos de substituição, como nos primeiros exemplos apresentados na introdução, de (1) a (5). Entretanto, razões a serem discutidas na apresentação dos grupos de fatores levaram a uma relativização deste critério.

Por fim, ressalto que esta análise propõe que as variantes em análise constituem frutos legítimos da mudança que se processa em nosso sistema, que perpassa por fenômenos como reanálise e apagamento de clítico *se*, surgimento do nulo genérico, substituição da forma *a gente* por *você* para marcação genérica e preenchimento da posição vazia do sujeito indefinido e não referencial. Para comprovação da hipótese da substituição de pronomes para referência genérica, assim como para subsidiar a análise da evolução e difusão do emprego dos pronomes genéricos como formas de indeterminação em tempo real de curta duração, é que os *corpora* são compostos por amostras de duas décadas.

Grupo 1: Variedades do Português:

- a) Português do Brasil (PB)
- b) Português Europeu (PE)

Este fator foi considerado para comprovar a hipótese de que o PB, ao, cada vez mais, preferência às formas inovadoras de indeterminar o sujeito, afasta-se do PE, que não privilegia tais formas, e tem a passiva com concordância como a forma preferida de indeterminação (Duarte, Kato e Barbosa, 2003) e, tendo em vista, ainda, as considerações de Tarallo (1993a) de que as mudanças no PB são evidências de que seria muito improvável e nada natural que o PE e o PB viessem a se encontrar de novo sintaticamente. Os exemplos a seguir, extraídos respectivamente dos *corpora* do PE e do PB, evidenciam este contraste entre as duas variedades do Português em relação ao emprego de estratégias de indeterminação do sujeito.

(103) Às vezes ajuda sei lá por exemplo: a a maneira como *se pronunciam* algumas palavras eu identifico se é do Norte, se é do Alentejo....⁸² (PE)

(104) Quanto mais *você_{gen}* distancia da natureza ... mais *você_{gen}* *você* perde a percepção a noção de que as coisas ... se dão em ciclos ... então ...⁸³ (PB)

Grupo 2: Tipo de Construção

- c) existenciais
- d) demais construções

⁸² Exemplo extraído do Projeto de Cooperação Internacional Brasil-Portugal (INF CAC A 3 – 2010)

⁸³ Exemplo extraído do *corpus* atual do PB – Informante (AAX – M1 -14).

Este fator foi selecionado para comprovar a hipótese de que as formas pronominais indefinidas (*a gente, você e cê*) estão ocupando a posição de sujeito gramatical, em contextos tanto de nulo genérico (105) em construções predicativas, como de nulo não referencial (106) em construções com *ter* existencial, nas quais parece haver maior propensão à inserção do pronome genérico de 2ª. pessoa.

(105) O desemprego é real, é um risco, então ***você_{gen}*** tem que se preparar para isso.⁸⁴

(106) E nesse momento ***você_{gen}*** **tem** a presidente da República, do PT, sofrendo processo de *impeachment*.

Grupo 3: Tipo de sentença que favorece o emprego de sujeitos de formas plenas:

e) construções finitas – tempo finito

f) construções infinitivas – tempo infinitivo

O trabalho de Duarte, M.E; Kato, M. e Barbosa, P. (2003) aponta que as formas plenas para indeterminação do sujeito apareceram com maior frequência nas sentenças finitas, enquanto o sujeito nulo e o clítico *se* aparecem preferencialmente nas infinitivas, razão por que este foi selecionado como possível fator linguístico favorecedor das variantes em questão.

(107) Os cariocas vão querer me matar, mas lá em São Paulo (...) tem lugar [pra Ø tocar], lugar [pra Ø ensinar as crianças]...⁸⁵

(108) Mas [para *se* usar o preto] (...) as fábricas de pigmentos tiveram que produzir o preto em barda.

A sentença (109) a seguir é um exemplo do emprego de forma pronominal plena para indeterminação do sujeito em construção infinitiva no PB contemporâneo:

(109) Então, o PDI, Plano de Desenvolvimento Institucional é extremamente importante pra ***você_{gen}*** traçar o rumo da universidade. (UFV- M2 –26).⁸⁶

⁸⁴ Exemplos (105) e (106) foram extraídos do Programa *Em Pauta* da Globo News, em: 03/01/2016.

⁸⁵ Exemplos (107) e (108) foram extraídos de Duarte, M.E; Kato, M.A.; Barbosa, P. (2003).

⁸⁶ Exemplos de (109) a (128) foram extraídos do *corpus* da década atual à exceção de (110) e (111) que são da década de 70.

Grupo 4: Grau de indeterminação

g) indeterminado (+) inclusivo

h) indeterminado (-) inclusivo

A seleção deste fator objetivou dar conta dos empregos de *a gente* na amostra, muitos dos quais, embora não consistam pronomes genéricos de fato, mas indeterminadores, não poderiam ser descartados, tendo em vista a hipótese aventada neste estudo da substituição de pronomes no PB para marcação genérica. Além disso, os empregos do pronome *nós* que integram os dados igualmente se dividem em indeterminados (+) ou (-) inclusivos, ou (+) ou (-) específicos, e consistiram no único indeterminador em algumas entrevistas do *corpus* da década de 70 do PB. Pelos exemplos a seguir pode-se evidenciar essa diferença.

Genérico (+) inclusivo:

(110) Mas... não, não **temos** metrô ainda, metrô tem que ser uma malha, certo? (SP-M1-343)

(111) É. O que **a gente** tem notado aqui no Rio de Janeiro, eu acho que o comércio progrediu muito aqui, ultimamente, quer dizer, em matéria de, de atrair consumidor, de, de, de sofisticar a oferta, de, de, isso realmente, principalmente na zona sul, não é? (RJ- M2-233).

Genérico (-) inclusivo:

(112) Quando **criamos** o curso de Medicina Veterinária, eles ah, vão acabar com o meu curso de Zootecnia e tal. (UFV- M3-29)

(113) Engenharia de produção tem corpo, **a gente** já pleiteou por sistemas de informação. (UFV- M2 – 26)

Grupo 5: Relação com a forma verbal da sentença (presente x passado)

i – presente

j – outros

t – passado

v – futuro do subjuntivo

No PB, o tempo presente não é tão usado com valor temporal, prestando-se mais à descrição de estados ou hábitos. Assim, os predicados genéricos, habituais e hipotéticos,

geralmente associados a tempos verbais no presente, apresentam uma leitura genérica, ou quase-universal, diferentemente dos predicados episódicos. Portanto, o emprego dos pronomes genéricos *você* e *a gente* com verbo no tempo presente pode resultar em uma situação em que se descreve um fato comum às pessoas em geral, não sendo direcionado a um ser específico. A repetição da mesma situação é que confere à sentença uma leitura genérica. Dessa forma, em uma construção como: “**Quando** *você_{gen}* trata bem o cliente, ele volta”, pode-se evidenciar o caráter habitual da ação que, aliado ao pronome *você_{gen}* na posição do sujeito, confere leitura genérica ao pronome.

Outro tempo verbal, tradicionalmente empregado em orações condicionais de caráter hipotético, o futuro do subjuntivo, concorre para o favorecimento da variante inovadora *você_{gen}*. Com base nesses pressupostos, presumo que a forma pronominal *você_{gen}* seja mais empregada em construções nos tempos do presente, por terem interpretação genérica e não episódica, e no futuro do subjuntivo.

Diferentemente, os predicados episódicos (com tempo verbal no pretérito) favorecem uma leitura referencial do pronome. Assim, pela propriedade dos pronomes *a gente* e *nós* de, além da referência genérica, referirem-se a pessoas em geral incluindo o falante mas não o ouvinte, portanto em contextos (-) inclusivos, suponho que o tempo verbal que favorece o seu emprego seja o pretérito. Os exemplos a seguir ilustram esses empregos nos tempos verbais: pretérito (114), presente (115) outros (116), e futuro do subjuntivo (117):

(114) Ah... eu acho que foi um órgão importante porque tem muitas decisões que antes *a gente* não tinha como bater o martelo... (UFV – M1-30).

(115) ... e hoje *você_{gen}* consegue bater por causa dessa nova legislação do campus. (UFV – M1- 30).

(116) Então acho que isso é *você_{gen}* ter uma certa coerência com o processo de ética também, *você* ter um processo de ética. (UFV – M3-29)

(117) Se *você_{gen}* tiver meia dúzia de alunos, *você_{gen}* mantém aquele curso com meia dúzia de aluno. (UFV – M3-29)

Grupo 6: Caráter hipotético da construção

8 – sim

9 – não

A observação da recorrência de *você_{gen}* em construções hipotéticas levou-me à postulação desse fator como outro contexto favorecedor da variante inovadora. Sabe-se que

além do tempo/modo futuro do subjuntivo para marcar construções condicionais, o caráter condicional-hipotético “não realizável” pode ser marcado por outros recursos linguísticos, como conjunções, principalmente a **se**, exemplo (118). Considera-se também que conjunções como *quando*, *quanto mais*, *para*, exemplos (119), (120) e (121) conferem caráter hipotético à construção por apresentarem uma condição para que algo ocorra. Os exemplos a seguir ilustram o que considero como caráter hipotético da construção:

(118) **Se** *você_{gen}* tem um animal e uma pessoa ali precisando de cuidado, *cê* vai escolher o que? (AAX – F1- 1)

(119) Eu acho que o que difere é que **quando** *você_{gen}* é mãe *você_{gen}* tem uma responsabilidade bem maior... (AAX – F1- 1)

(120) Eles usam o dólar, então hoje com o dólar mais valorizado *você_{gen}* acaba perdendo um pouco, **quanto mais** *você_{gen}* usa, *você_{gen}* melhora de categoria, *você_{gen}* converte os pontos.⁸⁷

(121) Nada pode ser 8 ou 80 né, *cê* sempre tem que tá ali no meio, olhar os dois lados... **para** *você_{gen}* não ser tão exigente ou pouco exigente. (AAX – F1- 1)

Grupo 7: Presença de SADVs e SPs locativos e temporais

m) presença

n) ausência

De acordo com Tarallo, o conhecimento dos contextos favorecedores das variantes só pode ser obtido mediante observação do vernáculo. Conforme justificativa para seleção do grupo 6, a análise dos contextos de maior produtividade da variante inovadora *você_{gen}* mostrou que, em sua maioria, trata-se de situações hipotéticas, frequente nas argumentações e exemplificações, as quais são normalmente iniciadas por sintagmas adverbiais e preposicionais, assim como SPs locativos e temporais, razão por que esse fator foi considerado como possível motivador de seu emprego. As construções a seguir exemplificam esses contextos:

(122) **Hoje em dia às vezes** *você_{gen}*... *você_{gen}* não tem mais isso numa cidade grande *você* é filho de rico *você* também está no mesmo status... do teu pai. (SP-M1-343)

⁸⁷ Exemplo extraído do Programa *Em Pauta* da Globo News, em: 28/11/2015.

(123) Por isso que eu falo, **às vezes quando** *você_{gen}* entra no cargo de gestão *você_{gen}* acaba tendo que discutir o sistema. (UFV – M1-30).

(124) **Antigamente** *você_{gen}* fazia umas reuniões com todos os professores, esse ano já teve uma... daqui pra frente voltar a ter essas reuniões, porque aí todo mundo participa, às vezes surge algumas ideias interessantes. (CEFET – M2 – 17).

(125) Mas **aqui** *você_{gen}* não tinha essa nomeação, essa eleição. (UFV – M1-24).

Grupo 8: Paralelismo – Presença x ausência de indeterminador anterior

k) presença

l) ausência

O fator paralelismo, que é a tendência da repetição da mesma forma numa sequência discursiva, sendo livre a escolha da primeira referência, tem se mostrado relevante em pesquisas sobre a alternância das formas *nós* e *a gente*, como na de Omena (2003).

(126) Então aquele negócio, se *você_{gen}* quanto mais *você* distancia da natureza ... mais *você*, *você* perde a percepção a noção de que as coisas ... se dão em ciclos ... (SP-M1-343)

Grupo 9: Intercambialidade das variáveis

y– sim

z– não

Relacionado ao fator grau de indeterminação das variantes (grupo 4), o fator intercambialidade foi elencado para selecionar, dentre as variantes com possibilidade de marcação (+) / (-) inclusiva, os empregos em que elas são intercambiáveis com as demais e os em que não são. Assim, em seus empregos mais específicos, as variantes *nós* e *a gente* não devem ser consideradas como possuindo o mesmo valor de verdade de *você_{gen}*, conforme exemplo (127) abaixo. Já no exemplo (128), a variante *a gente* é intercambiável com *você_{gen}*.

(127) E ***a gente*** monitorava o tempo todo e elas não sabiam disso. (AAX – F2- 6)
*E *você_{gen}* monitorava o tempo todo e elas não sabiam disso.

(128) Então ***a gente*** tenta fazer com que dê certo a distância. (AAX – F2- 6)
Então *você* tenta fazer com que dê certo a distância.

Grupo 10: Período do *corpus*

- 6– década atual
- 7 – década de 70

Grupo 11: Gênero

- o) feminino
- p) masculino

Grupo 12: Faixa Etária

- q) de 25 a 35 anos
- r) de 36 a 55 anos
- s) acima de 56 anos

5. ANÁLISE E DISCUSSÃO DOS DADOS

Tendo em vista uma análise empírica da evolução do emprego das estratégias de indeterminação no PB e no PE, em especial o pronome *você_{gen}*, foco desta pesquisa, tanto em tempo aparente como em tempo real de curta duração, que compreende as décadas de 70 e a atual, ambas amostras são constituídas por *corpora* de dois períodos de tempo, num interstício de aproximadamente cinco décadas.

A expectativa é de que a comparação dos resultados do processamento quantitativo-estatístico dos dados pelo pacote de programas *Goldvarb X*, entre as décadas e variedades investigadas, seja reveladora quanto à hipótese de que as mudanças sintáticas relativas às inovadoras estratégias de indeterminação que estão desencadeando alteração de valor do PSN trata-se de um fenômeno que acomete somente o PB, bem como do percurso dessas estratégias no PB, tomando-se por base três hipóteses básicas norteadoras deste estudo: o apagamento gradual da partícula *se*, que atesta a supressão de clíticos no PB (Nunes, 1990), favorece o surgimento do nulo genérico na posição de sujeito; depois de implementada uma das propriedades de LSNs, o preenchimento dos sujeitos referenciais, o sistema começa a caminhar no sentido do preenchimento dos sujeitos não referenciais (Duarte, 2003), assim como dos genéricos; a emergência do emprego da forma pronominal de 2^a. pessoa para referência genérica atesta esta tendência.

As Tabelas 1 e 2 a seguir mostram os resultados gerais das ocorrências das variantes respectivamente em PB e PE por década.

Tabela 1: Ocorrência das estratégias de indeterminação no PB por década:

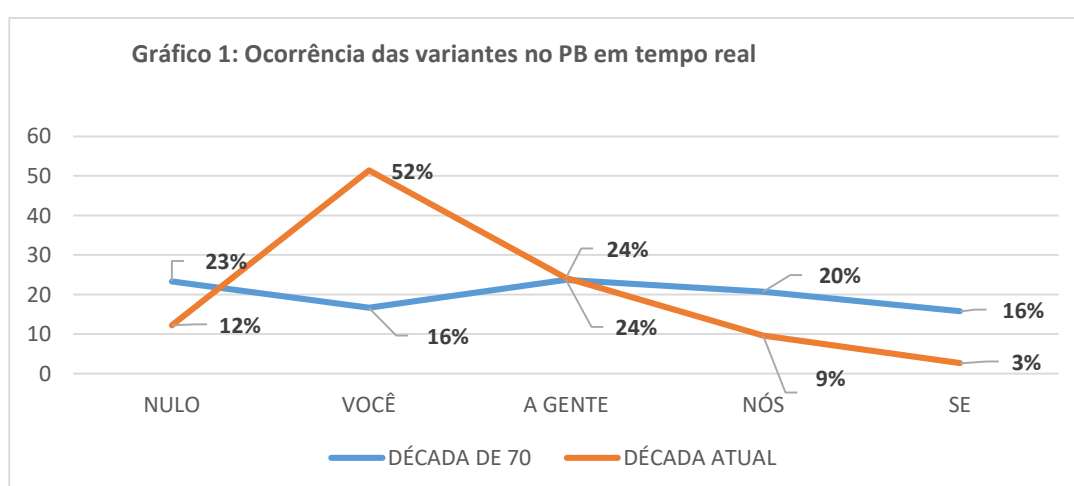
Variantes	Década de 70		Década atual	
	Realização	%	Realização	%
Nulo	501 /2137	23,0	340 / 2758	12,0
Você	351 /2137	16,0	1425 / 2758	52,0
A gente	511 /2137	24,0	662 / 2758	24,0
Nós	436 /2137	20,0	260 / 2758	9,0
Se	338 /2137	16,0	71 / 2758	3,0

Tabela 2: Ocorrência das estratégias de indeterminação no PE por década:

Variantes	Década de 70		Década atual	
	Realização	%	Realização	%
Nulo	50/303	16,5%	100/628	16%
Você	- ⁸⁸		-	
A gente	11/303	3,5%	19/628	3%
Nós	117/303	38,5%	314/628	50%
Se	124/303	41%	195/628	31%

Enquanto os resultados do PB da década de 70 indicam uma competição equilibrada entre todas as variantes, com menor incidência sobre o pronome *você_{gen}* e o clítico *se*, os resultados da década atual mostram um caminho oposto em relação à ocorrência dessas variantes: se por um lado *você_{gen}* apresenta expressivo aumento de emprego no período analisado, de 16 para 52%, por outro, o clítico *se* indeterminador mostra decréscimo de emprego, de 16 para 3%. Assim, por esses resultados, pode-se comprovar a supressão do *se* ao longo dessas cinco décadas, e depreender que tanto o apagamento do *se* quanto o aumento na frequência de inserção do pronome *você_{gen}*, são fenômenos encaixados que concorrem para o processo de mudança por que passa o PB.

Essa comparação é melhor visualizada no Gráfico 1 a seguir:



⁸⁸ Não ocorrência da variante.

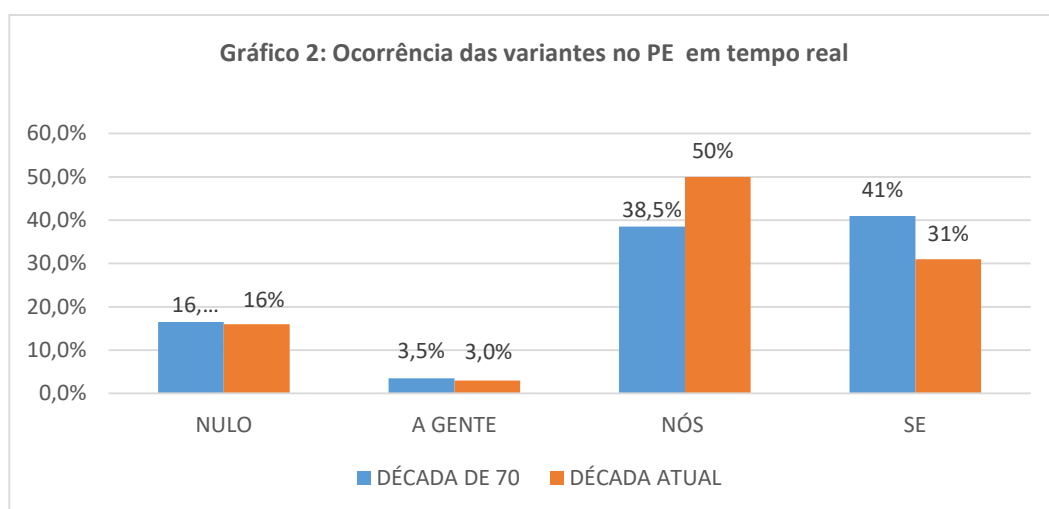
Os dados do PE (Tabela 2) confirmam que as duas variedades do Português seguem por caminhos divergentes em relação ao emprego das estratégias de indeterminação. Enquanto no PB o emprego de *você_{gen}* lidera o emprego das estratégias de indeterminação, os dados do PE confirmam a não ocorrência do pronome *você* com marcação genérica nas duas décadas analisadas.

O *corpus* da década de 70 mostra que o pronome *se* era a estratégia de indeterminação mais empregada no PE à época, ocorrendo em 41% do *corpus*, seguido de *nós*, com 38,5%, e o pronome *a gente* tem emprego inexpressivo, aparecendo em apenas 3,5% das estruturas analisadas. Dessa forma, esses dados validam o que se vem discutindo (Nunes, 1990), (Duarte, 2000), acerca das divergências em relação às estratégias de indeterminação entre PB e PE, onde a passiva com concordância é a estratégia preferida, embora elas coexistam com as construções sem concordância (Duarte, 2003).

Deve-se ressaltar que as ocorrências de pronome nulo no PE referem-se ao nulo em oposição ao *se* em construções infinitivas, conforme exemplo (132) a seguir. Em construções finitas, não foram encontrados nas amostras analisadas exemplos de ocorrências de nulo genérico nem expletivo, tendo em vista o não emprego de *ter* com sentido existencial em PE.

Já na década atual, surpreendentemente, os dados revelam uma alteração em relação às estratégias favoritas para indeterminação no PE. De acordo com a amostra analisada, na atualidade, é o pronome *nós* que tem maior percentual de emprego, 48,4% das construções indeterminadas se realizam com esse pronome, seguido de *se*, com 29,5% das ocorrências. O pronome *a gente* se mantém pouco produtivo: em apenas 2,4% das construções do *corpus* atual esse pronome indeterminador é realizado na posição de sujeito.

O Gráfico 2 permite uma melhor visualização da distribuição das estratégias de indeterminação no PE nas duas décadas.



Exemplos extraídos do *corpus* atual do PE mostram o emprego desses pronomes indeterminadores, na sequência: nós (129), passiva com concordância (130), passiva sem concordância (131), *nulo* (132) e *a gente* (133).

