

**UNIVERSIDADE ESTADUAL DO SUDOESTE DA BAHIA (UESB)
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA (PPGLIN)**

GILVANEI DE OLIVEIRA SOUZA

**VOCÊ E A GENTE NO PORTUGUÊS FALADO DE VITÓRIA DA CONQUISTA-BA:
UMA ANÁLISE SOCIOFUNCIONALISTA DO FENÔMENO DE
INDETERMINAÇÃO DO SUJEITO**

VITÓRIA DA CONQUISTA – BA

2018

GILVANEI DE OLIVEIRA SOUZA

**VOCÊ E A GENTE NO PORTUGUÊS FALADO DE VITÓRIA DA CONQUISTA-BA:
UMA ANÁLISE SOCIOFUNCIONALISTA DO FENÔMENO DE
INDETERMINAÇÃO DO SUJEITO**

Dissertação apresentada ao Programa de Pós-Graduação em Linguística (PPGLin), da Universidade Estadual do Sudoeste da Bahia (UESB), como requisito parcial e obrigatório para obtenção do título de Mestre em Linguística.

Área de Concentração: Linguística

Linha de Pesquisa: Análise e descrição de Línguas Naturais

Orientador: Valéria Viana Sousa

VITÓRIA DA CONQUISTA – BA

2018

Souza, Gilvanei de Oliveira.

S715v Você e a gente no Português falado de Vitória da Conquista – Ba: uma análise sociofuncionalista do fenômeno de indeterminação do sujeito. / Gilvanei de Oliveira Souza, 2018.

121f.

Orientador (a): Dra. Valéria Viana Sousa.

Dissertação (mestrado) – Universidade Estadual do Sudoeste da Bahia, Programa de Pós-Graduação em Linguística – PPGLin, Vitória da Conquista, 2018.

Inclui referência F. 107 – 111.

1. Gramaticalização – Você e A gente. 2. Indeterminação do sujeito. 3. Sociofuncionalismo. 4. Português falado – Vitória da Conquista - BA. I. Sousa, Valéria Viana. II. Universidade Estadual do Sudoeste da Bahia, Programa de Pós-Graduação em Linguística. T. III.

CDD: 469.3

Catálogo na fonte: Juliana Teixeira de Assunção – CRB 5/1890
UESB – Campus Vitória da Conquista – BA

Título em inglês: *Você e a gente* in the spoken Portuguese of Vitória da Conquista: a sociofunctionalistic analysis of indeterminacy of the subject.

Palavras-chave em inglês: Sociofuncionalism. Indetermination. Grammaticalization. *Você* and *A gente*.

Área de concentração: Linguística

Titulação: Mestre em Linguística

Banca examinadora: Prof^ª Dr^a Valéria Viana Sousa (Presidente-Orientadora); Prof. Dr. Jorge Augusto Alves da Silva (UESB); Prof. Dr. Camilo Rosa Silva (UFPB)

Data da defesa: 06 de abril de 2018

Programa de Pós-Graduação: Programa de Pós-Graduação em Linguística.

GILVANEI DE OLIVEIRA SOUZA

**VOCÊ E A GENTE NO PORTUGUÊS FALADO DE VITÓRIA DA CONQUISTA-BA:
UMA ANÁLISE SOCIOFUNCIONALISTA DO FENÔMENO DE
INDETERMINAÇÃO DO SUJEITO**

Dissertação apresentada ao Programa de Pós-Graduação em Linguística (PPGLin), da Universidade Estadual do Sudoeste da Bahia (UESB), como requisito parcial e obrigatório para obtenção do título de Mestre em Linguística.

Data da aprovação: 06 de abril de 2018.

Banca Examinadora:

Profa. Dra. Valéria Viana Sousa (Presidente)
Instituição: UESB

Ass.: Valéria Viana Sousa

Prof. Dr. Jorge Augusto Alves da Silva
Instituição: UESB

Ass.: Jorge Augusto Alves da Silva

Prof. Dr. Camilo Rosa da Silva
Instituição: UFPB

Ass.: Camilo Rosa da Silva

Aos meus pais (Seny e Juarez), meus exemplos de força, perseverança e resistência. Sou grato por terem me garantido o acesso à educação.

AGRADECIMENTOS

São tantos nomes para ocupar esse espaço que provavelmente eu vou me esquecer de alguns. Mas, mesmo correndo o risco do esquecimento, não posso deixar de reconhecer que muitas mãos escreveram esse texto comigo ao longo desses dois anos.

À Universidade Estadual do Sudoeste da Bahia (UESB) e ao Programa de Pós-Graduação em Linguística (PPGLin), por abrir espaço e investir no nosso/meu crescimento acadêmico.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo apoio e financiamento das atividades do PPGLin da UESB.

Agradeço a Profa. Dra. Valéria Viana Sousa por ter sido tão humana, compreensiva, acessível e por estar sempre disponível para dar o suporte necessário quando necessário. A leveza das aulas e orientações combinada com a sua voz calma fizeram com que eu encarasse a árdua jornada da pesquisa com mais serenidade. Obrigado por me acolher, minha orientadora.

Ao Prof. Dr. Jorge Augusto Alves da Silva e à Profa. Dra. Elisângela Gonçalves da Silva pelas contribuições dadas no exame de qualificação e ao Prof. Dr. Camilo Rosa da Silva pela disponibilidade em participar da banca de defesa e pelas contribuições que com certeza darão ao trabalho mais sustentação.

Aos professores do PPGLin que, com empenho e competência, fizeram com que eu ampliasse a minha visão acerca da linguagem humana e me tornasse mais capaz de dar conta da pesquisa.

Aos funcionários do PPGLin pela atenção de sempre.

Sou grato aos meus pais que, desde muito cedo, lutaram com tudo o que puderam para manter os cinco filhos na escola. Sou grato a eles por abrirem essa porta e por terem garantido que ela nunca se fechasse até eu conseguir andar com as próprias pernas. Vocês são o meu exemplo de força.

Sou grato aos meus irmãos pelo incentivo e pela união que me traz o conforto de saber que sempre vou ter com quem contar.

Agradeço a Heloísa, minha amiga/irmã que esteve e está sempre ao meu lado. Sem a sua crença na minha força e capacidade eu não encararia tudo o que encaro do mesmo modo. Sou muito grato a você por ter sido a minha família desde a graduação.

Agradeço a Geovana, uma das melhores pessoas que conheci em Conquista e na vida. Sempre com uma explicação para tudo, você foi o meu ponto de equilíbrio. Eu levo muito de você comigo, tenha certeza!

Agradeço também a André pelo auxílio constante e pela disponibilidade em me direcionar sempre quando precisei.

Agradeço a todos os meus colegas de curso, especialmente a Nayara e Savanna por compartilharem as angústias e pesos da vida acadêmica. Entre uma disciplina e outra, um evento e outro a força de vocês foi imprescindível.

Sou grato a Warley e Vânia Raquel por terem sido o meu socorro quando precisei. Obrigado pelas dicas e pelas dúvidas sanadas.

Por fim, sou grato a Deus pela felicidade de poder contar com tantas pessoas boas na minha vida. Reconheço em cada uma dessas pessoas a presença do próprio Deus.

Enfim, minha gratidão a todos e todas que, de algum modo (por meio de alguma palavra de incentivo ou simplesmente por torcer pelo meu sucesso), contribuíram para a realização desse sonho.

RESUMO

O presente trabalho é resultado de um estudo que tem como objetivo primário investigar o modo como acontece a indeterminação do sujeito por meio das formas gramaticalizadas *você* e *a gente* no Português Culto e Popular falado de Vitória da Conquista - BA. Pretendemos, aqui, ampliar a abordagem acerca do fenômeno em questão, verificando como, em seus processos de mudança via gramaticalização, as formas *você* e *a gente* adquiriram a função de indeterminar o sujeito, considerando a influência de variáveis linguísticas (*superestrutura textual, mudança ou manutenção do referente e tempo verbal*) e extralinguísticas (*sexo, faixa etária e grau de escolaridade*) no processo de escolha do falante por qualquer uma das formas de indeterminar supracitadas. Para dar conta desses objetivos, ancoramo-nos na orientação de pesquisa sociofuncionalista, que resulta da junção de pressupostos teóricos-metodológicos de duas correntes linguísticas: a Sociolinguística e o Funcionalismo Norte-Americano, formando, assim, o Sociofuncionalismo. A amostra usada é composta por 24 entrevistas extraídas dos *corpora* de Português falado de Vitória da Conquista -BA (*corpora* PPVC e PCVC) e os dados localizados passaram por análise estatística probabilística por meio do programa *GoldVarb X*. No processo de análise, foram localizadas 1261 ocorrências das duas formas em posição de sujeito na amostra, sendo 637 delas de sujeito indeterminado. Os resultados obtidos a partir da análise desses dados permitiram inferir que (i) tanto o *você* quanto o *a gente* atuam enquanto estratégias de indeterminação do sujeito no Português falado de Vitória da Conquista – BA; (ii) essa atuação sofre influência de fatores linguísticos e sociais; (iii) há contextos específicos em que a função indeterminadora é empregada e outros também específicos nos quais a função prototípica das formas prevalece; (iv) a função de indeterminar, comum às duas formas, é resultado do processo de gramaticalização pelo qual passaram, já que observamos uma relação entre a atuação de princípios como a *descategorização semântica* e a *persistência* (típicos da mudança via gramaticalização) e a função inovadora agregada às duas formas, evidenciando o paralelo entre gramaticalização e indeterminação.

PALAVRAS-CHAVE

Sociofuncionalismo. Indeterminação. Gramaticalização. Você e A gente.

ABSTRACT

The present work is the result of a study that has as primary objective to investigate the way in which the indetermination of the subject happens through the grammaticalized forms *você* and *a gente* in the Formal and Popular Portuguese spoken of Vitoria da Conquista - BA. We intend here to broaden the approach to the phenomenon under study by establishing a direct relation between the process of change through grammaticalization by which the forms *você* and *a gente* have passed and the function of indeterminating the subject that both acquired, considering the influence of linguistic variables (text superstructure, change or maintenance of referent and verbal tense) and extra linguistic (gender, age group and schooling level) in the process of choosing from the speaker by any of the above indeterminate forms. In order to fulfill these goals, we are anchored in the orientation of sociofunctionalist research, which results from the junction of theoretical and methodological assumptions of two linguistic currents: Sociolinguistics and North American Functionalism and, thus, Sociofunctionalism. The sample used is composed by 24 interviews extracted from the corpora of Portuguese spoken of Vitória da Conquista -BA (corpora PPVC and PCVC) and the data located underwent by probabilistic statistical analysis through the program GoldVarb X. In the process of analysis, 1261 occurrences of both forms in subject position were located, 637 of them being of indeterminate subject. The results obtained from the analysis of these data allowed us to infer that (i) both *você* and *a gente* act as strategies of indeterminacy of the subject in the spoken Portuguese of Vitoria da Conquista - BA; (ii) this act is influenced by linguistic and social factors; (iii) there are specific contexts in which the indeterminating function is used and others, also specific, in which the prototypical function of forms prevails; (iv) the indeterminate function, common to both forms, is a result of the grammaticalization process they have gone through, since we observe a relationship between the principles acting as semantic decategorization and persistence (typical of the change through grammaticalization) and an innovative function aggregated to the two forms, evidencing the parallel between grammaticalization and indetermination.

KEYWORDS

Sociofuncionalism. Indetermination. Grammaticalization. *Você* and *A gente*.

LISTA DE GRÁFICOS

Gráfico 01 – Distribuição geral das ocorrências de <i>você</i> e <i>a gente</i> em posição de sujeito nos corpora PCVC e PPVC.....	81
Gráfico 02 – <i>Você</i> e <i>a gente</i> : sujeito indeterminado nos corpora PCVC e PPVC.....	81
Gráfico 03 – <i>Você</i> e <i>a gente</i> : distribuição geral e relação determinado x indeterminado.....	83
Gráfico 04 – Ocorrências de <i>você</i> determinado e indeterminado nos corpora PCVC e PPVC.....	85
Gráfico 05 - Ocorrências do <i>a gente</i> determinado e indeterminado nos corpora PCVC e PPVC.....	90

LISTA DE QUADROS

Quadro 1 - Parâmetros de Lehmann.....	36
Quadro 2 - Os princípios de Hopper.....	37
Quadro 03 - Sociolinguística e Funcionalismo: o que une as teorias.....	40

LISTA DE TABELAS

Tabela 01 - <i>Você e a gente</i> indeterminados nos corpora PCVC e PPVC.....	82
Tabela 02 - <i>Você</i> indeterminado em relação à <i>superestrutura textual</i>	86
Tabela 03 - <i>Você</i> indeterminado em relação à <i>mudança ou manutenção do referente</i>	88
Tabela 04 - <i>A gente</i> indeterminado em relação à <i>superestrutura textual</i>	94
Tabela 05 - <i>A gente</i> indeterminado em relação à <i>faixa etária</i>	96
Tabela 06 - <i>A gente</i> indeterminado em relação à <i>escolaridade</i>	97
Tabela 07 - <i>Superestrutura textual x escolaridade</i> : combinando fatores.....	97
Tabela 08 - <i>A gente</i> indeterminado em relação ao <i>tempo verbal</i>	99
Tabela 09 - <i>Tempo verbal x superestrutura textual</i> : combinando fatores.....	101

SUMÁRIO

1 INTRODUÇÃO	14
2 SOCIOLINGUÍSTICA + FUNCIONALISMO = SOCIOFUNCIONALISMO: O QUE SE QUER DA(S) TEORIA(S).	17
2.1 Sociolinguística Variacionista: a língua na perspectiva social	17
2.1.1 A teoria da variação e mudança.....	20
2.2 O estudo da língua no Funcionalismo: a linguística centrada no uso.....	23
2.2.1 A origem do Funcionalismo linguístico	24
2.2.2 O Funcionalismo norte-americano: de onde falamos?	25
2.2.3 Gramática e interação no Funcionalismo	27
2.2.3.1 A gramática emergente de Hopper	27
2.2.3.2 O modelo de interação verbal de Dik.....	29
2.2.4 Princípios da gramaticalização aplicados à indeterminação.....	31
2.2.4.1 Iconicidade e marcação - Givon.....	32
2.2.4.2 Os princípios de Hopper.....	36
2.3 Interesses afins = um “casamento”, o Sociofuncionalismo	37
3 – INDETERMINAÇÃO DO SUJEITO: DOS CONCEITOS ÀS CLASSIFICAÇÕES, DOS DICIONÁRIOS ÀS GRAMÁTICAS	42
3.1 Indeterminação e conceitos afins nos dicionários.....	42
3.2 A indeterminação do sujeito nas gramáticas	45
3.2.1 Nas gramáticas históricas	46
3.2.2 A indeterminação na Tradição Gramatical.....	48
3.2.3 A indeterminação na Tradição Linguística.....	51
3.2.3.1 O que dizem as gramáticas descritivas?.....	51
3.2.3.2 O que dizem as pesquisas?	54
3.3 A gente e você – gramaticalização, referência e indeterminação.....	59
3.3.1 De Vossa Mercê a você: um breve histórico pela rota da mudança	59
3.3.2 De Gente a A gente – relação entre gramaticalização e indeterminação.....	62
4 METODOLOGIA.....	66
4.1 Descrição dos corpora.....	66
4.2 Passo a passo de análise	67
4.3.1 Variável dependente.....	68
4.3.2 Variáveis independentes linguísticas	69

4.3.2.1 <i>Superestrutura textual</i>	69
4.3.2.2 <i>Tempo verbal:</i>	70
4.3.2.4 <i>Mudança ou manutenção do referente</i>	71
4.3.3 <i>Variáveis independentes extralinguísticas ou sociais</i>	71
4.3.3.1 <i>Faixa etária</i>	72
4.3.3.2 <i>Grau de escolaridade</i>	72
4.3.3.3 <i>Sexo</i>	72
5 VOCÊ E A GENTE EM POSIÇÃO DE SUJEITO NO PORTUGUÊS FALADO DE VITÓRIA DA CONQUISTA: UMA ANÁLISE SOCIOFUNCIONALISTA DO FENÔMENO DE INDETERMINAÇÃO	73
5.1 <i>Primeira rodada: você e a gente indeterminado x determinado</i>	76
5.2 <i>Você e a gente: o que favorece a indeterminação?</i>	79
5.2.1 <i>Você indeterminado x você determinado</i>	80
5.2.1.1 <i>Superestrutura textual</i>	81
5.2.1.2 <i>Mudança ou manutenção do referente</i>	83
5.2.2 <i>A gente indeterminado x a gente determinado</i>	85
5.2.2.1 <i>Superestrutura textual</i>	87
5.2.2.2 <i>Faixa etária</i>	90
5.2.2.3 <i>Escolaridade</i>	90
5.2.2.4 <i>Tempo verbal</i>	93
CONCLUSÕES.....	96
REFERÊNCIAS	100
ANEXOS	103
Anexo 1 – Roteiro de perguntas para entrevista	103
Anexo 2 – Ficha social do informante	106

1 INTRODUÇÃO

A indeterminação do sujeito é um fenômeno linguístico que, salvo a abordagem simplificada da Tradição Gramatical, é de difícil definição e delimitação. Tal fato decorre das múltiplas possibilidades de análise e classificação do fenômeno e, acima de tudo, dos conceitos que são evocados quando o objeto é posto em análise, afinal *de qual sujeito falamos quando estudamos indeterminação?* A resposta para essa pergunta pode levar a diferentes definições de indeterminação, uma vez que a definição de sujeito não é só uma. As Gramáticas Tradicionais (GTs) que, na verdade, beberam na fonte da Filosofia (PONTES, 1986), por exemplo, oscilam entre uma definição de sujeito como sendo (i) *termo da oração* ou (ii) *um ser sobre o qual se faz uma declaração/ sobre o qual recai a ideia contida no predicado*. Sobre a indeterminação, afirma-se ser um fenômeno que ocorre *quando não se quer ou não se pode identificar a quem o predicado da oração se refere* (BECHARA, 2009; ROCHA LIMA, 2010; CUNHA; CINTRA, 2008). Assim, há uma definição de indeterminação para dois conceitos de sujeito e, obviamente, ela não se aplica a todos os casos, pois assumir qualquer uma das duas definições de sujeito acima levaria a conceitos diferentes de indeterminação. Indeterminar um termo não é o mesmo que indeterminar um ser...

Em relação às estratégias de indeterminação do sujeito, que é o ponto central da nossa análise, as gramáticas, em sua maioria, fazem menção a duas estratégias através das quais o fenômeno pode ser codificado no Português Brasileiro: (1) quando o verbo é colocado na terceira pessoa do plural, sem que se refira a nenhum termo identificado anteriormente, e (2) quando o verbo surge acompanhado da partícula *se*, que atua como índice de indeterminação do sujeito, como se vê nos exemplos a seguir: (1) “**Reputavam-no** o maior comilão da cidade” (C. dos Anjos) (2) “**Precisa-se** do carvalho; não se precisa do caniço..” (C. dos Anjos).¹

A definição apresentada por parte desses manuais não contempla nenhum outro recurso de indeterminação do sujeito, seja na fala ou na escrita. Desse modo, os conceitos e exemplos apresentados neles acabam por tornar o trato desse fenômeno limitado a esses dois casos, apenas, sem levar em consideração outras possibilidades, principalmente na língua em uso.

Atraídos pela necessidade de ampliação da abordagem acerca do fenômeno de indeterminação, vendo em *você* e *a gente* duas estratégias indeterminadoras atuantes e percebendo uma possibilidade de alinhamento entre as trajetórias de mudança via gramaticalização das duas formas e os seus usos como estratégias indeterminadoras, propomo-

¹ Exemplos extraídos de Cunha e Cintra (2001).

nos, com essa análise, descrever como a indeterminação do sujeito vem ocorrendo no Português Culto e Popular de Vitória da Conquista – BA por meio, especificamente, de *você* e *a gente*.

O nosso objetivo é descrever o uso de *você* e *a gente* como sujeito indeterminado, observando como pressões sociais, contextuais e de uso condicionam tal uso. Na medida em que a análise se afunila, buscaremos respostas para os seguintes questionamentos norteadores (i) como e em quais contextos *você* e *a gente* atuam enquanto estratégias indeterminadoras? (ii) como se deu o processo de gramaticalização pelo qual as duas formas passaram e que garantiu a ambas a função de indeterminar o sujeito? (iii) qual o grau de influência de condicionadores linguísticos e sociais para uso desse itens linguísticos com a função de indeterminação do sujeito?

A fim de ter um ponto de partida, que pode ou não ser confirmado, lançamos as seguintes hipóteses (i) a função de indeterminar é inovadora nas duas formas e, portanto, essa função só entra em jogo em contextos específicos, sendo, em outros contextos, a função prototípica ainda largamente empregada; (ii) essa função inovadora, tanto do *você* quanto do *a gente*, é resultado do processo de mudança via gramaticalização e da atuação de princípios que geram, entre outros efeitos, o desbotamento/descoloramento semântico ocasionado pela rotinização do uso das formas e a persistência de traços das formas-fonte nas formas-alvo; (iii) tanto fatores linguísticos quanto extralinguísticos/sociais influenciam o emprego de *você* e *a gente* como sujeito indeterminado.

A fim de pôr à prova as hipóteses das quais partimos e de cumprir com os objetivos propostos, ancoramo-nos nos pressupostos teóricos de duas correntes: a Sociolinguística, uma vez que buscaremos no extralinguístico explicações/motivações para os usos de *você* e *a gente* como estratégias de indeterminação do sujeito e o Funcionalismo Norte-americano, já que entendemos a indeterminação do sujeito por meio de *você* e *a gente* como um resultado da mudança via gramaticalização das duas formas e, assim, do Sociofuncionalismo, orientação de pesquisa que se embasa na junção de pressupostos teóricos e metodológicos da Sociolinguística e Funcionalismo.

A análise proposta será realizada a partir de dados que espelham situação de interação real de uso da língua. Esses dados serão extraídos de 24 (vinte e quatro) entrevistas que seguem o padrão da Sociolinguística Quantitativa e o modelo de entrevista DID – Diálogo entre Informante e Documentador - que compõem os *corpora* do Português falado de Vitória da Conquista: Português Culto de Vitória da Conquista, *Corpus* PCVC, e Português Popular de Vitória da Conquista, *Corpus* PPVC, estratificados em *sexo*, *faixa etária* e *escolaridade*.

O nosso trabalho é constituído da Introdução e de outras 4 seções. Na introdução, apresentamos a dissertação de modo geral. Na seção *Sociolinguística + Funcionalismo = Sociofuncionalismo: o que se quer da(s) teoria(s)*, apresentamos os pressupostos da Sociolinguística Quantitativa a partir, principalmente, dos postulados de Labov (2008 [1972]); os pressupostos do Funcionalismo em sua vertente norte-americana a partir dos postulados de Hopper (1987); Hopper e Traugott (1993); Traugott (2014), Givon (1984) e Lehmann (1982); e, por fim, trataremos do “casamento” entre as duas teorias a partir de Tavares (2003; 2013).

Na seção *Indeterminação do sujeito: dos conceitos às classificações, dos dicionários às gramáticas*, trataremos da indeterminação desde a ótica dos dicionários com Bluteau (1789), Dubois (1973), Crystal (1941) e Camara Junior (1997), discutindo definições que envolvem o entendimento do que é o sujeito indeterminado; perpassando pelas Gramáticas Históricas com os postulados de Bueno (1944), Cruz (1948) Pereira (1945), Almeida (1961) Said Ali (1964;1971), Mattos e Silva (1989); tratando, também, da abordagem da Tradição Gramatical a partir de Rocha Lima (1962), Sacconi (1982), Cipro Neto e Infante (1998; 2003), Cunha e Cintra (2001), Bechara (2006; 2009) e encerrando a seção com a abordagem da Tradição Linguística a partir das gramáticas, com Perini (2010), Castilho (2010; 2016), Azeredo (2008), Ferrarezi Junior (2008) e com as pesquisas mais recentes que abordaram o fenômeno de indeterminação: Carvalho (2010), Assunção (2012) e Teixeira (2014). Além disso, trataremos, também nessa seção, do processo de gramaticalização de *você* e *a gente* e da descrição que diferentes teóricos e gramáticos fazem acerca das duas formas pronominais.

Na seção *Metodologia*, apresentaremos a composição dos *corpora* e da seleção da amostra em análise, do passo a passo metodológico da análise; descreveremos o envelope de variação, sua variável dependente e variáveis independentes linguísticas e extralinguísticas, acompanhadas de suas respectivas hipóteses.

Por fim, na seção *Você e a gente em posição de sujeito no português falado de Vitória da Conquista: uma análise sociofuncionalista do fenômeno de indeterminação*, apresentaremos os resultados das análises realizadas a partir das rodadas dos dados no programa *GoldVarb X*, acompanhados das interpretações da atuação e comportamento de cada variável e seus respectivos pesos relativos em relação à ocorrência do fenômeno de indeterminação do sujeito a partir de *você* e *a gente*.

Na conclusão, faremos um apanhado geral acerca de cada seção e da discussão como um todo, a fim de apresentar as hipóteses confirmadas e não confirmadas com a presente análise.

2 SOCIOLINGUÍSTICA + FUNCIONALISMO = SOCIOFUNCIONALISMO: O QUE SE QUER DA(S) TEORIA(S).

Nesta seção, trataremos das duas correntes teórico-metodológicas que norteiam e dão base para a análise, aqui, realizada: a Sociolinguística Variacionista e o Funcionalismo norte-americano, fazendo um apanhado geral sobre as suas histórias e trajetórias, os seus lugares no campo da linguística, suas perspectivas de análise da língua(gem) e alguns dos principais conceitos e princípios. Com isso, definiremos, em linhas gerais, o lugar de onde falamos e o arcabouço teórico e metodológico que utilizaremos para dar conta dos objetivos aqui colocados.

A discussão será iniciada com a delineação dos campos da Sociolinguística e do Funcionalismo, na subseção 2.1 – *Sociolinguística Variacionista: a língua na perspectiva social* e na subseção 2.2 – *O estudo da língua no Funcionalismo: a linguística centrada no uso*, subseções nas quais serão arrolados aspectos relacionados ao surgimento das teorias, suas propostas de investigação da língua(gem), alguns dos principais conceitos e princípios. Por fim, na subseção 2.3 – *Interesses afins e um “casamento”, o Sociofuncionalismo*, discorreremos sobre a orientação de pesquisa Sociofuncionalista, da sua proposta de investigação dos fenômenos de variação e mudança da língua e dos pontos comuns existentes entre a Sociolinguística e o Funcionalismo que fundamentaram a junção dos postulados das duas teorias.

Pela ótica da Sociolinguística, conseguimos captar a atuação das variáveis sociais na escolha de uma ou outra estratégia de indeterminação do sujeito (*a gente* e *você*). Paralelamente, a análise pelo viés funcionalista permitirá observar em que medida os princípios típicos da mudança via gramaticalização possibilitam o uso das formas *a gente* e *você* com a função de indeterminar o sujeito. Acreditamos ser essa abordagem integrada, propiciada pelo Sociofuncionalismo, a mais adequada para dar conta dos objetivos aqui propostos.

2.1 Sociolinguística Variacionista: a língua na perspectiva social

“Todo linguista reconhece que a língua é um fato social, mas nem todos dão a ênfase a esse fato” (LABOV, 2008, p. 302).

O reconhecimento de que a língua é um fato social é, de certa forma, comum na linguística, mas quando as investigações esbarram no fenômeno de mudança “[...] encontramos um grau muito diferente de preocupação com o contexto social em que essas mudanças ocorrem.” (LABOV, 2008, p. 302). Alguns linguistas, segundo Labov (2008), ampliam as suas

visões, a fim de abarcar o máximo de fatos sobre os falantes e seus comportamentos extralinguísticos, outros estreitam essa visão, por questões metodológicas, excluindo-os o máximo possível. Nesse sentido, percebemos que a consciência acerca da existência do fenômeno da mudança (e também da variação) linguística, de alguma forma, sempre existiu entre linguistas, mas há certos entraves quando as explicações dos referidos fenômenos envolvem aspectos de natureza extralinguística, sociais. Esses entraves se justificam pela natureza, escolha e prioridades metodológicas de cada corrente linguística ou, de um ponto de vista mais crítico, pela existência de “[...] barreiras ideológicas para o estudo da língua na vida diária” (LABOV, 2008, p. 13).

No despontar da Linguística enquanto ciência da linguagem, fato que se deu no início do século XX com a publicação póstuma do *Curso de Linguística Geral*, a concepção preponderante de língua estabelecida era de sistema autônomo, abstrato, dissociado de tudo o que é externo, em outras palavras, extralinguístico. Por esse caráter autônomo, é que a língua ocupou o posto de objeto de estudo da linguística em detrimento da fala² que foi concebida, no *Curso de Linguística Geral*, como a realização da língua, o lado individual da linguagem e, portanto, variável.

A realidade do estudo da relação língua-sociedade e do tratamento das influências que os fatores externos à língua exercem sobre o seu uso mudou a partir da década de 1960, quando os postulados do linguista William Labov deram origem à corrente linguística que se tornou referência no estudo da língua na perspectiva social: a Sociolinguística. O termo *Sociolinguística* surgiu pela primeira vez, segundo Calvet (2002), na década de 1950, mas sua consolidação enquanto corrente linguística se deu a partir de 1960, tendo como ponto de partida os trabalhos de Labov, Gumperz, Dell Hymes e William Bright que, na conferência intitulada *Sociolinguistics*, publicada em 1966, afirmaram estar o escopo da Sociolinguística na demonstração da existência de uma covariação sistemática entre a estrutura linguística e a estrutura social, rompendo, assim, com a visão imanente de língua vigente até então. Nesse contexto, duas obras foram cruciais para a consolidação da Sociolinguística enquanto corrente, a saber: *Fundamentos empíricos para uma teoria da mudança linguística*, publicado em 1968

² Na visão de Saussure (2006 [1916]), se ocuparia do estudo da fala a Estilística, a Linguística Externa. Saussure institucionalizou a distinção entre uma Linguística Interna oposta a uma Linguística Externa e foi essa institucionalização que dividiu os estudos linguísticos nas duas orientações que conhecemos hoje: a formal e a contextual (ou funcional), estando sob o escopo da primeira a maioria dos estudos desenvolvidos acerca da língua à época da publicação do *Curso*, fazendo com que o estudo da fala e, conseqüentemente, dos fatores sociais que lhes são intrínsecos fossem enfocados posteriormente.

por Uriel Weinreich, William Labov e Marvin Herzog e *Padrões sociolinguísticos*, publicado em 1972 por Labov.

É, ainda, válido ressaltar que, com a divisão proposta por Saussure e, como consequente, opção metodológica realizada, a relação língua-sociedade, embora já fosse reconhecida nos estudos estruturalistas, não foi, como era de se esperar, incluída nas análises de cunho formalista. O estudo acerca dessa relação só ganhou destaque na década de 1960 com o surgimento da Sociolinguística. Antes desse período, na verdade, já havia interesse acerca dessa inter-relação, e, por extensão, acerca do fenômeno de variação. Segundo Coelho *et al.* (2010) e Alkmin (2001), alguns autores do século XX já divergiam da proposta estruturalista saussureana e postulavam uma concepção social de língua. Entre esses teóricos estão Marr (1865-1934), Bakhtin (1895-1975) e Meillet (186-1936)³. Esses linguistas que procuraram demonstrar a relação entre língua, cultura e sociedade podem e, quiçá, devem, de certo modo, até serem considerados precursores da Sociolinguística.

