



UNIVERSIDADE ESTADUAL PAULISTA  
“JÚLIO DE MESQUITA FILHO”  
Câmpus de São José do Rio Preto

Marcelo Henrique Vieira de Faria

**Abordagem construcional dos pronomes “a gente” e “você” em construções  
de indeterminação do sujeito**

São José do Rio Preto  
2022

Marcelo Henrique Vieira de Faria

**Abordagem construcional dos pronomes “a gente” e “você” em construções  
de indeterminação do sujeito**

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Estudos Linguísticos, junto ao Programa de Pós-Graduação em Estudos Linguísticos, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de São José do Rio Preto.

**Financiadora:** CAPES

**Processo n.º.** 88887.481575/2020-00

**Orientador:** Prof. Dr. Sebastião Carlos Leite Gonçalves

São José do Rio Preto

2022

F224a

Faria, Marcelo Henrique Vieira de

Abordagem construcional dos pronomes "a gente" e "você" em construções de indeterminação do sujeito / Marcelo Henrique Vieira de Faria. -- São José do Rio Preto, 2022

125 p. : il., tabs.

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Instituto de Biociências Letras e Ciências Exatas, São José do Rio Preto

Orientador: Sebastião Carlos Leite Gonçalves

1. Linguística. 2. Funcionalismo (Linguística). 3. Língua Portuguesa. 4. Sujeito indeterminado. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de Biociências Letras e Ciências Exatas, São José do Rio Preto. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

Marcelo Henrique Vieira de Faria

**Abordagem construcional dos pronomes “a gente” e “você” em construções  
de indeterminação do sujeito**

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Estudos Linguísticos, junto ao Programa de Pós-Graduação em Estudos Linguísticos, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de São José do Rio Preto.

Financiadora: CAPES

Processo nº. 88887.481575/2020-00

**Comissão Examinadora**

Prof. Dr. Sebastião Carlos Leite Gonçalves - Titular - Orientador  
UNESP – Câmpus de São José do Rio Preto

Prof<sup>a</sup>. Dr<sup>a</sup>. Marize Mattos Dall’Aglío-Hattnher - Titular - Membro interno  
UNESP – Câmpus de São José do Rio Preto

Prof<sup>a</sup>. Dr<sup>a</sup>. Taísa Peres de Oliveira - Titular - Membro externo  
UFMS – Câmpus de Três Lagoas

Prof<sup>a</sup>. Dr<sup>a</sup>. Lucia Regiane Lopes Damásio – Suplente  
UNESP – Câmpus de Assis

Prof<sup>a</sup>. Dr<sup>a</sup>. Vânia Cristina Casseb Galvão - Suplente  
UFG – Câmpus de Goiânia

São José do Rio Preto  
20 de maio de 2022

## **AGRADECIMENTOS**

Ao orientador, Professor Dr. Sebastião Carlos, pela orientação firme e pelo exemplo de competência, um amigo que também me orienta para as coisas da vida!

Às Professoras Dra. Marize Mattos Dall'Aglio-Hattner e Dra. Taísa Peres de Oliveira, pela leitura atenciosa e pelos apontamentos tão pertinentes.

À Taissa, minha irmã, que sempre acreditou em mim.

Aos amigos que estiveram comigo nessa jornada.

Ao Grupo de Estudos Sociofuncionalistas – GESF – pelas discussões em nossas reuniões tão proveitosas.

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

## RESUMO

Neste trabalho, analisamos a indeterminação do sujeito no português brasileiro (PB), objetivando descrever duas estratégias de indeterminação não canônicas instanciadas pela referência genérica, não dêitica, das construções pronominais [*a gente*] e [*VOCÊ*]. Buscamos reunir evidências empíricas para a hipótese de que a referenciação genérica no uso dessas construções pronominais é resultante de processos de *mudança construcional pós-construcionalização*, dadas suas trajetórias na história do português, diferentemente de outras construções pronominais que também operam a mesma estratégia, mas que sempre se mantiveram estáveis no sistema pronominal. [*A gente*], por ser originalmente expressão nominal referencial genérica não-dêitica, no tempo, passa a expressão pronominal genérica, servindo de estratégia de indeterminação do sujeito, e, só mais tardiamente, é usada como expressão pronominal dêitica de referência à primeira pessoa do plural. [*VOCÊ*], sendo em sua origem forma de tratamento dêitica de referência específica, no tempo, amplia seu caráter dêitico, passando a ser usada como expressão pronominal de referência à segunda pessoa do singular, e, só mais tardiamente, é usada como proforma para indeterminação do sujeito. Adotamos como referencial teórico os *Modelos Baseados no Uso* (BARLOW; KEMMER, 2000) e a abordagem construcional de mudança linguística (TRAUGOTT; TROUSDALE, 2013). A metodologia da pesquisa contempla duas etapas: (i) investigar os contextos de mudança por que essas construções passaram, recorrendo a trajetórias de mudanças já amplamente documentadas por trabalhos diacrônicos (LOPES, 2003; RUMEU, 2008), e interpretá-las na abordagem construcional da mudança, e (ii) com base em amostras do PB falado no interior paulista (GONÇALVES, 2007), fazer um levantamento quantitativo dessas construções a fim de analisar qualitativamente o grau de sedimentação de cada uma delas à luz de parâmetros de forma e de função/significado. Consideradas as trajetórias diferentes de mudança das construções pronominais [*a gente*] e [*VOCÊ*], a hipótese confirmada é a de que a construção com [*VOCÊ*] é mais produtiva do que a com [*a gente*], o que se confirma pelos resultados discutidos neste trabalho. [*VOCÊ*] apresenta maior frequência de ocorrência e de tipo em todos os contextos analisados, apresentando maior número de colocados em seu esquema.

**Palavras-chave:** Indeterminação do sujeito. Pronomes. Modelos Baseados no Uso.

## ABSTRACT

In this dissertation, we analyze indeterminate subject in Brazilian Portuguese (BP), aiming at descriptions of two non-canonical strategies of subject indetermination instantiated by the generic, non-deictic reference of the pronominal constructions [*a gente*] and [*VOCÊ*]. We seek to gather empirical evidence for the hypothesis that generic referencing in the use of these pronominal constructions results from post-constructionalization constructional change processes, given their trajectories in Portuguese history, unlike other pronominal constructions that also operate the same strategy, but which have always remained stable in the pronominal system. [*A gente*], as it is a non-deictic generic referential nominal expression, in time, it becomes the generic pronominal expression, serving as a strategy of indetermination of the subject, and, only later, it is used as a deictic pronominal expression of reference to the first person do plural. [*VOCÊ*], being in its origin a form of deictic treatment of specific reference, without time, it expands its deictic character, starting to be used as a pronominal expression referring to the second person singular, and, only later, is it used as a proforma for indeterminacy of the subject. We adopted as a theoretical framework Usage-Based Models (BARLOW; KEMMER, 2000) and a constructional approach to language change (TRAUGOTT; TROUSDALE, 2013). The research methodology comprises two steps: (i) to investigate the contexts of change that these constructions have gone through, using trajectories of change already widely documented by diachronic works (LOPES, 2003; RUMEU, 2008), and interpreting them in the constructional approach of change, and (ii) based on the BP spoken in the countryside of São Paulo (GONÇALVES, 2007), carry out a quantitative survey of these constructions in order to qualitatively analyze the degree of productivity of each one of them. Considering the different trajectories of change of the pronominal constructions [*a gente*] and [*VOCÊ*], we expected that this would be more productive than that, which is confirmed by the results discussed in this dissertation. [*VOCÊ*] have a higher frequency of occurrence and type in all contexts of form, higher number of elements in your schema.

**Keywords:** Indeterminate subject. Pronouns. Usage-Based Models.

## LISTA DE QUADROS, TABELAS, ILUSTRAÇÕES E ABREVIATURAS

### Lista de quadros

<b>Quadro 1</b> - Mercado teórico da GC nos dias de hoje .....	21
<b>Quadro 2</b> - Noções sobre indeterminação do sujeito em gramáticas tradicionais .....	31
<b>Quadro 3</b> - Aspectos formais e semânticos das estratégias de indeterminação com <i>a gente</i> e <i>você</i> .....	36
<b>Quadro 4</b> - Construcionalização de <i>a gente</i> no português .....	46
<b>Quadro 5</b> - Construcionalização de <i>você</i> no português .....	51
<b>Quadro 6</b> - Evolução dos paradigmas pronominal e flexional do português .....	52
<b>Quadro 7</b> - Escala de conexão discursiva.....	62
<b>Quadro 8</b> - Identificação (em destaque) dos perfis sociais da Amostra Censo do Banco de dados Iboruna que compõem a subamostra utilizada na pesquisa .....	70
<b>Quadro 9</b> - Feixe dos traços semântico-pragmáticos dos predicados das construções de indeterminação do sujeito .....	75

### Lista de tabelas

<b>Tabela 1</b> - Distribuição de estratégias de indeterminação em sentenças finitas por período de tempo.....	55
<b>Tabela 2</b> - Distribuição de estratégias de indeterminação em sentenças não-finitas por período de tempo .....	56
<b>Tabela 3</b> - Distribuição de sujeito de referência [+arb] por período de tempo em sentenças finitas .....	58
<b>Tabela 4</b> - Frequência de tempo e modo verbal das construções de indeterminação do sujeito .....	88
<b>Tabela 5</b> - Frequência das construções de indeterminação do sujeito em diferentes tipos de sentenças não-finitas .....	92
<b>Tabela 6</b> - Frequência do grau de referência do sujeito das construções de indeterminação .....	95
<b>Tabela 7</b> - Frequência token e type dos feixes de traços semântico-pragmáticos dos predicados das construções de indeterminação do sujeito .....	99
<b>Tabela 8</b> - Força colostrucional de tipos de predicados, com nível de significância entre 1 e 5, para as construções de indeterminação de sujeito com <i>A GENTE</i> e <i>VOCÊ</i> .....	103
<b>Tabela 9</b> - Frequência das construções de indeterminação do sujeito por sequência discursiva .....	106
<b>Tabela 10</b> - Formato Oracional X Grau do referente da construção com [ <i>VOCÊ</i> ].....	108
<b>Tabela 11</b> - Feixe semântico do predicado X Grau de referência da construção com [ <i>VOCÊ</i> ].....	109
<b>Tabela 12</b> - Feixe semântico X Sequência discursiva da construção com [ <i>VOCÊ</i> ] .....	110

### Lista de figuras

<b>Figura 1</b> - A estrutura simbólica de uma construção .....	22
<b>Figura 2</b> - Representação de Langacker da rede construcional .....	25
<b>Figura 3</b> - Exemplo de gradiência de relações hierárquicas entre construções, segundo Traugott e Trousdale (2021) .....	28
<b>Figura 4</b> - Rede de referenciação das estratégias canônicas de indeterminação do sujeito .....	79
<b>Figura 5</b> - Reconfiguração das redes de construções afetadas pela mudança do paradigma pronominal do PB ..	81
<b>Figura 6</b> - Construção de indeterminação do sujeito pronominal .....	84

## Lista de gráficos

<b>Gráfico 1</b> - Percurso histórico da pronominalização de <i>a gente</i> .....	43
<b>Gráfico 2</b> - Ocorrência total de sujeitos nulos .....	53
<b>Gráfico 3</b> - A representação plena das estratégias de indeterminação eles, nós, a gente e você.....	59
<b>Gráfico 4</b> - Frequência das construções de indeterminação do sujeito por tipo de amostra .....	87
<b>Gráfico 5</b> - Frequência das construções de indeterminação do sujeito em sentenças finitas e não-finitas .....	91

## **Lista de abreviaturas e siglas**

Esq. – Esquema

F – Forma

imp. – Impessoal

inf. – Infinitivo

Micro-Cxn – Microconstrução

MT – Modo e Tempo

NP – Número e Pessoa

p – Pessoa

PB – Português brasileiro

pl. – Plural

PP – Pessoa do Plural

PS – Pessoa do Singular

R – Radical

Ref. – Referência

S – Significado

sg. – Singular

SN – Sintagma Nominal

SNPro – Sintagma Nominal Pronominal

V – Verbo

## SUMÁRIO

<b>INTRODUÇÃO</b> .....	12
<b>CAPÍTULO I - FUNDAMENTAÇÃO TEÓRICA</b> .....	17
<b>1.1. Os Modelos Baseados no Uso e a Gramática de Construções</b> .....	17
<b>1.2. Abordagem construcional da mudança linguística</b> .....	26
<b>1.3. A noção de indeterminação do sujeito</b> .....	30
1.3.1. A indeterminação com os pronomes <i>a gente</i> e <i>VOCÊ</i> .....	33
<b>1.4. Referenciação</b> .....	36
<b>1.5. Dêixis e a categoria de <i>pessoa</i></b> .....	39
<b>CAPÍTULO II – A MUDANÇA NA SINTAXE DO PORTUGUÊS BRASILEIRO</b> .....	41
<b>2.1. A emergência das construções pronominais [a gente] e [VOCÊ] no português</b> .....	41
2.1.1. A emergência de [a gente] .....	41
2.1.2. A emergência de [VOCÊ] .....	46
<b>2.2. A representação do sujeito no português brasileiro na perspectiva Sociolinguística Paramétrica</b> .....	51
2.2.1. Os trabalhos de Duarte (1993, 1995) .....	51
2.2.2. O trabalho de Cavalcante (1999) .....	54
2.2.3. O trabalho de Vargas (2010) .....	57
<b>2.3. A representação do sujeito no português brasileiro na perspectiva Funcionalista</b> .....	60
2.3.1. Os trabalhos de Paredes Silva (1988, 2003) .....	60
2.3.2. O trabalho de Guerra (2017) .....	64
2.3.3. O trabalho de Oliveira (2018) .....	66
<b>CAPÍTULO III – PROCEDIMENTOS METODOLÓGICOS</b> .....	68
<b>3.1. Corpus de análise</b> .....	68
<b>3.2. Composição da subamostra</b> .....	69
<b>3.3. Parâmetros de análise</b> .....	70
3.3.1. Parâmetros de forma .....	70
3.3.2. Parâmetros de significado .....	71
3.3.3. Exclusão de dados .....	76
<b>CAPÍTULO IV – RESULTADOS E ANÁLISES</b> .....	77
<b>4.1. Reinterpretando a mudança da rede de indeterminação do sujeito no quadro dos Modelos Baseados no Uso</b> .....	77
<b>4.2. Caracterizando a construção sob análise</b> .....	83
<b>4.3. Distribuição geral dos dados de indeterminação do sujeito o <i>corpus</i></b> .....	86
<b>4.4. Parâmetros de forma</b> .....	87
4.4.1. Tempo e modos verbais .....	87
4.4.2. Formato oracional .....	90
<b>4.5. Parâmetros de significado</b> .....	94
4.5.1. Grau de referência do sujeito .....	94
4.5.2. Feixes semânticos dos predicados .....	97
4.5.3. Tipos de sequência discursiva .....	106
<b>4.6. Quadro geral da produtividade de construção de indeterminação com [VOCÊ]</b> .....	107
4.6.1. Formato oracional vs. Grau de referência do sujeito .....	108
4.6.2. Feixe semântico do predicado vs. Grau de referência do sujeito .....	108
4.6.3. Feixe semântico do predicado vs. Sequência discursiva .....	110
<b>CONCLUSÕES</b> .....	112
<b>REFERÊNCIAS</b> .....	116

## INTRODUÇÃO

O paradigma pronominal do português brasileiro (PB) tem sido um campo produtivo para as investigações de diversos processos de variação e mudança linguística. Um desses processos diz respeito a mudanças que se desencadearam com a franca entrada dos pronomes *você* e *a gente* no quadro pronominal do PB. A esse respeito, Duarte (1993) observa que a reorganização do quadro pronominal afetou consideravelmente o paradigma flexional dos verbos, que apresentava seis formas distintas, mais dois sincretismos, e atualmente apresenta apenas três formas. No entanto, se levarmos em consideração formas de concordância não padrão, o paradigma flexional dos verbos se reduz a apenas duas formas distintas que opõem a 1ª. pessoa do singular às demais pessoas, na maior parte dos tempos e modos verbais. A autora defende que, devido ao empobrecimento da morfologia flexional dos verbos, que pode acarretar ambiguidades referenciais, os falantes do PB têm preferido cada vez mais a estrutura de preenchimento do argumento sujeito de referência determinada, fenômeno que vem alterando a marcação do *Parâmetro do Sujeito Nulo* no PB, de um valor positivo, [+ *pro-drop*], para um valor negativo, [- *pro-drop*], conforme *Modelo de Princípios e Parâmetros* da Teoria da Gramática Gerativa. Como bem observam Duarte (1995), Cavalcante (1999) e Vargas (2010), as mudanças não se interrompem aí: o preenchimento obrigatório da posição de sujeito não vem afetando apenas os contextos de sujeito de referência determinada, mas também os de referência indeterminada (ou arbitrária, nos termos das autoras).

Considerando esse quadro de mudanças, neste trabalho de pesquisa, tomamos como fenômeno de investigação a indeterminação do sujeito, a fim de analisar e descrever construções instanciadas pelos pronomes *a gente* e *você*, cuja propriedade primeira é a de assumirem, no uso, referência dêitica às pessoas do discurso, mas que têm sofrido modificações semânticas, servindo também de estratégia de indeterminação do sujeito. Evidentemente, essa mudança é reflexo de mudanças mais profundas na gramática do PB, que integram um conjunto de mudanças sintáticas mais amplas e que permitem delinear a emergência de uma gramática brasileira (DUARTE, 1993).

É precisamente dentro desse panorama investigativo que esta pesquisa se insere, focalizando o comportamento das construções pronominais [*a gente*] e [*VOCÊ*],<sup>1</sup> que vêm

---

<sup>1</sup> A exemplo da representação de Scherre *et al.* (2011), neste trabalho, grafamos a construção [*VOCÊ*] em caixa-alta para representar a forma genérica que abarca indistintamente as alternantes *você*, *ocê* e *cê*, que também operam

passando por processos de mudança pós-construcionalização, ou seja, por mudanças depois que já se construcionalizaram como formas pronominais. Interessa-nos aqui, especialmente, examinar construções pronominais do tipo das exemplificadas na ocorrência em (1).

- (1) Inf.: quando a gente ama alguém você tem compromisso... você tem responsabilidade ((latidos ao fundo))... amar é um sentimento passivo... num é um sentimento agressivo (BDI - AC-114; RO: L. 627-629).

A ocorrência em (1) é bastante elucidativa do fenômeno aqui investigado, porque revela claramente a anulação da referência dos pronomes aos papéis do discurso, em favor da instanciação, em um mesmo contexto, de construções de indeterminação do sujeito, que, de tão generalizadas, tornam possível a alternância entre formas pronominais distintas no alcance de um mesmo valor referencial. Com base nessa ocorrência, deve ficar claro que nosso objeto de investigação só se circunscreve no contexto da construção mais ampla em que ocorrem as construções pronominais em análise, o que requer que delimitemos esse contexto como sendo o da predicação formada pelas construções pronominais e uma construção predicado, aqui referenciada pelo esquema [[SNPro] [PREDICADO]], como ocorre em (1) em *[[a gente] [ama alguém]], [[você] [tem compromisso]], [[você] [tem responsabilidade]]*, todas construções argumentais com sujeitos pronominais de referência indeterminada.

Como objetivo mais geral desta pesquisa, pretendemos demonstrar que os mecanismos de mudança que atuam nas construções pronominais *[a gente]* e *[VOCÊ]* não foram interrompidos quando elas passaram a integrar o paradigma pronominal. Mais especificamente, a forma pronominal *você*, por ter se desenvolvido do uso originalmente dêitico de *Vossa Mercê*, com valor referencial altamente específico, invariável e contextualmente definido (com o referente disponível na situação comunicativa), ao longo do tempo, tem sua foricidade abstratizada, em direção a usos dêiticos de sentido constante (referência à segunda pessoa do discurso) e referência variável (qualquer interlocutor), com valor genérico *default*, que só se especifica no contexto situacional, ou se mantém indefinido, no contexto linguístico, servindo,

---

a mesma estratégia de indeterminação de sujeito. Não ignoramos, no entanto, que essa construção se submete a mudanças de forma pós-construcionalização, e a consideração das alternantes e seus respectivos significados passam a ser tema relevante para outra pesquisa. Além disso, usamos as duas construções pronominais entre colchetes para mostrar que se trata de uma construção, mas, quando nos referimos metalinguisticamente aos pronomes, eles são grafados normalmente e fora de colchetes.

só mais recentemente, à estratégia de indeterminação do sujeito. A forma pronominal *a gente*, por sua vez, em direção contrária a *você*, por ter se desenvolvido do uso não dêitico do SN pleno *a gente*, com valor referencial genérico e indefinido, advindo do traço plural coletivo que o nome comporta, ao longo do tempo se abstratiza e adquire traços dêíticos, mantendo o traço de coletividade que permite, no contexto situacional, referenciar deiticamente as pessoas envolvidas no discurso (referência à primeira pessoa do plural), ou operar, no contexto linguístico, referência indefinida, não incluindo necessariamente falante e ouvinte, e também servindo, portanto, de estratégia de indeterminação do sujeito.

A hipótese aventada para nossa investigação é a de que, a partir da trajetória individual de mudança de ambas as construções, [VOCÊ] seja a construção mais produtiva como estratégia de indeterminação do que [a gente]. Na verificação dessa hipótese, temos por objetivo específico investigar, por meio da frequência de uso, o grau de sedimentação na rede linguística de indeterminação do sujeito das duas construções em foco, e, desse modo, compreendermos qual a relação, em termos de forma e de significado, que elas mantêm entre si na mesma rede que passam a integrar após suas construcionalizações como pronomes.

Diante desses objetivos e hipótese, três noções se tornam necessárias para o desenvolvimento desta pesquisa. A primeira diz respeito à *referenciação*, para a qual adotamos os postulados de Marcuschi e Koch (2006), que a entendem como um processo ligado a objetos de discurso, e, como tal, um processo sociocognitivo exclusivamente discursivo; a segunda é a noção de *dêixis* e a de *categoria de pessoa*, aqui assumidas, seguindo Benveniste (1989), como fundadas no processo de comunicação decorrente das relações de interlocução estabelecidas entre as pessoas do discurso, da qual o enunciado é apenas produto; assim, o discurso comporta unicamente duas pessoas: a primeira pessoa, o EU (o locutor), e a segunda pessoa, o NÃO-EU (o alocutário), permanecendo a terceira pessoa fora das manifestações de subjetividade da linguagem, por representar a NÃO-PESSOA ou o mundo objetivo.

A terceira noção está diretamente relacionada ao objeto de análise deste trabalho, a *indeterminação do sujeito*. Entendemos que a indeterminação do sujeito diz respeito ao fato de não ser possível, dentro de um contexto discursivo, se alcançar uma identidade referencial para o argumento sujeito, esteja ele expresso ou não nos limites da oração. Se a referência não é recuperável no discurso multiproposicional, ela é, então, uma referência indeterminada. Desse modo, parece-nos mais apropriado tratar a estratégia de indeterminação do sujeito como um

expediente mais pragmático do que sintático, porque o que torna o sujeito indeterminado é, em primeiro lugar, a intenção do locutor, que, por uma razão qualquer, não sabe ou não quer individualizar, precisar, apontar o referente do argumento sujeito previsto na estrutura argumental do predicado. A partir daí, a noção de sujeito que adotamos aqui não se limita apenas a uma categoria puramente sintática. Se assim o fosse, seria contraditório tratar como indeterminados casos de sujeito expressos por pronomes nominativos, como é, por exemplo, o caso de sujeitos pronominais que operam indeterminação. Assumimos, portanto, que o sujeito expresso sempre tem uma função referencial, seja determinada ou indeterminada. Há certa distinção entre pessoa verbal e referente extralinguístico a que se reporta um pronome, porque, como salienta Setti (1997), o que não se determina é o referente.

Por se tratar de um trabalho alinhado aos *Modelos Baseados no Uso* (BARLOW; KEMMER, 2000), é imperativo utilizar amostras de uso efetivo da língua. Desse modo, por questões metodológicas, que serão mais bem esclarecidas no capítulo 3, buscamos dados sincrônicos do PB contemporâneo disponíveis no Banco de dados Iboruna (GONÇALVES, 2007), que registra a variedade do PB falado no interior paulista. Todavia, para tratarmos dos processos de construcionalização, como parte da metodologia, recorreremos a trabalhos diacrônicos já realizados por outros autores, e propomos uma reinterpretação de seus resultados sob o arcabouço dos Modelos Baseados no Uso.

Uma outra justificativa para a realização desta pesquisa se ampara no fato de que as mudanças aqui retratadas são bastante exploradas no quadro da Teoria da Gramática Gerativa (cf. DUARTE, 1993; 1995, CAVALCANTE, 1999; VARGAS 2010), mas carecem ainda de uma explicação funcional-cognitiva, como a da abordagem construcional da mudança, que tem se dedicado mais a processos de formação de construções específicas às línguas e pouco a mudanças sintáticas mais amplas, como as aqui retratadas.

Para o registro da investigação aqui proposta, este trabalho compreende, após esta seção de introdução, mais quatro capítulos e as considerações finais. No capítulo 1, apresentamos os principais postulados dos *Modelos Baseados no Uso da Língua*, da Gramática de Construções e da abordagem construcional da mudança linguística, principais subsídios teóricos nos quais se assenta nossa investigação. No capítulo 2, fazemos uma breve incursão pela literatura que trata da inserção de *a gente* e *você* no sistema pronominal do PB e das conseqüentes mudanças na sua sintaxe, dentre as quais a mudança na rede de indeterminação do sujeito aqui investigada.

O capítulo 3 traz os procedimentos metodológicos para o cumprimento dos objetivos aqui propostos. É no capítulo 4 que discutimos e expomos os resultados obtidos. Por fim, apresentamos as considerações (ainda preliminares) a que chegamos, seguidas das referências utilizadas neste trabalho.

## CAPÍTULO I - FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, esboçamos a fundamentação teórica que balizam esta pesquisa. Na primeira seção, apresentamos os principais pontos dos Modelos Baseados no Uso e da Gramática de Construções, em que abordamos, principalmente, as noções de língua e de construção adotadas nesta dissertação. Na seção subsequente, focalizamos os princípios da abordagem construcional da mudança que fundamentam as discussões sobre mudanças tratadas aqui. Nas próximas seções, mobilizamos conceitos referentes às noções de *indeterminação do sujeito*, focalizando a indeterminação instanciada pelos pronomes *a gente* e *ocê*, de *referenciação*, de *dêixis* e a categoria de *pessoa*.

### 1.1 Os Modelos Baseados no Uso e a Gramática de Construções

Este trabalho tem suas bases teóricas ancoradas nos *Modelos Baseados no Uso da Língua* (daqui em diante, MBU) (LANGACKER, 1987; BARLOW; KEMMER, 2000; BYBEE, 2016), vertente que conjuga postulados teóricos advindos de diferentes teorias de base cognitiva e funcional (CROFT, 2001; GOLDBERG, 1995; TRAUGOTT; TROUSDALE, 2021 [2013]) ao associar, na descrição das línguas, funções da linguagem a aspectos cognitivos. Esse modelo teórico se consolida, portanto, na consideração de trabalhos anteriores desenvolvidos por representantes da Linguística Funcional da Costa Oeste dos Estados Unidos, como Talmy Givón, Paul Hopper, Sandra Thompson, Joan Bybee, Elizabeth Traugott, dentre outros, e por representantes da Linguística Cognitiva, como George Lakoff, Ronald Langacker, Adele Goldberg, John Taylor, William Croft, dentre outros. Essa junção passa a constituir um novo quadro teórico, porque subjaz a ela a crença de que aspectos estruturais e comunicativos que regem o funcionamento da linguagem humana só podem ser mais bem compreendidos à luz de processos que regem a cognição geral (BYBEE, 2016).

Como pressuposto fundamental, os MBU assumem que a gramática é de natureza flexível, relativamente instável e multifacetada, e que, a partir das negociações entre os falantes na interação verbal, emerge a padronização de usos regularizados na dimensão social e cultural. Por aí, chega-se ao entendimento de que o discurso e a gramática devem ser estudados em conjunto e se tornam centrais na compreensão de como a língua, na situação concreta de

intercomunicação, é estruturada. Assim, a gramática e o discurso interagem e se influenciam mutuamente.

Barlow e Kemmer (2000) admitem que há uma estreita relação entre os eventos experimentados pelos falantes e a representação abstrata da gramática. Os oito princípios que põem em evidência a relação entre estrutura e uso, são resumidos abaixo da seguinte maneira:

(2) **Princípios dos MBU:**

(i) **Importância da frequência de uso:** o modo do processamento de uma unidade linguística é afetado pela sua rotinização; assim, frequência de uso é tanto resultado quanto força de moldagem do sistema;

(ii) **Compreensão e produção integradas:** a estrutura e os atos do processamento mental não são vistos separadamente; o funcionamento do sistema linguístico é estruturado pelo uso e, assim, performance é parte da competência;

(iii) **Foco na experiência durante a aquisição de linguagem:** não se admitem estruturas inatas, pois as crianças abstraem padrões linguísticos mais gerais do uso da língua; dessa maneira, produção e compreensão linguística são significantes em fase de aquisição da linguagem;

(iv) **Emergência das representações linguísticas:** unidades linguísticas não são “estocadas” em qualquer localização neural particular, porque são sempre emergentes de rotinas cognitivas de padrões mentalmente ativados;

(v) **Importância dos dados de uso:** é no uso efetivo da língua (produção e compreensão) que teorias linguísticas devem se pautar, buscando dados empíricos em *corpora* e rejeitando intuições como fonte de pesquisa;

(vi) **Uso, variação sincrônica e mudança diacrônica:** quanto mais interagem, mais os usuários tendem à padronização; o uso é o *locus* da mudança. O falante é fonte de micromudanças diacrônicas. Em cada estágio da mudança, operam as mesmas motivações que afetam a percepção e a produção;

(vii) **Sistema linguístico e processos cognitivos gerais:** processos de abstração linguística não diferem de processos de outros domínios cognitivos. A estrutura linguística é parte de estruturas conceituais;

(viii) **Papel do contexto:** o significado não está na estrutura em si, mas nos contextos de uso. Há uma complexa interação entre contexto e representações cognitivas abstratas da experiência.

(BARLOW; KEMMER, 2000, com adaptações).

O que fica evidente nesses oito princípios é a concepção de que as estruturas linguísticas refletem outras capacidades cognitivas. Segundo Tomassello (2009), do mesmo modo como o falante categoriza objetos do mundo externo a partir de suas experiências, também as construções linguísticas devem ser concebidas como esquemas cognitivos, apreendidos e armazenados a partir da experiência comunicativa. Bybee (2013) postula que as construções linguísticas são concebidas como esquemas cognitivos de mesmos tipos dos que estão na base de outras habilidades não linguísticas. Esses são pontos basilares do modelo.

Os pressupostos mais centrais que norteiam esta pesquisa estão assentados, então, nas seguintes noções dos MBU: a rejeição da autonomia da sintaxe, o reconhecimento da integração da semântica e da pragmática às análises, o contínuo entre léxico e gramática, a relação intrínseca entre estrutura linguística e uso, a noção de que a língua deve ser analisada em contextos naturais de comunicação, pois tem natureza emergente e adaptável às pressões advindas do uso, e a noção de que as motivações semântico-pragmáticas levam à emergência das estruturas linguísticas e a padrões de regularização de novos significados. Subjacentes a esses pressupostos centrais, estão a concepção de língua como um sistema adaptativo e complexo (BYBEE, 2016) e a de que sua estrutura emerge à medida que ela é usada. Consoante essa concepção de língua, a gramática é entendida como resultado da estruturação de fatores cognitivos e comunicativos da língua (TRAUGOTT, 2004) e como um conjunto de esquemas/processos simbólicos que são utilizados na produção e na organização do discurso. Assim, os MBU não põem o foco de análise exclusivamente nas estruturas, pois assumem que os fenômenos estruturais se originam a partir de processos cognitivos de domínio geral e na relação simbólica entre forma e significado.

Para o desenvolvimento de estudos filiados aos MBU, Bybee (2016) propõe que se investigue a relação entre cinco processos da cognição humana, que são de domínio geral, e processos de domínio específicos à língua: (i) categorização, (ii) *chunking*, (iii) *memória enriquecida*, (iv) *analogia* e (v) *associação transmodal*, os quais seguem resumidamente explicitados em (3).

### (3) Processos cognitivos de domínio geral e específico à língua

- (i) a *categorização*, na cognição humana, é responsável por criar vários tipos de categorias perceptuais, a partir da experiência, e independe da língua; no domínio específico da língua, construções, de um modo geral, se emparelham, por identidade, e são reconhecidas e associadas a representações já estocadas na mente; assim, as categorias do sistema linguístico resultam desse processo, sejam elas unidades sonoras, morfemas, palavras, sintagmas ou construções;
- (ii) *chunking*, na cognição humana, é o processo que possibilita entender por que as pessoas se aprimoram em certas atividades cognitivas e neuromotoras com a prática; no domínio específico da linguagem, é o processo que explica como sequências de palavras são “embaladas” juntas na cognição, de modo que são acessadas como unidade simples;
- (iii) a *memória enriquecida* tem relação com a estocagem mental de detalhes da experiência, o que, no domínio específico da língua, inclui a experiência com detalhes fonéticos para palavras e sintagmas, contextos de uso, significados e inferências associadas a enunciados. É por meio de exemplares construídos com base em experiências linguísticas consideradas idênticas que a memória linguística é representada; no domínio geral, memórias não linguísticas também têm impacto nas representações cognitivas e na estrutura neurológica;

- (iv) a *analogia*, no domínio geral, tem sido estudada em termos de estruturas relacionais sobre estímulos visuais, como cenas, formatos e cores; no domínio específico da linguagem, é o processo pelo qual enunciados novos emergem com base em enunciados de experiências prévias;
- (v) em termos de domínio geral, a capacidade para fazer *associações transmodais* se refere a experiências co-ocorrentes que tendem a ser cognitivamente associadas. Isso quer dizer que processos implícitos, automáticos, na detecção de padrões ocorrem tipicamente entre os níveis encadeados mais altos de nódulos ativados; no domínio específico da linguagem, é o processo pelo qual o elo entre significado e forma se estabelece. Isso implica dizer que inferências contextuais de enunciados particulares também podem vir a ser associadas a sequências particulares, dando surgimento a mudanças no significado.

(BYBEE, 2016; com adaptações)

Ao relacionar esses cinco processos cognitivos, a autora chega ao entendimento de que a mente humana não é organizada em módulos compartimentados, independentes, a exemplo de como formalistas concebem a organização da mente, com módulo específico para a gramática. Com base nesse entendimento, a mente é organizada de modo que processos de domínio geral e específicos se inter-relacionam e operam em várias áreas da cognição humana. Sendo assim, é por meio desses processos cognitivos que os MBU discutem o modo como suas aplicações interativas no uso da língua criam as categorias e as unidades linguísticas, estruturas sequenciais, tais como construções e constituintes.

Pinheiro e Alonso (2018), em artigo em homenagem aos 30 anos da Gramática de Construções, argumentam que os estudos conhecidos como Gramática de Construções (GC) começaram a ser desenvolvidos na década de 1980, quando houve um reavivamento do interesse pelos padrões sintáticos, simultaneamente idiossincráticos e produtivos. É com o trabalho de Fillmore (1988), intitulado *The mechanisms of “Construction Grammar”*, que se pode considerar que um novo modelo teórico, batizado de *Construction Grammar*, nasce. A partir daí, diversos modelos de GC foram elaborados, desde os mais formais, até os que adotam os MBU. O quadro 1, retirado dos autores, demonstra os diversos estudos construcionistas e seus principais alinhamentos teóricos.