(129) ...porque ***estamos*** habituados na construção civil é construir prédios. (OEI-C- 3H)⁸⁹

(130) Mas já ***se ouvem*** muitas coisas que não *se ouviam* não é? Coisas de arrastões à praia e coisas assim... (LISBOA-B3-M)

(131) ... porque o grupo era muito grande eu dava me com muita gente e realmente foi ***se perdendo os contactos*** eu acho que tem a ver muito com isso com...cada um depois segue na sua vida. (CAC-A3-M)

(132) Nesse aspecto pode ser bom mas em termos pessoais para ***Ø*** viver não não me atrai não me atraia não me atrai muito. (CAC-A3-H)

(133) Eles têm muito mais opções por onde ***a gente*** os possa encaminhar eles possam...ter várias experiências onde antes de decidir alguma coisa em termos de vida escolar deles futura eles podem experimentar muitas situações. (CAC-B3-M)

Considerando que as diferenças entre o PB e PE ultrapassam os limites da variação em análise, uma vez que o PE é LSN e o PB, LSNP, conforme já esperado, o cruzamento dos dados do PE não se mostrou relevante em todos os grupos de fatores elencados, tendo em vista principalmente a não ocorrência da variante inovadora, objeto desta pesquisa, razão por que os dados do PE não constarão de todas as análises.

Assim, a partir das evidências mostradas pelos dados, busca-se descrever um dos sintomas da mudança na marcação do sujeito nulo no PB: a emergência da estratégia de indeterminação com a forma pronominal plena – *você_{gen}*, em variação com as formas canônicas de indeterminação do sujeito: *se* indeterminador, *a gente* e *nós*, nulo genérico e nulo não referencial.⁹⁰

⁸⁹ Os exemplos citados neste trabalho referentes ao *corpus* do PE obedecem a seguinte convenção: as letras iniciais são siglas da cidade de origem informante (OEI= Oeiras/ CAC=Cacém); a seguir as letras A, B e C indicam a idade; na sequência tem-se as letras H e M que indicam o gênero e o número 3 que se refere ao nível de instrução.

⁹⁰ Embora o pronome nulo não referencial não constitua forma de indeterminação, ele está sendo considerado neste estudo como variante por ser intercambiável com as demais variantes em construções de *ter_{ex}*. Ressalto que não foram selecionados para análise empregos de nulo referencial.

5.1 Sujeitos nulos *versus* plenos

Os resultados apresentados na Tabela 1 confirmam que o PB vem demonstrando uma perda na opção pelo sujeito nulo não referencial e genérico em oposição ao expressivo aumento do preenchimento da posição de sujeito nesses contextos.

Com relação à frequência de sujeitos nulos x preenchidos no período analisado, a Tabela 1 mostra que, do total de 2.137 dados computados na década de 70, 23% (501) exibem sujeitos nulos (genéricos e expletivos) e 77%, sujeitos preenchidos (1.636). Quanto aos dados atuais, os resultados apresentados na Tabela 1 confirmam o declínio dos sujeitos nulos, com 12% das ocorrências (340), em oposição aos sujeitos preenchidos com formas pronominais plenas, que totalizam 88% das ocorrências (2.418).

5.2 Sujeitos nulos em variação com sujeitos genéricos pronominais

Verificada a gradual perda do nulo não referencial e genérico ao longo das décadas - 1970 e 2015, e a previsível compensação desta perda com o aumento do emprego das formas pronominais plenas nessa posição, resta analisar esta variação.

Conforme Galves (1987) e Nunes (1990), o nulo genérico tornou-se possível com a supressão do *se* indefinido. Mais especificamente, Nunes (1990) considera que a supressão do clítico *se* em construções de indeterminação é uma inovação na língua que se seguiu à reanálise do *se* passivo como *se* indeterminador, conforme discutido na seção 3.1. Galves (1987, p. 37) é pioneira em apontar esse emprego por meio do exemplo (134):

(134) Nos dias atuais não **Ø** usa mais saia.

A partir da identificação da possibilidade de referência indefinida do pronome nulo de 3^a. pessoa, conforme apontado por Galves (1987), Nunes (1990) levanta duas hipóteses para essa inovação: a primeira se justifica pelo fato de a concordância no PB admitir paradigmas como ele/eles fala; já a segunda hipótese busca apoio no quadro geral de supressão de clíticos no PB (Tarallo, 1993; Duarte, 1986).

Para Nunes (1990), o caminho para a escolha da hipótese mais adequada leva em conta a exclusão/inclusão do enunciador em construções como em (135) e (136) a seguir. Como aponta Maurer Jr. (1951, apud Nunes, 1990, p.100), enquanto a 3^a. pessoa plural designa um

sujeito indeterminado, mas não se liberta inteiramente da ideia primitiva de 3ª. pessoa do plural, na forma passiva inclui-se comumente a pessoa que fala, ou aquela com quem se fala.

(135) Nos dias atuais não **usam** mais saia.

(136) Nos dias atuais não **se** usa mais saia.

Nunes (1990) sugere, então, que já que a referência indeterminada pode incluir o enunciador, construções de nulo com a 3ª. pessoa do singular com referência indeterminada podem ser analisadas como evolução das construções com *se* indeterminador.

A intercambialidade entre sentenças com nulo genérico na posição de sujeito com *se* indeterminador, como em (137) e (138), extraídas do *corpus* atual do PB, evidencia essa proposta.

(137) a. Num outro cenário no futuro, quem sabe, *Ø_{gen}* ressuscita esse, *Ø_{gen}* tira esse projeto da gaveta. (UFV - M1 – 23)⁹¹

b. Num outro cenário no futuro, quem sabe, ressuscita-*se* esse, tira-*se* esse projeto da gaveta.

(138) a. E isso é uma situação que quanto mais *Ø_{gen}* faz parceria, mais surgem oportunidades, é uma coisa crescente. (UNI - M3 – 27)

b. E isso é uma situação que quanto mais *se* faz parceria, mais surgem oportunidades, é uma coisa crescente.

A consideração de construções de sujeito indeterminado sem *se* como evolução das construções com *se* indeterminador é interessante ainda para este trabalho por corroborar a hipótese de que o nulo que tende a permanecer na língua é para referência arbitrária, nos seguintes termos. Uma vez que no processo de reanálise do *se* a categoria vazia das passivas pronominais passa a ser interpretada como um pronome referencial e não mais como um expletivo, o qual é um *pro*⁹², esse fato explica não só o bloqueio do alçamento do argumento interno para a posição de sujeito, como também da inserção de um pronome pleno indefinido nesta posição, como no exemplo (141), considerando-se que (140) é “evolução” de (139):

⁹¹ Todos exemplos utilizados na análise, à exceção de (129) a (132), de (134) a (13) e de (137) a (140) foram extraídos dos *corpora* desta pesquisa.

⁹² Esta proposta de que o nulo arbitrário é considerado referencial, portanto *pro*, encontra respaldo em Kato (2000) que, embora descarte *pro* como categoria descritiva, admite que, em línguas como o PB, *pro* é mantido para sujeitos nulos arbitrários e anafóricos.

(139) Passa-*se* cartão.

(140) *Ø_{arb}* passa cartão.

(141) **Você_{gen}* passa cartão.

Assim, considerando-se a proposta de Nunes (1990) de que construções de nulo de 3ª. pessoa singular com leitura indeterminada são evolução de construções de passiva pronominal sem concordância e sem o *se*, nas quais o nulo é considerado referencial (*pro*) - portador de papel temático do argumento externo e de Caso nominativo, pode-se propor que em tais construções sem o *se*, o nulo é portador dos mesmos traços de *se*, daí o caráter arbitrário do pronome nulo nessas construções.

Diferentemente, construções com nulo expletivo (não referencial) e nulo genérico são bem-formadas com a inserção de um pronome genérico, como se pode constatar em (142) e (143):

(142) a. E nesse momento *Ø_{exp}* tem a presidente da república do PT, sofrendo processo de impeachment.⁹³

b. E nesse momento *você_{gen}* tem a presidente da república do PT, sofrendo processo de impeachment.

(143) a. Para ser jornalista, não *Ø_{gen}* pode ter muita vergonha não. (AAX-M1-14)

b. Para ser jornalista, *você_{gen}* não pode ter muita vergonha não.

Entretanto, há outros fatores que complicam o quadro, como construções em que a presença do suposto *pro* (nulo arbitrário) não bloqueia a inserção de *você_{gen}*, como no exemplo a seguir:

(144) Quando *você_{gen}* passa cartão, *você_{gen}* acumula pontos.⁹⁴

O que se constata, a partir de exemplos como o (144), é que há outros elementos responsáveis pela leitura genérica do sujeito das sentenças. De fato, os predicados genéricos e habituais, diferentemente dos episódicos, apresentam por si só uma leitura genérica, por meio de predicados hipotéticos (tempos e modos verbais), verbos modais e sintagmas adverbiais.

⁹³ Exemplos (142) e (143) foram extraídos do Programa *Em Pauta* da Globo News.

⁹⁴ Exemplo da fala espontânea.

Dessa forma, se se acrescentam a construções como (141) elementos que tornam o predicado habitual ou hipotético, como *quando*, *sempre que*, *se*, a leitura do predicado passa a genérica em vez de arbitrária, e a inserção do pronome é aceitável.

Constata-se, assim, que a marcação genérica do pronome inserido em construções de indeterminação sem *se* é em parte atribuída pela genericidade do predicado. Da mesma forma, a leitura genérica do sujeito nulo de referência indeterminada está associada a predicados genéricos, habituais ou hipotéticos, como se pode notar nos exemplos (137) e (138). É por isso que privilegio aqui, além de mostrar uma possível mudança em progresso, os contextos sintáticos que favorecem a leitura genérica do sujeito nulo.

Ainda em relação à supressão do clítico *se* – que resultou em sentenças de nulo genérico na posição de sujeito, concorrendo com esta mudança surge na língua uma competição entre o pronome nulo e o *você_{gen}*, na qual se observa que o PB passa a contar com estratégia alternativa de indeterminação do sujeito com uma forma pronominal genérica, *você*, além dos pronomes indeterminados *a gente* e *nós*, como se pode observar nas sentenças (145) e (146) a seguir, criadas partir do exemplo (137).

(145) Num outro cenário no futuro, quem sabe, *você_{gen}* ressuscita esse, *você_{gen}* tira esse projeto da gaveta. (UFV - M1 – 23)

(146) Num outro cenário no futuro, quem sabe, *a gente/ nós* ressuscita(mos) esse, *a gente/ nós* tira(mos) esse projeto da gaveta.

Sob outra perspectiva teórica, a proposta de Holmberg (2010) adotada neste estudo sobre o traço de definitude (D), exigido nos traços-*phi* das LSN consistentes, e não necessário em LSNPs, como uma propriedade daquelas línguas, é elucidativa quanto à leitura indefinida do pronome *nulo*, e aceitabilidade de construções como (137). Conforme abordado no referencial teórico, em LSNPs, sem D em T, a relação entre sonda (T) e alvo local (sujeito nulo ϕ P) não supre o valor de definitude, resultando em uma interpretação genérica do pronome nulo.

A presença/ausência desse traço de definitude em T, que pode co-ocorrer ou não com o traço de pessoa, possibilita ainda uma explicação para a distinção de um T com *você* segunda pessoa definida de outro com *você* genérico, e torna aceitáveis construções como (144). Ou seja, uma vez que o pronome genérico inclui todo mundo, sua referência é irrestrita, o que significa que ele não tem traços-*phi*, razão por que ele é intercambiável com o nulo genérico com o mesmo valor de verdade.

Ainda com respaldo na teoria de Holmberg (2010) acerca da incorporação dos traços do nulo indefinido em T e da projeção de SpecTP, depreende-se uma explicação para a emergência do preenchimento da posição de sujeito em construções de nulo genérico no PB.

Retomando esta proposta, em uma língua que não tem traços D em T, LSNP, o sujeito nulo indefinido, genérico, cujos traços não podem ser avaliados por um A-tópico, tem seus traços incorporados em vP, enquanto o nulo definido não é incorporado e é atraído pelo EPP para SpecTP. Dessa forma, se o sujeito nulo definido está em SpecTP, ele checka o EPP, enquanto o genérico que está em SpecvP não o checka. Para satisfação do EPP nesses casos, um elemento precisa ocupar a posição de SpecTP, normalmente via movimento de expressões adverbiais.

Pode-se considerar, assim, que, da mesma forma que o Filandês, língua parcialmente *pro-drop* (Holmberg, 2010), o PB tem uma condição EPP, na maior parte das vezes satisfeita pelo sujeito, mas que pode ser satisfeita por outras categorias, como os advérbios circunstanciais. Entretanto, no PB, a categoria que preferencialmente está satisfazendo esta condição é a dos pronomes indefinidos foneticamente realizados, via inserção e não via movimento. Os dados da tabela 1 evidenciam esse fato.

As construções (147) e (148), extraídas do *corpus* atual do PB, de nulo genérico na posição de sujeito (a) como evolução do *se* indeterminador (b), evidenciam que a inserção do pronome genérico de 2ª. pessoa é licenciada nessa posição, sem causar estranhamento, conforme se pode constatar nos exemplos em (c):

- (147) a. Começa na infância inclusive, porque \emptyset não vive mais aquela infância.
(AAX – F2- 6)
b. Começa na infância inclusive, porque não *se* vive mais aquela infância.
c. Começa na infância inclusive, porque *você_{gen}* não vive mais aquela infância.
- (148) a. A partir dos anos 60, toda mulher começava a sentir um calorzinho \emptyset já dava estrogênio pra ela. (AAX – M2- 8)
b. A partir dos anos 60, toda mulher começava a sentir um calorzinho já *se* dava estrogênio pra ela.
c. A partir dos anos 60, toda mulher começava a sentir um calorzinho *você_{gen}* já dava estrogênio pra ela.

É possível notar pelos exemplos em (148) que os elementos ‘a partir dos anos 60’ e ‘toda mulher’ atribuem um valor habitual, usual, ao predicado, tornando-o genérico, no sentido

de se aplicar a qualquer elemento de um grupo, possibilitando, assim, a inserção do pronome genérico *você* na posição de sujeito, como em (148)c.

Esta constatação levou à postulação dos fatores: *Presença de SADVs e SPs locativos e temporais* e *Caráter hipotético da construção* como possíveis favorecedores da forma pronominal inovadora *você_{gen}* na posição de sujeito como marcador de leitura indeterminada/genérica do predicado. A análise dos dados poderá comprovar a relevância desses fatores nas seções 5.7 e 5.8.

5.3 Nulo não referencial x *você_{gen}* em construções existenciais com *ter*

O preenchimento da posição do sujeito gramatical no PB com *você_{gen}* em construções de nulo não referencial, especificamente no caso de construções de *ter_{ex}*, mais que uma inovação na língua, constitui um dos contrastes mais relevantes entre as variedades brasileira e europeia do Português, evidenciando ainda mais a discrepância entre essas gramáticas, haja vista que no PE o verbo *ter* não instancia sentido existencial, conforme assinalam trabalhos como os de Avelar (2009), Avelar e Callou, (2011) e Marins (2013), que mostram que o verbo existencial canônico no PE é *haver*. Aliado a isso, o fato de que o pronome *você* com marcação genérica não consta da amostra analisada do PE, faz com que este seja um fator de destaque a ser investigado nesta pesquisa.

Dessa forma, a variação entre os pronomes nulo expletivo e *você_{gen}* em construções de *ter_{ex}* está sendo analisada como uma inovação encaixada em um conjunto mais amplo de mudanças na gramática do PB em relação ao PE, evidenciando uma divergência marcante entre as duas variedades da língua.

Portanto, a análise desse emprego nos dados coletados pode ser reveladora quanto ao fato de o PB estar caminhando para se tornar uma LSNN, em vista da tendência à perda do nulo expletivo com o preenchimento por formas pronominais plenas com marcação indefinida/genérica. Pode convalidar, assim, a hipótese de Duarte (2000) de que após perder o nulo referencial, seria natural esperar que os sujeitos não referenciais começassem também a se realizar foneticamente, apresentando o sistema do PB um conjunto de estruturas nessa posição, a qual, antes vazia, passaria a se mostrar preenchida.

Ainda que não cause estranhamento, o licenciamento de um pronome genérico na posição de sujeito de construções existenciais suscita investigação, visto que essas construções não apresentam sujeito semântico referencial, e o requerimento de preenchimento da posição de sujeito no PB, em princípio, deveria afetar apenas as construções com sujeitos

referencialmente definidos. Além disso, conforme discutido na seção 3.3.2 sobre a derivação de construções com nulo genérico e nulo expletivo, em construções existenciais com *ter* sem qualquer constituinte na posição de sujeito, SpecTP é projetado e ocupado por um *pro_{exp}* capaz de satisfazer o EPP, o que inviabiliza a proposta de que a inserção de um pronome nessas construções se deve à satisfação desse princípio, como no exemplo (7) apresentado na introdução desta tese e aqui repetido para efeito de esclarecimento.

(7) Não *Ø_{exp}* tem investimento na economia hoje.

Avelar & Callou (2011) apontam que uma possibilidade para elucidar a emergência de existenciais com *você_{gen}* tem a ver com o que Kato & Duarte (2003) chamam “condição anti-V1”, que é a tendência no PB a inserir elementos argumentais ou não argumentais na posição de sujeito gramatical, evitando a ocorrência do verbo em posição inicial absoluta.

Quanto às razões por que as construções com *ter* e pronome na posição pré-verbal desencadeiam sentido existencial, em vez de posse, uma explicação apontada por Avelar (2009) seria a de que *você_{gen}* não é argumento de *ter*; nas existenciais, esse pronome funciona como um sujeito gramatical que não interage tematicamente com o verbo. Portanto, por não constituir um sujeito semântico de *ter*, o pronome não recebe papel temático relacionado à interpretação possessiva atribuída pelo verbo, razão por que esse pronome pode ser licenciado em construções tipicamente impessoais. Assim, em uma sentença como

(149) *Você_{gen}* tem muitos castelos na Europa.

tendo em vista o valor genérico do pronome, a interpretação natural é existencial, equivalente a:

(150) Existem muitos castelos na Europa.

Conforme mostrado na seção 3.3.2 sobre a *sintaxe das construções existenciais com ter*, a proposta de Avelar (2009a) que considera que *ter* existencial e possessivo possuem a mesma estrutura de base, por dar conta do licenciamento de pronomes foneticamente realizados na posição de sujeito das construções existenciais com *ter*, é adotada neste estudo.

Contudo, pelas propriedades de *ter_{ex.}*, relativas à possibilidade de emprego de outros mecanismos sintáticos para um relativo “esvaziamento” da posição do sujeito, e de

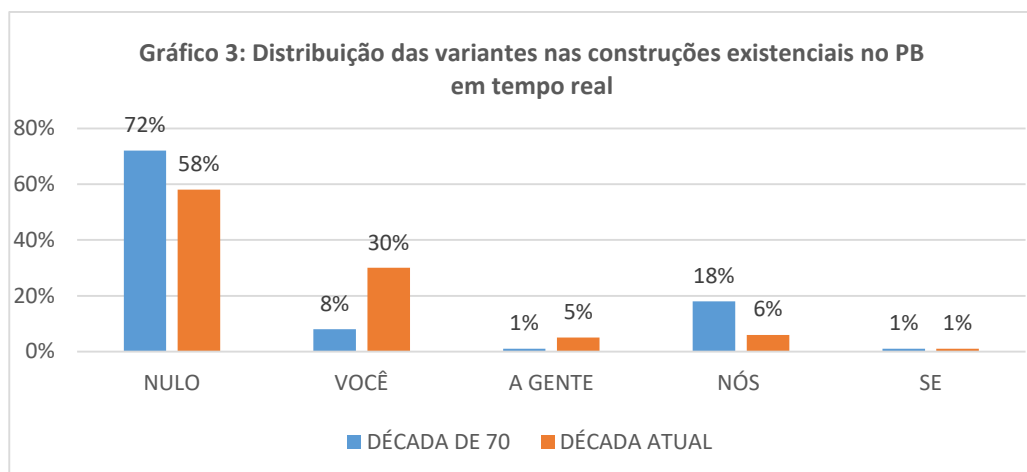
concordância do verbo com o sujeito gramatical, no caso do pronome indeterminado *nós*, testifica-se que as existenciais com *ter* não integram as chamadas construções impessoais, por não se caracterizarem, necessariamente, pela impessoalidade do verbo. Considero mais pertinente considerá-las detematizadas (cf. Viotti, 1998 e Franchi *et al*, 1998), dada sua incapacidade de atribuir papel temático ao sujeito gramatical, mas não necessariamente impessoais.

Deve-se considerar ainda a proposição de um avanço em relação à proposta de Avelar (2009) quanto aos tipos de pronomes que se realizam na posição de sujeito das sentenças existenciais no PB. Em Avelar (2009) e Avelar e Callou (2011) são analisadas somente as construções existenciais com *ter* em que a posição de sujeito é preenchida por *você_{gen}*.

A realização de outras formas pronominais, além de *você_{gen}*, em construções existenciais, é abordada em pesquisas como as de Duarte (2003), Berlinck, Duarte e Oliveira (2009) e Vitorio (2013), que confirmam que os pronomes utilizados nessas construções são os mesmos que representam o sujeito indeterminado, predominando o uso de *você*, *nós*, *a gente* e *eu*. É interessante notar que o trabalho de Duarte (2003) aponta que não só houve um aumento no emprego de pronomes indefinidos nas construções existenciais, como também aumentam os tipos de pronomes utilizados: na amostra de 80, só havia realizações de *você* e *a gente*, enquanto na de 2000, aparecem outros como *eu*, *nós*, *ele/ela* e *se*. Em trabalho realizado em 2013, sobre a fala culta alagoana, Vitorio (2013) constata que os pronomes *você*, *a gente*, *eu*, *nós*, *se* e *ele/ela* são realizados na posição de sujeito das construções existenciais.