Em suma, não se pode alegar que a relação entre língua e sociedade foi de todo ignorada no início do século XX, tampouco que teve a relevância devida, pois não foi alvo de interesse das correntes teóricas que nasceram com o alvorecer da Linguística Moderna, conforme supracitado. As correntes (Estruturalismo e Gerativismo), que encabeçaram esses primeiros momentos da linguística enquanto ciência, viam/veem a língua por um prisma em que os fenômenos de variação e mudança não são determinantes para o estudo do sistema diferentemente do que vem a propor a Sociolinguística no meio do século. O estudo desenvolvido por essas vertentes teóricas é, por natureza, interno, um *estudo imanente da língua*, segundo Costa (2011, p. 115), o que fez com que a dimensão externa da língua fosse investigada de modo mais profundo a partir de correntes/disciplinas como a Pragmática, que estuda a língua em uma dimensão contextual; a Sociolinguística Paramétrica, que agrega pressupostos da Sociolinguística Variacionista e do Gerativismo; e a Sociolinguística em si que se consolidou a partir da segunda metade do século XX, buscando investigar o modo segundo o qual condicionadores sociais influenciam o uso e evolução da língua, conforme discorreremos na subseção 2.1.1.

³ Aos teóricos mencionados por Coelho *et al.* (2010), Alkmim (2001, p. 24) acrescenta Marcel Cohen, Émile Benveniste e Romam Jacobson, que, segundo a autora, também se ocuparam de pensar a questão do social no campo da linguística.

2.1.1 A teoria da variação e mudança

Com base no que foi exposto na subseção anterior, a Sociolinguística pode ser definida como uma corrente que, partindo do pressuposto de que a linguagem é um fenômeno social, busca, nas variações dos contextos sociais, explicações para as variações inerentes ao sistema linguístico, conforme postula Camacho (2001). Seguindo essa definição, a proposta da Sociolinguística é a de lançar luz sobre a influência dos condicionadores sociais (*sexo, faixa etária, grau de escolaridade* do falante, entre outros) no uso da língua, divergindo do princípio das correntes formalistas que, para atender aos seus propósitos de pesquisa, priorizaram o estudo da língua enquanto forma, enquanto sistema autônomo.

Nessa perspectiva, a Sociolinguística “redefine” o objeto de estudo até então predominante na Linguística, passando a abarcar, ao lado da perspectiva linguística, o lado social (extralinguístico) da língua. Daí, ocupa-se de questões como, entre outras, bilinguismo, contato linguístico, língua minoritárias, variação e mudança linguística (CALVET, 2002). De todas essas questões que estão sob o escopo da Sociolinguística, dois fenômenos em particular nos interessam: os fenômenos de variação e mudança.

A variação linguística é comum a todas as línguas naturais. Ela pode ser definida como um fenômeno que consiste na existência de mais de uma forma desempenhando uma mesma função em um mesmo contexto linguístico. A mudança, por outro lado, é a consolidação da variação, é o fenômeno que marca o fim da concorrência entre as formas e a entrada de uma para o contexto antes “disputado”. Variação e mudança são, de modo geral, o carro-chefe da corrente Sociolinguística e qualquer discussão acerca desses fenômenos implica outras noções mais gerais que sustentam a teoria como (i) visão de língua e sua natureza, (ii) objeto de estudo, (iii) noção de falante, (iv) comunidade de fala e, de modo geral, conceitos e princípios que estão envolvidos na ocorrência desses fenômenos.

A primeira noção que se tem quando se propõe analisar a língua na perspectiva da Sociolinguística é a de que os fenômenos de variação e mudança são inerentes ao sistema. Por variação inerente, entendemos que, “[...] como o sistema linguístico é heterogêneo, (i) a variação é uma propriedade regular do sistema; (ii) o falante tem competência linguística para lidar com regras variáveis. ” (COELHO ET AL, 2010, p. 22). É por essa razão que, na Sociolinguística, há a defesa de que as variações devem ser sempre levadas em consideração na análise linguística.

Por serem vistas como inerentes, o objetivo do Sociolinguista é verificar as motivações internas e externas dessa variação e o estado em que a mudança se encontra, podendo ser:

- (i) inicial (variação estável) - Ponto em que duas formas ou duas funções ocorrem simultaneamente, figurando em um mesmo contexto linguístico sem que necessariamente haja predomínio de uma sobre a outra. Nesse ponto, há a variação linguística em si. A título de exemplo, podemos citar a indeterminação do sujeito que, na língua em uso, pode ser codificada de diferentes formas: *uso da terceira pessoa do plural sem referente anterior*, *uso de alguns sintagmas nominais com referência genérica*, *uso do infinitivo sem sujeito* e de formas pronominais como *a gente* e *você* com referência genérica, sendo que algumas dessas formas podem registrar maior ou menor número de ocorrências em alguns contextos, não significando, no entanto, que uma está sendo ou será suplantada por outra.
- (ii) Mudança em curso ou mudança em progresso - Quando uma variante começa a se sobressair em relação à outra, sendo empregada em maior número de contextos, indicando, portanto, que uma forma está perdendo espaço e que está sendo substituída por outra.
- (iii) Mudança final ou mudança efetiva (mudança linguística) - Ponto em que a mudança já foi implementada e que uma forma foi suplantada, cedendo espaço para outra forma inovadora que passa a figurar naquela posição, ou seja, já há substituição absoluta.

Nesse processo de monitoramento da mudança linguística, a Sociolinguística parte do pressuposto da inexistência de homogeneidade. Estudar cada estágio da variação e mudança mencionados anteriormente pelo viés sociolinguístico pressupõe reconhecer, acima de tudo, que a língua é heterogênea por natureza, uma vez que a comunidade de fala também é. Nessa mesma direção, Weinreich, Labov & Herzog (1968 [2006]) afirmam que um dos pressupostos da abordagem da Sociolinguística é de que “[...] numa língua que serve a uma comunidade complexa (i.e., real), a ausência de heterogeneidade estruturada é que seria disfuncional. ” (WEINREICH, LABOV & HERZOG, (2006 [1968], p. 100-101).

Essa comunidade de fala, a propósito, é, segundo Alkmim (2001. p. 31), dentro dos pressupostos da Sociolinguística quantitativa, um “[...] conjunto de pessoas que interagem verbalmente e que compartilham um conjunto de normas com respeito aos usos linguísticos”. Ela se caracteriza

[...] não pelo fato de se constituir por pessoas que falam do mesmo modo, mas por indivíduos que se relacionam por meio de redes comunicativas diversas, e que

orientam seu comportamento verbal por um mesmo conjunto de regras (ALKMIM, 2001, p. 31).

De modo complementar, Coelho *et al.* (2010) afirmam que

Não existe uma comunidade de fala homogênea, nem um falante-ouvinte ideal. Pelo contrário, a existência de variação e de estruturas heterogêneas nas comunidades de fala é um fato comprovado. Existe variação inerente à comunidade de fala – não há dois falantes que se expressam do mesmo modo, nem mesmo um falante que se expresse da mesma maneira em diferentes situações de comunicação. (COELHO ET AL. 2010, p. 17).

Nos moldes da Sociolinguística, o falante é real (nunca ideal) e a comunidade de fala é sempre heterogênea. Por essa razão, é que o falante é identificado a partir do seu perfil/estratificação social que, no contexto da pesquisa sociolinguística, recebe o nome de *grupo de fatores sociais* ou *extralinguísticos*. Uma explicação concisa para o monitoramento do uso da língua a partir dessa estratificação pode ser encontrada nas palavras de Sousa (2008) que afirma:

Ao se mapear uma variante, em particular, observamos que há uma tendência de que a frequência da ocorrência dessa variante esteja sendo usada, em um ato comunicativo, em indivíduos que concentrem certas características sociais, tais como o pertencimento ao mesmo gênero/sexo, faixa etária, grau de escolaridade, profissão etc. (SOUSA, 2008, p. 64).

Do ponto de vista prático, no estudo do fenômeno de variação feito pelo sociolinguista, são observados o nível de influência que cada fator social (*sexo, faixa etária, grau de escolaridade*) exerce sobre a escolha de determinada forma em relação à outra] e [a coincidência de uso entre falantes com o mesmo perfil. É por esse viés de combinação/busca de semelhança entre a variação na estrutura linguística e na estrutura social que a Sociolinguística investiga a língua.

No processo de pesquisa sociolinguística, o perfil social do falante é controlado a partir da definição do *envelope de variação*. O envelope de variação constitui o detalhamento de uma *variável*, denominada *dependente*, e suas respectivas *variantes*, isto é, as formas que concorrem à expressão da variável e os contextos linguísticos em que podem ou não ocorrer cada uma dessas variantes. A *variável dependente* é o ponto de partida da análise do fenômeno em variação, ela é o “espaço” em que ocorre a variação em si, espaço esse que pode ser ocupado por cada uma das variantes. Os fatores que podem influenciar essa ocupação, ou seja, a escolha do falante pelo uso de qualquer uma das variantes disponíveis, são monitorados por meio das *variáveis independentes*. A variável independente pode ser *linguística* (tempo e modo verbal,

posição do sujeito, natureza do verbo, paralelismo sintático e/ou semântico etc) ou extralinguística/social (*sexo, faixa etária, escolaridade* etc). Desse modo, na Sociolinguística, há a compreensão de que exista uma combinação entre os fatores linguísticos e sociais que motivam a ocorrência da variação. A escolha das variáveis independentes, em geral, depende da natureza da discussão, do fenômeno em estudo e dos interesses e hipóteses do pesquisador.

Observamos, até aqui, que, de modo geral, a Sociolinguística investiga a correlação existente entre variação linguística e variação social. Para tanto, o pesquisador parte da análise da fala distensa: o *vernáculo*, que é alcançado no momento em que o falante se preocupa mais no que diz e menos em como diz. Ou seja, no momento em que ele execute o mínimo de julgamentos acerca das variáveis produzidas.

Segundo Labov (2008), esses julgamentos de um falante acerca da língua podem ser conscientes ou inconscientes e podem ser de três tipos: os *estereótipos*, os *marcadores* e os *indicadores*. Os *estereótipos* são julgamentos conscientes relacionados a traços socialmente marcados, podendo ser socialmente estigmatizados, o que conduz para a rapidez da mudança linguística e, conseqüentemente, para a extinção da forma que sofre estigma. Os *marcadores* são julgamentos que consistem em uma avaliação subjetiva do falante acerca do uso de determinadas formas em variação, embora o uso dessas formas, que podem ser julgadas como “feias” ou “erradas”, seja feito de modo inconsciente por esse falante. Os *indicadores*, por sua vez, referem-se a um julgamento inconsciente acerca de elementos cuja variação está mais associada à estratificação social do que à variação estilística. Desse modo, a avaliação do falante não é forte por se tratar de variações “comuns”. Em suma, no processo da pesquisa sociolinguística, os julgamentos inconscientes e conscientes de um falante acerca da língua também são sistematizados e contribuem para o entendimento da variação e mudança.

Nessa subseção, tratamos da Sociolinguística enquanto corrente teórica que se ocupa, entre outras coisas, do estudo da variação e mudança linguísticas. Na subseção seguinte, discorreremos sobre o Funcionalismo norte-americano que também lança luz sobre o fenômeno da mudança.

2.2 O estudo da língua no Funcionalismo: a linguística centrada no uso

Nesta subseção, conforme anunciamos, faremos um apanhado geral acerca da segunda corrente teórica que alicerça a análise feita sobre o fenômeno de indeterminação do sujeito: o Funcionalismo norte-americano. A discussão se mantém no mesmo viés da subseção anterior e nos preocuparemos em discorrer sobre o modo como o Funcionalismo olha para a língua(gem)

e para os fenômenos linguísticos, sobre qual noção de gramática se alicerça e quais os princípios que regem a mudança linguística segundo essa corrente.

2.2.1 A origem do Funcionalismo linguístico

A corrente Funcionalista, embora seja divergente das formalistas em termos teóricos e metodológicos, tem suas raízes no Estruturalismo. Segundo Lyons (1981, p. 207, grifo do autor), “[...] o **funcionalismo** é mais corretamente visto como um movimento particular dentro do estruturalismo.”. Na verdade, o seu surgimento se configura como uma reação aos postulados difundidos pelos estruturalistas e que predominavam na linguística na segunda década do século XX. A origem dessa reação, e de onde surgiram os primeiros funcionalistas, segundo Mattos e Silva (2008), foi o Círculo Linguístico de Praga que foi fundado em 1926, quando se tornou particularmente influente na linguística europeia em período precedente à segunda guerra mundial, conforme também postula Lyons (1981).

O interesse pelo estudo da mudança linguística, que, até então, não era ponto central na perspectiva estrutural, foi um dos indicativos de rompimento com o formalismo. A origem desse interesse se deu, segundo Sousa (2008), já em 1920, quando, mesmo estando o Funcionalismo sombreado pelos estudos das funções da linguagem e pelo próprio Estruturalismo, os funcionalistas de Praga já demonstravam, no estudo acerca das funções da linguagem, uma visão de língua diferente da que predominava até então, anunciando a existência de, por exemplo, códigos e subcódigos. Nessa direção, Matos e Silva (2008) afirma que, no artigo de 1960, no qual a teoria das funções da linguagem foi reformulada e publicada, Jakobson, mesmo estando na condição de estudioso/pesquisador estruturalista, “[...] vai além do estruturalismo, uma vez que leva em conta a interação dos que participam de um ato de fala, ou seja, considera a língua no seu uso, no seu contexto comunicativo e não apenas a estrutura.” (MATTOS E SILVA, 2008, p. 71).

Os funcionalistas que surgiram do Círculo Linguístico de Praga e os seus sucessores se dividem em diferentes vertentes dentro da própria corrente. Mas, de modo geral, eles se unem por estarem em volta de um mesmo propósito teórico “[verificar] a(s) **línguas(s) em uso** no seu **contexto social** e **sociointeracional**, ou seja, no **processo comunicativo** entre **falantes**.” (MATTOS E SILVA, 2008, p. 72, grifos nossos). Embora todos os modelos partam desse mesmo objeto e propósito, a vertente funcionalista que nos dá base na presente análise é a norte-americana. Nas próximas subseções, discorreremos sobre alguns dos conceitos e princípios que norteiam essa vertente.

2.2.2 O Funcionalismo norte-americano: de onde falamos?

O Funcionalismo pode ser definido como “[...] uma corrente linguística que [...] se preocupa em estudar a relação entre a estrutura gramatical das línguas e os diferentes contextos comunicativos em que elas são usadas.” (FURTADO DA CUNHA, 2011, p. 157). Assim, segundo a autora,

[...] a abordagem funcionalista apresenta não apenas propostas teóricas distintas acerca da natureza geral da linguagem, mas diferentes concepções no que diz respeito aos objetivos da análise linguística, aos métodos nela utilizados e ao tipo dos dados utilizados como evidência empírica. (FURTADO DA CUNHA, 2011, p. 157).

Do mesmo modo como ocorreu com a Sociolinguística, a corrente Funcionalista “redefine” o modo de olhar para a língua e para os fenômenos linguísticos em geral. As palavras de Mattos e Silva (2008), transcritas no fim subseção anterior, definem, em resumo, o olhar que o Funcionalismo norte-americano lança sobre a língua, uma vez que: (i) o seu objeto de estudo é a língua em uso; (ii) o contexto social e sociointeracional são tidos como determinantes, pois as pressões que eles exercem sobre o uso é fator considerado motivador nas mudanças linguísticas; (iii) o argumento primordial é o de que é, no processo comunicativo, que os falantes negociam, por exemplo, a função que determinada forma assume (essa noção, de certo modo, está relacionada com a ideia de *gramática emergente* defendida por Hopper (1987), da qual falaremos adiante); e, por fim, (iv) o papel dos falantes na teoria funcional, por priorizar o trabalho com a língua em uso, é evidentemente relevante, uma vez que, no Funcionalismo, a análise linguística se dá com “[...] dados reais de fala ou escrita retirados de contextos efetivos de comunicação, evitando lidar com frases inventadas, dissociadas de sua função no ato da comunicação.” (FURTADO DA CUNHA, 2011, p. 158).

Nesse sentido, a corrente funcionalista está na mesma rota⁴ em que a Sociolinguística se encontra: buscando, também, no extralinguístico, as motivações para a compreensão dos fenômenos linguísticos.

Para evidenciar de que lado essa teoria olha a linguagem, registramos que a primeira das divergências existentes entre o Funcionalismo e as outras teorias reside no modo de ver a linguagem. Segundo Furtado da Cunha, Costa e Cezario (2015)

⁴ Embora investiguem a língua de uma perspectiva semelhante, Funcionalismo e Sociolinguística possuem, também, distinções, conforme postula Tavares (2003). Trataremos dos aspectos relacionados a essa questão na subseção seguinte.

O FUNCIONALISMO LINGUÍSTICO contemporâneo difere das abordagens formalistas - estruturalismo e gerativismo – primeiro **por conceber a linguagem como instrumento de interação social** e segundo porque seu interesse de investigação linguística vai além da estrutura gramatical [...] (FURTADO DA CUNHA; COSTA; CEZARIO, 2015, P.21, grifo nosso).

Nessa direção, o Funcionalismo define e caracteriza a linguagem como atividade, como uso. A análise da linguagem e dos seus fenômenos se dá sempre pelo ponto de vista da interação. Ainda segundo Furtado da Cunha (2011), “Os funcionalistas norte-americanos advogam que uma dada estrutura da língua não pode ser proveitosamente estudada, descrita ou explicada sem referência à sua função comunicativa [...]” (FURTADO DA CUNHA, 2011, p. 163). Essa posição defendida pelo Funcionalismo é, de modo geral, um dos pontos centrais da corrente.

Um outro aspecto da linguística funcional caro à nossa discussão é o princípio de iconicidade. Segundo Furtado da Cunha, Costa e Cezario (2015), a iconicidade “[...]é definida como a correlação natural entre forma e função, entre código linguístico (expressão) e seu *designatum* (conteúdo)” (FURTADO DA CUNHA, COSTA E CEZARIO 2015, p. 21). Assim, é consenso entre os funcionalistas a defesa de que há a existência de um reflexo entre a estrutura da língua e a estrutura da experiência. A língua, na visão desses teóricos, é, portanto, motivada ao passo que o signo linguístico pode ser (ou não) arbitrário⁵.

Mesmo compartilhando uma visão semelhante acerca da língua, há, entre os funcionalistas, posicionamentos diferentes em relação à iconicidade. Alguns teóricos assumem uma posição mais radical e outros assumem, por seu turno, uma posição mais branda no que se refere à relação entre forma e função. O primeiro caso, segundo Furtado da Cunha, Costa e Cezario (2015), pode ser exemplificado no *isomorfismo* radical de Bolinger que diz ser condição natural da língua preservar uma forma para um sentido e um sentido para uma forma, posição que, segundo Neves (1997, p. 105), não pode ser tomada como geral, uma vez que tanto código quanto mensagem estão sujeitos a modificações diante do uso. Diferentemente de Bolinger, há teóricos que assumem posição moderada e defendem a existência de uma arbitrariedade relativa o que, segundo Martelotta e Wilson (2008), caracteriza os casos de *motivação*. A motivação, segundo os autores pode ser definida como

[...] a relação de necessidade estabelecida entre uma palavra e seu sentido ou, aproveitando a própria estrutura do termo, como um fenômeno característico de

⁵ A discussão acerca da arbitrariedade do signo linguístico, embora não se desse por meio dessa terminologia, existe desde os filósofos gregos, segundo Martelotta e Wilson (2008). Essa discussão apenas foi retomada pela linguística, com Saussure, e se estendeu às correntes que surgiram após o Estruturalismo.

determinadas palavras que refletem um *motivo* para assumirem uma forma em vez de outra. (MARTELOTTA; WILSON, 2008, p. 75, grifo do autor).

Em suma, o princípio de iconicidade que norteia os estudos funcionalistas e que nos dá respaldo para discutir a indeterminação do sujeito está presente na definição de Martelotta e Wilson (2008). Defendemos que, na língua em uso, a escolha do falante pelas formas *a gente* e *você*, como estratégia para indeterminar o sujeito é, de algum modo, motivada. A título de exemplificação, a escolha do *a gente* pode ter motivação herdada do significado histórico do termo que, antes de passar por processo de gramaticalização, era usado para se referir a um grupo de pessoas, estando associado à noção de coletividade (*as gentes* = *as pessoas*), conforme exemplo registrado no século XVIII⁶: “Quen viu o mundo qual o eu já vi, e viu *as gentes* que eran enton...”. Essa persistência⁷ do traço de coletividade do *a gente*, a nosso ver, é o que favorece, atualmente, o emprego do termo com sentido amplo e referência indefinida/genérica, possibilitando a indeterminação do sujeito.

2.2.3 Gramática e interação no Funcionalismo

Nesta subseção, discorreremos com brevidade sobre gramática e interação numa perspectiva funcionalista, uma vez que não é nosso interesse discutir os pormenores da teoria, mas apresentá-la enquanto ótica para a nossa análise. Trataremos, em linhas gerais, da Gramática Emergente de Hopper e do conceito de interação que adotamos em nossa análise (o Modelo de Interação Verbal de Dik), a fim de definirmos, em termos teóricos, como vamos olhar para o nosso objeto. Para garantir uma análise uniforme e, ao mesmo tempo, sucinta, apresentaremos os postulados de Hopper (1987) sobre a noção de gramática emergente seguidos das abordagens realizadas por Furtado da Cunha, Costa e Cezario (2015) além de Tavares (2013). Feito isso, discorreremos com base em Lima-Hernandes (2011) sobre o modelo de interação verbal de Dik.

2.2.3.1 A gramática emergente de Hopper

De modo geral, a gramática, no sentido em que é usada por Hopper (1987, p. 142), “[...] emerge continuamente no discurso, não havendo, portanto, gramática, mas gramaticalização,

⁶ Excerto extraído de Lopes (2004)

⁷ Faremos uma discussão mais ampla sobre o princípio da persistência de Hopper ainda nessa seção.

um movimento em direção à estrutura, que nunca se completa totalmente.”. Portanto, na concepção de Hopper (1987), a gramática está sempre incompleta, em processo constante de construção. Ela pode ser entendida como “[...] um conjunto de regularidades decorrentes de pressões cognitivas e, sobretudo, de pressões de uso.” (FURTADO DA CUNHA, COSTA E CEZARIO, 2015, p. 41 - 42). Ou ainda, na releitura de Tavares (2003, p. 15), ela é “uma atividade em tempo real, *on-line*, que emerge cotidianamente no discurso⁸”.

Das definições que apresentamos acima, duas noções são basilares na análise que desenvolvemos acerca da indeterminação do sujeito: a gramática como sendo (i) **decorrente do uso** e como sendo (ii) **atividade em tempo real e incompleta**. A primeira noção é fundamental por defendermos que as formas gramaticalizadas *a gente* e *você* atuam como estratégias indeterminadoras justamente pelo “desgaste semântico” motivado pela alta frequência de uso à qual foram submetidas nos seus processos de mudança⁹. A segunda, por defendermos que tanto o *a gente* quanto o *você* não possuem características intrínsecas de indeterminação (Cf. Castilho, 2010) em termos formais, mas adquirem esses traços quando os falantes estão em processo de interação e inseridos em um contexto discursivo que homologue, no sistema linguístico, o uso dessas formas para essa função. Em suma, é quando a língua acontece que se torna possível negociar a referência dessas duas formas entre os falantes envolvidos no ato de comunicação. É especificamente essa dinâmica de alterações sofridas pelas formas *a gente* e *você*, que surgem devido à necessidade de uma maior expressividade linguística, que Hopper (1987) denomina a gramática como Gramática Emergente.

Por esse ângulo “A gramática não é estável, fechada e auto-contida, mas sim aberta, fortemente suscetível à mudança e intensamente afetada pelo uso que lhe é dado no dia-a-dia.” (TAVARES, 2003, p. 16). A gramática é, nessa perspectiva, maleável e suas regras são passíveis de modificações ocasionadas pelo uso. *Você* e *a gente*, por exemplo, podem ser usados com referências distintas, embora tradicionalmente designem apenas a segunda pessoa do singular e a primeira pessoa do plural, respectivamente. Essa possibilidade de negociação da referência, de certo modo, está associada à frequência de uso e a pressões e exigências do próprio contexto. Sobre esse aspecto, Tavares, (2003) afirma sobre a gramática emergente que

As situações comunicativas se desenrolam em meio a pressões facilitadoras da inovação, quais sejam as exigências pragmáticas e sociolinguísticas *on-line*. Essas

⁸ Na perspectiva funcionalista, o termo **discurso** se refere “às estratégias criativas utilizadas pelo falante para organizar funcionalmente seu texto para determinado ouvinte em determinada situação comunicativa” (FURTADO DA CUNHA, COSTA E CEZARIO, 2015, p. 42). Em termos práticos, se refere à língua em uso.

⁹ Falaremos mais pontualmente sobre esse processo na próxima subseção quando tratarmos dos princípios subjacentes à mudança via gramaticalização.

exigências levam ao ajuste de diferentes experiências, negociando-se fórmulas lingüísticas à medida que a troca comunicativa avança. (TAVARES, 2003, p. 17).

No caso específico da indeterminação do sujeito, uma hipótese possível é a de que o falante use *você* e *a gente* como estratégia para indeterminar o sujeito com o objetivo de se isentar de uma opinião, no contexto de uma argumentação com o ouvinte, assunto que será discutido posteriormente na análise dos dados.

Em suma, as ideias que giram em torno da gramática emergente de Hopper e que compartilhamos na nossa análise são (i) a gramática como sendo maleável e não estática; (ii) a gramática como sendo uma atividade em tempo real; (iii) a ideia de que as alterações são negociadas no contexto e licenciadas pelo sistema linguístico; e (iv) a premissa de que essas alterações são decorrentes do uso e cedem a pressões determinadas e diversas na/da interlocução.

2.2.3.2 O modelo de interação verbal de Dik

Estando a noção de gramática a ser adotada na discussão definida, cabe-nos agora discorrer brevemente sobre o que entendemos por interação na análise aqui desenvolvida. Segundo Lima-Hernandes (2011, p. 21), “O pressuposto básico da abordagem funcionalista é que o estudo do sistema linguístico leva em consideração o usuário da língua e, como consequência, todo o entorno suficiente para a produção de sua fala tornou-se alvo de interesse”. Tendo em vista que o objetivo da Gramática Funcional é o de descrever as línguas focalizando aspectos pragmáticos e psicológicos, segundo a autora, discussões acerca do entorno comunicativo são centrais.

É apoiando-se nisso que, para a autora, a proposta de Dik se fundamenta. Ou seja, na ideia de que o usuário da língua é parte do sistema e, sendo assim, a cadeia linguística por si só se torna insuficiente para compor o discurso, necessitando de determinados conhecimentos à parte que são definidos por Dik como (i) *conhecimentos de longo termo linguísticos* (lexicais, gramaticais e pragmáticos) e não linguísticos (domínio de informações do mundo real e dos mundos possíveis). Os (ii) *conhecimentos de curto termo situacionais* (ligados à dêixis do evento comunicativo) e *textuais* (ligados às referências textuais e estados das coisas).

De modo sucinto e específico, interessa-nos, no modelo de interação verbal de Dik, a ideia de que todos esses conhecimentos entram em jogo na comunicação a fim de produzir sentido. Usando as palavras de Lima-Hernandes (2011, p.23), a ideia de que existe um jogo de

“[...] antecipação de interpretação por parte do falante e de reconstrução por parte do ouvinte”. Falantes e ouvintes cooperam a fim de, em um contexto específico, seja possível compartilhar a “mesma informação”.

No modelo de interação verbal de Dik reproduzido por Lima-Hernandes (2011),

[...] todo ato comunicativo incorpora a intenção do falante. E o ouvinte (num papel não passivo de receptor) interpreta a partir do *output* linguístico gerado pelo falante. O jogo comunicativo se estabelece, assim, entre a expressão linguística e a interpretação à luz das informações pragmáticas do ouvinte. (LIMA-HERNANDES, 2011, p. 24).

Aplicando esse modelo de interação a um exemplo do nosso objeto, retirado do *Corpus PPVC*, como em: “Eu nem queria, porque quando A GENTE qué A GENTE pode, né?, querê é podê, né?, então assim eu quero muito, porque eu falo assim eu quero aprendê pra mim, num é nem pra ninguém.” (E.S.P. – PPVC), teríamos, como resultado, a seguinte análise.

A situação é de interação entre dois falantes e o locutor utiliza por duas vezes a forma gramaticalizada *a gente* sendo que as duas ocorrências estão, a nosso ver, com valor genérico. Supomos que o ouvinte tenha consciência de que o *a gente* possui esse valor genérico, porque falante e ouvinte estão imersos em um contexto em que a superestrutura textual predominante na fala do locutor é argumentativa e, logicamente, a sua intenção é de convencer o interlocutor de uma ideia que defende, uma ideia que deseja que seja apresentada como universal e não particular, como seria o caso se o *a gente* fizesse referência apenas a ele próprio. Desse modo, informações pragmáticas e linguísticas são acionadas a ponto de o ouvinte ter noção de que o *a gente*, no exemplo 1, não se refere a *ele+o falante*, mas que tem referência imprecisa/genérica devido ao teor da conversa, havendo, portanto, uma reinterpretação à luz do contexto interacional.¹⁰ É, entre outros, por esse caminho que acreditamos que pode funcionar a indeterminação do sujeito por meio do uso das formas gramaticalizadas *a gente* e *você*, a partir de uma interpretação do modelo de interação verbal de Dik. É “o ir e vir” do ato comunicativo, na troca de informações pragmáticas entre os interlocutores, que há a possibilidade de as duas formas em estudo funcionarem como indeterminadoras.

¹⁰ É válido ressaltar que, na análise realizada acima, consideramos o fato de o *a gente* estar inserido em um ditado popular/provérbio (querer é poder) o que reforça o seu valor genérico.

2.2.4 Princípios da gramaticalização aplicados à indeterminação

A gramaticalização é entendida, dentro do campo da linguística funcional, como um processo em que há, segundo Meillet (1965, p. 131), a “[...] atribuição de um caráter gramatical a um termo ou palavra anteriormente autônomo”. Para Hopper e Traugott (1993) ela é “[...] o processo pelo qual itens e construções gramaticais passam, em determinados contextos linguísticos, a servir a funções gramaticais, e, uma vez gramaticalizados, continuam a desenvolver novas funções gramaticais.” (HOPPER e TRAUGOTT, 1993, p. 15).

De modo semelhante, Lehmann (1982) “[...] define gramaticalização como um processo que transforma lexemas em formativos gramaticais e formativos gramaticais em mais gramaticais ainda” (GONÇALVES E CARVALHO, 2007, p. 70). Uma definição que, de certo modo, amplia o escopo do conceito de gramaticalização é apresentada por Traugott (2014). A autora define gramaticalização como sendo “[...] o desenvolvimento de expressões procedurais, entre elas as expressões de tempo, aspecto, modalidade, caso, pronomes pessoais, complementizadores e outros conectivos.” (TRAUGOTT, 2014, p. 98). Além disso, a autora pondera que, a depender da visão de gramática que se adota, a gramaticalização pode incluir também o desenvolvimento de marcadores pragmáticos, ou seja, a gramaticalização não se refere apenas a mudanças que ocorrem do léxico para a gramática ou dentro da própria gramática, mas de um fenômeno que assume contornos mais amplos a ponto de incluir aspectos pragmáticos e semânticos da língua. Em suma, a gramaticalização, pelo que observamos nas definições apresentadas, constitui um processo de mudança que, no geral, envolve mobilidade entre formas e funções que estão em um mesmo domínio funcional.

Há inúmeras discussões, como as encontradas em Gonçalves e Carvalho (2007) e Neves (1997), acerca da existência de relação entre a mudança via gramaticalização e outros tipos de mudanças linguísticas, assim como em relação ao próprio uso do termo gramaticalização e outros termos paralelos para se referir a um tipo específico de mudança. E, nesse contexto, há os teóricos que defendem ser a gramaticalização um tipo particular de mudança e há, também, os que defendem que a maioria das mudanças envolve gramaticalização.