**Quadro 1** - Mercado teórico da GC nos dias de hoje

Modelos da GC	Alinhamentos	Incorpora processos associados à cognição geral?	Contempla os efeitos do uso sobre a representação subjacente?
Gramática de Construções de Berkeley (FILLMORE, 1988; FILLMORE; KAY; O'CONNOR, 1988)		Não	Não
Gramática de Construções Baseada no Signo (BOAS; SAG, 2012)		Não	Não
Gramática de Construções Cognitiva (GOLDBERG, 2006; a sair)		Sim	Sim
Gramática Cognitiva (LANGACKER, 1987; 1991)		Sim	Sim
Gramática de Construções Radical (CROFT, 2001)		Sim	Sim
Gramática de Construções Fluida (STEELS, 2011)		Sim	Sim
Gramática de Construções Corporificada (BERGEN; CHANG, 2005)		Sim	Sim

**Fonte:** Pinheiro; Alonso (2018, p. 21, com adaptações).

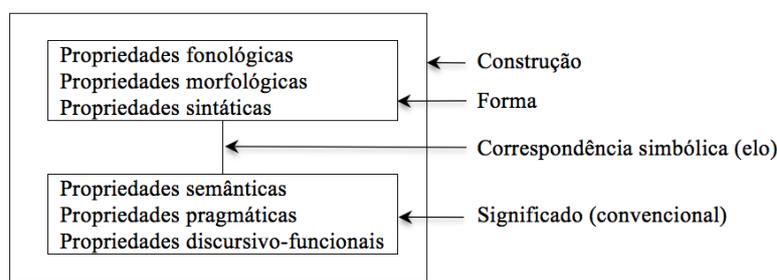
A verdade é que a GC surge em reação ao modelo de organização da gramática gerativa, para a qual cada componente (fonológico, sintático e semântico) opera independentemente, tendo, o sintático, proeminência sobre os demais. Assume-se, portanto, como pressuposto fundamental da GC a noção de que todo conhecimento gramatical que reside na mente do falante pode ser representado como uma construção, desde um morfema até padrões mais gerais de regras sintáticas e semânticas. Traugott e Trousdale (2021) postulam que, em geral, abordagens construcionistas compartilham os seguintes quatro princípios:

- a) A unidade básica da gramática é a construção, um pareamento convencional entre forma e significado (cf., p. ex., LAKOFF, 1987; FILLMORE; KAY, O'CONNOR, 1988; GOLDBERG, 1995, 2006);
- b) A estrutura semântica é mapeada diretamente na estrutura sintática superficial, sem derivações (cf. GOLDBERG, 2002; CULICOVER; JACKENDOFF, 2005);
- c) A língua, como outros sistemas cognitivos, é uma rede de nós e elos entre nós; as associações entre alguns desses nós tomam a forma de hierarquias de herança (relações taxonômicas que capturam o grau em que propriedades de construções de nível mais baixo são previsíveis a partir de construções mais gerais) (cf. LANGACKER, 1987; HUDSON, 1990, 2007).
- d) A variação translinguística (e dialetal) pode ser explicada de vários modos, incluindo processos cognitivos de domínio geral (BYBEE, 2010; GOLDBERG, 2013) e construções específicas da língua (cf. p. ex., CROFT, 2001; HASPELMATH, 2008).
- e) A estrutura linguística é moldada pelo uso da língua (cf. p. ex., BARLOW; KEMMER, 2000; BYBEE, 2010).

(TRAUGOTT; TROUSDALE, 2021, p. 27-28).

A GC reconhece que não há um nível da gramática independente e principal; há, ao contrário, o entendimento de que os diversos níveis linguísticos operam conjuntamente na construção, configurando propriedades dela. Nesse contexto, a gramática é concebida como simbólica e organizada em família de construções, compondo uma rede construcional, com diferentes níveis. A proposição de Goldberg (1995; 2006), de que a construção é qualquer unidade simbólica com forma e significado pareados, é sintetizada por Croft (2001) num esquema bastante conhecido e que explicita em que consistem os polos da forma e do significado, conforme se observa na figura 1, reproduzida do autor.

**Figura 1** – A estrutura simbólica de uma construção



**Fonte:** Croft (2001, p. 18; tradução própria)

A representação na figura 1 indica um dos princípios do modelo, que consiste em admitir que a forma básica da estrutura sintática é a construção, concebida como uma estrutura gramatical complexa, composta pelo pareamento simbólico de forma e significado. Desse modo, Croft (2007) postula que as dimensões *forma* e *significado* estão relacionadas às propriedades básicas que as constituem e que devem ser consideradas para a descrição de uma construção. Chega-se, a partir desses princípios, à noção de que esse modelo de representação gramatical é apropriado por conseguir apreender fenômenos empíricos diversos das línguas, já que se assume que a gramática é um inventário estruturado de unidades simbólicas, caracterizadas pelo pareamento de *forma* e *significado*, sendo o significado definido por todos os aspectos convencionalizados da função de uma construção, incluindo propriedades da situação, do discurso e do contexto pragmático da interação verbal (CROFT, 2007).

Um clássico estudo dentro do modelo da GC, e bastante conhecido entre os construcionistas, é o de Goldberg (2006), no qual ela investiga construção de estrutura argumental (transitiva, intransitiva, bitransitiva, etc.), para apontar que a interpretação e formação de padrões de uma estrutura argumental não são feitas exclusivamente a partir das especificações do verbo principal. Verbos não prototípicos de uma construção bitransitiva, isto é, verbos que não selecionam três argumentos, podem atuar dentro dessa construção e, conseqüentemente, exercer a mesma função dos verbos prototípicos. O que Goldberg (2006) postula é um dos fundamentos da GC: o sentido da construção é captado a partir do pareamento entre *forma e função*,<sup>2</sup> pois verbos que inicialmente não instauram uma construção bitransitiva, podem ser empregados em uma construção desse tipo exercendo a mesma função. Desse modo, o postulado principal é o de que o valor de uma construção está na combinação de propriedades formais com a função comunicativa.

Na GC, uma rede taxionômica de construções mostra como elas estão todas interligadas, e cada nível de representação da construção, do mais esquemático e geral aos mais específicos, constitui um nó de uma rede taxionômica, com construções interligadas nos planos horizontal e vertical, em um nível apropriado da rede. A elaboração de redes de construções, ligadas por elos de herança, deve obedecer, segundo Goldberg (1995), a princípios psicológicos de adequação da gramática, os quais seguem apresentados em (4).

(4) **Princípios relevantes para organização de redes construcionais** (GOLDBERG, 1995; p. 67-68)

(i) **Princípio da motivação maximizada:** entre duas construções sintaticamente relacionadas deve haver motivações semânticas e pragmáticas.

(ii) **Princípio da não-sinonímia:** entre construções sintaticamente diferentes deve haver também diferenças semânticas e/ou pragmáticas.

*Corolário A:* duas construções sintaticamente distintas e semanticamente sinônimas devem ser pragmaticamente distintas. Aspectos pragmáticos particulares da construção envolvem estrutura da informação (Tópico e Foco) e aspectos estilísticos, como registro.

*Corolário B:* duas construções sintaticamente distintas e pragmaticamente sinônimas devem ser semanticamente distintas.

(iii) **Princípio do poder expressivo maximizado:** o inventário de construções é maximizado para atender a propósitos comunicativos.

(iv) **Princípio da economia maximizada:** diante do Princípio (iii), o número de construções distintas é minimizado tanto quanto possível.

---

<sup>2</sup> Entendemos *função* como equivalente a *significado*, que recobre as propriedades semânticas, pragmáticas e/ou discursivos-funcionais que Croft (2001, p. 18) atribui ao polo do *meaning*.

Com base nesses princípios de adequação psicológica da organização de redes, Goldberg (1995) propõe a verificação de tipos de elos relacionais entre construções *por polissemia* (relação entre um sentido específico e alguma extensão desse sentido); *por subpartes* (partes de uma construção mais ampla existem independentemente como construções particulares); *por instanciação* (elementos fixos de uma construção específica podem instanciar outra construção); e *por extensão metafórica* (projeção de sentido de uma construção pode construir construção de um outro domínio semântico). No plano vertical da rede, as construções se ligam por elos de herança: um nó de nível apropriado da rede herda propriedades de nós dominantes, numa relação assimétrica, porque uma construção A, de nível mais alto, motiva uma construção B, de nível mais baixo, se, e somente se, B herda traços de A.

A metáfora de rede foi proposta por Hudson (1984), que entende a língua como uma rede conceitual. É central esse entendimento para a Linguística Cognitiva porque subjaz a ele a noção de que a organização da língua não difere da organização de outros aspectos da cognição. É em Langacker (2008) que encontramos uma definição da arquitetura da língua representada em redes. Segundo o autor,

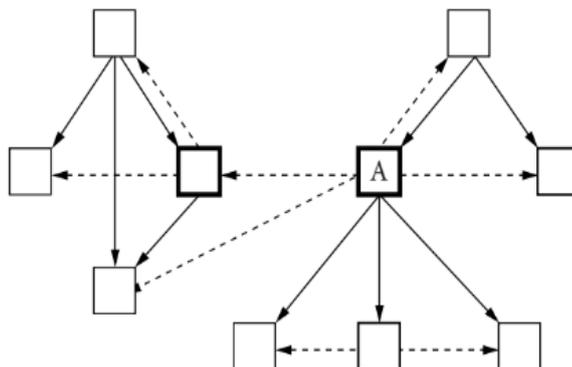
Podemos descrever uma língua como um inventário **estruturado** de unidades linguísticas convencionais. Essa estrutura – a organização das unidades em redes e conjuntos – está intimamente relacionada ao uso da língua, tanto a moldando quanto sendo moldado por ela. (LANGACKER, 2008, p. 222, grifo no original; tradução nossa).<sup>3</sup>

A figura 2, a seguir, também retirada do autor, exemplifica as redes cognitivas.

---

<sup>3</sup> No original: “[We] can describe a language as a structured inventory of conventional linguistic units. This structure — the organization of units into networks and assemblies — is intimately related to language use, both shaping it and being shaped by it”.

**Figura 2** - Representação de Langacker da rede construcional



**Fonte:** Langacker (2008, p. 226).

A representação metafórica da rede linguística na figura 2 exemplifica como as construções se associam. Cada caixa representa uma construção, ou seja, nós, e a caixa em negrito com a letra A representa o protótipo da categoria ao qual se ligam outras construções por elos diversos, organizados em diferentes graus de complexidade. Langacker (2008) argumenta que as construções destacadas em negrito representam a construção que é mais facilmente ativada pelo usuário, numa relação oposta com as construções não destacadas, cuja relação é a de mais distante do protótipo. As setas representam como os usuários ativam as construções, e cada tipo mostra o grau de sedimentação na rede: as setas contínuas ativam construções por elaboração, ou seja, a construção de partida é convencionalizada, e a construção de chegada é uma particularidade do evento de uso categorizado pela construção dominante. As setas descontínuas mostram como as construções são ativadas por relação de extensão.

Como Traugott e Trousdale (2021) afirmam, uma representação em redes como a da figura 2, por ser bidimensional, não dá conta de captar as multidimensões esperadas de uma rede linguística, visto que esta é conceptualizada em termos de redes neuronais. Assumimos nesta dissertação a visão desses autores por concordar a ideia de que os elos que ligam as construções “são possíveis em direções múltiplas entre funções semânticas, pragmáticas, discursivas, sintáticas, morfológicas e fonológicas de cada nó” (TRAGOTT; TROUSDALE, 2021, p. 105).

## 1.2 Abordagem construcional da mudança linguística

De uma perspectiva de mudanças que afetam as construções em rede de uma língua, a análise linguística precisa dar conta de descrever e explicar como a configuração da rede é afetada pela mudança e como ela se reconfigura face a essas mudanças. Na explicação da abordagem construcional da mudança, Traugott e Trousdale (2021), assumindo os pressupostos teóricos dos MBU, consideram como os processos cognitivos do domínio geral operam na língua e na mente do falante para criação e na convencionalização de construções. Os autores também assumem a construção como (i) *unidade básica de análise*, com pareamento de forma e significado; (ii) *convencional*, por ser compartilhada por determinada comunidade de fala; (iii) *emergente de processos de mudança construcional*, quando a mudança atinge apenas uma face da construção (ou a forma ou o significado), e de *processos de construcionalização*, quando uma construção com novo significado é criada, a partir de mudanças na forma e no significado de construções já existentes na gramática.

Traugott e Trousdale (2021) elaboram o esquema em (5), que representa o pareamento entre *forma* e *significado* de uma construção.

(5) **Esquema de representação da construção** (TRAUGOTT; TROUSDALE, 2021, p. 36)

[[F]] ↔ [[S]]

Sintetizado de Croft (2001) (cf. figura 1), esse esquema dos autores também prevê que uma construção é ligada (↔) pelo polo da Forma (F), que abriga propriedades da Fonologia, da Morfologia e da Sintaxe, e pelo polo do Significado (S), que contém propriedades do Discurso, da Semântica e da Pragmática. Os autores assumem que construções estão interligadas em rede, por diversas relações, sendo uma delas os traços de herança.<sup>4</sup>

Para um modelo que assume a construção como unidade básica de análise e a língua como um inventário de construções, torna-se imperioso explicar como as construções são organizadas e ligadas em rede e também como traços de construções mais específicas herdam traços de padrões mais gerais. Aqui encontramos grande vantagem em adotar a GC para descrever o fenômeno que investigamos, uma vez que estamos tentando captar, na sincronia

---

<sup>4</sup> Importante observar que, segundo os autores, o termo “herança” não implica a noção de diacronia. Refere-se a relações taxonômicas estritamente sincrônicas. (TRAUGOTT; TROUSDALE, 2021, p. 117; nota de rodapé).

atual do PB, micropassos da mudança linguística que levam as construções pronominais [*a gente*] e [*VOCEÊ*] a atuarem na rede de indeterminação de sujeito, e entender as propriedades de que essas construções são formadas nos ajuda na apreensão do próprio processo que as constituem.

Um ponto fundamental dos MBU é o tratamento da mudança linguística numa perspectiva holística, pois, com a adoção da noção de construção, aspectos da forma e de significado se implicam mutuamente. Assim, Traugott e Trousdale reconhecem dois processos da mudança que atingem as construções: *mudança construcional* e *construcionalização*. De modo geral, esses dois processos preveem a gradualidade da mudança, porque a *mudança construcional* constituiu-se numa série de micropassos de mudança que, em geral, precedem a *construcionalização*. A *construcionalização* é o resultado de processos anteriores de mudança e pode ser procedural (gramatical) ou lexical (conteúdo). A *mudança construcional* envolve mudanças que afetam ou a face da forma ou a face do significado, levando à criação de um nó na rede construcional, ou seja, uma nova construção, o que é caracterizado como *construcionalização*.

Traugott e Trousdale (2021) postulam três propriedades intervenientes na identificação desses dois processos de mudança: *esquematicidade*, *produtividade* e *composicionalidade*. Compreender essas três propriedades da construção é importante, pois, a partir delas, é possível “(a) apreender as propriedades mais gerais e específicas das construções; (b) capturar as semelhanças e diferenças entre as construções; (c) entender como construções passam a atrair novas construções; e (d) ajudar a definir em que consiste a relação entre forma e significado” (ROBUSTE, 2018, p. 33).

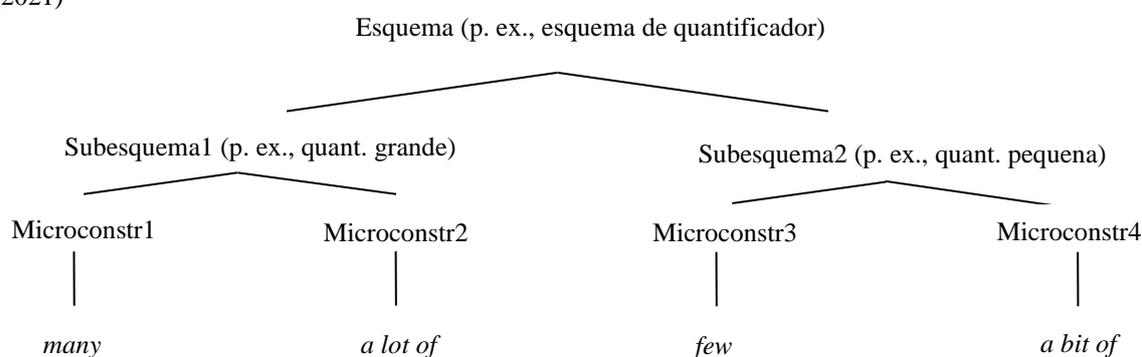
A *esquematicidade* é um conceito que envolve primeiramente abstratização, e, portanto, o processo cognitivo de categorização. Segundo os autores, um esquema deve ser entendido em termos taxonômicos, em que é possível generalizar categorias, sejam estas linguísticas ou não. Para eles, em termos de abstração, não há distinção entre construções procedurais e conteudísticas, e esquemas linguísticos consistem em abstrações de grupos de construções semanticamente gerais. Assim, dentro de um esquema, há elementos que apresentam níveis de generalização e especificidades diferentes.<sup>5</sup> Linguisticamente, é por meio da esquematicidade

---

<sup>5</sup> A título de exemplificação de abstratização de categorias, a categoria não linguística *animal* é mais abstrata e inclusiva do que *canis*, cuja generalização, por sua vez, é mais abstrata do que a de *cão*.

que padrões mais gerais da língua são capturados a partir de uma série de construções mais específicas. A figura 3, extraída de Traugott e Trousdale (2021), mostra um exemplo de como as construções se organizam em níveis mais gerais e específicos.

**Figura 3** - Exemplo de gradiência de relações hierárquicas entre construções, segundo Traugott e Trousdale (2021)



**Fonte:** Traugott; Trousdale (2021, p. 50).

Seguindo as noções dos autores, e como pode ser visto na figura 3, uma rede taxonômica de construções apresenta diferentes graus de esquematicidade: *esquema*, *subesquema*, *microconstruções* e *construtos*. O nível do *esquema* é o mais alto na hierarquia, por ser mais abstrato e virtual; no nível do *subesquema* encontram-se construções específicas com funções semelhantes; no nível imediatamente inferior, o das *microconstruções*, representam-se diversos *types* ou padrões construcionais individuais que a rede comporta; e no nível dos *construtos* (não representado na rede) ou dos *tokens*, alocam-se as ocorrências de uso real (p.ex., *many clothes*, *a lot of people*, *a few things*, *a bit of chocolate*).<sup>6</sup> Considerando essa forma de organização de redes, o conhecimento linguístico, para os MBU, é adquirido de modo ascendente ou *bottom-up*, ou seja, é a partir de dados empíricos (*construtos* ou *tokens*) que as representações esquemáticas de níveis mais altos vão sendo abstraídas (TAYLOR, 2002, p. 592 *apud* TRAUGOTT; TROUSDALE, 2013, p. 47).

Por *composicionalidade*, os autores entendem a relação entre a extensão do significado de uma construção, seu grau de transparência ou opacidade semântica e seu grau de

<sup>6</sup> Exemplos coletados no *Cambridge Dictionary* on-line, disponível em: <https://dictionary.cambridge.org> (acesso em out.2021).

convencionalização do significado. Se o significado das subpartes de uma construção é recuperável, ela é composicional e, se o significado decorre de seu todo, ela é não-composicional, opaca. Em termos sintáticos, quanto mais bem formadas e recursivamente complexas, mais composicionais são as construções. Isso ocorre com estruturas muito parecidas superficialmente e que devem ser interpretadas diferentemente pelo falante, porque o valor semântico de determinada construção pode não corresponder à sintaxe, sendo ela considerada, portanto, menos composicional.

Bybee (2016) argumenta que a produtividade tem estreita relação com a frequência de uso, porque uma construção é produtiva se ela apresenta capacidade de se estender a novos usos, na medida em que sua frequência *type* se amplia. Ao aumento de produtividade, relacionam-se os mecanismos cognitivos de *rotinização* e de *automatização*, que se instauram com o aumento da frequência de uso de uma construção. A produtividade tem impacto nos processos de mudança linguística porque tem a capacidade de expandir as posições esquemáticas das construções, que, por analogização, podem atrair novas construções para o esquema mais genérico.

Sob essa perspectiva, Traugott e Trousdale (2021) definem a produtividade como a capacidade de uma construção de atrair construções menos esquemáticas. Assim, quando uma construção consegue atrair outra construção, por meio de processos analógicos, essa nova construção passa, a partir daí, a integrar o quadro de construções que estão ligadas a um padrão construcional mais geral e abstrato e já devidamente convencionalizado na língua. A produtividade de uma construção pode ser avaliada de duas maneiras: (i) a primeira diz respeito à gama de itens lexicais diferentes que são captados em um determinado *slot* de uma construção, o que Baayen (2009) chama de “produtividade realizada”; (ii) outra interpretação que se dá à produtividade é o que Barðdal (2008) chama de “extensibilidade”, que se refere à propriedade de uma construção ser estendida a novos itens lexicais que não foram atestados nela. Dessas noções, Perek (2016) argumenta que é possível, comparando a produtividade realizada em diferentes pontos através do tempo, captar a extensibilidade de uma construção, ou seja, uma construção que se encontra estendida entre dois diferentes períodos pode ser considerada extensível a partir do período anterior.

Resumimos, a seguir, os parâmetros para verificação de mudança construcional e de construcionalização.

## (6) Parâmetros para verificação de mudança construcional e de construcionalização

**Esquematicidade:** noção ligada à categorização e que necessariamente envolve abstração: um esquema resulta de uma generalização taxonômica de categorias, tanto linguísticas quanto não linguísticas. No que respeita a esquemas linguísticos, os autores postulam ser eles de natureza abstrata e semanticamente geral, quer se trate de construções procedurais (ou gramaticais) quer de construções contedísticas (ou lexicais), e, nessa organização esquemática, há elementos que apresentam níveis de generalização e especificidades diferentes.;

**Produtividade:** é uma propriedade gradiente relativa ao potencial que tem uma construção de atrair construções menos esquemáticas, via analogização. Está ligada à frequência de uso de novos padrões colocacionais. Como medida de produtividade, a interpretação da frequência na abordagem construcional depende do nível de esquematicidade recortado como objeto de análise: enquanto frequência *type* apura a diversidade de tipos de padrões particulares (ou subesquemas) de uma construção mais geral, frequência *token* apura a frequência de construtos (ocorrências) instanciados por um *type* da construção;

**Composicionalidade:** os autores entendem a relação entre a extensão do significado de uma construção, seu grau de transparência semântica e de convencionalização de seu significado. Se o significado das subpartes de uma construção é recuperável, ela é composicional ou semanticamente transparente; se decorre do seu todo, é não-composicional ou opaca.

**Fonte:** (Traugott e Trousdale, 2021 [2013]; com adaptações).

### 1.3 A noção de indeterminação do sujeito

A indeterminação do argumento sujeito é um fenômeno bastante estudado entre os linguistas. Os trabalhos, de orientação teórica diversa, recortam para si a noção de indeterminação que mais lhes interessam. Neste trabalho, não será diferente. A partir de diversas noções encontradas em gramáticas, elegemos a noção que nos interessa aqui. Sem o intuito de sermos exaustivos, reunimos, no quadro 2, as noções de sujeito indeterminado retratadas nas principais gramáticas normativas.

**Quadro 2** - Noções sobre indeterminação do sujeito em gramáticas tradicionais

Autores	Definições
Cunha; Cintra (2001, p. 142).	“algumas vezes o verbo não se refere a uma pessoa determinada, ou por se desconhecer quem executa a ação ou por não haver interesse no seu conhecimento. [...] Nestes casos [...] <b>o sujeito não vem expresso na oração nem pode ser identificado</b> ”.
Rocha Lima (2000, p. 289).	“[o sujeito é indeterminado] <b>se não pudermos ou não quisermos</b> especificá-lo”.
Luft (2002, p. 46)	“[ocorre sujeito indeterminado] quando não se exprime o agente, que, no entanto, existe na ideia [...]. Há um sujeito (agente humano) que, <b>por não querer ou não poder</b> , não se declara”.
Infante (1995, p. 350)	“quando <b>não se quer ou não se pode</b> identificar claramente a quem o predicado da oração se refere”.
Bechara (2009, p. 35)	“ <b>não queremos ou não sabemos determinar com precisão o sujeito da oração</b> . Temos certeza de que a ação verbal foi praticada por um ser, mas nem por isso desejamos ou podemos apontá-lo. A indeterminação do sujeito não quer sempre dizer que não o conhecemos; serve também de manha inteligente da linguagem, quando não nos interessa tomá-lo patente àquele com quem falamos”.

Fonte: elaboração própria (grifos acrescidos).

Podemos observar que não há alinhamento total entre os gramáticos, e, conseqüentemente, certos problemas se apresentam nessas definições de sujeito indeterminado. O que mais se destaca é que os níveis morfossintáticos, semânticos e pragmáticos se misturam em meio às próprias definições: no nível morfossintático (marcas morfológicas no verbo; sujeito não vem expresso na oração), no nível semântico (não é possível identificar o agente; não é possível saber quem realizou a ação), no nível pragmático (não se quer ou não se pode explicitar o sujeito). Embora autores como Cunha e Cintra (2001) e Infante (1995) considerem a noção de *referência* em suas definições, noção que tangencia a adotada neste trabalho, eles ainda continuam apresentando apenas duas formas para indeterminar o sujeito (verbo na 3ª pessoa do plural ou verbo na 3ª pessoa do singular acompanhado do clítico *se*). Nesse ponto, concordamos com Setti (1997, p. 5), que afirma que, “se a indeterminação é relacionada ao referente, não há razões para serem apresentadas apenas duas formas de indeterminar o sujeito. Deve-se enfatizar o caráter genérico do referente do sujeito, e não a forma de apresentação”.

Assumimos, também, as posições de Milanez (1982) e Rollemberg *et al.* (1991), para os quais, tratar da indeterminação do sujeito implica, na realidade, ter de lidar com a indeterminação da referência do sujeito. Milanez (1982) argumenta que, mesmo que se aponte o significado de não-determinação de um agente verbal para esse expediente sintático, o que está em jogo é o fato de se desconhecer ou não poder determinar a referência do sujeito,

referência que, nos estreitos limites da oração, não está precisada, estabelecida. É o que Castilho (2010) argumenta sobre o *sujeito referencial*, definindo-o como aquele que destaca determinado referente de um conjunto dos referentes possíveis que compartilham as propriedades indicadas pelo sintagma nominal-sujeito. A indeterminação, portanto, vai em direção contrária a essa definição, isto é, não se destaca, ou não é possível destacar, um referente de um conjunto de referentes possíveis.

Milanez (1982) estabelece um paralelo entre sujeitos identificados e específicos com a noção de *generalização*. Para a autora, sujeitos que não são identificados e específicos estão dentro de construções cuja referência não se estabelece com um ou mais seres específicos e identificáveis, mas que pode envolver, indistintamente, qualquer uma das pessoas do discurso, inclusive a não-pessoa, ou as pessoas do discurso e a não-pessoa de uma só vez. Cabe, neste ponto, a distinção apresentada pela autora entre *indeterminação* e *indefinição*. Uma aparente semelhança envolve esses conceitos, em razão de, por detrás deles, estar a noção de *imprecisão*. Entretanto, há diferenças que particularizam a indeterminação e a indefinição. Sintaticamente, as estratégias de indeterminação são bem distintas entre si, enquanto apenas formas lexicais de terceira pessoa marcam a indefinição. O traço de generalização é uma característica essencial da indeterminação, enquanto para a indefinição ele é apenas uma possibilidade. Assim, Milanez (1982) afirma que:

[...] mesmo quando generaliza, a indefinição distingue-se da indeterminação pela natureza *limitada* dessa generalização. Ou seja, os recursos de indefinição *sempre* pressupõem um conjunto “fechado” de elementos que pode ser expresso ou na sua totalidade [...] ou no seu esvaziamento [...] ou parcialmente. (MILANEZ, 1982, p. 39; grifos no original).

Setti (1997), a partir dos estudos de Milanez (1982) e de Menon (1994), apresenta quatro importantes pontos na distinção entre indefinição e indeterminação:

- a. Diferentemente da indefinição, que abrange exclusivamente formas lexicais de 3.a pessoa, a indeterminação pode se referir às 1.a, 2.a ou 3.a pessoas, de forma isolada ou simultânea;
- b. A característica mais marcante da indeterminação é a generalização, sendo, portanto, questão categórica o fato de o referente ser generalizado. Já na indefinição, a generalização é apenas uma possibilidade, conforme MILANEZ (1982: 38);
- c. De acordo com MENON (1994), quando houver generalização na indefinição, poderemos observar que há o aspecto quantitativo, ou seja, a indefinição pressupõe um conjunto de elementos não definidos mas delimitáveis. A indeterminação, ao contrário, faz alusão a um grupo ilimitado, cujo referente, por não ser determinado, não tem quantificação.

d. As locuções a partir de *artigo* mais *substantivo* têm no uso do artigo uma relevante distinção entre indefinição e indeterminação. Para MENON (1994), que tratou a construção [*artigo definido* + *substantivo*] como *forma nominal*, o uso de artigos indefinidos provocam [sic.] a determinação, com caráter indefinido (p. ex.: *Um homem precisa de companhia*) e os artigos definidos são propícios à indeterminação, uma vez que compõem uma expressão generalizadora (p. ex.: *O homem precisa de companhia*). No primeiro exemplo a referência é a alguém determinável, embora não definido, e no segundo exemplo a referência é indeterminada: trata-se do ser humano de uma forma generalizada. (SETTI, 1997, p. 6-7). [grifos no original].

Entendemos, portanto, que a *indeterminação do sujeito* diz respeito ao fato de não ser possível, dentro de um contexto discursivo, se alcançar uma identidade referencial para o argumento sujeito, esteja ele expresso ou não nos limites da oração. Se a referência não é recuperável, porque não foi precisada ou determinada na totalidade do discurso, ela é, então, uma referência indeterminada. Desse modo, nos parece mais apropriado tratar a estratégia de indeterminação do sujeito como um expediente mais pragmático do que sintático, porque o que torna indeterminado o sujeito é, em primeiro lugar, a intenção do falante, que, por uma razão qualquer, não sabe ou não quer individuar, precisar, apontar o referente do argumento sujeito previsto na estrutura argumental do predicado. A partir daí, a noção de sujeito que adotamos aqui não se limita apenas a uma categoria puramente sintática. Se assim fosse, seria contraditório tratar como indeterminados casos de sujeito expresso lexicalmente, como é, por exemplo, o caso de sujeitos pronominais que operam indeterminação. Assumimos, portanto, que o sujeito expresso sempre tem uma função referencial, seja determinada ou indeterminada. Há certa distinção entre pessoa verbal e referente extralinguístico a que se reporta um pronome. Como salienta Setti (1997), o que não se determina é o referente.

### **1.3.1 A indeterminação com os pronomes *a gente* e *VOCÊ***

Em seu estudo intitulado *Recursos de indeterminação do sujeito*, Milanez (1982) investiga variadas formas de indeterminação do sujeito de que o PB dispõe. Aqui, focalizaremos apenas duas formas estudadas pela autora, a saber as formas *a gente* e *você*. Nosso objetivo aqui é o de mostrar as especificidades formais e funcionais dessas duas construções.

A autora argumenta que os recursos constituídos de pronomes pessoais, o que é o nosso caso aqui, não apresentam problemas sintáticos, dada a ordem SVO fixa considerada não marcada no português. A autora afirma que, em termos semânticos, só é possível que ocorram em estruturas de indeterminação verbos que subcategorizam o sujeito com o traço [+ humano]. Os exemplos a seguir, retirados da autora, exemplificam essa afirmação.

- (6) a. \*Late-se muito aqui.  
b. \*Ventou-se ontem. (MILANEZ, 1982, p. 37).

A autora argumenta que, nos dados encontrados em seu *corpus* de análise, *a gente* é o pronome mais frequente como estratégia de indeterminação, equivalendo ao pronome *se*. Mas afirma que essa equivalência não é absoluta por considerar que nem sempre os dois pronomes envolvem as mesmas pessoas: “*a gente* sempre inclui à primeira, condição essa que não é comum ao *se*” (op. cit., p. 54). Reproduzimos abaixo exemplo da autora a esse respeito.

- (7) a. Aqui só *se* come carne; eu não como, pois sou vegetariano.  
b. \*Aqui *a gente* só come carne, eu não pois sou vegetariano. (MILANEZ, 1982, p. 54).

Outra característica semântica que a autora observa em relação ao pronome *a gente* é que seu uso compromete menos o locutor do que os usos de *eu* e de *nós*, mesmo sempre envolvendo a 1ª pessoa, justamente por ser capaz de estender a unicidade da 1ª pessoa a outras que não estão determinadas no contexto. Sob essas afirmações, a autora postula que *a gente* opera indeterminação pela intersecção do plano da determinação com o da indeterminação, dado sua presença obrigatória da 1ª pessoa mesmo quando generaliza. Comparando *a gente* com as formas canônicas de indeterminação, a autora afirma que “embora o referido pronome abranja as três pessoas do discurso, o comprometimento do locutor é inevitável, e nesse sentido esse recurso não apresenta o mesmo grau de ‘abstratização’ dos outros três” (MILANEZ, 1982, p. 27): (i) verbo na 3ª pessoa do plural; (ii) verbo na 3ª pessoa do singular mais o *se*; e (iii) verbos na forma não finita).

Em relação ao pronome *você*, argumenta a autora que, por também operar estratégias de indeterminação pela sua capacidade de amplitude de abrangência, ele pode ser interpretado como referente a qualquer das pessoas do discurso. Diferentemente de *a gente*, que faz uma intersecção entre os planos de determinação e indeterminação, *você* opera indeterminação por meio de projeção, ou seja, o sentido de 2ª pessoa se projeta para fazer referência a qualquer pessoa do discurso.

Milanez (1982) postula que *você* é normalmente empregado como estratégia de indeterminação quando o falante quer falar sobre ações que são comuns a todos, se referir a suas impressões pessoais, e não às do interlocutor, apresentando-as como características de muitos. Esse tipo de pronome é usado, segundo a autora, como uma 2ª pessoa fictícia, para fins apenas de ilustração de um fato. A ocorrência a seguir é dada pela autora como evidência desse seu argumento.

- (8) Doc. – O que você acha do clima de São Paulo?  
Inf. – Miserável.  
Doc. – Por que?  
Inf. – Por que? há... principalmente pra mim que... às vezes saio o dia inteiro, então *você* tem que pensar que roupa *você* vai pôr de manhã, pra quando chegar meio-dia *você* não estar padecendo de calor, [...]. Eu chego em Campinas e está aquele calor que *você* morre. Então é esse problema, né, de *você* ter uma roupa adequada pra aguentar as três variações de um só dia. (MILANEZ, 1982, p. 58 – DID 37).

Para a autora, o uso do pronome *você* só é possível em contextos em que a 2ª pessoa tem o potencial de estar incluída. Vejam-se os exemplos e os testes de substituição que a autora propõe, para advogar que, nos segundos exemplos de cada par, não é possível a substituição por não incluírem o ouvinte.

- (9) a1. *Dizem* por aí que este ano vai haver eleições.  
a2. \**Você* diz por aí que este ano vai haver eleições.  
b1. Agora, doce, *se* come muito pouco doce na minha casa. (DID 11).  
b2. \*Agora, doce, *você* come muito pouco doce na minha casa.  
c1. Meu trabalho no jornal é na linha assim de algumas coisas que *a gente* acredita. (D2 255).  
c2. \*Meu trabalho no jornal é na linha assim de algumas coisas que *você* acredita.

Assim, a partir do que a autora postula, apresentamos o quadro abaixo com as características gerais das estratégias de indeterminação do sujeito empregadas com os pronomes *a gente* e *você*.

**Quadro 3** - Aspectos formais e semânticos das estratégias de indeterminação com a gente e você

Estratégias	
<i>A gente</i>	<i>Você</i>
<ul style="list-style-type: none"><li>• Ordem fixa: SVO;</li><li>• Traço semântico do sujeito: [+ humano];</li><li>• Sempre envolve a 1ª pessoa;</li><li>• É menos comprometedor;</li><li>• Generaliza por meio de intersecção.</li></ul>	<ul style="list-style-type: none"><li>• Ordem fixa: SVO;</li><li>• Traço semântico do sujeito: [+ humano];</li><li>• Normalmente usado quando o falante quer informar sobre ações que são comuns a todos;</li><li>• Refere-se às impressões pessoais do falante, e não do interlocutor, apresentando-as como características de muitos;</li><li>• Só é possível o uso num certo contexto em que a 2ª pessoa está envolvida;</li><li>• Generaliza por meio da projeção.</li></ul>

**Fonte:** elaboração própria, com base em Milanez (1982).

## 1.4 Referenciação

Neves (2002) afirma que, desde os filósofos gregos, o interesse na relação que os elementos da língua mantêm com a realidade é matéria de preocupação dos que estudavam a linguagem. Naquele momento, a autora elucida que ficaram conhecidas duas hipóteses; a primeira como “naturalista”, advinda de Platão, segundo o qual as palavras refletem, por natureza, a realidade que nomeiam, e a segunda, conhecida como “convencionalista”, postulada por Aristóteles, tinha como tese o entendimento de que o significado da palavra é resultado do acordo estabelecido entre os homens.

Podemos observar que o principal objetivo desses postulados que se propõem a pensar o potencial de referenciação dos signos linguísticos é o de rastrear como, por meio deles, vamos da língua às coisas do mundo que ela é capaz de significar. Neste trabalho, entretanto, não adotamos essa noção de referenciação assentada na relação entre língua e mundo. Aqui, nos valem dos estudos de Neves (2006), Koch e Elias (2009), Marcuschi e Koch (2006), dentre outros, para os quais, a referenciação é um processo ligado ao mundo do discurso, e, como tal, um processo exclusivamente sociocognitivo. É por meio da construção e da reconstrução de referentes no discurso que os objetos de discurso podem categorizar e recategorizar a entidades da realidade.

É nesse sentido que Marcuschi e Koch (2006) entendem que as entidades que designamos no momento em que criamos uma situação discursiva referencial são vistas como objetos de discurso e não como objetos de mundo. Para esses autores, o que vemos e o que dizemos não coincidem com o próprio real; é antes uma interpretação do real e é, portanto, na

atividade sociocognitiva que reelaboramos a realidade, sempre por meio da referência que ativamos no desenvolvimento do discurso. Em relação à referenciação do sujeito, alinhamo-nos ao que Florencio (2015) afirma quanto às intenções do produtor do texto, que enquadra uma estrutura como de indeterminação do sujeito estabelecendo “relação entre a expressão referencial e o objeto de discurso que ela representa. Esse objeto de discurso não deve encontrar referente específico no contexto [...]” (FLORENCIO, 2015, p. 67).