De forma semelhante, nos *corpora* desta pesquisa relativos ao PB, os pronomes indeterminados *a gente* e *nós* aparecem realizados na posição de sujeito de construções existenciais.

Vislumbrando essa tendência no PB à realização de construções existenciais com o pronome *você_{gen}* na posição de sujeito gramatical, além dos indeterminados *a gente*, *nós* e *se*, e, diante da competição entre as variantes pronominais para marcação indefinida, e dos indicadores da Tabela 1 que apontam *você_{gen}* como o pronome mais frequentemente empregado para referência genérica no PB atual, realizo no Gráfico 3 uma análise comparativa entre as duas décadas pesquisadas quanto à evolução do emprego dessas variantes na posição de sujeito gramatical das construções existenciais com *ter*, as quais, por não selecionarem argumento externo, constituem construções de sujeito nulo expletivo, conforme discutido na seção 3.4.3.



Os dados do Gráfico 3 revelam perda na opção pelo sujeito nulo também nas construções existenciais, de 72 para 58%, além do aumento do preenchimento do sujeito com a inserção de *você_{gen}* nessas construções ao longo do período. Se na década de 70, o pronome *você_{gen}* era inserido em cerca de 8% das orações existenciais, na década atual, o pronome aparece em 30% dessas construções. O estudo em tempo real revela assim que na medida em que diminui a ocorrência de construções existenciais com *ter* com a posição de sujeito vazia, há implementação de *você_{gen}* nessas construções, o que sugere que este pronome está ocupando a posição do nulo expletivo em construções existenciais.

No seguinte trecho de uma entrevista televisiva da atualidade atesta-se a competição entre o nulo não referencial e *você_{gen}*, em que o emprego do pronome pleno é predominante:

(151) *Você_{gen}* numericamente ali *você_{gen}* **tem** a questão do número oficial, é lógico com a força do governo por trás, isso acaba dando uma vantagem pro Leonardo Piciane, mas por outro lado *você_{gen}* **tem** o Eduardo Cunha trabalhando nos bastidores e aí \emptyset **tem** um poder de pressão altíssimo.⁹⁵

Outra constatação a partir do Gráfico 3 é que, embora a frequência das formas pronominais *nós* e *a gente* não seja relevante como a de *você_{gen}* nas construções de *ter_{ex}* na década atual, tais formas são realizadas em existenciais desde a década de 70, quando o emprego de *nós* superou o de *a gente* e *você* juntos nesses contextos. Já o emprego de *se* em sentenças de *ter_{ex}* se manteve estável ao longo das décadas.

Considerando esses resultados e comparando-se as Tabelas 1 e 3 e Gráficos 1 e 3, confirma-se a previsão deste estudo de que a mudança na marcação do PSN no PB afetou não só a representação dos sujeitos pronominais de referência definida, como os de referência

⁹⁵ Exemplo (151) foi extraído do Programa *Em Pauta* da Globo News, em: 16/02/2016.

indefinida, atuando também em sentenças existenciais, com a inserção de um pronome pleno - *você/ cê, a gente e nós* - na posição de sujeito gramatical com marcação genérica/indeterminada em substituição ao nulo não referencial.

A questão que se coloca aqui é: em que diferem as construções existenciais empregadas com os pronomes *a gente* e *nós*, de um lado, e *se* e *você_{gen}* de outro, considerando-se a indefinição referencial de *se* e de *você_{gen}*, e o avanço do pronome *a gente* em direção à referência mais específica, assim como *nós*?

Em Avelar (2004), o autor estabelece um paralelo entre as sentenças com *você_{gen}* e aquelas em que se tem um pronome caracterizando uma estratégia de indeterminação do sujeito, e sugere que a relação que se estabelece entre tais pronomes e o verbo *ter* é de posse. Em outro trabalho, Avelar (2009b) sugere outra possibilidade de representação formal para as sentenças com *você_{gen}*, e defende que as sentenças com *ter* em que *você_{gen}* ocupa a posição de sujeito devem ser interpretadas como existenciais, mas não aborda as construções com *ter* e os pronomes de indeterminação *a gente* e *nós*.

A proposta de Avelar (2009b), conforme já discutido, para a localização de *você_{gen}* em sentenças em que *ter* tem interpretação existencial como em (149) *Você_{gen}* tem muitos castelos na Europa, é a de que o pronome genérico é inserido no interior do constituinte locativo ‘na Europa’, onde recebe seu papel temático, e desse locus é movido para a posição de sujeito gramatical (Spec-T/Infl).

Embora esta análise seja interessante para explicar por que a inserção desse pronome não desencadeia, nas orações com *ter_{ex}*, interpretação de posse indeterminada, entendo que o caráter genérico do pronome *você* não permite uma coindexação com o sintagma locativo.

De outra forma, deve-se notar que, embora sejam indeterminados, os pronomes *a gente* e *nós* expressam mais especificidade do que os pronomes *se* e *você_{gen}*. Assim, ao empregar os pronomes *a gente* ou *nós* em construções existenciais, o enunciador remete geralmente ao lugar ao qual se refere, incluindo-se na enunciação. No exemplo a seguir (152), extraído de uma entrevista da amostra do PB atual,

(152) **Nós** não *tínhamos* recursos humanos na época, né, tivemos que correr atrás.
(AAX – M3)

percebe-se que se trata de uma construção existencial, por ocorrer com constituinte interno e sem constituinte sujeito definido e o predicado bloquear a leitura possessiva, além de a construção poder ser parafraseada por *haver*:

(153) Não **havia** recursos humanos na época, né, tivemos que correr atrás.

E, ainda, o pronome pode ser eliminado sem prejuízo para a manutenção do valor existencial da oração, como pode-se perceber em (154):

(154) Não **Ø tinha** recursos humanos na época, né, tivemos que correr atrás.

Nesse caso, nota-se que o pronome *nós* em (152) não recebe papel temático do verbo, por não constituir seu sujeito semântico. Embora o constituinte locativo não esteja presente na sentença, ele pode ser recuperável, uma vez que o falante faz referência ao local no qual se insere como participante da relação de existência estabelecida pelo verbo *ter*, o que faz pressupor que o pronome *nós* tem ligação com o lócus. Pode-se entender pela relação estabelecida que o enunciador dessa sentença deve ser alguém que pertença ou tenha no seu conjunto de conhecimento as características daquele lugar. Ou seja, a sentença (152) pode ser entendida como uma narração na perspectiva do enunciador.

Assim, à luz da proposta de Avelar (2009b) de um *link* do sujeito gramatical de sentenças de *ter* existencial com o constituinte locativo, pode-se considerar que os pronomes indeterminados *a gente* e *nós*, que também aparecem realizados na posição de sujeito gramatical dessas construções, têm imersão com o constituinte locativo para depois serem movidos para a posição de sujeito, Spec-TP, uma vez que, por não serem selecionados pelo verbo, este não lhes atribui papel temático. A relação semântica estabelecida é entre o pronome e o lócus, ou ancoragem locativa, quando este não está presente na sentença.

Por sua vez, *você_{gen}*, assim como o pronome *se*, expressa maior indefinição referencial; por incluir todas as pessoas, sua referência é irrestrita. A sugestão, portanto, é que o pronome *você_{gen}* tem imersão distinta dos pronomes *a gente* e *nós*: da mesma forma que o clítico *se*, *você_{gen}* é inicialmente inserido em INFL, e posteriormente movido para Spec TP. Essa discussão está explicitada na seção 3.4.2 sobre *o caráter indeterminado de você_{gen} em existenciais*.

Note-se que uma sentença como (149) *Você tem muitos castelos na Europa* pode ser empregada por qualquer falante que mantenha ou não algum tipo de relação com o lócus Europa, enquanto que, se na posição de sujeito figurar um dos pronomes *a gente* ou *nós*, como em (155) e (156),

(155) **A gente** tem muitos castelos na Europa.

(156) **Nós** temos muitos castelos na Europa.

o sentido existencial só é instanciado se a sentença for empregada por um falante que mora, é originário da Europa ou esteja lá no momento da enunciação.

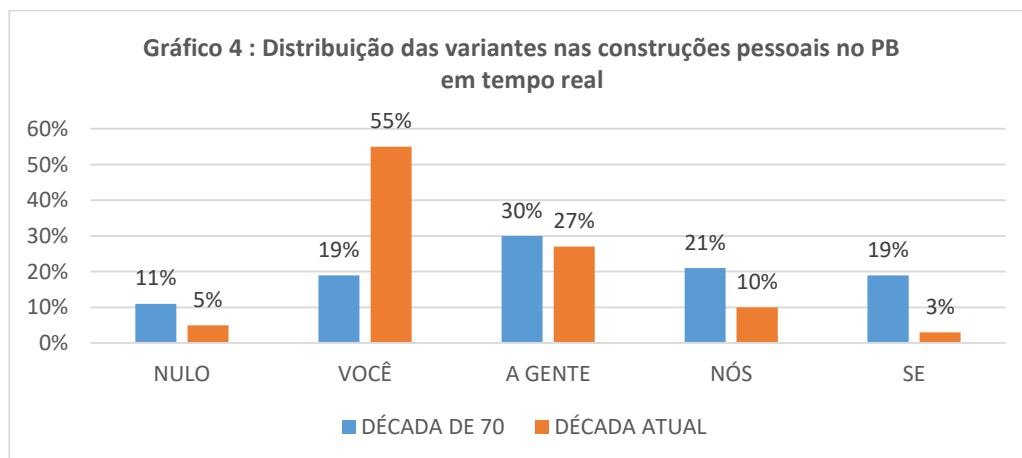
Além disso, deve-se ressaltar que há ocorrências no *corpus* do PB atual de construções de *ter_{ex}* com a realização de *você_{gen}* na posição de sujeito sem a presença de um sintagma locativo ou mesmo uma ancoragem locativa, conforme se pode evidenciar pelos exemplos a seguir:

(157) Afinal de contas como é que *você_{gen}* pode ter uma ciência que não tem um método que a caracterize? *Você_{gen}* pode dizer, mas eu não estudo linguística de *corpus*... então *você_{gen}* tem a linguagem. (UFU – M3 – 30)

(158) *Você_{gen}* tem uma outra droga, a fluoxetina, paciente com pânico e uma paciente que quer emagrecer, por ansiedade, ela vai para fluoxetina. (AAX-M2- 15)

Por esta razão, um outro problema que esta proposta de Avelar (2009b) sobre a imersão do pronome genérico no lócus em estruturas de *ter_{ex}* pode apresentar é de que o tratamento dado a essas estruturas é na presença de um sintagma locativo, mas não se aplica ou não esclarece as construções com *ter_{ex}* sem ancoragem locativa. A ocorrência dessas construções sem a presença de um constituinte locativo será retomada no item 5.8 que analisará o seu favorecimento na presença x ausência de SADVs e SPs locativos e temporais.

Contrariando a expectativa de que o pronome *você_{gen}* seja mais recorrente em construções existenciais, em função da produtividade de seu emprego observada na pesquisa empírica, os dados indicam que é nas demais construções que a frequência desse pronome ganha destaque, conforme se pode verificar no Gráfico 4.



Nas sentenças pessoais, a análise em tempo real chama a atenção para a significativa diferença percentual do emprego de *você_{gen}* no período analisado, passando de 19 para 55% de

frequência. Esse resultado evidencia o rápido avanço no processo de propagação da forma inovadora, em curto período de tempo, suplantando o emprego de *a gente*.

Em relação às demais variantes, enquanto os pronomes *nulo*, *nós* e *se* exibem decréscimo linear de emprego com marcação indefinida em construções pessoais, *a gente* se mantém estável ao longo do período. A competição entre as formas *você_{gen}* e nulo genérico em sentenças pessoais é evidenciada no exemplo (159) abaixo:

(159) *Cê* não pode considerar só aspectos físicos, é importante o que a Sandra fala, *cê* tem que ter empatia, *Ø_{gen}* tem que ter inteligência, acho que vários outros atributos porque *você_{gen}* pode tá ali com miss Brasil do teu lado, se não *Ø* tem empatia, não *Ø* tem aquela troca, não adianta né?⁹⁶

5.4 *A gente* cede lugar a *você*: o que dizem os dados

A proposta de Kato (2000) sobre pronomes fracos e fortes igualmente lança luz sobre a consideração de que os sujeitos nulos genérico e não referencial vêm sendo cada vez mais preenchidos pelos falantes do PB. Segundo esta teoria, uma das propriedades que identificam LSNN está no surgimento de um paradigma de pronomes fracos, o que pode ser visto como uma estratégia para compensar a perda da propriedade de licenciar sujeito nulo e para satisfação do EPP da sentença em contextos nos quais a língua licenciaria sujeito nulo.

Uma evidência de que *você_{gen}* pode ser categorizado como pronome fraco está em que os pronomes fortes são mais estáveis fonologicamente se comparados com os pronomes livres e os clíticos, o que pode ser notado, por exemplo, pelas formas pronominais fortes e fracas do paradigma pronominal do português (*cf.* Duarte, 2008):

EU > eu > ô (Duarte, 2008, p. 34, exemplo 9)
VOCÊ > ocê > cê

Para o presente estudo, a relevância da consideração de *você_{gen}* como pertencente ao paradigma de pronomes fracos está em que sua forma foneticamente reduzida – *cê* – é empregada similarmente ao pronome *você* na marcação genérica em posição de sujeito

⁹⁶ Exemplo extraído do Programa *Em Pauta* da Globo News, exibido em: 30/11/15.

gramatical no PB⁹⁷, cujo comportamento sintático está sendo comparado com o clítico *se*, os quais podem ser usados em contextos semelhantes com similar função indeterminadora, dando ensejo à postulação do percurso de indeterminação de *se* a *cê*.

Por esta proposta de Kato (2000), é possível estabelecer ainda a distinção entre *você* 2ª pessoa definida e *você* pronome genérico, como pertencentes a paradigmas distintos, respectivamente de pronomes fortes e fracos.

Esta proposta correlaciona-se, ainda, com a hipótese de que o emprego de *você_{gen}* está suplantando o emprego de *a gente*, em função da perda da propriedade de marcação genérica deste pronome, por estar sendo empregado mais recorrentemente como pronome de 1ª. pessoa plural, portanto menos inclusivo (OMENA, 2003) e, como forma de compensação dessa perda, o sistema lança mão de outra estratégia, *você_{gen}*, que surge na língua no paradigma de pronomes fracos.

Pela relação dos fenômenos do surgimento do paradigma de pronomes fracos e da substituição dos pronomes *a gente* por *você*, pode-se evidenciar o encaixamento linguístico, considerando-se com Labov (2008) que toda mudança produz reflexos que propiciam o surgimento de outras estruturas a ela associadas.

Na esteira do princípio do uniformitarismo, a distribuição das variantes nas construções pessoais nas duas décadas, conforme Gráfico 4, evidencia que as forças que operaram no passado são as mesmas que operam no presente, uma vez que, de forma semelhante ao processo de substituição de *nós* por *a gente* para referência indeterminada no passado, o emprego do pronome *você* com valor genérico está suplantando o de *a gente* (55% x 27%), o qual, de alguma forma, perdeu esse valor, seja porque ele tem sido mais usado com referência específica para se referir a 1ª pessoa plural, ou porque assim como o ‘*on*’ do Francês, ele se sobrecarregou, uma vez que a alternância *a gente* x *nós* pode ser observada em outras funções além da de sujeito, tanto no nível do sintagma verbal como no interior de sintagmas nominais.

A variação entre *nós* e *a gente* no PB é muito estudada quando a alternância das formas ocorre em posição de sujeito, posição sintática mais produtiva à entrada de formas inovadoras, mas esta alternância pode ser observada em outras funções no nível do sintagma verbal, como no exemplo (160). Isso pode ter sobrecarregado o pronome, dificultando a referência genérica na posição de sujeito.

⁹⁷ Deve-se esclarecer que o emprego da forma reduzida *cê* foi observado nos dados do estado de Minas Gerais, não podendo ser generalizada como característica do PB.

(160) Então, a cultura *da gente* no geral era... tornava-se muito mais difícil, *nós* tínhamos que fazer a cultura por nós mesmos.

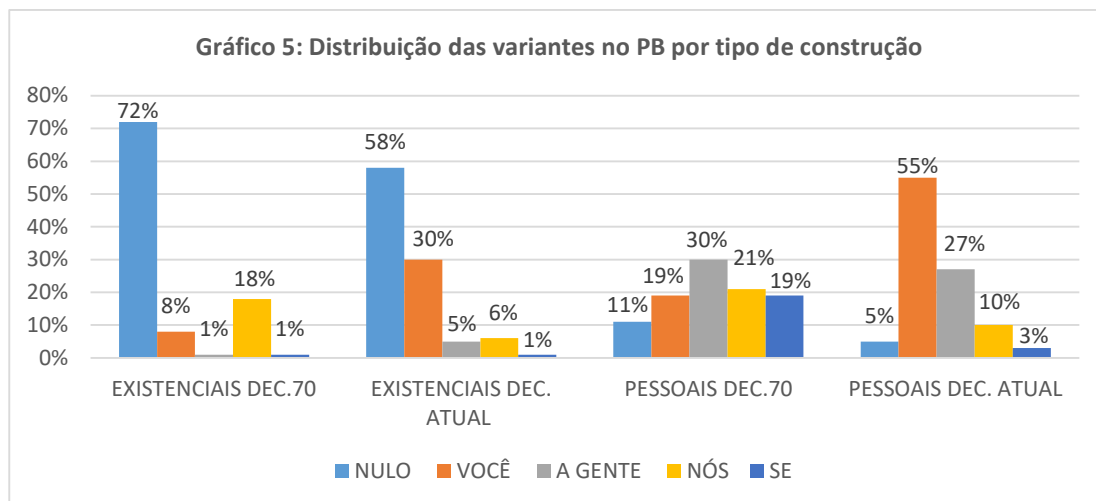
O emprego de *a gente* em outras funções como complemento do verbo e do nome foi pesquisado por Vianna (2011) na perspectiva da sociolinguística laboviana. O estudo concluiu que a forma ‘inovadora’ está ocupando os espaços de *nós* no nível oracional de maneira geral e não apenas na função de sujeito, como é constatado em inúmeras pesquisas do PB. O estudo identificou, ainda, que nessas funções o pronome *a gente* assume o caráter identificador de 1ª pessoa plural, explicitando o referente. Segundo Vianna (2011), ao dizer, por exemplo, “Ele saiu com *a gente*”, o falante elucida que a ação foi realizada com a pessoa X e não com Y.

Assim, a entrada da forma *a gente* em funções gramaticais diferentes da de sujeito pode tê-la descaracterizado como pronome de referência genérica, uma vez que, nas funções de complemento, *a gente* se comporta como pronome de 1ª pessoa do plural indeterminado, mas não genérico ou indefinido.

Se *você* e *a gente* genérico são formas variantes intercambiáveis em contextos (+) inclusivos, fica evidenciado pelos Gráficos 4 e 5, a seguir, que está havendo um espaço para referência genérica na forma *a gente*, que está permitindo a entrada da forma pronominal *você* em seu domínio, visto que a forma *a gente* como estratégia de indeterminação do sujeito tem sido pouco produtiva relativamente a *você*, 27% contra 55%, em construções pessoais, e 5% *versus* 30% em existenciais, considerando-se ainda que esses percentuais incluem os contextos de referência mais específica, para os quais o pronome *a gente* tem se expandido.

Constato, assim, que, simultaneamente a esse processo de difusão de *a gente* para referência mais específica, a implementação do pronome *você* com referência genérica tem se expandido largamente na língua falada no PB. Portanto, é possível inferir que a sobreposição das formas *você_{gen}* e *a gente* para marcação genérica está quase completa no PB.

O Gráfico 5 a seguir ilustra a comparação da distribuição das variantes nos dois tipos de construção – existenciais e pessoais, em cada um dos períodos analisados.



Semelhantemente ao que ocorreu em Francês e Espanhol, conforme mostrado em 3.3.1, o emprego da 2ª pessoa *você* para referência genérica no PB é recente e, pela observação dos resultados dos dados investigados, Gráfico 5, é possível afirmar que ele vem sendo incrementado a partir das últimas décadas.

Na década de 70, é perceptível pelos dados que o pronome preferido para referência genérica em construções pessoais ainda era *a gente*. Já na década de 90, tem-se o estudo de Duarte (1995) que atesta que *você* é mais empregado do que *a gente* para referência indeterminada, e na década de 2000, o estudo de Silva e Almeida (2010) com falantes semicultos feirenses confirma que a forma preferida para indeterminação é *você* com 61% de ocorrências, seguida de *a gente*, com 24%; 3ª pessoa do singular sem o ‘se’, 4%; e 3ª pessoa do singular com o ‘se’, menor ocorrência, 1%. Evidencia-se assim a recente difusão do emprego de *você_{gen}* no PB.

A competição entre os pronomes *você* e *a gente* para referência genérica no PB atual pode ser constatada por meio dos seguintes trechos de entrevistas televisivas⁹⁸:

(161) *Você_{gen}* tem um grande momento de incertezas, *você_{gen}* vai fechar o ano com essa incerteza e vai chegar em 2016 com essa incerteza. Por outro lado, *você_{gen}* tem o seguinte, se a presidente é afastada pelo senado, e aí *a gente* pensa que o presidente da câmara está efetivamente sendo investigado, *a gente* tem um grande imbróglia que parece que ninguém se salva.

(162) Com a presidente Dilma, é diferente, e *você_{gen}* tem um processo que não, é, vai ser como o do Collor, então o que *você_{gen}* pode ver, o que é certo é que *você_{gen}* vai ter muita oscilação no mercado financeiro. *A gente* vai viver momentos de euforias e de pessimismos.