Sem nos ater às discussões que giram em torno da definição de gramaticalização nem às divergências terminológicas, assumimos, aqui, ser a gramaticalização um tipo específico de mudança. O que nos interessa, nessa análise, é que, nessa mudança, há a atuação de princípios específicos que fazem com que haja o deslocamento de uma forma seja do léxico para a gramática, seja dentro da própria gramática e, nesse processo de deslocamento, novas funções entram em jogo. Dentre elas, acreditamos estar a função de indeterminar o sujeito, em se

tratando especificamente das formas pronominais *a gente* e *você*. De todos os princípios que atuam nessa mudança, daremos destaque aos defendidos por Givon (1984) e, de modo complementar, associaremos a esses os postulados de Lehmann (1982). Além disso, consideraremos também os princípios de Hopper (1993).

Na análise que fazemos acerca do fenômeno de indeterminação do sujeito, os princípios que atuam na gramaticalização nem sempre são aplicáveis de modo efetivo, mas a hipótese que norteia a investigação do nosso objeto é a de que as propriedades indeterminadoras das formas *a gente* e *você* são resultado dos seus históricos de mudança (mais especificamente, devem-se ao resultado da atuação dos princípios aqui mencionados). Nesse sentido, trataremos dos princípios que julgamos ter contribuído, no processo de mudança, para que as duas formas sejam usadas como recurso indeterminador no português falado.

2.2.4.1 Iconicidade e marcação - Givon

Givon (1984) trata o princípio da *iconicidade* de forma branda e a partir de três subprincípios: o da **quantidade**, do **grau de integração** e da **linearidade**¹¹.

- (i) O subprincípio da **quantidade** se refere à correlação e à correspondência entre a quantidade de forma e de informação. Assim sendo, a complexidade do pensamento, certamente, estará refletida na complexidade da expressão.
- (ii) O subprincípio do **grau de integração** prevê que há uma relação de integração entre o cognitivo e o codificado, o que está cognitivamente próximo é colocado sintaticamente próximo.
- (iii) O subprincípio da **linearidade**, por sua vez, está ligado a uma espécie de hierarquia: o conteúdo mais importante ocupa o primeiro lugar na cadeia sintática.

O segundo princípio defendido dentro do quadro teórico funcionalista e que também nos dá respaldo para identificar como a indeterminação do sujeito ocorre no português falado do Sudoeste da Bahia é o de *marcação*. Defendido principalmente por Givon (1984), o princípio de marcação foi herdado da linguística estrutural e estabelece três critérios que diferenciam *formas marcadas* das *não marcadas* em contraste binário.

¹¹ Estamos, nesta análise, considerando a iconicidade em dois sentidos (i) enquanto princípio que rege a língua (ii) enquanto princípio que rege mudanças em particular. Interessa-nos, no estudo da indeterminação do sujeito, a noção mais geral de iconicidade, uma vez que o fenômeno será analisado pelo ponto de vista da referência o que, de certo modo, invoca a discussão da relação entre a forma e o que ela designa e, conseqüentemente, esbarrando na noção de iconicidade. Nesse sentido, os subprincípios que atuam nas mudanças em particular não serão aplicados efetivamente na análise, mas serão explicados.

Segundo Furtado da Cunha, Costa e Cezario (2015), a marcação diz respeito a um conjunto de combinações (*complexidade estrutural*, *complexidade cognitiva* e *distribuição de frequência*) que tende a opor binariamente duas formas que desempenham uma mesma função.

- (i) O critério da *complexidade estrutural* diz que a estrutura marcada tende a ser mais complexa ou maior do que a não marcada.
- (ii) O da *distribuição de frequência* diz ser a estrutura marcada menos frequente do que a estrutura não marcada.
- (iii) Por fim, o da *complexidade cognitiva* diz ser a estrutura marcada cognitivamente mais complexa do que a não marcada.

É a combinação ou não desses fatores que define uma forma como marcada ou não marcada. A coincidência entre a complexidade estrutural, complexidade cognitiva e a baixa frequência, por exemplo, indica que determinada estrutura é marcada em oposição à outra que não combina esses três critérios. Givon (1995) postula que o princípio da marcação é relativo, isto é, há formas que são marcadas em um contexto e não em outro, fazendo com que o princípio da marcação deva ser investigado considerando fatores de ordem sociocultural, biológico, cognitivo etc. Além disso, o princípio da marcação não se restringe apenas a categorias linguísticas e opera em um contínuo dado o caráter fluido e criativo da língua. Desse modo, faz-se necessária uma análise não só binária do princípio, mas gradativa e hierárquica que parta de um nível mais marcado para um nível menos marcado.

Dos três critérios mencionados, que, dentro do princípio da marcação, atuam no processo de gramaticalização, o que dialoga mais com a análise que fazemos da indeterminação do sujeito é o critério da *distribuição de frequência*. Mais especificamente, os efeitos que a distribuição de frequência (o uso) de uma forma lhe causa no decorrer da mudança linguística. Segundo Neves (1997), alguns autores associam mudanças que envolvem perdas¹² fonéticas e descoloramento semântico à gramaticalização. Diante disso, amparamo-nos no fato de que o uso em larga escala das formas em análise favoreceu o seu desbotamento semântico e, consequentemente, o seu emprego com valor genérico

Embora não adotemos o conceito de perda, percebemos que, no decorrer do processo de gramaticalização, as formas *você* e *a gente* deixaram de ter referências específicas devido à alta frequência de emprego. Em consequência, o *você*, que em termos paradigmáticos, refere-se à segunda pessoa do singular passou a fazer, também, referência genérica. Essa alteração

¹² Para Neves (1997), as noções de *perda*, *desbotamento*, *enfraquecimento* ou *descoloramento* são contrariadas por alguns autores, entre eles (Sweester, Traugott e König).

referencial é o que permite que a forma funcione como recurso de indeterminação do sujeito na língua em uso. O mesmo processo ocorre com a forma *a gente*.

Assim como Givon (1984), Lehmann (*apud* Gonçalves e Carvalho, 2007), na investigação da gramaticalização de itens lexicais, defende a existência da atuação de determinados parâmetros tanto no nível *sintagmático* quanto no *paradigmático* que permitem captar a mudança em seu estágio inicial.

Segundo Gonçalves e Carvalho (2007), o ponto de partida para estabelecer os parâmetros de Lehman é a combinação de seis parâmetros formais nos eixos sintagmáticos e paradigmáticos que medem o grau de autonomia de um item e, consequentemente, medem o seu grau de gramaticalidade, uma vez que, segundo os autores, “[...] a autonomia de um signo é postulada como contrária ao seu estatuto de ‘gramatical’.” (GONÇALVES E CARVALHO, 2007, p. 70, grifo dos autores). É a partir da combinação dos aspectos *peso*, *coesão* e *variabilidade* com os dois eixos *sintagmático* e *paradigmático* que se obtém cada parâmetro e é, a partir da combinação de um parâmetro do nível sintagmático com outro do nível paradigmático, que se mede o grau de autonomia do item. Em termos práticos, para que um item seja considerado autônomo, o aspecto *peso* deve ser combinado no sintagma e no paradigma em que participa o que vai ser resultado da combinação de dois parâmetros: *peso* no sintagma = **escopo** e no paradigma = **integridade**, logo, confrontou/combinou-se **integridade** e **escopo** e o mesmo ocorre com cada um dos outros parâmetros. Lehmann (1982) defende que, embora sejam dependentes um do outro para que seja possível medir o grau de autonomia, cada parâmetro pode ser avaliado isoladamente no sintagma e no paradigma, mas defende, também, que exista correlação significativa entre os parâmetros. Todos os seis parâmetros, segundo o autor, levam em consideração o item em estágio avançado sendo dificultada a identificação na gramaticalização em curso. Os parâmetros são os que constam no quadro:

Quadro 1 - Parâmetros de Lehmann

I - Parâmetros paradigmáticos

Analisa-se nesse eixo a integração (quantidade) dos traços semânticos, seu grau de participação no campo semântico e da sua possibilidade de escolha em relação a outros itens.

1- *Peso vs. Paradigmatização: integridade* - Esse parâmetro se refere, grosso modo, ao tamanho (fonológico) do item e à quantidade de traços semânticos que agrega. No estágio inicial, ele pode ser polissílabo e agregar muitos traços semânticos, mas pelo processo de *atrição fonológica* ou *dessemantização* ele pode chegar a ser monossilábico e portar poucos traços semânticos. Segundo os autores, esse resultado pode ser ocasionado pela repetição/frequência do uso do item no discurso que pode gerar automatização, o que não ocorre com itens menos frequentes. A persistência leva o item a desgastar seu valor semântico original e, conseqüentemente, a desempenhar novas funções

2- *Coesão vs paradigmatização: paradigmaticidade* - Diz respeito à relação que o item mantém com outros dentro de um paradigma. O que está em jogo, então, é o tamanho do paradigma em si e a relação com os outros itens..

3- *Variabilidade paradigmática* - Diz respeito à possibilidade de uso de outro item no lugar do que está em processo de gramaticalização, sendo que pertencem a um mesmo paradigma.

II - Parâmetros sintagmáticos

Diz respeito à relação entre itens da sentença, suas colocações e mobilidade.

1- *Peso vs. Sintagmatização: escopo* - esse aspecto, grosso modo, está ligado à capacidade de projeção que o item tem em relação à sentença, sendo que itens não gramaticalizados têm capacidade de projetar sentenças mais complexas ao passo que itens gramaticalizados se articulam em menor nível.

2- *Coesão vs sintagmatização: conexidade* - esse aspecto trata da relação de dependência ou não entre os itens do sintagma. O item gramaticalizado tende a ser mais contraído.

3- *Variabilidade sintagmática* – esse aspecto refere-se à possibilidade de um item assumir uma posição fixa (o que aponta para um processo mais avançado de gramaticalização) ou não na sentença.

Fonte: Elaborado pelo próprio pesquisador com base nos postulados de Gonçalves e Carvalho (2007) e Lehmann (1982).

Dos parâmetros defendidos por Lehman (1982), interessa-nos, em particular, o de *integridade*, mais especificamente, a **dessemantização** que, segundo Gonçalves e Carvalho (2007), é causada pela alta frequência de uso que gera **automatização** e, conseqüentemente, leva ao **desgaste do valor original** do item, pois acreditamos que é o que ocorre com o *você* e *a gente* quando são usados com referência genérica.

Em suma, é conciliando os postulados de Lehman (1982), no que se refere ao parâmetro da *integridade*, e os de Givon (1984), no que se refere ao subprincípio da *distribuição de*

frequência, que associaremos a mudança pela qual as formas *a gente* e *você* passaram aos seus usos como estratégia para indeterminar o sujeito.

2.2.4.2 Os princípios de Hopper

Hopper (1993) defende a existência de cinco princípios que atuam na gramaticalização. Esses cinco princípios dão conta do processo de forma gradual, considerando não o pertencimento de um item à gramática, mas o seu grau mais ou menos gramatical, conforme esclarecem Gonçalves e Carvalho (2007). Os princípios propostos por Hopper (1993) são os que constam no quadro 2:

Quadro 2 – Os princípios de Hopper

- 1- *Estratificação* - O princípio da estratificação está ligado à coexistência de uma forma gramaticalizada nova e uma forma-fonte antiga em um mesmo domínio funcional sem que, necessariamente, haja uma substituição imediata ou, segundo os autores, essa substituição pode até mesmo não vir a acontecer. Trata-se, portanto, da convivência de formas novas e antigas com valor aproximado, atuando em um mesmo contexto, como é o caso das formas *nós* e *a gente* que atuam em conjunto fazendo, entre outras, referência à terceira pessoa do plural.
- 2- *Divergência* - Está ligada à autonomia de um item que dá origem ao processo de gramaticalização. Esse item, segundo Hopper, continua sujeito a outras mudanças inclusive a um novo processo de gramaticalização, o que ocorreu com a forma *gente* que, passou, adjungida ao determinante *a*, da categoria de substantivo para a de pronome, mas permaneceu na língua, estando disponível para outros processos de mudança.
- 3- *Especialização* - Diz respeito ao estreitamento da escolha de uma forma em detrimento da outra que pertence ao mesmo domínio funcional. Isto é, uma forma passa a ser preferencial em termos de escolha pelo falante em relação a outra. Um exemplo de especialização pode ser encontrado no uso da forma *você* que, em contextos específicos, é usado posposto ao sintagma *sabe quando* com referência genérica. Ou seja, a tendência é que o falante use o *você* na sentença “*sabe quando você...*” sempre com referência genérica.
- 4- *Persistência* - Esse princípio “[...] é o que prevê a manutenção de alguns traços semânticos da forma-fonte na forma gramaticalizadas o que pode ocasionar restrições sintáticas para o uso da forma gramaticalizadas” (GONÇALVES E CARVALHO, 2007, p. 83). O traço de coletividade do substantivo *gente* foi preservado na forma gramaticalizada *a gente*, por exemplo, fazendo com que a forma pronominal seja usada com sentido indeterminado por se

referir a um grupo grande de pessoas. A restrição sofrida pela forma inovadora se deve ao fato de ela não aceitar ser modificada por quantificadores como numerais.

- 5- *Descategorização* - Grosso modo, a descategorização está ligada à perda de autonomia discursiva do item em gramaticalização. Os nomes, por exemplos, perdem a propriedade de identificar participantes do discurso e os verbos de reportar eventos, passando a assumir papéis de afixos, desinências, clíticos etc. Nos processos de gramaticalização das formas *a gente* e *você*, é perceptível a atuação do princípio de descategorização. Com o *a gente*, por exemplo, houve a passagem de substantivo para pronome com a perda dos traços de diminutivos e plural. Com o *você*, houve a passagem da forma nominal composta pelo pronome pessoal *vossa* + substantivo *mercê* para a categoria de pronome. Em ambos os casos, houve a descategorização morfossintática e o que observamos agora é que duas formas continuam suas trajetórias de mudança sob a atuação de um outro tipo de descategorização, a semântica, uma vez que as suas referências estão sendo ampliadas de específica para genérica.

Fonte: Elaborado pelo pesquisador com base nos postulados de Hopper (1993) e de Gonçalves e Carvalho (2007)

Dos princípios defendidos por Hopper (1993), os de *estratificação e persistência* são os que mais se revelam na análise aqui desenvolvida. No primeiro caso, o da *estratificação*, por defendermos que, mesmo não operando da mesma forma, diversas estratégias de indeterminação do sujeito coocorrem no português brasileiro, tanto as defendidas pela Tradição Gramatical quanto as que são apresentadas na linguística descritiva, conforme apresentaremos na seção 2. No segundo caso, o da *persistência*, por defendermos que as formas *a gente* e *você* são empregadas como estratégias para indeterminar o sujeito porque, de certo modo, alguns traços pertencentes às formas-fonte foram mantidos (o traço da pessoa, o traço de coletividade, em se tratando do *a gente*, conforme supracitado).

2.3 Interesses afins = um “casamento”, o Sociofuncionalismo

Nesta subseção, trataremos da junção dos pressupostos teóricos e metodológicos das duas teorias sobre as quais discorreremos nas subseções anteriores, Sociolinguística e Funcionalismo norte-americano, que deu origem à orientação de pesquisa tomada como base para a análise que desenvolvemos acerca do fenômeno de indeterminação do sujeito: o Sociofuncionalismo. A escolha pela orientação de pesquisa sociofuncionalista se deve, conforme já mencionamos, ao fato de termos como objetivo observar a influência que o perfil social do falante exerce na escolha de uma ou outra estratégia de indeterminação do sujeito (*a*

gente e você) e, em paralelo, observar em que medida a atuação de princípios ligados à gramaticalização tornou possível que as duas formas atuassem como estratégias para indeterminar o sujeito. Essa dupla visão torna-se possível a partir do Sociofuncionalismo que une pressupostos da Sociolinguística e do Funcionalismo norte-americano.

Segundo Tavares (2003, 2013), o Sociofuncionalismo é uma orientação de pesquisa que resulta do “casamento” teórico de duas correntes linguísticas: a Sociolinguística Variacionista e o Funcionalismo norte-americano, principalmente a partir dos postulados de Labov, no que se refere à primeira e de Hopper, Heine e Givon, no que se refere à segunda. O termo Sociofuncionalismo foi usado pela primeira vez, segundo Neves (1999), no maior e mais antigo grupo que desenvolve pesquisa na linha funcionalista: o PEUL-RJ (Programa de Estudos sobre o Uso da Língua) que, conforme a autora, tem uma orientação variacionista dominante e trata do fenômeno de variação linguística sob a ótica da função discursiva das variantes. Desde então, estudos nessa perspectiva vêm se popularizando no Brasil lançando uma luz dupla sobre os fenômenos de variação e mudança da língua. Assim, aconteceu, posteriormente, na UFSC com a Prof.^a Dr.^a Edair Gorski; na UFRN, com as Prof.^{as} Dr.^{as} Maria Angélica Furtado da Cunha e Maria Alice Tavares; na UESB com os Profs. Drs. Valéria Viana Sousa e Jorge Augusto Alves da Silva.

A proposta de engate entre as duas correntes que constituem essa nova orientação tem como base os pontos de vista e interesses comuns que compartilham. Mas, para Tavares (2003), a tarefa de conciliar pressupostos de duas teorias até então distintas não é simples, uma vez que

Não basta associar preceitos de diferentes fontes teóricas por parecerem idênticos, semelhantes ou complementares, sem que se verifique se o que parece comum de fato o é e sem que se decida como lidar com as diferenças quando estas surgirem. Além disso, é importante que se estabeleça uma discussão mais refinada acerca do que fundamenta a junção e de quão profundamente esta se dá. Uma teoria é um pacote completo e talvez nem todos os seus pressupostos sejam encaixáveis em outra teoria, o que traz implicações para o casamento teórico que devem ser visitadas. (TAVARES, 2003, p. 98).

Embora estejam em uma mesma rota no campo dos estudos linguísticos, necessário se faz rever os limites de combinação das abordagens das duas teorias (Sociolinguística e Funcionalismo), assim como as suas divergências, a fim de garantir consistência e solidez ao “casamento”. Para enxugarmos a discussão e apresentarmos o nosso interesse acerca da orientação de pesquisa sociofuncionalista, trataremos, especificamente, dos aspectos presentes na ressalva de Tavares (2003, 2013) sobre a junção das teorias: **o que possibilita a união**. Para

tanto, traremos para o nosso texto o recorte feito pela autora. Esse recorte está adaptado no quadro 3. Necessário se faz ressaltar que os conceitos gerais relacionados ao Funcionalismo e Sociolinguística dos quais a autora trata antes de entrar no mérito da união das duas teorias já foram, de certo modo, fazendo as focalizações de nosso interesse nesse estudo, discutidos nas subseções anteriores não havendo necessidade de serem retomados.

Quadro 3 - Sociolinguística e Funcionalismo: o que une as teorias¹³

- Prioridade atribuída à língua em uso, cuja natureza heterogênea abriga a variação e a mudança (cf. WEINREICH; LABOV; HERZOG, 2006 [1968]; GIVÓN, 1995; entre outros);
- A língua não é estática. Ao contrário, sofre alterações constantes (cf. GUY, 1995; GIVÓN, 1995, 2001; HOPPER, 1987);
- O fenômeno da mudança linguística recebe um lugar de destaque, e é entendido como um processo contínuo e gradual (cf. WEINREICH; LABOV; HERZOG, 2006 [1968]; GIVÓN, 1995, 2001; HOPPER; TRAUGOTT, 2003);
- Dados sincrônicos e diacrônicos são tomados complementarmente com o intuito de obtenção de prognósticos de mudança mais refinados e confiáveis (cf. LABOV, 1994; HEINE; CLAUDI; HUNNEMEYER, 1991);
- Crença no princípio do uniformitarismo, segundo o qual as forças linguísticas e sociais que agem hoje sobre a variação e a mudança são em princípio as mesmas que atuaram em épocas passadas (cf. LABOV, 2008 [1972a]; HOPPER; TRAUGOTT, 2003);
- A frequência das ocorrências recebe destaque. Na perspectiva funcionalista, a frequência é fundamental para o estabelecimento e a manutenção da gramática, e, além disso, a difusão linguística e social da mudança pode ser captada através do aumento da frequência de uso em diferentes contextos (BYBEE, 2010). Na perspectiva variacionista, o aumento de frequência é compreendido como índice de difusão sociolinguística (LABOV, 2008 [1972a]), além do que as variantes devem ter certa recorrência para que possam ser comparadas por meio de instrumental estatístico.
- Há relação entre os fenômenos linguísticos e a sociedade que usa a língua. A mudança se espalha de forma gradual ao longo do espectro social, considerando-se fatores como região, geração, classe social etc. É comum haver diferença entre falantes mais velhos e mais jovens, no caso de mudança em progresso. (cf. LABOV, 2008 [1972a]; LICHTENBERK, 1991);
- Fatores de natureza interacional têm papel importante na variação e na mudança linguística. No âmbito da sociolinguística, Labov (2008 [1972a]) compreende a variação estilística como uma adaptação da linguagem do falante ao contexto imediato do ato de fala. E no âmbito do funcionalismo, Traugott (2002) não só defende que a mudança é motivada por práticas discursivas e sociais, como acredita que os estudos funcionalistas de gramaticalização orientados para o falante podem contribuir para o estudo sociolinguístico da variação intrafalante.
- A gramaticalização, processo de mudança responsável pela migração de formas linguísticas para a gramática, vem recebendo grande destaque nos estudos funcionalistas (cf. HARDER, BOYE, 2011). Para o funcionalismo, vários casos de variação e mudança morfossintática podem ser explicados através desse processo (BYBEE, 2010; NEVALAINEN; PALANDER-COLLIN, 2011). No âmbito da sociolinguística, Labov (2010) também aponta a gramaticalização como uma possível fonte de explicação para a mudança morfossintática, e cita autores funcionalistas que vêm trabalhando nessa perspectiva, como Heine e Kuteva (2005), Hopper e Traugott (2003) e Haspelmath (2004)

¹³ Quadro de autoria própria, elaborado a partir dos postulados de Tavares (2003, 2013). À discussão desenvolvida pela autora, acrescentamos a noção de comunidade de fala que nos norteia e que também é comum entre Sociolinguística e Funcionalismo.

□ A comunidade de fala é o lugar no qual tanto o sociolinguista quanto o funcionalista, e o sociofuncionalista, encontram o seu objeto de estudo: a “língua falada, observada, descrita e analisada em seu contexto social, isto é, em situações reais de uso.” (ALKMIM, 2001, p. 31). Ou, nas palavras de Cezario e Votré (2011), no vernáculo, na fala natural “em que supostamente, o falante se preocupa mais com o *que* dizer do que com o *como* dizer.” (CEZARIO E VOTRE, 2011, p. 149, grifos dos autores). Chegar a esse objeto, a um estado de comunicação em que o falante se preocupe mais com o que dizer é tarefa complexa e cercada de paradoxos, conforme veremos na seção 4, que trata da metodologia da pesquisa em Sociolinguística de modo mais específico.

Fonte: Elaborado pelo pesquisador com base em Tavares (2003, 2013)

De modo geral, todos os postulados teóricos e metodológicos discutidos por Tavares (2003, 2013) que estão adaptados no quadro 3, e aos quais acrescentamos a noção de comunidade de fala, dão-nos respaldo para analisar o fenômeno de indeterminação do sujeito no português falado do Sudoeste da Bahia. De todos esses postulados, alguns, de modo mais específico, estão mais alinhados com os princípios que consideramos para a nossa análise e dialogam mais com as hipóteses das quais partimos para analisar os dados.

Em primeira mão, e esse é o pressuposto, talvez, mais relevante defendido pelas duas teorias, interessa-nos o fato de ambas trabalharem com a *língua em uso* e com a premissa de que ela *não é estática*, uma vez que assumimos a posição de que a indeterminação ocorre com as formas gramaticalizadas *a gente* e *você* justamente na língua em uso, na interação. Sem essa visão, uma análise estritamente formal não daria conta dos objetivos e hipóteses que aqui colocamos.

Outro postulado destacado por Tavares (2013) e que também nos é relevante está relacionado com a importância que a *frequência de uso* tem nas duas teorias. Ao analisarmos a indeterminação do sujeito por meio de *a gente* e *você*, acreditamos que uma possibilidade para a efetivação do fenômeno por meio dessas formas é resultado de um *desbotamento semântico* ocasionado pela frequência de uso. Nos respaldamos, para tanto, em princípios defendidos por Givon (1984), Lehmann (1982), sobre os quais já tratamos nessa seção.

Outro aspecto que está intimamente relacionado com a discussão que fazemos diz respeito à atuação de *fatores de ordem interacional* na mudança. Trouxemos esse postulado para a nossa análise quando adotamos o Modelo de Interação Verbal proposto por Dik, que prevê, na interação, a movimentação de diversos conhecimentos em um *continuum* que vai do linguístico ao pragmático e que garantem o compartilhamento de informações entre os interlocutores envolvidos em um ato comunicativo.

Por fim, e esse é o nosso grande “guarda-chuva”, é o interesse comum pela mudança linguística que as duas correntes compartilham que alicerça a nossa análise. O nosso foco é o de perceber como a mudança linguística envolvendo as duas formas em análise, *a gente* e *você*,

pode ser vista sob a luz das duas vertentes teóricas em conjunto. No que se refere à Sociolinguística, verificaremos como a pressão de fatores sociais influenciam o uso de *a gente* e *você* como estratégias de indeterminação do sujeito em variação, bem como os condicionantes linguísticos que favorecem essa nova função para esses termos já existentes na língua. Já em relação ao Funcionalismo, buscaremos um evidenciar o ponto de intercessão entre gramaticalização e indeterminação, por acreditarmos que a atuação de determinados princípios típicos da mudança via gramaticalização favoreceu, no percurso de mudança das duas formas, os seus usos como estratégias indeterminadoras. Estabelecendo essa relação variação-gramaticalização-indeterminação e lançando mão de um arcabouço teórico e metodológico também comum à Sociolinguística e ao Funcionalismo, acreditamos ser possível “cercar” o fenômeno dentro dos limites da relação teórica, o Sociofuncionalismo, dando um tratamento mais amplo à indeterminação do sujeito.

3 INDETERMINAÇÃO DO SUJEITO: DOS CONCEITOS ÀS CLASSIFICAÇÕES, DOS DICIONÁRIOS ÀS GRAMÁTICAS

Esta seção, na qual temos o propósito de discutir a gênese das controvérsias que envolvem o estudo da indeterminação do sujeito, a definição e a classificação, está subdividida em três subseções. Na primeira, *Indeterminação e conceitos afins nos dicionários*, apresentaremos e discutiremos os conceitos de *indeterminação*, *determinação*, *definido/definição* e *indefinido/indefinição* em alguns dicionários da Língua Portuguesa, a saber Bluteau (1789), Dubois (1973), Crystal (1941) e Camara Junior (1997). Na segunda, *A indeterminação do sujeito nas gramáticas*, trataremos da abordagem do fenômeno nas gramáticas históricas, Tradição Gramatical e Tradição Linguística. Por fim, na terceira subseção, *A gente e você – gramaticalização, referência e indeterminação*, faremos um apanhado geral acerca do histórico de mudança de *você* e *a gente* e apresentaremos a abordagem realizada por alguns teóricos acerca das duas formas com o objetivo de observar quais aspectos relacionados à mudança das duas formas possibilitam os seus usos como estratégias para indeterminar o sujeito.

3.1 Indeterminação e conceitos afins nos dicionários

Um dos problemas que cria entraves no que se refere ao entendimento da indeterminação do sujeito parte da falta de consenso, principalmente entre as gramáticas prescritivas, no que se refere à definição do fenômeno em si e da sua classificação. Tal fato se deve, conforme veremos na subseção 2.3, à mistura de critérios que paira sobre a abordagem do fenômeno. Partindo dessa problemática, buscaremos, nesse momento, a definição de *indeterminação/determinação* em dicionários¹⁴ da Língua Portuguesa, a fim de ampliar a discussão para além do escopo da gramática prescritiva. Além da definição de *indeterminação* x *determinação*, buscaremos também a definição dos verbetes *definido/definição* x *indefinido/indefinição*, uma vez que discutiremos a indeterminação do sujeito pelo viés da referência das formas *a gente* e *você* (que pode ser definida ou indefinida), fazendo com que as definições desses termos também entrem em pauta.

¹⁴ O nosso interesse, ao elencar e discutir as definições presentes nos dicionários, é o de verificar se há relação entre os termos, portanto, por não ser objeto de nosso interesse, não levamos em consideração, para esse estudo, o propósito de cada lexicólogo com a elaboração dos dicionários.

De início, buscamos as definições dos termos supracitados no *Diccionario da Lingua Portuguesa* de Bluteau (1789). Os significados de *indeterminação* e *determinação* apresentados pelo autor não têm relação com a noção de sujeito, assim como as expressões *definido* e *indefinido*. A definição das palavras *indeterminado* e *indefinido* está relacionada a “não determinado” (BLUTEAU, 1789, p. 710). O termo *indeterminação*, por sua vez, refere-se ao “sentido vago de uma palavra” (BLUTEAU, 1789, p. 711) e, também, não faz referência a sujeito. Se considerarmos a definição proposta por Bluteau (1789), há certa correlação entre *indefinir* e *indeterminar*, no entanto, nenhum dos conceitos se volta para o sujeito.

No *Dicionário da Língua Brasileira de Pinto* (1832), *indeterminação* é definida como falta de *determinação*. A definição de *determinação* está ligada à vontade, significado mais comum e, assim como em Bluteau (1789), não se relaciona com a noção de sujeito. No compêndio, não se encontram os conceitos de *definido* e *indefinido*.

No *Dicionário de Linguística* de Dubois (1973), as definições são mais amplas e se aproximam da noção de sujeito. O autor não apresenta uma definição para o termo *determinação*, mas, sobre o termo *indeterminado*, postula: “[...] diz-se que um termo é *indeterminado* quando a noção expressa não está relacionada a circunstâncias definidas.” (DUBOIS, 1973, p. 338, grifo do autor). Buscando, no mesmo autor, o conceito do termo *definido*, a fim de esclarecer a definição de *indeterminado*, encontramos duas acepções:

1. Em gramática tradicional, o *artigo definido* especifica que o substantivo que segue designa uma coisa ou uma pessoa precisa. O *artigo indefinido* indica a ausência de uma especificação precisa. Em português, o *artigo definido* é *o(s)*, *a(s)* e o *indefinido* *um (uns)*, *uma (s)*.
2. Em gramática gerativa, chama-se *definido* um traço inerente a certos artigos por oposição ao traço – *definido* inerente a outros artigos (V. INDEFINIDO); esse traço “definido” é interpretado semanticamente como criando uma referência precisa ou como tendo um valor genérico. Assim, *A pessoa (que você sabe) telefonou* opõe-se a *Uma pessoa telefonou* ou *O homem é um animal que fala* opõe-se a *Um homem não poderia agir assim*. (DUBOIS, 1973, p. 167).

A definição de Dubois (1973) para o termo *indefinido* (encontrado na página 337) é a mesma apresentada quando trata do termo *definido*. O autor se vale, de modo complementar, da oposição entre os pronomes *ele*, que, nas suas palavras, é definido em relação a *a gente*, que não é; e o artigo *o*, definido, com relação a *um*, indefinido. Com base nas definições da gramática tradicional, Dubois (1973) ainda postula que há pronomes substantivos que são indefinidos por enumeração, caso dos vocábulos: *nenhum*, *outro*, *outrem*, *cada um*, *o mesmo*,

*ninguém, alguém*¹⁵, *todos* etc. No entanto, embora discorra sobre a natureza indefinida dos referidos pronomes, reconhece, prudentemente, que entre essas palavras, algumas são indeterminantes (indefinidas), mas outras exprimem uma determinação completa. A título de exemplificação, o autor cita o emprego do termo *o mesmo* na sentença “*Quero o mesmo livro que você*” que, apesar de ser considerado indefinido em termos de forma e, portanto, indeterminante, segundo o autor, exprime uma determinação completa. Desse modo, supomos que o traço de indefinição que integra os pronomes indefinidos mencionados pelo autor é relativo/discutível e dependente do contexto de emprego na sentença.