Para Marcuschi e Koch (2006), a indeterminação do sujeito está intimamente relacionada a formas particulares de organização do texto em que esses objetos de discurso aparecem. Os autores entendem que, por uma necessidade discursiva, designar referencialmente a existência de certos referentes não é necessário. Assim, seguindo Castilho (2010), casos de indeterminação e determinação não devem ser levados em conta sem as propriedades semânticas que os regem. Segundo o autor, “O traço de definitude é definido no texto, ou seja, nenhuma classe passível de figurar como sujeito será intrinsecamente determinada ou indeterminada” (CASTILHO, 2010, p. 298). Nesse sentido, para Castilho, nenhum caso formal de sujeito indeterminado pode ser estabelecido previamente; é somente no texto, e contextualmente, que esses casos de determinação ou indeterminação devem ser validados.

Dik (1989) assume que o falante pode se referir a uma entidade por meio de um termo, e a referenciação é um processo pragmático e cooperativo do falante, no qual discursivamente se aponta uma entidade sobre a qual algo pode ser predicado, seja essa entidade do mundo mental ou extramental. Corroborando essa noção, Neves (2006) postula que

No processo da língua em uso, os participantes de um discurso negociam o universo de discurso de que falam, e, dentro dele, num determinado momento, escolhem referir-se a algum (alguns) indivíduo(s) cuja identidade estabelecem – ou não – segundo queiram – ou não – garantir sua existência nesse universo. Isso significa que referenciação envolve interação, e, conseqüentemente, intenção. (NEVES, 2006, p. 75).

E também

A captação da referência envolve o universo discursivo, nascido de uma negociação entre os interlocutores para estabelecimento das entidades que nele devem existir, e um componente importante desse processo é a intenção que o falante tem de referir-se a algum indivíduo (NEVES, 2006, p. 80).

Diversos autores, como Dik (1989) e Du Bois (1980), por exemplo, estudaram a noção de referencialidade e propuseram a distinção entre *definido x indefinido*, *identificável x não-identificável*, *específico x genérico*. Embora esses traços compartilhem noções entre si, nos ateremos aos traços *específico x genérico*, que é o que mais nos interessa neste trabalho. Para Dik (1989), se se tem em mente um referente particular dentro o conjunto designado pelo termo, esse termo é *específico*. Se se pode interpretar um referente arbitrariamente como qualquer um dentro aqueles designados pelo termo, esse termo é *genérico*.

Camacho, Dall’Aglio-Hattner e Gonçalves (2008), ao tratarem do substantivo no português falado culto no Brasil, evidenciam que estratégias sintáticas, semânticas e discursivas podem revelar graus de referência que um determinado nome pode assumir. A entidade referenciada pode deixar de ser individualmente especificada com a perda gradativa da referencialidade, a depender da função que o nome exerce na cadeia sintagmática. Pode-se extrair dessa descrição dos autores que o substantivo, no seu funcionamento discursivo, comporta **graus de referencialidade**, que vão desde uma **referencialidade nula**, quando em posição predicativa atuam como classificadores ou qualificadores, como em (10), até uma referencialidade **descritiva específica**, como em (12), passando por casos intermediários de **referencialidade genérica**, como em (11). De (10) a (12), sistematizamos os exemplos dados pelos autores em um *continuum* de graus de referencialidade do nome, de nula a específica e definida.

(10) **Referencialidade nula do nome** [-genérica; - definida]

a. Aquilo me chocou [eu] *era tão criança*. [DID SP 234] (qualificador)

b. *Eu era aluna da Maria Ulineva*. [DID SP 234] (classificador). (CAMACHO *et al.*, 2008, p. 36; grifos acrescidos).

(11) **Referencialidade genérica** [+ genérica; +/- definida]

a. Dizem que *o baiano* [+ genérica; + definida] *é preguiçoso*... ((risos)) que *o:cearense* [+ genérica; + definida] *é preguiçoso*... porque realmente você depois de comer aquilo tudo cê tem que ter uma hora pra descansar [DID RJ 328]

b. *a gente toma vinho* [+ genérica; - definida] de acordo também...com o tipo de comida...se é carne...aqueles hábitos que a gente tem...*se é carne é vinho tinto* [+ genérica; + definida]...se é peixe *a gente usa vinho branco* [DID RJ 328]. (CAMACHO *et al.*, 2008, p. 36-37; grifos acrescidos).

(12) **Referencialidade descritiva específica** [+ específica; + definida]

a. tive a oportunidade de trabalhar *fazer uma cena com o:: o balê russo* [DID SP 234]

b. *Eu era aluna da Maria Ulineva* [DID SP 234]. (CAMACHO *et al.*, 2008, p. 36; grifos acrescidos).

Lyons (1977) postula que, nas línguas naturais, a referência nominal a indivíduos abriga três tipos de expressões, conforme tipologia do autor dada em (13).

(13) **Tipologia de referenciação** (LYONS, 1977)

- (i) *Expressões genéricas*: se referem à classe de indivíduos;
- (ii) *Expressões definidas*: se referem a algum indivíduo específico ou a alguma classe específica de indivíduos;
- (iii) *Expressões indefinidas*: não se referem nem a algum indivíduo específico nem a alguma classe de indivíduos.

A partir do que expomos até aqui sobre referenciação, podemos dizer que no processo de construção de referentes na cadeia discursiva, os casos de indeterminação do sujeito se encaixam em: (i) segundo Dik (1989), referentes genéricos, cuja referência pode ser interpretada arbitrariamente como qualquer um dentre aqueles designados pelo termo; e (ii) segundo Lyons (1977), combinação de traços de expressões genéricas, que se referem a classe de indivíduos, e de expressões indefinidas, cuja referência não se ancora nem em algum indivíduo específico nem em alguma classe de indivíduos.

### **1.5 Dêixis e a categoria de *pessoa***

Pode-se dizer que, desde Saussure, a linguística moderna estava bastante preocupada, sobretudo, ora com a estrutura, corrente que ficou conhecida como Estruturalismo, ora com a capacidade humana de gerar a sintaxe das línguas, o Gerativismo. É a partir de Benveniste, e também de Jakobson, que elementos considerados à parte do sistema linguístico, como o sujeito que realiza a enunciação, foram sendo considerados em bases teóricas sobre a língua. Esses autores entenderam que a Linguística do Discurso, como ficou conhecida, era passível de ser analisada como sistema, apesar de envolver uma diversidade de particularidades.

Sob tal perspectiva, categorias como, pessoa, tempo e espaço, ganham destaque por parte desses estudiosos. É por isso que Fiorin (1999) argumenta que “compreender os mecanismos da temporalização, espacialização e actorização é fundamental para entender o processo de discursivização” (FIORIN, 1999, p. 15). Ficam implicadas aí as categorias de pessoa, tempo e espaço de que trata Benveniste (1989). Embora o termo *dêixis* não tenha um denominador comum a todos os autores que tratam do assunto, perpassam pelo conceito três

noções básicas, eu/tu, aqui e agora, basilares de sua constituição (KOELLING, 2003). Como o título desta seção já sugere, interessa-nos focalizar aqui apenas a categoria de pessoa.

Buscando uma definição de dêixis que mais se aproxima da noção de que tratamos aqui, citamos uma definição de Câmara Junior (2002):

Dêixis - faculdade que tem a linguagem de designar mostrando, em vez de conceituar. A designação dêítica, ou mostrativa, figura assim ao lado da designação simbólica ou conceptual em qualquer sistema lingüístico. Podemos dizer que o SIGNO lingüístico apresenta-se em dois tipos – o SÍMBOLO, em que um conjunto sônico representa ou simboliza, e o SINAL, em que o conjunto sônico indica ou mostra (...). O pronome é justamente o vocabulário que se refere aos seres por dêixis em vez de o fazer por simbolização como os nomes. Essa dêixis se baseia no esquema lingüístico das três pessoas gramaticais que norteia o discurso: a que fala, a que ouve e todos os mais situados fora do eixo falante-ouvinte. (CÂMARA JUNIOR, 2002, p. 90).

No que diz respeito à categoria de pessoa, segundo Benveniste (1989), como o processo de comunicação decorre das relações de interlocução estabelecidas entre as pessoas do discurso, da qual o enunciado é apenas produto, o discurso comporta, então, unicamente duas pessoas: a primeira pessoa, o EU (o locutor), e a segunda pessoa, o NÃO-EU (o alocutário), permanecendo a terceira pessoa fora das manifestações de subjetividade da linguagem, por representar a NÃO-PESSOA ou o mundo objetivo. Assumindo essa definição de pessoa, é importante ter claro que, ao tratar da estrutura das relações de pessoa no verbo, o processo de pluralização a que se submetem as pessoas do discurso não representa, pela sua singularidade, o EU e o TU quantificados, mas apenas ampliados ao alcance do NÃO-EU e/ou da NÃO-PESSOA (BENVENISTE, 1989). O autor assim se expressa

Estamos na presença de uma classe de palavras, os “pronomes pessoais”, que escapam ao *status* de todos os outros signos da linguagem. A que, então, se refere o “eu”? Algo muito singular, que é exclusivamente lingüístico: “eu” se refere ao ato de discurso individual no qual é pronunciado, e lhe designa o locutor. É um termo que não pode ser identificado a não ser dentro do que, noutro passo, chamamos uma instância de discurso, e que só tem referência atual. (BENVENISTE, 1989, p. 288).

A linguagem é tomada aí como uma entidade que porta uma organização que dá condições aos locutores, cada qual a seu tempo, de no momento enunciativo designarem-se como “eu”. É esse “eu” que serve de base para a constituição do espaço e do tempo discursivo, sendo, portanto, o centro da referência discursiva.

## CAPÍTULO II - A MUDANÇA NA SINTAXE DO PORTUGUÊS BRASILEIRO

Neste capítulo, tratamos das mudanças sintáticas por que o português vem passando, especificamente relacionadas ao quadro pronominal, e o acarretamento que isso gerou para a sintaxe do português. Primeiramente, fazemos uma resenha dos pontos principais de trabalhos sobre o processo de gramaticalização dos pronomes *a gente* e *você* (OMENA; BRAGA, 1996; LOPES, 2003; MENON, 2006; RUMEU 2008), para então tomá-los como ponto de partida para interpretá-los dentro do quadro dos MBU, no próximo capítulo. Posteriormente, recorreremos a trabalhos que flagram as mudanças que a franca entrada desses pronomes no quadro pronominal desencadeou no sistema gramatical do PB, como os trabalhos, em perspectiva da Sociolinguística Paramétrica, que tratam de mudanças que afetam o *Parâmetro de sujeito nulo* (DUARTE, 1993, 1995; CAVALCANTE, 1999; VARGAS, 2010).<sup>7</sup> Na última seção deste capítulo, apresentamos como a representação do sujeito é tratada em perspectiva funcionalista (PAREDES SILVA, 1988; OLIVEIRA, 2018; GUERRA, 2017), para assim dispormos uma visão diversificada de como nosso fenômeno se ajusta ao que já se conhece sobre a indeterminação do sujeito.

### 2.1 A emergência das construções pronominais [*a gente*] e [VOCÊ] no português

#### 2.1.1 A emergência de [*a gente*]

Embora a gramaticalização de nome a pronome não seja processo tão produtivo se comparado ao quadro de conectivos que se gramaticalizam a partir de diferentes bases lexicais ou mesmo gramaticais (nomes, verbos, adjetivos, advérbios), há na literatura linguística inúmeros estudos dedicados aos processos de gramaticalização de que resultaram os pronomes *a gente* e *você*. Nesta seção, tratamos da formação do pronome *a gente*, cuja origem é encontrada no substantivo *gente* (OMENA; BRAGA, 1996; LOPES, 2003).

Etimologicamente, a palavra *gente* origina-se do latim, *gens*, *gentis*, e possuía vários significados: raça, estirpe, linhagem, família, raça de povo, povo (em ordem decrescente de significado: gente, nação, cidade) ou ainda povo de uma cidade, de um país, de uma região ou

---

<sup>7</sup> Souza (2013) também partiu desses mesmos trabalhos para investigar o seu objeto de tese. Seguimos esse mesmo raciocínio, porque acreditamos que nosso trabalho dialoga diretamente com esses trabalhos, mesmo que em linha teórica diferente.

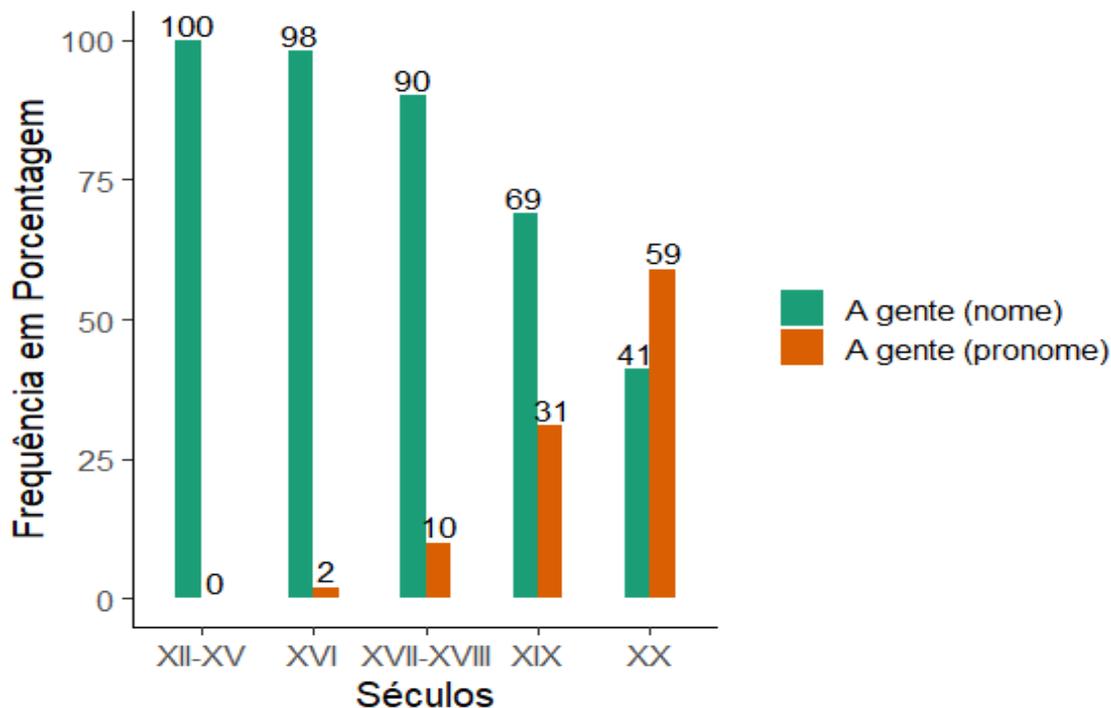
de um lugar; bárbaros; todos esses significados têm em comum a referência ao gênero humano ou à raça humana (GAFFIOT, 1934, *apud* MENON, 1996).

Lopes (2003) apresenta exemplos do português, dados em (14) e (15), do uso de *gente* como substantivo, tanto no singular, quanto no plural, em textos do século XIII.

- (14) No que o moço cantava | o judeu meteu mentes, e levó-o a ssa casa, | poi se foram as gentes. (LOPES, 2003, p. 9; *Cantigas de Santa Maria*, século XIII).
- (15) mas o monge lla cuidou fillar, mas disse-ll' a gente. (LOPES, 2003, p. 9; *Cantigas de Santa Maria*, século XIII).

Conforme afirma a autora, a forma singular *gente* começa a se cristalizar a partir do século XVI, e o uso pluralizado entra, gradativamente, em desuso. Paralelamente a essa mudança, segundo a autora, a palavra *homem*, usada como pronome de referência indefinida, começa a desaparecer no português arcaico, em razão da entrada, no português, da forma cristalizada *a gente* com função de designar os seres de maneira coletiva, genérica, indeterminada, entre os quais o falante poderia estar incluído. Assim, a forma adquire traços semânticos de primeira pessoa do discurso, pois sua interpretação passa a ser “todo mundo (todas as pessoas), inclusive eu” (LOPES, 2003, p. 64). Os contextos ambíguos serviram de gatilho para que os usos de *a gente* começassem a ascender, a partir do século XVI, e, entre os séculos XVII e XIX, crescessem progressivamente, marcando um período de transição, até a forma se implementar definitivamente como pronome no século XX, tornando-se variante do pronome *nós*, mas preservando a concordância com o verbo em terceira pessoa. Essa trajetória de mudança do nome *gente* a pronome *a gente* pode ser vista no gráfico 1, extraído de Lopes (2003).

**Gráfico 1** - Percurso histórico da pronominalização de *a gente*



**Fonte:** Lopes (2003, p. 104; com adaptações).

Uma observação bastante pertinente é em relação ao caráter genérico e globalizante que *a gente* possui é feita pela autora, que afirma a esse respeito

O caráter genérico e globalizante que *a gente* herdou do substantivo *gente* levou diversos pesquisadores a analisar esse uso da forma como um recurso para indeterminar o sujeito (Rollemberg *et alii*; 1991 e Cunha, 1993). Entretanto, não se pode, no atual estágio evolutivo, considerar *a gente* como pronome indefinido, mas sim como pronome pessoal, uma vez que, como aponta Nascentes (1953:170), nas classes incultas no Brasil, o verbo deixa de estabelecer a concordância formal para fazer concordância semântica com a primeira pessoa do plural, pois “a pessoa que está falando tem em mente a sua pessoa e as mais, com ela associadas”: “*a gente vamos hoje*”; “*a gente tínhamos de voltar*”. (LOPES, 2003, p. 11).

É sob tal perspectiva que a autora propõe uma trajetória pela qual o pronome *a gente* passou na história do português, mostrada em (16) abaixo.

(16) *GENTE* [NOME GENÉRICO] → *A GENTE* [PRONOME INDEFINIDO] → *A GENTE* [PRONOME PESSOAL]. (LOPES, 2003, p. 11).

Em um clássico artigo, Omena e Braga (1996), seguindo uma perspectiva funcionalista da língua, objetivam demonstrar a gramaticalização de *a gente*. Para tanto, as autoras aplicam os princípios postulados por Hopper (1991), quais sejam: *estratificação*, *divergência*, *especialização*, *persistência* e *decatégorização*.

O primeiro princípio, *estratificação*, refere-se à coexistência de variantes dentro de um mesmo domínio funcional. Assim, como é sabido, a partir de vários estudos (LOPES, 1993; OMENA, 1986; dentre outros), as formas *nós* e *a gente* ainda competem entre si, ou seja, a substituição de *nós* por *a gente* ainda não se realizou por completo, principalmente no PB. O princípio da *divergência* considera que a unidade lexical que dá origem ao processo de gramaticalização pode manter suas propriedades originais, preservando-se como item autônomo. É o que acontece, apontam Omena e Braga (1996), com a forma *a gente*, cujas propriedades fonológicas e semânticas ainda são conservadas no substantivo *gente*, como em *muita gente*, *esta gente*. A *especialização* diz respeito ao estreitamento da escolha de formas pertencentes a um mesmo domínio funcional. Omena e Braga (1996) evidenciam usos mais especializados de *a gente* em posição de adjunto adverbial entre os falantes não-cultos. O princípio da *persistência* refere-se à manutenção de traços da forma fonte na forma gramaticalizada. A esse princípio, as autoras relacionam o traço da referência indeterminada de *gente* que permanece no pronome. Por fim, a *decatégorização* diz respeito à perda dos marcadores opcionais de categorialidade e de autonomia discursiva. É o que afirmam as autoras em relação à mudança de *a gente*; há uma forte integração da forma gramaticalizada ao paradigma pronominal que a dissocia da forma substantiva, na medida em que, como pronome, *a gente* restringe a aplicação de processos morfossintáticos característicos dos nomes (flexão, modificação, quantificação etc.).

A construção *a gente* foi estudada, dentro dos MBU, por Soares (2018),<sup>8</sup> e apresentamos aqui os resultados aos quais a autora chegou. Embora seu trabalho dialogue diretamente com o de Lopes (2003), justifica-se a sua relevância pela necessidade de explicar, diz a autora, alguns pontos não antes questionados nos trabalhos sobre o fenômeno *a gente*. Assim, partindo da noção de contextos atípicos e ambíguos de Traugott (2015), Soares busca explicar quais foram

---

<sup>8</sup> A autora investiga a emergência da microconstrução [a gente] dando bastante atenção ao item. Em seu trabalho, apesar de investigar os contextos que serviram de gatilho para a construcionalização de [a gente] no quadro pronominal, pouca atenção é dada à relação que a forma pronominal estabelece com o predicado. Aqui, damos um tratamento à construção de indeterminação do sujeito focalizando a predicação da qual as formas pronominais em foco participam.

esses contextos que propiciaram, tanto pragmática como distribucionalmente, a emergência da construção pronominal *a gente* no PB, e o porquê de ter sido essa a forma recrutada, dentre outras de semântica semelhante, como *o povo* e *o mundo* (SOARES, 2018). A autora também justifica a necessidade de um novo olhar para esse processo de mudança, evocando uma das noções mais caras à abordagem construcional da mudança linguística, a noção de que a construcionalização não se dá em um item lexical isoladamente, mas em toda a construção e na rede linguística de uma sincronia.

Soares (2018) parte de uma detalhada revisão dos trabalhos anteriores (OMENA; BRAGA, 1996; MENON, 1997; TAMANINE, 2002; LOPES, 2003; dentre outros), e ancorada nos MBU, a autora chega aos resultados que resumimos a seguir.

- i. apesar de o vocábulo *gente* ter sentido coletivo, de grupo de pessoas, os vocábulos *povo* e *mundo* apresentam significado ainda mais amplo. [...] Sendo assim, *gente* pode ter sido escolhido por se referir a um número menor de pessoas;
- ii. desde o século XVI, a forma *gente* tinha frequência mais alta em relação a *mundo* e *povo*. Diferente destas formas com semântica semelhante, o vocábulo *gente* já trazia o sentido de pessoa [+eu], partindo do indivíduo para dar a ideia de coletivo, o que levou o ouvinte a se ver como parte desse grupo e a poder interpretar a forma de maneira inclusiva;
- iii. o uso com *gente* vai se especializando na inclusão do sujeito (com futura restrição e uso pronominal), enquanto os usos com *povo* e *mundo* seguem em outra direção, tendo em vista que *o povo* tende à exclusão do falante e *todo mundo* tende à inclusão do falante, porém indicando necessariamente a ideia de totalidade. Assim posto, não se tornaram pronomes;
- iv. não só o substantivo *gente* foi mais frequente em relação a *mundo* e *povo* ao longo dos séculos, mas a construção acompanhada pelo artigo *a* já era a forma preferida em detrimento de outras, como *aquela gente*, *sua gente*, *esta gente*, *toda a gente* etc., o que propiciou o surgimento de [a gente] pronominal. Diferentemente de *a gente*, as demais formas no discurso estabeleciam uma referência anafórica, ou seja, se referiam a um grupo de gente específico e já mencionado anteriormente (*aquela gente da Capitania* etc.), porém não inseriam o falante e o ouvinte na situação comunicativa;
- v. diferentemente das formas *povo* e *mundo*, os usos com a forma *gente* antecedida de pronome demonstrativo, como *aquela gente*, *esta gente*, *essa gente*, foram muito frequentes nos dados analisados. A referência anafórica desses pronomes pode ter propiciado a construcionalização de *a gente*, visto que os pronomes foram responsáveis por mudanças linguísticas;
- vi. outra razão que propiciou a construcionalização de *a gente* em detrimento das demais com pronomes anafóricos foi o fato de o artigo *a* apresentar sentido mais abstrato e procedural do que os pronomes *essa*, *esta*, *aquela*, *sua* etc., que são mais referenciais. O artigo *a* é um elemento sintático menos referencial e esvaziado de sentido. Todavia, a alta frequência de uso dos pronomes com *gente* foi essencial para a propagação desta construção, conforme mencionado acima;
- vii. apesar de ser composta por dois vocábulos morfológicos, a forma [a gente] é formada por apenas um vocábulo fonológico, o que a distingue das outras formas, como [esta gente], [nossa gente] etc., constituídas cada qual por dois vocábulos fonológicos. Sob o ponto de vista cognitivo, por ser a mente humana holística, um único vocábulo fonológico contribui para identificarmos a construção *a gente* como

uma unidade e, portanto, acessarmos como um bloco cognitivo também num plano morfossintático. Esse link entre fonologia e morfologia é um dos fatores que explica o porquê de *a gente* ter sido recrutada para assumir a função pronominal, ao invés de *esta gente, aquela gente*, por exemplo. (SOARES, 2018, p. 146-151).

A autora, então, a partir do conceito de que construção é um pareamento de forma-significado, mostra que a construcionalização gramatical é acompanhada por mudanças no polo tanto da forma quanto do significado, e que a mudança não afeta apenas um item isoladamente, mas toda a construção, abrangendo suas propriedades formais, de significado e contextuais. O quadro 3 traz resumidamente as propriedades de forma e de significado que foram afetadas no processo da mudança e as que resultaram desse processo.

**Quadro 4** - Construcionalização de *a gente* no português

<b>[gente] – coletivo de pessoas</b>	<b>[a gente] – pronome de 1PP</b>
<p><b>Propriedades de Forma</b></p> <ul style="list-style-type: none"> <li>- Pode ser modificada por determinante, adjetivo, possessivo e quantificador (<i>a gente, gente boa, a boa gente, nossa gente, toda gente</i>);</li> <li>- Permite flexão de número e grau (<i>as gentes, gentinha</i>);</li> <li>- Permite derivação (<i>gentalha, gentarada</i>);</li> <li>- Flexibilidade sintagmática.</li> </ul>	<p><b>Propriedades de Forma</b></p> <ul style="list-style-type: none"> <li>- Perda da analisabilidade: a construção não é mais processada por [Det] mais [Nome] (processo cognitivo de domínio geral: <i>chunking</i>);</li> <li>- Perda das propriedades de nome: a construção não pode mais ser modificada por determinante, adjetivo, possessivo e quantificador; não permite mais flexão de número e grau; não permite mais derivação.</li> <li>- Não aceita material interveniente;</li> <li>- Especialização na posição de sujeito.</li> </ul>
<p><b>Propriedades de Significado</b></p> <ul style="list-style-type: none"> <li>- Propriedade referencial de nome: refere-se a um conjunto de pessoas.</li> </ul>	<p><b>Propriedades de Significado</b></p> <ul style="list-style-type: none"> <li>- Propriedade referencial de pronome: perde o caráter generalizante e ganha propriedade dêitica, passando a se referir à primeira pessoa do plural, concorrendo com o pronome <i>nós</i>.</li> </ul>

**Fonte:** elaboração própria, com base em Soares (2018).

Embora sob a mesma perspectiva dos MBU, nosso trabalho se diferencia do de Soares (2018) por investigar mudanças pós-construcionalização de *a gente* como forma pronominal que interagem com outros aspectos gramaticais do PB.

### 2.1.2 A emergência de *você*

Sob variados enfoques (histórico, sociolinguístico, da mudança), diversos autores, como Cintra (1972), Faraco (1996), Vitral (1996), Vitral e Ramos (2006), Menon (2006), Lopes e

Duarte (2007), Lopes e Rumeu (2007), Rumeu (2004; 2008), dentre outros, desenvolveram estudos aprofundados sobre a forma a emergência de *você* no quadro pronominal do PB.

Ao estudar historicamente as formas de tratamento em português a partir do contexto histórico-social da sociedade portuguesa do século XIII, Cintra (1972) postula que o português herdou do latim tardio um sistema bifurcado de referência à 2ª pessoa do discurso – *Tu* e *Vós*. Segundo o autor, esses pronomes se distinguiam pelo tratamento cortês para este, e íntimo para aquele. O pronome *vós* representou, até o começo do século XV, o preferido pela realeza no que diz respeito ao tratamento cerimonioso, que, entre os séculos XV e XVI, deixa de ser empregado como demonstração de polidez para com a realeza portuguesa. A explicação para esse fenômeno é vista em Cintra (1972), com base no qual Rumeu (2004) também oferece a seguinte uma explicação:

Os motivos para o abandono do pronome de tratamento *Vós* como uma demonstração de polidez para com a realeza portuguesa devem ser entendidos, [...] a partir do *contexto histórico-social* de Portugal nos séculos XV e XVI. Em termos sociais, há de se admitir que a *corte* e a *nobreza* foram as responsáveis pela adoção e posterior degradação das polidas formas nominais de tratamento cortês na sociedade portuguesa, visto que desejavam estabelecer, linguisticamente, uma diferenciação hierárquica entre as diferentes camadas da pirâmide social. Essa tentativa de consolidação de uma sociedade hierarquizada expressa o desejo de manutenção das *relações sociais de poder* entre a realeza portuguesa e seus súditos a partir da efetiva instauração de um relacionamento interpessoal *assimétrico*. Em termos linguísticos, observa-se que, em virtude do desuso do pronome *Vós* para o tratamento cerimonioso de interlocutor, surgiram formas nominais de tratamento e, entre elas, está a forma nominal de tratamento *Vossa Mercê* que, lenta e gradualmente, evoluiu até constituir a forma pronominal *Você* em português. (RUMEU, 2004, p. 28; itálicos no original).

De acordo com Rumeu (2004), a forma de tratamento *vossa mercê* era dirigida à pessoa do rei por fazer referência à sua benignidade. Etimologicamente, o lexema *mercê* em português originou-se do “substantivo latino *mercēs*, *-ēdis* com o sentido de *salário*, *prêmio*, assumindo, em língua portuguesa, conforme Moraes Silva (1813), Nascentes (1956) e Cunha (1982), o sentido de *graça*, *benefício*, *proteção*, *mercadoria*” (RUMEU, 2004, p. 29). É nesse sentido que Said Ali (1937) afirma que os súditos, ao apresentarem suas queixas e solicitações ao rei, pediam por sua *mercê* e “punham frequentemente *vossa mercê* por *vós*, referindo-se não à pessoa do soberano, e sim a graça e o favor que dele dimanava” (SAID ALI, 1937, p. 139 *apud* RUMEU, 2004, p. 29). Assim, a forma nominal de tratamento *vossa mercê* era empregada indiretamente para a segunda pessoa, pois a intenção dos súditos era, em primeiro lugar, pedir

a proteção e a graça que emanava do rei. Desse modo, a forma se cristalizou composta por um pronome possessivo de cerimônia *vossa* e o substantivo *mercê*.

A partir de 1490, conforme atesta Cintra (1972), a forma de tratamento *vossa mercê* deixa de ser empregada para se dirigir, única e exclusivamente, à pessoa do rei, descendo a hierarquia social e aplicando-se também a duques e infantes, e, em seguida, a fidalgos e, no século XVI, a pessoas da burguesia. Essa disseminação da forma *vossa mercê* nas classes mais baixas da hierarquia social portuguesa se deu devido à própria reorganização social de fim do século XV e início do século XVI, quando, além do clero e da nobreza, a burguesia também passou a compor a parte mais alta da pirâmide social portuguesa. Havia um desejo da monarquia de se distinguir linguisticamente das demais classes sociais, e, portanto, essa vontade se refletiu na reestruturação das relações interpessoais. A distinção da hierarquia social era tão prestigiada que, na segunda metade do século XV, foram promulgadas leis, que dispunham sobre o tratamento de cortesia, o que aconteceu primeiramente na sociedade espanhola, por Felipe II, e em 1597, por Felipe I, em Portugal. As formas nominais de tratamento de cortesia objeto das leis reservavam *Vossa Majestade* para o Rei e para Rainha, *Vossa Alteza*, para Príncipes e seus sucessores, *Vossa Excelência*, para sucessores dos Infantes e para o Duque de Bragança, e *Vossa Senhoria*, para o Clero e autoridades do Império português. O autor ainda informa que *vossa mercê* não era forma de tratamento especificada nas leis de cortesia, mas, àquela altura, ela já fora generalizada entre a nobreza e a burguesia funcionando como estratégia de tratamento formal que se opunha ao informal *tu*.

Menon (2006) evidencia também que, a partir do século XVI, a forma nominal *vossa mercê* se vulgariza, passando de um tratamento honorífico a um tratamento comum e, posteriormente, a tratamento vulgar. Na tentativa de demonstrar esse declínio do uso na escala social, a autora apresenta três momentos do percurso de vulgarização por que *vossa mercê* passou: (i) primeiro, os que estavam no escalão superior da hierarquia da corte começam a exigir essa forma de tratamento aos seus subordinados imediatos; (ii) havia dois tipos de nobreza, uma mais ligada à corte e outra mais distanciada, e esta passa a exigir o mesmo tratamento dispensado aos que não são nobres; (iii) por fim, os que não possuíam título de nobreza, mas possuíam dinheiro, reclamaram para si, de seus empregados ou de qualquer outra pessoa que estava abaixo na escala social, o tratamento respeitoso de *vossa mercê*.

Em Lopes (2001) e Lopes e Duarte (2003b), encontramos a trajetória da evolução de *vossa mercê* a *você* no português entre os séculos XVIII ao XIX. Segundo as autoras, a forma *vossa mercê*, em determinado momento, perde seu caráter cortês e passa a fazer referência à segunda pessoa do discurso, concorrendo com a forma informal *tu*. Lopes (2001) postula que, a partir do século XIX, a forma *você*, por um processo lento e gradual de mudança categorial, sofre mudanças sintáticas no que diz respeito à posição mais fixa na sentença, especializando na função sintática de sujeito em posição pré-verbal, em oposição a *vossa mercê* que tem maior variabilidade sintagmática. Lopes e Duarte (2003b) afirmam que, em fins do século XIX, na variedade brasileira do português, a forma *você* entra de forma robusta e crescente no quadro pronominal concorrendo com *tu*, concorrência iniciada já no século XVIII, porém de maneira tímida. As autoras evidenciam que, no século XVIII, o uso de *vossa mercê* é pouco frequente, e praticamente desaparece no século seguinte; a essa ausência significativa, as autoras atribuem a baixa frequência do uso de *vós*, que resvala para a dessemantização de *vossa mercê* em favor do aumento da frequência do uso de *você*. Rumeu (2004) resume os resultados aos quais as autoras chegam da seguinte forma:

Na tentativa de integrar a perspectiva variacionista discutida por Weinreich *et alii* (1968) aos princípios funcionais discutidos por Hopper (1991), as autoras chegaram às seguintes conclusões: 1<sup>a</sup>) nas relações simétricas, prevalecem, nas cartas *setecentistas* e *oitocentistas*, o uso mútuo do pronome *Tu* (T/T) – 72% –, contrapondo-se as relações sociais [- íntimas], nas quais se detecta maior frequência de *Vossa Mercê* – 61% – em relação as outras formas de tratamento – *Vossa Senhoria*, *Vossa Excelência* – 39%; 2<sup>a</sup>) em virtude do processo de *dessemantização* sofrido por *Vossa Mercê*, a sua contraparte gramaticalizada, *Você*, mostrou-se produtiva nas relações sociais assimétricas de superior para inferior, principalmente nas cartas pessoais; 3<sup>a</sup>) a forma nominal de tratamento *Vossa Mercê* mostrou-se frequente nos séculos XVIII e XIX nas relações sociais assimétricas de inferior para superior; 4<sup>a</sup>) a rigidificação da ordem SV, a *mistura de formas tratamentais* nas cartas pessoais e a presença de co-referentes de 2<sup>a</sup> pessoa fornecem indícios de que a gramaticalização de *Vossa Mercê* se inicia, ainda que lentamente, no século XVIII, implementando-se de forma mais acelerada no final do século XIX, principalmente em relação ao pronome pessoal *Vós*. (RUMEU, 2004, p. 32-33; itálicos no original).

Sob essa mesma perspectiva da mudança, Faraco (1996) indica que *vossa mercê* foi afetada por um rápido processo de simplificação fonética tornando-se os pronomes *você/vocês* de uso corrente hoje, e que essa simplificação fonética produziu diversas formas paralelas. Nascentes (1956) apresenta dezoito registros simplificados que resultaram da forma *vossa mercê*: “cê, mecê, mincê, ocê, oncê, sucê, suncê, vacê, vainicê, vancê, vansmincê, vassuncê, voncê, vosmecê, vossemecê, vosmincê, vossuncê, ucê” (NASCENTES, 1956, p. 119-121), além de *você*, podendo haver mais.

A evolução da forma *vossa mercê* a *you* parece não se interromper aí. É o que Vitral (1996) sugere ao apresentar argumentos a favor de que a forma pronominal tem passado por um processo de cliticização, isto é, *you* passa de pronome a *cê* clítico. Segundo Hopper e Traugott (1993), a cliticização é um processo mais geral de gramaticalização, em que duas propriedades se apresentam: (i) perda da autonomia lexical, sendo mais contextualmente dependente; (ii) tendência a “uma significação mais geral” (HOPPER; TRAUGOTT, 1993, p. 133) em que, conforme se vê em Croft (1990, *apud* HOPPER; TRAUGOTT, 1993), o significado referencial segue um percurso do tipo mostrado em (17).

(17) [+definido/+referencial] → [-definido/+referencial] → [-definido/-referencial]. (CROFT, 1990, *apud* HOPPER; TRAUGOTT, 1993, p. 157).