⁹⁸ Exemplos (161) e (162) foram extraídos do Programa *Em Pauta* da Globo News, em 02/12/16.

Extraídos do *corpus* da década de 70 do PB, os exemplos a seguir mostram o emprego genérico do pronome *a gente* em construções nas quais, na atualidade, possivelmente, a referência genérica do sujeito seria realizada pelo pronome *você_{gen}*, dada a sua indefinição referencial.

(163) Eu acho que o Fellini sempre ele aborda temas bastante atuais, não sofisticado, então aquilo que *a gente* vê no filme. (RJ-F3-259)

(164) Então isso já é um grande melhoramento, *a gente* poder assistir o filme tomando um café ou tomando um drinque já melhora ainda mais, né, principalmente quando *a gente* vai pra curtir, curtir mesmo. (RJ-F3-259)

(165) Então no, no dicionário dele vai ter muita gíria, então como é que *a gente* vai evitar de falar gíria quando está conversando normalmente? (RJ-F3-259)

Desta análise resta ressaltar que, embora seja possível inferir que o PB está percorrendo o mesmo percurso que o Francês quanto à substituição de pronomes de referência genérica pela 2ª pessoa, é possível constatar que esse processo se dá em estágios distintos em relação à remarcação do parâmetro *pro-drop*: enquanto em Francês, o emprego do pronome de 2ª pessoa com referência genérica surge após a efetivação da mudança do *status* de língua *pro-drop* para não *pro-drop*, no PB esse processo acontece concomitantemente a outras alterações de propriedades associadas ao parâmetro *pro-drop*, mais especificamente, a perda do clítico *se* e do nulo genérico e expletivo, caracterizando-se o pronome de 2ª. pessoa genérico como uma estratégia de indeterminação que está preenchendo a posição vazia do sujeito.

Por essa razão, este fenômeno linguístico parece evidenciar uma característica de LSNNs, a qual, no caso do PB, surge em estágio de alteração paramétrica, na perspectiva da sociolinguística paramétrica, ou como uma propriedade de LSNPs, visto sob o prisma gerativista.

Sob a perspectiva sociolinguística, a correlação deste fenômeno no PB com a substituição em Francês do *on* pelo *tú* esbarra com outra divergência quanto ao grau de formalidade, uma vez que *on* é padrão, formal e aceito em todos os registros (COVENEY, 2003), diferentemente de *a gente*, que, da mesma forma que *você* em sua acepção genérica, é um pronome informal e não aceito no registro escrito padrão.

Além disso, estudos sobre o emprego de *tú* genérico em Francês apontam que se trata de uma mudança de baixo para cima, em que a variante inovadora *tú* é mais empregada pela classe trabalhadora, mas não só, geralmente na linguagem coloquial do dia-a-dia (COVENEY,

2003). Portanto, a informalidade do pronome *tú* em Francês aparece como um segundo fator motivador da variação.

Já no PB, a substituição é de um pronome informal por outro igualmente informal, e os contextos de emprego do pronome de 2ª pessoa sugerem que ele é produtivo em discursos argumentativos, mais propensos a se proferirem argumentações, explanações, generalizações, suposições. Além disso, tendo em vista que para a seleção dos informantes que comporiam os *corpora* desta pesquisa, o grau de escolaridade superior foi preestabelecido, os dados foram extraídos de entrevistas com falantes considerados cultos⁹⁹.

A despeito das diferenças apontadas, a proposta postulada neste estudo de que o PB está passando por processo semelhante ao Francês de substituição de pronomes referenda-se no fato de que *você_{gen}* pode estar suprimindo a carência de um pronome para referência genérica, causada pelo avanço de *a gente* em direção à referência mais específica. Para atestar essa hipótese, foi elencado o fator ‘grau de indeterminação das variantes’, cujos resultados serão mostrados a seguir.

5.5 Grau de indeterminação

O fator ‘grau de indeterminação das variantes’ teve como objetivo averiguar se a indeterminação das variantes é (+) ou (-) inclusiva, considerando a escala de gradação de arbitrariedade (96), conforme proposto por Duarte *et al* (2015) e apresentada no item 3.4.4 do referencial teórico: *Referência arbitrária x referência genérica*.

(96) (arbitrário/ menos genérico) eles < eu < nós/ a gente < se/ zero/ você (mais genérico)

Correlacionando as propostas de Holmberg (2015) e de Duarte *et al*, propus que quanto mais inclusiva for a marcação do pronome, mais indefinida será sua referência e, de forma oposta, quanto menos inclusiva a marcação, menos indefinida será sua referência. Assumindo essa correlação, adoto nesta análise as denominações ‘mais inclusivo’ para os pronomes *se/ zero/ você* e ‘menos inclusivo’ para os que se encontram na posição intermediária da escala

⁹⁹ Retomo aqui o conceito de falante culto conforme os critérios adotados pelo Projeto NURC (Projeto de Estudo da Norma Linguística Urbana Culta). Nas pesquisas do Projeto, classifica-se o falante culto como o indivíduo que possui nível universitário, ou seja, o falante instruído formalmente e que, portanto, foi exposto à norma explícita no ambiente escolar.

proposta por Duarte *et al*: *nós* e *a gente*. Ressalto, entretanto, que os pronomes *a gente* e *nós* realizam no PB marcação indeterminada tanto menos como mais inclusiva.

Por essa razão, a criação deste fator fez-se necessária para dar conta dos empregos de *a gente* na amostra, os quais, em sua maioria, embora não consistam pronomes genéricos propriamente, mas indeterminadores, não poderiam ser descartados, tendo em vista a hipótese aventada neste estudo da substituição de pronomes - *a gente* por *você* - no PB para marcação genérica. Este fator também é importante para convalidar a hipótese de que os nulos arbitrários tendem a permanecer na língua. Os resultados dessa classificação estão apresentados na Tabela 5 a seguir:

Tabela 5: Grau de indeterminação das variantes por período no PB

Variantes	Década de 70		Década atual	
	(+) inclusivo	(-) inclusivo	(+) inclusivo	(-) inclusivo
Nulo	500 27%	1	341 16,5%	-
Você	334 18,5%	17 4%	1407 68%	17 2,5%
A gente	335 18,5%	176 56%	168 8%	494 71%
Nós	308 17%	128 40%	77 4%	183 26,5%
Se	337 19%	-	69 3,5%	-

Os dados da Tabela 5 são relevantes para a comprovação da hipótese de que na atualidade o pronome *a gente* é mais produtivo para referência indeterminada (-) inclusiva, portanto (+) específica. Conforme se pode evidenciar, na década atual, o pronome *a gente* aparece em 71% do total das ocorrências de pronomes indeterminados (-) inclusivos, em oposição a somente 8% do total de pronomes empregados para referência indeterminada (+) inclusiva. A análise em tempo real mostra uma diminuição da referência (+) inclusiva do pronome *a gente* ao longo do período, que cai de 19% nos anos 70 para 8% de emprego na atualidade. Inversamente, observa-se um aumento no emprego desse pronome para referência (-) inclusiva, que salta de 51% nos anos 70 para 71% de emprego nos dados atuais.

Pela leitura horizontal da Tabela tem-se que 75% dos empregos de *a gente* no *corpus* atual apresentam marcação (-) inclusiva, razão por que sua ocorrência ainda é expressiva, se se considera que para leitura indeterminada (+) inclusiva *você* está suplantando *a gente*, conforme vem sendo discutido neste trabalho.

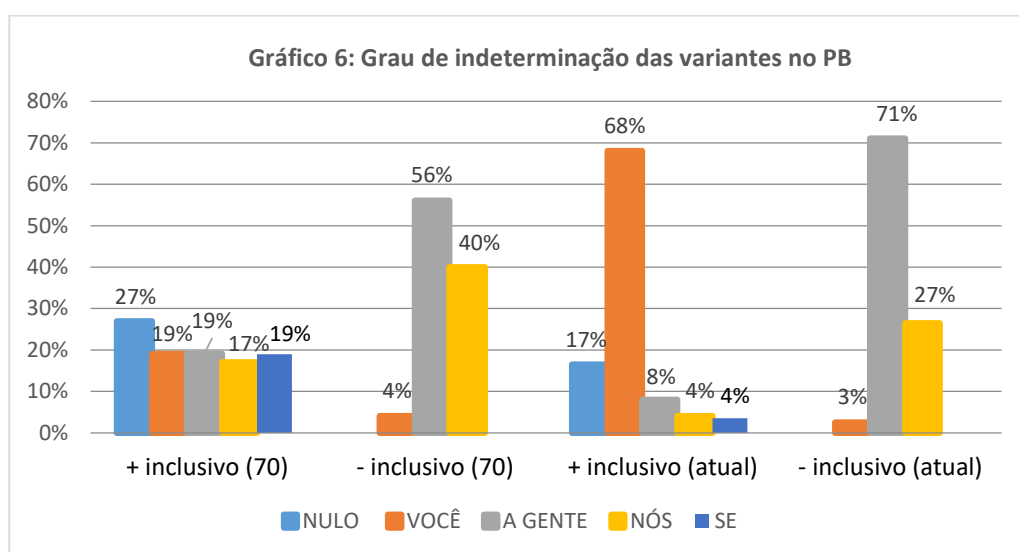
Assim, pela Tabela 5 é possível confirmar a hipótese de que *você_{gen}* está suprimindo a gradual carência de um pronome para referência genérica no PB, causada pelo avanço de *a gente* em direção à referência mais específica.

Por meio dos exemplos a seguir extraídos do *corpus* atual do PB, pode-se evidenciar essa diferença de marcação em relação ao pronome *a gente*: (-) inclusivo, exemplo (166) e (+) inclusivo exemplo (167):

(166) Tinha uma época há muito tempo atrás ***a gente*** tinha um caixa dois aqui, esse caixa dois era um problema seríssimo, e ele acabou sendo colocado sobre (*judis*). (CEFET – M3 – 18)

(167) Mas o menino antigamente entrava cruzinho, né? Sem saber nada, era o normal, e ***a gente*** tinha que alfabetizar certo período, né? (AAX -F3 – 13)

Para melhor visualização da evolução das variantes em relação ao grau de indeterminação ao longo do período, os resultados estão ilustrados no Gráfico 6.



A diferença percentual de emprego de *você_{gen}* (+) inclusivo e *a gente* (+) inclusivo na atualidade, 68% e 8%, dá conta do problema da transição, conforme postulado por Labov (2008), que é o de encontrar o caminho pelo qual um estágio de uma mudança linguística evolui a partir de um estágio anterior. A partir desse resultado é possível afirmar que a propagação de *você_{gen}* se deu pela saturação do pronome *a gente*, quando este assume diversas funções na língua, perdendo os traços de pronome indeterminado (+) inclusivo. Aponto, assim, essa como uma das possíveis causas provocadoras da mudança (implementação de *você_{gen}* no PB), analisadas pelo estudo da mudança linguística em andamento.

Exatamente na contramão desse resultado, *você_{gen}* responde por somente 3% dos pronomes empregados com referência indeterminada (-) inclusiva. Embora esses empregos menos inclusivos de *você_{gen}* não sejam percentualmente relevantes, é interessante mostrá-los diante da expansão de sentidos que esse pronome vem assumindo na língua, conforme exemplo (168) do *corpus* atual, em que *você_{gen}* é empregado para fazer referência à 1ª. pessoa do singular.

(168) E tem, muitas vezes os filhos da gente cobram coisas que *você_{gen}* fez de errado com eles e eu falei pra ele, até é o meu filho mais velho, esse que é médico. (AAX – F3 – 9)

Outra constatação a partir deste Gráfico 6 diz respeito à diminuição no emprego do pronome nulo (-) inclusivo ao longo das décadas. Conforme discutido em 3.4.4, a previsão era de que o nulo indefinido que tende a permanecer na língua seria para realização de uma leitura arbitrária que o pronome *você_{gen}* não é capaz de realizar.

A sentença (169) exemplifica o emprego do nulo genérico (-) inclusivo, ou arbitrário:

(169) Aquela fruta de conde que aqui no Rio é caríssima lá... *Ø_{arb}* vende assim com preço baratíssimo... (RJ 328-F2)¹⁰⁰

Entretanto, na amostra analisada, o emprego do nulo arbitrário, (-) inclusivo, é inexpressivo. Deve-se considerar que o baixo emprego de nulo arbitrário pode estar, ainda, relacionado ao problema previsto no método sociolinguístico laboviano relativo à raridade das formas gramaticais com a qual o pesquisador pode se deparar. Nos dizeres de Labov (2008, p. 224), “qualquer tentativa de especificar regras sintáticas implica inevitavelmente formas que podemos esperar não ouvir numa investigação limitada”. Este parece ser o caso do nulo arbitrário. Aliado a este fator, conforme discutido em 5.2, a leitura genérica do sujeito nulo está associada a predicados genéricos ou habituais, os quais predominam em entrevistas orais, especialmente de cunho argumentativo.

Pelo fato de não haver ocorrências na amostra desta pesquisa de sentenças isoladas como “*Ø Passa cartão*”, pode-se afirmar, por conseguinte, que não há ocorrência de pronome nulo arbitrário.

Assim, à exceção das construções existenciais, de nulo expletivo, em todas as demais ocorrências nos *corpora* do PB os pronomes nulos são considerados genéricos, sobretudo, por

¹⁰⁰ Exemplo extraído do *corpus* do PB da década de 70.

serem intercambiáveis com as variantes deste estudo. Portanto, a hipótese de que o nulo que tende a permanecer no PB é o arbitrário é refutada pelos dados desta pesquisa.

Em contrapartida, os dados sinalizam, conforme Gráfico 5, que o nulo mais proeminente no PB na atualidade continua sendo o expletivo, cujo emprego em sentenças de *ter_{ex}* é predominante nas duas décadas analisadas. Por este fato depreendo que os nulos que permanecem na língua são os expletivos em construções existenciais com *ter*, apesar de, nos dados atuais, uma relevante parcela delas (42%) apresentarem a posição do sujeito preenchida, conforme Gráfico 3.

Este resultado em relação ao emprego do nulo não referencial revela que o sistema do PB ainda tem a capacidade de interpretar a categoria vazia na posição estrutural de sujeito de verbos existenciais com *ter* como um pronome nulo expletivo. Deve-se ressaltar que a possibilidade de ocorrência de sujeito nulo expletivo é apontada por Duarte (1995) como uma das propriedades das línguas *pro-drop*, embora não seja exclusiva delas, conforme mostram Jaeggli & Safir (1989 apud Duarte, 1995) sobre o alemão. Nesse sentido, pode-se afirmar que o PB ainda exhibe propriedades de LSN.

Para uma análise da evolução do emprego dos pronomes nulos genérico e expletivo em relação às demais variantes em tempo real de curta duração, considerando as hipóteses iniciais deste trabalho, comparo seus empregos em construções existenciais e pessoais no Quadro 1 a seguir, baseado nos Gráficos 3 e 4.

Quadro 1: Frequência dos pronomes nulo genérico e expletivo no PB nas duas décadas em relação às demais variantes		
	Década de 70	Década atual
Nulo genérico (construções pessoais)	11%	5%
Nulo expletivo (construções existenciais)	72%	58%

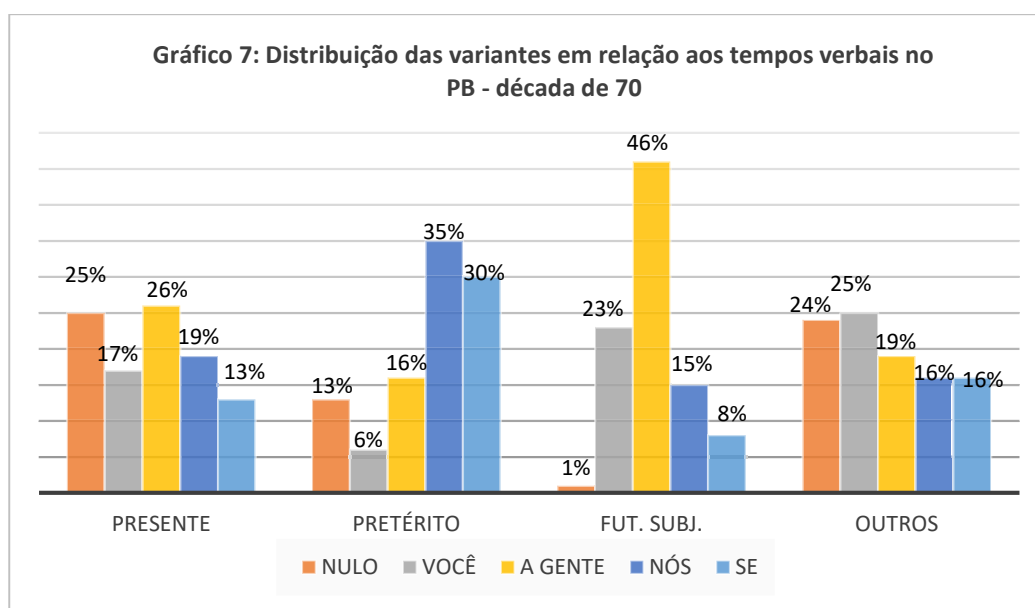
Pelo quadro 1 pode-se evidenciar que o nulo expletivo é produtivo na fala atual enquanto o genérico é inexpressivo. Com realização em 72% das construções existenciais na década de 70, o emprego do nulo expletivo é mantido na maioria destas construções na década atual: 58%. Já o nulo genérico, que ocorre em apenas 11% das construções pessoais em 70, tem seu emprego reduzido para apenas 5% dessas construções na atualidade.

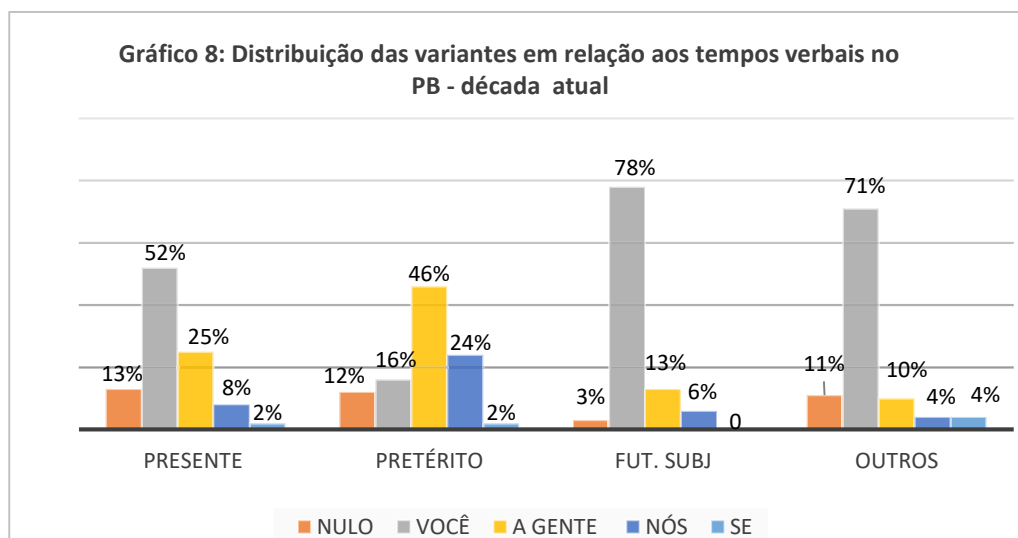
5.6 O tempo verbal

O tempo verbal foi elencado como um possível fator favorecedor das variantes em análise, tendo em vista que a leitura genérica do predicado é em parte atribuída pelo tempo verbal no presente e no futuro do subjuntivo. A explicação está em que predicados genéricos, habituais e hipotéticos associados ao presente apresentam uma leitura genérica, ou quase-universal, podendo resultar em uma situação em que se descreve um fato comum às pessoas em geral, e não a um ser específico. A repetição da mesma situação é que confere à sentença uma leitura genérica, como se pode ver em exemplos como o (170):

(170) Então isso aí acho que me valeu muito e serviu de exemplo para mim, *você_{gen} sempre tem que* incentivar o aluno, né? (AAX -F3 – 13)

Por outro lado, pela propriedade dos pronomes *a gente* e *nós* de, além da referência genérica, referirem-se a pessoas em geral, incluindo o falante mas não o ouvinte, portanto em contextos (-) inclusivos, presumo que o tempo verbal que favorece o seu emprego na atualidade seja o pretérito. Atribuo essa expectativa ao fato de que pronomes indeterminadores (-) inclusivos, portanto mais específicos, são mais empregados em narrativas, tipo textual que favorece construções no tempo passado. Apresento, assim, nos Gráficos 7 e 8 a seguir (Tabela 6 - Apêndice), a frequência das variantes em relação ao tempo verbal da sentença nas duas décadas avaliadas.





Os Gráficos 7 e 8 confirmam que o tempo verbal é um fator relevante como favorecedor das variantes em análise.

Na década atual, as realizações de *você_{gen}* foram mais expressivas com predicados no presente, aparecendo em 52% das sentenças nesse tempo, e no futuro do subjuntivo, totalizando 78% das formas de indeterminação empregadas em sentenças nesse tempo verbal. Dessa forma, conforme previsto, o futuro do subjuntivo, por ser um tempo/ modo verbal tradicionalmente empregado em orações condicionais de caráter hipotético, exemplo (171), foi o tempo que se mostrou mais favorecedor de *você_{gen}*, na posição de sujeito.

(171) E lá eles são muito rigorosos com isso, **se *você_{gen}* for adquirir** um animal...
(AAX – F1- 1)

O emprego de *você_{gen}* na década atual ganha evidência também nos tempos verbais incluídos em “outros”, com realização em 71% dessas ocorrências. Dentre os tempos verbais incluídos em outros, destaca-se o infinitivo. A inserção de *você_{gen}* em construções infinitivas será analisada no item 5.9.