O que se depreende, até aqui, das definições de Dubois (1973) é que o conceito de indeterminado está diretamente ligado a uma relação opositiva entre *definido* e *indefinido*. Essa oposição é exemplificada, no contexto das definições e com bases nos conceitos da gramática tradicional, por meio de vocábulos – artigos e pronomes – que podem ou não integrar esse valor de definição e estendê-lo para o substantivo, na sentença, criando uma designação precisa ou imprecisa de uma pessoa ou coisa.

Percebemos que, de modo geral, a classificação dada aos pronomes indefinidos, exemplos dos quais o autor lança mão para apresentar as definições, gira em torno da noção de **referência**, que pode ser genérica ou não, conforme observamos na oposição que o autor faz entre as formas pronominais *ele* e *a gente*, sendo a primeira considerada definida em oposição à segunda. É dessa noção de referência *indefinida* x *definida* ou *definida* x *genérica*, possíveis de serem encontradas nas formas pronominais *a gente* e *você* que partimos quando analisamos a indeterminação do sujeito. O que difere as definições propostas por Dubois (1973), com base na Tradição Gramatical, do norte que seguimos é o fato de acreditarmos que essa referência não é rigidamente inerente às formas pronominais *a gente* e *você*, mas negociáveis no contexto de uso. Então, entendemos a relação *referência definida* x *referência indefinida* dentro de um contínuo em que é possível para o falante, através de uma única forma, transitar entre esses valores, o que depende do contexto de interação.

Camara Junior (1997), em consonância com Dubois (1973), relaciona o conceito do termo *indeterminado* com *indefinido*, indicando, também, correlação desses com o sujeito sintático. O sujeito indeterminado ou indiferenciado, para o autor, é aquele que faz referência “[...] à massa humana indiferenciada.” (CAMARA JUNIOR, 1997, p. 229). Pelo que se pode

¹⁵ Tanto na Tradição Gramatical (ROCHA LIMA, 1962) quanto na linguística descritiva (TEIXEIRA, 2014), é possível encontrar autores que defendem o uso de pronomes indefinidos (*alguém, ninguém, etc...*) como estratégias de indeterminação do sujeito. No entanto, há, também autores (BECHARA, 2009) que não consideram esse uso como caso de indeterminação, mas de impessoalização.

observar, as definições de Camara Junior (1997) também estabelecem relação entre os as noções de *indeterminado* e *indefinido*. Além disso, embora o autor não faça relação direta, a ideia de sujeito indeterminado está atrelada à de referência genérica de um termo.

Além dos autores mencionados, buscamos os conceitos referidos também em Crystal (2000 [1941]). No *Dicionário de linguística e fonética* do autor, a definição de *indeterminação/indeterminado* não se relaciona diretamente com a noção de sujeito. As definições de *definido* e *indefinido* encontradas em Dubois (1973) são semelhantes à de Crystal (2000 [1941]). Os dois autores definem os dois termos como estando opostos, embora (Crystal (2000 [1941] defenda que uma distinção entre eles não é de todo clara por envolver aspectos de ordem extralinguística. No que se refere a *indefinido*, especificamente, o autor diz ser um “Termo usado na GRAMÁTICA e na SEMÂNTICA para caracterizar uma entidade (ou classe de entidades) que não tem uma identificação específica” (CRYSTAL, 2000 [1941], p. 145, grifos do autor). A noção de indefinido, segundo ele, é transmitida por meio de artigos e pronomes indefinidos (elementos, também, citados por Dubois (1973).

A definição dos termos *indefinido* e *definido* como estando relacionados a uma **entidade** que pode possuir ou não **identificação específica** proposta por Crystal (2000 [1941]) está, de certo modo, correlacionada com a noção de referência que adotamos aqui, uma vez que consideramos ser *a gente* e *você* indeterminados ou não quando possuem referência definida ou não, ou seja, quando identificam/especificam ou não os seus referentes no mundo.

Em suma, pelas definições arroladas, percebemos que, de modo geral, os conceitos de *indeterminação* e *indefinição*, *definição* e *determinação* são, até certo ponto, intercambiáveis e invocam as noções típicas do processo de indeterminação do sujeito: os sentidos de *vago* e *genérico*. Nessa perspectiva, a ideia de sujeito indeterminado que adotamos como sendo a de um termo ou sintagma de referência genérica/indefinida em oposição a um sujeito determinado, cuja referência seja definida/específica encontra, por esse viés, respaldo nas definições propostas pelos autores consultados. Buscaremos, a partir disso, na subseção seguinte, verificar se as gramáticas, tanto históricas e prescritivas quanto descritivas, tratam a indeterminação pelo mesmo viés.

3.2 A indeterminação do sujeito nas gramáticas

Nesta subseção, faremos, conforme anunciamos, um apanhado geral sobre a abordagem do fenômeno de indeterminação do sujeito em três perspectivas: das *gramáticas históricas*, da *Tradição Gramatical* e da *Tradição Linguística*, incluindo, nesta, as gramáticas descritivas e

pesquisas recentes. Discutiremos, dessa forma, os modos como o fenômeno é definido e classificado por alguns autores das três perspectivas, buscando evidenciar os pontos comuns e divergentes entre as abordagens.

3.2.1 Nas gramáticas históricas

Bueno (1944) define o sujeito como sendo “[...] a parte da oração da qual se declara alguma coisa.” (BUENO, 1944, p. 286). Em relação ao sujeito indeterminado, o autor afirma: “Diz-se que o sujeito é *indeterminado* quando realmente não existe, sendo o verbo impessoal: *Chove – Troveja – Faz dez dias.*” (BUENO, 1944, p. 287). Observamos que o gramático trata a indeterminação por um viés estritamente sintático, uma vez que define o sujeito como uma parte da oração e a indeterminação como um processo que consiste na ausência do sujeito, ou seja, na omissão ou impossibilidade de realização dessa parte da oração. Nesse sentido, segundo a classificação proposta pelo autor, a indeterminação em nada está ligada à relação entre um termo e seu referente, mas ao preenchimento lexical ou não do termo. O que Bueno (1944) classifica como sujeito indeterminado é reconhecido/classificado, pela Norma Gramatical Brasileira (NGB) e pela Tradição Gramatical, como oração sem sujeito, por se tratar de verbos que indicam fenômeno da natureza, em se tratando de *chove* e *troveja*, e do verbo *fazer* com sentido de *haver*, que é impessoal.

Em Cruz (1948, p. 80), o sujeito é definido como “[...] o ser de que se afirma alguma coisa”. O gramático afirma que ocorre o sujeito indeterminado “[...] quando se não pode determinar nem encontrar [*o sujeito*]. (CRUZ, 1948, p. 80, grifo nosso). As estratégias de indeterminação mencionadas pelo autor incluem apenas o uso do verbo *haver* e o uso da partícula *se*, conforme os exemplos: *há homens* e *vive-se bem aqui*, retirados do próprio autor. Assim como ocorre em Bueno (1944), Cruz (1948) não faz distinção entre a ocorrência de sujeito indeterminado e de oração sem sujeito, ressaltando o uso dos verbos impessoais como um recurso indeterminador.

Na *Gramática Expositiva*, Pereira (1945, p. 216) afirma que o sujeito é definido como “[...] o membro da proposição do qual se declara alguma coisa.”. Em relação à indeterminação, o gramático diz ocorrer “[...] quando [*o sujeito*] não é enunciado nem conhecido” (PEREIRA, 1945, p. 220, grifo nosso). As estratégias de indeterminação reconhecidas pelo autor são, assim como em Bueno (1944) e Cruz (1948), o uso do verbo impessoal e, diferentemente dos autores mencionados, o uso de certos verbos no modo infinitivo.

Almeida (1961, p. 348) diz ser o sujeito indeterminado “[...] quando de difícil identificação” e, diferentemente dos gramáticos acima apresentados, estabelece distinção entre os casos de indeterminação do sujeito e oração sem sujeito. Quanto às estratégias de indeterminação, o gramático reconhece duas possibilidades com verbos (i) ativos, acidentalmente impessoalizados na 3ª pessoa do plural: *dizem que ele vem* e (ii) acidentalmente impessoalizados na passiva: *precisa-se de um datilógrafo*.

Já na *Gramática histórica da língua portuguesa*, Said Ali (1964) trata o sujeito indeterminado sob a terminologia de *indefinido*¹⁶ e em oposição a *definido*. Para o gramático, o sujeito indefinido “[...] é o que indica ente humano que não podemos ou não queremos especificar.” (SAID ALI, 1964). Em relação às estratégias de codificação do sujeito indefinido, o autor destaca três possibilidades (i) o uso do verbo na terceira pessoa, (ii) na forma reflexiva ou (iii) verbo na forma ativa dando-lhe por sujeito um pronome indefinido.

A abordagem feita por Said Ali em 1971 é, de modo geral, equivalente à realizada na década anterior. O sujeito é também tratado a partir da oposição *definido* x *indefinido*, sendo indefinido, segundo o gramático, nos casos em que se deixa de “[...] nomear por conveniência ou por ser difícil especificá-lo” (SAID ALI, 1971, p. 268). No que se refere às estratégias de indeterminação, além das consideradas anteriormente, o autor postula também que o vocábulo *homem*, na linguagem antiga, atuava como estratégia para realizar o sujeito indefinido, quando era usado na função de pronome indefinido. Modernamente, *homem* não mais atua como pronome indefinido, mas como substantivo, sendo, por vezes, empregado acompanhado de um determinante como em *um homem* que possui, também, valor genérico.

É interessante, nesse momento, abrir uma espécie de parênteses para comentar que Mattos e Silva (1989) trata da existência de um sujeito **não-determinado** no Português Arcaico que se realiza em dois tipos de estruturas: *estruturas em que o sujeito é nulo* e *estruturas em que ele é genérico, não especificado*. No primeiro caso, a codificação se dá, segundo a autora, (i) por meio de predicado representado pelos verbos *aver* e *seer* com valor existencial; (ii) uso de *aver* indicando tempo corrido; e (iii) estruturas com verbos ou com lexias com o verbo *fazer* que expressam fenômenos naturais. No segundo caso, a autora considera a realização (i) por meio do emprego do verbo em P6; (ii) pelo vocábulo *homem* quando funciona com valor genérico, neste caso, conforme o postulado por Said Ali (1971). Essa segunda estratégia, segundo a autora, desapareceu na história do português em proveito do pronome *se* que,

¹⁶ O autor ressalta que está fazendo uso do termo *indefinido* em lugar de *indeterminado*, que é a nomenclatura adotada pela N.G.B.

sabemos, é uma das estratégias de indeterminação apresentadas pelas gramáticas prescritivas e descritivas, conforme veremos adiante.

Ao que parece, as estratégias com *estruturas de sujeito nulo* apresentadas por Mattos e Silva (1989) correspondem, hoje, ao que as gramáticas prescritivas denominam recorrentemente de oração sem sujeito, assim como vimos também em Bueno (1944), Cruz (1948) e Pereira (1945). As segundas estratégias, as de *sujeito genérico*, incluem uma que é defendida pelas gramáticas prescritivas e descritivas, a saber, o emprego do verbo em P6 e uma possibilidade que é amplamente defendida nas gramáticas e estudos descritivos e desconsideradas pelas prescritivas: o uso de *homem* como *pronome de valor genérico*, que, segundo a autora, veio a desaparecer no decorrer da história da língua, cedendo lugar ao uso do *se*. Atualmente, pesquisas¹⁷ dão conta de que há vocábulos acompanhados de determinantes (denominados de sintagmas) que exercem esse papel de indefinição como nos casos de *o cara*, *a pessoa*, *o sujeito*, *o povo*, *o pessoal*. Desse modo, observamos que o uso de vocábulo com valor genérico¹⁸ - estratégia existente no Português Arcaico - atua no português moderno.

Pelo que observamos diante do que foi exposto, o fenômeno de indeterminação nem sempre foi visto e descrito de modo consensual. Na abordagem de alguns gramáticos, Bueno (1944), Cruz (1948), Pereira (1945) e Mattos e Silva (1989), o sujeito indeterminado é classificado dentro dos moldes do que as gramáticas atuais consideram como oração sem sujeito. Indeterminar o sujeito, na história do português, se considerarmos a abordagem das gramáticas apresentadas, era um processo de omissão/não realização do sujeito na sentença e não um fenômeno que envolve referência, que é uma concepção mais atual. É a partir da abordagem proposta por Said Ali (1964) e Matos e Silva (1989) que há, de certo modo, distinção entre oração com sujeito indeterminado e oração sem sujeito. Outro aspecto notado nas abordagens de Said Ali (1964) e Matos e Silva (1989) é a explicação do fenômeno a partir da referência, que é a concepção de indeterminação mais difundida na Tradição Linguística, conforme discorreremos ainda nesta seção.

3.2.2 A indeterminação na Tradição Gramatical

¹⁷ No campo da linguística Descritiva, Carvalho (2010), em dissertação de mestrado, analisou o uso de tais estruturas como estratégias de indeterminação do sujeito. Além disso, Perini (2010) também discorre sobre esses usos.

¹⁸ Ressaltando o fato de que, atualmente, esse vocábulo vem acompanhado de um determinante.

As abordagens acerca da indeterminação do sujeito realizadas por diferentes autores da Tradição Gramatical, de modo geral, seguem um mesmo roteiro. As diferenças encontradas estão em dois aspectos: (i) no uso dos critérios de definição, pois alguns autores lançam mão de critérios estritamente sintáticos, outros de critérios semânticos, pragmáticos ou discursivos para definir o fenômeno e para classificá-lo; e (ii) na enumeração de algumas estratégias, ponto em que nem sempre há consenso entre os gramáticos.

Essas abordagens são geralmente postas em discussão devido a esses dois fatores e, além disso, às definições de sujeito encontradas nos manuais dão margem para que se entenda o fenômeno de indeterminação por outros ângulos e criam possibilidades para outras estratégias de ocorrência além das tradicionalmente prescritas, conforme discutiremos a seguir.

Rocha Lima (1962, p. 226) afirma que o sujeito é o “[...] ser de quem se diz algo” e que ele é indeterminado “[...]se não pudermos ou não quisermos especificá-lo.” (ROCHA LIMA, 1962, p. 226). Segundo o gramático, a língua se vale das seguintes estratégias para indeterminar o sujeito: (i) empregar o verbo na terceira pessoa do plural, ex.: **mataram** um guarda; (ii) usá-lo na 3ª pessoa do singular acompanhado da partícula *se*, desde que ele seja intransitivo, ou traga complemento preposicional, ex.: vive-**se** bem aqui; (iii) dar-lhe por sujeito um pronome indefinido, ex.: **Alguém** bateu à porta. A partir das definições do gramático, inferimos que os casos de indeterminação envolvem uma questão de identidade, uma vez que o sujeito é definido como um *ser* e a indeterminação é a não especificação desse ser.

Para Sacconi (1982, p. 258) o sujeito é “[...] o ser ao qual se atribui a idéia contida no predicado.” E ocorre indeterminação quando ele “não existe como elemento na oração, sendo a sua identidade desconhecida realmente, ou escondida propositadamente.” (SACCONI, 1982, p. 260). O autor cita duas das estratégias já mencionadas: (i) emprego do verbo na 3ª pessoa do plural ou (ii) verbo+*se*, exceto se ele for transitivo direto e faz ressalva quanto ao uso de pronomes indefinidos como um recurso que não indetermina o sujeito oracional, contrariando o que postula Rocha Lima (1962). Nesses casos, segundo o gramático, trata-se de sujeito simples, pois o sujeito existe como elemento, embora a sua identidade não seja conhecida. Observamos, na definição proposta por Sacconi (1982) que o fenômeno de indeterminação pode ser tanto a não existência do sujeito enquanto elemento da oração quando o desconhecimento ou omissão da identidade do sujeito, definições que, evidentemente, estão em dois domínios diferentes.

Cipro Neto e Infante (1998; 2003), diferentemente dos gramáticos mencionados, definem o sujeito como o termo que está em concordância com o verbo da oração. O sujeito indeterminado, na visão dos autores, ocorre quando não se quer ou não se pode identificar a

quem o predicado da oração se refere. Os gramáticos consideram apenas as duas estratégias de indeterminação mencionadas por Sacconi (1982).

Já Cunha e Cintra (2001), assim como Rocha Lima (1962) e Sacconi (1982), definem o sujeito como um *ser* e afirmam que ele é indeterminado quando “[...] 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.” (CUNHA E CINTRA, 2001, p. 128). Em relação às estratégias que podem ser utilizadas para produzir esse tipo de sujeito, os gramáticos destacam as duas possibilidades mais recorrentes (i) empregando o verbo na terceira pessoa do plural e (ii) empregando-o na terceira pessoa do singular seguido do *se*. O conceito de indeterminação e os recursos linguísticos apresentados para produzir esse tipo de sujeito na língua, presentes em Cunha e Cintra (2001), são dois mais recorrentes já mencionados. Em Cunha e Cintra (2010), a abordagem encontrada é semelhante.

Em sua *Gramática Escolar da Língua Portuguesa*, Bechara (2006) define as orações com sujeito indeterminado como aquelas que “[...] não apresentam nenhuma unidade linguística para ocupar a casa ou função de sujeito [...]” (BECHARA, 2006, p. 21). Segundo o gramático, nesses casos, há uma referência a sujeito, mas essa referência é imprecisa, indeterminada. Em relação às estratégias de indeterminação do sujeito, o autor considera três possibilidades: (i) o uso do verbo na terceira pessoa do plural sem referência a termo anterior ou posterior que lhe sirva de sujeito; (ii) uso de verbo no infinitivo ou (iii) verbo na terceira pessoa do singular com valor de terceira pessoa do plural; e (iv) uso de verbo na terceira pessoa do singular, seguido da partícula *se*, que atua como índice de indeterminação do sujeito.

Bechara, em 2009, por sua vez, não define de modo preciso o fenômeno da indeterminação, mas deixa subentendido que se trata de uma referência imprecisa do sujeito. No que se refere às estratégias que geram o fenômeno, considera, as mesmas possibilidades elencadas anteriormente. A abordagem proposta por Bechara (2009), conforme observamos, em pouco se difere da apresentada em (2006).

Dessa forma, em uma visão geral da Tradição Gramatical, a respeito da indeterminação, temos, quanto à definição, algumas estratégias que dão conta da indeterminação como (i) *um processo de omissão do sujeito em si*; (ii) *a omissão da identidade do sujeito (sem especificação de que sujeito se trata)*; e (iii) *a não referência a uma pessoa determinada*. Quanto às estratégias de codificação, há duas que são encontradas em todos os autores da Tradição Gramatical consultados, a saber, (i) *o uso do verbo em terceira pessoa do plural sem termo referente anterior ou posterior* e (ii) *uso do verbo mais a partícula se*.

Nesse ponto, vale ressaltar os três casos em que foram consideradas outras estratégias além das duas mais recorrentes. O primeiro caso está em Rocha Lima (1962), que considera o uso de pronomes indefinidos (alguém, ninguém...) como uma estratégia de indeterminação, o que é contrário às abordagens tanto de Saconni (1982) quanto de Bechara (2006) que excluem essa possibilidade. Para Saconni (1982), as construções com pronomes indefinidos são “construções especiais” que não correspondem ao sujeito indeterminado. O segundo caso encontra-se em Bechara (2006; 2010) que considera tanto o uso de verbos no infinitivo quanto o uso de verbos na terceira pessoa do singular com valor de terceira pessoa do plural como possibilidades para indeterminar, o que não é encontrado ou comentado por nenhum outro autor da Tradição Gramatical consultado para esse estudo. Sobre o caso do uso de verbos no infinitivo, é pertinente destacar que constitui uma estratégia defendida por autores da Tradição Linguística, conforme veremos na próxima subseção.

3.2.3 A indeterminação na Tradição Linguística

Esta subseção que se inicia está dividida em duas partes. Na primeira, traremos para a discussão a abordagem acerca do fenômeno de indeterminação do sujeito feita pelas gramáticas descritivas. Em seguida, observaremos como o fenômeno vem sendo estudado por pesquisadores da área da linguística em análises mais recentes vinculadas a programas de pós-graduação *stricto sensu*. Trataremos, neste último caso, de análises distintas em relação aos aportes teóricos utilizados, a fim de garantir uma abordagem ao mesmo tempo panorâmica e verticalizada do fenômeno.

3.2.3.1 O que dizem as gramáticas descritivas?

Na linguística descritiva, o fenômeno de indeterminação do sujeito é tratado de modo mais amplo e as possibilidades de codificação não se restringem às duas preconizadas pela Tradição Gramatical. Essa diferença de abordagem se deve à própria definição de sujeito indeterminado que as gramáticas descritivas apresentam, conforme discutiremos.

Perini (2010) define a indeterminação como “[...] o fenômeno que consiste em entender mais ou menos esquematicamente a referência de um sintagma” (PERINI, 2010, p. 83). Desse modo, segundo o autor, um sintagma pode ser mais determinado ou menos determinado (que equivale ao indeterminado). O primeiro caso ocorre quando o próprio sintagma oferece meios para que o receptor identifique a pessoa de quem se trata. O segundo ocorre quando o sintagma

assume uma referência mais geral, tornando o sujeito, portanto, menos determinado. Desse modo, segundo Perini (2010), quanto menos individualizada for a referência do sintagma, mais indeterminado será o sujeito. Há, portanto, para o linguista, graus de indeterminação que são marcados gramaticalmente pelas línguas.

Em relação às estratégias de indeterminação, o autor elenca cinco possibilidades: (i) a utilização do sintagma nominal sem um determinante - *ex.: criança suja muito o chão*; (ii) o uso do verbo sem sujeito na terceira pessoa do plural - *ex.: quebraram a janela*; (iii) o verbo sem sujeito na terceira pessoa do singular - *ex.: nessa fazenda planta café e milho*; (iv) o uso do infinitivo sem sujeito - *ex.: nadar é bom para a saúde*; (v) o uso de determinados itens lexicais (*o cara, o sujeito, a pessoa*) com referência restrita a seres humanos e pronomes pessoais como *você, a gente, eles, tu* – *ex.: ou você pertence àquele grupo de críticos ou te ignoram completamente.*¹⁹

Castilho (2010; 2016) analisa o sujeito indeterminado pelo ponto de vista da referência. Para o autor, “Um sujeito /referencial/ é aquele que destaca determinado referente dentre o conjunto de referentes possíveis que compartilham as propriedades indicadas pelo sintagma nominal-sujeito.” (CASTILHO, 2010, p. 297). Correlacionada com a noção de referência, está a questão da *definitude* do sujeito. Com base em Dubois (1973), Castilho (2016) postula que o sujeito pode ter referente *definido*, quando esse ocupa papel de destaque no texto ou –*definido* quando passa fugazmente pelo texto. Em ambos os casos, segundo o autor, a identificação é feita na dinâmica do discurso e por meio do rastreio dos sujeitos sentenciais do texto.

Castilho (2010) afirma ser a indeterminação a propriedade semântica do sujeito mais explorada na Gramática Tradicional. O autor discorre sucintamente sobre três estratégias possíveis para a realização do sujeito indeterminado, considerando o traço de definitude já mencionado: (i) Sujeito expresso por pronomes pessoais com referência genérica, estratégia que é exemplificada por Castilho (2016) com o uso do *você* com valor genérico; (ii) sujeito expresso pelo pronome *se* e (iii) sujeito elíptico, com verbo na terceira pessoa do plural. Em relação às estratégias mencionadas, o autor ressalta que “[...] nenhuma classe possível de figurar como sujeito será intrinsecamente determinada ou indeterminada” (CASTILHO, 2016, p. 298). Nesse sentido, todas elas devem ser validadas no texto, nas palavras do autor. A análise que desenvolvemos acerca do fenômeno de indeterminação está, de certo modo, alinhada com a ressalva de Castilho (2016), uma vez que defendemos ser as formas gramaticalizadas *a gente* e

¹⁹ Exemplos extraídos de Perini (2010).

você funcionam como estratégias de indeterminação em contextos específicos e, acima de tudo, na língua em uso.

Na *Gramática Houaiss da Língua Portuguesa*, Azeredo (2008) trata o fenômeno de indeterminação como o não preenchimento do sujeito. Segundo o autor, isso se deve “[...] à impossibilidade de se estabelecer qualquer correspondência entre o sujeito da oração e algum item léxico da língua.” (AZEREDO, 2008, p. 224) e o seu emprego se deve a razões cognitivas e discursivas diferentes, que abarcam *desconhecimento da identidade do ser de quem se fala ou conveniência na ocultação de sua identidade*.

Azeredo (2008) destaca três estratégias mais comuns para a realização do sujeito indeterminado: (i) o uso do pronome se; (ii) o emprego de infinitivos; e (iii) o emprego da terceira pessoa do plural. Segundo o autor, “No registro informal típico da conversação espontânea, a indeterminação do sujeito faz raro uso do *se*, dando preferência ao emprego da terceira pessoa do plural ou mesmo do singular.” (AZEREDO, 2008, p. 226). De modo complementar, afirma que o uso do *se* em qualquer variedade da língua “[...] permite que o enunciador se inclua, não como indivíduo, mas como parte do grupo social, na **referência genérica ou indeterminada** do sujeito.” (AZEREDO, 2008, p. 226, grifo nosso). No entanto, ressalta que

[...] na interação espontânea, a omissão da identidade do sujeito se exprime corretamente mediante o emprego de sintagmas nominais de significação genérica ou indeterminadora como o *a gente, muita gente, todo mundo* no papel de sujeito simples. (AZEREDO, 2008, p. 226).

Até certo ponto, a abordagem realizada por Azeredo (2008) é semelhante à das gramáticas prescritivas. No entanto, assim como Perini (2010) e Castilho (2010; 2016), o autor se difere das abordagens prescritivas por tratar da indeterminação do sujeito como uma questão que envolve referência e, de certo modo, por envolver questões cognitivas e discursivas e não unicamente sintáticas. Além disso, estratégias que não são consideradas pela tradição, como o uso de sintagmas nominais com referência genérica, são tidas pelo autor como verdadeiras possibilidades de omissão da identidade do sujeito.

Encontramos, na *Gramática do Brasileiro de Ferrarezi Junior* (2008), a definição de indeterminação do sujeito que mais se distancia de uma visão puramente sintática proposta pelas gramáticas prescritivas. O autor afirma que

A primeira coisa que se deve notar é que, no chamado sujeito indeterminado, o que é indeterminado não é o sujeito (o termo sintático), mas o referente ao qual ele

pretensamente se refere. Se o sujeito sintático fosse indeterminado, o verbo enlouqueceria. (FERRAREZI JUNIOR, 2008, p. 192).

No que se refere às estratégias de indeterminação, Ferrarezi Junior (2008) menciona apenas duas: (i) uso da palavra *se* com alguns verbos na terceira pessoa do singular; e (ii) o emprego de verbos na terceira pessoa do plural. Nesse aspecto, conforme se observa, a abordagem do autor em nada se difere da abordagem prescritiva. No entanto, a definição de sujeito indeterminado proposta por ele está em consonância com a maioria das gramáticas descritivas e, nela, está exposta a natureza do fenômeno de indeterminação tal qual adotamos aqui, que a indeterminação não é do termo, uma vez que, enquanto termo sintático, o sujeito não pode ser considerado indeterminado. A questão, então, está muito mais voltada para a referência do que para a sintaxe em si. Ou seja, o fenômeno de indeterminação do sujeito é uma questão de ligação entre a estrutura linguística e um referente no mundo. Assim, quando essa ligação é possível de ser estabelecida e o referente no mundo é claramente identificável, não há indeterminação, mas, quando o referente do mundo não é claramente associado à estrutura linguística, há o sujeito indeterminado.

Dessa forma, a ideia de indeterminação transcende claramente o nível sintático e se arranja na semântica, pragmática ou discurso. Essa questão se mostra evidente quando tomamos como exemplo de indeterminação o uso de verbos na terceira pessoa do plural como no excerto: *derrubaram o muro*. Nesse exemplo, a indeterminação não pode ser vista apenas pelo viés sintático, uma vez que pela desinência verbal o sujeito sintático é preenchido pelo pronome *eles*. A questão, então, é referencial e trata-se de saber quem é o referente de *eles* do mundo que praticou a ação de derrubar. A ligação entre o praticante da ação e o pronome *ele*, nesse sentido, não é evidente, mas o sujeito sintático é claramente detectável. Por essa razão, a definição de indeterminação deve levar em consideração, acima de tudo, (i) a questão da referência e, em segunda mão, (ii) o fato de que se trata de questões outras que não de sintaxe pura.

3.2.3.2 O que dizem as pesquisas?

O fenômeno de indeterminação do sujeito no Português Brasileiro tem sido amplamente investigado no campo da Linguística Descritiva. As pesquisas realizadas, principalmente em programas de pós-graduação (mestrado e doutorado), têm dado conta do fenômeno por diferentes perspectivas. Considerando essa diversidade de abordagens, buscamos observar os resultados de três diferentes análises. Os trabalhos revistos são os de Carvalho (2010), Assunção (2012) e Teixeira (2014) e a escolha dessas três pesquisas se justifica pela perspectiva que cada

autor adota para tratar o fenômeno de indeterminação e pelo diálogo que há entre as hipóteses que lançamos e as que foram consideradas por cada um deles:

- Carvalho (2010) se ampara em pressupostos do Funcionalismo e da Sociolinguística e chega à conclusão de que as formas pronominais *você* e *a gente* são as estratégias de indeterminação mais produtivas entre falantes de Salvador-BA, além de considerar o grau de escolaridade como um forte condicionador da escolha entre as estratégias disponíveis;
- A abordagem de Assunção (2012) se dá apenas pelo viés da Sociolinguística e a pesquisadora considera a indeterminação como um fenômeno ligado à referência, conforme fazemos aqui;
- Teixeira (2014), que lança mão de pressupostos do Funcionalismo em sua análise, considera, assim como nós, a atuação de fatores de ordem pragmática e vê a indeterminação como um fenômeno que é codificado no contexto de interação emissor-receptor.

Desse modo, observaremos em que aspectos os resultados até então obtidos pelos pesquisadores supracitados estabelecem diálogo com as nossas hipóteses e resultados.

Carvalho (2010), em dissertação de mestrado, investigou os recursos dos quais lançam mão os falantes do português popular e culto de Salvador para codificar a indeterminação do sujeito. Ancorado no quadro teórico da Sociolinguística Quantitativa e considerando, também, pressupostos do Funcionalismo, o autor teve como objetivo investigar se havia e quais seriam as estratégias alternativas às apresentadas pelas gramáticas tradicionais para codificar a indeterminação. Além disso, investigou os condicionamentos sociais e linguísticos que se mostraram motivadores em relação aos usos encontrados nos *corpora* que foram extraídos do PEPP - Programa de Estudos sobre o Português Popular Falado de Salvador e do NURC/SSA.