Vital (1996) objetiva, então, investigar a correlação entre perda de substância fonética e perda de conteúdo semântico no ciclo de gramaticalização, a fim de verificar que a forma *cê* seja a forma usada não referencialmente, chegando a expletivo, portanto. Nesse estudo, o autor constatou que a forma *you*, e não a forma *cê* ou zero, configura como expletivo, sendo um contraexemplo de que a perda de substância fonética e perda de conteúdo semântico fossem simétricas. Em estudos posteriores, Vital e Ramos (2006) e Othero (2013) revisitam o *status* do pronome *cê* e chegam à conclusão, a partir de testes sintáticos e prosódicos, de que *cê* não pode ser considerado um pronome tônico e tampouco um pronome clítico. Os autores evidenciam que *cê* tem um comportamento sintático e prosódico próprio: “ele é um pronome especializado para a função de sujeito do verbo em PB, que tem força acentual intermediária entre um tônico e um clítico” (OTHERO, 2013, p. 152). Othero (2013) advoga, então, que o pronome *cê* é um pronome novo na língua “provavelmente uma partícula com as mesmas propriedades prosódicas que outras formas inovadoras, como *es*, *eis*, *a’ente*, *ente*, *tá* e *num* (estudados por CASTILHO, 2010)” (op. cit., p. 152). De qualquer forma, segundo o autor, é matéria de interessante investigação porque é não usual, dentro dos estudos de mudança, constatarmos a emergência de uma nova forma em uma classe gramatical fechada, como é a classe de pronomes.

A partir dos trabalhos aqui referenciados, elaboramos o quadro a seguir no qual focalizamos as mudanças tanto de forma quanto de significado da construção reinterpretando como um processo de construcionalização, como se vê no quadro a seguir.

**Quadro 5** - Construcionalização de *você* no português

[vossa][mercê] – referência indireta à realeza	[você] – pronome de 2P
<b>Propriedades de Forma</b> - A construção é formada por um pronome possessivo de cerimônia [vossa] mais um nome [mercê]; - Variabilidade sintagmática.	<b>Propriedades de Forma</b> - Perda de material fonético e fonológico; - Especialização na posição de sujeito.
<b>Propriedades de Significado</b> - Propriedade referencial de pronome de tratamento: refere-se indiretamente à pessoa do rei/rainha.	<b>Propriedades de Significado</b> - Propriedade referencial de pronome pessoal: perde o caráter de tratamento cerimonioso e amplia seu caráter dêitico, passando a se referir à segunda pessoa do singular, concorrendo com o pronome <i>tu</i> .

Fonte: elaboração própria, com base em Rumeu (2008).

## 2.2 A representação do sujeito no PB na perspectiva da Sociolinguística Paramétrica

### 2.2.1 Os trabalhos de Duarte (1993, 1995)

As pesquisas de Duarte (1993, 1995; trabalhos posteriores) têm dado bastante atenção a um tipo de mudança que vem acontecendo na sintaxe do PB. Trata-se, segundo a teoria gerativa, da mudança do parâmetro *pro-drop* ou *parâmetro do sujeito nulo*. Duarte (1993) afirma que esse é um dos tópicos que mais se tem discutido dentro do quadro do gerativismo e que tem atraído inúmeras investigações que visam correlacionar a riqueza flexional dos paradigmas verbais de uma língua e a possibilidade de omissão do sujeito em sentenças finitas. Faz-se necessário trazer esses trabalhos neste capítulo, e os demais estudos que deles decorrem, porque, como a própria autora afirma, a relação entre a crescente preferência pelo sujeito pronominal pleno e a redução nos paradigmas flexionais, redução atribuída à perda da segunda pessoa direta *tu* e à franca entrada das proformas *você* e *a gente*, acarretou mudanças significativas na sintaxe do PB que o distanciam do português europeu (PE). Recorrer a esses resultados obtidos a partir de um outro quadro teórico justifica-se por duas razões: (i) a ausência de pesquisas no quadro dos MBU envolvendo mudanças sintáticas mais amplas para as quais

contribui o fenômeno da indeterminação do sujeito; (ii) o lançamento de um novo olhar teórico-metodológico, o dos MBU, sobre fenômeno semelhante.

Partindo de um *corpus* composto por trechos de peças de teatro popular escritas nos séculos XIX e XX, Duarte (1993) busca evidências que atestassem, para o PB, a relação entre a crescente preferência por sujeito pronominal pleno e a redução dos paradigmas flexionais. A autora constata que no PB há uma crescente simplificação dos paradigmas flexionais, que evolui de um sistema com seis formas distintivas, mais dois sincretismos (representados pela segunda pessoa indireta, que sempre emprega as formas verbais de terceira pessoa), para um paradigma de quatro formas, coexistente com um paradigma de apenas três formas. A autora atribui essa simplificação, como dito, à perda da segunda pessoa direta e à perda do pronome de primeira pessoa do plural *nós*, substituído na fala dos mais jovens por *a gente*. O quadro abaixo exemplifica essas mudanças.

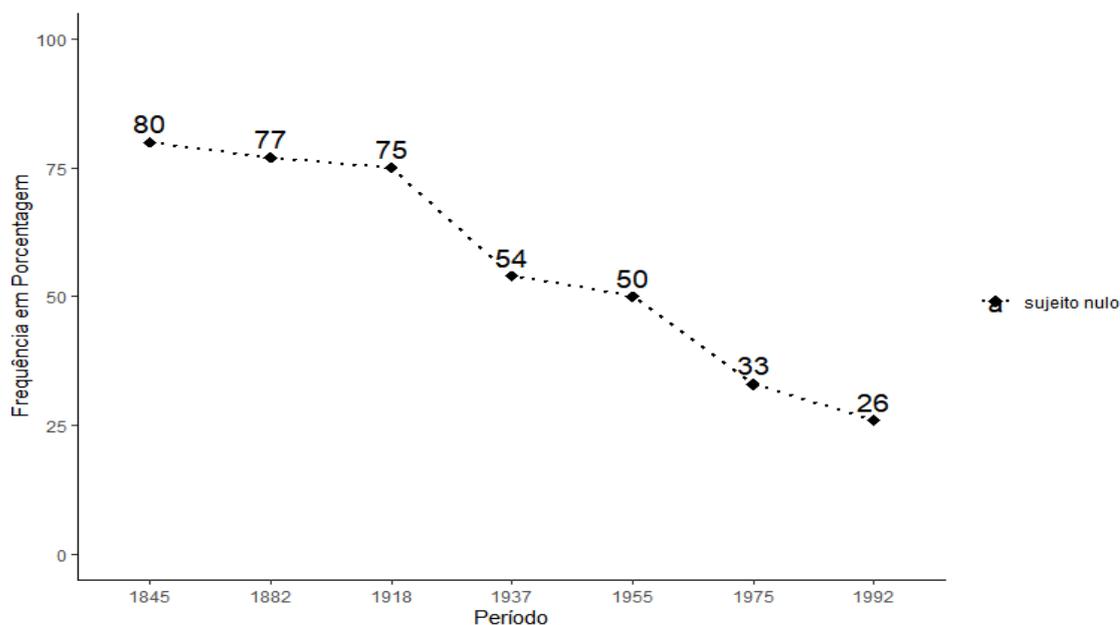
**Quadro 6** – Evolução dos paradigmas pronominal e flexional do português

Número	Pessoa	Paradigma 1	Paradigma 2	Paradigma 3
Singular	1ª ( <i>eu</i> )	<i>cant-o</i>	<i>cant-o</i>	<i>cant-o</i>
	2ª ( <i>tu</i> )	<i>canta-s</i>	.-	.-
	2ª ( <i>você</i> )	<i>canta-Ø</i>	<i>canta-Ø</i>	<i>canta-Ø</i>
	3ª ( <i>ele</i> )			
1ª. ( <i>a gente</i> )				
Plural	1ª. ( <i>nós</i> )	<i>canta-mos</i>	<i>canta-mos</i>	<i>canta-Ø</i>
	2ª ( <i>vós</i> )	<i>canta-is</i>	.-	.-
	2ª ( <i>vocês</i> )	<i>canta-m</i>	<i>canta-m</i>	<i>canta-m</i>
	3ª ( <i>eles</i> )	<i>canta-m</i>	<i>canta-m</i>	<i>canta-m</i>
		<i>6 formas</i>	<i>4 formas</i>	<i>3 formas</i>

Fonte: Duarte (1993, p. 109; com adaptação).

A autora, então, investiga se essa simplificação do paradigma flexional impactou a representação do sujeito. Diante dos resultados dessa investigação da autora expostos no gráfico 2, a seguir, ela considera que é inevitável não os associar à simplificação do paradigma flexional.

**Gráfico 2** - Ocorrência total de sujeitos nulos



**Fonte:** Duarte (1993, p. 112; com adaptações).

Duarte (1995) chega a resultados muito interessantes. No gráfico 2, reproduzido da autora, podemos observar que nos três primeiros períodos, referente ao uso do paradigma 1, o índice de ocorrência de sujeito nulo é bem superior ao do último período, quando estão em uso os paradigmas 2 e 3. Afirma a autora que é a partir do período de 1937, quando se perde a segunda pessoa direta, que ocorre súbita queda de sujeitos nulos. No último período, em 1992, a forma *a gente* é concorrente do pronome *nós*. A partir desses resultados, a autora observa que a mudança não atua uniformemente por todas as pessoas gramaticais e que o comportamento da terceira pessoa revela uma assimetria. Duarte (1995) assim se expressa:

Se, de um lado, parece estarmos perdendo a opção pelo sujeito nulo na representação dos sujeitos de primeira e segunda pessoas, de outro, continuamos a usufruir dessa opção para representar os de terceira, segundo o tipo de amostra utilizado. Tal fato, contudo, não invalida a hipótese levantada. De fato, a riqueza funcional do paradigma se perdeu, o que significa que teremos cada vez menos sujeitos nulos referenciais licenciados por Agr [concordância]. A primeira e a segunda pessoas são a clara evidência disso. Quanto à terceira, pode-se presumir que, não sendo mais inteiramente realizada através de Agr, a identificação do sujeito nulo esteja ancorada na sua coindexação com um SN numa posição acessível, seja no contexto discursivo, seja em estruturas subordinadas com sujeitos correferentes. (DUARTE, 1995, p. 21).

A partir dos resultados de suas pesquisas, Duarte chega a um entendimento de que, de fato, a redução no quadro de desinências verbais alterou significativamente a propriedade de

que o PB dispõe para o preenchimento ou não do sujeito. Duarte (1993) ressalta que essa não é uma mudança isolada a que o PB está se submetendo, ao contrário, faz parte de “um conjunto de mudanças sintáticas que permitem delinear a emergência de uma gramática brasileira” (DUARTE, 1993, p. 124). É nesse sentido que a autora afirma que a mudança na representação do sujeito de referência definida afetou também a representação dos sujeitos de referência arbitrária (ou indeterminada, como assumimos aqui) em sentenças tanto finitas quanto não-finitas (DUARTE, 2008).

### **2.2.2 O trabalho de Cavalcante (1999)**

O trabalho de Cavalcante (1999) tem por objetivo identificar as formas de indeterminação do sujeito (formas pronominais e usos do *se*) em orações finitas e não-finitas, em textos escritos da imprensa carioca nos dois últimos séculos. Segundo a autora, o estudo se justifica porque ela procura relacioná-lo a outros fenômenos de mudança em curso no PB, como, por exemplo, a preferência pelo preenchimento do sujeito de referência definida que acaba por afetar também a realização do sujeito de referência indeterminada. A hipótese da autora é a de que as mudanças que ocorrem na fala se refletem na variedade escrita padrão, “cedendo a pressões estruturais por que passa o nosso sistema pronominal” (CAVALCANTE, 1999, p. 40).

Cavalcante (1999) parte da ideia de que, se o quadro pronominal do PB tem passado por mudanças, pode-se prever que essas mudanças também estejam refletidas em outros tipos de sujeitos pronominais, não só de referência definida, mas também os de referência indeterminada (arbitrária). O modelo de análise adotado pela autora é o da Sociolinguística laboviana associado ao da Gramática Gerativa, e o *corpus* por ela utilizado é constituído de editoriais, textos opinativos e crônicas publicados em jornais da cidade do Rio de Janeiro, entre 1848 e 1998, e por ela organizados por períodos: *I – 1848-1869; II – 1891-1910; III – 1935-1942; IV – 1964-1968; V – 1996-1998*. Dos resultados da autora, apresentamos nesta seção apenas aqueles que tratam da distribuição das estratégias de indeterminação em sentenças finitas e não-finitas durante os períodos, pois é o que mais interessa a nosso trabalho. É da autora os resultados apresentados na tabela 1 com a distribuição das formas de indeterminação do sujeito através dos séculos em sentenças finitas.

**Tabela 1** - Distribuição de estratégias de indeterminação em sentenças finitas por período de tempo

<b>Período</b>	<b>Período I (1848-1869)</b>	<b>Período II (1891-1910)</b>	<b>Período III (1935-1942)</b>	<b>Período IV (1964-1968)</b>	<b>Período V (1996-1998)</b>
<b>Pronome</b>					
<i>a gente</i>	1%	1%	1%	0,5%	8%
<i>eles</i>	7%	16%	2%	7%	8%
<i>nós</i>	47%	37%	43%	31%	18%
<i>se</i>	45%	46%	55%	61%	62%
<i>você</i>	-	-	-	0,5%	4%

**Fonte:** Cavalcante (1999, p. 66; com adaptações).

Como observa Cavalcante, é possível notar em seus resultados diferentes estratégias de indeterminação além daquelas prescritas pela norma gramatical, observação que é relevante, pois a autora lida com a modalidade escrita padrão, e seus resultados mostram que mesmo a língua escrita se distancia das prescrições gramaticais. A partir da tabela, podemos verificar que, no período I, há uma forte concorrência entre as formas *nós* e *se*, que começa a se desfazer, numa inversão, a partir do período II, em que o uso de *se* cresce e o de *nós* decresce. Segundo a autora, esse decréscimo no uso de *se* pode ter sua explicação buscada na redução do paradigma flexional do PB, conforme Duarte (1993; 1995) atesta. Como mostrado na seção anterior, Duarte (1993) evidencia que, nos dois últimos períodos, houve uma redução de uso do pronome *nós* com referência definida e sua substituição crescente por *a gente*, principalmente entre a faixa etária mais jovem. Desse modo, Cavalcante (1999) argumenta que essa redução do uso de *nós* com referência definida pode estar afetando também seu uso em contextos de referência arbitrária.

Em relação ao uso de *se* na escrita padrão, a autora constata que seu aumento pode estar relacionado a uma outra mudança sintática por que passa o PB, como é o caso, por exemplo, da redução no uso de clíticos acusativos (cf. CYRINO, 1994; DUARTE 1989). “Com o uso de *se* apassivador, elimina-se automaticamente o uso do clítico acusativo”, sustenta a autora (CAVALCANTE, 1999, p. 67).

No que diz respeito ao uso de *eles*, a autora argumenta que a terceira pessoa do plural é estratégia preferida para indeterminar, porque exclui o falante. Segundo Cavalcante (1999), todas as ocorrências dessa estratégia nos editoriais do último período provêm de textos narrativos. A forma *a gente* se mantém com percentuais bem baixos durante todos os períodos. A forma *você*, a mais frequente das formas de indeterminação na fala, só apareceu, aponta a

autora, a partir do período IV com apenas 2 dados (0,5%) chegando a 4% no período V (21 dados). Esse resultado, advoga a autora, revela que há uma implementação na escrita padrão de estratégias mais usadas na fala.

A distribuição das estratégias de indeterminação nas sentenças não-finitas pode ser observada na tabela 2 abaixo.

**Tabela 2** - Distribuição de estratégias de indeterminação em sentenças não-finitas por período de tempo

<b>Período</b>	<b>Período I (1848-1869)</b>	<b>Período II (1891-1910)</b>	<b>Período III (1935-1942)</b>	<b>Período IV (1964-1968)</b>	<b>Período V (1996-1998)</b>
<b>Estratégias</b>					
gerúndio com <i>se</i>	4%	8%	8%	3%	6%
infinitivo com <i>se</i>	23%	11%	16%	10%	19%
pronominais	5%	7%	7%	7%	8%
sem <i>se</i>	68%	74%	69%	80%	67%

**Fonte:** Cavalcante (1999, p. 84; com adaptações).

Com os resultados obtidos, a autora nota que a preferência pela posição vazia de sujeito é a mais frequente em todos os períodos. Os resultados apresentados na tabela revelam que o uso dessa estratégia se encontra em variação estável com as demais estratégias. A segunda estratégia mais usual é a com o *se* em sentenças infinitivas, que, segundo a autora, está em distribuição complementar com a posição vazia, "ou seja, quando uma diminui, a outra automaticamente aumenta" (CAVALCANTE, 1999, p. 84). Entretanto, a hipótese da autora era a de que haveria um crescente aparecimento desses usos ao longo dos períodos, o que não se confirmou. Essa hipótese foi aventada a partir do trabalho de Nunes (1990), que constatou o aumento do uso de *se* em sentenças infinitivas. Insta dizer que mais adiante a autora mostra que o que se ampliou foram os contextos que favorecem tais usos.

As formas *se* com gerúndio e pronominais apresentam índices bastante discretos, e demonstram, para a primeira, estratégia pouco usada e muito provavelmente de uso bem específico, e para a segunda, um aumento, mesmo que pequeno, dos usos das formas *a gente*, *nós*, *eles* e *você*. A autora entende que essa subida pode indicar um alargamento dos contextos de uso dessas formas.

Com o objetivo de observar as diferentes estratégias de indeterminação (formas pronominais e usos do clítico *se* em sentenças finitas e não-finitas) na escrita padrão em textos

jornalísticos, que circularam entre os séculos XIX e XX, a autora conclui, dentre outros resultados, que, a partir do período III (1935-1942), quando o início da mudança começa a se delinear em direção ao preenchimento do sujeito nas peças de teatros populares (DUARTE, 1993), nos textos jornalísticos ocorre um enrijecimento da norma: “há a diminuição do uso de sujeitos plenos e da não-concordância entre verbo e argumento interno plural nas construções com *se*, ao passo que há o aumento do uso de *se* e de *nós*” (CAVALCANTE, 1999, p. 108). Assim, a autora argumenta que, se é possível verificar que na língua padrão ainda há um maior prestígio da norma prescritiva, seus resultados mostram indícios de implementação de uma mudança já atestada para a fala nas formas pronominais, principalmente nas crônicas. O entendimento final é o de que: (i) a implementação e a propagação da mudança do parâmetro do sujeito nulo encontram lugar primeiramente nas crônicas; (ii), de modo geral, é confirmada a hipótese central de que as mudanças que acontecem primeiramente na língua falada estão por afetar a língua escrita padrão, inclusive no que diz respeito à regra de uso de sujeitos de referência arbitrária.

### **2.2.3 O trabalho de Vargas (2010)**

Vargas (2010), também sob viés da Sociolinguística paramétrica, busca investigar a distribuição de estratégias de indeterminação do sujeito entre os séculos XIX e XX, mais especificamente entre os anos de 1845 e 1992, mesmo período retratado por Cavalcante (1999), mas buscando evidências empíricas em textos de peças de teatro de cunho popular, gênero que busca representar a fala da época que retrata. Também motivada pelos trabalhos de Duarte (1993, 1995), a autora busca evidências para a implementação da mudança em direção aos sujeitos de referência indeterminada plenos, especialmente observando os usos das formas pronominais nominativas, como possíveis substitutos do uso de *se* indefinido (apassivador/indeterminador) e de sujeito nulo com verbo na terceira pessoa do plural. A autora visa, portanto, investigar se a realização do sujeito pleno de referência indeterminada acompanha o preenchimento dos sujeitos de referência definida no PB. Assim como na seção anterior, apresentaremos apenas os resultados relacionados à distribuição da realização desses sujeitos plena ou nula, pois é o que nos interessa nessa discussão.

A autora parte de um conjunto de 958 dados de estratégias com sujeitos indeterminados em sentenças finitas e não-finitas, distribuídas pelos sete períodos considerados, conforme Duarte (1993; 1995): I – 1837-1847; II – 1862-1882; III – 1918-1920; IV - 1933-1938; V –

1945-1955; VI – 1975-1984; VII – 1990-1992. Os dados em sentenças finitas apresentam um total de 859 ocorrências e nas sentenças não-finitas um total de 99 ocorrências. A tabela 3 traz a distribuição dos sujeitos de referência arbitrária por período de tempo.

**Tabela 3** - Distribuição de sujeito de referência [+arb] por período de tempo em sentenças finitas

Estratégias \ Período	Período							
	Período I (1837-1847)	Período II (1862-1882)	Período III (1918-1920)	Período IV (1933-1938)	Período V (1945-1955)	Período VI (1975-1984)	Período VII (1990-1992)	
<i>se</i>	40%	56%	52%	62%	36%	5%	4%	
<i>eles</i>	58%	38%	35%	29%	45%	33%	32%	
<i>nós</i>	1%	1%	5%	4%	5%	13%	1%	
<i>a gente</i>	1%	4%	11%	6%	16%	23%	42%	
<i>você</i>	-	-	-	-	-	26%	18%	
<i>zero</i>	-	-	-	-	-	-	2%	

**Fonte:** Vargas (2010, p. 60; com adaptações)

A autora chama a atenção para a competição acirrada, até o período V (1945-1955), entre as formas alternantes de indeterminação representadas por *se* e pela terceira pessoa do plural, referida como *eles*. Até o primeiro período (1837-1847), a terceira pessoa lidera a preferência no uso em relação a *se*, situação que se inverte a partir do período seguinte (1842-1882). É a partir do período V que a estratégia de terceira pessoa supera a do clítico *se*. Essa competição fica mais clara, aponta Vargas, quando se verifica que grande parte das estruturas com a terceira pessoa do plural opera indeterminação que exclui o falante, podendo ser substituída pelo uso de *se*.

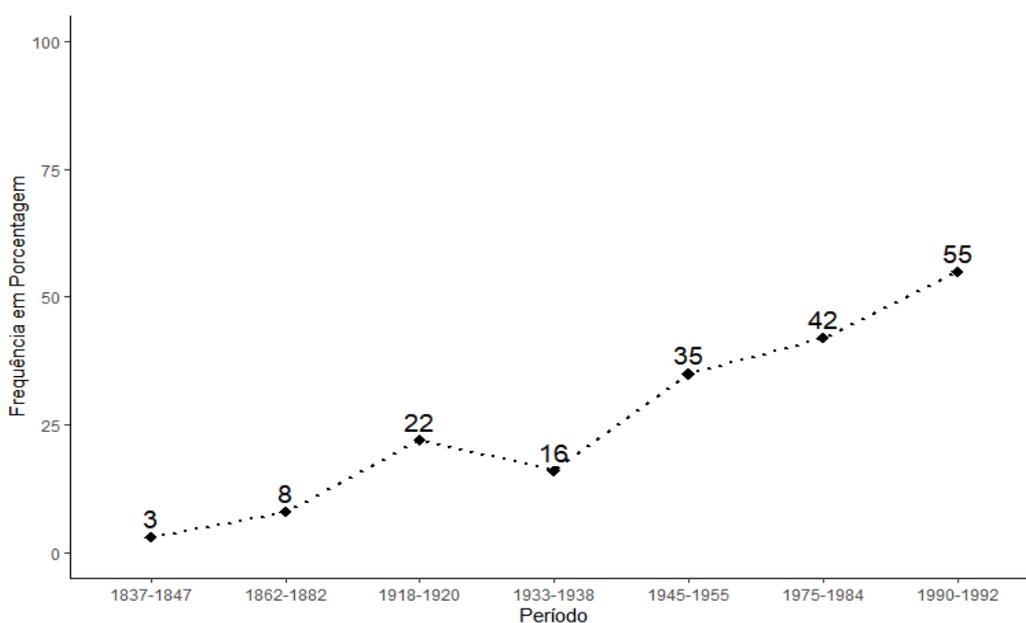
A estratégia de primeira pessoa do plural apresenta índice bem tímido nos primeiros períodos, com *a gente* apresentando ligeira vantagem sobre *nós*. A estratégia com *a gente* se acentua a partir do período V (1945-1955), divide a preferência com a estratégia emergente *você*, no período VI (1975-1984), e chega ao período final (1990-1992) como a preferida sobre todas as demais, resultado interpretado pela autora como devido à mudança que se verifica na fala, como já haviam constatado Duarte (1993; 1995) e Cavalcante (1999).

A forma de terceira pessoa do singular retomada por posição vazia, não associada a qualquer índice de indeterminação (a que a autora se refere como *zero*), aparece apenas no período VII, com apenas 2%. Segundo a autora, Duarte (1995; 2003) e Cavalcante (2006) já

constatarem que essa é uma estratégia empregada em contextos bem específicos, como relatos de procedimentos ou de manifestação das categorias de aspecto e modalidade.

Vargas (2010) confirma a hipótese de que as peças mais antigas apresentariam índices maiores do uso de *se* e de terceira pessoa do plural. Confirma também que a entrada de outras estratégias de indeterminação acompanha as mudanças relacionadas à redução do paradigma flexional do PB (DUARTE,1993, 1995). Com vistas a verificar se a representação plena de sujeitos com referência arbitrária acompanha a representação plena de sujeitos com referência definida, de que trata Duarte (1993, 1995), a autora apresenta o gráfico reproduzido aqui como gráfico 6, que mostra a distribuição geral das formas plenas de indeterminação, representadas por *eles, nós, a gente e você*.

**Gráfico 3** - A representação plena das estratégias de indeterminação eles, nós, a gente e você



Fonte: Vargas (2012, p. 67; com adaptações).

O gráfico que a autora apresenta é bem eloquente, ao evidenciar baixa frequência de uso das formas pronominais operando indeterminação de sujeito nos períodos iniciais e aumento considerável da frequência a partir do período V (1945-1955). A autora relaciona esses dados aos de Duarte (1993), que podem ser consultados no gráfico 2, na seção 3.1<sup>9</sup> deste trabalho. Na

<sup>9</sup> Os dados devem ser interpretados a partir da diferença entre os valores do gráfico 3 (VARGAS, 2010) e do gráfico 2 (DUARTE, 1993).

comparação desses resultados, verifica-se que, no período V (1945-1955), há uma competição bem acirrada entre os sujeitos nulos e plenos de referência definida, com índices de 50% para cada um, o que não acontece com os sujeitos de referência arbitrária, que começam a ter uma competição mais acentuada no período VI (1975-1984), quando a estrutura de sujeitos definidos plenos ultrapassa a de sujeitos definidos nulos com índices de 67%, o que acontece com os sujeitos de referência arbitrária apenas no último período (1990-1992), com índices de 55%. Com isso, a autora evidencia que a escolha por realizar o sujeito pleno ocorre primeiramente com os sujeitos de referência definida, com uma diferença de dois períodos, enquanto a competição acirrada entre o sujeito nulo e as formas *se*, *eles*, *a gente* e *você*, só acontece no último período. Assim, a autora demonstra o encaixamento dessas mudanças e como através do tempo se dá sua implementação no sistema.

Os trabalhos apresentados nesta seção serviram para mostrar a evolução na mudança da representação do sujeito no PB. Como esta pesquisa se dedica às estratégias pronominais de indeterminação do sujeito, foi necessário apresentar desde os estudos pioneiros de Duarte (1993, 1995) sobre a mudança no Parâmetro do Sujeito Nulo no que diz respeito à evolução na realização dos sujeitos plenos de referência definida. A partir disso, os trabalhos de Cavalcante (1999) e Vargas (2010) são de importante referência aqui por terem tratado especificamente das estratégias de indeterminação do argumento sujeito e terem demonstrado que a realização plena acompanha a mudança por que passa os sujeitos de referência definida.

## **2.3 A representação do sujeito no PB na perspectiva da Funcionalista**

### **2.3.1. Os trabalhos de Paredes Silva (1988, 2003)**

Como se pôde ver nas seções anteriores, a investigação sobre o fenômeno do sujeito pronome no PB é um tema muito discutido no âmbito da linguística formal, principalmente de orientação chomskyana. Por se tratar de um tema que suscita relevantes indagações, Paredes Silva (1988, 2003), alinhada a uma outra corrente predominante na linguística, a linguística funcional, associando-a ao tratamento variacionista laboviano, investiga o fenômeno variável do pronome sujeito no PB, no que diz respeito a sua presença e ausência, defendendo a posição de que a escolha da realização plena do sujeito e seu apagamento se correlacionam a motivações discursivo-funcionais, especialmente a variáveis linguísticas como *conexão discursiva*, *ambiguidade*, *ênfase ou contraste* e *número gramatical*. Desse modo, tomando a referência como um domínio funcional, a autora investiga como esse domínio requer, pressionado por

fatores de natureza funcional, maior ou menor codificação gramatical, nos termos de Givón (1995b), baseado no grau de previsibilidade da informação. Sob essa perspectiva, a autora busca comprovar que “quanto mais previsível o referente de um sujeito, menor a necessidade de explicitá-lo” (PAREDES SILVA, 2003, p. 97).

Em sua tese de doutoramento, Paredes Silva (1988) investiga o fenômeno aqui delineado a partir de amostras de escrita informal e chega a conclusões que evidenciam um comportamento bastante diferenciado entre as pessoas gramaticais. A primeira pessoa apresenta índices altos de omissão do pronome, enquanto a segunda pessoa e a terceira pessoa apresentam taxas elevadas de sujeito explícito, sendo a segunda pessoa a que ocupa a primeira posição em termos de frequência de preenchimento do sujeito.

Em trabalho posterior, Paredes Silva (2003) retoma a discussão de sua tese a fim de realizar um estudo do tipo Tendência e do tipo Painel (cf. LABOB, 1994), utilizando, dessa vez, um *corpus* de língua falada. É sobre esse trabalho que passamos a discorrer nesta seção, apresentando os fatores internos que foram controlados pela autora, quais sejam: *conexão discursiva, tipo de oração, contraste ou ênfase, ambiguidade e número gramatical*.

Paredes Silva (2003) evidencia que uma série de estudos desenvolvidos já na década de 1980 postula que a variável *mudança de referência* é fator de alta relevância para os estudos sobre expressão do sujeito pronominal. O que esses diversos estudos mostram, como já é bastante difundido nos estudos sobre a temática, tanto para o português, quanto para o espanhol (BENTIVOGLIO, 1980; SILVA-CORVALÁN, 1982; LIRA, 1982), é que a motivação para a escolha de uma estrutura com o sujeito explícito pelo pronome tem estreita ligação com a não-manutenção do mesmo referente como sujeito. Entretanto, desde sua tese, a autora parte de um entendimento de que é necessário investigar mais detalhadamente o contexto discursivo anterior, com o objetivo de dar maiores explicações acerca da relação entre um referente e sua menção prévia no discurso e procede assim por compreender que existem distinções mais imperceptíveis que não são bem captadas pela simples verificação do referente/sujeito anterior. A partir dessa constatação, Paredes Silva (2003) propõe, então, como variável uma escala de *conexão discursiva*, ao invés de *mudança de referência*, como estudos anteriores vinham realizando. Sua argumentação vai na direção de que a variável *conexão discursiva* é mais apropriada por dar conta de fenômenos mais amplos, tais como continuidades ou não de ações, manutenção ou não do plano discursivo e a própria mudança de tópico discursivo,<sup>10</sup>

---

<sup>10</sup> Tópico discursivo é entendido pela autora no sentido do assunto ou tema (PAREDES SILVA, 2003, p. 104).

diferentemente do controle de *mudança de referência*, capaz apenas de distinguir a identidade ou não de referentes em orações adjacentes. Os conceitos referentes a cada tipo de grau dessa escala são fornecidos no quadro 7.

**Quadro 7** - Escala de conexão discursiva

Nível na escala	Definição
Grau 1	Corresponde à permanência, na função de sujeito, do mesmo referente/tópico, no mesmo plano discursivo (manifestado pela manutenção do sistema de tempo-aspecto-modo verbal).
Grau 2	Indica uma mudança de plano: embora se mantenha o mesmo referente/tópico como sujeito, há uma mudança de plano discursivo, do tipo figura-fundo, fato-opinião, realidade-irrealidade, refletida em modificações no sistema de tempo-aspecto-modo das orações adjacentes.
Grau 3	Considera-se a manutenção do mesmo referente no sujeito, apesar da mudança de turno.
Grau 4	A interferência de orações de sujeito impessoal entre o sujeito em causa e sua última menção representa uma quebra na continuidade de um referente como sujeito, mas não chega a afetá-lo muito.
Grau 5	Trata da interferência de orações com sujeitos de outras pessoas gramaticais, estes sim mais capazes de desviar a atenção do referente/tópico previamente mencionado.
Grau 6	Retomada, na função do sujeito, de referente que ocorreu anteriormente em outra função. Trata-se, portanto, de mudança de função, pela passagem a tópico.
Grau 7	A mudança de tópico/subtópico discursivo representa um desvio nos rumos do discurso, ou mesmo uma digressão.

**Elaboração:** Paredes Silva (2003).

A hipótese da autora por trás da variável *conexão discursiva* é a de que a necessidade de explicitação do sujeito é mitigada à medida que a conexão entre um referente/sujeito e sua menção prévia se dilui no contexto discursivo, medida a partir da escala apresentada no quadro 7, cujos níveis estão qualificados em graus que vão de 1 a 7, em que, no grau 1, a conexão é considerada “ótima”, e no outro extremo da escala, grau 7, a conexão é considerada “não ótima”, comportando, portanto, graus intermediários.

A autora apresenta, em seu estudo de Tendência e de Painel, dados relativos à primeira pessoa. De todas as variáveis analisadas por ela, a variável *conexão discursiva* se revelou o condicionante mais relevante, selecionado em primeiro lugar pelas rodadas do VARBRUL. Paredes Silva (2003) observa que, na passagem do grau 1 ao grau 2, os sujeitos expressos praticamente dobram, mesmo que ambos apresentem os mesmos referentes. Essa constatação coloca em relevo que somente olhar para a identidade do referente anterior não basta, e que é preciso levar em conta a natureza discursiva. Os resultados a que a autora chega corroboram sua hipótese e evidenciam que, de fato, levar em consideração a natureza da sequência discursiva é uma via mais promissora para a investigação do fenômeno variável do pronome sujeito.

O dado em (18), da autora, demonstra que a mudança de um plano em que fatos são expressos para um plano em que se expressa um desejo “traz alguma descontinuidade na sequência capaz de justificar o aumento da expressão pronominal do sujeito” (PAREDES SILVA, 2003, p. 107).

(18) Eu sempre morei em casa, nunca morei em apartamento. *Estou* satisfeita. Eu não queria morar em apartamento não [Nad00: 2,44]. (PAREDES SILVA, 2003, p. 105).

Em relação à segunda variável, *tipo de oração*, a autora parte de uma classificação sintática, mas atribui, a partir da perspectiva da teoria da estrutura retórica de Mathiessen e Thompson (1988), uma interpretação funcional, entendendo que a combinação de orações reflete a estrutura retórica do discurso. Nesse grupo, há seis fatores de tipo de oração selecionados pela autora: *independentes, coordenadas não iniciais, principais antepostas, principais pospostas, subordinadas e adjetivas*. O resultado para essa variável é o seguinte: as percentagens do tipo *coordenadas e independentes* divergem consideravelmente, mas, como atesta a autora, os pesos relativos guardam certa proximidade, revelando os contextos em que a expressão do sujeito é fraca; as *adjetivas*, por outro lado, encontram-se no outro extremo, apresentando maior expressão de sujeitos realizados. A esse resultado, a autora atribui o fato de que as orações adjetivas apresentam informações suplementares, de segundo plano, conforme a estrutura retórica.

Em relação à variável *contraste* ou *ênfase*, também de natureza discursivo-funcional, a autora mostra que seus fatores apresentam pesos relativos extremamente consistentes. Como o nível prosódico não é abordado em sua pesquisa, ela adota como critérios a presença de marcas formais no discurso para considerar um uso contrastivo, tais como conectivos contrastivos, antonímias entre verbos, oposições no espaço e no tempo etc. O que fica evidente a partir dessa variável é que o papel contrastivo do pronome se revela constante no aumento da expressão do sujeito explícito.

Sobre a variável *número gramatical*, a autora destaca que essa é uma variável que se mostra mais influente para a primeira pessoa, justamente por conta da alternância *nós/a gente*. Desse modo, a argumentação vai na seguinte direção: a proforma *nós*, por acarretar uma flexão verbal bem marcada, favorece a ausência do sujeito; a forma inovadora *a gente*, por sua vez,

favorece um contexto de possível ambiguidade referencial, pois apresenta uma flexão verbal de terceira pessoa, sendo, portanto, necessário sua explicitação.

Os trabalhos de Paredes Silva (1988, 2003) são pertinentes por apontar que é igualmente relevante tratar do fenômeno de expressão do sujeito levando em consideração motivações de natureza discursivo-funcional. A maior contribuição desses trabalhos, a nosso ver, e dada a seleção da variável *conexão discursiva* como a mais relevante dentre as variáveis controladas, é a de que compreender mais a fundo a escolha por uma estrutura com sujeito expresso ou não, aparentemente um fenômeno de natureza morfossintática, depende sobretudo de fenômenos de ordem discursivo-pragmática cujas pressões se revelam na escolha do falante por uma estrutura ou outra, considerando a predizibilidade da informação, conforme postula a autora.