Já na década de 70, o pronome mais favorecido em construções nos tempos verbais do presente e futuro do subjuntivo foi *a gente*, ocorrendo, respectivamente, em 25% e 46% dessas sentenças. Esse resultado corrobora a proposição deste estudo de que ‘*a gente* está cedendo lugar a *você*’. Outro fator relevante para a comprovação da substituição de *a gente* por *você* para referência genérica, em relação ao Gráfico 7, está na constatação de que na década de 70 o pronome *a gente* não foi favorecido nas construções no passado, com realização em apenas 16% delas. Esse emprego é exemplificado pela construção (172):

(172) Mas naquele tempo não tinha televisão, só tinha rádio, **a gente** escutava e acreditava integralmente no locutor.

Por sua vez, os resultados do Gráfico 8 confirmam a previsão de que o pretérito condiciona a realização de *a gente* em face de *você* na atualidade. Enquanto nesse *corpus*, 46% das construções no pretérito têm *a gente* na posição de sujeito, como no exemplo (173), *você_{gen}* é realizado em somente em 16% dessas construções. Juntos, os empregos dos pronomes de 1ª pessoa do plural, *a gente* e *nós*, totalizam 70% das sentenças no pretérito. Embora em percentual baixo, a realização de *você_{gen}* em construções no pretérito são significativas para demonstrar a produtividade do pronome na língua, com ocorrências até mesmo em contextos que não o favorecem, conforme se pode observar no exemplo (174):

(173) É *você_{gen}* tentar ver se a pessoa estava bem ou não naquele setor. **A gente** não fez muitas alterações. (CEFET – M2 – 17)

(174) O que *a gente* percebe é que ele manteve o discurso da reforma, *você_{gen}* teve ali manifestação, mas foi algo mais pacífico. Na sexta-feira sim, *você_{gen}* teve ali uma mobilização, uma estratégia de bloquear as vias.¹⁰¹

A alternância das formas *a gente* e *você_{gen}*, em exemplos como (175), evidencia que o tempo verbal favorece o emprego de um ou de outro pronome de indeterminação:

(175) Ah eu acho que foi um órgão importante porque tem muitas decisões que **antes a gente não tinha** como bater o martelo e **hoje você_{gen} consegue** bater por causa dessa nova legislação do *campus*. (UFV – M1- 30)

Dessa forma, os dados atestam que enunciados episódicos, representados pelo pretérito, não favorecem o emprego do pronome genérico *você*, bem como que a genericidade do predicado é mais forte nas construções nos tempos presente e futuro do subjuntivo. Esse fator pode indicar que este é um mecanismo gramatical que se relaciona à estrutura das sentenças genéricas e episódicas no PB.

¹⁰¹ Exemplo extraído do Programa Em Pauta da Globo News.

5.7 O caráter hipotético da construção

Além do tempo verbal, outros fatores que conferem genericidade ao predicado podem favorecer o emprego da 2ª pessoa genérica. Portanto, o “caráter hipotético da construção” foi elencado como outro possível fator favorecedor desta variante, considerando-se que a marcação genérica do pronome *você* inserido em construções de indeterminação sem *se* é atribuída em parte pela presença de elementos que as tornam hipotéticas, a partir de exemplos como o (144) *Quando você_{gen} passa cartão, você_{gen} acumula pontos*, apresentado em 5.2.

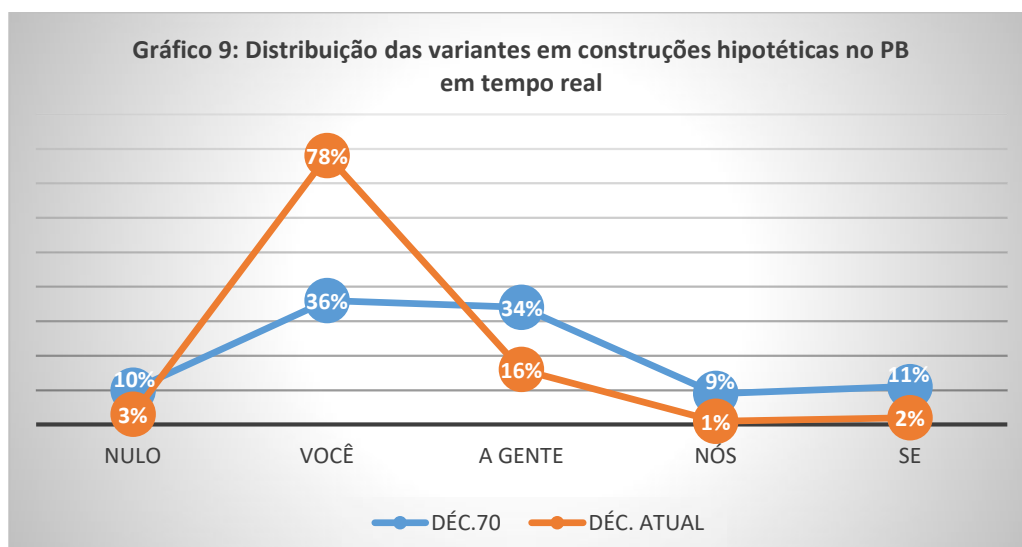
Assim, é possível constatar que, além da construção condicional no tempo-modo *futuro do subjuntivo*, o caráter condicional-hipotético pode ser marcado por outros recursos linguísticos, como as conjunções *se*, *quando*, *quanto mais*, *para*, exemplos (176), (177), (178) e (179), como relacionadas à construção hipotética, por apresentarem uma condição para que algo ocorra. Os resultados do fator “caráter hipotético da construção” estão apresentados no Gráfico 9.

(176) A aula é mais ou menos assim, é um jogo mas **se** *você* vacilar, aí *cê* perdeu, aí *cê* não dá volta mais no respeito, ele vai vir de outras formas. (AAX – F1 – 2)

(177) Então **quando** *você_{gen}* faz uma coisa, uma coisa te toca, não, isso não tá certo... (AAX – F1- 1)

(178) Mas aí **quanto mais** *você_{gen}* trabalha com outras coisas pra *linkar* com isso acho que fica um pouco melhor. (AAX – F1 – 2)

(179) Por que antecipar as situações positivas? **Para que** *você_{gen}* possa usufruir ao máximo desse ponto forte. (UNIPAM - F2 – 28)



Como se pode evidenciar pelo Gráfico 9 (Tabela 6 - Apêndice), o pronome genérico mais produtivo nas construções de caráter hipotético foi *você_{gen}* nas duas décadas, com expressiva gradação de emprego ao longo das décadas, de 36 para 78%. Por seu turno, *a gente*, que na década de 70 teve equivalência de emprego com *você_{gen}* em sentenças hipotéticas, na faixa de 30%, nos dados atuais apresentou declínio de 18% nessas sentenças. Portanto, esse fator também vem validar a hipótese de que ‘*a gente* cede lugar a *você*’ para referência genérica (+) inclusiva. As demais variantes, pronomes *nós*, *se* e *nulo*, não são favorecidos nesses contextos.

Por estes resultados do Gráfico 9 também se constata que no período analisado o pronome nulo é pouco produtivo na presença de elementos que marcam o caráter condicional-hipotético. Consequentemente, os pronomes plenos são mais empregados, e, conforme discutido em 5.2, em construções de nulo arbitrário em que não a inserção de um pronome na posição gramatical de sujeito não é licenciada, na presença de elementos que contribuem para o caráter hipotético da construção, essa inserção é possível. Ou seja, esses contextos hipotéticos favorecem o emprego de pronomes genéricos plenos, mesmo em construções em que a presença do suposto *pro* (nulo arbitrário) bloqueia a inserção de *você_{gen}*, como em (145), repetido aqui.

(145) Num outro cenário no futuro, quem sabe, *você_{gen}* ressuscita esse, *você_{gen}* tira esse projeto da gaveta.

5.8 Presença x ausência de SADVs e SPs locativos e temporais

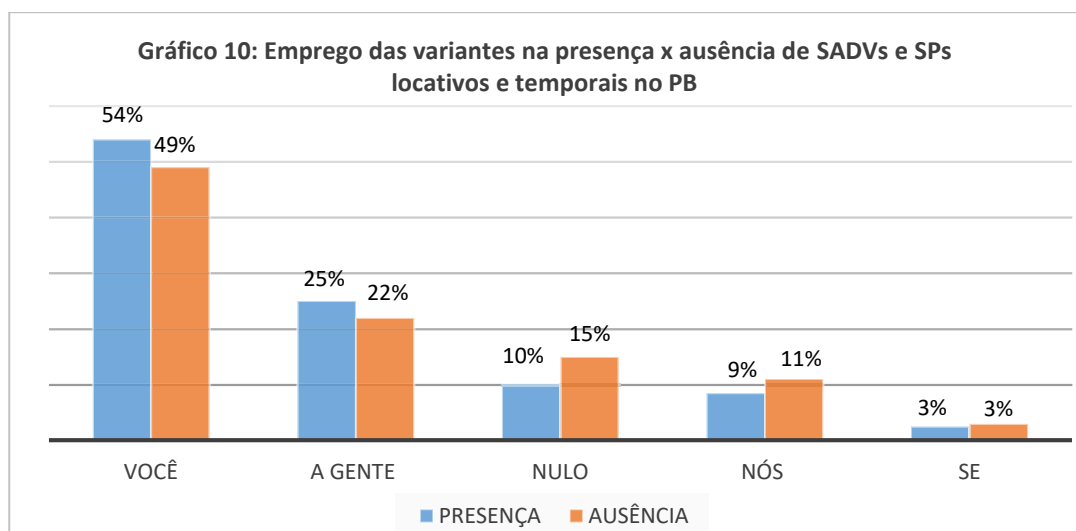
A análise dos contextos em que é empregado o pronome *você_{gen}* mostrou que ele é mais produtivo em situações hipotéticas, conforme mostram os Gráficos 7, 8 e 9, as quais são frequentes em argumentações e exemplificações, normalmente iniciadas por sintagmas adverbiais e preposicionais, assim como SPs locativos e temporais, exemplos de (180) a (183). Por essa razão, o fator ‘Presença de SADVs e SPs locativos e temporais’ foi elencado como outro possível favorecedor da variante inovadora *você_{gen}*, cujos resultados estão apresentados no Gráfico 10.

(180) Então **ali** *você_{gen}* tem uma rodovia, se tira essa rodovia e coloca, sei lá, um tráfego de outros, uma ferrovia passando ali, aí vai demandar um estabelecimento, um ponto de apoio. (AAX – F1 – 7)

(181) Hoje não, **hoje** *você_{gen}* não tem essa proximidade. (AAX – F2- 6)

(182) **Aqui** *você_{gen}* tem bons professores, é da índole deles realmente buscar todo, ah eu estou procurando um curso, estou melhorando isso. (CEFET – M2 – 17)

(183) Então aqui é muito bom por isso, porque **dia de prova** *cê* tem três horas pra estudar. (AAX – F1- 1)



Os resultados do Gráfico 10 comprovam que o pronome de 2ª. pessoa é o mais favorecido na presença de SADVs e SPs locativos e temporais em relação às demais variantes nesses contextos. Por outro lado, se comparados com os empregos na ausência de SADVs e SPs locativos e temporais, percebe-se que não há favorecimento desses elementos para o emprego de *você_{gen}*.

Entretanto, em vez de serem considerados irrelevantes para esta análise, esses resultados podem trazer significativa contribuição, visto que as ocorrências de *ter_{ex}* na ausência de SPs locativos contraria o pressuposto de que a presença do locativo em construções com *ter_{ex}* sem sujeito é que garante a sua boa formação.

Conforme discutido na seção 3.3.3, de acordo com Avelar (2009), sentenças com *ter_{ex}* apresentam restrições que revelam uma conexão entre a presença de um sintagma locativo e a interpretação existencial. Entretanto, exemplos de construções com *ter_{ex}* sem a presença de locativo encontrados no *corpus* atual contrariam essas restrições em relação ao sintagma locativo:

(184) Mas também **Ø tem** o MEC com o seu sistema regulador. (ESAMC–M2–20)

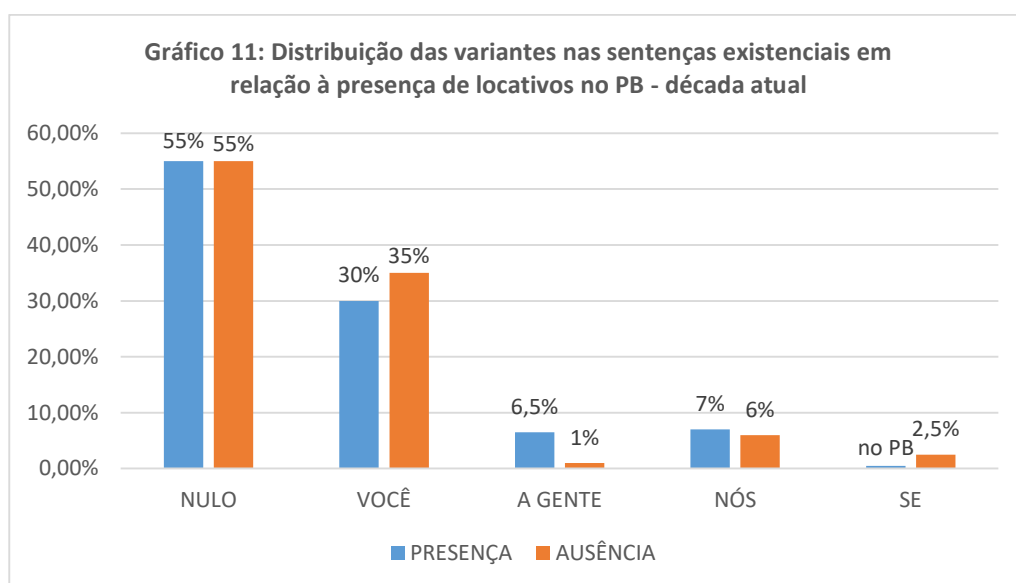
(185) Ninguém quer trabalhar pesado mais porque estão entrando na faculdade. Só que não **Ø tem** emprego para as pessoas que formam. (AAX -F1– 21)

Já com relação às construções de *ter* existencial com *você_{gen}* na posição gramatical de sujeito, Avelar (2009b) propõe um *link* entre o pronome e a locução locativa. A principal razão para esta proposta está no fato de que o pronome *você_{gen}*, de acordo com o autor, é licenciado somente se a sentença existencial apresentar uma ancoragem locativa:

Se esta visão é correta, a presença de *você* não entra em conflito com o caráter impessoal da sentença existencial, uma vez que, como esperado, não há relação temática entre o verbo existencial e o pronome, mas entre o pronome e a categoria predicativa que encabeça o sintagma locativo.¹⁰² (AVELAR, 2009, p. 3)

Nesses contextos, de construções de *ter_{ex}* com *você_{gen}* na posição de sujeito gramatical, também há ocorrências no *corpus* atual do PB de *ter_{ex}* sem a ancoragem de um locativo, como se pode constatar na seção 5.3, por meio dos exemplos (157) e (158).

Tendo em vista a possibilidade dessas ocorrências, as construções existenciais do *corpus* atual foram selecionadas e codificadas quanto à presença ou não de SPs locativos, e posteriormente processadas com cruzamento dos fatores ‘tipo de construção’ e ‘presença x ausência de SPs’, conforme Gráfico 11 a seguir.



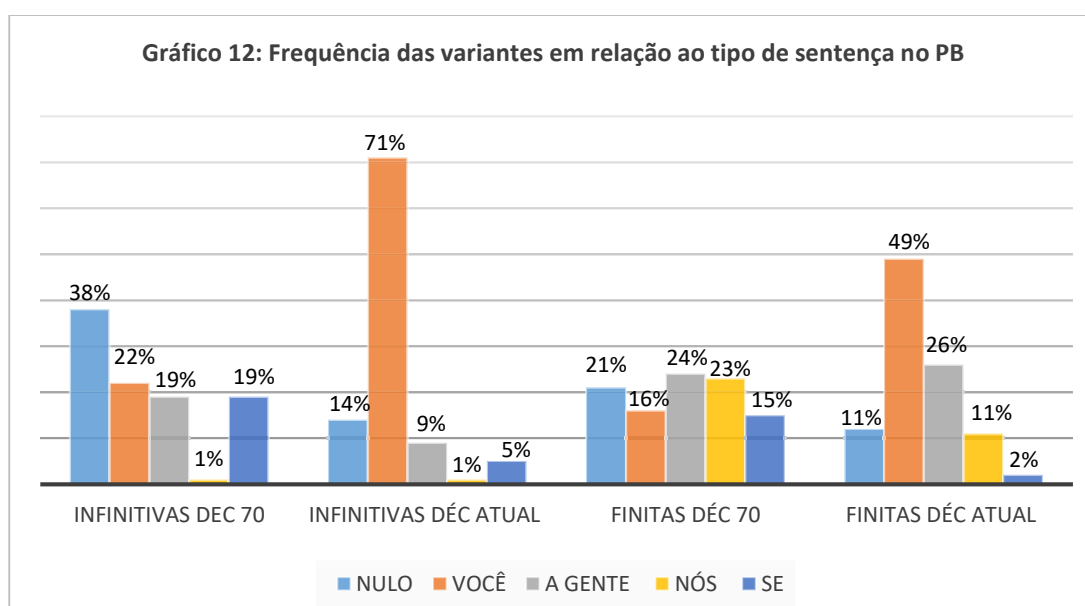
O Gráfico 11 confirma que em termos percentuais as ocorrências de construções existenciais com *ter* sem a presença de SPs locativos são equivalentes às ocorrências com esses

¹⁰² Original: “If this view is correct, the presence of *você* does not conflict with the impersonal character of the existential sentence, given that, as expected, there is no thematic relation between the existential verb and the pronoun, but between the pronoun and the predicative category heading the locative frase”.

sintagmas, contrariando o pressuposto de que tais sentenças só são aceitáveis no PB na presença de um constituinte locativo na sentença (AVELAR, 2009b). Deve-se ressaltar, entretanto, que a codificação em relação à presença de locativos não levou em conta a existência de ancoragem locativa no contexto no qual a sentença se insere.

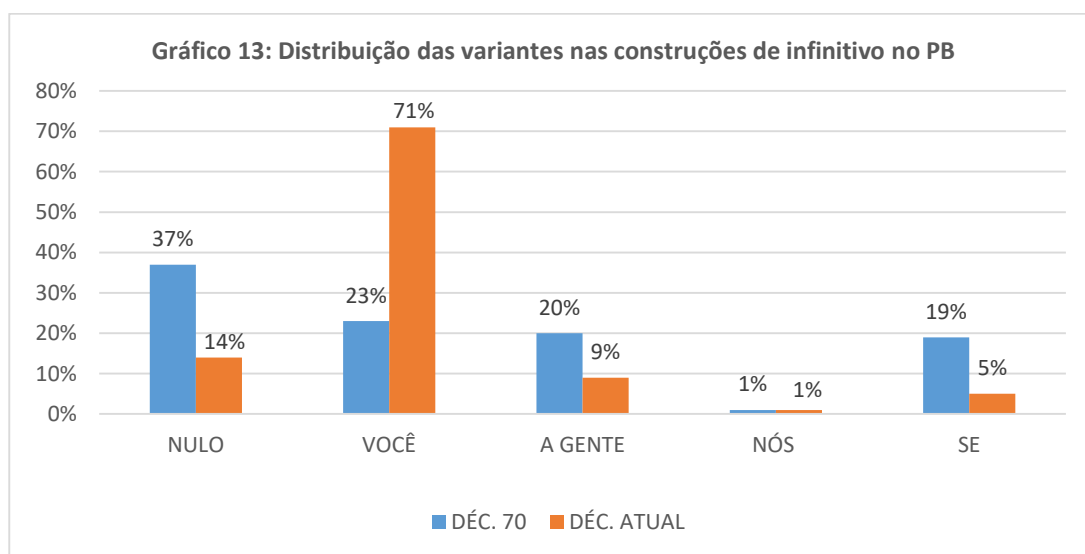
5.9 Construções finitas x infinitivas

Tomando-se por base o trabalho de Duarte, M. E; Kato, M e Barbosa, P. (2003) que aponta que as formas plenas para indeterminação do sujeito apareceram com maior frequência nas sentenças finitas, enquanto o sujeito nulo e o clítico *se* aparecem preferencialmente nas infinitivas, “o tipo de sentença - finita e infinitiva” foi elencado como possível fator linguístico favorecedor das variantes. Os resultados estão apresentados no Gráfico 12 a seguir:



Considerando que a maioria dos *corpora* é constituída por sentenças finitas, o número de ocorrências de todas as variantes em tais sentenças supera em muito nas infinitivas (Tabela 9 – Apêndice). Entretanto, em termos percentuais, pelo expressivo aumento no emprego de *você_{gen}* em sentenças de infinitivo ao longo do período analisado, percebe-se que essa forma verbal favorece seu emprego. 71% das sentenças de infinitivo do *corpus* atual têm *você_{gen}* na posição de sujeito, contra 22% na década de 70. Já nas sentenças finitas, embora *você_{gen}* também lidere entre as variantes de indeterminação, o pronome aparece em percentual menor, 49% das ocorrências de sentenças finitas na década atual.

Como o infinitivo é atribuído genericamente à totalidade das pessoas, e a nenhuma em particular, essa forma, por si só, já exprime a ausência de um sujeito determinado, dispensando, portanto, até mesmo a partícula *se*, que carregaria a mesma informação. Contrariando essa previsão, os dados atestam a perda gradual nesta posição do pronome nulo (de 38 para 14%) e do clítico *se* (de 19 para 5%), com o consequente aumento do preenchimento da posição de sujeito gramatical de sentenças no infinitivo. No Gráfico 13 a seguir é possível visualizar essa tendência.