Na análise, o pesquisador controlou as variáveis **linguísticas** *tempo e modo verbal, tipo de oração, tipo de verbo, forma antecedente, mudança/manutenção do referente, preenchimento do sujeito e grau de indeterminação* e as variáveis **extralinguísticas** *escolaridade, faixa etária e sexo/gênero*. Após o cruzamento dos dados, o autor concluiu que os falantes de Salvador de fato utilizam estratégias complementares às recorrentemente apresentadas pela Gramática Tradicional para marcar o sujeito indeterminado. Dentre essas estratégias, as principais são as formas pronominais *você* e *a gente* que, segundo o pesquisador, são influenciadas por fatores tanto linguísticos quanto extralinguísticos. Além das formas pronominais, as formas nominais (o cara, o indivíduo, o sujeito, nego, o povo etc.) foram também largamente utilizadas nos *corpora* analisados. Do total de 2.595 ocorrências de sujeito

indeterminado encontradas pelo autor, 873 (33%) corresponde ao uso da forma *você*, seguida de 668 (25,7%) de ocorrência da forma *a gente*. A forma pronominal *nós*, que também é objeto de investigação nesse estudo, ocorreu 115 vezes (4,4%). O autor concluiu também que há preferência no uso de formas pronominais (*eu, tu, você, nós, a gente*) como estratégia de indeterminação em relação às não pronominais pelos falantes da capital baiana. No que se refere à atuação dos fatores extralinguísticos, Carvalho (2010) concluiu que quanto maior é o grau de escolaridade, maior é o número de estratégias utilizadas para indeterminar o sujeito em Salvador.

Também em pesquisa de mestrado, Assunção (2012), partindo do pressuposto de que há mais estratégias de indeterminação em uso no português brasileiro além das propostas pela GT, realizou estudo da variedade linguística de Feira de Santana - BA, com base no arcabouço teórico e metodológico da Sociolinguística Variacionista e com dados extraídos do *corpus* do projeto *A língua portuguesa falada no semiárido baiano*, do Núcleo de Estudos da Língua Portuguesa (NELP-UEFS). A autora considerou a atuação de variáveis **linguísticas** *tempo e modo verbal, tipo de oração, tipo de verbo e preenchimento ou não do sujeito* e **extralinguísticas** *sexo/gênero, faixa etária e escolaridade*, chegando à conclusão de que os falantes de Feira de Santana fazem uso de estratégias alternativas às apresentadas pela GT para indeterminar o sujeito, entre elas *eu, você, a gente, nós, eles, formas nominais (FNs), Ø+3PS e o Ø+VINF*, sendo as mais preferidas, segundo a autora, as formas inovadoras *a gente* e *você*.

A pesquisadora chegou, dentre outras, à conclusão de que a indeterminação não se restringe às estruturas de sujeito lexical não preenchido, por tratar-se de um fenômeno linguístico pertencente aos níveis semântico e discursivo, uma vez que o que é indeterminado é a referência do sujeito, o que faz com que o fenômeno seja dependente do contexto. No que se refere à atuação das variáveis, Assunção (2012) concluiu que as variáveis linguísticas não foram determinantes, mas, no que se refere ao *você*, o tempo e modo verbal futuro do subjuntivo foi o que mais condicionou a escolha. Os fatores sociais, em contrapartida, foram selecionados como favorecedores do uso do *você*. Segundo a autora, os falantes jovens mostraram maior sensibilidade ao uso da forma. Além disso, o estudo demonstrou que há tendência no maior uso da forma *você* por homens, mas influência menor da variável *escolaridade* na escolha tanto da forma *você* quanto *a gente*, embora mulheres com ensino fundamental I tenham feito uso mais frequente da forma *a gente* para indeterminar.

Fazendo ponderações acerca da atuação das variáveis, Assunção (2012) chegou à conclusão final de que as variantes *você* e *a gente* mostram-se como preferidas pelos feirenses

para indeterminar o sujeito, mas tendo o *você* maior número de escolhas em relação ao *a gente*, o que, segundo a autora, sinaliza uma mudança em curso.

Já Teixeira (2014), em tese de doutorado, motivada pela dificuldade de aliar a noção tradicional de classificação do sujeito indeterminado à prática de ensino de língua e com o objetivo de demonstrar que a indeterminação é um processo notadamente pragmático de diferentes formas de codificação morfossintática, investigou o modo como a indeterminação do sujeito ocorre no português brasileiro, em textos do tipo argumentativo, veiculados no meio jornalístico. A autora trabalhou com *corpus* estruturado a partir de diversos gêneros textuais de base argumentativa (editorial, artigo de opinião, carta do leitor e entrevista), e assumiu a posição de que a indeterminação é uma estratégia acionada pelo emissor no contexto da interação emissor-receptor, o que fica mais evidente em gêneros textuais desse tipo. É válido ressaltar que a nossa abordagem, embora seja feita por um viés diferente, dialoga com a de Teixeira (2014) uma vez que a autora considerou ser o contexto de argumentação favorecedor da indeterminação, hipótese da qual também lançamos mão quando consideramos como variável a *superestrutura textual* (argumentativa, narrativa, descritiva), conforme discorreremos na próxima seção.

Partindo do pressuposto de que a indeterminação do sujeito é um fenômeno pragmático, Teixeira (2014) considerou que o fenômeno é motivado pela pragmática e codificado morfossintaticamente. Após a montagem do *corpus*, cujos dados foram extraídos dos cadernos publicados no *Jornal Folha de São Paulo*, a autora realizou a leitura dos textos selecionados e a localização das formas codificadoras de indeterminação do sujeito. Posteriormente, as formas foram agrupadas e as ocorrências encontradas foram, então, avaliadas sob a luz da teoria funcionalista.

A análise desenvolvida por Teixeira (2014) levou à identificação de três tipos específicos de indeterminação: (1) a indeterminação do tipo 1, na qual o próprio emissor desconhece o referente; (2) indeterminação do tipo 2, na qual o receptor desconhece o referente; e (3) indeterminação do tipo 3, na qual ambos desconhecem o referente. Para essa análise, conforme observamos, a pesquisadora considerou a relação entre emissor e receptor, no contexto comunicativo e o fato de que quem codifica a indeterminação é sempre o emissor. Desse modo, a indeterminação parcial do tipo 01 (+id; -s) ocorre quando o emissor codifica morfossintaticamente o seu desconhecimento do referente. Já na indeterminação parcial do tipo 02 (-id; +s), o emissor parte do pressuposto de que o receptor não consegue identificar o referente. Por fim, na indeterminação do tipo 03 (-id; -s), o emissor codifica a impossibilidade, tanto por parte do receptor, quanto por ele mesmo (o emissor) de identificar o referente.

Diante da investigação, a autora concluiu que a indeterminação do tipo 3, na qual o emissor pressupõe tanto o seu desconhecimento quanto o desconhecimento do interlocutor em relação ao referente, mostrou-se mais produtiva nos dados analisados, representando 93% das ocorrências. Por fim, e a fim de apresentar um conceito para o fenômeno de indeterminação, que foi um dos objetivos da pesquisa, a pesquisadora concluiu que a indeterminação, formulada pelo emissor como uma estratégia argumentativa, define-se como a falta de referência (**pragmática**) e, por vezes, de designação (**semântica**) de uma forma linguística de 3ª pessoa na função sintática de sujeito (**morfossintática**), demonstrando que o fenômeno, se considerado como relacionado à referência, não é apenas de natureza sintática, mas também semântica e pragmática.

As conclusões às quais os pesquisadores mencionados chegaram nos ajudam a lançar luz sobre o fenômeno de indeterminação da perspectiva em que analisamos.

Em primeira mão, a confirmação de Teixeira (2014) de que a indeterminação se dá no contexto de interação emissor-receptor, a partir do compartilhamento de informações/conhecimentos pragmáticos e semânticos no ato comunicativo nos direciona no sentido de adotar o modelo de interação verbal de Dik, sobre o qual discorreremos na seção 1, como um dos nortes da nossa análise. Assim, verificamos se esse intercâmbio de informações pode, de algum modo, ser associado à atuação das variáveis linguísticas e extralinguísticas, o que nos garantirá uma abordagem mais respaldada do fenômeno de indeterminação.

Em seguida, a constatação de Carvalho (2010) de que as formas pronominais *a gente* e *você* estão entre as estratégias de indeterminação mais recorrentes na fala dos soteropolitanos e que o grau de *escolaridade* influencia no uso de um número maior de estratégias fortalecem, de certo modo, uma de nossas hipóteses (sobre a qual discorreremos na próxima seção), a de que os falantes cultos usam o *você* e *a gente* como formas indeterminadoras em maior número de contextos do que os falantes do português popular.

E, por fim, a confirmação de Assunção (2012) de que a indeterminação está ligada à referência e não simplesmente ao preenchimento do sujeito e de que o *a gente* e *você* são as formas de indeterminação preferidas pelos feirenses, assim como em Carvalho (2010), também fortalecem as hipóteses que levantamos no nosso estudo, de modo geral, uma vez que as formas canônicas (uso do *se* e emprego do verbo na terceira pessoa do plural) não foram as mais numerosas nas três análises que consultamos ao passo que o uso de *a gente* e *você* foi numeroso e, assim, produtivo.

3.3 A *gente* e *você* – gramaticalização, referência e indeterminação

Nesta subseção, faremos um apanhado geral acerca das formas gramaticalizadas *a gente* e *você*, já que analisaremos o fenômeno de indeterminação quando ocorre por meio desses dois itens linguísticos. Trataremos, aqui, de alguns aspectos, propriedades e traços que, de algum modo, podem lançar luz sobre os motivos da escolha dos falantes por essas formas como estratégias de indeterminação do sujeito. Nessa perspectiva, a nossa intenção com essa abordagem é a de mostrar como os autores/pesquisadores em geral tratam o *a gente* e *você* enquanto formas gramaticalizadas não importando a posição que elas ocupam na sentença. Assim, tendo essa visão geral acerca dos dois itens em questão, teremos, conseqüentemente, um fio condutor que nos auxiliará na observação do modo como eles atuam enquanto recursos de indeterminação.

3.3.1 De *Vossa Mercê* a *você*: um breve histórico pela rota da mudança

O *Vossa Mercê*, forma fonte do atual *você*, surgiu, na Língua Portuguesa, como uma alternativa linguística para atender a uma demanda social: a forma de tratamento entre os indivíduos pertencentes a diferentes classes. *Vossa Mercê* surgiu em um contexto histórico em que outras formas de tratamento (o *vós*, por exemplo) não eram mais suficientemente respeitadas para que os vassalos se dirigissem aos seus reis. *Vossa* (pronome possessivo) junto a *Mercê* (substantivo que era sinônimo de graça, favor, merecimento e generosidade), segundo Sousa (2008), a partir do seu uso rotinizado, passou a ser a forma ideal para que as classes inferiores se referissem aos seus superiores. Ou seja, foi a busca de determinada classe social por nova expressão digna para a referência a outra que colocou a expressão *Vossa Mercê* na trajetória da gramaticalização, possibilitando a chegada da expressão à atual *você*, com mudanças na forma e na referência.

Segundo Santos Luz (1958), a expressão *Vossa Mercê* existiu na Língua Portuguesa entre no período de 1331 a 1481/1482 e, para Cintra (1986), ela deixa de existir por volta de 1490. Esse período marca o fim do seu uso direcionado ao rei, o que se justifica pelo fato de não ser mais uma forma expressiva (e digna) de referência à majestade. A partir desse período, o *Vossa Mercê* começa a cair no gosto popular e ser usado com referência mais ampla e não mais restrita a uma classe. Segundo Nascentes (1956, p.116), “*Vossa Mercê* agradava a todo mundo.” E isso fez com que a classe mais humilde se apoderasse da forma e a usasse extensivamente.

Em outras palavras, podemos dizer que o ponto alto da existência do *Vossa Mercê* na Língua Portuguesa é o período em que a sua referência é restrita ao rei, em substituição a outra forma, até então, vigente: o vós. Após esse período, a expressão continua na sua rota de mudança, conforme supracitado, passando a ser amplamente usada com referência mais popular, ou seja, cedendo a novas pressões sociais. Nessa trajetória de mudança, segundo Nascentes (1956, p. 114), “[...] [o Vossa Mercê] degradou-se, fonética e semanticamente, a tal ponto que mutilou extraordinariamente a sua forma e, de tratamento real, pronominalizando-se, chegou a tratamento empregado para inferiores.” Essa mudança de uso, a nosso ver, por ter ampliado as possibilidades de referência da expressão teve como um dos desdobramentos o uso da atual forma *você* com referência genérica e, quando na posição de sujeito, com valor indeterminado. Ou seja, o histórico de mudança do *Vossa Mercê* até o *você* possibilitou a existência de uma referência menos afunilada e, assim, mais espraiada da forma, que é o que favorece a indeterminação do sujeito.

Atualmente, o *você* é amplamente empregado no Português Brasileiro. Tal popularidade pode ser explicada, nas palavras de Sousa (2008), pelo fato de funcionar como uma espécie de *pronome curinga*. Ou seja, as suas possibilidades de referência são amplas, podendo se direcionar a uma pessoa, a um grupo particular, a um grupo mais amplo e, acrescentamos, ser usado com valor genérico para indeterminar o sujeito.

O processo de gramaticalização pelo qual passou o *você* possibilitou a sua chegada ao estágio atual. Conforme mencionamos, e com base em Nascentes (1956), nesse longo processo histórico de mudança, o *você* sofreu degradação tanto formal quanto semântica e, acima de tudo, adquiriu propriedades e traços não pertencentes à sua forma-fonte entre os quais julgamos estar a propriedade indeterminadora.

Encontramos, em Neves (2015; 2000), a abordagem bastante pontual acerca da propriedade indeterminadora do *você*. A autora postula que “O **pronome VOCÊ**, embora seja forma de pessoa envolvida no discurso (**segunda pessoa**), pode indicar referência genérica. A indeterminação, nesse caso, é muito forte (**VOCÊ** = uma pessoa, seja qual for) (NEVES, 2000, 463). Neves (2015) descreve a indeterminação do sujeito considerando o uso de diferentes formas pronominais (*você, eu, eles, nós, a gente*), além das formas canônicas (uso do *se* e emprego do verbo na terceira pessoa do plural). Em todos esses casos, a autora trata do fenômeno considerando a existência de uma referência precisa ou genérica, que é a posição que assumimos na nossa análise. No que se refere especificamente aos pronomes, a autora faz menção a uma “anulação de referência” das formas em relação aos papéis do discurso. Ou seja, há o emprego da forma pronominal em determinado contexto linguístico, mas a sua referência

é, grosso modo, esvaziada podendo essa forma representar um referente distinto do que designa usualmente ou, no caso da ocorrência de indeterminação do sujeito, atingir um grau de generalização, conforme notamos na análise encontrada em Neves (2015):

[...] Antigamente *você* ia ao Cine Ipiranga eram umas poltronas ótimas tinha lá em cima *você* ficava bem acomodado [DID SP 234].

[...] vocês retiram... por intervenção na paciente... *você* está com... um câncer... avançado... já muitas vezes ou então se faz a mastectomia alargada... se retira a glândula... se retira a grande peitoral... ficando somente os músculos intercostais... vocês vejam como é traumatizante na intervenção [EF SSA 49].

Aí o pronome *você* (“*você* está com... um câncer... avançado”) seguramente não inclui o ouvinte (o aluno ao qual a aula é ministrada), o que se pode considerar o grau máximo de generalização para um pronome que, tipicamente, representa o alocutário.
(NEVES, 2015, p.35)

Como podemos observar, a autora reconhece que, no contexto comunicativo, o *você*, mesmo representando tipicamente o alocutário (segunda pessoa) não o inclui, passando a um grau máximo de generalização que, na prática, configura o sujeito indeterminado.

Essa possibilidade de emprego do *você* com valor indeterminado, quando está em posição de sujeito, acreditamos e defendemos ser fruto do seu processo de gramaticalização, conforme mencionamos nas subseções anteriores. A ampliação do emprego do *você* aliada à sua gramaticalização que, em um processo “corrosivo”, garantiu à forma novas funções teve como uma das consequências também a alteração da sua referência e isso possibilita o seu uso como estratégia de indeterminação. Tal constatação pode ser confirmada na diversidade de estudos acerca da gramaticalização da forma que foram realizados no português brasileiro.

Dentre esses estudos, destacamos os de Sousa (2008) que, em análise do português falado na Paraíba, comprovou a ampliação da referência do *você*, reconhecendo a existência de três possibilidades: P1, P2 e genérico. Destacamos também a análise desenvolvida por Rocha (2017)²⁰, que, no estudo da variação entre *você* e *cê* em uma perspectiva sociofuncional no vernáculo conquistense também constatou a existência de uma referência genérica para *você* e *cê*.

Desse modo, defendemos ser essas possibilidades de referências do *você*, entre as quais se inclui a referência genérica (que configura o sujeito indeterminado), resultado direto do seu processo de mudança via gramaticalização.

²⁰ Orientado por Sousa estabelecendo, também, um estudo sobre o mesmo objeto em outra região e com perspectiva diferenciada.

3.3.2 De Gente a A gente – relação entre gramaticalização e indeterminação

A atual forma *a gente* teve origem no substantivo *gente* que, ao ser adjungido ao *a*, artigo, passou a desempenhar a função de pronome com referência (entre outras) à primeira pessoa do plural juntamente com o *nós*. O uso atual da forma é resultado de um longo processo de mudança via gramaticalização. Nesse percurso de mudança do *a gente*, vários princípios típicos da gramaticalização podem ser captados e, na nossa visão, os resultados da atuação desses princípios, e da gramaticalização em si, são os grandes motivadores do uso da forma como estratégia de indeterminação do sujeito.

A fim de fundamentar essa hipótese, faremos, a seguir, um apanhado breve e geral acerca do processo de mudança e uso do *a gente* no Português Brasileiro, com base em pesquisas, que buscaram investigar a gramaticalização da forma e realizando abordagem de gramáticas prescritivas e descritivas.

Para percebermos como se deu a gramaticalização do *a gente*, bem como a atuação de alguns princípios envolvidos nesse processo, tomaremos como base as considerações de Lopes (2004). Em estudo realizado acerca da gramaticalização do *a gente*, em *corpora* escrito do português, a autora constatou que

[...] Na gramaticalização de *gente* (nome) > *a gente* (pronome) ocorre o mesmo [atuação do princípio da persistência]. Nem todas as propriedades formais do nome *gente* foram perdidas, assim como não foram assumidas todas as propriedades intrínsecas aos pronomes pessoais. A forma gramaticalizada mantém do nome *gente* o traço formal de 3ª pessoa, embora acione uma interpretação semântico-discursiva de 1ª pessoa [+ EU]. Mesmo que o verbo em concordância com *a gente* permaneça na 3ª pessoa do singular, se pressupõe a existência de um “falante + alguém”, numa frase do tipo *a gente* precisa comprar a nossa própria casa. Tal comportamento remete-nos ainda ao princípio da decategorização (HOPPER, 1991) que consiste na neutralização das marcas morfológicas e propriedades sintáticas da categoria-origem (nome) e adoção dos atributos da categoria-destino (forma pronominal). (LOPES, 2004, p 54).

Em relação à atuação do princípio de persistência, a autora afirma que, além dos traços de concordância verbal, “Mantém-se inclusive o caráter indeterminador e coletivo do nome primitivo” (LOPES, 2004, p. 64). A hipótese que norteia a nossa análise é, entre outras, a de que o uso da forma *a gente* como estratégia de indeterminação do sujeito se deve, entre outros aspectos, à persistência de determinados traços típicos da forma-fonte *gente*, conforme postula Lopes (2004). Entre esses traços preservados, está o traço de coletividade que, em se tratando de persistência, acreditamos ser uma das bases para que o falante use o *a gente* como estratégia de indeterminação. Esse traço se manifesta na codificação do sujeito indeterminado quando o

falante emprega o *a gente* com valor genérico, sem referência a qualquer pessoa ou com referência difícil de ser especificada.

No que se refere ao princípio de decategorização mencionado, Lopes (2004) afirma que, na gramaticalização de *gente* para *a gente*, houve tanto a adoção de propriedades típicas dos pronomes, como, em se tratando do traço de gênero, a correlação com adjetivos femininos e masculinos (*a gente* está bonito/ *a gente* está bonita), quanto à perda. Nesse segundo caso, a autora destaca a perda do traço de plural que, na análise realizada, mostrou ter sido significativa já no século XVI e chegando a 100% no século XX. Curiosamente, Lopes (2004, p. 54, grifo nosso) destaca que o período de perda desse traço de plural de *gente*, que indica avanço no seu processo de gramaticalização, “[...] coincide com o desaparecimento do emprego de *homem* como indefinido.” O emprego de *homem* com valor indefinido, segundo Said Ali (1971), era, na língua antiga, uma das estratégias de indeterminação do sujeito. Tal relação existente entre a coincidência da saída do sintagma *homem* com valor genérico e a perda acelerada do traço de plural de *gente*, de certo modo, estreita ainda mais a relação entre gramaticalização e indeterminação que propomos aqui.

Além da perda do traço de plural e da adoção do traço de gênero, conforme especificamos, a análise de Lopes (2004) apontou que, por meio da descategorização, houve, também, a perda da possibilidade de *gente* ser determinado por anteposição e posposição (**aquela** gente saiu/ gente **interessante** saiu), que é um traço dos substantivos, adquirindo atributo dos pronomes que é o de não determinação no sintagma nominal. Essa oposição entre determinação e não determinação “[...] seria o principal fator que oporia uma classe à outra, determinando sua referenciabilidade” (LOPES, 2004, p. 65).

No que se refere especificamente ao traço [+/- determinado] do *gente/a gente*, Lopes (2004) afirma que

Herdando talvez o traço indeterminado do substantivo *gente*, *a gente* integra-se ao sistema pronominal concorrendo com *nós*. Na verdade, postula-se que *a gente* resultou do seguinte processo: *gente* [nome genérico] → *a gente* [pronome indefinido] → *a gente* [substituto virtual do pronome pessoal *nós*]. (LOPES, 2004, p. 73, grifos da autora).

Nessa perspectiva, o traço de [-determinado], que observamos no *a gente* em nossa análise, no estudo de Lopes (2004) é parte integrante do processo de gramaticalização de *gente* para *a gente*. A decategorização e persistência que atuaram nesse processo de passagem nome-pronome, a nosso ver, integram e reforçam esse traço encontrado atualmente no *a gente*, forma-

alvo. Assim, a possibilidade de indeterminar o sujeito com o *a gente* é, de certo modo, resultado do processo de gramaticalização pelo qual passou a forma, que, ao ter saído do domínio do substantivo e passado para o do pronome, ampliou as suas possibilidades de figurar em posição de sujeito e pela persistência do traço de coletividade pode ser empregado com valor genérico/referência imprecisa, configurando, portanto, o sujeito indeterminado. Como podemos observar, os estudos desenvolvidos por Lopes (2004) endossam a hipótese de estreitamento entre o processo de mudança da forma *a gente* e o seu uso para indeterminar o sujeito que levantamos na presente análise.

Feitas as considerações acerca da gramaticalização do *a gente*, seguiremos, discutindo o tratamento que é dado à forma no Português Brasileiro, com o objetivo de perceber se e em que medida as abordagens de gramáticas e pesquisas em geral nos auxiliam no entendimento do modo como o *a gente* funciona enquanto estratégia de indeterminação do sujeito.

- ✓ Na *Gramática de usos do português*, Neves (2000, p. 469, grifos da autora) reconhece que “Na linguagem coloquial o **sintagma nominal A GENTE** é empregado como um **pronome pessoal**”. Segundo a autora, a forma pode ser empregada:
 - a) com referência à **primeira pessoa** do **plural** (= NÓS)
 - b) para fazer referência genérica, incluindo todas as **pessoas** do discurso
 No que se refere a esse último caso, Neves (2000) postula que, em exemplos específicos, o uso do *a gente* e do *se* tem praticamente as mesmas características quanto à indeterminação do sujeito.
- ✓ Em Neves (2012, p. 25), encontramos a mesma classificação do *a gente* presente em Neves (2000), no entanto, a autora não menciona aspectos relacionados à referência da forma, que é o que nos interessa.
- ✓ Bechara (2006, p. 136) afirma que “O substantivo *gente*, precedido do artigo *a* e em referência a um grupo de pessoas em que se inclui a que fala, ou a esta sozinha, passa a pronome e se emprega fora da linguagem cerimoniosa.”
- ✓ Para Azeredo (2008, p. 176), “[...] os brasileiros empregam em geral a forma *a gente*, especialmente na língua falada semiformal e informal, como equivalente de nós, seja com um valor genérico/indeterminado (como o do pronome *se*: *não se sabe/a gente não sabe*) seja para a referência dêitica situacionalmente identificada”
- ✓ Para Koch (2001, p. 210), “A *gente* funciona, sobretudo na língua falada, como pronome e exige terceira pessoa do singular, designando normalmente um grupo em que o falante também se inclui”,

- ✓ Em Neves (2015), o *a gente* é incluído no quadro dos pronomes pessoais do Português Culto. Encontramos, na abordagem feita pela autora, a correlação mais direta entre o *a gente* e o fenômeno de indeterminação do sujeito. Na análise realizada acerca do uso de *nós* e *a gente* no Português Culto falado do Brasil, a autora constatou que o *a gente* é usado com referência mais indeterminada do que o *nós*.

Pelo que percebemos nas abordagens dos teóricos mencionados, embora não tenha sido efetivamente incorporado no paradigma pronominal do Português Brasileiro (tendo em vista que as considerações feitas por alguns autores são apresentadas em forma de notas e/ou observações), o *a gente* vem sendo reconhecido como uma forma que atua com características pronominais. No que se refere a um traço do *a gente* pronome, a referência, que é o cerne da nossa análise, observamos que, tanto no estudo da gramaticalização desenvolvido por Lopes (2004), quanto nas abordagens dos demais teóricos, a referência do *a gente* não se restringe à 1ª pessoa do plural. Alguns desses autores, Neves (2000), Azeredo (2008), reconhecem a possibilidade de emprego do *a gente* com referência mais generalizada e, inclusive, consideram que certos usos equivalem, em termos de indeterminação, ao uso da partícula *se* como nos exemplos (i) *não se sabe/ a gente não sabe* (ii) *não SE pode falar desse assunto com Carlinhos – A GENTE quer fazer um bem, vira pecado mortal*, encontrados em Azeredo (2008) e Neves (2000), respectivamente. Em Neves (2015), a associação entre o uso do *a gente* e a indeterminação, conforme citado, é mais direta, uma vez que a forma já é reconhecida pela autora como uma estratégia de indeterminação, juntamente com o *você*, sobre o qual discorreremos na subseção anterior.

Diante do que foi exposto nessa seção, partindo do princípio de que o processo de mudança via gramaticalização de *você* e *a gente* possibilitou uma ampliação das possibilidades de referenciação das duas formas, analisaremos a indeterminação ponderando que:

- (i) o *você* em posição de sujeito pode fazer referência tanto à primeira e segunda pessoa do singular quanto possuir uma referência genérica e
- (ii) o *a gente*, no mesmo sentido, quando posto em posição de sujeito, pode fazer referência à primeira pessoa do singular, à primeira pessoa do plural e, também, fazer uma referência imprecisa/genérica.

Tanto no que se refere ao *você* quanto ao *a gente*, os casos em que fazem referência genérica são os que constituem, de fato, o sujeito indeterminado. Esses casos serão, aqui, analisados levando em consideração a influência de fatores linguísticos e extralinguísticos, conforme discorreremos na seção que se segue.

4 METODOLOGIA

Apresentado o suporte teórico sobre o qual ancoramos a nossa análise (seção 2) e tendo feito o rastreio do fenômeno de indeterminação do sujeito nas gramáticas e dicionários (seção 3), procederemos, nesta seção, com a descrição do passo a passo metodológico adotado na nossa análise. Essa seção está subdividida em três subseções: na subseção 4.1 – *Descrição dos corpora*, na qual trataremos da composição/estruturação dos *corpora* PPVC e PCVC e da sua estratificação; na subseção 4.2 - *Passo a passo de análise*, na qual delinearemos o caminho percorrido no tratamento dos dados; e, por fim, na subseção 4.3 – *O envelope de variação: descrição das variáveis*, na qual apresentaremos as variáveis linguísticas e extralinguísticas controladas e as hipóteses que lançamos sobre as suas respectivas atuações.

4.1 Descrição dos corpora

Os *corpora* utilizados na nossa análise foram elaborados a partir de entrevistas coletadas entre os anos de 2011 e 2015 pelos pesquisadores do Grupo de Pesquisa em Linguística Histórica e Sociofuncionalismo CNPq– Grupo Janus, coordenado pela Prof.^a Dra. Valéria Viana Sousa e pelo Prof. Dr. Jorge Augusto Alves da Silva. Os *corpora* são constituídos de um total de 48(quarenta e oito) entrevistas, sendo 24 (vinte e quatro) entrevistas do *Corpus* Português Culto de Vitória da Conquista, *Corpus* PCVC, e 24 (vinte e quatro) entrevistas do *Corpus* do Português Popular de Vitória da Conquista, *Corpus* PPVC.

A constituição dos *corpora* segue os pressupostos metodológicos da Sociolinguística Quantitativa Laboviana, conforme delineamos na seção 2, em que a coleta de dados se dá a partir de diálogo entre informante e documentador – entrevistas do tipo DID - tendo sempre como ponto de partida um questionário previamente elaborado composto por questionamentos gerais acerca do informante (Cf. Anexo 1). Embora seja baseada em um roteiro, a entrevista sofre variações, uma vez que o objetivo do pesquisador é sempre o de alcançar um estado de

fala em que o informante esteja o mais distenso possível, ou seja, o vernáculo em termos labovianos.

Após a realização das entrevistas, que foram gravadas e duraram, em média, 60 (sessenta) minutos, os pesquisadores procederam com a transcrição dos dados. Para isso, tomaram como base a chave de codificação adotada pelo Grupo Vertentes²¹ do Português Popular do Estado da Bahia, coordenado pelo Prof. Dr. Dante Lucchesi – UFBA. Após esse processo, os dados transcritos passaram por processo de revisão e foram disponibilizados para a realização de pesquisas. Além das entrevistas, foram preenchidas, inicialmente, fichas sociais (Cf. Anexo 2), que acompanham as entrevistas transcritas, nas quais estão contidos dados gerais acerca de cada informante.

As 48 entrevistas que compõem os *corpora* PPVC e PCVC estão estratificadas em *sexo*, *faixa etária* e *grau de escolaridade*. Das 24 (vinte e quatro) entrevistas que compõem cada um deles, 12 (doze) são de informantes do *sexo* feminino e 12 (doze) de informantes do *sexo* masculino. Em relação à *faixa etária*, foram considerados três grupos: o grupo I, composto por falantes jovens; o grupo II, composto por informantes adultos; e o grupo III, composto por informantes com idade mais avançada. Em relação ao grau de *escolaridade*, para a formação do *Corpus* PCVC, foram realizadas entrevistas com informantes que possuíam 11 anos ou mais de *escolaridade*; já a formação do *Corpus* PPVC se deu com informantes que não tinham *escolaridade* ou que tinham até 5 anos de *escolaridade*.

4.2 Passo a passo de análise

A princípio, das 48 (quarenta e oito) entrevistas às quais nos referimos acima, selecionamos, para compor a nossa análise, um total de 24 (vinte e quatro), sendo 12 (doze) do *Corpus* PCVC e 12 (doze) do *Corpus* PPVC. A nossa seleção seguiu a estratificação dos *corpora* em si, sendo selecionadas 12 (doze) entrevistas de cada *sexo* e, entre elas, informantes das três faixas etárias: grupo I (informantes de 15 a 35 anos); grupo II (informantes de 36 anos a 49 anos) e grupo III (informantes com 50 anos ou mais).

Após a composição da amostra, procedemos com a seleção das ocorrências de *você* e *a gente*, mantendo cada forma dentro de um contexto considerável, a fim de garantir o máximo de coerência na classificação e codificação das ocorrências. Faz-se necessário ressaltar que, no processo de seleção, foram encontradas ocorrências de *a gente* e *você* tanto em *posição de sujeito* quanto em *posição de objeto*. Nesse caso, consideramos, para a análise, em função dos

²¹ <http://www.vertentes.ufba.br/>

objetivos da pesquisa, ora dissertação, apenas as ocorrências das formas em *posição de sujeito*, excluindo, assim, as formas em *posição de objeto*, já que não são o foco da presente análise.