### 2.3.2. O trabalho de Guerra (2017)<sup>11</sup>

O trabalho de Guerra (2017) trata, em perspectiva diacrônica, do grau de transparência do sistema de referência por expressão pronominal e desinencial do argumento-sujeito de 1ª e 2ª pessoas no português brasileiro. Esse trabalho, que dialoga, em linhas gerais, com a nossa pesquisa, tem por objetivo analisar a variação diacrônica do grau de transparência da expressão do sistema de referência do sujeito de 1ª e 2ª pessoas no PB, correlacionando-a a três mudanças, quais sejam: (i) o aumento da frequência de sujeito pronominal explícito; (ii) o aumento do uso do pronome *você*; e (iii) o aumento do uso do pronome *a gente*. Escolhido o modelo da Gramática Discursivo Funcional (GDF) (HENGEVELD; MACKENZIE, 2008), a autora investiga como, na diacronia, o grau de transparência e opacidade varia através do tempo, motivado por questões funcionais, como a *iconicidade* e *economia*.

Importante ressaltar que os termos *transparência* e *opacidade* são utilizados com sentidos bastante típicos da teoria à qual a autora se filia. O modelo da GDF, cuja arquitetura é bastante elaborada em termos de camadas de representação do componente gramatical, a questão de transparência e opacidade deve ser medida a partir da correspondência entre as camadas, ou até mesmo dentro da mesma camada. Considerando que a estrutura da oração é composta por quatro níveis de interface, organizados hierarquicamente, a GDF distingue, então, os seguintes níveis, do mais alto na arquitetura, ao mais baixo,: nível interpessoal (NI), nível representacional (NR), nível morfossintático (NM), e o nível fonológico (NF). É a partir do

---

<sup>11</sup> Para uma análise mais bem detalhada da tese de Guerra (2017), remetemos o leitor à resenha de Tescari Neto (2019) sobre esse trabalho da autora, publicada na revista *Letras Escreve*.

cruzamento entre as representações das orações em cada nível que a transparência e opacidade devem ser investigadas. Desse modo, baseado no objeto de análise da autora, o fenômeno de referência pronominal e desinencial do argumento-sujeito de 1ª e 2ª pessoas recebe atributos nos três primeiros níveis, NI, NR e NM, sendo codificados como subato de referência, argumento, pronome e morfologia flexional verbal em cada um desses níveis, respectivamente.

A hipótese aventada pela autora é a de que o grau de transparência do sistema de expressão do sujeito pronominal e desinencial de 1ª e 2ª pessoas segue uma mudança não direcional, e, portanto, varia a depender de motivações comunicativas relacionadas à economia e à iconicidade. A hipótese é confirmada. Guerra evidencia que o grau de transparência do PB varia ora em direção à diminuição, ora em direção ao aumento de frequência da transparência.

Com o intuito de verificar que tipo de fenômeno interferiria na transparência/opacidade, Guerra (2017) investiga, primeiramente, como o aumento na expressão do argumento sujeito se relaciona a esse tema. Nesse ponto, o trabalho da autora é de crucial importância para este trabalho, pois, a partir de uma visão da linguística funcional, a autora demonstra com seus dados a mudança por que o PB vem passando: de uma frequência que gira em torno de 10% a 20% de sujeitos expressos no século XIX passa a expressivos 70%, aproximadamente, de sujeitos pronominais explícitos no século XXI. Com esse resultado, o trabalho de Guerra (2017), além de confirmar o que trabalhos anteriores já vinham discutindo sobre o fenômeno, soma-se também à lista de importantes pesquisas que delineiam a emergência de uma gramática brasileira (ROBERTS; KATO, 1993).

Outro ponto de contato entre este trabalho e o da autora é sobre a expressão de sujeito por meio das formas pronominais *a gente* e *você*. A autora investiga como o PB passou a preferir o uso de *você* em detrimento de *tu*, e de *a gente* em detrimento de *nós*. Para a variação diacrônica nas frequências de uso de *tu* e *você*, a autora mostra que a frequência de uso de *tu* tem uma queda de 100% de uso no século XXI,<sup>12</sup> enquanto a frequência de uso de *você*, do século XIX para o século XXI, tem uma subida para 81%. Para as formas *nós/a gente*, os resultados diferem um pouco: no século XIX, o pronome *nós* apresenta uma frequência de uso de 96,5%, caindo para 39% no século XXI; a forma inovadora *a gente* apresenta percentagem de 3,5% no século XIX e sobe para 61% no século XXI.

---

<sup>12</sup> A autora entende que a frequência nula do pronome *tu* no século XXI em seu estudo não significa desuso definitivo; significa, apenas, que a frequência de uso desse pronome teria caído de forma muito expressiva. (GUERRA, 2017, p. 119).

Como mencionado anteriormente, Guerra (2017) consegue demonstrar como diversas mudanças na gramática do PB estão relacionadas. Nosso trabalho, busca, na mesma esteira da autora, demonstrar que isso de fato se confirma, agora com o olhar para o fenômeno de indeterminação do sujeito e filiados a um outro quadro teórico que tem se mostrado promissor para investigações de fenômenos linguísticos.

### **2.3.3. O trabalho de Oliveira (2018)**

Oliveira (2018) analisa a expressão do sujeito com referente identificável no PB a partir do arcabouço teórico da GDF. As questões tratadas no trabalho da autora tangenciam a discussão apresentada por Guerra (2017), uma vez que Oliveira analisa o grau de transparência e opacidade e tem por objetivo mostrar que o PB está desenvolvendo um sistema mais transparente em comparação com sincronias mais pretéritas e com outras línguas românicas. A diferença entre os dois trabalhos está no fato de que, enquanto Guerra (2017) faz um estudo diacrônico, considerando a referência de pronomes de 1ª e 2ª pessoas, Oliveira (2018) realiza um estudo sincrônico, observando todas pessoas gramaticais.

Partindo de uma série de estudos realizados sobre o quadro pronominal, como os já apresentados neste trabalho, Oliveira (2018), a partir de 1500 sentenças coletadas no banco de dados Iboruna (GONÇALVES, 2007), confirma que a estrutura preferida pelos falantes é a expressa pela marcação verbal combinada com um pronome expresso, contabilizando 61% dos dados, enquanto a estrutura realizada apenas pela marcação verbal tem índices de 39%. Observando a distribuição da expressão do sujeito entre as pessoas gramaticais, a autora postula que essa diferença também é verificada nas diferentes pessoas gramaticais, exceto a referência do pronome *nós*, sujeito especificado pronominalmente combinado com morfologia verbal, com índices de 75% de expressão apenas pela morfologia verbal. Desse modo, a autora coloca em relevo que no PB há duas possibilidades para codificar sujeitos: referência cruzada ou referência única. Entretanto, a autora argumenta que a forma mais frequente no PB é aquela em que o sujeito é especificado por meio da referência única. Quando o referente é especificado pelo pronome, a autora chama a atenção para os dados que apresentam morfologia zero de terceira pessoa no verbo, como é o caso dos pronomes *você* e *a gente*, por exemplo. A autora assevera que esse tipo de referência conta com apenas um subato referencial realizado pelo pronome, evidenciando uma situação de transparência, uma vez que existe uma relação de um-para-um entre todos os níveis gramaticais envolvidos. A autora entende que esses casos não

podem ser tratados como concordância, uma vez que a forma verbal zero, de terceira pessoa, não contém informação sobre pessoa ou número e não pode ser acionada por mecanismo de cópia.

Dessa forma, a autora apresenta os seguintes dados para defender que o PB prefere estruturas mais transparentes: referência única – 73%, referência cruzada – 27%. Segundo a autora, a referência cruzada é usada apenas para a 1ª pessoa do singular e plural e para a 3ª pessoa do plural; a referência única, por sua vez, é atestada em todas as demais pessoas gramaticais, sendo o meio mais frequente para expressar o sujeito, diz a autora. Oliveira postula, então, que a referência cruzada está se tornando a estratégia marcada para a realização do sujeito no PB, o que confirma a hipótese inicial da autora de que o PB está adotando um sistema mais transparente para a referência de sujeitos pronominais.

## CAPÍTULO III – PROCEDIMENTOS METODOLÓGICOS

Em concordância com o aparato teórico que sustenta as análises feitas neste trabalho, é imprescindível que as ocorrências sejam buscadas em um *corpus* baseado em instâncias de uso real da língua. Este trabalho busca caracterizar o fenômeno de estratégias pronominais de indeterminação do sujeito realizadas pelas construções pronominais *a gente* e *você* do PB contemporâneo falado no interior paulista e, portanto, é de natureza sincrônica. Entretanto, por se tratar de um trabalho que se insere dentro dos estudos da mudança linguística, é essencial que se recorra à diacronia, pois só assim a mudança pode ser atestada. Desse modo, a metodologia do trabalho contempla duas etapas: (i) numa primeira fase, investigamos os contextos de mudança por que passaram as construções *a gente* e *você* recorrendo a trabalhos diacrônicos já amplamente documentados na literatura (LOPES, 2003; RUMEU, 2008, OMENA; BRAGA, 1996; CINTRA 1972, FARACO, 1996; dentre outros), passo já cumprido no capítulo 2 deste trabalho; (ii) posteriormente, com base em amostras de fala do PB contemporâneo disponíveis no banco de dados Iboruna (GONÇALVES, 2007), fizemos levantamento de natureza quantitativa dessas construções a fim de analisá-las qualitativamente quanto ao grau de esquematicidade, produtividade e composicionalidade em função dos contextos linguísticos. A justificativa por recorrermos a trabalhos diacrônicos está assentada no fato de que essas construções passaram por um processo de *construcionalização* e que experimentam, no PB contemporâneo, *mudanças construcionais pós-construcionalização*. Nosso compromisso aqui é o de descrever a mudança que vem afetando essas construções após elas terem entrado para o quadro pronominal do PB, pois, como dito anteriormente, a literatura linguística conta com diversos trabalhos sobre a mudança categorial dessas construções. O que ainda pouco se considera é processos de mudança que continuam operando pós-construcionalização.

### 3.1 *Corpus* de análise

É especialmente na modalidade falada da língua que processos de mudanças podem ser bem mais captados, como bem demonstramos no capítulo precedente recorrendo aos trabalhos de Duarte (1993, 1995), Cavalcante (1999) e Vargas (2010). Assim, os dados de uso real da língua analisados neste trabalho são provenientes de uma fonte principal, o banco de dados Iboruna, de responsabilidade do Projeto ALIP (Amostra Linguística do Interior Paulista),

sediado no Instituto de Biociências, Letras e Ciências Exatas – UNESP, *campus* de São José do Rio Preto e documentado em <https://www.alip.ibilce.unesp.br/projeto> (último acesso em ago.2021). O banco de dados Iboruna conta com amostras de fala coletadas na região de São José do Rio Preto, a Noroeste do estado de São Paulo, e compõe-se de dois tipos de amostra: *Amostra Censo* ou *Amostra Comunidade* (AC, daqui em diante) e Amostra de Interação Dialógica (AI, daqui em diante).

AC se compõe de 151 entrevistas sociolinguisticamente controladas e orientadas, por um roteiro de entrevista, para a obtenção de cinco tipos textuais: (i) narrativa de experiência pessoal (NE); narrativa recontada (NR); relato de procedimento (RP); relato de descrição (DE) e relato de opinião (RO). AI é constituída de 11 interações dialógicas, envolvendo de dois até cinco informantes, gravadas secretamente, em contextos livres de interação social. Considerando os dois tipos de amostra, o banco de dados conta com pouco mais de 1,5 milhões de palavras, considerado, portanto, um banco de porte médio (GONÇALVES, 2007). Para a AC, os perfis sociais dos informantes foram pré-definidos pelo entrecruzamento das variantes de *sexo/gênero*, cinco *faixas etárias*, quatro *níveis de escolaridade* e quatro faixas de *renda familiar*, do qual resultam os 151 perfis sociais pertencentes a uma das sete cidades de abrangência do Projeto: Bady Bassit, Cedral, Guapiaçu, Ipiguá, Mirassol, Onda Verde e São José do Rio Preto.<sup>13</sup> Para AI, dado o caráter secreto das gravações, com consentimento posterior dos participantes, não houve controle de perfis sociais dos participantes, pois o objetivo da amostra é o de captar a fala natural e distensa usada no cotidiano em contextos reais de interação social.

### 3.2 Composição da subamostra

Para a análise das construções pronominais *a gente* e *você* funcionando como estratégias de indeterminação do sujeito, montamos uma subamostra de AC, com 38 entrevistas, e incluímos todas as 11 interações dialogadas de AI. Seguindo a mesma metodologia de Rubio (2012), para a subamostras de AC, a seleção dos informantes foi feita atendendo-se a um balanceamento dos perfis sociais. Assim, selecionamos entrevistas com duração de 25 a 40

---

<sup>13</sup> Importante destacar que, do cruzamento das variantes sociais, AC deveria ser composta de 152 perfis sociais; no entanto, o perfil identificado como AC-060 (mulher, 26 a 25 anos, até 4 anos de escolaridade, renda familiar de 11 a 24 salários-mínimos, de São José do Rio Preto) não foi localizado na comunidade, razão de AC contar, ao final, com 151 perfis sociais (GONÇALVES, 2007),

minutos e provenientes de 19 homens e 19 mulheres, estratificados pelas cinco faixas etárias e pelos quatro níveis de escolaridade controlados pelo Projeto ALIP.

No quadro 8 dado a seguir, dos 151 perfis sociais de AC, destacamos os selecionados para a composição de nossa subamostra, com o desprezo da variável *renda familiar*.

**Quadro 8** – Identificação (em destaque) dos perfis sociais da Amostra Censo do Banco de dados Iboruna que compõem a subamostra utilizada na pesquisa

CARACTERIZAÇÃO DA SUBAMOSTRA (AC)											
Faixa Etária e Escolaridade	Sexo/gênero									Subtotal de inf.	Total de inf.
		M	F	M	F	M	F	M	F		
7 a 15 anos	1º ciclo EF	<b>001</b>	002	003	<b>004</b>	005	006	007	008	8	24
	2º ciclo EF	009	<b>010</b>	<b>011</b>	012	013	014	015	016	8	
	Ensino Médio	017	018	019	020	<b>021</b>	<b>022</b>	023	<b>024</b>	8	
16 a 25 anos	1º ciclo EF	025	026	027	028	<b>029</b>	<b>030</b>	031	032	8	32
	2º ciclo EF	033	034	<b>035</b>	<b>036</b>	037	038	039	040	8	
	Ensino Médio	<b>041</b>	<b>042</b>	043	044	045	046	047	048	8	
26 a 35 anos	Superior	049	050	<b>051</b>	<b>052</b>	053	054	055	056	8	32
	1º ciclo EF	057	058	059	060	061	<b>062</b>	<b>063</b>	064	8	
	2º ciclo EF	065	<b>066</b>	<b>067</b>	068	069	070	071	072	8	
36 a 55 anos	Ensino Médio	<b>073</b>	<b>074</b>	075	076	077	078	079	080	8	32
	Superior	081	<b>082</b>	<b>083</b>	084	085	086	087	088	8	
	1º ciclo EF	<b>089</b>	<b>090</b>	091	092	093	094	095	096	8	
+ 55 anos	2º ciclo EF	097	098	<b>099</b>	<b>100</b>	101	102	103	104	8	32
	Ensino Médio	<b>105</b>	<b>106</b>	107	108	109	110	111	112	8	
	Superior	113	<b>114</b>	<b>115</b>	116	117	118	119	120	8	
+ 55 anos	1º ciclo EF	<b>121</b>	<b>122</b>	123	124	125	126	127	128	8	32
	2º ciclo EF	<b>129</b>	130	131	<b>132</b>	133	134	135	136	8	
	Ensino Médio	137	<b>138</b>	<b>139</b>	140	141	142	143	144	8	
Superior	145	<b>146</b>	<b>147</b>	148	149	150	151	152	8		
Total por idade		Total por gênero				Total por escolaridade				Total geral	
7 a 15 anos = 6		Masculino = 19				1º ciclo EF = 10				38	
16 a 25 anos = 8						2º ciclo do EF = 10					
26 a 35 anos = 8		Feminino = 19				Ensino Médio = 10					
36 a 55 anos = 8						Ensino Superior = 8					
+ 55 anos = 8											

Fonte: Gonçalves (2007, com adaptações).

### 3.3 Parâmetros de análise

#### 3.3.1 Parâmetros de forma

##### a) Tempo e modo verbal

A escolha dessa variável deve-se ao fato de que certos tempos e modos verbais podem favorecer o afastamento da realidade, indicando situações atemporais, como é o caso, por exemplo, do presente do indicativo. A indeterminação do sujeito se ancora mormente no presente do indicativo porque, como afirma MENON (1994, p. 163), “o presente tem um valor

aspectual neutro pois ele denota uma ação sem fazer referência à sua duração e é utilizado por indicar fatos habituais”.

### **b) Formato da sentença**

Um conjunto de estudos voltados aos aspectos formais da língua tem se dedicado a investigar as mudanças referentes à realização de pronomes na posição de sujeito em estruturas de indeterminação do sujeito. Duarte (1995, 2003, 2008), por exemplo, investiga a realização do sujeito em sentenças tanto finitas quanto não-finitas, sob a hipótese de que o crescimento de sujeitos expressos começa pelas sentenças finitas e só mais tardiamente atinge as não-finitas. Objetivamos controlar o formato da sentença a fim de verificar quais índices são alcançados de sujeitos expressos por *[a gente]* e *[VOCÊ]* na sincronia atual como estratégia de indeterminação, e, assim, poder comparar os resultados a que a autora chegou e observar qual o caminho que a mudança tem percorrido.

### **3.3.2 Parâmetros de significado**

#### **a) Tipo de sequência discursiva**

Para darmos um tratamento semelhante de análise a ocorrências extraídas das duas amostras que compõem o *cópus* deste trabalho, de modo que possam ser analisadas pelos mesmos critérios, já que as amostras são de naturezas estilísticas diferentes no que respeita aos gêneros nelas retratados (entrevista sociolinguística e interações dialógicas espontâneas), seguiremos a metodologia de Freitag (2014) e Nardelli (2021), controlando a sequência discursiva em que ocorre a indeterminação de sujeito operada por *[a gente]* e *[VOCÊ]*.

Em relação à AC, a coleta das entrevistas semiestruturadas é guiada por um roteiro que compele o participante a produzir cinco tipos de texto oral: narrativa de experiência pessoal (NE), narrativa recontada (NR), relato de descrição (RE), relato de opinião (RO) e relato de procedimento (RP), assim classificados, pela predominância definida pelo tipo de pergunta dirigida ao participante pelo documentador. Tratar as ocorrências em termos de predominância de tipo significaria não ignorar a mescla de tipos que pode haver, já que não é de se esperar textos orais de tipo genuinamente único. No entanto, esse critério não seria adequado à análise de ocorrências de AI, porque, devido à natureza metodológica de sua coleta, torna-se

praticamente impossível lidar com predominância de tipo textual, em razão da liberdade que os participantes têm na escolha dos tópicos discursivos que surgem espontaneamente na interação, tornando assim inoperante a comparação de ocorrências provenientes das duas amostras. Assim, a opção por sequência discursiva contorna o problema da não comparabilidade de tipos textuais das duas amostras. Como afirma Nardelli (2021), tanto nas entrevistas de AC quanto nas interações dialógicas de AI podem ser identificadas sequências discursivas variadas, e, ainda que se considere a predominância de tipos textuais, não significa coincidência com sequência discursiva. Desse modo, a análise desse parâmetro será feita a partir de microporções textuais onde se encontram as ocorrências das construções de indeterminação do sujeito com *[a gente]* e *[VOCÊ]*, a partir das especificidades descritas por Nardelli (2021) e dadas a seguir.

#### **Sequência discursiva**

- a) **Narrativa:** relato verbal de um fato ou de uma história no passado, quase sempre no pretérito;
- b) **Descritiva:** trechos em que um fato, um objeto ou uma pessoa são descritos detalhadamente em suas peculiaridades e contornos, sem progressão temporal;
- c) **Argumentativa:** trecho em que se fundamentam opiniões ou se defendem pontos de vista. É frequentemente iniciada por construções: “eu acho (que)”, “eu penso (que)”, entre outras;
- d) **Injuntiva:** trechos que caracterizam por orientar ou instruir um procedimento. Essas sequências geralmente apresentam verbos no imperativo, infinitivo ou presente do indicativo com sujeitos não determinados;
- e) **Expositiva:** trechos em que são expostas informações sobre um fato, uma pessoa, ou um objeto, enumerando suas características de forma clara. (NARDELLI, 2021, p. 93-94).

Segundo Milanez (1987), a função de indeterminação do sujeito com pronomes nominativos é a de exemplificação, isto é, o falante utiliza essa estratégia para exemplificar situações genéricas, comuns a qualquer pessoa. O controle de sequência discursiva parte da hipótese de que, na construção de indeterminação do sujeito, a tendência de uso em sequências discursivas descritiva, injuntiva e expositiva alcance índices maiores do que em sequências narrativa e argumentativa, pelo modo elas se configuram estruturalmente.

Lidar com os dados a partir da noção de sequência discursiva parece ser uma via mais produtiva para investigar determinados fenômenos em textos de língua falada. O dado abaixo exemplifica como as sequências discursivas alteram constantemente na mesma fala de uma única pessoa, evidenciando assim a fluidez do fenômeno em análise.

- (19) Inf.: bom eu fiz um curso né?... porque como você falô(u) eu desfilo então [*a gente tem que... aprendê(r) a se maquiá::(r) né?... então:: a gente tem que... éh::... prime(i)ro você tem que passá(r) corretivo né?*]. [BDI-AC-042; RP: L. 196-198].

Como podemos observar no trecho parentetizado, esse turno de fala de uma única participante começa com uma sequência narrativa (*bom eu fiz um curso né?... porque como você falô(u) [eu desfilo]*) e, em seguida, alterna para uma injuntiva, onde ocorrem as formas pronominais *a gente* e *você* (*a gente tem que... aprendê(r) a se maquiá::(r) né?*]... *então:: a gente tem que... éh::... [prime(i)ro você tem que passá(r) corretivo né?*). Esse dado exemplifica muito bem como adotar as *sequências discursivas* é uma via de análise mais produtiva do que tipo textual predominante.

## b) Grau de indeterminação do referente sujeito

Conforme aponta Milanez (1987), a indeterminação do sujeito comporta graus quanto a sua abrangência, com o grau mais alto expresso pelas formas despronominalizadas: verbo na terceira pessoa do plural, verbo na terceira pessoa do singular seguido do pronome *se* e verbo na forma não-finita não-pessoal. Ainda segundo a autora, os recursos a *a gente* e a *você* “não apresentam um efeito tão genérico como o das três formas anteriores, porque, embora envolvam pessoas não determinadas pelo contexto, ainda mantêm referências *específicas* aos protagonistas da situação dialogal” (MILANEZ, 1987, p. 81; destaque no original). Concordamos com a autora no que diz respeito aos graus de indeterminação, mas discordamos no ponto em que ela afirma que *a gente* e *você* sempre mantêm referências *específicas*. As ocorrências em (20) abaixo corroboram essa nossa discordância.

- (20) a. só que num é por aí ***a gente num tem que dá(r) o troco...*** fazem(d)o o que a pessoa fez... ***a gente vai fazê(r) muito pior porque a gente é PIOR ainda que ela...*** tem que de(i)xá(r) passá(r) conversá(r) largá(r)... então se num tá bem larga... e aí com o tempo o que a pessoa fez com você ela vai tê(r) o troco... ***num é a gente que dá o troco*** é o tempo e a vida. [BDI-AC-024; RO: L. 401-405].

b. e aí num tem mais aquela coisa de **você tê(r) uma pessoa que cê pode contá(r) um segredo** a pessoa num vai falá(r) pa ninguém:... porque no o(u)tro dia a rodinha inte(i)ra já tá saben::(d)o... tá difícil. [BDI-AC-024; RO: L. 386-388].

Como se pode observar, em (20), há claramente a anulação das pessoas do discurso. O falante, em ambas as ocorrências, argumenta sobre fatos comuns e que servem de aconselhamento para qualquer pessoa. Nas ocorrências, o grupo de pessoas é genérico e não-específico, e, portanto, é muito possível os falantes usarem pronomes nominativos também para esse tipo de indeterminação, não apenas as três estratégias apontadas por Milanez (1987). O que defendemos aqui é que é possível identificar dois graus de generalização: um genérico definido e outro, genérico indefinido. Pretendemos controlar o grau de referência genérica e verificar em quais contextos as estratégias com *[a gente]* e *[VOCÊ]* são mais produtivas. Para tanto, recorreremos à categorização proposta por Rubio (2012), a qual resumimos a seguir.

(i) *referência genérica e indefinida*: [...] o pronome remete a uma categoria generalizada e indeterminada de indivíduos, geralmente com referência a pessoas ou a grupos;

(ii) *referência genérica e definida*: [...] o pronome remete a uma categoria generalizada, mas determinada de indivíduos. [...] fica claro que o falante tem consciência de determinado grupo de indivíduos, no qual ele próprio está incluso, por exemplo, as pessoas do trabalho, do futebol, da família, do bairro. (RUBIO, 2012, p. 167).

### c) Feixes semânticos do predicado da construção

Por meio desse parâmetro, almejamos investigar até que ponto as construções de indeterminação de sujeito com *[a gente]* e *[VOCÊ]* são produtivas e como o *slot* de predicado dessas construções é capaz de se estenderem a variados feixes semânticos. Para tanto, utilizamos a proposta de Bybee (2016), sobre feixe de exemplares, associada à escala de traços semântico-pragmáticos verbais de Tavares e Freitag (2010), reproduzida no quadro 9.<sup>14</sup>

---

<sup>14</sup> Para a investigação desse parâmetro, baseamos-nos na metodologia de Prezotto Júnior (2020), aplicada ao estudo diacrônico das construções “deixar” e “parar” na expressão de aspecto final no português.

**Quadro 9** – Feixe dos traços semântico-pragmáticos dos predicados das construções de indeterminação do sujeito

<b>TIPO</b>	<b>CONCEITO</b>	<b>VERBOS</b>
<b>1. Momentâneo</b>	Atividade repentina de curta duração.	<i>saltar, chutar, bater, derrubar etc.</i>
<b>2. Atividade específica</b>	Evocação de imagem específica.	<i>escrever, jogar, beber, desenhar etc.</i>
<b>3. Dicendi</b>	Registro de fala reportada.	<i>dizer, falar, responder, perguntar etc..</i>
<b>4. Atividade difusa</b>	Não evoca de imagem específica.	<i>aposentar-se, trabalhar, estudar etc..</i>
<b>5. Instância</b>	Posição corporal estática.	<i>deitar(-se), recostar(-se), sentar(-se) etc.</i>
<b>6. Estímulo mental</b>	Sujeito com papel de estímulo da experiência mental de outrem.	<i>impressionar, agradar, surpreender, assustar, espantar, aborrecer.</i>
<b>7. Evento transitório intencional</b>	Indicação de permanência ou não do referente do em certo lugar.	<i>permanecer, residir, situar, estar (em algum lugar) etc.</i>
<b>8. Evento transitório não intencional</b>	Ações não intencionais.	<i>cair, desmaiar, adormecer, acordar, quebrar (não intencional).</i>
<b>9. Processo</b>	Mudança não intencional sofrida por um corpo (+ ou - animado).	<i>deteriorar, crescer, amadurecer, transformar, ferver, congelar.</i>
<b>10. Experimentação mental</b>	Sujeito com papel de experienciador.	<i>adorar, odiar, desejar, pensar, lembrar, entender.</i>
<b>11. Atenuação</b>	Distanciamento ou suavização da opinião	<i>achar, pensar.</i>
<b>12. Relacional</b>	Processo de percepção da realidade: identidade, analogia, comparação, posse, causa, finalidade, consequência etc.	<i>depende de, merecer, precisar, servir como, assemelhar-se, causar, igualar, ter (posse), determinar, faltar (algo), errar, resultar de/em, relacionar-se com, custar.</i>
<b>13. Sensação corporal</b>	Sensação física.	<i>machucar-se, doer, ferir, sentir, sofrer.</i>
<b>14. Existência</b>	Denota existência de algo.	<i>ter, haver, existir.</i>
<b>15. Estado</b>	Denota um estado de alguém ou de algo.	<i>ser, estar, parecer, ter (olhos azuis)</i>

**Fonte:** Tavares e Freitag (2010, p. 108).

O quadro 8 apresenta a escala proposta pelas autoras, que a organizam hierarquizando os predicados numa escala crescente de abstratização, que vai de tipos com traços mais concretos, o ponto 1 da escala, a tipos com traços mais abstratos, o ponto 15 da escala. Considerar os tipos de predicados das construções aqui analisadas permitirá mostrar a produtividade dessas construções em termos da extensibilidade do *slot* reservado ao predicado. Queremos verificar qual das duas construções está mais sedimentada na rede linguística de indeterminação do sujeito da sincronia investigada. Faremos esse controle de tipos semântico de predicados recorrendo ao controle de frequência *type* (de tipo) e frequência *token* (de ocorrência). A conjugação desses dois critérios de frequência permitirá quantificar quantos tipos diferentes de verbos uma dada classe de predicados comporta (*types*, portanto) e quantas vezes cada tipo ocorre (*token*, portanto).

Procedemos também a uma análise colostrucional objetivando investigar quais tipos de predicados específicos são mais atraídos para cada uma das construções. A análise foi feita por meio do programa estatístico R, considerando o número total de palavras do *corpus* utilizado nesta pesquisa, conforme metodologia sugerida por Gries e Stefanowitsch (2004). A contagem do número total de palavras foi feita com auxílio da ferramenta *AntConc*. Processamos os dados no programa R, utilizando o *script* fornecido por Stefan Gries (2014) e seguindo o passo a passo para de uma análise colexêmica simples.

### **3.3.3 Exclusão de dados**

Construções de marcadores conversacionais com o pronome *você* (“agora *você* vê”, por exemplo) não foram consideradas nas análises devido a sua função pragmática ser de outra natureza. Nessa função de marcador discursivo, a construção é cristalizada, formando uma sequência inalisável em termos de forma e de conteúdo, escapando, portanto, à função de indeterminação do sujeito.

## CAPÍTULO IV – RESULTADOS E ANÁLISES

Neste capítulo, encontram-se a análise e as discussões dos resultados a que chegamos conforme investigação empírica que procedemos. Na primeira seção, tratamos de reinterpretar, dentro do quadro do MBU, as mudanças que afetaram a rede de indeterminação, foco do nosso trabalho, e como ela se reconfigurou a partir de diversas mudanças sintáticas relacionadas com a entrada de *a gente* e de *você* no quadro pronominal. Na sequência, após evidenciarmos a emergência do subesquema que abriga construções pronominais na rede de indeterminação do sujeito, focalizamos duas microconstruções que se subordinam a esse subesquema, a saber, [[*a gente*] [PREDICADO]] e [[*VOCÊ*] [PREDICADO]], e mostramos como elas se relacionam e quais suas especificidades na rede a que pertencem.

### 4.1 Reinterpretando a mudança da rede de indeterminação do sujeito no quadro dos Modelos Baseados no Uso

A primeira noção a ser refinada para os processos de mudança de *a gente* e de *VOCÊ* dentro do quadro dos MBU é a noção de *gramaticalização*, que, segundo Gisborn e Patten (2011), não deve ser considerada como superada. Os MBU adotam uma terminologia diferente para os processos de mudança que decorrem de gramaticalização. Hopper e Traugott (2003) definem a gramaticalização como um processo de mudança que, apreendido sincrônica ou diacronicamente, revela como “itens lexicais e construções passam a servir a funções gramaticais [...] e, uma vez gramaticalizados, continuam a desenvolver novas funções gramaticais” (p. 18).<sup>15</sup>

Traugott e Trousdale (2021) chamam a atenção para os termos *construção* e *construcional* que aparecem em Hopper e Traugott (2003) e afirmam que esses termos não indicam o mesmo sentido adotado na abordagem construcional da mudança, mas um sentido pré-teórico de constituinte, pois, enquanto a gramaticalização volta sua atenção ora para a forma, ora para o significado, a abordagem construcional, consoante a definição de *construção*, privilegia igualmente forma e significado.

Dados os propósitos deste trabalho, não pretendemos aqui fazer uma exposição exaustiva das diferenças e das semelhanças entre processos de gramaticalização e

---

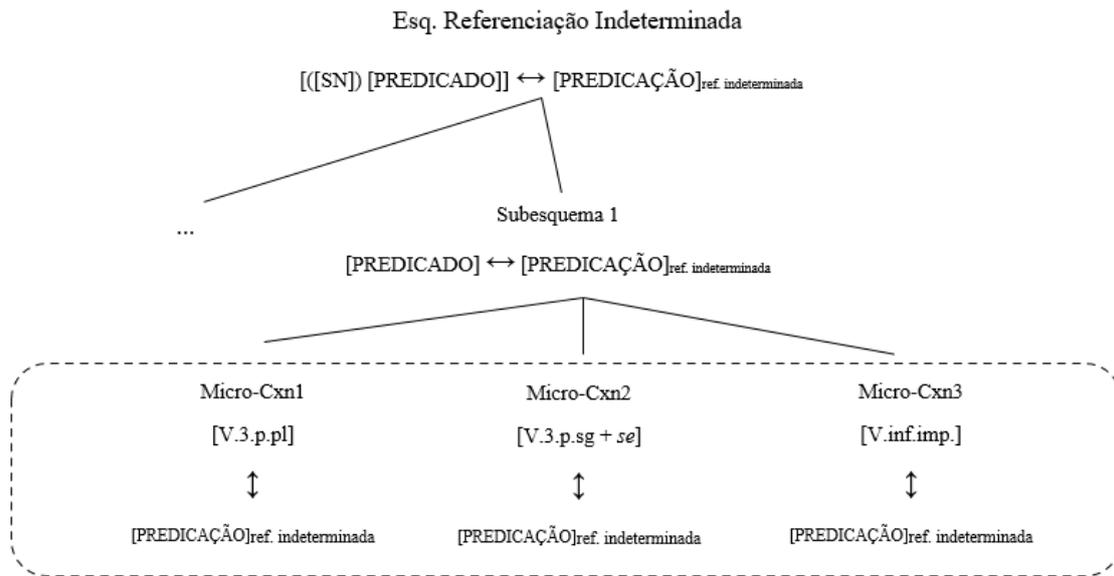
<sup>15</sup> No original: “lexical items and constructions come [...] to serve grammatical functions and, once grammaticalized, continue to develop new grammatical functions”.

construcionalização de [*a gente*] e [*VOCÊ*]. O que deve ficar claro é que, na abordagem construcional, a mudança não é estudada tomando o item lexical isoladamente, mas sob a ótica das propriedades de forma e de significado integrantes da *construção*. Desse modo, nosso foco aqui está voltado para as mudanças que ocorrem após a emergência de [*a gente*] e [*VOCÊ*] na rede pronominal, mudanças *pós-construcionalização* que continuam atuando e levando ao surgimento de novo padrão construcional na rede de indeterminação do sujeito.

Segundo os MBU, uma rede de construções é estruturada a partir de sua base, onde se situa o *locus* das inovações e onde se encontram os *construtos*, isto é, os usos efetivos da língua a partir dos quais representações esquemáticas são abstratizadas nos níveis mais altos da rede, uma representação *bottom-up*, portanto (TRAUGOTT; TROUSDALE, 2021; TAYLOR, 2002). No nível das *microconstruções* (Micro-Cxn), alocam-se construções individuais instanciadas a partir dos construtos. O próximo nível, o do *subesquema*, é reservado à representação de construções que guardam um conjunto de similaridades abstraídas das microconstruções. No nível mais alto da rede, o do *esquema construcional*, representam-se padrões construcionais altamente abstratos, genéricos e complexos, com diversas possibilidades de preenchimento previstas a partir dos níveis inferiores.

Assumindo esse modo de organização da rede, propomos, na Figura 4, a rede de referência das estratégias canônicas de indeterminação do sujeito, para, a partir dela, propormos, mais adiante na Figura 5, a rede de referência indeterminada atual do PB. Subjaz a essa proposição o entendimento de que a rede representada na figura 4 é historicamente anterior à representada na figura 5, dada a emergência mais recente de [*a gente*] e [*VOCÊ*] em construções de indeterminação no PB, como atestam os trabalhos diacrônicos de Duarte (1993; 1995), de Cavalcante (199) e de Vargas (2010) apresentados no capítulo 2 desta dissertação.