Ainda com relação ao apagamento do pronome *se*, conforme discutido em Holmberg (2010), na ausência do pronome ou de algum antecedente possível, a interpretação do sujeito nulo de um verbo na terceira pessoa em LSNPs é de indeterminação. Assim, pode-se afirmar que o *se* indeterminado não tem mais razão de ser, ou sob o ponto de vista de Galves (1987), a ausência do traço */+ pronominal/* na concordância no PB faz com que o *se* seja dispensável para expressar a indeterminação em sentenças finitas.

Já nas infinitivas, a situação se inverte, de acordo com Galves (1987) e Nunes (1990), que mostram que parece estar em curso no PB dois processos de mudança antagônicos: enquanto em um processo, o *se* indeterminador é elidido em sentenças finitas, no outro, o *se* é inserido em sentenças no infinitivo.

Conforme mostrado por Galves (1987, 2001), a possibilidade de o tópico ser antecedente do sujeito nulo do infinitivo faz com que o emprego do *se* seja necessário na sintaxe dessas construções no PB, como marcador de indeterminação. Em uma sentença como (186),

extraída de Galves (1987), é possível entender a razão da inserção do *se* antes de infinitivo no PB:

(186) João é difícil de pagar¹⁰³.

Galves (1987) dá a seguinte explicação para a ambiguidade dessa sentença (186) no PB: ela pode significar que João paga dificilmente ou que é difícil pagá-lo. No PE, só a segunda interpretação é possível. Por essa razão é que no PB também é possível dizer:

(187) João é difícil de *se* pagar.

Entretanto, os dados do PB atual não atestam esse fato. O que se evidencia pelo Gráfico 13 é um decréscimo no emprego de *se* indeterminador nas construções de infinitivo no período analisado, de 19 para 5%. Apesar dessa redução no emprego do clítico *se* ao longo do período analisado, na década atual percebe-se maior emprego desse pronome nas construções infinitivas em relação às finitas, com realização respectivamente em 5% e 2% dos dados analisados, conforme Gráfico 12.

Considerando-se a hipótese deste estudo de que o pronome *você* está exercendo no PB função de elemento indeterminador, em competição com o *se* indeterminador, os dados do Gráfico 13 sugerem que a diminuição no emprego de *se* nas construções de infinitivo está sendo compensada pelo emprego de *você*, que aumentou significativamente na década atual. Este resultado consubstancia a proposta deste estudo do percurso da indeterminação no PB: de *se* a *cê*. Os exemplos a seguir do *corpus* atual evidenciam esse emprego:

(188) Porque sem isso não tem condição de *você_{gen}* abrir uma pós-graduação.
(UFV- M2 – 26)

(189) Então, o PDI, Plano de Desenvolvimento Institucional é extremamente importante pra *você_{gen}* traçar o rumo da universidade. (UFV- M2 – 26)

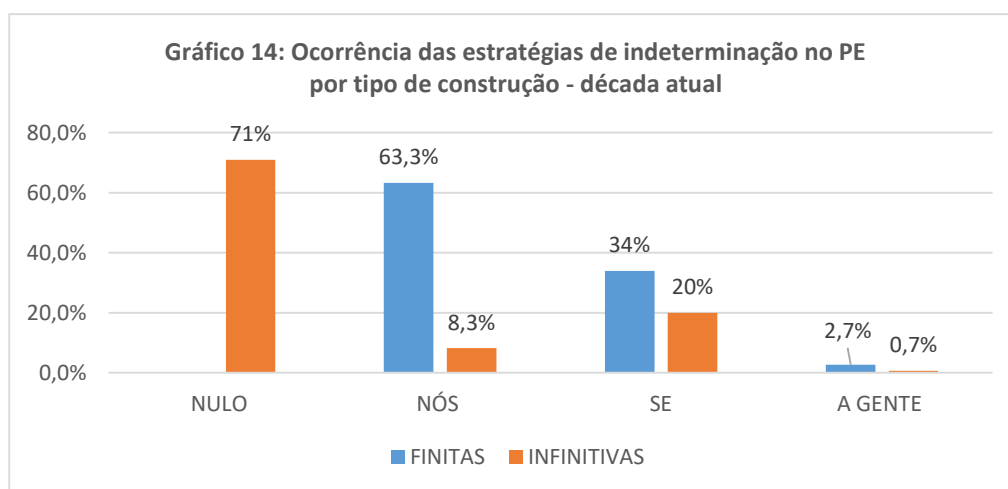
(190) Para *você_{gen}* construir alguma coisa, uma casa, um prédio, qual é a estratégia a se desenvolver para *você_{gen}* poder atingir isso? (UNI - M3 – 27)

Além do pronome *você_{gen}* que não consta da variedade europeia do Português, outra estratégia de indeterminação que revela divergência marcante entre as duas variedades da língua diz respeito ao clítico *se*, que ainda é bastante produtivo em PE. Conforme discutido acima,

¹⁰³ Exemplos (186) e (187) foram extraídos de Galves (2001, p.51).

Galves (1987) mostra que, além da baixa produtividade do clítico *se* no PB, seu emprego ainda se distingue do PE em relação à inserção do *se* indeterminador em sentenças com infinitivo, onde esse emprego não é canônico. Se no PB o *se* estaria sendo inserido nas construções infinitivas para desfazer ambiguidade, em que o sujeito da oração matriz pode ser interpretado como sujeito da oração encaixada, no PE essa ambiguidade é inexistente.

Para averiguação da suposta não ocorrência do clítico *se* em construções de infinitivo no PE foi elaborado o Gráfico 14.



Os dados do Gráfico 14 também revelam uma curiosidade em relação ao *se* vem sendo discutido acerca do emprego do clítico *se* no PE: esse pronome é igualmente produtivo em construções finitas e infinitivas, com realização em respectivamente 31,2% e 20% dessas construções na atualidade. Os exemplos a seguir sinalizam que a inserção do clítico *se* em construções de infinitivo no PE não é para desfazer ambiguidades.

(191) Quer dizer, de *se ver* a alegria da criança ao ver-me por exemplo quer dizer é uma coisa que enfim penso que é verdadeira. (OEI-C- 3H)

(192) Se gosto do Cacém para *se viver*, não é assim não acho muito atraente não é como digamos... (CAC-A-3H)

(193) Não sei se é possível *criar-se* um grupo de psicólogos para professores que acompanhe essas pessoas de uma maneira particular (CAC- C- 3H)

(194) ... que é muito fácil *se aprender* a língua quando *se* é novinho não é? (CAC-C- 3M)

(195) Para *se poder* fazer a tal prevenção diária em casa Ø fazer (higienizações) vulgarmente também denominadas por () aplicações de () aplicações de flúor. (CAC-A-3M)

5.10 Intercambialidade

Na esteira da proposta do percurso de *se* a *cê*, o fator intercambialidade das variantes corrobora a hipótese de concorrência dos pronomes indeterminadores *se* e *você_{gen}*. Embora constitua condição *sine qua non* para a seleção das variantes a serem investigadas que elas sejam intercambiáveis entre si, conforme discutido em 5.4 “Grau de indeterminação das variantes”, foram considerados para esta análise empregos ‘menos inclusivos’ de *a gente* e *nós*, nos quais esses pronomes não apresentam intercambialidade com as demais variantes.

O fator intercambialidade foi proposto, então, para selecionar os empregos em que as variantes têm o mesmo valor de verdade, utilizando-se diagnósticos de substituição. Assim, com raras exceções, praticamente em 100% dos empregos constantes do *corpus* atual do PB, *você_{gen}* é intercambiável com as demais variantes, conforme pode-se evidenciar na Tabela 10.

Tabela 10: Intercambialidade das variantes na década atual

Intercambialidade	Você_{gen}	A gente	Nós	Se	Nulo
sim	1416/2331 61%	405/2331 17,5%	99/2331 4%	67/2331 3%	336/2331 14,5%
não	6/427 1%	257/427 60%	161/427 37%	4/427 1%	5/427 1%

Por esses resultados da Tabela 10, em que o fator intercambialidade se mostrou presente em 99% dos empregos de *você_{gen}*, é possível afirmar que a equivalência dos pronomes *se* e *você* pode ser comprovada pela intercambialidade das variantes. O exemplo (196) a seguir da década atual do PB ilustra a competição das variantes de indeterminação *se* e *você* em uma mesma sentença com mesmo valor de verdade.

(196) ...ou *se* expandia e pegava *campus* fora ou *você_{gen}* não conseguia dinheiro.
(UFV-M1-24)

Atestando a intercambialidade de *você_{gen}* e *se*, no *corpus* atual do PB há exemplos que admitem a inserção somente de um desses pronomes na posição sujeito gramatical de construções tanto pessoais, como em (197), quanto existenciais, (198), sem que a semântica apresentativa de *ter_{ex}* seja alterada.

(197) *Cê* não governa o país sem um congresso.

(198) *Você_{gen}* pode ter algumas coisas interessantes, empresas interessantes na bolsa, *você_{gen}* **tem** empresas que tiveram valorização excelente.

Por outro lado, a não intercambialidade de *nós* e *a gente* com *você_{gen}* pode ser exemplificada pelas construções (199) e (200):

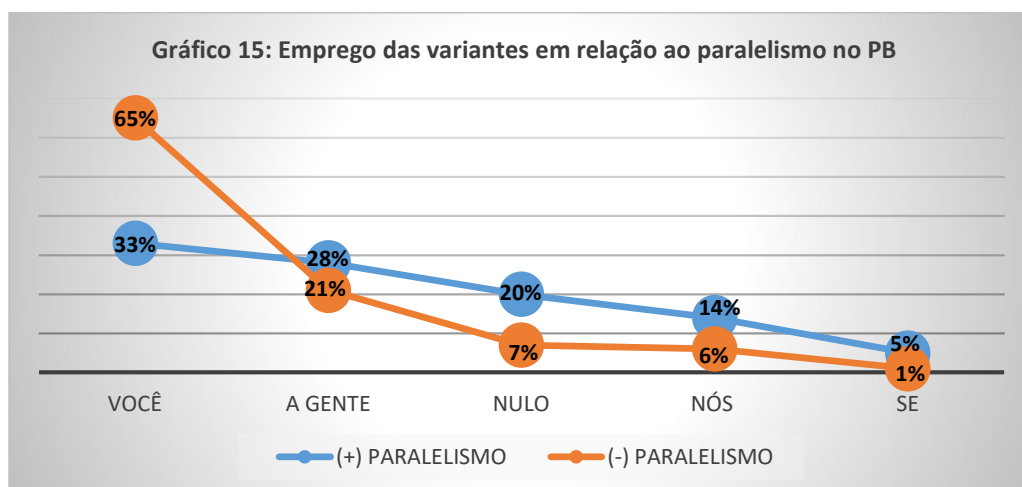
(199) *Nós (*você_{gen})* tivemos várias pessoas que vieram e foram embora, então, além da comunidade da cidade, a comunidade acadêmica. (UFV – M1- 24)

(200) Só que *a gente (*você_{gen})* começou a ter alguns problemas que a Semana do Fazendeiro era uma referência que o professor Baião tinha sucesso que todo ano ocorria. (UFV – M1- 30)

5.11 Paralelismo

Outro fator elencado como possível favorecedor das variantes em análise foi o paralelismo, por se mostrar relevante em diversas pesquisas sobre a alternância das formas pronominais plenas *nós* e *a gente* na posição de sujeito, como na de Omena (2003). Neste trabalho, Omena (2003, p.71) mostra que “há um princípio discursivo já bastante observado nos estudos variacionistas, o de que, no desempenho do falante, uma forma provoca outra semelhante. É o princípio do Paralelismo, que atua fortemente na seleção das variáveis em estudo”.

Os dados comprovam esta previsão, conforme mostrado no Gráfico 15. Como na década de 70 este fator não se mostrou relevante (cf. Tabela 11–Apêndice), somente a década atual foi considerada para elaboração deste gráfico.



Pelo Gráfico 15, evidencia-se que a variante *você_{gen}* é especialmente favorecida na atualidade pela tendência de repetição da mesma forma numa sequência discursiva, com 65% das sentenças com pronomes indeterminados em situação de paralelismo. A atuação do paralelismo se verifica igualmente em relação ao pronome nulo. Já em relação às variantes *a gente* e *nós*, o paralelismo não se mostrou relevante. Os exemplos a seguir extraídos do *corpus* atual do PB sinalizam o efeito desse fator discursivo no emprego das variantes.

(201) Essa coisa do capitalismo de *Ø* ter que trabalhar, de não *Ø* poder prover a casa, óbvio todo mundo precisa, mas valores como *cê* sentar, *cê* tentar ouvir, *cê* ensinar uma coisa nova, *cê* aprender uma coisa nova também. (UFV-M1-24)

(202) Então para o ano que vem *a gente* mudou e *a gente* já levou isso pra uma reunião de diretoria. (CEFET – M3 – 18)

5.12 As forças sociais exercidas sobre as formas em variação

Esta pesquisa parte do entendimento de que o estudo do mecanismo da mudança linguística não terá avanços significativos se não levar em conta os fatores sociais que motivam sua evolução. Assim, a questão do encaixamento linguístico da mudança deve considerar também a matriz do comportamento social da comunidade de fala em investigação como correlacionada com mudanças sociais. O principal caminho, segundo Labov (2008, 193), está na descoberta das correlações entre elementos do sistema linguístico e do sistema não linguístico de comportamento social. O pesquisador possui meios de analisar o comportamento linguístico e eleger os aspectos do contexto social da língua que se vinculam mais estritamente com a variação linguística.

Assim, a estratificação social desta pesquisa foi obtida pela combinação de nível de escolaridade, gênero e idade. A escolaridade de nível superior foi preestabelecida tendo em vista, além da regularidade com o *corpus* do Projeto NURC que é constituído por falantes da norma urbana culta, a observação preliminar da variável inovadora no vernáculo de falantes graduados, especialmente em discursos argumentativos, tipo textual em que predominam as generalizações (palestras, aulas, entrevistas).

Ao se atribuir motivação social para a variação linguística deve-se correlacionar o comportamento linguístico com a medição do *status* atribuído ou adquirido pelos falantes. Se educação se trata de um *status* adquirido, os falantes de nível de escolaridade superior são considerados de prestígio, por conseguinte, atribui-se significado social às variações por eles

empregadas. É dessa forma que as mudanças na língua podem estar relacionadas às mudanças dos subgrupos com os quais o falante se identifica.

De acordo com Labov (2008, p. 281), “na fala monitorada, as mulheres usam menos formas estigmatizadas do que os homens e são mais sensíveis do que os homens ao padrão de prestígio, ou valores sociolinguísticos explícitos”. Por serem mais sensíveis às formas de prestígio, elas avançam mais rápido no curso da mudança linguística, embora não se possa afirmar que elas sempre a liderem. Em vários estudos apontados por Labov (2008), as mulheres são retratadas como iniciadoras de mudanças linguísticas. Vejamos o que dizem os dados da Tabela 13 a esse respeito.

Tabela 13: Estratificação das variantes em relação a gênero no PB

Realização x gênero		<i>Você_{gen}</i>	<i>A gente</i>	<i>Nós</i>	<i>Se</i>	<i>Nulo</i>
Década de 70 (2137)	Feminino	139 12%	324 29%	288 26%	137 12%	239 21%
	Masculino	212 21%	187 18%	148 15%	201 20%	262 26%
Década atual (2758)	Feminino	737 70%	160 15%	83 8%	13 1%	66 6%
	Masculino	688 40,5%	502 29,5%	177 10,5%	58 3,5%	274 16%

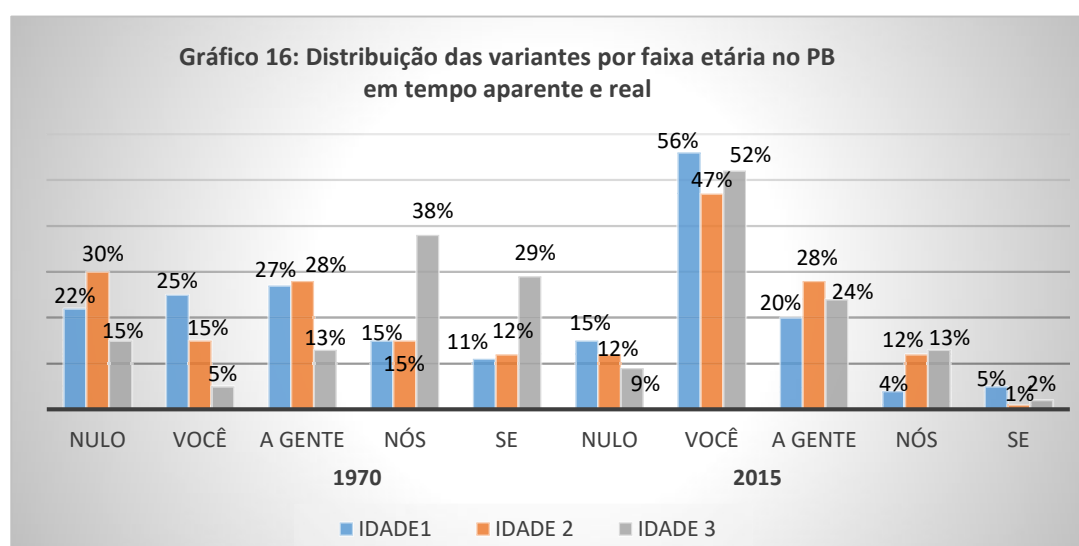
Os dados da Tabela 13 confirmam a hipótese de Labov (2008). Na década de 70, as mulheres usavam menos a forma inovadora *você_{gen}* do que os homens, que certamente não era a de prestígio, e mais a forma canônica de indeterminação *a gente*, num percentual de 29% das sentenças indeterminadas com este pronome, contra 12% com *você_{gen}*.

Diante do gradual acréscimo de emprego de *você_{gen}* na década atual por graduados universitários, considerados falantes de prestígio (Labov, 2008, p.168), é possível inferir que a variante inovadora *você_{gen}* é considerada uma forma de prestígio¹⁰⁴. Daí o avanço das mulheres na trilha da forma de prestígio. Enquanto o uso de *você_{gen}* está mais avançado entre as mulheres, aparecendo em 70% das sentenças indeterminadas proferidas por elas, esse percentual cai para 40% de ocorrência nas mesmas sentenças proferidas por homens, mostrando que as mulheres de fato são mais influenciadas do que os homens no emprego de formas de prestígio.

¹⁰⁴ Labov (2008, p. 168), em estudo sobre a introdução do *-r* em Nova York, correlaciona idade/ classe social à capacidade de usar formas prestigiadas, e mostra que “os graduados universitários alcançaram certo grau de segurança no seu uso do Inglês, em parte graças ao intenso contato com falantes de prestígio”.

No caminho oposto, a variante *a gente* aparece em apenas 15% das falas femininas e em 29% das masculinas. Se se considera o prognóstico de que as mulheres usam menos as formas estigmatizadas, pode-se supor que o pronome *a gente* como indeterminador está se tornado a forma estigmatizada.

Com relação ao outro fator extralinguístico elencado para seleção da amostra, faixa etária, a variável inovadora não se mostrou sensível à idade dos falantes na década atual. O Gráfico 16 exhibe o desenvolvimento das variantes em tempo aparente e real.



Surpreendentemente, enquanto nos anos 70 poder-se-ia falar em mudança em andamento em função dos grupos etários, em que a variante inovadora *você_{gen}* exhibe um aumento regular em faixas etárias sucessivas, dos mais velhos para os mais novos, sendo que os mais velhos se mostram mais conservadores, na década atual não se registra essa tendência. As variantes exibem distribuição uniforme entre as faixas etárias.

Se no fenômeno variável analisado, *você_{gen}* caracteriza-se como um novo padrão de prestígio, identificado pela escolaridade dos informantes da amostra, graduados universitários, portanto, considerados falantes cultos que exibem maior segurança linguística, era de se esperar maior percentual de emprego entre os mais jovens.

No entanto, a distribuição das variantes por faixa etária não mostra correlação da variável inovadora com a idade. Esta não tendência confronta-se com outras comprovações da teoria laboviana de que o uso da variante de prestígio aumenta entre as pessoas mais jovens (Labov, 2008, p. 79). Ou seja, a distribuição por idade não mostra nenhuma comprovação de mudança em andamento em relação à variável inovadora.

Deve-se considerar ainda com Labov (1998, p. 336) que “nem toda mudança linguística está vinculada a um grupo social particular”. Para exemplificar essa hipótese, Labov mostra a fusão do *o* breve aberto e do *o* longo aberto em algumas áreas dos Estados Unidos que parece afetar todas as pessoas, predizendo ser este um caso de mudança avançada.

Dessa forma, ainda que não se possa falar que a variação em análise caracteriza mudança em andamento, sob a perspectiva da mudança aparente, é possível depreender que, semelhantemente ao fenômeno descrito por Labov (1998), parece tratar-se de uma mudança avançada, que é quando não se pode ver o padrão em pirâmide através das diferentes faixas etárias. Se esse entendimento estiver no caminho correto, afirmo que nesse estágio em que se encontra a variação em análise, os maiores determinantes sociais são gênero e escolaridade (a qual pode pressupor o tipo textual em decorrência do tipo discursivo predominante na fala), e o fator idade desempenha um papel secundário.

O fenômeno detectado da não correlação da variante inovadora com idade nos dados atuais pode ser explicado, ainda, pelos seguintes pressupostos labovianos:

À medida que o traço linguístico se desenvolve dentro do grupo de origem ele se generaliza. Com o curso do tempo (três ou quatro décadas), isso pode desenvolver um espectro mais amplo de subclasses condicionadas, bem como ambientes mais extremos (menos favorecidos). Além disso, a simetria estrutural do sistema leva a generalização a membros da mesma classe natural, que começam a se mover na mesma direção (ou na oposta). (LABOV, 2008, p. 366).

Assim, a surpreendente regularidade de emprego da variante inovadora *você_{gen}* entre os falantes das três faixas etárias pode sinalizar uma menor diversidade linguística entre os subgrupos da comunidade de fala – falantes da norma culta, que, de fato, compartilham um conjunto comum de normas linguísticas.