Feitas as seleções e descartadas as ocorrências que não se enquadraram na análise que propomos, procedemos com a codificação dos dados, aplicando, para cada ocorrência, os códigos específicos de cada variável linguística e extralinguística, preparando, assim, os dados para serem tratados no programa *GoldVarb X*, utilizado como suporte para obtermos as estatísticas e os pesos relativos.

4.3 O envelope de variação: descrição das variáveis

A escolha de *você* e/ou *a gente* como estratégias de indeterminação do sujeito é, a nosso ver, influenciada tanto por fatores de ordem linguística quanto extralinguística. Em ambos os casos, na pesquisa sociolinguística e funcionalista, essa influência é monitorada a partir da descrição do envelope de variação, no qual são apresentadas a *variável dependente* e as *variáveis independentes*, também denominadas de grupos de fatores, que podem ser *linguísticas(os)* e *extralinguísticas(os)* e suas respectivas variantes, o que faremos nas subseções a seguir.

4.3.1 Variável dependente

Conforme discutimos na seção 2, *Sociolinguística* + *funcionalismo* = *sociofuncionalismo*: o que se quer da(s) teoria(s), no Português Brasileiro, há diversas e diferentes formas das quais os falantes podem lançar mão para codificar o sujeito indeterminado. Entre essas possibilidades, há duas que são mais recorrentes tanto na Tradição Gramatical quanto nas pesquisas em geral: a (i) *flexão do verbo na terceira pessoa do plural*; (ii) o uso da *partícula se*, entre outras, sendo que cada uma dessas formas envolve processos e mecanismos linguísticos específicos.

Há, além dessas, outras estratégias de codificação do sujeito indeterminado, como as descritas por Carvalho (2010), Perini (2010), Assunção (2012). Entre todos esses recursos de indeterminação, apenas dois compõem a nossa variável dependente: (i) o *você* e (ii) o *a gente*, em casos como os exemplificados nos excertos 01 e 02 retirados dos *corpora* analisados:

(01)Tipo assim VOCÊ tem uma pessoa que VOCÊ acha que é sua amiga... por um certo momento é aí depois por causa de um mal entendido a pessoa no caso...no caso eu vou contá como foi a história... (S.J.S.)

(02)(GIs/APh1+ Sempre... sempre reagi bem com isso, *num* me afeta muito não, acho que é ordem natural das coisas, entendeu? A GENTE se apega as pessoas e A GENTE fica triste quando elas vão mas é inevitável, né? (F.S.L.B.)

A escolha pelo *you* e *a gente* se deu por ser do nosso interesse, conforme mencionamos na seção anterior, verificar como aconteceu o processo de mudança via gramaticalização que garantiu às duas formas a função indeterminadora. Na nossa análise, monitoramos a escolha entre uma e outra forma a partir dos condicionadores linguísticos e extralinguísticos descritos nas duas subseções que seguem.

4.3.2 Variáveis independentes linguísticas

A escolha das variáveis linguísticas a serem controladas na análise se deu com base na leitura de alguns autores que, direta ou indiretamente, desenvolveram investigação acerca da indeterminação do sujeito ou da referência das formas *you* e/ou *a gente*, são eles: Ponte (2008), Sousa (2008), Carvalho (2010) e Rocha (2017). No total, três variáveis foram controladas: 1 - *superestrutura textual*, 2 - *tempo verbal*, 3 - *mudança ou manutenção do referente*. Elencaremos, a seguir, as variantes de cada uma delas e as hipóteses que norteiam a nossa discussão.

4.3.2.1 Superestrutura textual

A *superestrutura textual* está associada ao que conhecemos na literatura por *tipos textuais* que, segundo Marcuschi (2006), diz respeito a certa sequência de enunciados que se caracteriza por aspectos linguísticos em geral (tempos e modos verbais, estruturas sintáticas e relações lógicas) presentes na composição dos gêneros textuais. Os tipos textuais são cinco: *descritivo*, *narrativo*, *expositivo*, *argumentativo* e *injuntivo*. Na nossa análise, controlaremos a variável *superestrutura textual* na perspectiva de Van Dick (2000) que trata de três tipos: (i) **superestrutura descritiva**, que se caracteriza pela presença de um tema, subtemas e expressões caracterizadoras desses elementos; (ii) **superestrutura narrativa**, que segue um esquema de resumo, orientação, complicação, solução, avaliação e coda, segundo (LABOV, 1967), e (iii)

superestrutura argumentativa, que, grosso modo, conta com a presença de uma tese e de argumentos, segundo Sousa (2008).

Em todos os três casos, a *superestrutura textual* é de difícil delimitação pelo fato de, nos textos, as superestruturas se interpenetrarem e, principalmente, pelo fato de o falante, no ato interativo, usar dessa prerrogativa, e transitar entre uma e outra tipologia textual. Desse modo, consideraremos, em nossa análise, a presença dos traços básicos e dominantes de cada uma delas em torno do contexto em que se encontram as ocorrências em análise.

A nossa hipótese, em relação à atuação dessa variável, é de que o uso tanto do *você* quanto do *a gente* como sujeito indeterminado seja predominante em *superestruturas argumentativas*, pois, durante o ato comunicativo, a nosso ver, tende a ser comum a não identificação de um referente tanto pela falta de necessidade, quanto pela tentativa do falante de se “eximir da responsabilidade” de assumir um posicionamento frente a determinados temas e, ainda, pelo fato de, por meio da indeterminação, apresentar um espécie de sentimento coletivo (e não individual) ao argumento, o que faz parte do jogo comunicativo. Nas superestruturas narrativa e descritiva, por outro lado, tanto há a necessidade de identificação dos personagens envolvidos nos acontecimentos mencionados pelo falante quanto há a necessidade de especificar o que ou quem está sendo descrito, o que favorece a identificação dos referentes de *você* e *a gente*. O monitoramento da atuação dessa variável mantém, de certo modo, relação direta com outra variável linguística, o *tempo verbal*, conforme discorreremos abaixo.

4.3.2.2 *Tempo verbal*:

A variável *tempo verbal* será analisada segundo a abordagem das Gramáticas Tradicionais. Serão consideradas, então, as ocorrências dos tempos (i) **presente** (ii) **pretérito** ou (iii) **futuro**) em verbos que se articulam com o *você* e o *a gente* na oração em que se encontram. Essa variável está, de certo modo, articulada com a *superestrutura textual* e a nossa hipótese é a de que o verbo no *presente* favoreça o uso de *você* e *a gente* como sujeito indeterminado, pois esse *tempo verbal* é mais comum em superestruturas argumentativas. Os verbos no pretérito, em nosso ponto de vista, são mais comuns em superestruturas narrativas e descritivas e favorecem mais o uso de *você* e *a gente* com referente identificável/específico e, portanto, sujeito determinado, o que pode ser explicado pelo fato de se tratar de relatos envolvendo pessoas específicas cujas identidades sejam conhecidas pelo informante e codificadas por ele no ato comunicativo.

Em relação ao tempo futuro, há tanto a possibilidade de favorecer o uso de *você* e *a gente* como sujeito indeterminado ou não indeterminado, uma vez que pode ser usado em argumentações, conforme mencionamos, ou na revelação/descrição de roteiros e/ou planos que envolvem pessoas específicas, inclusive o informante.

Esclarecemos, nesse ponto, que não defendemos que a ocorrência de verbos no presente, pretérito ou futuro estejam intrinsecamente associadas ao uso das superestruturas mencionadas, tratando-se, portanto, de uma hipótese restrita à nossa análise que pode ou não ser comprovada com a análise dos dados.

4.3.2.4 *Mudança ou manutenção do referente*

Levaremos em consideração, com essa variável, a manutenção ou não do referente de uma forma quando vem imediatamente após outra igual. Ponderaremos, portanto, duas possibilidades (i) **manutenção do referente** (ii) **mudança do referente**, sendo os casos de primeira ocorrência de uma sequência ou de ocorrência isolada considerados, para efeito de análise, como os que o fator não se aplica. A nossa hipótese é de que o referente seja mantido entre as formas que se mantêm dentro de uma sequência, de um mesmo contexto e/ou *superestrutura textual*. Assim sendo, haveria relação entre esse grupo de fatores e o anterior, ou seja, há sempre uma sequência de formas iguais dentro de determinado contexto e nessa sequência o referente será, hipoteticamente, preservado.

Ressaltamos que, para a análise dessa variável, consideramos a existência de uma série de ocorrências iguais desde que (i) o informante não seja interrompido pelo entrevistador e (ii) o assunto/tópico em discussão não seja modificado. Ou seja, quando o entrevistador interrompe o informante e quando o tópico em discussão é modificado, a forma subsequente será considerada como nova.

4.3.3 *Variáveis independentes extralinguísticas ou sociais*

Seguindo os pressupostos da Sociolinguística Quantitativa presentes, principalmente, nos postulados de Labov (2008 [1972]), consideraremos, na nossa análise, também a atuação de variáveis sociais na escolha de *você* e *a gente* como estratégias de indeterminação. Para tanto, monitoramos a atuação das variáveis *faixa etária*, *grau de escolaridade* e *sexo*.

4.3.3.1 Faixa etária

A variável *faixa etária* é constituída de três fatores: (i) **faixa I: de 15 a 25 anos**; (ii) **faixa II: de 26 anos a 50 anos**; e (iii) **faixa III: mais de 50 anos em diante**. Em relação à hipótese levantada acerca da atuação desse grupo de fatores, partimos do princípio de que, pelo fato de a função indeterminadora de *você* e de *a gente* ter sido resultado dos seus históricos de mudança via gramaticalização, esse uso seria inovador e, por isso, predominante na fala dos informantes que pertençam, principalmente, à *faixa etária* I, que, em tese, teriam “mais conhecimento” sobre o funcionamento dessa nova função das duas formas. Na fala dos informantes pertencentes às faixas II e III, em contrapartida, haveria a predominância do uso de *você* e *a gente* com referência específica e, portanto, a ocorrência do sujeito não indeterminado.

4.3.3.2 Grau de escolaridade

A variável *grau de escolaridade* está composta por dois fatores (i) **[+escolarizado]**, formado por informantes com até 5 (cinco) anos de *escolaridade* e (ii) **[-escolarizado]**, formado por informantes com 11 (onze) anos ou mais de *escolaridade*. A decisão de controlar a atuação dessa variável parte da hipótese de que os informantes mais escolarizados tendem a fazer uso de *você* e *a gente* como estratégias de indeterminação por terem, devido às competências desenvolvidas ao longo da sua formação, mais habilidade no encaixamento das duas formas em contextos específicos em que elas funcionem com referência indefinida, configurando sujeito indeterminado. Tal fato, entretanto, não indica que falantes menos escolarizados não façam uso das mesmas estratégias, mas que isso ocorra em menor escala.

4.3.3.3 Sexo

No campo da Sociolinguística, as diferenças linguísticas existentes entre os sexos são apresentadas a partir da premissa de que as mulheres são mais sensíveis que os homens em relação ao padrão de prestígio, conforme postula Labov (2008 [1972]), o que é mais evidente nos estereótipos. Além disso, há também a premissa de que as mulheres podem ser ora conservadoras das formas linguísticas, ora líderes das mudanças (SOUSA, 2008). Nesse sentido, a fim de verificar se as mulheres apresentam índices de maior uso de *você* e *a gente* com a função inovadora, função de indeterminar, em oposição aos homens que,

hipoteticamente, usariam mais as duas formas com os seus valores característicos, monitoramos a atuação da variável *sexo*, que é composta dos fatores (i) **homem** e (ii) **mulher**.

Para essa variável, assim como a anterior, ressaltamos que hipotetizamos um uso maior sendo realizado pelas *mulheres*, no entanto, não excluimos a possibilidade de indeterminação com o *você* e o *a gente* por parte do *sexo* masculino.

É, portanto, a partir de cruzamento entre as variáveis linguísticas *superestrutura textual*, *tempo verbal* e *mudança ou manutenção do referente* com as variáveis extralinguísticas *sexo*, *faixa etária* e *grau de escolaridade* que buscaremos entender quais fatores internos ou externos à língua estão em jogo quando a indeterminação do sujeito é codificada por meio das formas *você* e *a gente* no Português falado de Vitória da Conquista - Bahia. Discorreremos sobre os resultados dessa análise na próxima seção.

5 VOCÊ E A GENTE EM POSIÇÃO DE SUJEITO NO PORTUGUÊS FALADO DE VITÓRIA DA CONQUISTA: UMA ANÁLISE SOCIOFUNCIONALISTA DO FENÔMENO DE INDETERMINAÇÃO

Conforme anunciamos, nesta seção, apresentaremos os resultados da análise das ocorrências de *você* e *a gente* em posição de sujeito, doravante PS, em 24 (vinte e quatro) entrevistas dos *corpora* PPVC e PCVC, a fim de demonstrar como essas duas formas gramaticalizadas atuam enquanto estratégias de indeterminação do sujeito. A discussão que se segue resulta de um conjunto de análises (que chamaremos, aqui, de *rodada*) probabilísticas dos dados realizado a partir do programa de tratamento de dados *GoldVarb X*, que foi a ferramenta “quantificadora” usada para realizar a análise da atuação das variáveis linguísticas e extralinguísticas, descritas na seção anterior a fim de que, a partir disso, possamos e observar o comportamento de cada variável no processo de escolha do *você* ou do *a gente* como estratégias indeterminadoras ou não.

Apresentaremos, de início, uma abordagem mais geral que possibilitará uma visualização panorâmica das ocorrências das duas formas. Discutiremos, nesse primeiro momento, dados que mostram a quantidade de ocorrências que configuram o sujeito determinado e indeterminado, opondo, em termos de números, uma e outra; em seguida, daremos enfoque às ocorrências de *você* e *a gente* consideradas indeterminadas e, por fim, trataremos da atuação das variáveis linguísticas e extralinguísticas com relação à variável dependente.

Faz-se necessário, de antemão, esclarecer o critério que usamos para classificar cada ocorrência como de sujeito indeterminado ou não. Conforme especificamos nas seções anteriores, tratamos o fenômeno de indeterminação, na nossa análise, pelo ponto de vista da referência. Desse modo, no processo de codificação dos dados, uma ocorrência foi considerada como um sujeito indeterminado quando o seu referente não era identificável ou era de difícil especificação. Nesse entremeio, obviamente, questões contextuais interferem na análise, uma vez que a identificação de um referente pode ser dúbia a depender do ponto de vista. Isso nos obriga a assumir critérios bem definidos para a classificação de cada dado em função de tal peculiaridade. Optamos, então, por fazer assim:

- A partir dos estudos consultados e do primeiro olhar que lançamos sobre os dados, observamos que o *você* pode assumir três referências: **(i) P1, (ii) P2 e (iii) genérico**, conforme os exemplos presentes nos excertos 03, 04 e 05, respectivamente

(03) É tipo assim, prestar atenção quando o encarregado fosse explicar o que tinha que fazer, as veys A GENTE não entendia direito, o que era já era costumado ali cum pai {init} orientano você, aí quando a pessoa ia explica VOCÊ tinha que gravar tudo até pá saber fazer, ai essa foi < barra > dificuldade... (J.S.R.)

(04) [...] aí ela disse: “ô, meu filho, não é assim, VOCÊ tem que fazer o vestibular... (L.S.S)

(05) Com certeza! Muito importante! Hoje em dia, mais ainda, porque se VOCÊ num estuda, VOCÊ num tem nada. (E.J.R)

- Do mesmo modo, o *a gente* também conta com múltiplas possibilidades: **(i) P1; (ii) P1+P2; (iii) P1+P3 ou P6; (iv) grupo de contornos indefinidos [estudantes,**

mulheres, conquistenses etc.] e (v) genérico, conforme os exemplos que se seguem, na respectiva ordem apresentada:

(06) Do que eu mais gostava? Então... eu acho que das amizades que A GENTE fazia por que... ou melhor das amizades que eu fiz... por que... até hoje A GENTE lembra né, das pessoas... que a gente... que eu convivi né, os colegas que a gente... que A GENTE passou... né, por as mãos de algum... de algum professor, pelo alguns colegas e que eu não esqueço até hoje a gente... eu tenho alguns colegas que foi do ensino fundamental né, e também do ensino médio. (A.S.A)

(07) Guilherme tem um monte de livro antigo que ele comprava que ele apaixonava, hoje ele não dá muito importância *O Diário de um banana*, só que ele qué dois, aí eu falei: “vai fica difícil” “tá vendo aí, quê que a gente lê e fica aí escolhendo pelo preço” eu falei: “tá bom, Guilherme, depois A GENTE conversa” eu pego emprestado do meu sobrinho pra ele lê, porque se realmente ele gostá, ele lê um inteiro... (L.S.S.)

(08) A GENTE se conheceu quando eu tinha treze anu. Meu esposo chego de de (deixa conta a história toda)... Ele morava aqui em conquista, ele é catingueiro daqui de conquista, e eu morava em Macarani, ai ele chego... ele foi trabalhá pá meu pai, chego lá meu pai tava com muita gente fui trabalha meu pai nisso A GENTE noivo e casô. (M.C.A.O.)

(09)[...] um parque que ficasse ali permanente né, um parque de diversões permanente não só crianças, mas nós também adultos... com certeza A GENTE iria gostar... (R.F.)

(10) Eu acho isso, meio que... Brasil, na sua formação, desde sua formação, ele sempre foi um país mestiço. Então, é uma... é mais hipocrisia da população, porque A GENTE sabe que o país é mestiço, que o país tem pessoas diferentes. (L.C.S)

A partir dessa constatação, analisamos cada dado considerando que:

- O *você* se referindo a **P1** e **P2** são determinados em oposição à referência **genérica**, que assumimos como sendo sujeito indeterminado.

- O *a gente* configura sujeito indeterminado apenas quando a sua referência for a um **grupo de contornos indefinidos** ou **genérica**. Os outros casos, em contrapartida, foram considerados como sujeito determinado.

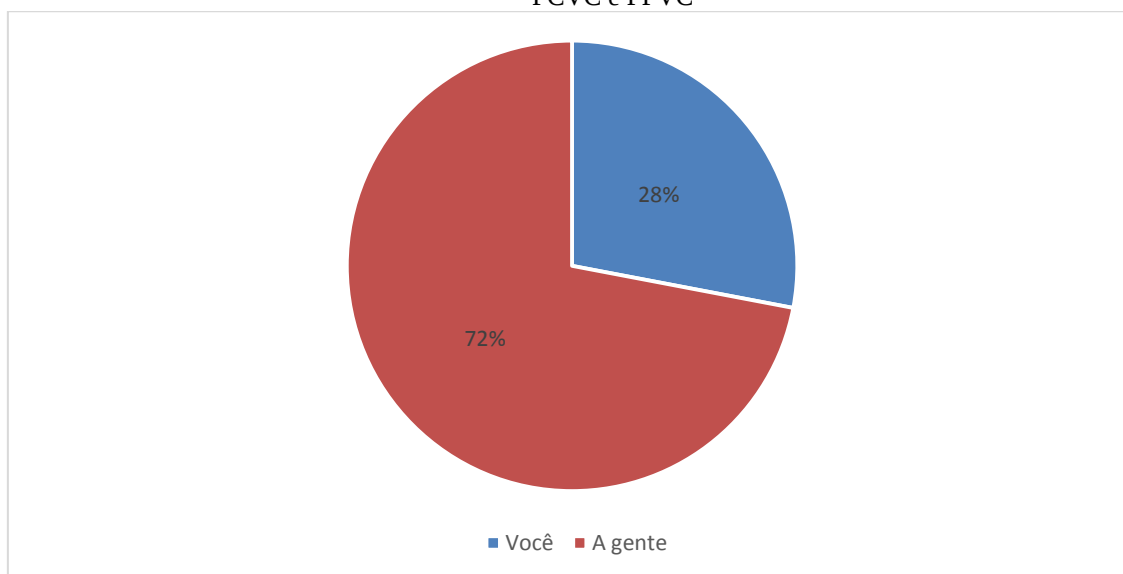
Feitas essas ressalvas, apresentaremos, na subseção seguinte, os resultados da nossa análise, seguidos das respectivas discussões.

5.1 Primeira rodada: *você e a gente* indeterminado x determinado

Nas 24 (vinte e quatro) entrevistas que compõem a nossa amostra, foram registradas 1261 ocorrências das formas *você* e *a gente* em PS, ressaltando que, nessa rodada inicial, contabilizamos todas as ocorrências das duas formas (tanto as que representam sujeito indeterminado quanto as que representam sujeito determinado), a fim de opor uma e outra e verificar, em um contexto geral, o comportamento das duas formas.

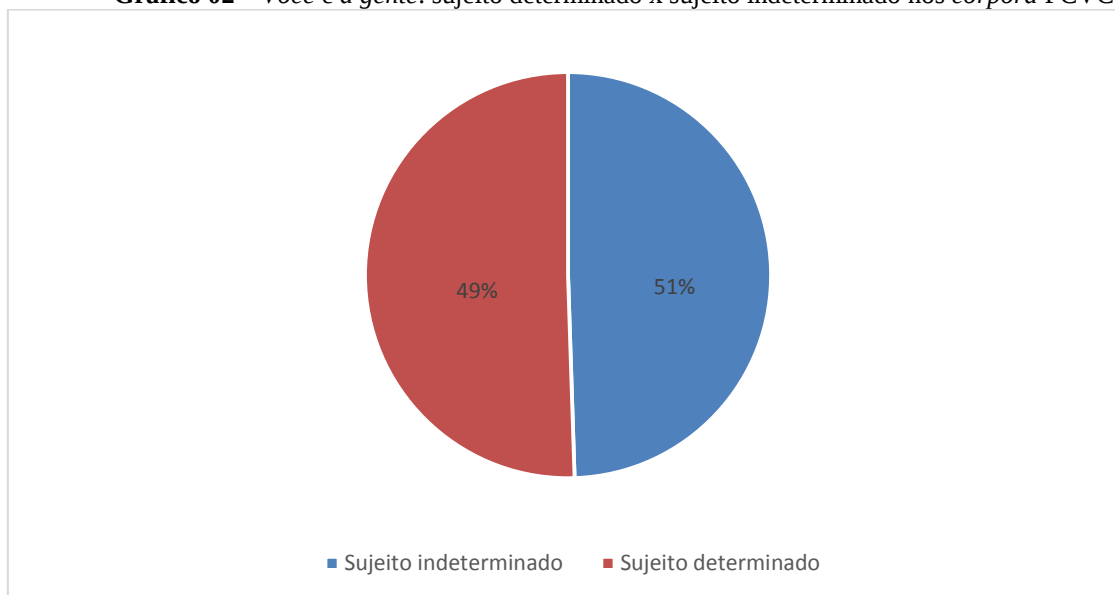
Das 1261 ocorrências encontradas nos *corpora*, 908 (72%) são de *a gente* e 353 (28%) de *você*, conforme exposto no Gráfico 01:

Gráfico 01 – Distribuição geral das ocorrências de *você* e *a gente* em posição de sujeito nos *corpora* PCVC e PPVC



Fonte: Elaborado pelo pesquisador.

Desse total geral de ocorrência das duas formas, 624, que correspondem a 49% dos dados, são de sujeito indeterminado por terem referência indefinida em oposição a 637 de sujeito determinado, 51% dos dados, como consta no Gráfico 02:

Gráfico 02 – *Você e a gente*: sujeito determinado x sujeito indeterminado nos corpora PCVC e PPVC

Fonte: Elaborado pelo pesquisador.

Esse resultado, obtido a partir da primeira rodada dos dados, confirma a impressão inicial que motivou a presente análise: tanto *você* quanto *a gente* atuam enquanto estratégia de indeterminação do sujeito no Português falado de Vitória da Conquista, embora as duas formas não tenham a função indeterminadora como exclusiva e, do ponto de vista prescritivo, como a função prototípica, já que há certo equilíbrio no uso de ambas as formas como sujeito determinado e indeterminado, conforme confirmam os dados.

Das 624 ocorrências de sujeito indeterminado localizadas nos corpora, 373 são de *a gente*, representando 60% dos casos e 251 são de *você*, representando 40%. A Tabela 01 ilustra a distribuição das ocorrências de *você* e *a gente* como sujeito indeterminado. Vejamos:

Tabela 01: *Você e a gente* indeterminados nos corpora PCVC e PPVC

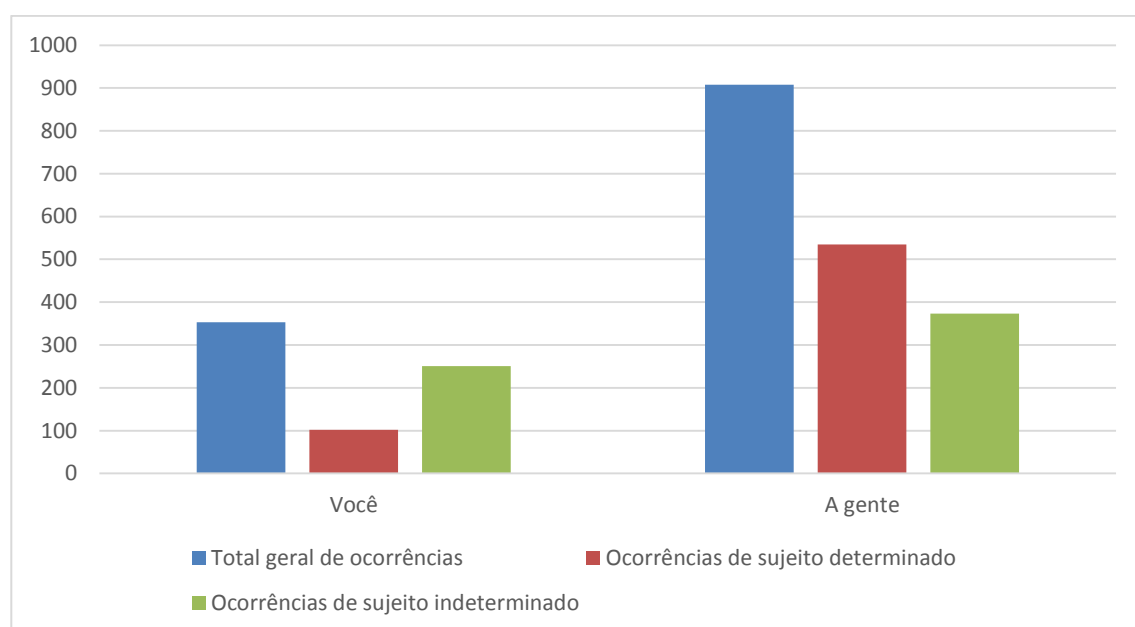
Variante	Total	%
Você	251/624	40%
A gente	373/624	60%

Fonte: Elaborada pelo pesquisador

Em primeiro plano, podemos observar que, embora o número geral de ocorrências do *a gente* (908) tenha sido superior ao número de ocorrências do *você* (353) nos corpora (conforme exposto no Gráfico 1), quando se trata do sujeito indeterminado, o *você* apresenta um maior número do que o *a gente*. Assim, essa rodada inicial demonstra que o *você* está sendo usado com menor frequência em relação ao *a gente* quando a análise não distingue sujeito determinado

de indeterminado, mas, quando a análise é afunilada para o processo de indeterminação, o *você* indeterminado está se mostrando mais produtivo que o *você* determinado, enquanto o *a gente* determinado se sobressai em relação a *gente* indeterminado. O Gráfico 03 ilustra essa diferença, constatemos:

Gráfico 03 – *Você e a gente*: distribuição geral e relação determinado x indeterminado



Fonte: Elaborado pelo pesquisador.

Conforme observamos, as ocorrências de *a gente* enquanto sujeito determinado são superiores às de sujeito indeterminado, enquanto o contrário ocorre com a forma *você*. Essa diferença é um indício de que o processo de gramaticalização do *você* pode estar mais avançado em relação ao *a gente*, no que diz respeito à frequência associada ao valor de indeterminação, e já provocou o que Lehmann (1982) denomina de *dessemantização*, caracterizada pela perda de traços gerada pela automatização do uso de uma forma, ou seja, pela alta frequência de seu emprego.

No caso do *você*, esse processo provocou a perda do traço *+específico* e agregou à forma a função indeterminadora que vem sendo amplamente usada pelos falantes conquistenses. Esse tipo de alteração recebe respaldo também no princípio de *descategorização* de Hopper (1993), conforme discorremos na seção 4, segundo o qual uma forma em processo de gramaticalização

tende a perder determinadas propriedades e assumir outras. Podemos, por meio dos dados, inferir, com base na *descategorização* (em nível semântico e não sintático), que *você* “perde”²² a propriedade exclusiva de identificar referentes específicos no discurso e adquire a propriedade indeterminadora. Dentro da linguística funcional, conforme já destacamos nas seções anteriores, esse processo não é visto apenas como perda, mas como a “substituição” de uma função por outra inovadora. É válido ressaltar ainda que esse processo de “convivência” de duas funções para uma forma (tal qual ocorre com *você* e *a gente*) remonta ao princípio da *estratificação*, já que a adesão de uma nova função não implicou a perda da função antiga, ou seja, *você* e *a gente* podem ser usados como sujeito determinado e indeterminado em contextos específicos.

Em suma, os dados presentes no Gráfico 02 permitem inferir que o processo de dessemantização e, conseqüentemente, de gramaticalização que está ocorrendo com o *você* ocorre também com a forma *a gente* (o que é comprovado pelo seu uso como estratégia de indeterminação). No entanto, esses dados indicam também que esse processo de mudança pode estar em estágio mais avançado na forma *você*, já que o seu emprego como sujeito indeterminado, que é uma função inovadora, supera o emprego como sujeito determinado, o que não ocorre com o *a gente*. Tal fato pode levar os falantes a priorizarem o uso do *você* como estratégia de indeterminação, conforme constatação de pesquisas recentes como a de Carvalho (2008), na qual o *você* figura como forma de indeterminação mais usada pelos soteropolitanos, seguida do *a gente*.

Em um cenário mais geral, os resultados da primeira rodada demonstraram que: (i) tanto o *você* quanto o *a gente* funcionam como recurso de indeterminação do sujeito; (ii) o número de ocorrências em que o *você* configura sujeito indeterminado é superior ao número de ocorrências em que a forma codifica o sujeito determinado em oposição ao *a gente*, que apresenta maior número de ocorrências para o sujeito determinado e menor para o sujeito indeterminado. Estando observado o comportamento das duas variantes de modo mais geral, direcionaremos, na subseção que segue, a análise da atuação das variáveis independentes linguísticas e extralinguísticas.

5.2 *Você* e *a gente*: o que favorece a indeterminação?

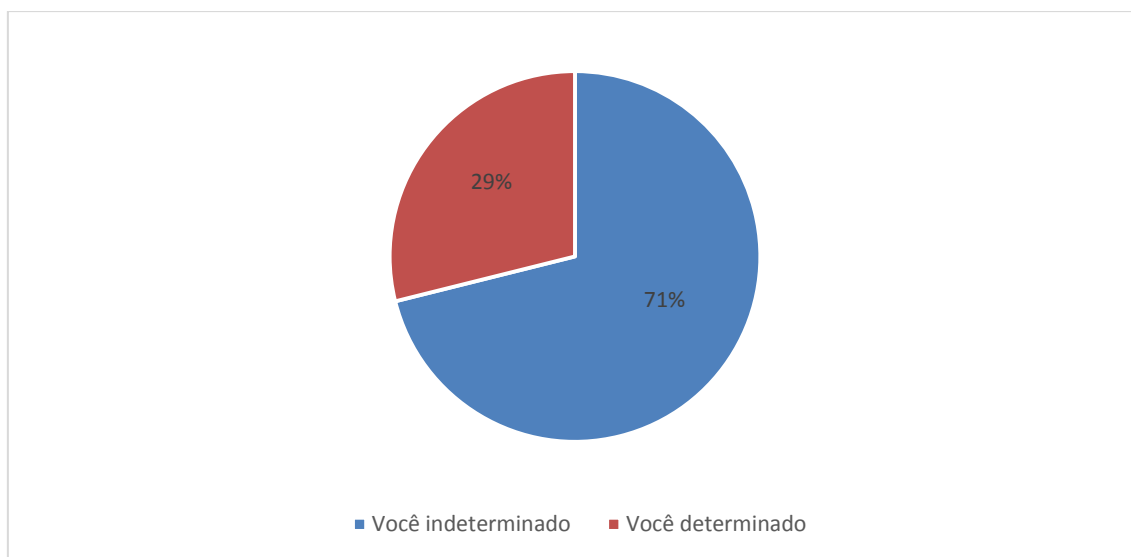
²² No caso específico do *você*, não há, ainda, uma perda da propriedade de identificar referentes específicos, mas os dados analisados indicam que o número de ocorrências com referente específico é menor que o número de ocorrências com referentes genéricos.