**Figura 4** - Rede de referência das estratégias canônicas de indeterminação do sujeito



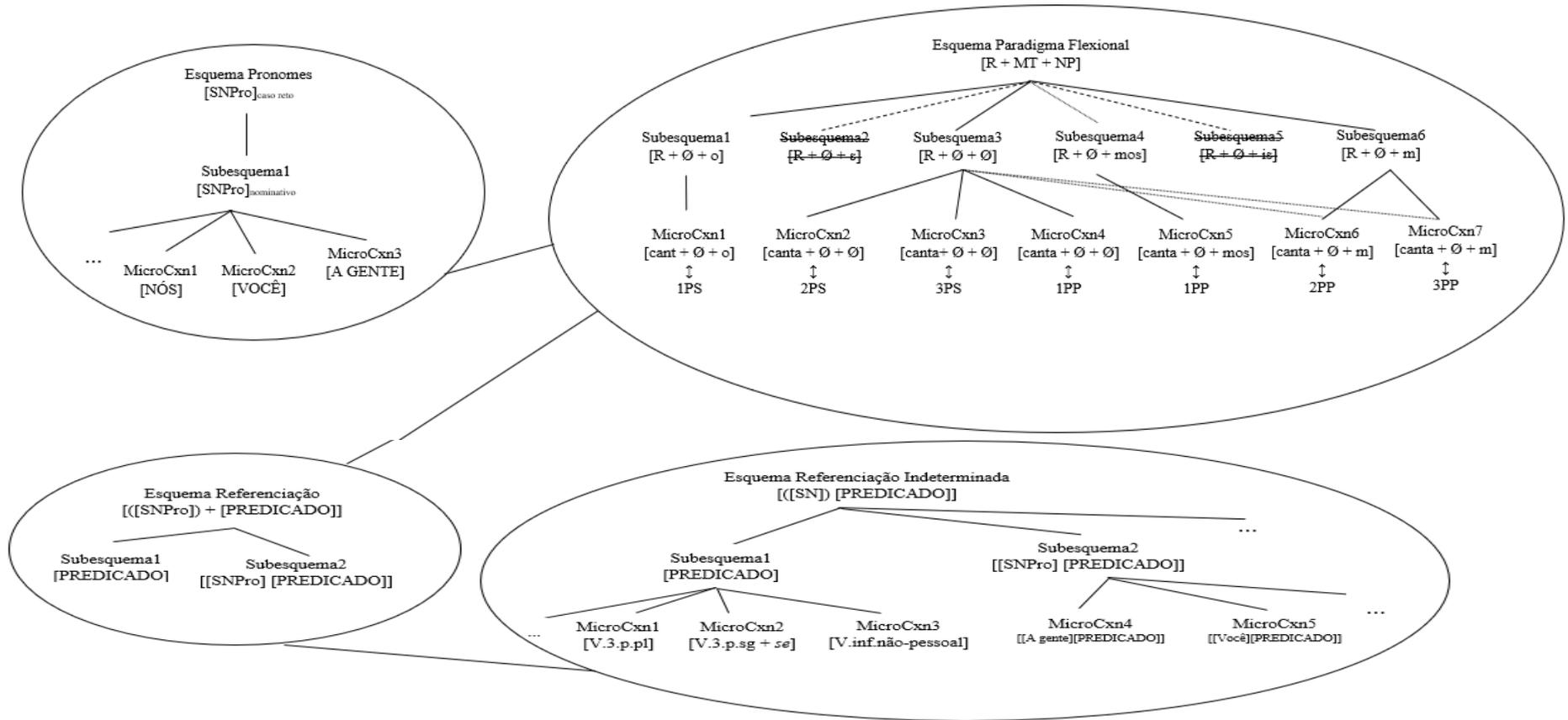
**Fonte:** elaboração própria.

Essa proposta de rede canônica de indeterminação do sujeito no português prevê que é no nível das Micro-Cxn que captamos o que tradicionalmente os gramáticos postulam como estratégias de indeterminação do argumento sujeito: (i) a Micro-Cxn1 representa a estratégia por meio de verbo na terceira pessoa do plural (p.ex.: *falaram muito de você*); (ii) a Micro-Cxn2 se realiza com o verbo na terceira pessoa do singular seguido do clítico *se* (p.ex.: *fala-se muito de você*); (iii) e a Micro-Cxn3 é formada por verbos impessoais na forma infinitiva (p.ex.: *falar é fácil*). Assim considerando, chegamos ao entendimento de que, no nível do subesquema, a indeterminação do sujeito se dá a partir de um [PREDICADO] que opera referência indeterminada na predicação que ele forma. Notamos que a rede não comporta apenas um subesquema, o que está representado pelas reticências (...), para dar conta de mais subesquemas que participam dessa rede, como é o caso de estratégias de indeterminação do sujeito operada por sintagmas nominais genéricos, como *o cara*, *o pessoal*, *a humanidade* etc., a partir de um subesquema do tipo [[SN]<sub>genérico</sub> [PREDICADO]] ↔ [PREDICAÇÃO]<sub>ref. indeterminada</sub> omitido na figura. Assim considerando, o esquema geral, representado no nível mais alto da rede, precisa dar conta do pareamento de forma e significado da diversidade de subesquemas e de microconstruções que a rede comporta e, desse modo, ele pode ser representado por uma construção que contém duas subpartes: uma reservada a ([SN]) pleno de referência genérica,

que, representado entre parênteses, pode ou não se manifestar nos subesquemas do nível imediatamente abaixo, opção mostrada pelas reticências; a outra reservada ao [PREDICADO] que, responsável pela formação da predicação, necessariamente, se realiza em qualquer dos níveis abaixo.

A partir das mudanças sintáticas pelas quais o PB tem passado, dentre as quais nos interessa focalizar as construções pronominais [*a gente*] e [*VOCÊ*] em construções de indeterminação, defendemos que a construcionalização dessas formas como pronomes referenciais dêiticos vem afetando a rede de indeterminação do sujeito, com a emergência de um padrão construcional que antes não estava disponível e cuja forma não é compatível com o subesquema 1 [[PREDICADO] ↔ [PREDICAÇÃO]]<sub>ref, indeterminada</sub>, representado na Figura 4, mas inclui, dentro da rede mais ampla de referenciação, também um novo subesquema representado por [[SNPro] [PREDICADO] ↔ [PREDICAÇÃO]]<sub>ref, indeterminada</sub>. É o que tentamos captar na Figura 5, na qual representamos a rede mais ampla de referenciação no português, instanciada pelo esquema genérico [[(SN)] [PREDICADO] ↔ [PREDICAÇÃO]]<sub>ref, indeterminada</sub>, ao qual se subordinam subesquemas de referência tanto determinada quanto indeterminada.

**Figura 5** - Reconfiguração das redes de construções afetadas pela mudança do paradigma pronominal do PB



Fonte: elaboração própria.<sup>16</sup>

<sup>16</sup> Por questões de economia, não representamos, nesta figura, o polo do significado dos esquemas por entender que o significado está explícito no próprio título de cada esquema.

A partir do que foi discutido no capítulo 2 sobre a mudança na sintaxe do PB, propomos, na figura 5, como a rede de referência do PB é impactada a partir da entrada dos pronomes *a gente* e *você* no quadro pronominal, o que afetou consideravelmente outras redes da língua. Essas mudanças corroboram fortemente uma das noções mais caras aos modelos de gramática que entendem a língua como redes interligadas, e que uma mudança em determinada rede pode afetar diversas outras redes interligadas por nós horizontais e verticais, representando o sistema neural do cérebro humano.

Como se pode observar na figura, estão representadas três redes que experimentaram uma mudança após a emergência na rede de pronomes dos nós que abrigam as construções pronominais [*a gente*] e [*VOCÊ*]. A partir daí, a rede do paradigma flexional também sofre forte alteração: um esquema que contava com seis subesquemas, atualmente conta com quatro, resistindo os subesquemas 1, 3, 4 e 6, representados pelo tracejado contínuo. Importante considerar que o subesquema 4, flexão de 1PP com a marca de número e pessoa “-mos”, tem sido cada vez menos usual, por conta da presença cada vez maior de *a gente*, sua forma concorrente (vd. RUBIO, 2012); por essa razão, representamos o elo vertical com o subesquema mais geral com um tracejado descontínuo, porque entendemos que esse subesquema está seguindo o mesmo caminho dos subesquemas 2 e 5. A esses subesquemas, estão subordinadas as microconstruções referentes aos pronomes pessoais do PB. Outra observação importante a se fazer é a de que, embora ainda haja uma distinção entre a 2PP e 3PP, subordinadas ao subesquema 6, na fala não padrão, essas construções se compatibilizam com o subesquema 3, o que pode encaminhar o PB a ter apenas duas distinções flexionais, característica de línguas com o preenchimento do sujeito obrigatório, com exceção do francês, por exemplo.

Essa simplificação do paradigma flexional levou a uma outra alteração em outra rede da língua, a do preenchimento ou não do sujeito, representada na figura pelo esquema de referência. O PB dispõe de dois padrões para a formação da predicação: (i) com sujeito vazio, marcado apenas na flexão verbal do predicador; e (ii) com sujeito preenchido lexicalmente, replicado na flexão verbal (GUERRA, 2017). Com a simplificação do paradigma flexional, o padrão com o sujeito vazio tem apresentado um decréscimo de uso cada vez maior, pois os falantes têm preferido cada vez mais, a fim de evitar ambiguidades referenciais, a posição do sujeito preenchida por pronomes (DUARTE, 1993, 1995). Com isso, a frequência de estruturas com sujeitos expressos por pronomes nominativos, representada pelo subesquema 2, aumentou significativamente, deixando alguns contextos bem específicos em que o

preenchimento do sujeito ainda é nulo, como é o caso das orações encaixadas com sujeito correferencial ao da matriz.

A partir daí, a rede de referenciação indeterminada, por analogização, começou a atrair para a predicação de referência indeterminada pronomes nominativos. Argumentamos, nesse sentido, que houve a criação de um padrão na rede de referenciação indeterminada, que contava apenas com esquemas em que era possível sintagmas nominais plenos, como *o cara, o pessoal, as pessoas*, ou só o predicado, como o verbo da 3PP, por exemplo, como representado na figura 4. A partir desses vários processos de mudança, a rede de referenciação indeterminada passa a abrigar mais um subesquema, representado pelo subesquema 2, que agora abriga o padrão de indeterminação com sintagma nominal pronominal (SNPro).<sup>17</sup>

A partir da formulação dessas redes, extraímos delas, para investigação mais aprofundada, as microconstruções *[[a gente] [PREDICADO]]* e *[[VOCÊ] [PREDICADO]]*, ambas operando estratégias de indeterminação do sujeito, subordinadas ao subesquema 2 da rede de referenciação indeterminada apresentada na figura 5.

Quanto ao grau de esquematicidade, as microconstruções *[[a gente] [PREDICADO]]* e *[[VOCÊ] [PREDICADO]]* são caracterizadas, então, por serem parcialmente esquemáticas, pois ambas são compostas de dois *slots*, o primeiro, preenchido pelos respectivos pronomes, e o segundo, aberto em [PREDICADO] e, portanto, altamente abstrato, por poder ser preenchido por predicados de tipos variados. Em termos de composicionalidade, essas construções são semanticamente não-composicionais, porque os pronomes que ocupam o primeiro *slot* perdem sua propriedade de referenciar sujeitos específicos da interação e, na combinação com predicados, passam a operar estratégia de indeterminação de sujeito na predicação formada. Em termos sintáticos, elas são composicionais, porque é possível reconhecer em suas partes a relação entre um sujeito e um predicado. Essa característica sintática parece convergir com um sistema gramatical em que o preenchimento do sujeito é cada vez mais frequente, o que coloca

---

<sup>17</sup> É importante evidenciar que afirmação feita aqui parte de análise geral dos trabalhos diacrônicos mobilizados nesta dissertação, carecendo, portanto, de uma pesquisa que leve em consideração pressupostos dos MBU para uma comprovação diacrônica mais assertiva. Além do que afirmamos aqui, há uma forte hipótese que merece investigação, qual seja: verificar se a emergência de um padrão construcional na rede de indeterminação instanciado por pronomes nominativos está relacionada, de fato, com o aumento do padrão de sujeito preenchido de referência determinada, como afirmamos, ou se já havia disponível na língua construções de indeterminação com os outros pronomes do quadro pronominal, e, assim, os pronomes *a gente* e *VOCÊ*, ao se construcionalizar, apenas sofrem pressão paradigmática, e passam a participar da rede motivados por elos horizontais.

em destaque a mudança na rede que passa a abrigar um padrão em que é possível indeterminar o sujeito com pronomes nominativos expressos.

#### 4.2. Caracterizando a construção sob análise

Assumimos que a construção de indeterminação do sujeito com pronome nominativo expresso constitui um novo pareamento de forma e sentido na rede de indeterminação do sujeito. A emergência dessas construções pode ser entendida, de forma geral, como um efeito colateral de diversas mudanças por que a gramática do PB vêm passando, conforme apontado na seção 4.1.

O que deve ficar claro aqui é que nosso objeto de análise não se restringe apenas aos pronomes *a gente* e *você*, mas recai sobre toda a construção de indeterminação da qual eles fazem parte. Embora o fenômeno que pretendemos investigar se relacione com o processo de mudança por que os pronomes passaram, ou seja, é uma mudança desencadeada pela entrada de *a gente* e *você* no quadro pronominal, não estamos interessados em olhar especificamente as possíveis mudanças em curso dos pronomes *a gente* e *você*.<sup>18</sup> Desse modo, a construção sob enfoque aqui é a representada na figura abaixo:

**Figura 6** - Construção de indeterminação do sujeito pronominal



**Fonte:** elaboração própria.

Nosso entendimento sobre tratar a construção de indeterminação como um esquema de predicação se justifica pelo fato de que, se considerarmos as estratégias canônicas de indeterminação do sujeito (cf. figura 4), podemos observar que a unidade de informação no contexto discursivo de indeterminação do argumento sujeito se dá por meio da predicação apenas. O que queremos dizer é que não há um único elemento dentro da predicação, seja o predador, sejam os participantes ou os termos predicados, capaz de atualizar o sentido de indeterminação do sujeito por si só. Em se tratando de pronomes nominativos em construções de indeterminação, a dependência da predicação verbal se acentua ainda mais, dada a natureza

---

<sup>18</sup> As mudanças que estamos fazendo referência são aquelas relacionadas à redução fonética da forma *você*, por exemplo.

fórica dos pronomes, que só ativam sua função de referenciação no uso, e, portanto, num esquema predicativo.

Desse modo, conforme Pina (2009) coloca em relevo, as construções de indeterminação, por servir a funções mais pragmáticas, distinguem-se das construções de sujeitos determinados porque, em termos de figura-fundo, o sujeito indeterminado é sempre fundo em relação ao predicado que, por sua vez, é sempre figura. Assim, constatamos que o sujeito indeterminado é menos saliente, desconhecido e não específico. A constatação de que é o predicado que tem saliência cognitiva em construções de indeterminação do sujeito encontra-se justificada em exemplos, fartamente observáveis no nosso dia-a-dia, como as construções em (21), realizadas sem sujeito, mas com complemento nulo definido, cuja identidade é recuperável apenas no contexto.

- (21) a. Aluga-se.  
b. Vende-se.

Partimos do entendimento de Langacker (1991), para quem uma melhor compreensão da estrutura oracional só é alcançada a partir da relação gramatical entre outros elementos da oração e o sujeito, cujo valor conceptual é o responsável por explicar tanto sua função discursiva como seu comportamento gramatical. O autor argumenta que a organização informacional figura-fundo é o parâmetro mais importante para definir o sujeito. Pina (2009) considera, e seguimos a autora, que os sujeitos podem ser fundo, como o caso dos sujeitos indeterminados, mas o sujeito canônico é figura, como os sujeitos determinados.

Nossos dados corroboram essas afirmações, uma vez que, por meio do teste estatístico,<sup>19</sup> constatamos que o tipo de construção mais fortemente atraído para a construção de indeterminação com *[VOCÊ]*, a mais frequente, evidencia como esse tipo de construção é orientado para o evento, desfocalizando o sujeito, que é genérico, e focalizando a predicação.

- (22) pra treiná(r) entendeu? [Doc.: hum::] porque:: cê tem que tê(r) uma:./ que nem voleibol *você tem que tê(r) vinte e quatro meninas* né?...

---

<sup>19</sup> O teste estatístico da análise colostrucional será detalhado na seção (4.4.2).

Caracterizamos esse tipo de construção dado em (22) como *relacional modal deôntica*, uma vez que a proposição está relacionada à necessidade ou possibilidade de atos realizados por agentes moralmente responsáveis (LYONS, 1977). Seguindo Hengeveld (2004), numa abordagem funcional das modalidades, entendemos que há dois alvos de avaliação da modalidade deôntica: o participante e o evento. Dall Aglio Hattnher (2009) resume essa distinção da seguinte forma:

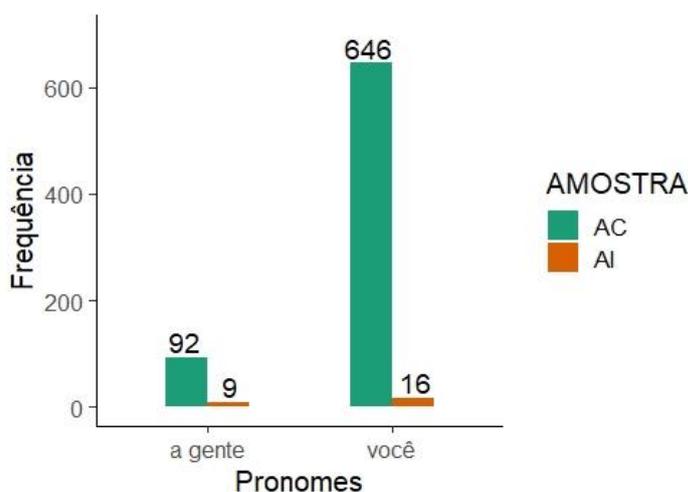
Na primeira, o participante é qualificado como aquele que tem a obrigação ou a permissão para se engajar no evento descrito pelo predicado. Na segunda, um evento é caracterizado como obrigatório ou permitido dentro de um sistema de convenções morais e legais. Na medida em que a modalidade orientada para o evento não incide sobre um participante específico, a obrigação ou permissão é representada como regra geral de conduta. (DALL AGLIO HATTNHER, 2009, p. 184).

Como se vê em (22), a modalidade expressa pela construção é deôntica e orientada para o evento, pois o falante está descrevendo a regra geral para montar um time de voleibol, regras permitidas dentro de um sistema de convenção. Dados como esse exemplificam como a construção de indeterminação do sujeito de fato tem por regra a desfocalização do sujeito, mesmo expresso, e focaliza o predicado. Por essa razão, entendemos que a construção de indeterminação deve ser analisada a partir de todo o esquema predicativo, constituído por dois *slots* altamente abstratos: o primeiro, representado por um sintagma nominal pronominal [SNPro], e o segundo, pelo predicado [PREDICADO], que inclui, em seus limites, a representação tanto de predicadores verbais e não verbais (transitivo, intransitivo, predicativo) como de seus complementos, quando é o caso. Outra constatação importante, já postulada por outros autores, é a de que, no polo formal da construção, a ordem dos constituintes é fixa, ou seja, em construções de indeterminação do sujeito no PB, a ordem sempre é SVO. Desse modo, argumentamos que de fato há criação de um nó na rede de indeterminação do sujeito, pois a construção com pronomes nominativos passa por mudanças em sua forma e em seu significado.

#### **4.3. Distribuição geral dos dados de indeterminação do sujeito no *corpus***

Um primeiro resultado trazido aqui se refere à frequência geral dos dados de construções de indeterminação do sujeito com [*a gente*] e [*VOCÊ*], apurada nas duas amostras, AC e AI.

**Gráfico 4** - Frequência das construções de indeterminação do sujeito por tipo de amostra



**Fonte:** elaboração própria.

A partir do gráfico 7, pode-se observar que os 763 dados de indeterminação de sujeito se distribuem, na seguinte ordem decrescente de frequência nas duas amostras: (i) em AC, [VOCÊ] (646 dados), [a gente] (92 dados), totalizando 738 dados; (ii) em AI, [a gente] (9 dados) e [VOCÊ] (16 dados), totalizando apenas 25 dados. Na totalidade dos dados, verificamos que [VOCÊ] e variantes somam 662 dados, contra 101 de [a gente], resultado que, em termos de produtividade, já confirma nossa hipótese de que o pronome [VOCÊ] é o recurso mais empregado como estratégia de indeterminação.

A grande diferença revelada na distribuição dos dados nas duas amostras se justifica em razão da extensão e das diferenças estilísticas das amostras. Como mencionamos no capítulo de metodologia, consideramos a totalidade dos inquéritos de AI, apenas 11, ao passo que, na composição de nossa subamostra de AC, selecionamos 38 das 151 entrevistas. O primeiro resultado em termos de produtividade em frequência bruta já confirma nossa hipótese de que o pronome [VOCÊ] é o recurso mais empregado como estratégia de indeterminação, totalizando, em ambas as amostras, 662 ocorrências, e [a gente] com 101 dados.

## 4.4 Parâmetros de Forma

### 4.4.1 Tempo e Modo Verbais

Um conjunto de estudos interessados no fenômeno de indeterminação do sujeito evidencia que a variável *tempo e modo verbal* é um forte favorecedor de usos de construções de indeterminação de sujeito. É o que se verifica em trabalhos como o de Menon (1994) e o de Setti (1997), cujos resultados indicam que, dentre os tempos e modos verbais, o *presente do indicativo* é o mais empregado em construções cujo argumento sujeito tem referência indeterminada. Em relação aos dados analisados neste trabalho, a tabela que segue traz os resultados a que chegamos.

**Tabela 4** - Frequência de tempo e modo verbal das construções de indeterminação do sujeito

[SNPro] Tempo e modo verbal	[A gente]		[VOCÊ]		Σ	
	nº oco	%	nº oco	%	nº oco	%
Presente do Indicativo	80	14,3	476	85,61	556	<b>72,87</b>
Pretérito Perfeito do Indicativo	-	-	17	100	17	2,22
Pretérito Imperfeito do Indicativo	3	7,69	36	92,3	39	5,11
Futuro do Presente do Indicativo	7	19,44	29	80,55	36	4,71
Futuro do Pretérito do Indicativo	-	-	3	100	3	0,39
Presente do Subjuntivo	-	-	4	100	4	0,52
Pretérito Imperfeito do Subjuntivo	4	50	4	50	8	1,04
Futuro do Subjuntivo	2	7,4	25	92,59	27	3,53
Forma não-finita	5	6,75	69	92,24	74	9,69
Σ	101	13,23	662	86,76	763	100

**Fonte:** Elaboração própria.

A partir dos dados apresentados na tabela 4, certificamos que, nas construções de indeterminação do sujeito em exame, o tempo verbal que apresenta a maior frequência é o presente do indicativo ( $556/763 = 72,87\%$ ). Os resultados demonstram que a construção  $[[VOCÊ] [PREDICADO]]$  tem índices de ocorrência muito superiores no presente do indicativo em relação aos outros tempos e modos verbais que ocorrem com essa construção, como, por exemplo, a forma não finita, que ocupa o segundo lugar de tipo mais frequente, com 69/74 dados. Com a construção  $[[a gente] [PREDICADO]]$ , o presente do indicativo também é o mais frequente, chegando a índices expressivos de quase 80% (80/101) do total de dados. Como se vê, o segundo tempo e modo mais frequente nesse caso é o futuro do presente do indicativo, registrando 7/101 ocorrências (6,93%), com uma diferença de mais de 70% em relação ao presente do indicativo.

Interessante notar que, na distribuição dos tempos e modos, a construção *[[a gente] [PREDICADO]]*, no que diz respeito ao presente do indicativo, apresenta uma proporção maior (90/101 = 80%) do que a construção *[[VOCÊ] [PREDICADO]]* (476/662 = 71,9%). Esse resultado para a predicação com *[a gente]* pode ser explicado pelo fato de que essa construção pronominal, como aponta Neves (2008), mesmo podendo efetuar referência indeterminada, frequentemente inclui a primeira pessoa do discurso e, assim, empregá-la em tempo e aspecto neutros pode ser uma estratégia do falante de escapar da expressão de sua subjetividade. O mesmo não acontece com a predicação com *[VOCÊ]*, pois a expressão da pessoa do discurso sempre é apagada quando esse pronome é usado como indeterminação do sujeito.

De fato, esperávamos encontrar esse resultado porque, como afirma Menon (1994), o presente do indicativo tem valores temporal e aspectual neutros, e a indeterminação pode se realizar mais facilmente em contextos em que situações atemporais e fatos habituais, por exemplo, são temas no momento da interação verbal, como exemplificam as ocorrências em (23) e (24).

- (23) Inf.: acredito que sim... informação dentro de casa... né?... e:: conselho também dentro de casa você sabe por quê?... lá fora assim no mundo de hoje... (fora) nas esco::las na televisão o que **a gente mais vê** é informação... mas eu acho que falta assim mais um apoio dos pais né? [BDI-AC-036; RO: L. 338-340].
- (24) porque se **você vive c'uma pessoa** que num tem condições de trazê(r) uma uma estrutura BOA... **você vai aprendê(r) aquilo que ela faz de errado na vida... você é crian::ça... você tá forman(d)o teu caráter** tá forman(d)o a sua personalidade... e **você FORma tua/ teu caráter ou sua personalidade...** de acordo com aquilo que **você vivencia naquele momento...** então isso é um:: lado muito sério. [BDI-AC-099; RO: L. 505-509].

De fato, o uso do presente do indicativo presta-se mais à descrição de estados ou de hábitos, como se vê nas ocorrências acima. Tanto em (23) quanto em (24), os falantes descrevem situações que podem ser vivenciadas por qualquer pessoa devido a seu caráter atemporal, que dá à sentença uma leitura genérica. Embora não se possa desprezar a quantidade de dados que ocorrem com outros tempos e modos verbais, como em (25) a seguir, o presente do indicativo é o tempo e modo verbal que mais favorece o uso das construções de indeterminação. Parece mesmo que a relação que o presente do indicativo estabelece com a construção tem por função enfatizar o evento em si, em que o foco no sujeito verbal é deixado em segundo plano.

- (25) Inf.: é então é:: põe uma música mais escurinha... [Doc.: aham ((concordando))] pra você podê(r) lê(r)... bem a letra aí você canta né?... aí depois que você cantô::(u) aí começa tudo né? um culto... [BDI-AC-106; DE: L. 498-500].

Já em (25), em que a falante está descrevendo como se realiza o culto em sua igreja, o sujeito indeterminado da construção tem a leitura de *genérico e definido*, variável que será considerada mais adiante, pois a informante tem em mente o grupo específico de pessoas que frequentam o culto e do qual ela faz parte. Nesse caso, o tempo e modo verbais são empregados para descrever a sucessão de ações episódicas temporalmente situadas, uma em relação a outra (*pôr a música → ler a música → cantar a música → começar o culto*), numa clara ordenação que é iconicamente motivada.

O que se depreende também desses resultados, considerando o que Himmelmann (2004) chama de expansão da classe hospedeira, é a maior variabilidade de tempo e modo com a construção *[[VOCÊ] [PREDICADO]]*, o que revela que a sua alta frequência de ocorrência se correlaciona com um maior número de colocados, diferentemente do arranjo colocacional de *[[a gente] [PREDICADO]]*, cujos tempos e modos verbais se restringem a 6 tipos e com frequência de ocorrência bastante baixa. Nesse sentido, o resultado apresentado para a variável de tempo e modo corrobora a afirmação de que a alta frequência da construção *[[VOCÊ] [PREDIÇÃO]]* propicia a maior variabilidade de tipos de tempos e modos verbais, um tipo de correlação prevista por Bybee (2005) e que revela a expansão dos recursos disponíveis no esquema.

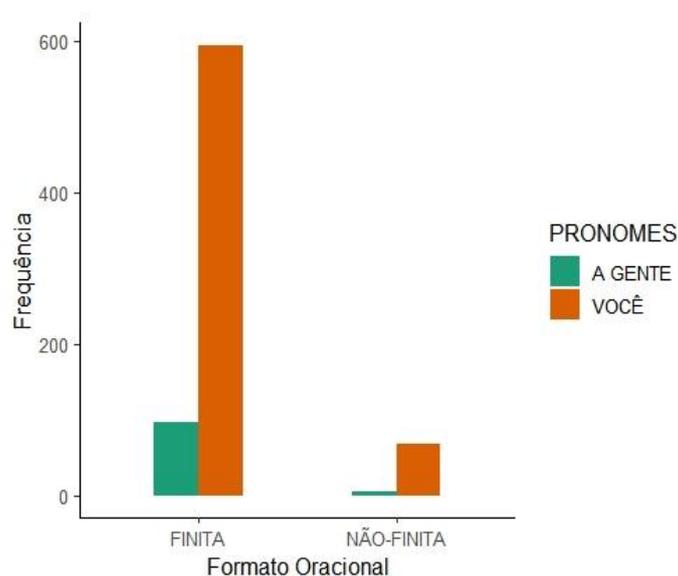
#### 4.3.2 Formato Oracional

Cavalcante (1999) constata que, em sentenças finitas, há uma maior frequência de uso de sujeitos preenchidos, mas que essa tendência não se restringe apenas a esse formato oracional. Duarte (2008) entende que sentenças não-finitas com sujeito preenchido se caracteriza como um novo padrão sentencial do PB, e que parece não haver notícias de que, em línguas ocidentais, tal fenômeno aconteça. As autoras afirmam que sujeitos pronominais nominativos de referência indeterminada começam a surgir no PB em contextos de sentença finita e que, só mais tardiamente, se expandem para sentenças não-finitas. Paralelamente a essas mudanças, as autoras constatam que, enquanto o uso de formas pronominais nominativas em

posição de sujeito tem um intenso crescimento, o clítico *se*, forma prescrita na norma gramatical para uso de indeterminação do sujeito, tem experimentado um decréscimo em seu uso.

Se levarmos em consideração as asserções das autoras, os resultados deste trabalho para o parâmetro *formato oracional* convergem para o fato de que sujeitos pronominais nominativos de referência indeterminada em sentenças finitas, de fato, ainda estão se expandindo para os contextos sentenças não-finitas. O gráfico a seguir demonstra o resultado.

**Gráfico 5** - Frequência das construções de indeterminação do sujeito em sentenças finitas e não-finitas



**Fonte:** elaboração própria.

Como se pode observar, os índices alcançados para o tipo de sentença finita são bem maiores do que os de sentenças não-finitas, para cada tipo de construção, com 90,30% (= 689/763) de sentenças finitas e apenas 9,69% (= 74/763) de não-finitas. Considerando cada tipo de construção, tanto *[[VOCÊ] [PREDICADO]]* quanto *[[a gente] [PREDICADO]]* se concentram acentuadamente em sentenças finitas, com índices respectivos de 89,57% (593/662) e 95,04% (= 96/101). Em contextos de sentença não-finita, essa proporção se mantém, com a construção *[[VOCÊ] [PREDICADO]]* sendo mais frequente (10,42% = 69/662) do que *[[a gente] [PREDICADO]]* (4,95% = 5/101).

O que chama atenção nesses resultados é o fato de que a construção de indeterminação com pronome nominativo se expandiu para usos em sentenças não-finitas, e, como constatado por Duarte (2008), esse é um novo padrão sentencial que o PB desenvolveu, especificamente com [*a gente*] e [*VOCÊ*]. Esse resultado corrobora a proposição de Bybee e McClelland (2005) de que, ao emergirem, as construções se expandem paulatinamente e, portanto, suas frequências vão aumentando no decorrer do tempo.

Focalizando apenas esse tipo de estrutura, apresentamos no gráfico a seguir a frequência de cada construção nos tipos de sentenças não-finitas encontradas no *corpus*.

**Tabela 5** - Frequência das construções de indeterminação do sujeito em diferentes tipos de sentenças não-finitas

[SNpro] Tipo de Oração não-finita		[ <i>A gente</i> ]		[ <i>VOCÊ</i> ]		$\Sigma$	
		Freq.	Proporção (%)	Freq.	Proporção (%)	Freq.	Proporção (%)
Completiva	Adjetival	2	40%	15	21,73%	17	22,97%
	Nominal	-	-	11	15,94%	11	14,86%
	Verbal	1	20%	10	14,49%	11	14,86%
Adverbial		1	20%	27	75%	28	37,83%
Coordenada		1	20%	5	7,24	6	8,10%
Relativa		-	-	1	1,44%	1	1,35%
$\Sigma$		5	100%	69	100%	74	100%

**Fonte:** elaboração própria.

Considerando os tipos de sentenças não-finitas, observamos que a predicação com [*VOCÊ*] é a mais frequente em todos os seis tipos de orações encontrados, enquanto a com [*a gente*] ocorre em quatro tipos, o que revela que, em termos de expansão de contextos, construções de indeterminação com [*VOCÊ*] está se sedimentando na gramática do PB antes das construções com [*a gente*].

Construções de indeterminação com [*a gente*], como mostram as ocorrências em (26), podem ocorrer nos seguintes contextos sintáticos: de oração adverbial, como em (26.a), de completiva adjetival, como em (26.b) e (26.c),<sup>20</sup> de coordenação (sindética), como em (26.d) e de completiva verbal (26.e.).

<sup>20</sup> A classificação desses tipos oracionais segue a proposta de Duarte (1993), dado que o quadro da Gramática Gerativa, por ela adotado, não reconhece a existência de sujeitos oracionais (KATO; MIOTO, 2000), como, por exemplo, outros pesquisadores de orientação funcionalista reconhecem para esse mesmo tipo de oração complexa

- (26) a. cê tá consideran(d)o o CAro e os bens materiais que a pessoa tem... e não é por aí... porque num é sempre na vida que **a gente vai tê(r) um enCOSTo pa gente ficá(r) encostan(d)o** e conseguin(d)o as coisa através das pessoas. [BDI-AC-024; RO: L. 374-376].
- b. Doc.: ah mas o professor vai cobran(d)o também... se ele <sup>31</sup>[num] <sup>31</sup>[Inf.: é] cobra realmente **é difícil a gente fazê(r)**. [BDI-AI-011; L. 171-172].
- c. Doc.: éh:: bom **o importante é** isso né? **a gente vivê(r) num lugar...** que é... confortável  
Inf.: é. [BDI-AC-036; DE: L. 257-258].
- d. mas **a gente tendo Deus** cê vence tudo... tu::do cê vence... tanto a doen::ça que nem::... coisas que acontece na vida porque acontece mesmo né?... mas graças a Deus eu tô bem contente... [BDI-AC-122; RO: L. 486-488].
- e. éh a gente tem que usá(r) o coração tudo bem... mas a raZãO é que **faz a gente cresce(r)**... não adianta. [BDI-AC-106; RO: L. 873-874].

A predicação com [VOCÊ], por sua vez, apresenta uma distribuição mais diversa: em sentenças adverbiais (27/69 = 39,13%), na forma de infinitivo ou de gerúndio, como em (27.a) e (27.b.), contextos mais frequentes; em completivas adjetivais (15/69 = 21,73%), como em (27.c); em completivas nominais (11/69 = 15,94%), como em (27.d); em completivas verbais (10/69 = 14,49%), como em (27.e); em estruturas de coordenação (5/69 = 7,24%), como em (27.f); ; e em relativa (1/69 = 1,44%), como em (27.g). Assim, podemos observar que o maior número de orações não-finitas se concentra em infinitivas.

- (27) a. um curso que eu gostaria de aprendê(r) bastante é Letras [...] porque **eles dão um novo idioma pra você falá(r)**... além de *você* (entendê(r)) seu verdade(i)ro idioma que é a raiz a origem que é o Português... [BDI-AC-035; RO: L. 461-465].
- b. eu por exemplo moro num bairro que fica muito longe aqui do centro... que **é uns quarenta minutos você andan(d)o a pé** né? [BDI-AC-074; RO: L. 43-44].
- c. Inf.: então eu acho que **é melhor às vezes... você fumá(r) um cigarro** do que matá(r) um próprio pai... e uma mãe... como a gente vê ultimamente na televisão. [BDI-AC-036; RO: L. 410-411].
- d. um curso que eu gostaria de aprendê(r) bastante é Letras... porque:: é um curso que é diferente dos o(u)tros eles dão **oportunidade de você... se torná(r) uma pessoa mais além do que você já é né?** [BDI-AC-035; RO: L. 459-463].
- e. mas... eu aceito ele na dele ele qué(r) saí(r) ele sa::i ele tem suas coisas ele fa::z... não sô(u) contra agora t/ assim... ví(r) morá(r)... com a/ na minha casa com o namorado isso eu já num aceito... (porque) **cê num justifica você criar um FILHO.. achando que:: cê vai tê(r) NEtos..** vai tê(r) várias coisas e não é nada disso então éh ele POde seguí(r) a vida de::le ele continua do jeito que ele qué(r) mas... lá fora... entendeu? [BDI-AC-052; RO: L. 300-305].

(GONÇALVES, 2017; GONÇALVES et al., 2008). Também nesse tipo incluem-se as orações subordinadas predicativas. Mantemos a classificação de Duarte para o simples fim de comparação de resultados.

f. Inf.: pra você vê uma cobra... e eu já penso uma cobra dentro da tua casa... **cê saí(r) daqui e í(r) lá onde ela tá** cê matá(r) ela tudo bem **mas ela éh vim dentro da sua casa** e **você tê(r) essa/**  
Doc.: í(r) com cuidado <sup>19</sup>[de tirá(r) a cobra]  
Inf.: <sup>19</sup>[**cê tê(r) esse cuidado**] [BDI-AC-063; NR: L. 527-531].

g. porque:: os melhores amigos MESmo né?... pode sê(r) considerado amigo... é os PAis né?... que são aqueles que sempre que deu fo::rça... apo::io né?... **tudo que cê precisá(r)**... o::... eu vô(u) citá(r) um eu vô(u) citá(r) um exemplo né?... [BDI-AC-011; RO: L. 257-260].