Entretanto, é a combinação das observações em tempo aparente e real que constitui o método fundamental de análise de mudanças em curso (LABOV, 1994). Assim, se se considera o tratamento dos dados em tempo real (cinco décadas), os resultados sinalizam um notável avanço da variável *você_{gen}* entre os grupos etários, o que pode indicar uma mudança correspondente na estrutura da comunidade de fala no período. Proponho, portanto, que o estudo de tendência permite depreender a direcionalidade da variação como em curso na comunidade linguística, e que as mudanças na configuração social se refletem na estabilização do processo de mudança, uma vez que se nota estabilidade e semelhança no comportamento dos diferentes grupos etários em tempo aparente.

A distribuição em tempo aparente mostra também que as duas variantes de 1ª. pessoa plural, *nós* e *a gente*, ainda estão sendo realizadas entre os falantes das três faixas etárias, com

frequência equivalente à de cinco décadas atrás, portanto, a comunidade se mantém estável em relação a essas variantes, não se processando qualquer tipo de avanço com uma exceção. Com maior decréscimo de emprego, a variante *nós* na idade 3, passa de um de percentual de emprego de 38% nos anos 70 para 12,7% na atualidade, evidenciando uma tendência da indeterminação no PB. Semelhantemente, a variante *se* apresenta maior redução de emprego entre os falantes mais velhos, de 25% em 70 para 2% na década atual.

Portanto, a partir das características estruturais apresentadas nesta análise, considero, com base em Tarallo (1994, p. 26), que o sistema, quase que preditivamente, caminhou na direção de *você_{gen}* e não na direção *a gente* ou *se* para referência genérica no PB.

5.13 O percurso da indeterminação no PB

A partir das discussões apresentadas e da análise dos dados que integram esta pesquisa, considero que *você/ cê* genérico é empregado no PB tanto em construções existenciais com *ter* como nas demais construções predicativas como elemento indeterminador com características semelhantes ao clítico indeterminador *se*, o que pode ser evidenciado pela equivalência entre as construções com clítico *se*, com o pronome nulo e com o pronome *você_{gen}/cê* e na posição de sujeito, sugerindo o percurso dessas estratégias de indeterminação no PB, conforme sequência proposta a seguir:

- i) E depois, se o projeto foi aprovado para construir num lugar, não *se* pode trazer para outro.
- ii) E depois, se o projeto foi aprovado para construir num lugar, não *Ø* pode trazer para outro.
- iii) E depois, se o projeto foi aprovado para construir num lugar, *você_{gen}* não pode trazer para outro. (UFV - M1 – 23)

A análise dessas construções a partir da sequência proposta permite-me inferir que a supressão do clítico *se* em construções de indeterminação deu origem a construções de nulo genérico, licenciadas no PB em função do traço de (in)definitude de LSNPs (Holmberg, 2010) que faculta interpretação indefinida ao pronome nulo de 3^a. pessoa, e que, na direção da perda de licenciamento de sujeitos nulos no PB, essa posição vazia pela supressão do clítico *se* passa a ser preenchida por um pronome foneticamente realizado com leitura genérica - *você_{gen}*, o qual emerge na língua com o surgimento do paradigma de pronomes fracos.

Deve-se ressaltar aqui que, nesse período de transição da língua, de acordo com a teoria de Dobrovie-Sorin (1998, apud Duarte, M.E; Kato, M.A.; Barbosa, P., 2003), o *se* nominativo e o acusativo podem estar sendo afetados de forma diferente. Assim, nos casos de *se* nominativo, este é considerado um clítico sujeito, argumental, o qual não pode ser omitido e, em uma língua em mudança como o PB, tende a ser substituído por outra forma pronominal, como *a gente*, *você*, entre outros. Os exemplos abaixo, apresentados pelas autoras, ilustram a impossibilidade da omissão do *se* nominativo:

(203) *É feliz aqui.

(204) *É frequentemente traído por falsos amigos.

Com manutenção de sentido indeterminado, *você_{gen}* apresenta-se como uma forma alternativa que vem substituir o pronome impessoal *se*:

(203) a. *Você_{gen}* é feliz aqui.

(204) a. *Você_{gen}* é frequentemente traído por falsos amigos.

Portanto, a etapa (ii) da sequência acima proposta não contempla todas as construções de indeterminação no PB, visto que em contextos de *se* nominativo não pode haver omissão do pronome nominativo. Nessas construções, considerando-se a supressão do clítico *se* e a sua substituição em contextos de impossível apagamento, e os dados desta pesquisa, atesto que o pronome *você* está exercendo a função de elemento indeterminador em substituição ao *se*. A construção (205) extraída do *corpus* do PB atual exemplifica sentenças de *se* nominativo que não licenciam uma categoria vazia (nulo genérico) na posição de sujeito.

(205) Então acaba que dentro dos institutos *você_{gen}* trabalha tanto as propostas, quanto as discussões. (UFV – M1- 30)

Já a perda do *se* passivo não obriga necessariamente o preenchimento da posição do sujeito, visto que esta é uma posição não temática e na reanálise da passiva pronominal, o expletivo da posição de sujeito foi tomado como um pronome nulo referencial - *pro* (NUNES, 1990).

Com respaldo nesta proposta de Nunes (1990), a expectativa desta pesquisa era de que construções de *se* passivo sem concordância e sem o *se* não licenciariam a inserção de um pronome genérico na posição estrutural de sujeito, o que foi invalidado pelos dados, tendo em

vista que o emprego de alguns recursos linguísticos que tornam o predicado genérico propiciam a sua inserção.

Em relação às construções existenciais com *ter*, a proposta de implementação no PB de um elemento indeterminador com função similar ao clítico *se*, um pronome fraco nos termos de Kato (1999), na posição de sujeito gramatical, sustenta a proposição em 3.3.3, de que ao ser empregado em construções com sentido existencial, o verbo *ter* não passa propriamente por um processo de impessoalização, mas de detematização.

Considerando, portanto, a proposta de Avelar (2009a) de que a inovação no PB não envolve exatamente uma reanálise de *ter* como um verbo existencial, mas o uso da estrutura possessiva para expressar sentido existencial, argumento que construções como (206),

- (206) O Cefet é o seguinte, como *você_{gen}* **tem** vários *campi*, *você_{gen}* precisa de dar uma unidade para isso. (CEFET – M2 – 17).

revelam uma posição de sujeito, mas mantendo a função apresentativa de verbo existencial com interpretação indeterminada.

Nessa perspectiva, pode-se considerar que a generalização do emprego do verbo *ter* no PB tanto para predicados possessivos como existenciais – *ter* com sentido existencial – reflete uma solução da língua, facultada pela estrutura possessiva, para compensar a perda da propriedade de licenciamento do sujeito nulo, no caso não referencial, pela possibilidade de inserção de um sintagma pronominal na posição de sujeito, sem atribuição de papel temático, não comprometendo, dessa forma, o caráter apresentativo do verbo existencial.

Constitui evidência a favor da proposta de aproximação das propriedades sintáticas de construções de indeterminação e das existenciais, o paralelismo sintático e interpretativo de ambas construções – de nulo expletivo e de nulo genérico – com o clítico *se* e com *você_{gen}* na posição de sujeito, como nos exemplos (83) e (84), aqui repetidos em (207) e (208):

- (207) *Ø_{expl}* Não *tem* mais burca na síria.
a. Não *se* *tem* mais burca na síria.
b. *Você_{gen}* não *tem* mais burca na síria.¹⁰⁵
- (208) Não *Ø_{gen}* *usa* mais burca na síria.
a. Não *se* *usa* mais burca na síria.
b. *Você_{gen}* não *usa* mais burca na síria.

⁹⁹ Exemplo extraído do Programa *Em Pauta* da Globo News, em 26/02/2016.

Diante desses fatos, parece estar em curso no PB três processos concomitantes em relação ao licenciamento do pronome pleno *você* para marcação genérica. De um lado, tem-se a substituição pronominal: a) de *se* por *você*, em contextos de *se* nominativo nos quais não pode haver omissão do clítico, exemplo (205); b) de nulo expletivo por *você* em construções existenciais, exemplo (206); c) de *a gente* por *você*, em contextos de indeterminação (+) inclusiva (208)b. De outro lado, tem-se inserção do pronome *você_{gen}* na posição de sujeito nulo genérico, esvaziada pela supressão do clítico *se*. Nesse grupo estão as construções de indeterminação com nulo genérico, nas quais a inserção do pronome não é necessária para a gramaticalidade da sentença, como no exemplo (209) a seguir. Por último, tem-se as construções de *se* passivo sem concordância e sem o *se*, interpretadas como indeterminadas, em que a inserção do pronome *você_{gen}* só é licenciada na presença de elementos que tornem o predicado habitual ou hipotético (211); do contrário, a inserção do pronome na posição de sujeito é bloqueada e a marcação do sujeito é de nulo arbitrário, como em (210). Incluem nesse grupo de construções de *se* passivo aquelas construções em que a presença de expressões como “não mais”, exemplo (211), dentre outras como em (144), tornam o predicado habitual, usual, atribuindo-lhe uma leitura genérica, e, por isso, permitindo a inserção de um pronome genérico na posição de sujeito.

(209) Para resolver os problemas do Brasil hoje ***você_{gen}*** precisa de investimentos.¹⁰⁶

(210) * ***Você_{gen}*** amola faca.

(211) ***Você_{gen}*** **não** amola **mais** faca nos dias de hoje.

Em vista dos fatos apresentados, a propriedade indeterminadora do pronome *você_{gen}*, é evidenciada tanto pela possibilidade de sua inserção em construções de nulo genérico; de substituição ao nulo expletivo em construções de *ter_{ex}* sem perda da semântica existencial, e ao pronome *a gente* por perda de sua propriedade de genericidade; bem como pela não restrição de substituição ao clítico *se* em contextos de *se* nominativo, nos quais não pode haver omissão do pronome nominativo, confirmada pela intercambialidade entre as variantes. Comprovo, assim, pelos processos de inserção e substituição pronominal, o percurso de *se* a *(vo)cê*.

¹⁰⁶ Exemplo extraído do Programa *Em Pauta* da Globo News, em 06/02/2016.

6. CONCLUSÕES

As discussões propostas nesta tese visaram correlacionar a emergência no PB de um pronome foneticamente realizado para marcação genérica – *você*, e para ocupar a posição de sujeito gramatical de construções de *ter_{ex}*, a outros fenômenos sintáticos em processo de mudança na língua, tais como: o apagamento do clítico *se*, o surgimento de um paradigma pronominal fraco, a perda de genericidade do pronome *a gente*, e a perda da propriedade da língua de licenciar sentenças em que a posição estrutural de sujeito aparece vazia, manifestando o encaixamento da mudança de marcação paramétrica que se processa na gramática do PB.

A tese mostrou que concorrendo com os pronomes nulo não referencial e genérico, emerge na língua a forma pronominal de 2ª pessoa *você* com referência genérica, a qual pode ser vista como a solução encontrada pelo sistema para que se evite a posição de sujeito vazia, com a perda de clíticos no PB e da propriedade de LSN de licenciamento do sujeito nulo, que aponta para a perda do nulo não referencial e do nulo genérico. Como o pronome nulo genérico constitui uma das especificidades que caracterizam o PB como uma LSNP, a perda dessa propriedade associada ao preenchimento da posição estrutural do sujeito de construções existenciais com *ter*, sinalizam que o PB caminha na direção de se tornar uma LSNN.

As análises mostraram que a implementação da variante inovadora *você_{gen}* no PB apresenta, a princípio, duas motivações, sendo uma por inserção e outra por substituição pronominal. A inserção do pronome *você_{gen}* em construções de indeterminação, em que o nulo genérico é incorporado em SpecvP, e SpecTP não é projetado, atenderia à satisfação do princípio EPP. Já em construções existenciais com *ter*, a substituição do pronome *Ø_{exp}* por *você_{gen}* se deve à tendência dos falantes no PB a inserir elementos lexicalizados na posição de sujeito gramatical, evitando a ocorrência do verbo em posição inicial, de acordo com Duarte e Kato (2003), e não à projeção de SpecTP, e, ainda, em atendimento ao parâmetro ϕ -independente de EPP. Por fim, tem-se a substituição de *se* por *você* em construções de *se* nominativo, em que a omissão do pronome nominativo não é permitida, evidenciando o percurso da indeterminação no PB: de *se* a *(vo)cê*.

A análise dos dados buscou evidências empíricas para a emergência no PB de um pronome com função indeterminadora semelhante ao *se* indeterminador, que vem preencher a posição dos pronomes *se*, esvaziada pela supressão de clíticos na língua, e do nulo genérico, além de substituir os pronomes nulo expletivo e *a gente*, que perde a propriedade de marcação genérica ao ser empregado em contextos mais específicos. Os dados confirmaram que *você* é o

pronome mais frequentemente empregado na atualidade para marcação genérica. Confirmaram ainda a hipótese da função indeterminadora de *você_{gen}*, o fato de que os percentuais de emprego de *você_{gen}* aumentaram paralelamente ao apagamento do clítico *se*, e à diminuição de emprego do pronome *a gente* em contextos (+) inclusivos, e dos pronomes nulo genérico e nulo expletivo.

Entretanto, as análises atestam a manutenção de um tipo de sentença em que a posição estrutural de sujeito permanece vazia. Pode-se afirmar, assim, que, apesar das mudanças processadas no PB no processo de remarcação do PSN, em que o PB deixa de licenciar o sujeito nulo de referência definida, o sistema caminha para a perda do nulo de referência genérica, mas continua a licenciar sentenças com nulo expletivo: as existenciais, o que constitui uma propriedade de LSNs.

Sobre a questão dos condicionamentos linguísticos que concorrem para a implementação de *você* no PB, os dados apontam o tempo verbal no presente e no futuro do subjuntivo, construções de infinitivo e o caráter hipotético da construção, além do fator intercambialidade que foi determinante para comprovar a hipótese de que *você_{gen}* ocupa no PB a função de elemento indeterminador em substituição ao clítico *se*, que está sendo suprimido da língua falada.

Além dessas comprovações quanto à maior frequência da variante inovadora *você_{gen}* para marcação genérica e de sua substituição aos pronomes *Ø_{gen}*, *Ø_{exp}*, *a gente* e *se*, outro ponto relevante deste trabalho em relação ao licenciamento do pronome *você_{gen}* diz respeito à proposição de que o interior de INFL é o lócus onde o pronome é inicialmente inserido antes de ser movido para a posição de sujeito gramatical das construções existenciais, tendo em vista a não relação temática do pronome com o verbo, e que Spec-T/Infl é a posição naturalmente destinada ao sujeito gramatical em línguas como o Português (Avelar, 2009).

Um fator surpreendente das análises diz respeito à não correlação da variante inovadora com a faixa etária mais jovem, que, de acordo com a teoria laboviana, emprega mais a variante de prestígio (Labov, 2008, p. 79), o que pressupõe não se tratar de uma mudança em andamento. Entretanto, sob a perspectiva da mudança aparente, parece tratar-se de uma mudança avançada, que é quando não se pode ver o padrão em pirâmide através das diferentes faixas etárias.

Já o fator social gênero comprovou que a variante inovadora, por ser mais empregada pelas mulheres, pode ser considerada uma forma de prestígio, tendo em vista que de acordo com a teoria laboviana elas são mais sensíveis às formas de prestígio, e, portanto, avançam mais rápido no curso da mudança linguística. Corrobora para a atribuição de um padrão de prestígio

à variante inovadora, o fato de que os informantes da amostra possuem formação superior, sendo, portanto, considerados falantes da norma urbana culta. Dessa forma, o estudo comprova que as mulheres tendem a liderar os processos não estigmatizados de mudança linguística.

Diante disso, é possível constatar nesse processo de propagação de *você_{gen}* na língua falada do PB, a atuação de dois importantes fatores: a difusão crescente do pronome *você_{gen}* não só entre jovens, mas também em todas as faixas etárias nos últimos 50 anos, e o fato de não haver estigma associado ao uso da forma inovadora no desempenho oral dos falantes, mesmo entre os considerados cultos.

No PE, por outro lado, a amostra analisada comprovou a inexistência de *ter* em contextos existenciais, bem como da estratégia de indeterminação com o pronome *você*, indicando que essas possibilidades não parecem constar da gramática dos falantes considerados cultos do PE.

Por fim, é possível pressupor, pelos resultados da pesquisa apresentados nesta tese, que o PB caminha para se tornar uma LSNN, mas ainda mantém traços de LSN. É evidente que esse processo que pode levar à perda das propriedades de sujeito nulo tem sido gradual e constitui parte de um extenso conjunto de mudanças sofridas pelo sistema pronominal do PB, considerando-se, ainda, que uma mudança paramétrica implica um longo período de variação antes que uma nova opção seja estabelecida.

Deve-se ressaltar ainda que, como o tema desta pesquisa é bastante amplo e emergente, ele não se esgota com esta tese. Novas perspectivas surgirão provavelmente, ou mesmo esta poderá ser ampliada como agenda futura de trabalho, tendo em vista questões que são despertadas pelos resultados da pesquisa, como, por exemplo, o que pode estar concorrendo para a permanência do nulo expletivo nas construções de *ter_{ex}* no PB.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- ADAMS, M. From Old French to the theory of prodrop. *Natural Language and Linguistic Theory*, 5, 1987, p. 1-32.
- ALMEIDA, G. *Resgatando a contribuição da sociolinguística laboviana*. D.E.L.T.A., São Paulo, 1989, v.5, n.1, p.71-99.
- AVELAR, J. O. *Dinâmicas morfossintáticas com ter, ser e estar em Português Brasileiro*. Dissertação de Mestrado. UNICAMP, 2004.
- _____. On the emergence of “ter” as an existential verb in Brazilian Portuguese. In: CRISMA, P. & LONGOBARDI, G. *Historical syntax and linguistic theory*. Oxford: Oxford University Press, 2009a.
- _____. On the status of the (supposed) expletive in Brazilian Portuguese existential clauses. In: Danièle Tork & Leo Wetzels (eds.). *Romance Language and Linguistic Theory 2006*. Amsterdam: John Benjamins, 2009b, p. 17-32.
- AVELAR, J. O. & CALLOU, D. *Sobre ter e haver em construções existenciais: variação e mudança no português do Brasil*. Gragoatá, 2000.
- _____. Sobre a emergência do verbo possessivo em contextos existenciais no português brasileiro. In: CASTILHO, A. et al. (Org.) *Descrição, história e aquisição do português brasileiro*. São Paulo: Editora Pontes, 2007, p. 375-402.
- _____. Sentenças existenciais e preenchimento de sujeito: indícios de mudança em progresso na fala culta carioca. In: SILVA, A.; TORRES, A.; GONÇALVES, M. (Org.). *Línguas Pluricêntricas - Variação Linguística e Dimensões Sociocognitivas*. Braga: Aletheia, 2011, p. 287-300.
- AVELAR, J. O. & GALVES, C. Tópico e concordância em português brasileiro e português europeu. In: COSTA, A.; FALÉ, I. & BARBOSA, P. *Textos Seleccionados –XXVI Encontro Nacional da Associação Portuguesa de Linguística*. Lisboa: Associação Portuguesa de Linguística, APL, 2011, p. 49-65
- BERLINCK, R.; DUARTE, E.; OLIVEIRA, M. Predicação. KATO, M.; NASCIMENTO, M. *Gramática do português culto falado no Brasil*. São Paulo: Editora da Unicamp, 2009.
- BIBERAWER, Theresa. Semi null-subject languages, expelives and expletive *pro* reconsidered. In: BIBERAUER, T.; HOLMBERG, A; ROBERTS, I. *Parametric Variation: Null Subjects in Minamalist Theory*. Cambridge University Press: Cambridge, UK, 2010. p. 153-199.
- BLONDEAU, H. The dynamics of pronouns in the Québec languages in contact. In: MEYERHOFF, M.; NAGY, N. *Social Lives in Language – sociolinguistics and multilingual speech communities*. V 24. John Benjamin Publishing Company: Amsterdam, Philadelphia, 2008, p. 260.
- BORGES NETO, J. A incomensurabilidade e a “compatibilização” de teorias. In: *Ensaio de filosofia da linguística*. São Paulo: Parábola Editorial, 2004.
- BURZIO, L. *Italian Syntax*. Dordrecht. D Reidel Publishin Company, 1986.

CALLOU, D. e MARTINS, L. *Mudança em tempo aparente e tempo real: construções ter e haver existenciais*. Anais do 5º Encontro da Celsul, Curitiba – PR, 2003.

CAVALCANTE, S. R. *O sujeito nulo de referência indeterminada na fala culta carioca*. Diadorim (Rio de Janeiro), v.2, 2007.

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

_____. A minimalism program for linguist theory. In: K. Hale & S.J. Keyser (eds). *The View from Building 20. Essays in Linguistic in Honor of Sylvain Bromberger*. The MIT Press; Cambridge, Mass, 1993, 1-52.

_____. *The minimalist program*. Cambridge, MA: The MIT Press, 1995.

COVENEY, A. Anything you can do, *tú* can do better: *tú* and *vous* as substitutes for indefinite ‘on’ in French. In: *Journal of Sociolinguistics*, 7/ 2. Blackwell Publishing, 2003.

DUARTE, F. B. Distribuição de pronomes fortes, fracos e afixos em línguas de sujeito nulo. *Revista do GEL*, S. J. do Rio Preto, v. 5, n. 1, p. 9-30, 2008.

DUARTE, M. E. L. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In: Ian Roberts & Mary Kato (orgs.) *Português Brasileiro: uma viagem diacrônica*. Campinas, SP. Editora da Unicamp, 1993, p. 107-28.