Nesta subseção, apresentaremos os resultados das rodadas em que opomos as ocorrências de *você* indeterminado e determinado e *a gente* indeterminado e determinado. Com essa análise, observaremos se os fatores condicionadores da indeterminação selecionados pelo programa *GoldVarb X* são os mesmos para as duas formas ou não, já que, conforme vimos no Gráfico 02, o uso do *você* e *a gente* indeterminados e determinados não estão numa mesma proporção.

5.2.1 *Você indeterminado x você determinado*

Conforme especificamos na subseção anterior, do total de ocorrências identificadas nos *corpora*, 353 são de *você*. Desse quantitativo, 251 correspondem ao *você* indeterminado e 102 ao *você* determinado, conforme expomos no Gráfico 04:

Gráfico 04: Ocorrências do *você* determinado e indeterminado nos *corpora* PCVC e PPVC



Fonte: Elaborado pelo pesquisador.

Conforme exposto no Gráfico 04, o número de ocorrências de *você* indeterminado (71%) é significativamente maior que o *você* determinado (29%). Nesse sentido, a fim de verificar quais fatores favorecem essa diferença, realizamos uma rodada opondo as ocorrências de *você* indeterminado e *você* determinado, assim como faremos com o *a gente* posteriormente.

Ao realizarmos a referida rodada, o *GoldVarb X* apontou como significativos apenas dois dos seis grupos de fatores considerados na análise, excluindo, portanto, quatro grupos (*tempo verbal*, *sexo*, *faixa etária* e *escolaridade*). Pela ordem de relevância, os grupos selecionados pelo programa foram:

- (i) *Superestrutura textual* e
- (ii) *Mudança ou manutenção do referente*

Os pesos relativos obtidos para cada uma das variáveis selecionadas serão apresentados a seguir.

5.2.1.1 *Superestrutura textual*

A variável *Superestrutura textual* nos serviu, na análise dos dados, como uma espécie de contexto. Ao controlá-la, esperávamos que o uso do *você* sofresse mudança de acordo com a superestrutura predominante. Assim, conforme expomos na seção 4, em contextos de *narração* e *descrição*, a hipótese é de que o *você* seja usado em maior número como sujeito determinado, já que o falante marcaria personagens específicos envolvidos nos fatos narrados e situações descritas. Em contextos de argumentação, em contrapartida, a hipótese era de que houvesse maior emprego do *você* enquanto sujeito indeterminado, sendo ele empregado como recurso para “eximir” o falante da tomada de posição e para atribuir um caráter menos pessoal à argumentação, reforçando, assim, o caráter de pronome “curinga”, nas palavras de Sousa (2008), que o *você* tem. Os resultados obtidos para essa variável constam na Tabela 02. Para a nossa análise, observemos:

Tabela 02: *Você* indeterminado em relação à *superestrutura textual*

	Total	%	P. R.
Narrativa	7/81	9%	.02
Descritiva	74/87	85%	.64
Argumentativa	170/185	91%	.77

Fonte: Elaborada pelo pesquisador.

Conforme observamos na Tabela 02, a hipótese aventada acerca da atuação da variável *superestrutura textual* foi comprovada para o *você* indeterminado. Com o peso relativo de .77, os dados analisados comprovam que o *você* indeterminado é mais recorrente em contextos de argumentação. Em oposição a isso, com o peso relativo de .02, os casos de *você* indeterminado nas superestruturas narrativas não mostraram alto rendimento. Já os contextos de descrição apresentaram um peso relativo de .64, também significativo. Abaixo, elencamos excertos extraídos dos *corpora* para os três casos: *narração*, *descrição* e *argumentação*, respectivamente.

(11) [...] eu nunca sofri na escola, mas na universidade aquilo começou a me causá muito mal, de chega assim, de sentir uma falta de ar tão absurda que é... eu pensei.... sabe quando VOCÊ fala: “eu vô morrê ou eu vô cair dura aqui e muito mal mesmo eu senti isso na prova de Pedro e senti em mais umas duas provas, mas nunca senti medo como senti na prova de Pedro. (L.S.S.)

(12) Eh... os métodos eu questiono, eu questiono até as mudanças, muita mudança de método, muita... muita eh... moda... moda... moda de métodos eh... num sei... num sei, mudô muito. Num é a mesma coisa. Agora se você me perguntá se no passado era mais eficiente, talvez fosse mais eficiente no passado, mas em compensação era uma outra realidade. VOCÊ tinha uma gama de conteúdos muito pequena. em relação à gama de conteúdo que VOCÊ tem hoje, que precisa sê estudado, que precisa sê repetido, cê tá entendendo? As informações hoje são muito maiores, são muito mais abrangentes. (A.I.R.M)

(13)[...] quando você vamos supô assim se você trabalhá... e *num* vê ali que VOCÊ tá sendo recompensado mais pelo aquilo que VOCÊ tá *fazeno* ‘cê *num* vai tê muita vontade de fazê aquelas, mas se VOCÊ está *fazeno* aquilo ali, tá *veno* que VOCÊ tá seno recompensado pelo aquilo ali, VOCÊ vai ter mais ânimo pra, né, agí... (E.S.P)

Conforme podemos perceber nos excertos (11), (12) e (13), há possibilidade de uso do *você* como sujeito indeterminado nas três superestruturas textuais que consideramos na análise. No entanto, os dados atestam menor número de ocorrências nos casos de narração, quando o falante faz uso do *você* com sua função prototípica: a de marcar um referente específico; e maior número de ocorrências em casos de argumentação, quando não há a obrigatoriedade de marcar referentes precisos, uma vez que a intenção do falante nesses contextos é de convencer o seu interlocutor e, por vezes, o convencimento é construído por meio do envolvimento do interlocutor, projetando-o na questão desejada através do uso do pronome *você*, nesse caso, com referência genérica. Observamos, também, a partir dos exemplos acima, que o caso de *você* indeterminado em contexto de narração encontra-se isolado, ao passo que na argumentação o falante faz uma série de repetições que, supomos, demonstra ser um meio de reforçar a tese que defende e enfatizar os argumentos expostos para o ouvinte. Essa repetição da forma com uma função inovadora (a de indeterminar) é, em nossa concepção, um impulso para que o *você* avance ainda mais em seu processo de mudança e, na língua em uso, exerça cada vez mais o papel de estratégia de indeterminação do sujeito, conforme discutimos na seção 02, com base em Neves (1997) e Lehmann (1982). Ou seja, quanto mais empregada uma forma é (no nosso caso, uma forma com uma função específica), mais rotinizada ela se tornará e mais avançada

ela estará no processo de gramaticalização, acarretando os acréscimos das novas funções e, quiçá, substituições.

É, ainda, válido lembrar, conforme mencionamos anteriormente, que as tipologias textuais se interpenetram, o que nos fez olhar atentamente para os excertos a fim de realizar a classificação adequada.

5.2.1.2 Mudança ou manutenção do referente

Na seção 04, justificamos o controle da variável *mudança ou manutenção do referente* com base na hipótese de que a atribuição de um referente a uma forma tende a se repetir dentro de uma série de formas iguais. Com base em resultados obtidos em diversas pesquisas sociolinguísticas, consideramos que o uso de uma forma leva automaticamente à sua repetição, o que já constitui uma tendência nos estudos e pesquisas dessa corrente linguística. Nesse sentido, lançamos a hipótese de que a manutenção da função também pode ser uma tendência quando se trata do *você* e *a gente*, ou seja, quando o falante usa o *você* com valor indeterminado, a tendência é que a função indeterminadora se repita nas formas subsequentes. Os resultados obtidos para essa variável estão expostos na Tabela 03. Vejamos:

Tabela 03: *Você* indeterminado em relação à *mudança ou manutenção do referente*²³

	Total	%	P. R.
Mudança do referente	18/42	43%	.23
Manutenção do referente	135/170	79%	.57

Fonte: Elaborada pelo pesquisador.

Os dados da Tabela 03, assim como ocorreu com o grupo de fatores anterior, confirmam a hipótese inicialmente lançada. Com o peso relativo de .57, os dados analisados demonstram que, em uma sequência, o uso do *você* indeterminado tende a levar a usos subsequentes da mesma forma com a mesma função, conforme ilustram os excertos (14) e (15):

- (14) [...] quando você... vamos supô... assim... se você trabalhá... e *num* vê ali que VOCÊ tá sendo recompensado mais pelo aquilo que VOCÊ tá *fazeno* ‘cê *num* vai tê muita vontade de fazê aquelas, mas se VOCÊ está *fazeno* aquilo ali, tá *veno* que VOCÊ tá seno recompensado pelo aquilo ali, VOCÊ vai ter mais ânimo pra, né, agí... (E.S.P)

²³ Esclarecemos que, nessa análise, desconsideramos as ocorrências em dois contextos (i) quando a forma era a primeira de uma série e (ii) quando a forma se encontrava isolada, uma vez que não havia um referente anterior. E, para a análise que a pesquisa demanda, apenas consideramos os casos em que a forma era antecedida de outra dentro de uma série.

- (15) Olha eu gosto dos dois porém eu sou... os dois são bem diferentes gosto de assistir pela televisão porque VOCÊ vê que que tá acontecendo, VOCÊ vê o gol ocê vê... realmente, VOCÊ enxerga tudo que tá acontecendo no campo. No rádio eu ach' também muito interessante porque tudo fica mais emocionante se abola passa a cinquenta metros de altura fora do gol o narrador tá ali "nossa foi pra fora", então VOCÊ não sabe com' é que tava a bola, então VOCÊ [diz] "poxa perdeu o gol" a emoção é outra, sabe, então se me agra... me agrada assistir nos dois jeitos. É bem legal assim é bem divertido você... VOCÊ acompanha de pertinho o que tá acontecendo também têm umas figuras inusitadas nos estádios, né, assim um pov' mei' doido dá pra rir bastante... (C.B.S.)

Já os casos em que o referente é modificado ocorrem em menor número, com um peso relativo de .23. Ou seja, a probabilidade de um falante fazer uso do *você* sucessivas vezes mudando o referente em cada uma delas é menor do que manter o mesmo referente para todas as formas repetidas. Em (16) e (17), estão transcritos alguns excertos que ilustram os casos de mudança do referente.

- (16) Sim o estudo hoje em dia ele se faz extremamente necessário na nossa vida, né, porque assim se VOCÊ almeja algo melhor uma vida estável, VOCÊ preci... necessita do... do estudo, inclusive meu pai costuma falá "olha VOCÊ estuda porque ate pra ser gari você precisa de curso superior hoje em dia"... (C.B.S)
- (17) Não! não tenho não, porque é... assim... é um lugá que eu me sinto muito bem eu acho que eu não me adaptaria a um local como esse que VOCÊ falou... assim que VOCÊ não conhece ninguém na rua, que VOCÊ num vê ninguém na rua... eu ach' que isso não me faria sentir bem morando. (C.B.S)

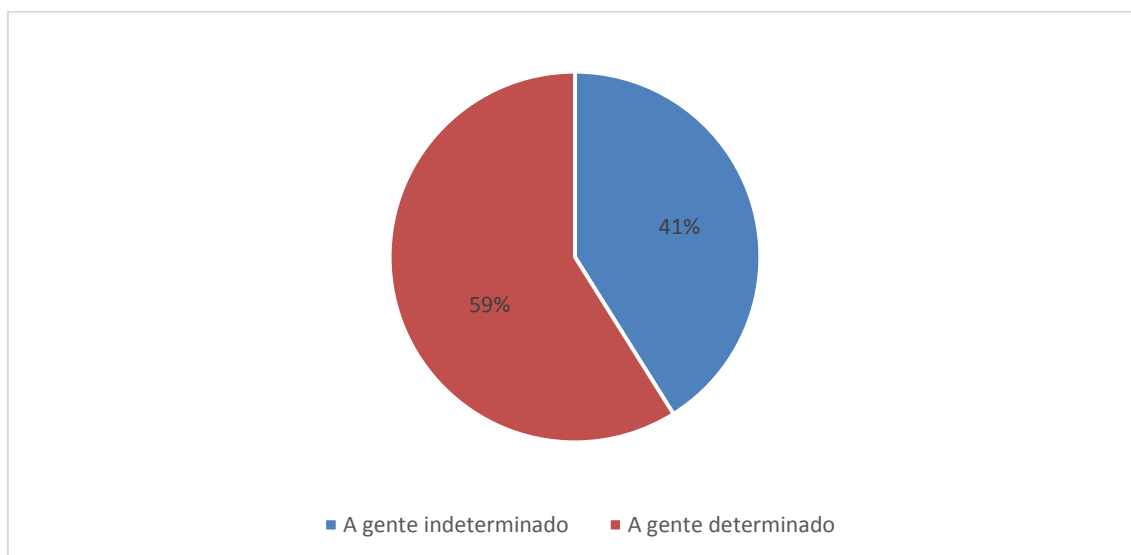
Comparando os excertos apresentados acima, podemos constatar que os casos em que o referente se mantém dentro de uma série de formas iguais são mais produtivos e comuns. Já os contextos em que há uma mudança de referente são mais precisos e pontuais. Nos exemplos em questão, um caso, excerto (16), ocorre quando o informante traz uma fala do pai como exemplo para reforçar a sua argumentação e o outro, excerto (17), ocorre quando o informante precisa retomar uma informação apresentada pelo documentador anteriormente. Em suma, embora não tenhamos controlado a variável *paralelismo formal*, uma análise qualitativa dos dados expostos acima aponta que, de fato, um uso de *você* indeterminado tende a levar a outros usos subsequentes.

Estando apresentados os dados que indicam os grupos de fatores favorecedores do *you* indeterminado em detrimento do *you* determinado, apresentaremos, na próxima subseção, conforme propusemos anteriormente, análise semelhante opondo *a gente* indeterminado e determinado, com o objetivo de verificar se as variáveis selecionadas pelo programa para uma forma correspondem às selecionadas para a outra.

5.2.2 A gente indeterminado x a gente determinado

Em relação ao *a gente*, foi computado um total de 908 ocorrências. Desse quantitativo, 373 representam sujeito indeterminado (41%) e 535 representam sujeito determinado (59%), conforme ilustramos no Gráfico 05:

Gráfico 05 - Ocorrências do *a gente* determinado e indeterminado nos corpora PCVC e PPVC



Fonte: elaborado pelo pesquisador

Ao contrário do que ocorreu com o *you*, podemos observar que o número de ocorrências do *a gente* determinado (59%) supera o número de *a gente* indeterminado (41%). Abriremos, aqui, um parêntese a fim de aventar possibilidades que possam esclarecer esse fato. A nosso ver, há duas explicações prováveis, sobre as quais já tratamos, mas retomaremos de modo mais específico para elucidar essa diferença numérica:

- (i) Primeiro, devemos considerar que *a gente* tem mais possibilidades de referência que o *you*, conforme discorreremos no início dessa seção. Isso, de certo modo, faz com que o falante o utilize em maior número de contextos para nomear pessoas/referentes específicos. Essas possibilidades de referência do *a gente* sempre podem incluir o falante e outras pessoas (irmãos, pais, familiares, amigos), tornando a forma rentável, por exemplo, em narrativas em que o falante relata situações vividas no

passado e que possam incluir esses referentes, o que não é sempre possível com o *você*, em situação de entrevista, já que o documentador (segunda pessoa) não presenciou ou fez parte desses acontecimentos. Assim, os usos do *você* ocorrem, em sua maioria, em contextos que cabem referência genérica ou em casos nos quais a forma esteja encaixada na reprodução de outra fala, conforme o exemplo (18)

- (18) Sim o estudo hoje em dia ele se faz extremamente necessário na nossa vida, né, porque assim se VOCÊ almeja algo melhor uma vida estável, VOCÊ preci... necessita do... do estudo, inclusive meu pai costuma falá “olha VOCÊ estuda porque ate pra ser gari você precisa de curso superior hoje em dia”... (C.B.S)

As três ocorrências destacadas no excerto (18) ilustram de modo claro os dois casos considerados na explicação acima. A primeira e segunda ocorrências são de um *você* genérico e a terceira refere-se ao próprio falante e está encaixada na reprodução da fala do seu pai.

- (ii) Em segunda mão, consideramos que o *você* pode estar em um estágio mais avançado de gramaticalização na função de indeterminação, conforme afirmamos anteriormente, o que pode justificar o fato de ele estar sendo mais largamente empregado em relação ao *a gente*.

Diante dos dados presentes na análise dos nossos *corpora*, não consideramos uma explicação como excludente da outra. Ao contrário, acreditamos serem as duas complementares entre si. O que os dados indicam estar acontecendo é justamente um encontro, conforme já mencionamos nas seções anteriores, entre *você* e *a gente* nas suas trajetórias de mudança via gramaticalização. Ou seja, há aí um ponto de interseção: a função indeterminadora que as formas compartilham e que não são funções prototípicas de ambas.

Portanto, a inferência que podemos fazer, a partir da leitura dos dados, é que o *você* alcançou a função de indeterminar antes do *a gente* e, portanto, está mais desligada da sua função prototípica e mais disponível para atuar na função inovadora. O *a gente*, em contrapartida, ainda é mais empregado com sua função prototípica por estar menos avançado do que o *você* em seu processo de mudança, o que se justifica, por entre outras questões, pelo fato de ter mais possibilidades de referência. Contanto, essas possibilidades de referências atualmente vigentes podem, posteriormente, passar por um afunilamento e, em um estágio mais avançado da mudança, ser mais empregado com referência genérica (função inovadora), que com referência específica (função prototípica), tal qual ocorre com o *você* atualmente, embora,

em consonância com o que já postulamos ao longo das seções anteriores, não seja possível comparar fielmente a mudança de uma forma com a da outra, já que ambas não concorrem para assumir a função de indeterminar, mas, sobretudo, coocorrem no sistema linguístico.

Por ora, a fim de verificar os grupos de fatores que favorecem a ocorrência do *a gente* indeterminado em detrimento do *a gente* determinado, realizamos uma rodada opondo as duas possibilidades, assim como fizemos com o *você*. Em resultado, o programa *GoldVarb X* selecionou quatro variáveis como sendo relevantes e, conseqüentemente, excluiu duas (*mudança ou manutenção do referente e sexo*). Por ordem de relevância, as selecionadas são:

- (i) *Superestrutura textual*
- (ii) *Faixa etária*
- (iii) *Escolaridade*
- (iv) *Tempo verbal*

De modo geral, as hipóteses que lançamos para cada variável em relação ao *você* são as mesmas que lançamos para o *a gente*, uma vez que nos interessa observar o que favorece o uso de cada uma das formas em uma oposição sujeito determinado x sujeito indeterminado. Desse modo, realizaremos a análise do *a gente* na mesma perspectiva que o *você*.

De antemão, já percebemos que o *a gente* indeterminado é favorecido pelo dobro de fatores que o *você* indeterminado, já que o *GoldVarb X* selecionou como favorecedor para o *você* indeterminado apenas dois grupos e para o *a gente* um total de cinco. Além disso, notamos ainda que apenas um dos grupos selecionados coincidiu para as duas formas, a saber, *superestrutura textual*. Tal resultado, pode ser um indício de que as trajetórias de mudança, com relação à indeterminação, das duas formas não estão, além de no mesmo estágio, no mesmo rumo, ou seja, sendo favorecidas pelos mesmos condicionantes, fato que já salientamos e retomaremos posteriormente.

Apresentaremos, nas próximas subseções, os resultados obtidos para todas as variáveis selecionadas como favorecedoras do *a gente* indeterminado acompanhados dos respectivos pesos relativos.

5.2.2.1 *Superestrutura textual*

Assim como ocorreu na análise do *você*, a variável *superestrutura textual* nos serviu como uma espécie de contexto. Assim, partimos do princípio de que, no ato interacional, falante e ouvinte estão imersos em um mesmo contexto: *argumentativo, descritivo ou narrativo* e, em cada um desses contextos, o *a gente* pode ser codificado com funções diferentes: ou

identificando referentes precisos, sendo o sujeito, portanto, determinado; ou fazendo uma referência genérica, sendo o sujeito indeterminado. É nesse ponto que adotamos a ideia do modelo interacional de Dik, sobre o qual discorreremos na seção 2, segundo o qual há um contexto com falantes em interação e, nesse contexto, há um jogo interpretativo por parte do ouvinte em relação ao que o falante codifica: falante codifica, ouvinte interpreta.

Desse modo, ao analisarmos essa variável, percebemos que há certa tendência no uso do *a gente* indeterminado em uma *superestrutura textual* específica, conforme apontam os dados presentes na Tabela 04, vejamos:

Tabela 04: *A gente* indeterminado em relação à *superestrutura textual*

	Total	%	P. R.
Narrativa	43/317	14%	.23
Descritiva	134/342	40%	.48
Argumentativa	196/249	79%	.82

Fonte: Elaborada pelo pesquisador.

Assim como esperávamos, o *a gente* indeterminado é fortemente favorecido pela superestrutura argumentativa que apresenta um peso relativo de .82. A superestrutura descritiva assumiu uma posição intermediária com um peso relativo de .48 e a narrativa com peso relativo de .23 não apresenta forte favorecimento na ocorrência do *a gente* indeterminado, confirmando, assim, as hipóteses iniciais que lançamos.

As justificativas para tais resultados, em consonância com o que já discorreremos na seção 4, residem na especificidade de cada *superestrutura textual* que serve de suporte para as ocorrências em combinação com as possibilidades de referência do *a gente*. Observemos o excerto abaixo:

- (19) [...] como foi o caso em Recife que a mulher chamou meu colega de piranguêro {risos} e não foi texto pra gente que A GENTE ficou boiando “Meu Deus o que é isso?” E aí ela explicô que piranguêro é uma pessoa que pechincha bastante e tenta compra as coisa mais barat’ possível {risos}... e aí foi que A GENTE entendeu e ela se assustô da gente não

conhecer a palavra piranguêro “como assim você não conhece piranguêro?” “Moça A GENTE não é daqui, A GENTE é da Bahia”. “Ah... pois é que A GENTE...” começou explicá...

Trata-se de um excerto em que predomina a superestrutura narrativa e, conforme podemos observar, os “personagens” que fazem parte do fato narrado são inicialmente identificados (o falante e seu amigo) e todas as ocorrências de *a gente* presentes no trecho fazem referência a eles. A identificação desses personagens é uma característica das narrativas, pois, de modo geral, narrar um fato “exige” a identificação dos participantes e, tendo o *a gente* a possibilidade de fazer essa identificação, a tendência é que ela entre mais em jogo nesse contexto. Ou seja, *a gente* com referência específica, que configura o sujeito determinado, predomina em superestruturas narrativas.

Na superestrutura argumentativa, por outro lado, o objetivo principal do falante no ato comunicativo é fazer com que o ouvinte seja persuadido de sua tese, não havendo a necessidade, pelo menos não no mesmo nível da narração, de identificar participantes/personagens, conforme ilustra o excerto 20:

(20) [...] principalmente a área da saúde que tá... eu acho que não é só... uma questão só de Conquista mais... eh... várias cidades do Brasil que passa por esses transtornos eh... a saúde é primordial e Conquista deixa muito a desejar né, e... essa questão também de infraestrutura eu acho que ainda tem muitos bairros ainda que sofre bastante com essa questão de asfalto... eh... esgoto né, e as vezes A GENTE vê assim a cidade um pouco mal cheirosa por conta disso então... é isso aí... fal... e assim... a questão também do lazer é muito difícil você encontra assim deveria mais... áreas verdes não tem né, se preocupa muito com a questão de shoppings construções de shoppings e falta investi na... na área vede que a única que A GENTE tem eh... tá um pouco devastada né, que é o Poço Escuro

Conforme podemos observar as duas ocorrências de *a gente* no excerto (20) não identificam referentes específicos, mas funcionam como uma estratégia para convencer o ouvinte de que as impressões, opiniões e sentimentos do falante com relação ao objeto, a cidade na qual residem, são consensuais. Em outras palavras, são coletivas. A nosso ver, a predominância de *a gente* determinado em contextos de argumentação se deve, entre outros fatores, a essa possibilidade que a forma tem de atribuir um caráter coletivo a um pensamento individual.

5.2.2.2 Faixa etária

A gramaticalização do *a gente*, assim como a do *você*, conforme discorremos nas seções anteriores, garantiu às duas formas novas funções. Por considerarmos a função de indeterminar como inovadora, partimos das hipóteses de que a forma *a gente* com valor de indeterminação: (1) seria mais comum na fala de informantes da *faixa etária* I, (2) apresentasse rendimento intermediário entre falantes da segunda *faixa etária* e (3) seria menos comum entre os falantes da terceira faixa. Os resultados obtidos com o controle dessa variável estão expostos na Tabela 05. Observemos:

Tabela 05: *A gente* indeterminado em relação à *faixa etária*

	Total	%	P. R.
Faixa I	83/350	24%	.27
Faixa II	122/288	42%	.55
Faixa III	168/270	62%	.73

Fonte: Elaborada pelo pesquisador.

Conforme podemos observar, os dados demonstram o contrário do que esperávamos em relação à atuação da variável *faixa etária*. Com um peso relativo de .73, a faixa etária III mostrou-se mais propensa a realizar o *a gente* indeterminado. A faixa etária I, contrariando a nossa hipótese, obteve um peso relativo de .27. Já a faixa etária II se manteve em um nível intermediário com um peso relativo de .55. É pertinente destacar que, embora tenha realizado menos a indeterminação por meio do *a gente*, os falantes da faixa etária I fizeram mais usos do *a gente* se comparado com os falantes das outras duas faixas etárias. Reconhecemos, diante disso, que por trás de tal resultado podem estar atuando outros fatores. O *a gente* indeterminado, talvez, seja uma estratégia que exija um maior amadurecimento cognitivo ou atenda a demandas mais presentes a partir dos 50 anos, contudo, reconhecemos que essas inferências carecem de um aprofundamento que pode ser realizado posteriormente em outros trabalhos.

5.2.2.3 Escolaridade

O uso do *a gente* como estratégia indeterminadora é, a partir dos estudos realizados em nossa análise, altamente dependente do contexto e, acima de tudo, parte da negociação típica da interação face a face entre falante e ouvinte. Nessa interação, o falante codifica o sujeito indeterminado e o ouvinte, por estar imerso na mesma situação interativa, “percebe” tal emprego a ponto de discernir o emprego do *a gente* genérico/indeterminado e do *a gente* específico/determinado. Considerando esses aspectos acerca da interação, sustentamos a hipótese de que falantes mais escolarizados tenham maior domínio/consciência do emprego do *a gente* indeterminado por terem grau de letramento mais elevado e a habilidade de interpretação dos usos das duas formas mais desenvolvida empregando, portanto, a forma com esse sentido/função mais vezes como estratégia discursiva.

Os resultados obtidos com o controle dessa variável estão expostos na Tabela 06. Observemos:

Tabela 06: *A gente* indeterminado em relação à escolaridade

	Ocorrências de indeterminação	%	P. R.
[+ Escolarizado]	254/465	55%	.63
[- Escolarizado]	120/443	27%	.35

Fonte: Elaborada pelo pesquisador.

Os dados da Tabela 06 evidenciam o que esperávamos da atuação da variável em questão. Com peso relativo de .63, os dados presentes em nossos *corpora* demonstram que, de fato, falantes mais escolarizados realizam mais o *a gente* indeterminado, embora o percentual de diferença seja de 10% entre o uso do determinado (45%) e o uso do indeterminado (55%). Com peso relativo de .35, os falantes menos escolarizados realizam mais o *a gente* determinado (73% dos dados). É pertinente destacar que o número geral de ocorrências do *a gente* para falantes mais escolarizados e menos escolarizados foi próximo (465 e 442), respectivamente, respaldando o fato de que o grau de letramento é um forte influenciador na codificação do *a gente* indeterminado.

Considerando e comparando os resultados obtidos para as variáveis *escolaridade* e *superestrutura textual* e vendo na argumentação um campo fértil para a ocorrência do *a gente* indeterminado, optamos por averiguar a existência de uma possível relação entre o grau de *escolaridade* e o uso do *a gente* indeterminado na superestrutura argumentativa, já que o processo de escolarização é um dos fatores responsáveis pelo desenvolvimento crítico do indivíduo cujos reflexos podem ser observados na língua em uso. Realizamos, portanto, a

primeira rodada combinando fatores dos dois grupos. Os resultados estão apresentados na Tabela 07. Vejamos:

Tabela 07: Superestrutura textual x escolaridade: combinando fatores

	Total	%	P.R
Argumentação + escolarizado	153/165	93%	.94
Argumentação - escolarizado	43/84	51%	.55
Descrição + escolarizado	79/161	49%	.58
Descrição - escolarizado	55/180	31%	.34
Narração + escolarizado	22/139	16%	.24
Narração - escolarizado	21/178	12%	.19

Fonte: Elaborada pelo pesquisador.

De modo geral, visualizando os resultados da combinação dos fatores das variáveis *Superestrutura textual* e *Escolaridade*, podemos perceber que a hipótese levantada foi confirmada. Com um peso relativo de .94, os dados demonstram que a realização do *a gente* indeterminado em superestrutura argumentativa é totalmente convergente com o fator *escolaridade*. Das 165 ocorrências de *a gente* realizadas por pessoas com grau de *escolaridade* mais elevado em contexto de argumentação, 153 representam sujeito indeterminado. Desse modo, podemos concluir que o *a gente* não é apenas uma estratégia de indeterminação do sujeito, mas também, nos exemplos analisados, configura-se como um recurso que auxilia o falante na construção da argumentação na língua em uso.

Além disso, podemos constatar também que, embora não esteja em um mesmo nível, em termos de número de ocorrências, os falantes com grau de *escolaridade* não tão elevados, mantêm um equilíbrio entre o uso do *a gente* determinado e indeterminado em superestrutura argumentativa com um peso relativo de .55.

Embora a nossa análise tenha “girado”, especialmente, em torno da argumentação, podemos perceber que os resultados relacionados à superestrutura narrativa demonstram certo equilíbrio. Ou seja, falantes escolarizados e não escolarizados fazem uso do *a gente* indeterminado (em narrações) em proporções relativamente próximas. Já na descrição, embora o uso do *a gente* entre um e outro perfil conte com diferença numérica não tão significativas como na argumentação, os falantes mais escolarizados demonstram indeterminar mais nesse contexto.

Considerando, então, essa rodada em que combinamos os fatores dos dois grupos e lançando um olhar especial sobre o uso do *a gente* indeterminado em superestruturas

argumentativas, que era o nosso foco, percebemos que o grau de *escolaridade* do falante favorece o uso da forma com a sua função inovadora. Além disso, constatamos também que esse uso se constitui em um recurso argumentativo, conforme consideramos ao apresentar o excerto (21), que pode elucidar a nossa abordagem.