Não é espantoso que esse tipo de dado aconteça no PB, porque, se levarmos em conta processos cognitivos, como rotinização e automatização (PAWLEY; SYDER, 1983; HAIMAN, 1994), podemos interpretar, de nossos resultados e dos resultados de Duarte (1995) e de Cavalcante (2006), que, a partir do uso cada vez maior de sujeitos indeterminados expressos por pronomes nominativos em sentenças finitas, os falantes, por pensamento analógico, procuram alinhar construções não-finitas não-pessoais, sem sujeito explícito, ao padrão de sujeito preenchido no PB, cada vez mais frequente. A partir daí, com o aumento de frequência, esse novo tipo de padrão se convencionaliza, e, por meio do mecanismo de analogização, um novo nó na rede se cria.

Em relação às construções em análise, considerando que a produtividade de uma construção é gradiente (TRAUGOTT, TROUSDALE, 2021), é possível afirmar que dos dois tipos de construções aqui enfocados e sancionados no novo esquema de indeterminação, a construção [[VOCÊ] [PREDICADO]] é o tipo mais produtivo, porque, além de mais frequente, ocorre também em mais contextos oracionais. Como argumenta Goldberg (1995), nem sempre a frequência de uso é suficiente para tratar da produtividade de uma construção, sendo, em alguns casos, necessário levar em consideração também a frequência de tipos de padrões específicos.<sup>21</sup>

## 4.4 Parâmetros de significado

### 4.4.1 Grau de referência do sujeito

Estudos sobre alternância pronominal, especificamente de *nós* e *a gente* (OMENA, 1986, 1996, 2003; LOPES, 1993; RUBIO; GONÇALVES, 2010; RUBIO, 2012; SCHERRE;

---

<sup>21</sup> A autora exemplifica esse postulado com o caso da construção com *way*, do inglês, que não tem frequência de ocorrência tão alta, mas conta com um grande número de colocados, e, por isso, deve ser considerada produtiva.

NARO; YACOVENCO, 2018; NARDELLI, 2021; dentre outros), evidenciam que, em relação à referência do sujeito, o emprego de *a gente* é mais frequente quando tem um valor referencial genérico e indefinido, ao se referir a um número grande de indivíduos. Ao tratar da variedade do interior paulista, Rubio (2012), em estudo sociolinguístico, confirma a hipótese de que *nós* seria mais empregado em contexto de referência específica e definida, e *a gente*, em contexto de referência genérica e indefinida, dada a persistência de traços semânticos de coletividade de sua forma de origem. Segundo os resultados do autor, é evidente a oposição entre os traços indefinido e definido do referente no emprego de *a gente* e de *nós*, com pesos relativos de 0,57 para o primeiro, e 0,61, para o segundo.

Esses resultados mostram que *a gente*, dada a atuação do *Princípio de persistência* (HOPPER, 1991), ainda carrega da forma original o traço de coletividade, o que possibilita uma leitura indeterminada (OMENA; BRAGA, 1996), ou genérica e indefinida, nos termos de Rubio (2012). Ao investigar essa mesma hipótese em nosso trabalho, esperávamos encontrar maior frequência de referência genérica e indefinida relacionada à predicação com [*a gente*]. Entretanto, os dados nos revelaram que a maior frequência de referência genérica e indefinida ocorre em proporção maior com a construção [[*VOCÊ*] [PREDICADO]], como mostram os resultados da tabela 6.

**Tabela 6** - Frequência do grau de referência do sujeito das construções de indeterminação

[SNpro]	[ <i>A gente</i> ]		[ <i>VOCÊ</i> ]		$\Sigma$	
	Freq.	Proporção (%)	Freq.	Proporção (%)	Freq.	Proporção (%)
Genérico definido	18	17,9	56	8,4	74	9,7
Genérico indefinido	83	82,1	606	91,6	689	90,3
$\Sigma$	101	100	662	100	763	100

**Fonte:** elaboração própria.

Os resultados apresentados na tabela 6 são bem interessantes. Considerando os significados de cada construção (cf. Quadro 3 da seção 1.3.1), enquanto a construção pronominal [*a gente*], mesmo quando empregada em contexto de indeterminação, sempre envolve, de algum modo, a 1ª pessoa (NEVES, 2008) e a generaliza por meio da interseção entre a determinação e a indeterminação (MILANEZ, 1982), a construção pronominal [*VOCÊ*], por sua vez, é mais afeita a contextos em que as informações da predicação se referem a situações comuns a todos, apresentando-as como características de muitos. Interessante notar que, do total de casos, a referência genérica e indefinida é mais frequente ( $689/763 = 90,3\%$ ),

e, nesse contexto específico, em termos de proporção, a estratégia com [VOCÊ] (606/689 = 88%) segue mais sedimentada do que a com [a gente] (83/689 = 12%). As ocorrências em (28) e (29) exemplificam, respectivamente, a referência genérica e indefinida de [a gente] e [VOCÊ] em construções de indeterminação de sujeito.

(28) a. acho que deveria proibí(r) SIM a arma sim porque a gente mesmo num pode se defendê(r) deles. [BDI, AC-030; RO: L. 143-145].

b. eu acho que se a gente parasse pra olhá(r) PRA GENTE e não só pra gente... *pra gente que eu digo é:: pessoas em plural e não no singular entendeu?...* [Doc.: uhum ((concordando))] éh:: a gente::... **poderia começá(r) a tê(r) uma educação...** [AC – 035; RO: L. 530-533].

(29) Doc.: ah então ele TAva mesmo perseguindo?  
Inf.: TAva... então ele tava porque... **é incrível você sair d’um bairro í(r) pa O(u)tro bairro e um moço ficá(r) te seguin(d)o se você atraVEssa uma avenida ele atravessa se você VIra uma rua ele vira... se você PÁra ele pára... se você olha pra trás a/ o moço disfarça...** dá MEdo né?... [BDI, AC-066; NE: L. 32-36].

Em (28.a), [a gente] é empregado no sentido de qualquer pessoa, a humanidade, sem referência a um grupo específico, ou seja, o falante diz que deveria proibir o uso de armas porque as pessoas em geral não conseguem se defender “deles”, os assaltantes, cuja referência também é genérica, mas, nesse caso, definida. O dado em (28.b) é contundente para o que estamos afirmando, uma vez que revela algum grau de consciência do falante acerca do valor referencial genérico que o pronome comporta, e, para não correr o risco de o interlocutor fazer uma associação errada, ele verbaliza que “a gente” são as pessoas em geral. Em (29), [VOCÊ] é empregado em uma situação em que a informante generaliza sua experiência pessoal, uma impressão pessoal, a qualquer pessoa que possivelmente passaria pela mesma situação por que ela passou, ou seja, ela não tem em mente um grupo específico de pessoas, mas um grupo genérico e indefinido.

Outro ponto importante a destacar nos resultados ocorre quando o grau do referente do [SNpro] sujeito é genérico e definido. Podemos verificar, nesse contexto, a proporção de uso de [a gente] é o dobro (18/101 = 17,9%) da de referência genérica e indefinida de [VOCÊ] (56/662 = 8,4%). Considerando mais uma vez o significado das duas construções pronominais como estratégias de indeterminação, [a gente] apresenta ainda a possibilidade de ocorrer em contextos em que é possível identificar um grupo específico do qual o falante faz parte (mas não sendo possível identificar um único referente), enquanto [VOCÊ], por operar indeterminação por meio de uma projeção, em que uma 2ª pessoa é “criada”, seu uso nesses

contextos encontra mais resistência. Em (30) e (31), apresentamos ocorrências exemplificativas de usos dos dois pronomes nesse contexto de referência genérica definida.

- (30) Inf.: os projetores... então as músicas são todas projetadas na parede em cima... então ***você tem a letra que tá cantan(d)o...*** então eles vão cantan(d)o e ***você vai acompanhan(d)o cantando*** porque ***você tem a letra... no projetor***. [BDI-AC-106; DE: L. 494-496].
- (31) a eleição vem aí... eu tô confiante... que isso aí dê certo... quanto a nosso paí::s geralmente... ***a gente fala o país nosso deveria sê(r) um país de prime(i)ro mundo*** tem tudo pra sê(r) de prime(i)ro mundo né? esse país tem:: esse país nosso é riquíssimo. [BDI-AC-139; RO: L. 551-554].

Como se pode observar, os usos de [*a gente*] e de [*VOCÊ*] em (30) e (31) diferem em algum grau dos usos apresentados em (28) e (29). Em (20), a falante está descrevendo como é o culto em sua igreja, e por isso, a referência do pronome é genérica, mas definida, referenciando o grupo de pessoas do culto de que ela participa, não sendo possível identificar um referente no qual o interlocutor possa estar incluído. Em (31), [*a gente*] se refere a um grupo também genérico, mas delimitado de pessoas, no caso, os brasileiros. O pronome possessivo de 1ª pessoa do plural associado ao nome "país" (em “nosso país” e “esse país nosso”) garante essa leitura genérica e definida, revelando que o informante tem em mente um conjunto delimitado de pessoas.

Como se observa pelos resultados do controle desse parâmetro, a indeterminação do sujeito não é algo absoluto; ao contrário, ela comporta graus (MILANEZ, 1982), como também esperamos ter demonstrado. Entretanto, se levarmos em consideração que o sentido mais básico da construção de indeterminação é a *generalização*, podemos ver como o significado de cada pronome contribui com o significado da construção de indeterminação e, conseqüentemente, para seu grau de produtividade. De fato, considerando o significado de [*VOCÊ*], o resultado apresentado evidencia que é em construções cuja posição de sujeito é ocupada por esse pronome que há maior grau de produtividade em termos de frequência de ocorrência. Os usuários da língua lançam mão dessa estratégia mais frequentemente porque, por meio dela, é possível desviar-se do traço de 1ª pessoa que a construção pronominal [*a gente*] pode ainda comportar, estando essa construção mais reservada a contextos específicos em que o interlocutor não tem a possibilidade de estar incluído, por exemplo.

#### 4.4.2 Feixes semânticos dos predicados

A produtividade de uma construção pode ser investigada a partir do arranjo colocacional dos esquemas abstratos que a compõem, o que Barðdal (2008) chama de *extensibilidade* do esquema. As duas microconstruções aqui investigadas contêm o *slot* de [PREDICADO] altamente esquemático. Por se tratar de um trabalho sincrônico não é possível mostrar a expansão nesse arranjo, porque a expansão de um determinado esquema só pode ser atestada diacronicamente. Entretanto, a partir das ocorrências encontradas no nosso *corpus*, podemos caracterizar os tipos de construções que ocupam esse *slot*, a fim de constatar o grau de consolidação das duas microconstruções na rede linguística, perseguindo a hipótese de que, à medida que elas são capazes de sancionar predicados mais abstratos, mais consolidadas elas estão. É o que pretendemos verificar nesta seção.

Para tratar dos tipos semânticos de predicados que ocupam o *slot* esquemático da construção, baseamos-nos na proposta de Bybee (2016), sobre feixe de exemplares, e na proposta de Tavares e Freitag (2010), sobre a escala de traços semânticos-pragmáticos verbais, cada um deles entendido aqui como um tipo de feixe semântico. A escala das autoras consiste na organização de traços semânticos de verbos com significado mais concreto, como “escrever”, de *atividade específica*, e no ponto mais abstrato, verbos com significado mais abstrato, como “ser”, de *estado*.

Bybee (2016) argumenta que circunstâncias do uso tem impacto direto na representação cognitiva da língua, e que representações de feixe de exemplares acompanham o uso e registram detalhes do conhecimento linguístico. Desse modo, são os feixes de exemplares os responsáveis por estocar na memória enriquecida todas as informações que o usuário da língua experimenta na percepção linguística, argumenta a autora. Essas informações dizem respeito a todos os níveis linguísticos e extralinguísticos, ou seja, desde detalhes fonéticos até propriedades do contexto social e físico. Em nosso caso, nosso objetivo é investigar qual a qualidade dos feixes semânticos, apurados na frequência de ocorrência (*token*) e frequência de tipo (*type*), que as construções sancionam.

Os resultados apresentados na tabela 7 trazem informações sobre a frequência de ocorrência (*token*) e de tipo (*type*) de cada uma das construções em relação aos feixes semânticos do *slot* [PREDICADO].<sup>22</sup>

---

<sup>22</sup> Para o estudo aqui proposto, foi necessário acrescentar à escala de Tavares e Freitag (2010) um traço semântico, o qual denominamos como “Relacional Modal”, ocupando, na escala, a posição subsequente do traço semântico do tipo “Relacional”.

**Tabela 7** - Frequência token e type dos feixes de traços semântico-pragmáticos dos predicados das construções de indeterminação do sujeito

	Feixes Semânticos do Predicado	[A GENTE]		[VOCÊ]	
		Tk	Ty	Tk	Ty
+ concreto ↑	Momentâneo	-	-	9	5
	Atividade específica	35	22	363	96
	Dicendi	3	1	7	4
	Atividade difusa	28	20	91	54
	Instância	1	1	2	2
	Estímulo mental	-	-	2	1
	Evento transitório intencional	-	-	5	3
	Evento transitório não intencional	-	-	1	1
	Processo	1	1	2	2
	Experimentação mental	5	3	36	14
	Atenuação	-	-	-	-
	Relacional	13	1	51	2
	Relacional modal	8	7	54	38
	Sensação corporal	1	1	6	4
	Existência	-	-	-	-
	+ abstrato ↓	Estado	6	2	33
	Subtotal	101	59	662	229
	Total	763			

Fonte: elaboração própria.

Tomando o feixe semântico *atividade específica* como exemplo, a leitura dessa tabela deve ser feita da seguinte maneira: com a predicação [*a gente*], encontramos 35 ocorrências (*tokens*) de predicados distribuídas em 22 tipos de verbos diferentes (*types*), o que significa que há tipos de verbos que ocorrem mais de uma vez. A construção [[*VOCÊ*] [PREDICADO]] desse mesmo feixe semântico apresenta frequência *token* de 363 ocorrências, agrupadas em 96 tipos diferentes de verbo.

Os resultados da tabela 7 permitem considerar que, de modo geral, as duas construções estão bem consolidadas na língua, uma vez que ambas sancionam feixes semânticos que vão de tipos mais concretos a mais abstratos. No entanto, a diferença entre elas no processo de sedimentação está no fato de que a construção [[*VOCÊ*] [PREDICADO]] apresenta maior variabilidade de feixes semânticos, sancionando 14/16 tipos da escala (exceto em *atenuação* e *existência*),<sup>23</sup> enquanto [[*a gente*] [PREDICADO]] sanciona 10/16 tipos, o que é mais um indício de que a primeira é mais produtiva do que a segunda.

<sup>23</sup> Em relação aos feixes semânticos de *existência*, a frequência nula dessas construções se justifica devido ao fato de que não interpretamos essas construções como de indeterminação do sujeito, mas como construções que estão se encaminhando para uma referência nula no PB, como o caso dos expletivos em línguas como o francês e o inglês. Estão, portanto, em um processo mais avançado de construcionalização. Os dados da ocorrência dada

Para cada construção, chama a atenção a diferença entre a frequência *token* e *type*. Na maioria dos feixes semânticos comuns às duas construções, a razão<sup>24</sup> *token/type* (= r) tende a ser sempre menor para [*a gente*] do que para [*VOCÊ*], como se observa, respectivamente, para os seguintes feixes: (i) de atividade específica, com 35/22 (r: 1,5) e 363/96 (r: 3,7); (ii) de atividade difusa, com 28/20 (r: 1,4) e 91/54 (r: 1,7); (iii) de experimentação mental, com 5/3 (r: 1,6) e 36/14 (r: 2,5); (iv) de relacionais, com 13/1 (r: 13) e 51/2 (r: 25,5); (v) de relacionais modais, com 8/7 (r: 1,1) e 54/38 (r: 1,4); (vi) e de sensação corporal, com 1/1 (r: 1) e 6/4 (r: 1,5). A razão *token/type* é maior para [*a gente*] do que para [*VOCÊ*] somente nos casos de verbos *dicendi* (3/1 = r: 3 e 7/4 = r: 1,7, respectivamente) e verbos de estado (6/2 = r: 3 e 3/3 = r: 1, respectivamente).

O que esses resultados parecem revelar é que a construção [[*VOCÊ*] [PREDICADO]], por sua alta frequência, tem *types* mais consolidados na memória enriquecida dos falantes, e esses *types*, como exemplares, é que estão servindo de base analógica para a atração daqueles não tão frequentes assim.

O que se pode notar também é que ambas as construções têm maior concentração de ocorrência em feixes semânticos mais concretos, especificamente os de *atividade específica* e *atividade difusa*. No ponto mais abstrato da escala, a maior concentração está nos feixes *relacional* e *relacional modal*. As ocorrências em (32) são exemplos dos feixes semânticos mais concretos.

- (32) a. as coisas eu preciso tá c'a receita grudada nos olhos pra fazê(r) certo... então... ela me ensinô(u)... que **a gente coloca... uma lata... de leite condensado... uma lata de leite...** [BDI-AC-036; RP: L. 266-269].  
→ [*atividade específica*]
- b. eu faço um bolo salgado que é muito bom... [Doc.: hum] cê pega o pão *Pullman*... tira as bordas [Doc.: uhum ((concordando))] aquele torrãozinho **você tira**... coloca em uma assade(i)ra em uma forma [Doc.: hum]... **você passa maionese**... [BDI-AC-066; RP: L. 324-326]. → [*atividade específica*]
- c. só que num é por aí a gente num tem que dá(r) o troco... fazen(d)o o que a pessoa fez... a gente vai fazê(r) muito pior porque a gente é PIOR ainda que ela... tem que de(i)xá(r) passá(r) conversá(r) largá(r)... então se num tá bem larga... e aí com o tempo o que a pessoa fez com você ela vai tê(r) o troco... num é

---

abaixo, semelhantes aos de Duarte (2003), representam esse tipo de construção existencial que, em contexto típico, prescindiria do pronome.

as músicas são todas projetadas na parede em cima... então **você tem a letra** que tá cantan(d)o... então eles vão cantan(d)o e você vai acompanh(an)d(o) cantando porque **você tem a letra**... no projetor. [BDI-AC-106; DE: L. 494-496]

Por terem estatuto diferente, consideramos que essas construções devem ser investigadas em outro estudo.

<sup>24</sup> A razão *token/type* revela a concentração média de *tokens* por *type*.

***a gente que dá o troco*** é o tempo e a vida [BDI-AC-024; RO: L. 401-405]. → [*atividade difusa*]

d. um mercado porque ele pegô(u)... um quilo de arroz pa levá(r) pa... pra fazê(r) po filho dele... e aí? como ***cê vai... incriminá(r) essa pessoa?***... aí cê... leva ele ele... puma cadeia... ao invés de ele aprendê(r) coisas boas... [BDI-AC-051; RO: L. 540-543]. → [*atividade difusa*]

e. Inf.: olha éh:: nunca aconteceu comigo... mas do contrário se aco/ chegasse a acontecê(r) comigo no caso eu acho que iria prime(i)ro sabê(r)... éh:: a opinião do professor né? porque cê num tem como ***você falá(r) alguma coisa*** sem sabê(r) o que aconteceu né? [BDI-AC-067; RO: L. 369-372]. → [*dicendi*]

Como podemos observar, em (32.a) e (32.b), os verbos são categorizados como de *atividade específica* porque evocam imagens específicas da atividade realizada, conceptualizadas nas cenas de “colocar uma lata de leite condensado” e “passar a maionese no pão”. Tipos como esses parecem ser condicionados pela natureza da sequência discursiva em que ocorrem, como abordaremos na próxima seção. No caso em questão, por exemplo, sequências discursivas do tipo *injuntiva*, em que o falante instrui como fazer algo, tem forte relação com o tipo semântico do verbo, pois dar uma instrução é fazer o interlocutor evocar uma imagem específica da atividade.

Em (32.c) e (32.d), os verbos exemplificam o feixe semântico *atividade difusa*, como *aprender, mendigar, trabalhar*, sem evocar uma imagem específica (TAVARES; FREITAG, 2010). Não é possível evocar uma imagem nem em (32.b), em que “dar o troco” é usado metaforicamente, nem em (32.d), com o predicado “incriminar” uma pessoa. Em (32.e), exemplificamos o ponto da escala de feixes semânticos mais concretos com um uso de verbos *dicendi*.

Em (33) seguem ocorrências de predicados dos feixes semânticos mais abstratos, os tipos *relacional* e *relacional modal*.

(33) Inf.: ai a distribuição de renda num/ num é TÃO assim correta né? se fô(r) vê(r) porque tem::... muitas pessoas que traBALham demais e ganham MUITO po(u)co... e é tudo que vai aumentan::do ... salário que num aumenta... só as coisas que aumentam... e por mais que a pessoa... seje bem de vida ela também passa necessidade pelas coisas que tá mu/ tudo muito ca::ro... pra ***você tê(r) um ca::rro*** [→ *relacional*] ***você tem que gastá(r) trinta reais por sema::na com gasoli::na***.... [→ *relacional modal*] aí a comida onde fi::ca? [BDI-AC-024; RO: L. 453-458].

Os feixes semânticos apresentados acima são exemplos dos tipos mais frequentes no ponto mais abstrato da escala. O feixe semântico *relacional* representa, como apontam Tavares e Freitag (2010), a relação de como se percebe a realidade. O verbo *ter* (posse) é um dos tipos verbais que dá condições para assinalar esse processo de percepção da realidade. Acrescentado

à escala das autoras, o tipo *relacional modal* refere-se à expressão de uma conduta (valor deôntico), em que o falante julga a necessidade, a obrigação, a capacidade ou a permissão de realização de um estado de coisas por um agente moralmente responsável (LYONS, 1977). É importante registrar que, dentre os feixes mais abstratos da escala, o tipo mais frequente é o *relacional modal*, em construção do tipo [TER + que + X], como exemplificado em (33).

Na relação de frequência entre tipos de predicados de feixes mais concretos e os de feixes mais abstratos, fica a hipótese forte de que o processo de sedimentação de construções de indeterminação com os dois pronomes se inicia com [VOCÊ] associado primeiramente a predicados mais concretos e, depois, a predicados mais abstrato, e se estende, por último, a construções [a gente], hipótese postulada em base sincrônica e que demanda comprovação diacrônica.

Para uma apuração mais bem aprofundada, optamos por fazer uma análise estatística que nos pudesse fornecer mais detalhes em termos da produtividade dos predicados que ocorrem na construção de indeterminação.

A força colostrucional indica o grau de atração ou de repulsão de verbos pela construção de indeterminação e é apurada por meio do Programa R que, nos cálculos estatísticos da significância da força de atração, considera correlações entre a frequência com que um tipo de verbo ocorre no *córpus*, independentemente do tipo de construção em que ocorre, e sua frequência de ocorrência na construção de indeterminação, fornecendo, ao final, informações sobre: (i) expectativa de ocorrência (maior a frequência de *tokens* no *córpus*, maior a expectativa de ocorrência na construção); (ii) atração ou repulsão do verbo pela construção; (iii) a força colostrucional, calculada pelo *log likelihood* (maior a diferença entre a expectativa de ocorrência e o total de ocorrências na construção e no *córpus*, maior a força); (iv) o nível de significância da força (variando de 0 a 5, sendo 0 referente a número de ocorrências insuficientes para o cálculo da força colostrucional).

Por meio de análise colostrucional de colexemas simples (verificação do preenchimento de apenas um slot da construção, o slot PRED de cada uma das construções tomadas isoladamente), interessa-nos, aqui, tratar dos tipos de verbos que são mais fortemente atraídos pelas duas construções de indeterminação. Antes, porém, é importante indicar que, no slot PRED da construção de indeterminação, foi apurada a possibilidade de ocorrência de 174 verbos, 162 dos quais ocorrem combinados somente com VOCÊ (174 – 12 verbos sem ocorrências), 52, somente com A GENTE (174 – 122 sem nenhuma ocorrência), e 40 são

comuns às duas construções, o que já aponta um forte indício de maior abstratização de VOCÊ, como estratégia de indeterminação, do que de A GENTE, por serem menores os contextos de restrições na formação da construção, podendo o pronome ocorrer com tipos mais variados de predicados.

Desses 40 tipos comuns, interessa-nos aqui a análise construcional somente daqueles que apresentam, **ao mesmo tempo**, algum nível de significância para as duas construções, de modo a ser possível uma comparação mais segura entre elas, em termos de força de atração. É o que mostramos na tabela 8 abaixo.

**Tabela 8** - Força construcional de tipos de predicados, com nível de significância entre 1 e 5, para as construções de indeterminação de sujeito com A GENTE e VOCÊ

Pronome	Slot PRED	Frequência no cópua	Frequência na construção	Expectativa de ocorrer na construção	Atração ou repulsão	Força construcional (log likelihood)	Nível de significância
---------	-----------	---------------------	--------------------------	--------------------------------------	---------------------	--------------------------------------	------------------------

VOCEÊ	ter X	704	60	1,6	attr	326,90027	5
VOCEÊ	ter	567	47	1,3	attr	252,49277	5
VOCEÊ	ver	638	28	1,5	attr	114,21245	5
VOCEÊ	estar	329	20	0,8	attr	94,17462	5
A GENTE	ter	567	13	0,2	attr	85,30725	5
VOCEÊ	fazer	1313	28	3	attr	76,14045	5
A GENTE	ver	638	10	0,2	attr	57,75441	5
A GENTE	ter X	704	8	0,2	attr	40,98597	5
VOCEÊ	chegar	433	12	1	attr	38,17567	5
VOCEÊ	ler	52	5	0,1	attr	28,07167	5
VOCEÊ	usar	102	6	0,2	attr	27,74237	5
VOCEÊ	escutar	31	4	0,1	attr	24,89824	5
VOCEÊ	saber	849	12	2	attr	23,73778	5
VOCEÊ	precisar	99	5	0,2	attr	21,60789	5
A GENTE	estar	329	4	0,1	attr	20,86591	5
VOCEÊ	aprender	63	4	0,1	attr	19,09133	4
VOCEÊ	deixar	40	3	0,1	attr	15,31668	4
A GENTE	fazer	1313	5	0,5	attr	15,07757	3
A GENTE	precisar	99	2	0	attr	12,39755	3
A GENTE	repetir	8	1	0	attr	9,91582	2
A GENTE	tremer	8	1	0	attr	9,91582	2
A GENTE	saber	849	3	0,3	attr	8,58757	2
VOCEÊ	escolher	43	2	0,1	attr	8,31282	2
A GENTE	amar	18	1	0	attr	8,2269	2
A GENTE	acordar	21	1	0	attr	7,91244	2
VOCEÊ	plantar	57	2	0,1	attr	7,22627	2
A GENTE	escutar	31	1	0	attr	7,12461	2
A GENTE	chegar	433	2	0,2	attr	6,69144	2
A GENTE	deixar	40	1	0	attr	6,61363	1
A GENTE	escolher	43	1	0	attr	6,46928	1
VOCEÊ	repetir	8	1	0	attr	6,15346	1
VOCEÊ	tremer	8	1	0	attr	6,15346	1
A GENTE	ler	52	1	0	attr	6,0913	1
A GENTE	plantar	57	1	0	attr	5,90942	1
A GENTE	aprender	63	1	0	attr	5,7117	1
A GENTE	usar	102	1	0	attr	4,76875	1
VOCEÊ	conseguir	115	2	0,3	attr	4,64933	1
A GENTE	conseguir	115	1	0	attr	4,53666	1
VOCEÊ	amar	18	1	0	attr	4,50367	1
VOCEÊ	acordar	21	1	0	attr	4,20096	1

Fonte: elaboração própria.

Como se observa na tabela 8, pela força colustrucional com que são atraídos para o slot PRED de cada uma das construções de indeterminação, os verbos se arranjam de modo diferente, do mais para o menos fortemente atraído, conforme mostramos em (34):

**(34) Arranjo dos tipos de verbo, segundo sua força colustrucional**

a. para construções com VOCÊ

*ter X > ter > ver > estar > fazer > chegar > ler > usar > escutar > saber > precisar > aprender > deixar >*

|-----[nível 5 de significância]-----| |-----nível 4 de significância-----|

*escolher > plantar >*

*repetir > tremer > conseguir > amar > acordar*

|--nível 2 de significância--|

|-----nível 1 de significância-----|

b. para construções com A GENTE

*ter > ver > ter X > estar >*

*fazer > precisar >*

|----[nível 5 de significância]----| |-----nível 3 de significância----|

*repetir > tremer > saber > amar > acordar > escutar > chegar > deixar > escolher > ler > plantar > aprender > usar > consegui*

|-----nível 2 de significância-----| |-----nível 1 de significância-----|

Enquanto a força atração de verbos para o slot PRED de construções de indeterminação com [VOCÊ] varia de 4,2 a 326,9, para construções com *a gente*, essa variação vai de 4,5 a 85,3, o que evidencia que o slot PRED da primeira tem mais força de atração do que o da segunda. Argumentamos que esse tipo de resultado está fortemente atrelado aos traços semânticos das construções pronominais, uma vez, pelo princípio da persistência (HOPPER, 1991), ainda persistem traços de primeira pessoa no pronome *a gente*, mesmo que bastante diluídos, enquanto, em construção de indeterminação do sujeito, já não há traço de segunda pessoa no pronome [VOCÊ]. De fato, isso contribui para confirmar nossa hipótese sobre a construção com o pronome *você* ser mais frequente, dada a sua trajetória rumo a uma maior abstratização, perdendo seu traço dêitico.

Atendo-nos aos tipos de verbos que atingem significância máxima, observamos, entretanto, que, dentro desses dois arranjos dados em (34), os cinco primeiros tipos de verbos, pela maior força colostrucional e nível máximo de significância que apresentam, conferem alta coesão dos dois tipos de construção na indeterminação do sujeito, porque são os mais fortemente atraídos para o slot PRED. É de se destacar ainda que a atração de *ter X > ter > ver > estar* é mais forte (força acima de 94) em construções de indeterminação com [VOCÊ] do que a primeira mais forte, *ter* (força de 85,3), para construção com *a gente*.

Ainda considerando o nível máximo de significância, o slot PRED de construções de indeterminação com [VOCÊ] vai atrair verbos que as construções com *a gente* atrairão somente abaixo desse nível. É o caso dos verbos *chegar > ler > usar > escutar > saber > precisar X*. Isso significa dizer que no nível máximo de significância, a variedade de tipos de verbo é maior

para construções com [VOCÊ] (11 tipos) do que para as com *a gente* (4 tipos, apenas). Variedade de tipos para *a gente* começa a ser maior somente abaixo do nível 3 de significância.

#### 4.4.3 Tipos de Sequência Discursiva

Considerando a proposta apresentada por Freitag (2014) elaborada a partir do trabalho de Oliveira e Silva e Macedo (1996), procedemos à análise dos dados sobre indeterminação do sujeito investigando a hipótese de que sequências discursivas do tipo *injuntiva* e *argumentativa* favorecem o uso de sujeito indeterminado: o tipo *injuntivo*, justamente por apresentar tempos menos marcados, como o presente do indicativo sendo empregado para orientar ou instruir um procedimento; o tipo *argumentativo*, por caracterizar trechos em que o falante, ao defender algum ponto de vista, que possa comprometê-lo, recorre a uma estratégia que generaliza sua opinião e, conseqüentemente, o isenta da responsabilidade pelo conteúdo enunciado. A tabela a seguir apresenta os resultados.

**Tabela 9** - Frequência das construções de indeterminação do sujeito por sequência discursiva

[SNpro]	[A gente]		[VOCÊ]		Σ	
	Freq.	Proporção (%)	Freq.	Proporção (%)	Freq.	Proporção (%)
Argumentativa	95	26,46	269	74,93	364	<b>47,70</b>
Injuntiva	2	0,85	231	99,14	233	<b>30,53</b>
Descritiva	4	2,39	163	97,6	167	21,88
Σ	101	13,23	662	86,76	763	100

Fonte: elaboração própria.

Diferentemente do que conjecturamos, ao supor que sequências injuntivas seriam as mais frequentes, a sequência discursiva em que mais se concentram ambas as construções é a *argumentativa* (47% = 359/763), seguida, em ordem, da *injuntiva* (30,5% = 233/763) e da *descritiva*, (21,9% = 167/763). As ocorrências apresentadas a seguir exemplificam casos de sujeito indeterminado com os três tipos de sequências discursivas: *argumentativa* (35.a), *injuntiva* (35.b) e *descritiva* (35.c).

- (35) a. éh num tô sendo pessimista sabe?... mas o nosso país ele vai melhorá(r) um dia... quem sabe?... éh:: isso depende não só deles... mas da gente também um po(u)co... acho que se ***a gente* fi/ começasse a prestá(r) mais atenção... no nos/ no nosso povo**::... na nossa::... na nossa conquista que a gente podia chegá(r) até hoje... aí ia melhorar ia ma::s/. [BDI-AC-035; RO: L. 509-512].

b. as coisas eu preciso tá c'a receita grudada nos olhos pra fazê(r) certo... então... ela me ensinô(u)... que **a gente coloca... uma lata... de leite condensado...** uma lata de leite.... [AC-036; RP: L. 266-269].

c. na sala tem:: [...]uma cade(i)ra do papai que sempre foi meu sonho de tê(r) [Doc.: ((risos))] e e:: então é uma cade(i)ra muito cômoda né? você... u/ u/ você **ce pode sentá(r) nela você pode deitá(r) nela você pode dormi(r)... e você se sente muito bem...** numa cade(i)ra desse tipo... e:: a minha cozinha tem um armário... éh:: éh:: desses pré-moldados... **um armário onde você:: define ele onde ele vai sê(r).** [BDI-AC-099; DE: L. 269-274].

Esse parâmetro de análise parece ser pouco esclarecedor em relação à produtividade das construções. Sua análise nos dá mais subsídios para a caracterização da construção em si e revela que as sequências discursivas que mais favorecem o uso de sujeitos indeterminados são, de fato, as *argumentativas* e as *injuntivas*, justamente pela própria definição delas, que preveem, para o primeiro tipo, a presença de marcadores de opinião em tempo de presente e na primeira pessoa, como *eu acho*, em (35.a), e, para o segundo, pelo uso de sujeitos genéricos e de tempos de presente. Como mencionado anteriormente, desses dois tipos de sequência discursiva mais frequentes, as injuntivas restringem mais a combinação do pronome de referência indeterminada com tipos de predicados semânticos mais concretos, pela própria natureza precisa das instruções codificadas nos predicados. Na próxima seção, cuidamos do cruzamento de parâmetros relevantes para melhor caracterização das duas construções.

#### **4.5 Quadro geral da produtividade de construção de indeterminação com [VOCÊ]**

Até aqui, exploramos os parâmetros que nos permitiram verificar o grau de sedimentação na rede de indeterminação do sujeito que cada construção apresenta, investigando isoladamente cada parâmetro de análise, mas já podemos assegurar que a construção [[VOCÊ][PREDICADO]] está mais sedimentada na rede linguística do que a predicação com *a gente*. Diante desse resultado, focalizamos nesta seção os contextos de maior produtividade dessa construção específica, por julgarmos que resultados da produtividade só se completam com a localização dos contextos mais específicos em que ela ocorre.

#### 4.5.1. Formato da sentença vs. Grau de referência

Com o cruzamento dos parâmetros *formato da oração* e *grau de referência do SNpro*, pretendemos verificar se, de fato, há correlação forte entre sentenças não-finitas e referência genérica e indefinida da construção. A tabela abaixo apresenta os resultados:

**Tabela 10** - Formato Oracional X Grau do referente da construção com [VOCÊ]

[SNpro]		[VOCÊ]	
		Freq.	Proporção (%)
Finito	Genérico definido	54/593	9,10%
	Genérico indefinido	539/593	<b>90,89%</b>
	Subtotal	593/662	89,58%
Não-finito	Genérico definido	2/69	2,89%
	Genérico indefinido	67/69	<b>97,10%</b>
	Subtotal	69/662	10,42%
$\Sigma$		662	100%

Fonte: Elaboração própria.

Como se pode observar, o traço genérico indefinido é o que predomina em tipo oracional tanto finito quanto não-finito. Embora a frequência de ocorrência de dados seja muito superior para orações finitas ( $593/662 = 89,52\%$ ), a porcentagem de referência genérica indefinida é maior para o tipo oracional não-finito ( $67/69 = 97,10\%$ ) do que para o finito ( $539/593 = 90,89\%$ ). Como postulamos anteriormente, o tipo de formato não-finito é o mais inovador na rede linguística de indeterminação do sujeito, e é relevante observar que esse tipo de contexto já aponta que o grau de referência genérico indefinido é o ambiente semântico mais propício para indeterminar o sujeito, revelando, assim, que há uma abstratização maior do grau de referência da predicação com [VOCÊ], o que também justifica essa construção ser a mais frequente.