_____. *A perda do princípio “Evite Pronome” no português brasileiro*. Tese de Doutorado. UNICAMP. Campinas, SP, 1995.

_____. Sociolinguística Paramétrica: perspectivas. In HORA, Demerval & CRISTIAN. *Estudos linguísticos: realidade brasileira*. João Pessoa, Ideia, 1997.

_____. A perda do princípio “Evite Pronome” no português brasileiro. In: KATO, M.; NEGRÃO, E. *Brasilian Portuguese and the null subject parameter*. Iberoamericana: Madrid, 2000. p. 223-258.

_____. O sujeito expletivo e as construções existenciais. In: RONCARATI, C. & ABRAÇADO, J. (orgs.) *Português brasileiro – contato linguístico, heterogeneidade e história*. Rio de Janeiro: 7 Letras, 2003. p. 123-131.

_____. *Avanço no estudo da mudança sintática associando a Teoria da Variação e Mudança e a Teoria de Princípios e Parâmetros*. Cadernos de Estudos Linguísticos (57.1), Campinas, Jan./Jun. 2015.

DUARTE, M. E; KATO, M. A. Mudança paramétrica e orientação para o discurso. Comunicação apresentada no XXIV ENCONTRO NACIONAL DA APL, Universidade do Minho, Braga, 2008.

DUARTE, M. E; KATO, M. A.; BARBOSA, P. Sujeitos indeterminados em PE e PB. In: CONGRESSO INTERNACIONAL DA ABRALIN, 2, 2003, Fortaleza. *Boletim da Abralín*. Número especial. Fortaleza: Imprensa Universitária. UfC, 2001, v.26, p.405-409.

DUARTE, M. E; MARINS, J. E.; SILVA, H.S. *Revisiting Duarte (1995): for a gradient analyses of indeterminate subjects in Brazilian Portuguese*. Paper apresentado no Encontro Intermediário do GT em Teoria da Gramática da ANPOL, UFMG, MG, 2015.

EVERETT, D. Why there are no clitics. An alternative perspective on pronominal allomorphy,. SIL and U of Texas. Arlington. *Publications in Linguistics*, 1996.

- FARACO, C. A. Estudos pré-saussurianos. In: MUSSALIM, F. BENTES, A. C. (orgs). *Introdução à linguística: fundamentos epistemológicos*. v.3, 5ed. São Paulo: Cortez, 2011.
- FERNÁNDEZ, S. Impersonality in Spanish personal pronouns. In: *Dêixis and Pronouns in Romance Languages*. John Benjamin Publishing Company, 2013.
- FIGUEIREDO SILVA, M.C. *A posição do sujeito no português brasileiro*. Campinas, SP: Editora da Unicamp, 1996.
- FIGUEROA, E. *Sociolinguistic metatheory*. University of Oxford, 1994.
- FRANCHI, C; NEGRÃO, E.V.; VIOTTI, E. Sobre a Gramática das Orações Impessoais com *Ter/Haver*: D.E.L.T.A, v. 14, n. spe, p. 00, 1998. Disponível em http://www.scielo.br/scielo.php?script=sci_arttext&pid=S010244501998000300009&lng=pt&nrm=iso
- FRASCARELLI, M. Subjects, topics and the interpretation of referential pro. *Latural Language and linguistic Theory*, 25. 2007, p. 691-734.
- FREEZE, Ray. Existential and other locatives. *Language*, 68, 1992, p. 553-595.
- GALVES, C. M. C. A sintaxe do Português Brasileiro. *Ensaio sobre as gramáticas do português*. Campinas, SP: Editora da Unicamp, 1987.
- _____. O enfraquecimento da concordância no português brasileiro. In: Mary Kato e Ian Roberts (orgs). *Português Brasileiro uma viagem diacrônica*. Campinas, SP: Editora da Unicamp, 1993.
- HAVERKATE, H. *La desfocalización referencial*. *Hispanic Linguistics* 2 (1): 1–21, 1985.
- HENRIQUES, F. P. *Construções com verbos de alçamento: um estudo diacrônico*. Dissertação de mestrado. UFRJ, 2008.
- HENRY, A. Variation and Syntactic Theory. In: J. K. Chambers et al. (Orgs.). *The Handbook of Language Variation and Change*. Oxford, Blackwell, 2002, p. 266-282.
- HOLMBERG, A. Null Subject Parameters. In: BIBERAUER, T.; HOLMBERG, A; ROBERTS, I. *Parametric Variation: Null Subjects in Minamalist Theory*. Cambridge University Press: Cambridge, UK, 2010. p. 88-124.
- _____. *Generic Pronouns and Phi-features: evidence from Thai*. Paper apresentado no Encontro Intermediário do GT em Teoria da Gramática da ANPOL, UFMG, MG, 2015.
- KATO, M. A. *The morpho-phonology of strong and weak pronouns in Brazilian Portuguese and the pro-drop parameter*. Paper presented at the at the Seminar on Morphology, USC, 1996.
- _____. *Strong pronouns, weak pronominals and the null subject parameter*. *PROBUS* 11 (1): 1-37. 1999.
- _____. The partial pro-drop nature and the restricted VS order in Brazilian Portuguese. In: KATO, M.; NEGRÃO, E. *Brazilian Portuguese and the null subject parameter*. Iberoamericana: Madrid, 2000. p. 223-258.
- _____. *Comparando o português da América com o português de Portugal e com outras línguas*. 2005. Disponível em: www.museulinguaportuguesa.org.br/files/mlp/texto_13 KATO, M.; DUARTE, M.E.L. *The loss of VS order in Wh-questions in Brazilian Portuguese and the null subject parameter*. Paper presented at NWAVE, Las Vegas, 1998.

- KATO, M. A. & Negrão, E.V. (eds.). *Brasilian Portuguese and the Null Subject Parameter*. Vervuert: Iberoamericana, 2000.
- KATO, M. A. & TARALLO, F. *Harmonia Trans-sistêmica: variação intra e interlinguística*. Preedição-5. Campinas: SP, 1989. 41p.
- _____. Anything YOU can do in Brazilian Portuguese. In: *Studies in Romance Linguistics*. Dordrecht: Foris, 1986, 343-358.
- Kitagawa, C. & Lehrer, A. Impersonal uses of personal pronouns. *Journal of Pragmatics* 14: 739-759, 1990.
- KROCH, A. Syntactic Change. In BALTIN and COLLINS (Eds). *The Handbook of contemporary Syntactic Theory*. Mass: Blackwell, 2001, p. 699-729.
- LABERGE, S; SANKOFF, G. Anything you can do. In: *The Social Life of Languages*. Philadelphia: University of Pennsylvania Press, 1979.
- LABOV, W. Is there a creole speech community? In: *Theoretical Orientations in Creole Studies*, 1980.
- _____. *The exact description of a speech community*. Fosold and Schiffins, 1989.
- _____. *Principles of Linguistic Change: internal factors*. Oxford: Blackwell, 1994.
- _____. Padrões Sociolinguísticos (1972). Tradução: Marcos Bagno; Maria Marta Pereira Scherre; Caroline Rodrigues Cardoso. São Paulo: Parábola, 2008.
- LOPES, C. R S. *Nós e a gente no português falado culto do Brasil*. In: DELTA, v.14, n.2. São Paulo: 1998, p.405-422.
- _____. *A inserção de “a gente” no quadro pronominal do português: percurso histórico*. Tese (Doutorado em Letras). Faculdade de Letras, UFRJ, Rio de Janeiro, 1999.
- LUNGUNHO, M.V. Partição de constituintes no português brasileiro: características sintáticas. In: SILVA, D. H. (Ed.). *Língua, gramática e discurso*. Goiânia: Cânone, 2006.p. 133-147
- MAIA, F. P. S. A gente vê por aqui: fronteiras da contemporaneidade na gramaticalização de uma forma pronominal. In: *Revista Memento*. n. 2, v. 1, jul.-ago. 2009. UNINCOR. ISSN 1807-9717.
- MARINS, J. E. As repercussões da remarcação do Parâmetro do Sujeito Nulo: um estudo diacrônico das sentenças existenciais com *ter* e *haver* no PB e no PE. Tese (Doutorado em Letras). Faculdade de Letras, UFRJ, Rio de Janeiro, 2013.
- MORALES, A. The Loss of the Spanish Impersonal Particle *se* Among Bilinguals: A Descriptive Profile. In: Silva-Corvalán, Carmen (Ed.). *Spanish in Four Continents: Studies in Language Contact and Bilingualism*, pp. 148-162, 1995.
- MUSSALIM, F. e BENTES, A. C. *Introdução à Linguística: fundamentos epistemológicos*. 5ed. São Paulo: Cortez, 2011.
- NARO, A. J. The genesis of the reflexive impersonal in Brazilian Portuguese: a study in syntactic change as a surface phenomenon. In: *Language* 52.4, 1976.
- NARO, A. J. Morphological constraints on subject deletion. In: D. Sankoff e H. Cedergen (orgs). *Variation Omnibus*. Edmonton, Alberta: Linguistic Research Inc, 1981.

NUNES, J. M. *O Famigerado SE: uma análise sincrônica e diacrônica das construções com se apassivador e indeterminador*. Dissertação de mestrado, UNICAMP, 1990.

_____. Se apassivador e Se indeterminador: o percurso diacrônico no português brasileiro. *Cadernos de Estudos linguísticos*. Campinas (20), 1991, p. 33-59.

_____. Bare X'-theory and the structure formed by movement. *Linguistic Inquiry* 29 (1), 1998, p. 160-68-59.

OLIVEIRA, M. A. *Algumas notas sobre o conceito de variável lingüística e sua dimensão nas descrições gramaticais*. ABRALIN: Boletim da Associação Brasileira de Lingüística n.8, p.87-95, 1986.

OMENA, N. P. A referência à primeira pessoa do plural: variação ou mudança? In: DUARTE, M. E.; PAIVA, M. C. (org.). *Mudança lingüística em tempo real*. Rio de Janeiro: Contra Capa livraria, 2003.

OMENA, N. P.; BRAGA, M. L. A gente está se gramaticalizando? In: MACEDO, A.T.; RONCARATI, C.E.; MOLLICA, M.C.M. (org.) *Variação e Discurso*. Rio de Janeiro: 1996.

PAGOTTO, E. G. Sociolinguística. In: C. C. PFEIFFER e J. H. NUNES (Orgs.). *Introdução às Ciências da Linguagem - Linguagem, História e Conhecimento*. Campinas. Pontes, 2006, p. 49-72.

PATRICK, Peter L. *The Speech Community*. In: CHAMBERS, J.K.; TRUGDILL, P.; SCHILLING-ESTES, N. (eds.). 2004 [2002]. p. 573-597.

PAUL, H. *Princípios fundamentais da história da língua*. Lisboa: Gulbenkian, 1970.

RAMOS, J. Mudança Sintática e Teoria Gramatical. *Cadernos de Estudos Linguísticos*, 1991.

RAMOS, J. *O uso das formas VOCÊ, OCÊ e CÊ no dialeto mineiro*. Paper presented at XI Encontro da ANPOLL, João Pessoa, 1996.

RISSANEM M. Variation and the study of English Historical Syntax. In: D. Sankoff (ed). *Diversity and Diachrony*. Amsterdam/ Philadelphia, John Benjamins P.C., 1986.

RIZZI, L. *Issues in Italian Syntax*. Foris: Dordrecht, 1982.

ROBERTS, I. *A formal account of grammaticalisation in the history of Romance futures*. Ms., University of Wales, 1992.

_____. *Verbs and Diachronic Syntax*. Dordrecht: Kluwer, 1993.

_____. A deletion analysis of null subject. In: BIBERAUER, T.; HOLMBERG, A; ROBERTS, I. *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge University Press: Cambridge, UK, 2010. p. 59-87.

ROMAINE, S. What is a speech community? In: *Sociolinguistic Variation in Speech Communities*. London: Edward Arnold, 1982.

_____. What is a speech community. In: *Sociolinguistic variation in speech communities*. London: Edward Arnold, 1992.

RUBENSTEIN, M. W. *Like You Do: a bilingual perspective on the indefinite 2^a person*. Yale College, 2010.

SANKOFF, G. *The Social Life of Language*. Philadelphia PA: University of Pennsylvania Press, 1980.

SEVERO, C. G. *A comunidade de fala na sociolinguística laboviana: algumas reflexões*. Santa Catarina: Voz das Letras: Universidade do Contestado, 2008, n.9.

SILVA, L. F.; ALMEIDA, N. L. F. Formas de indeterminação do sujeito no português semiculto de Feira de Santana: uma comparação com o português popular. *Anais do XIV Seminário de Iniciação Científica da Universidade Estadual de Feira de Santana*, UEFS, Feira de Santana, 18 a 22 de outubro de 2010.

SIGURÐSSON, H.A. Argument-drop in Old icelandic. *Lingua*, 89, 1993, p. 247-280.

TARALLO, F. *Fotografias Sociolinguísticas*. Campinas: Pontes/Ed. da UNICAMP, 1989.

_____. Diagnosticando uma gramática brasileira: O português d'aquém e d'além-mar ao final do século XIX. In: ROBERTS, I. e KATO, M. A. *Português Brasileiro - uma viagem diacrônica*. Editora da Unicamp: Campinas, 1993. 425p.

_____. *A pesquisa sociolinguística*. 5.ed. São Paulo: Ática, 1997. 96p.

TARALLO, F. & KATO, M. *Harmonia trans-sistêmica: variação inter e intralinguística*. In *Preedição* 5:315-353. Campinas, Unicamp, 1989.

VIANNA, J. B. de S. Semelhanças e diferenças na implementação de a gente em variedades do português. Tese de Doutorado. Faculdade de Letras/UFRJ, Rio de Janeiro, 2011.

VIOTTI, E. Uma história para ter e haver. In: *Cadernos de Estudos Linguísticos. Unicamp* 34. Campinas: Unicamp, 1998. p. 41-50.

VITÓRIO, E. G. S. L. A. As construções existenciais e a representação do sujeito pronominal. *Revista Linguística* do Programa de Pós-Graduação em Linguística da Universidade Federal do Rio de Janeiro. V 9, n 2, 2013. ISSN 1808-835X <http://www.lettras.ufrj.br/poslinguistica/revistalinguistica>

VITRAL, L. A forma CÊ e a noção de gramaticalização. *Revista de Estudos da Linguagem*, ANO V, n. 4,v,1, 1996, p.115-124.

WEINREICH, U. *Languages in contact*. New York: Publications of Linguistic Circle of New York, 1953.

WEINREICH, U. LABOV W. & HERZOG M. *Fundamentos empíricos para uma teoria da mudança linguística*, 1968. Tradução: Marcos Bagno. Parábola: São Paulo, 2009.

ZILLES, A. M. S. The development of a new pronoun: the linguistic and social embedding of a gente in Brazilian Portuguese. *Language Variation and Change*, v. 17, n. 1, p. 19-53, 2005.

_____. O que a fala e a escrita nos dizem sobre a avaliação social do uso de a gente? *Letras de Hoje*. Porto Alegre, v. 42, n. 2, p. 27-44, junho, 2007.

APÊNDICE

Tabela 3: Distribuição das variantes em construções existenciais no PB nas duas décadas

	Você_{gen}	A gente	Nós	Se	Nulo
Década de 70	33/429 8%	6/429 1%	76/429 18%	6/429 1%	308/429 72%
Década atual	116/ 383 30%	18/ 383 5%	25/ 383 6%	4/383 1%	221/383 58%

Tabela 4: Distribuição das variantes nas construções pessoais no PB nas duas décadas

	Você_{gen}	A gente	Nós	Se	Nulo
Década de 70	318/1708 19%	505/1708 30%	360/1708 21%	332/1708 19%	193/1708 11%
Década atual	1309/ 2374 55%	644/ 2374 27%	235/ 2374 10%	67/2374 3%	119/ 2374 5%

Tabela 6: Distribuição das variantes em relação aos tempos verbais no PB nas duas décadas

Período	Tempos verbais	Você	%	A gente	%	Nós	%	Se	%	Nulo	%
1970	Presente	256	17	408	26	294	19	209	13	392	25
	Pretérito	17	6	42	16	92	35	80	30	36	13
	Futuro do Subjuntivo	3	23	6	46	2	15	1	8	1	8
	Outros	75	25	55	19	48	16	48	16	72	24
2015	Presente	989	52	461	24	160	8,5	45	2	245	13
	Pretérito	52	16	145	46	77	24	7	2	38	12
	Futuro do Subjuntivo	24	78	4	13	2	6	0	0	1	3
	Outros	360	71	52	10	21	4	19	4	56	11

Tabela 7: Distribuição das variantes nas construções hipotéticas no PB nas duas décadas

	Você_{gen}	A gente	Nós	Se	Nulo
Década de 70	44/121 36%	41/121 34%	11/121 9%	13/121 11%	12/121 10%
Década atual	180/230 78%	36/230 16%	2/230 1%	4/230 2%	8/230 3%

Tabela 8: Emprego das variantes na presença x ausência de SADVs e SPs no PB

Presença de SADVs e SPs		Você_{gen}	A gente	Nós	Se	Nulo
Década de 70	sim	189/351 54%	251/511 49%	225/436 51%	179/501 36%	168/338 49%
	não	162/351 46%	260/511 51%	211/436 49%	322/501 64%	170/338 51%
Década atual	sim	873/1425 62%	407/662 62%	137/260 53%	163/340 48%	40/71 56%
	não	552/1425 38%	255/662 38%	123/260 47%	177/340 52%	31/71 44%

Tabela 8a: Distribuição das variantes em relação presença x ausência de SADVs e SPs

Presença de SADVs e SPs		Você_{gen}	A gente	Nós	Se	Nulo
Década de 70	sim	189/1012 19%	251/1012 26%	225/1012 22%	179/1012 17%	168/1012 16%
	não	162/1125 14%	260/1125 23%	211/1125 19%	322/1125 29%	170/1125 15%
Década atual	sim	873/1620 54%	407/1620 25%	137/1620 8,5%	163/1620 10%	40/1620 2,5%
	não	552/1138 49%	255/1138 22%	123/1138 11%	177/1138 15%	31/1138 3%

Tabela 9: Distribuição das variantes nas sentenças existenciais no PB em relação à presença de locativos na década atual

SPs locativos	Você_{gen}	A gente	Nós	Se	Nulo
Presença	59/196 30%	13/196 6,5%	14/196 7%	1/196 0,5%	109/196 55,5%
Ausência	62/177 35%	4/177 1%	11/177 6%	2/177 2,5%	98/177 55,5%

Tabela 11: Frequência das variantes em relação ao tipo de sentença no PB

Tipo de sentença		Você _{gen}	A gente	Nós	Se	Nulo
Década de 70 (2137)	finitas	297/1899 16%	464/1899 24%	434/1899 23%	292/1899 15%	412/1899 22%
	infinitivas	54/238 23%	47/238 20%	2/238 1%	46/238 19%	89/238 37%
Década atual (2758)	finitas	1192/2429 49%	632/2429 26%	258/2429 11%	54/2429 2%	293/2429 12%
	infinitivas	233/329 71%	30/329 9%	2/329 1%	17/329 5%	47/329 14%

Tabela 12: Distribuição das variantes em relação ao fator paralelismo no PB

Paralelismo		Você _{gen}	A gente	Nós	Se	Nulo
Década de 70	não	165/1127 14%	292/1127 26%	200/1127 18%	280/1127 25%	190/1127 17%
	sim	187/1010 18%	219/1010 22%	236/1010 23%	221/1010 22%	148/1010 15%
Década atual	não	392/1176 33%	327/1176 28%	168/1176 14%	231/1176 20%	58/1176 5%
	sim	1033/1582 65%	335/1582 21%	92/1582 6%	109/1582 7%	13 /1582 1%

Tabela 14: Distribuição das variantes por faixa etária: tempo aparente no PB

Faixa etária		Você _{gen}	A gente	Nós	Se	Nulo
Década de 70	IDADE 1	198 25%	212 27%	116 15%	88 11%	178 22%
	IDADE 2	130 15%	232 28%	127 15%	101 12%	247 30%
	IDADE 3	23 5%	67 13%	193 38%	149 29%	76 15%
Década atual	IDADE 1	498 56%	180 20%	33 4%	41 5%	135 15%
	IDADE 2	457 47%	267 28%	112 12%	10 1%	121 12%
	IDADE 3	470 52%	215 24%	115 13%	20 2%	84 9%

Tabela 15: Ocorrência das estratégias de indeterminação no PE por tipo de construção

	FINITAS		INFINITIVAS	
Variantes	Realização	%	Realização	%
Nulo genérico	25/513	4,9%	103/145	71%
A gente	14/513	2,7%	1/145	0,7%
Nós	314/513	61,2%	12/145	8,3%
Se	160/513	31,2%	29/145	20%

Tabela 16: Distribuição das variantes em relação a tempos verbais nas duas décadas em PE

Período	Tempos	Você	%	A gente	%	Nós	%	Se	%	Nulo	%
1970	Presente			11	5	97	42	90	39	33	14
	Pretérito			0		9	27	22	67	2	6
	Futuro do Subjuntivo			0		1	100				
	Outros			0		10	27	12	32,5	15	40,5
2015	Presente			15	3	311	58	136	25,5	72	13,5
	Pretérito			2	2	60	56	28	26	17	16
	Futuro do Subjuntivo			0		9		0			
	Outros					10	27	12	32,5	15	40,5

Tabela 17: Distribuição das variantes por faixa etária: tempo aparente em PE

Faixa etária		Você _{gen}	A gente	Nós	Se	Nulo
Década de 70	IDADE 1		5 6,5	19 25%	44 57%	9 12%
	IDADE 2		6 3%	74 42%	58 33%	39 22%
	IDADE 3		2 4%	24 49%	22 45%	1 2%
Década atual	IDADE 1		2 0,7%	122 43,8%	96 34,5%	59 21%
	IDADE 2		11 3,5%	164 54%	49 16%	81 26,5%
	IDADE 3		7 3%	123 49,5%	55 22%	63 25,5%