#### 4.5.2. Feixe semântico do predicado vs. Grau de referência

Mais uma vez com o intuito de identificar contextos semânticos mais propícios à indeterminação do sujeito, cruzamos os parâmetros *feixe semântico do predicado* e *grau de referência*. Os resultados são os mostrados na Tabela 11.

**Tabela 11** - Feixe semântico do predicado X Grau de referência da construção com [VOCÊ]

[SNPro]/Grau de referência Feixe Semântico do Predicado	[VOCÊ]	
	Genérico definido	Genérico indefinido
Momentâneo	-	9
Atividade específica	35	328
<i>Dicendi</i>	-	7
Atividade difusa	6	85
Instância	-	2
Estímulo mental	-	2
Ev. trans. intencional	-	5
Ev. trans. não intencional	-	1
Processo	-	2
Experimentação mental	2	34
Atenuação	-	-
Relacional	4	47
Relacional modal	9	45
Sensação corporal	-	6
Existência	-	-
Estado	-	33
$\Sigma$	56	606

Fonte: elaboração própria.

Mais uma vez, o resultado do cruzamento desses parâmetros aponta que o grau de referência das construções de indeterminação encontra sua preferência em tipo genérico indefinido. Conforme mostram os resultados da tabela 11, construções com grau de referência genérica indefinida tem maior variabilidade de tipos semânticos de predicado, o que não acontece com construções com grau de referência genérico definido, restritos apenas a cinco tipos semânticos de predicado. O tipo de referência genérica e indefinida encontra-se sedimentado em praticamente todos os tipos de feixes semânticos, não ocorrendo em apenas dois dos tipos da escala de Freitag e Tavares (2014), os quais também não aparecem com construções de grau genérico definido.

O que podemos observar também é que a concentração de ocorrência está em feixes semânticos de predicados mais concretos, como de *atividade específica*, para referência genérica definida e indefinida, e de *atividade difusa*, para referência genérica indefinida.

Em termos de distribuição de ocorrência em tipos de predicados, podemos observar que no ponto mais abstrato da escala, há uma certa concentração de ocorrência em tipos *relacional* e *relacional modal*.

#### 4.5.3. Feixe semântico vs. Sequência discursiva

O cruzamento dos parâmetros *feixe semântico do slot PRED* e *sequência discursiva* nos permite identificar se quais sequências discursivas favorecem mais a presença de certos tipos semânticos de predicados do que outras. A tabela 12 mostra o resultado desse cruzamento.

**Tabela 12** - Feixe semântico X Sequência discursiva da construção com [VOCÊ]

[SNPro]/Sequência Discursiva	[VOCÊ]			
	Argumentativa	Expositiva	Descritiva	Narrativa
Feixe Semântico do Predicado				
Momentâneo	3	3	3	-
Atividade específica	140	183	40	-
Dicendi	2	2	3	-
Atividade difusa	63	18	10	-
Instância	-	-	2	-
Estímulo mental	2	-	-	-
Ev. trans. intencional	1	1	3	-
Ev. trans. não intencional	1	-	-	-
Processo	1	-	1	-
Experimentação mental	19	10	6	1
Atenuação	-	-	-	-
Relacional	24	14	13	-
Relacional modal	27	15	11	1
Sensação corporal	3	1	2	-
Existência	-	-	-	-
Estado	15	11	6	1
$\Sigma$	301	258	100	3

Fonte: elaboração própria.

Os resultados apresentados na tabela 12 mostram que os vários tipos semânticos de predicados ocorrem em mais sequências discursivas *argumentativas* (301 dados). No ponto da escala de feixes semânticos mais concretos, o tipo semântico do predicado mais frequente é o de *atividade específica* em sequência discursiva *expositiva* (183 dados), seguida das sequências *argumentativa* (140 dados) e *descritiva* (40 dados). A frequência em sequência discursiva *narrativa* é nula, nesse caso.

A sequência discursiva *argumentativa* apresenta maior variabilidade de ocorrência com os feixes semânticos, deixando de apresentar apenas em três tipos: *instância*, *atenuação* e *existência*. A sequência *descritiva*, uma das que concentra menor número de ocorrências, ocupa

o segundo lugar em termos de variabilidade de tipos semânticos de predicados, com 12 tipos. Embora a sequência discursiva *expositiva* apresente o maior número de ocorrência em tipo semântico *atividade específica* (183 ocorrências), ela fica em terceiro lugar em relação a variabilidade de tipos semânticos, apresentando ocorrência em 10 tipos, enquanto a sequência *narrativa* é a que conta com o menor número de ocorrência, apresentando três dados e em três tipos semânticos.

Interessante observar que o tipo semântico *relacional modal* ocorre em todas as sequências discursivas. Mesmo com baixa frequência, isso aponta que é nesse tipo de feixe semântico que a construção é mais sedimentada por ser estar relacionada com a modalidade deôntica, em que a focalização, na construção de indeterminação do sujeito, é orientada para o evento.

## CONCLUSÕES

A fim de verificar o fenômeno de indeterminação do sujeito no português falado do interior paulista, nesta dissertação nos ocupamos da descrição e da análise das microconstruções de indeterminação do sujeito instanciadas por dois pronomes nominativos: *[[a gente] [PREDICADO]]* e *[[VOCÊ] [PREDICADO]]*. Nosso objetivo foi o de verificar, na sincronia atual, o grau de produtividade dessas duas construções, norteado pela hipótese de que a construção com *[VOCÊ]* seria mais produtiva do que a com *[a gente]*, dada as trajetórias de mudança diacrônica por elas experimentadas até se construcionalizarem como construções pronominais e continuarem a se submeter a mudanças pós-construcionalização.

Para alcançar nosso objetivo principal, conduzimos nossa investigação em dois momentos: primeiro, partimos dos estudos sobre processos de mudança envolvendo os pronomes em foco e seus reflexos na sintaxe do PB, especificamente no que diz respeito ao preenchimento da posição de sujeito, segundo resultados obtidos pelos vieses da Sociolinguística Paramétrica e do Funcionalismo. À luz dos MBU, reinterpretamos esses resultados como decorrentes da fixação do padrão de sujeito explícito na rede de referência determinada do PB, um esquema que, já disponível na gramática do PB, se sedimenta pela sua produtividade, se expande e atinge também construções de sujeito de referência indeterminada.

O modelo de análise adotado se mostrou apropriado para tratar do tema, uma vez que foi possível reinterpretar as mudanças concebendo a língua como redes interligadas em que as propriedades construcionais de esquematicidade, de produtividade e de composicionalidade puderam ser aplicadas às construções sob análise a partir dos parâmetros de forma e de significado. Desse modo, contribuímos com o modelo teórico na descrição de mudanças sintáticas mais amplas, ainda pouco investigadas, especificamente mudanças na rede de indeterminação do sujeito, e, assim, contribuímos também com outros estudos que delineiam a emergência de uma gramática brasileira, porém em perspectiva teórica diferente daquelas que mais comumente têm explorado esse fenômeno, como a da Sociolinguística paramétrica.

A mudança na rede de indeterminação do sujeito no PB pode ser atestada a partir da criação do subesquema *[[SNPro] [PREDICADO]]* que sanciona microconstruções com *[SNPro]* instanciados pelas construções *[a gente]* e *[VOCÊ]*. Em relação às propriedades intervenientes nessa construção no nível do Subesquema, constatamos que ela é altamente esquemática, pois, no *slot* de *[SNPro]*, é possível ocorrer, além dos pronomes aqui investigados,

diversos outros pronomes nominativos do quadro pronominal do PB, inclusive o *eu*, e no *slot* de [PREDICADO] diversos tipos semânticos de predicados.

Em termos de produtividade, por ser completamente esquemática, [[SNPro] [PREDICADO]] é altamente produtiva, uma vez que instancia diversos *types* no nível da microconstrução. Sua composicionalidade semântica é opaca, ou seja, é não-composicional, dada a presença de um pronome em seu esquema cuja propriedade primeira é a de fazer referência específica na interação verbal. Sintaticamente, ela é composicional, pois os falantes conseguem reconhecer aí a relação entre o sujeito e o predicado. A noção de composicionalidade sintática contribuiu para o entendimento de que o PB está se encaminhando para uma língua em que o padrão disponível com sujeito preenchido é o mais usado. O resultado da alta frequência de uso desse padrão só poderá ser visto futuramente, quando se constatará se o PB perderá o padrão de sujeito não-preenchido ou não.

A pesquisa empírica apresentou resultados que corroboram a hipótese aventada neste trabalho. A microconstrução [[VOCÊ] [PREDICADO]] é a mais produtiva em termos de frequência de ocorrência e de frequência de tipo em todos os contextos analisados, quais sejam: (i) com sentenças finitas e não-finitas; (ii) com as infinitivas de completivas adjetivais se sobressaindo aos demais tipos oracionais; (iii) com referência genérica e indefinida dos pronomes; (iv) com predicados de natureza mais concreta, como os do feixe semântico *atividade momentânea*, embora, em menor frequência, possa se combinar também com predicados de estado pertencentes ao feixe mais abstrato; (v) com predicados que expressam tempos do presente; (vi) em sequências discursivas argumentativa e injuntiva.

Argumentamos que a referência do sujeito indeterminado é realizada a partir do esquema predicador cujo pareamento da construção de indeterminação do sujeito no PB com pronomes nominativos é o [[SNPro][PREDICADO]↔[PREDICAÇÃO]]. Esse posicionamento corrobora a noção de Pina (2009) que afirma que, em estruturas de indeterminação do sujeito, o predicado é sempre focalizado, e o sujeito, por sua vez, sempre desfocalizado. A análise colostrucional apresentada na seção 4.4.2 se revelou bastante eficaz para sustentar essa argumentação, uma vez que revelou que o tipo de verbo mais atraído para o *slot* do predicado é o que chamamos *relacional modal deôntico*, que, nas construções sob análise, constrói uma predicação orientada para o evento, colocando o sujeito em segundo plano.

Diante desse resultado, comprovando nossa hipótese inicial, argumentamos que é justamente por conta da trajetória de mudanças que essas duas proformas percorreram que o pronome *VOCÊ* é o mais propício a ser recrutado para operar indeterminação do sujeito. O pronome *a gente*, antes de entrar para o quadro pronominal, já era usado como forma de indeterminação, como afirma (LOPES, 2003); ao integrar o paradigma pronominal, seu caráter generalizador se torna mais restrito à medida que ele adquire traços dêiticos de referência à 1ª. pessoa do discurso plural e, nesse ponto, se abstratiza como forma de referência indeterminada. *VOCÊ* em direção oposta, emerge de uso dêitico contextualmente definido e com referência específica e invariável, e, ao longo do tempo, tem sua foricidade abstratizada em favor de usos dêiticos de sentido constante, mas de referência variável na identificação da segunda pessoa do discurso; posteriormente, adquire traços que tornam ainda mais abstrato seu valor referencial, em favor de uma referência indeterminada.

Dada a complexidade do tema, em que vários fenômenos sintáticos estão interligados nessas mudanças, podemos esboçar alguns encaminhamentos futuros de pesquisa:

1. a investigação diacrônica que comprove hipótese forte por nós aventada de que construções de indeterminação emergem primeiramente com [*VOCÊ*] e só posteriormente com [*a gente*];
2. a investigação da competição entre os diversos construtos instanciados pela microconstrução [[*VOCÊ*] [PREDICADO]], a saber, *você*, *ocê* e *cê*. Nesse caso, poderia se investigar a rede de indeterminação do sujeito instanciada por pronomes nominativos, em que, ao esquema mais geral [[SNPro] [PREDICADO]], se subordinaria o subesquema [[*VOCÊ*] [PREDICADO]], que, por sua vez, sancionaria as microconstruções com [*você*], [*ocê*] e [*cê*], buscando-se, assim, possíveis contextos de especialização de cada uma dessas variantes;
3. a investigação diacrônica das construções em sentenças de infinitivo não-pessoais parece ser bastante promissora. De todos os efeitos causados na rede de indeterminação, esse é o mais inesperado. Duarte (2008) afirma que parece não haver notícias de línguas ocidentais com sujeitos pronominais nominativos expressos em sentenças de infinitivo impessoal, e que o PB desenvolveu tal padrão (cf. seção 4.3.2. deste trabalho). Investigar a emergência desse padrão dentro dos MBU é uma instigante tarefa que pode ampliar o entendimento da categoria de indeterminação do sujeito;

4. Duarte (2003) investiga construções existenciais no PB e apresenta dados em que os pronomes *a gente* e *you* recebem interpretação de referência nula, a exemplo de expletivos de línguas como o inglês e o francês, um tipo de construção que merece também investigação mais aprofundada.

O que constatamos até aqui é que, estudar a referencialidade desses pronomes não se encerra com os resultados alcançados nesta dissertação. Um estudo mais amplo contemplando as perspectivas acima aventadas pode contribuir para uma melhor caracterização desses pronomes em diferentes redes do PB.

## REFERÊNCIAS

- BARÐDAL, J. *Productivity: Evidence from Case and Argument Structure in Icelandic*. Amsterdã: Benjamins, 2008.
- BARLOW, M.; KEMMER, S. (ed.). *Usage-based models of language*. Stanford, CA: CSLI publications, 2000.
- BECHARA, E. *Moderna gramática portuguesa*. 37 ed. Rio de Janeiro: Nova Fronteira, 2009.
- BENVENISTE, É. O aparelho formal da enunciação. In: BENVENISTE, É. *Problemas de lingüística geral II*. Campinas: Pontes, 1989. p. 81-92.
- BERGEN, B. K.; CHANG, N. Embodied Construction Grammar in simulation-based language understand. In: OSTMAN, J.-O.; FRIED, M. (Eds.). *Construction Grammars: Cognitive grounding and theoretical extensions*. Amsterdam: John Benjamins, 2005, p. 147–190.
- BOAS, H. C.; SAG, I. *Sign-Based Construction Grammar*. Stanford: CSLI Publications, 2012.
- BYBEE, J. *Language, usage and cognition*. Cambridge: CUP, 2010.
- BYBEE, J. *Língua, uso e cognição*. tradução de Maria Angélica Furtado da Cunha; revisão de Sebastião Carlos Leite Gonçalves. São Paulo: Cortez, 2016.
- BYBEE, J. Mechanisms of change in grammaticalization: the role of frequency. In: JANDA, R.; BRIAN, J. (eds.). *Handbook of historical linguistics*. Oxford: Blackwell, 2005, p. 602-23.
- BYBEE, J. *Usage-based theory and exemplar representations of constructions*. In: The Oxford handbook of construction grammar. 2013.
- BYBEE, J.; McCLELLAND, J. L. Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *Special Issue of The Linguistic Review*, vol. 22, p. 381-410, 2005.
- CAMACHO, R. G.; DALL'AGLIO-HATTNER, M. M.; GONÇALVES, S. C. L. O substantivo. In: CASTILHO, A. T.; ILARI, R.; NEVES, M. H. M. (Org.) *Gramática do português culto falado no Brasil: classes de palavras e processos de construção*. Campinas: Editora da Unicamp, 2008. p. 21-84.
- CÂMARA JUNIOR, J. M. *Dicionário de Lingüística e Gramática: referente à língua portuguesa*. 13. ed. Petrópolis: Vozes, 2002.
- CASTILHO, A. T. *Nova gramática do português brasileiro*. São Paulo: Contexto, 2010.

CAVALCANTE, S. R. O. *A indeterminação do sujeito na escrita padrão: a imprensa carioca nos séculos XIX e XX*. 1999. Dissertação (Mestrado em Letras Vernáculas) Faculdade de Letras - UFRJ. Rio de Janeiro, Brasil, 1999.

CAVALCANTE, S. R. O. *O uso de se com infinitivo na história do português*. 2006. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, UNICAMP, Campinas, Brasil, 2006.

CINTRA, L. F. L. *Sobre "formas de tratamento" na língua portuguesa* (ensaios). Lisboa: Horizonte, 1972.

CROFT, W. *Typology and universals*. Cambridge, United Kingdom: CUP, 1990.

CROFT, W. Construction grammar. In: GEERAERTS, D.; CUYCKENS, H. (eds.). *The Oxford Handbook of Cognitive Linguistics*. Oxford: Oxford University Press, 2007a, p. 463-508.

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

CULICOVER, P. W.; JACKENDOFF, R. *Simpler Syntax*. Nova York: Oxford University Press, 2005.

CUNHA, A. G. *Dicionário etimológico Nova Fronteira da língua portuguesa*. Rio de Janeiro: Nova Fronteira Editora. 1982.

CUNHA, C. de S. *Indeterminação pronominal do sujeito*. 1993. Dissertação (Mestrado em Língua Portuguesa). Rio de Janeiro, Faculdade de Letras/UFRJ. 1993.

CUNHA, C.; CINTRA, L. *Nova gramática do português contemporâneo*. 3. ed. Rio de Janeiro: Nova Fronteira, 2001.

CYRINO, S. M. L. *O objeto nulo do português do Brasil: um estudo sintático-diacrônico*. 1994. Tese. (Doutorado em Linguística) Instituto de Estudo da Linguagem, UNICAMP. Campinas. 1994.

DIK, S. C. *The theory of functional grammar*. Dordrecht-Holland / Providence RI: Foris Publications, 1989.

DU BOIS, J. W. Beyond definiteness: the trace of identity in discourse. In: CHAFE, W. (ed.). *The pear stories*. Norwood: Ablex, 1980, p. 203-274.

DUARTE, M. E. L. *A perda do princípio "Evite Pronome" no português brasileiro*. 1995. Tese (Doutorado em Linguística), Instituto de Estudos da Linguagem, UNICAMP, Campinas, Brasil, 1995.

DUARTE, M. E. L. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In: ROBERTS, I.; KATO, M. A. (org.). *Português brasileiro: uma viagem diacrônica*. Campinas: Editora da UNICAMP, 1993. p. 107-128.

DUARTE, M. E. L. A evolução na representação do sujeito pronominal em dois tempos. In: PAIVA, M. C.; DUARTE, M. E. L. *Mudança Linguística em Tempo Real*. Rio de Janeiro: Contra Capa, 2003, p. 115-128.

DUARTE, M. E. L. O sujeito de referência indeterminada em sentenças infinitivas. *Revista do GEL*, v. 5, n. 1, p. 9-30, 2008.

FARACO, C. A. *O tratamento de você em português: uma abordagem histórica*. *Fragmenta*, Curitiba, n.13, p. 51-82, 1996.

FILLMORE, C. J. The mechanisms of “Construction Grammar”. *Proceedings of the 14th annual meeting of the Berkeley Linguistics Society*, p. 35-55, 1988.

FILLMORE, C. J.; KAY, P.; O’CONNOR, M. C. Regularity and idiomaticity in grammatical constructions: the case of let alone. *Language*, 64, 3, p. 501-538, 1988.

FIORIN, J. L. Sendas e veredas da semiótica narrativa e discursiva. *D.E.L.T.A.: Documentação de Estudos em Lingüística Teórica e Aplicada*, v. 15, p. 177-207, 1999.

FLORENCIO, J. H. N. *A indeterminação do sujeito em editoriais jornalísticos do Recife*. 2015. Dissertação (Mestrado em Letras) - Universidade Federal de Pernambuco, Recife, 2015.

FREITAG, R. M. K. Dissecando a entrevista sociolinguística: estilo, sequência discursiva e tópico. In: GORSKI, E., COELHO, I. L.; SOUZA, C. M. N. (ed.). *Variação estilística: reflexões teórico-metodológicas e propostas de análise*. Florianópolis: Insular, 2014, p. 125-141.

GAFFIOT, F. *Dictionnaire latin-français*. Paris: Hachette, 1934.

GISBORNE, N.; PATTEN, A. Construction grammar and grammaticalization. *The Oxford handbook of grammaticalization*, 2011.

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

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

GOLDBERG, A. E. Constructionist approaches. In: HOFFMANN, S.; TROUSDALE, G. (eds.). *The Oxford Handbook of Construction Grammar*. Nova York: Oxford University Press, 2013, p. 15-31.

GONÇALVES, S. C. L. *Banco de dados Iboruna: amostras eletrônicas do português falado no interior paulista*. 2007. Disponível em: <https://www.alip.ibilce.unesp.br>.

GONÇALVES, S. C. L.; GALVÃO, V. C. C.; SOUSA, G. C. As construções subordinadas substantivas. In: NEVES, M. H. M.; CASTILHO, A. T. (Orgs.) *Gramática do português falado culto no Brasil: classe de palavras e processos de construção*. 1ª ed. Campinas: Editora da Unicamp, 2008, p. 1021-1084.

GONÇALVES, S. C. L. O estatuto variável de construções com e sem alçamento; uma abordagem sociofuncionalista. *D.E.L.T.A. Documentação de Estudos em Linguística Teórica e Aplicada*, v. 33, p. 1291-1321, 2017.

GONÇALVES, S. C. L. Por uma abordagem de construções complexas em perspectiva construcional. *Working papers em Linguística*, Florianópolis, v. 21, n.1, p. 102-127, 2020.

GUERRA, A. R. *Diacronia do grau de transparência do sistema de referência por expressão pronominal e desinencial do argumento-sujeito de 1ª e 2ª pessoas no português brasileiro*. 2017. Tese (Doutorado em Estudos Linguísticos). Universidade Estadual Paulista “Júlio de Mesquita Filho”, Instituto de Biociências, Letras e Ciências Exatas, São José do Rio Preto, Brasil, 2017.

HAIMAN, J. Ritualization and the development of language. In: PAGLIUCA, W. (ed.). *Perspectives on grammaticalization*. Amsterdã: Benjamins, 1994, p. 3-28.

HASPELMATH, M. Parametric versus functional explanation of syntactic universalisals. In: BIBERAUER, T. (ed.). *The Limits of Syntactic Variation*. Amsterdã: Benjamins, 2008, p. 75-107.

HILMELMANN, N. P. Lexicalization and grammaticalization: Opposite or orthogonal? In: BISANG, W.; HIMMELMANN, N. P.; WIEMER, B. (eds.). *What makes Grammaticalization – A look from its Fringes and its Components*. Berlim: Mouton de Gruyter, 2004, p. 21-42.

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

HOPPER, P. J. On some principles of grammaticization. In: TRAUGOTT, E. C.; HEINE, B. (eds.). *Approaches to grammaticalization*, Volume I, Amsterdam/Philadelphia, John Benjamins Company. 1991.

HOPPER, P. J.; TRAUGOTT, E. C. *Grammaticalization*. Cambridge University Press. 1993.

HOPPER, P. J.; TRAUGOTT, E. C. *Grammaticalization*. Cambridge University Press, 2003.

HUDSON, R. A. *English Word Grammar*. Oxford: Blackwell, 1990.

HUDSON, R. A. *Word Grammar*. Oxford: Blackwell, 1984.

HUDSON, R.A. *Language Networks: The New Word Grammar*. Oxford: Oxford University Press, 2007.

INFANTE, U. *Curso de gramática aplicada aos textos*. Scipione, 1995.

KATO, M. A.; MIOTO, C. A inexistência de sujeitos oracionais. *Laços*, Rio de Janeiro, p. 61-90, 2000.

KAY, P. The kind of / sort of construction. *Proceedings of the 10th annual meeting of the Berkeley Linguistics Society*, p. 157-171, 1984.

KOCH, I. V.; ELIAS, V. M. *Ler e escrever: estratégias de produção textual*. São Paulo: Contexto, 2009.

KOELLING, S. B. Os dêiticos e a enunciação. *Revista Virtual de Estudos da Linguagem – ReVEL*, ano 1, n. 1, p. 1-12. 2003.

LABOV, William. *Principles of Linguistic Change*. Volume 1: Internal Factors. Cambridge: Blackwell Publishers. 1994.

LAKOFF, G. *Women, Fire and Dangeours Things – What Categories Reveal about the Mind*. Chicago: University of Chicago Press, 1987.

LANGACKER, R. W. *Cognitive Grammar*. A Basic Introduction. Nova York: Oxford University Press, 2008.

LANGACKER, R. W. *Foundations of Cognitive Grammar: Descriptive application*. Stanford: University Press, 1991.

LANGACKER, R. W. *Foundations of Cognitive Grammar: Theoretical prerequisites*. Stanford: University Press, 1987.

LOPES, C. R. S. *Nós e a gente no português falado culto do Brasil*. 1993. Dissertação (Mestrado em Língua Portuguesa) - Faculdade de Letras, Universidade Federal do Rio de Janeiro - UFRJ, Rio de Janeiro, 1993.

LOPES, C. R. dos S. *A inserção de “a gente” no quadro pronominal do português*. v. 18. Frankfurt/Madrid: Vervuert/Iberoamericana, 2003.

LOPES, C. R. Processo evolutivo de ‘Vossa Mercê’ > ‘Você’ (português) e ‘Vuestra Merced’ > ‘Usted’ (espanhol). *II Congresso Internacional da ABRALIN - Associação Brasileira de Lingüística*. Fortaleza: publicação on-line w [ww.lettras.ufrj.br/abralin](http://ww.lettras.ufrj.br/abralin). 2001.

LOPES, C. R. S.; DUARTE, M. E. L. Notícias sobre o tratamento em cartas escritas no Brasil dos séculos XVIII e XIX. In: RAMOS, J. M.; ALKMIN, M. (Orgs.) *Para a história do*

*português brasileiro: Estudos sobre Mudança Linguística e História Social*, v. 5, 2007, p. 329-357.

LOPES, C. R. S.; RUMEU, M. C. B. O quadro de pronomes pessoais do português: as mudanças na especificação dos traços intrínsecos. *Descrição, história e aquisição do português brasileiro*, v. 1, p. 419-436, 2007.

LOPES, C. R.; DUARTE, M. E. L. De ‘Vossa Mercê’ a ‘Você’: análise da pronominalização de nominais em peças brasileiras e portuguesas setecentistas e oitocentistas. In: BRANDÃO, S. F.; MOTA, M. A. (Orgs.) *Análise Contrastiva de Variedades do Português: Primeiros Estudos*. Rio de Janeiro: In-fólio, 2003b, p. 61-76.

LUFT, C. P. *Moderna gramática brasileira*. Globo Livros, 2002.

LYONS, John. *Semantics*. Cambridge: Cambridge University Press, 1977. [v.2]

MARCUSCHI, L. A.; KOCH, I. G. V. Referenciação. In: CASTILHO, A. T. (ed.). *Gramática do português culto falado no Brasil*. Campinas: Ed. da UNICAMP, v. 1, 2006, p. 381-399.

MENON, O. P. S. A gente: um processo de gramaticalização. *Estudos linguísticos*, São Paulo, n. 25, v. único, p. 622-628, 1996.

MENON, O. P. S. A indeterminação do sujeito no português do Brasil: NURC-SP e VARSUL. In: VANDRESEN, P. (Org.). *Varição, mudança e contato linguístico no português da região sul*. Pelotas: EDUCAT, 2006. p. 125-167.

MENON, O. P. S. *Analyse sociolinguistique de l'indétermination du sujet dans le portugais parlé au Brésil a partir des données du NURC/SP*. (Tese de Doutorado). Université de Paris VII, Paris, 1994.

MILANEZ, W. *Recursos de indeterminação do sujeito*. 1982. Tese (Doutorado em Linguística). Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 1982.

MORAES SILVA, A. *Diccionario de Língua Portuguesa*. Lisboa: Typografia Lacerdina. 1813.

NARDELLI, A. J. S. *Estudo socioestilístico da alternância pronominal entre nós e a gente no português falado no noroeste paulista*. 2021. Dissertação (Mestrado em Estudos Linguísticos). Universidade Estadual Paulista - UNESP, São José do Rio Preto, 2021.

NASCENTES, A. *O Linguajar Carioca*. Rio de Janeiro, Edição da Organização Simões/Coleção Rex. 1953.

NASCENTES, A. O tratamento de VOCÊ no Brasil. *Letras*. Curitiba: n. 5-6, p. 114-122, 1956.

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

NEVES, M. H. M. *A gramática: história, teoria e análise, ensino*. Unesp, 2002.

NEVES, M. H. M., Os pronomes. In: CASTILHO, A. T.; ILARI, R.; NEVES, M. H. M. (Org.) *Gramática do português culto falado no Brasil: classes de palavras e processos de construção*. Campinas: Editora da Unicamp, 2008. p. 507-616.

NUNES, J. *O famigerado se: uma análise sincrônica e diacrônica das construções com se apassivador e indeterminador*. 1990. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, UNICAMP, Campinas, 1990.

OLIVEIRA E SILVA, G. MACEDO, G. Análise sociolinguística de alguns marcadores conversacionais. In: MACEDO, A., RONCARATI, C.; MOLLICA, M. (Orgs.). *Variação e discurso*. Rio de Janeiro: Tempo Brasileiro, 1996, p.11-49

OLIVEIRA, T. P. Subject expression in Brazilian Portuguese. In: KEIZER, E.; OLBERTZ, H. (Org.). *Recent Developments in Functional Discourse Grammar*. 1ed. Amsterdam: John Benjamins Publishing, 2018, p. 208-232.

OMENA, N. P. A referência variável da primeira pessoa do discurso no plural. In: NARO, A. J. et alii: *Relatório Final de Pesquisa: Projeto Subsídios do Projeto Censo à Educação*, Rio de Janeiro, UFRJ, p. 286–319. 1986.

OMENA, N. P. As influências sociais na variação entre nós e a gente na função de sujeito. In: OLIVEIRA e SILVA, G. M.; SCHERRE, M.M.P. *Padrões sociolinguísticos*. Rio de Janeiro: Tempo Brasileiro, 1996, p. 309-323.

OMENA, N. P. A referência à primeira pessoa do plural: variação ou mudança? In: PAIVA, M. de C.; DUARTE, M. E. L. (orgs.) *Mudança linguística em tempo real*. Rio de Janeiro: Contra Capa Livraria, 2003, p. 60-78.

OMENA, N. P. de; BRAGA, M. L. *A gente está se gramaticalizando?* In: MACEDO, A. T.; RONCARATI, C.; MOLLICA, M. C. (Org.). *Variação e discurso*. Rio de Janeiro: Tempo Brasileiro, 1996. p. 75-83.

OTHERO, G. Revisitando o status do pronome *cê* no português brasileiro. *Revista de Estudos da Linguagem*, v. 21, n. 1, p. 135-156, 2013.

PAREDES SILVA, V. L. *Cartas cariocas: a variação do sujeito na escrita informal*. 1988. Tese (Doutorado em Língua Portuguesa). Faculdade de Letras, UFRJ, Rio de Janeiro. 1988.

PAWLEY, A.; SYDER, F. H. Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In: RICHARDS, J. C.; SCHMIDT, R. W. (eds.). *Language and Communication*. Londres: Longman, 1983, p. 191-225.

PEREK, F. 2016. Using distributional semantics to study syntactic productivity in diachrony: a case study. *Linguistics*, p. 149–188, 2016.

PINA, A. A. de. *Por uma nova "percepção" sobre um antigo problema: estudo cognitivo-construcional do sujeito indeterminado com se em português*. Tese (Doutorado) – Faculdade de Letras. Universidade Federal do Rio de Janeiro, Rio de Janeiro. 2009.

PINHEIRO, D.; ALONSO, K. 30 anos (ou mais) de Gramática de Construções: primeiros apontamentos para uma história do movimento construcionista (ou: 1988: o ano que não terminou), *Revista Linguística*. v. 14, n. 1, p. 6-29. 2018.

PREZOTTO JÚNIOR, J. R. *As microconstruções auxiliares com "deixar" e "parar" no português na expressão de aspecto final*. 2020. Dissertação (Mestrado em Estudos Linguísticos). Universidade Estadual Paulista (Unesp), Instituto de Biociências Letras e Ciências Exatas, São José do Rio Preto, 2020.

ROBUSTE, T. B. *Construções [VI+VER] no português brasileiro contemporâneo sob perspectiva construcional*. 2018. Tese. (Doutorado em Estudos Linguísticos). Universidade Estadual Paulista (Unesp), Instituto de Biociências Letras e Ciências Exatas, São José do Rio Preto, 2018.

ROCHA LIMA, C. H. *Gramática normativa da língua portuguesa*. Rio de Janeiro: José Olympio, 2000.

ROLLEMBERG, V. *et al.* Os pronomes pessoais sujeito e a indeterminação do sujeito na norma culta de Salvador. *Estudos Linguísticos e Literários*, Salvador, n. 11, v. único, p. 53-74, 1991.

RUBIO, C. F.; GONÇALVES, S. C. L. Opções metodológicas para o estudo de fenômenos variáveis relacionados à 1ª pessoa do discurso. *Gragoatá* (UFF), v. 29, p. 161-182, 2010.

RUBIO, C. F. *Padrões de concordância verbal e de alternância pronominal no português brasileiro e europeu: estudo sociolinguístico comparativo*. 2012. Tese (Doutorado em Estudos Linguísticos) – Universidade Estadual Paulista (Unesp), Instituto de Biociências Letras e Ciências Exatas, São José do Rio Preto, 2012.

RUMEU, M. C. B. *Para uma história do português no Brasil: formas pronominais e nominais de tratamento em cartas setecentistas e oitocentistas*. Rio de Janeiro, Dissertação (Mestrado em Língua Portuguesa), Faculdade de Letras/ UFRJ, Rio de Janeiro. 2004.

RUMEU, M. C. B. Traços formais e semântico-discursivos no processo de Gramaticalização de 'Vossa mercê' > 'você'. *Revista do GEL*, v. 3, p. 67-82, 2006.

RUMEU, M. C. B. *A implementação do 'Você' no Português Brasileiro Oitocentista e Novecentista: um estudo de painel*. 2008. Tese (Doutorado em Língua Portuguesa) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2008.

SAID ALLI, A. De “eu” e “tu” a “majestade” (Tratamentos de familiaridade e reverência). *Revista de Cultura*. n. 129. Rio de Janeiro: Vozes de Petrópolis, p. 137-151, 1937.

SCHERRE, M. M. P.; DIAS, E. P.; ANDRADE, C. Q.; LUCCA, N. N. G.; ANDRADE, A. L. V. S. *Tu, você, cê e ocê na variedade brasileira. Pávia (Brasília)*, v. Esp, p. 117-134, 2011.

SCHERRE, M. M. P.; NARO, A. J.; YACOVENCO, L. C. Nós e a gente em quatro amostras do Português Brasileiro: revisitando a escala da saliência fônica. *Diadorim*, Rio de Janeiro, vol. 20 – Especial, p. 428-457, 2018.

SETTI, A. C. R. *A indeterminação do sujeito nas três capitais do Sul do Brasil*. 1997. Dissertação. (Mestrado em Letras) Faculdade de Letras e Artes. Universidade Federal do Paraná, Paraná, 1997.

SOARES, B. G. *Mudança na rede construcional do sintagma nominal para pronome: a construcionalização de a gente*. 2018. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2018.

SOUZA, E. M. *Sujeitos de referência arbitrária: uma classe homogênea?* 2013. Tese. (Doutorado em Estudos Linguísticos). Faculdade de Letras da Universidade Federal de Minas Gerais. Belo Horizonte, 2013.

STEELS, L. *Design patterns in Fluid Construction Grammar*. Amsterdam: John Benjamins, 2011.

TAMANINE, A. M. B. *A alternância nós/a gente no interior de Santa Catarina*. Dissertação (Mestrado em Letras). Faculdade de Ciências e Letras, Universidade Federal do Paraná, 2002.

TAVARES, M. A.; FREITAG, R. M. K. Do concreto ao abstrato: influência do traço semântico-pragmático do verbo na gramaticalização em domínios funcionais complexos. *Revista Linguística*: Rio de Janeiro, revista do programa de pós-graduação em linguística da Universidade Federal do Rio de Janeiro, v. 6, n. 1, p. 103-119, jun. 2010

TAYLOR, J. R. *Cognitive Grammar*. Oxford: Oxford University Press, 2002.

TOMASELLO, M. *Constructing a language*. Harvard universitypress, 2009.

TRAUGOTT, E. C. *Exaptation and grammaticalization*. Linguistic studies based on corpora, 2004.

TRAUGOTT, E. C. Toward a coherent account of grammatical constructionalization. In: TRAUGOTT, E. C. (Org.). *Diachronic construction grammar*, 2015, p. 51-79.

TRAUGOTT, E. C., TROUSDALE, G. *Construcionalização e mudanças construcionais*. Tradução de Taísa Peres de Oliveira e Angélica Furtado da Cunha. Petrópolis, RJ: Vozes, 2021.

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

VARGAS, A. de S. C. *Estratégias pronominais de indeterminação: um estudo diacrônico*. 2010. Dissertação (Mestrado em Letras Vernáculas) Faculdade de Letras - UFRJ. Rio de Janeiro, Brasil, 2010.

VITRAL, L. A forma *Cê* e a noção de gramaticalização. *Revista de Estudos da Linguagem*, n. 4, v. 1. UFMG, p. 116-124, 1996.

VITRAL, L.; RAMOS, J. *Gramaticalização: uma abordagem Formal*. Rio de Janeiro: Tempo Brasileiro, 2006.

WEINREICH, U., LABOV, W.; HERZOG, M. I. Empirical foundations for a theory of language change. In: LEHMANN, W.; MALKIEL, Y. *Directions for historical linguistics*. University of Texas Press. 1968.