(21) [...] principalmente a área da saúde que tá... eu acho que não é só... uma questão só de Conquista mais... eh... várias cidades do Brasil que passa por esses transtornos eh... a saúde é primordial e Conquista deixa muito a desejar né, e... essa questão também de infraestrutura eu acho que ainda tem muitos bairros ainda que sofre bastante com essa questão de asfalto... eh... esgoto né, e as vezes A GENTE vê assim a cidade um pouco mal cheirosa por conta disso então... é isso aí... fal... e assim... a questão também do lazer é muito difícil você encontra assim deveria mais... áreas verdes não tem né, se preocupa muito com a questão de shoppings construções de shoppings e falta investi na... na área vede que a única que A GENTE tem eh... tá um pouco devastada né, que é o Poço Escuro.

Diante desse resultado, parece-nos possível ratificar a hipótese de que o uso do *a gente* indeterminado exige uma maior maturidade cognitiva.

5.2.2.4 Tempo verbal

Com o controle desse grupo de fatores, esperávamos que o tempo *presente* favorecesse mais a ocorrência de sujeito indeterminado, partindo do princípio de que o *pretérito* envolvesse, em sua maioria, relatos de fatos e acontecimentos nos quais o *você* e *a gente* fossem empregados com referência específica/determinada pela necessidade de identificar o sujeito/participante desses acontecimentos, conforme discorremos ao tratarmos da variável *superestrutura textual*, ao passo que o presente além do fato de ser apresentado no momento, ainda reflete o valor de um hábito. A Tabela 08 apresenta os resultados obtidos com o controle dessa variável:

Tabela 08: *A gente* indeterminado em relação ao tempo verbal

	Total	%	P. R.
Presente	280/465	60%	.61
Pretérito	84/412	20%	.39
Futuro	9/31	29%	.25

Fonte: Elaborada pelo pesquisador.

Percebemos que, de fato, o tempo presente representou o maior número dos casos de sujeito indeterminado com 60% das ocorrências e apresentando um peso relativo de .61. O pretérito, conforme esperado, foi empregado ao lado de *a gente* indeterminado em 20% dos casos, com um peso relativo de .39. O futuro, com peso relativo de .25, representou 29% dos casos. Abaixo estão transcritos excertos que representam cada um dos tempos verbais:

(22) Oh, emocionalmente pra mim... é... eu *ach'que* pra todas as pessoas que conseguem *chegá* na universidade depois dos 30, como eu entrei com 35 anos, A GENTE tá mais seguro com que A GENTE *qué* mais seguro do que *realm...* da opção que A GENTE faz, porque muitas vezes A GENTE faz *vestibulá* quando jovem só porque *qué entrá* na universidade ou só pra *agradá* os pais no curso que A GENTE faz... e aí às vezes entra aqui e diz: “ ah, não quero isso” faz outro *vestibulá*, outro curso a mente tá muito... então, assim... quando A GENTE tá mais madura, mais consciente do que A GENTE *qué* porque A GENTE *qué fazé* aquilo mesmo... (E.S.P.)

(23) Olha a cida... a educação em Conquista ela tá um pôco deixando a desejar, né, até mesmo po... esses índices que A GENTE viu agora do IDEB que saiu *eh...* a cidade ficou em um dos últimos lugares e isso é algo preocupante porque as pessoas que entram na universidade que são o futuro da... da nação da cidade vêm dessa... das escolas então... (C.B.S.)

(24) Praticamente, antigamente A GENTE tinha brincadêra, hoje não tem. Por quê? Por... pela... pela da tecnologia de hoje em dia, as pessoas preferem mais ficá em casa, ali no facebook, whatsapp no celulé e num tem praticamente brincadêra nenhuma hoje em dia. Antigamente sim, tinham brincadêras, hoje não, hoje são muito limitadas as brincadêras e as crianças de hoje num... num têm a mesma mentalidade de antigamente. (J.L.S.)

(25) Com certeza, como eu disse anteriormente, as áreas verdes né, precisa de muita área verde aqui em Conquista [...] o parque que ficasse ali permanente né, um parque de diversões permanente não só crianças mas nós também adultos... com certeza A GENTE iria gostar. (R.F.V.)

(26) Com certeza. Hoje em dia principalmente em Conquista e ôtros lugares, assim, alguns bairros, né? Mas, todos os bairros têm violência, mas em quantidade grande e quantidade baixa. Aqui mesmo temos poucos casos, mas em ôtros bairros assim, se A GENTE fô pegá no geral, tem casos de extrema violência e por isso muitas vezes os pais preferem deixá os

filhos mais em casa. Isso eu acho que atrapalha de uma forma psicológica o... o rendimento assim das crianças, entendeu? (J.L.S.)

No excerto (22), observamos uma série de ocorrências de *a gente* indeterminado acompanhado de verbos no tempo presente dentro do mesmo contexto. Nos excertos (23) e (24), as duas ocorrências destacadas, que estão acompanhadas de verbos no pretérito, encontram-se isoladas em seus respectivos contextos assim como ocorre nos excertos (25) e (26) com verbos no tempo futuro. Os exemplos em questão são um reflexo dos dados em geral, ou seja, *a gente* indeterminado acompanhado de verbo no presente tende a ser mais produtivo em relação aos tempos pretérito e futuro, confirmando a hipótese lançada inicialmente acerca da variável *tempo verbal*.

Com o avanço na análise dos dados, observamos certa tendência na ocorrência do *a gente* indeterminado acompanhado de verbos no presente inseridos em superestrutura argumentativa. Do mesmo modo, observamos essa mesma tendência na ocorrência do *a gente* determinado em contextos de narração. Diante disso, e como já esperávamos a existência dessa relação, conforme especificamos na seção 4, realizamos uma rodada a fim de combinarmos os fatores desses dois grupos. Os resultados obtidos constam na Tabela 09:

Tabela 09: tempo verbal x superestrutura textual: combinando fatores

	Total	%	P.R
Argumentação + presente	181/211	86%	.89
Argumentação + pretérito	8/24	33%	.55
Argumentação + futuro	7/14	50%	.67
Descrição + presente	82/198	41%	.56
Descrição + pretérito	50/130	38%	.44
Descrição + futuro	2/14	14%	.18
Narração + presente	17/56	30%	.37
Narração + pretérito	26/258	10%	.15
Narração + futuro ²⁴			

Fonte: Elaborada pelo pesquisador.

A rodada complementar, cujos resultados estão expostos na Tabela 08, confirmaram a hipótese que lançamos. Observando os itens [argumentação+presente] e [narração+pretérito], primeiro e último fatores da tabela, percebemos que os casos de *a gente* indeterminado

²⁴ Essa variável foi excluída da análise por ter representado um caso de KnockOut. A única ocorrência de verbo no futuro em contexto de narração ocorreu com o *a gente* determinado.

ocorreram, em sua maioria, em contextos de argumentação e acompanhados de verbos no presente, correspondendo a 86% das 211 ocorrências e apresentando um peso relativo de .89. O menor número de *a gente* indeterminado, em contrapartida, foi registrado em superestrutura narrativa e acompanhado de verbo no pretérito. Ou seja, verbo no presente e superestrutura argumentativa são fortes favorecedores da indeterminação do sujeito por meio da forma *a gente*, se opondo a verbos no pretérito e superestrutura narrativa que são fortes condicionadores do *a gente* enquanto sujeito determinado.

O segundo peso relativo mais alto obtido com essa rodada corresponde aos casos de superestrutura argumentativa + verbo no futuro, peso de .67. No entanto, a quantidade de dados está ligada a esse fator é relativamente baixa e não permite uma análise consistente e segura. Na mesma situação, está a combinação entre [superestrutura argumentativa + verbos no pretérito] que, embora esperássemos um peso relativo baixo, demonstrou certo equilíbrio (.55), mas com um total de apenas 24 ocorrências.

De modo específico, o que esperávamos confirmar especificamente com essa rodada era justamente a existência de um alinhamento entre *superestrutura textual* e *tempo verbal*, no sentido de comprovar que há uma polarização entre [argumentação+presente] e [narração+pretérito], o que foi suficientemente comprovado pelos dados da Tabela 08.

CONCLUSÕES

A indeterminação do sujeito é um fenômeno linguístico amplo cujas dimensões não se contêm no âmbito de uma teoria, perspectiva de análise e/ou investigação. Ao finalizarmos a análise dos dados, reconhecemos que ainda há inúmeros aspectos que envolvem o fenômeno a serem explorados e descritos. Reconhecendo o fato de que, com essa análise, não encerramos as discussões suscitadas, teceremos algumas considerações acerca dos resultados que

encontramos ao promover o confronto/encontro entre as teorias e os dados. Além disso, abriremos espaço para uma retomada do texto como um todo.

Na seção 2, *Sociolinguística + Funcionalismo = Sociofuncionalismo: o que se quer da(s) teoria(s)*, apresentamos a base teórica que fundamenta a nossa análise a partir da junção de alguns dos principais pressupostos da Sociolinguística Quantitativa e do Funcionalismo Norte-americano. Em relação à Sociolinguística, tratamos, principalmente, das noções de variação e mudança e dos sentidos e objetivos de se investigar a ocorrência de um fenômeno linguístico (a indeterminação do sujeito, no nosso caso), considerando as influências de variáveis sociais. Já em relação ao Funcionalismo, discutimos as noções de língua, gramática e interação que respaldam uma análise da natureza da nossa, além de apresentar princípios que regem a mudança via gramaticalização e que são, aqui, um meio para entendermos como a indeterminação do sujeito funciona a partir de *você* e *a gente*. Além disso, colocamos em debate as razões para se estabelecer um “casamento” entre as duas teorias supracitadas, dando espaço para a orientação de pesquisa que adotamos na análise, o Sociofuncionalismo.

Na seção 3, *Indeterminação do sujeito: dos conceitos às classificações, dos dicionários às gramáticas*, demos enfoque ao nosso objeto apresentando e discutindo as definições presentes nos dicionários, os conceitos e classificações encontrados nas Gramáticas Históricas, Tradicionais e Descritivas, além de expor outras perspectivas de análise a partir dos estudos linguísticos mais recentes. Toda essa busca foi realizada com o objetivo de garantir consistência à nossa análise a partir do olhar que outros lançaram sobre o nosso objeto.

Na seção 4, *Metodologia*, fizemos esclarecimentos relacionados ao processo de análise em si. Apresentamos os *corpora*, discurremos sobre o processo de seleção da amostra e descrevemos o envelope de variação com as suas variáveis, variantes e as respectivas hipóteses, a fim de mostrar o ponto do qual partimos. Além disso, tratamos também dos procedimentos adotados para a obtenção e interpretação dos resultados.

Por fim, na seção 5, *Você e a gente em posição de sujeito no português falado de vitória da conquista: uma análise sociofuncionalista do fenômeno de indeterminação*, apresentamos e discutimos os resultados obtidos a partir da análise dos 24 inquéritos que integram os *corpora* PCVC e PPVC. Observamos, de modo geral, como *você* e *a gente* vêm sendo usado no vernáculo conquistense, dentro dos limites da amostra, e verificamos a atuação das variáveis linguísticas e extralinguísticas controladas.

Em geral, observamos, a partir da discussão presente na seção 5, que, de fato, as referências tanto de *você* quanto de *a gente* são múltiplas e, entre elas, está a referência que garante às formas a função indeterminadora: a referência genérica. Assim, a primeira

constatação foi a de que, conforme esperado, os conquistenses fazem uso das duas formas gramaticalizadas como estratégia de indeterminação do sujeito na língua em uso, assim como constataram Carvalho (2010) e Assunção (2012). Isso nos leva a crer que, mesmo não estando as duas formas alinhadas historicamente e mesmo não sendo uma substituta da outra em qualquer contexto, há este ponto de interseção: a referência genérica, que, a nosso ver, não é fruto do acaso, mas resultado do processo de mudança via gramaticalização que as conduziu e, ainda, as conduz, nessa sincronia, fazendo com que, mesmo estando em rotas diferentes, cruzem-se e compartilhem um mesmo ponto. Em suma, o fato de serem formas diferentes e não concorrentes em todos os contextos, reafirma um outro fato: o de que a indeterminação com *você* e *a gente* é resultado da gramaticalização de ambas as formas.

Observamos, além dessa constatação geral, que, se postas em comparação, o *você* indeterminado é mais produtivo que o *você* determinado, ao passo que verificamos o contrário em relação ao *a gente*. Isso se mostra um indício de que o estágio da mudança do *você*, no que se refere à função de indeterminar, pode estar mais avançado em relação ao do *a gente* e que isso lhe garantiu, na amostra analisada, maior número de ocorrências considerando a oposição sujeito indeterminado x sujeito determinado.

Em relação à atuação das variáveis linguísticas, no que se refere ao *você*, constatamos, conforme esperado, que, em uma série de formas iguais dentro de um mesmo contexto, o uso da função indeterminadora aplicado na primeira forma tende a se repetir nas formas subsequentes, fazendo, dentro desse contexto, uma cadeia de formas com referências iguais. Esse resultado, além de comprovar que a repetição rotiniza determinados usos, mostra-nos também que um contexto pode ser um terreno fértil para a ocorrência do sujeito indeterminado, ao passo que outros nem tanto, o que se confirmou, complementarmente, a partir dos resultados encontrados com o controle da variável *superestrutura textual*. Com essa superestrutura, percebemos que o *você* indeterminado é mais produtivo quando falante e ouvinte estão imersos em um contexto de argumentação, que é propício para o uso da forma com valor genérico, conforme discutimos na seção 4, e menos produtivo em contextos narrativos que, no geral, exigem referência específica.

No que se refere ao *a gente*, constatamos, a partir da variável *superestrutura textual*, que a argumentação favorece mais a ocorrência da indeterminação em oposição à narração, que favorece o sujeito determinado, assim como ocorreu com o *você*. Além disso, atrelada à variável *superestrutura textual*, a hipótese acerca da variável *tempo verbal* também foi confirmada e reforça as constatações relacionadas ao contexto de uso para o *a gente* indeterminado, ou seja, os dados da amostra comprovaram que os falantes tendem a

indeterminar mais em contextos de argumentação e a forma indeterminadora, em sua maioria, tende a ser acompanhada por verbos no presente, opondo-se à determinação que é mais comum em contexto de narração estando a forma acompanhada de verbos no pretérito.

No tocante às variáveis sociais, percebemos que o grau de escolaridade, conforme esperávamos, influencia no uso do *a gente* indeterminado e, de modo complementar, comprovamos que há relação entre a escolarização e o uso do *a gente* indeterminado em contexto de argumentação. Ou seja, falantes mais escolarizados tendem a empregar mais o *a gente* indeterminado nas superestruturas argumentativas do que falantes não escolarizados.

Em relação à *faixa etária*, ao contrário do que esperávamos, os dados dos *corpora* demonstraram que falantes com 50 anos ou mais tendem a indeterminar mais por meio do *a gente* do que falantes das faixas etárias inferiores, o que abriu caminhos para aventarmos a hipótese de que a indeterminação com o *a gente* pode exigir maior amadurecimento cognitivo, justificando, por ora, os resultados encontrados com a análise da amostra.

De modo geral, com a discussão apresentada na seção 05, pretendíamos: (i) verificar se há um alinhamento na escolha do *a gente* e *você* indeterminados em oposição ao *a gente* e *você* determinados e (ii) se os grupos de fatores selecionados para uma seriam os mesmos selecionados para a outra. Para tanto, analisamos as duas formas considerando as mesmas variáveis linguísticas e extralinguísticas e, em resultado, verificamos que foram selecionadas, pelo Programa *GoldVarb X*, variáveis diferentes para o *você* e o *a gente*, havendo um único fator comum selecionado, a *superestrutura textual*. Tal resultado comprova o fato de que as formas gramaticalizadas estão alinhadas no sentido de que compartilham a função de indeterminar, mas seguem rumos diferentes nas suas mudanças via gramaticalização. Em relação ao fator comum selecionado, *superestrutura textual*, no entanto, observamos que os percentuais e pesos relativos obtidos para ambas as formas seguem uma mesma tendência: as situações interacionais que envolvem argumentação favorecem mais a ocorrência da indeterminação com *você* e *a gente* ao passo que as situações que envolvem narração favorecem a ocorrência do sujeito determinado.

Em suma, ao findar a análise dos dados, comprovamos que (i) a função de indeterminar comum a *você* e *a gente* é inovadora sendo, portanto, empregada em contextos específicos e não tendo substituído ainda a função prototípica; (ii) a função indeterminadora do *você* e *a gente* se deve à gramaticalização das duas formas, o que se comprova pela atuação de princípios como a *descategorização semântica* e a *persistência*; e (iii) a ocorrência da indeterminação com as formas é favorecida por fatores de ordem linguística e extralinguística, conforme esperávamos.

Mesmo reconhecendo que os resultados apresentados aqui não encerram a discussão suscitada, cabe-nos ressaltar a relevância desse trabalho para entender, a partir de um recorte, o modo como a indeterminação do sujeito ocorre no Português Brasileiro. Distantes de tentar abarcar o fenômeno como um todo, considerando a sua riqueza e complexidade, mostramos, a partir de um ponto de vista que se propôs a ser inovador, que, além das possibilidades prescritas pelas gramáticas, o fenômeno tem se renovado e acompanhado o fluxo de mudança da língua.

Nessa direção, esta dissertação se mostra como uma abertura para a análise do fenômeno de indeterminação de modo mais profundo e, no geral, contribui para a descrição do Português Falado no Brasil, objetivo comum às pesquisas que estão sob o escopo da Linguística Descritiva, tomando como referência o vernáculo dos informantes cultos e populares de Vitória da Conquista-BA.

REFERÊNCIAS

- ALKMIM, Tânia Maria. Sociolinguística. In: MUSSALIN, Fernanda. BENTES, Anna Christina (Org.). *Introdução à linguística: domínios e fronteiras*. São Paulo: Cortez, 2001.
- ALMEIDA, Napoleão Mendes. *Gramática Metódica da Língua Portuguesa*. 13. ed. São Paulo: Saraiva, 1961.
- ASSUNÇÃO, Janivam da Silva. *A indeterminação do sujeito na variedade linguística de Feira de Santana: um estudo variacionista*. Dissertação (Mestrado em Estudos Linguísticos). Universidade Estadual de Feira de Santana, Feira de Santana, 2012.
- AZEREDO, José Carlos de. *Gramática Houaiss da Língua Portuguesa*. São Paulo: Publifolha, 2008.
- BECHARA, Evanildo. *Gramática escolar da língua portuguesa*. Rio de Janeiro: Lucerna, 2006.
- BECHARA, Evanildo. *Moderna gramática portuguesa*. 37. ed. Rio de Janeiro: Nova Fronteira, 2009.
- BLUTEAU, Raphael. *Vocabulário portuguez & latino: aulico, anatomico, architectonico*. Coimbra: Collegio das Artes da Companhia de Jesus, 1712 -1728. 8 v. Versão digital disponível em: <<http://www.brasiliana.usp.br/en/dicionario/2/meio>>. Acesso em: julho. 2017.
- BUENO, Francisco da Silveira. *Gramática normativa da língua portuguesa*. São Paulo: Saraiva, 1944.
- CALVET. *Sociolinguística: uma introdução crítica*. São Paulo: Parábola, 2002.
- CAMACHO, Roberto Gomes. Sociolinguística. In: *Introdução à linguística: domínios e fronteiras*. MUSSALIN, Fernanda. BENTES, Anna Christina (orgs.). São Paulo: Cortez, 2001.
- CAMARA JUNIOR, Joaquim Mattoso. *Dicionário de Linguística e Gramática*. 29. ed. Petrópolis: Editora Vozes, 1997.
- CARVALHO, Valter de. *Você, a gente et alia indeterminam o sujeito em Salvador*. Dissertação (Mestrado em Letras) Universidade do Estado da Bahia. Salvador, 2010.

- CASTILHO, Ataliba de. *Nova gramática do português contemporâneo*. 2. ed. São Paulo: Contexto, 2010.
- CASTILHO, Ataliba de. *Nova gramática do português contemporâneo*. São Paulo: Contexto, 2016.
- CEZÁRIO, Maria Maura; VOTRE, Sebastião. Sociolinguística. In.: MARTELOTTA, Mario Eduardo (org.). *Manual de Linguística*. 2. ed. São Paulo: Contexto, 2011.
- CIPRO NETO, Pasquale; INFANTE, Ulisses. *Gramática da língua portuguesa*. São Paulo: Scipione, 2003.
- CINTRA, Luís F. Lindley. *Sobre “formas de tratamento” na língua portuguesa*. 2. ed. Lisboa: Livros Horizonte, 1986.
- CIPRO NETO; Pasquale; INFANTE, Ulisses. *Gramática da Língua Portuguesa*. São Paulo: Scipione, 1998.
- COELHO, Izete Lehmkuhl et. al. *Sociolinguística*. Florianópolis: LLV/CCE/UFSC, 2010.
- COSTA, Marcos Antonio. Estruturalismo. In. MARTELOTTA, Mario Eduardo (org.). *Manual de Linguística*. 2. ed. São Paulo: Contexto, 2011.
- CRUZ, Pe. Antônio da. *Prontuário de Análise Gramatical e Lógica*. 2. ed. São Paulo: Editora Vozes, 1948.
- CRYSTAL, David. *Dicionário de linguística e fonética*. 2. ed. Rio de Janeiro, Jorge Zahar Editora, 2000 [1941].
- CUNHA, Celso; CINTRA, Lindley. *Nova gramática do português contemporâneo*. 2. ed. Rio de Janeiro: Nova Fronteira, 1985
- CUNHA, Celso; CINTRA Lindley. *Nova gramática do português contemporâneo*. 3. ed. Rio de Janeiro: Nova Fronteira, 2001.
- CUNHA, Celso; CINTRA, Lindley. *Nova gramática do português contemporâneo*. 5 ed. Rio de Janeiro: Lexikon, 2008.
- DUBOIS, Jean. *Dicionário de Linguística*. São Paulo: Cultrix, 1973.
- FERRAREZI JUNIOR, Celso. *Gramática do brasileiro: uma nova forma de entender a nossa língua*. São Paulo: Globo, 2008.
- FURTADO DA CUNHA, Maria Angélica. Funcionalismo. In. MARTELOTTA, Mario Eduardo (org.). *Manual de Linguística*. 2. ed. São Paulo: Contexto, 2011.
- FURTADO DA CUNHA, Maria Angélica; CEZARIO, Maria Moura; COSTA, Marcos Antonio;. *Pressupostos teóricos fundamentais*. In: CUNHA, Maria Angélica Furtado da; MARTELOTTA, Mário Eduardo; OLIVEIRA, Mariangela Rios de. *Linguística Funcional: Teoria e prática*. 1. ed. São Paulo: Parábola Editoria, 2015.
- GIVÓN, Talmy. *Functionalism and grammar*. Amsterdam/Philadelphia: John Benjamins, 1995.
- GONÇALVES, Sebastião Carlos Leite; CARVALHO, Cristina dos Santos. Critérios de gramaticalização. In: GONÇALVES, Sebastião Carlos Leite; HERNANDES, Maria Célia; CASSEB-GALVÃO, Vânia Cristina (orgs.). *Introdução à gramaticalização: princípios teóricos & aplicação*. São Paulo: Parábola Editorial, 2007. (p. 67-90)
- HOPPER, Paul. *Emergent grammar*. In: *BLS*. v. 13, p. 139-157, 1987.
- HOPPER, Paul; TRAUGOTT, Elizabeth. *Grammaticalization*. Cambridge: CUP, 1993.
- KOCH, Mário Vilela Ingedore Villaça. *Gramática da Língua Portuguesa*. Coimbra, Livraria Almeida, 2001.
- LABOV, William. *Padrões sociolinguísticos*. Trad. de M. Bagno; M. M. P. Scherre; C. R. Cardoso. São Paulo: Parábola Editorial, 2008 [1972].
- LEHMANN, C. *Thoughts on grammaticalization. A programmatic sketch*. Köln: Arbeiten des Kölner universalien- Projekts, 1982.

- LOPES, Célia Regina dos Santos. *A gramaticalização de a gente em português em tempo real de longa e de curta duração: retenção e mudança na especificação dos traços intrínsecos*. Fórum Linguístico, Florianópolis, v. 4, n. 1, p. 47-80, 2004.
- LYONS, John. *Linguagem e Linguística: uma introdução*. Rio de Janeiro: Zahar Editores, 1982.
- MARCUSCHI, Luiz Antonio. Gêneros textuais: configuração, dinamização e circulação. In: KA; RWOSKI, Acir Mário; GAYDECZKA, Beatriz; BRITO, karin Siebeneicher (Orgs.) *Gêneros textuais: reflexão e ensino*. 2. ed. Rio de Janeiro: Lucerna, 2006.
- MATTOS E SILVA, Rosa Virgínia. *Caminhos da linguística histórica*. São Paulo, Parábola Editorial, 2008.
- MATTOS E SILVA, Rosa Virgínia. *Estruturas Trecentistas: Elementos para uma gramática do Português Arcaico*. São Paulo, Imprensa Nacional, 1989.
- MEILLET, Antoine. *Linguistique historique et linguistique générale*. Paris : Librairie Ancienne Honoré Champion, 1965.
- NASCENTES, A. *O tratamento de VOCÊ no Brasil*. Letras. Curitiba, 1956.
- NEVES, Maria Helena de Moura. *A gramática funcional*. 2. ed. São Paulo: Martins Fontes, 1997.
- NEVES, Maria Helena de Moura. *Gramática de usos do português*. São Paulo: Editora UNESP, 2000.
- NEVES, Maria Helena de Moura. *Guia de uso do português*. São Paulo: Editora UNESP, 2012.
- NEVES, Maria Helena de Moura. Os pronomes. In: ILARI, Rodlfo (org). *Gramática do português culto falado no Brasil: palavras de classe fechada*. São Paulo: Contexto, 2015.
- NEVES, Maria Helena de Moura. Estudos Funcionalistas no Brasil. *D.E.L.T.A.* Vol. 15. Nº Especial, 1999.
- PEREIRA, Eduardo Carlos. *Gramática expositiva curso superior*. 65. ed. São Paul: Companhia Editora Nacional, 1945.
- PERINI, M. A. *Para uma nova gramática do português*. São Paulo: Ática, 1985
- PERINI, Mário Alberto. *Gramática do português brasileiro*. São Paulo: Parábola Editorial, 2010.
- PINTO, Luiz Maria da Silva. *Diccionario da língua brasileira*. Ouro Preto: Typographia de Silva, 1832.
- PONTES, Eunice Souza Lima. *Sujeito: da sintaxe ao discurso*. São Paulo: Ática, 1986.
- ROCHA LIMA, Carlos Henrique. *Gramática normativa da língua portuguesa*. 7. ed. Rio de Janeiro: F. Briguiet & CIA, 1962.
- ROCHA LIMA, Carlos Henrique. *Gramática normativa da língua portuguesa*. 35. ed. Rio de Janeiro: José Olympio, 1998.
- ROCHA LIMA, Carlos Henrique da. *Gramática Normativa da Língua Portuguesa*. São Paulo: José Olmpio, 2010.
- ROCHA, Warley José Campos. *Você e cê: um estudo sociofuncional em uma comunidade do Sertão da Ressaca*. Dissertação (Mestrado em Linguística). Universidade Estadual do Sudoeste da Bahia. Vitória da Conquista, 2017.
- SACONNI, Luís Antônio. *Nossa gramática: teoria e prática*. São Paulo: Atual, 1982.
- SAID ALI, Manoel. *Gramática histórica da língua portuguesa*. 7. ed. Rio de Janeiro: Livraria Acadêmica, 1971.
- SAID ALI, Manoel. *Gramática secundária e Gramática histórica da língua portuguesa*. 3. ed. Brasília: Editora Universidade de Brasília, 1964.
- SANTOS LUZ, Marilina dos. Formas de tratamento. *Revista Portuguesa da Filologia*, Lisboa, v. VII, tomos I e II, p. 251-363, 1958.
- SAUSSURE, Ferdinand. *Curso de Linguística Geral*. 27. ed. São Paulo: Editora Cultrix, 2006 [1916].

- SOUSA, Valéria Viana. *Os (des)caminhos do você: uma análise sobre a variação e mudança na forma, na função e na referência do pronome você*. 2008. Tese (Doutorado) - UFPB, João Pessoa, 2008.
- TAVARES, M. A. *A gramaticalização de E, AÍ, DAÍ e ENTÃO: estratificação/variação e mudança no domínio funcional da sequenciação retroativo-propulsora de informações – um estudo sociofuncionalista*. Tese de Doutorado em Linguística – UFSC: Florianópolis, 2003.
- TAVARES, Maria Alice. Sociofuncionalismo: um duplo olhar sobre a variação e a mudança linguística. *Interdisciplinar-Revista de Estudos em Língua e Literatura*, v. 17, p. 27-48, 2013.
- TEIXEIRA, Maria Luiza de Sousa. *A indeterminação pragmática e semântica do sujeito*. Dissertação (Mestrado em Linguística) Universidade Estadual Paulista. São José do Rio Preto, 2014.
- TRAUGOTT, Elizabeth Closs. *Gramaticalização: uma entrevista com Elizabeth Closs Traugott*. *ReVEL*, vol. 12, n. 22, 2014. Traduzido por Gabriel de Ávila Othero e Ana Carolina Spinelli. [www.revel.inf.br].
- VAN DIJK, Teun A. *Cognição, discurso e interação*. 3. ed. São Paulo: Contexto, 2000.
- WEINREICH, Uriel; LABOV, William; HERZOG, Marvin. Empirical foundations a theory of language change. In: LEHMANN, W.; Y. (Eds.). *Directions for historical linguistics*. Austin: University of Texas Press, 1968.
- WILSON, Victoria; MARTELOTTA, Mário Eduardo. Arbitrariedade e iconicidade. In: MARTELOTTA et al. (orgs.). *Manual de linguística*. São Paulo: Contexto, 2008. (p. 71-85)

ANEXOS

Anexo 1 – Roteiro de perguntas para entrevista



Universidade Estadual do Sudoeste da Bahia – UESB
 Departamento de Estudos Linguísticos e Literários – DELL
 Grupo de Pesquisa em Sociofuncionalismo
 &
 Grupo de Pesquisa em Linguística Histórica e Aplicada

ROTEIRO DE PERGUNTAS PARA A ENTREVISTA

Abaixo, seguem sugestões de perguntas ,topicalizadas por tema, que podem ser feitas na entrevista com o informante selecionado previamente através do Perfil Social.

Infância:

1. Como foi a sua infância?
2. Você se lembra de alguma história interessante que tenha acontecido com você na sua infância? Ou você se lembra de alguma história interessante na sua infância com você, com sua família ou com alguém conhecido?
3. Como eram as brincadeiras na sua infância?
4. O que você acha das brincadeiras de sua época em relação às brincadeiras de hoje?
5. Na sua opinião, as crianças sempre se divertem independentemente do tipo de brinquedo, sempre se adaptam? Ou as crianças eram mais felizes antes que hoje ou o contrário disso são mais felizes hoje, porque têm mais recursos, do que antes?
6. No Brasil e na própria região onde moramos, há crianças que trabalham desde cedo. Você conhece alguém nessa situação?
7. O que você acha disso?

Localidade - bairro/ rua

1. Você sempre morou nesse bairro?
2. Você acha bom morar aqui? Por quê?
3. Você tem vontade de morar em outro lugar? Qual? Por quê?
4. Como é que é morar nesse bairro? É movimentado ou tranquilo? Você preferiria que fosse como?
5. E a vizinhança? Como ela é?

Localidade – cidade

1. O que você acha de morar em Vitória da Conquista? Por quê?
2. Você falou que gosta (ou que não gosta) de Micareta/ Carnaval. O que você costuma fazer nesse período?
3. Você falou que gosta (ou que não gosta) de São João. O que você costuma fazer nesse período?
4. Você tem vontade de morar em outro lugar? Qual? Por quê?
5. O que você acha do clima daqui? Você gosta ou não? Por quê?
6. E os conquistenses? O que você acha das pessoas que moram aqui?
7. Quando você viaja e passa um tempo fora, quando volta tem saudades de quê? Por quê?

Profissão:

1. Você trabalha em quê?
2. O que você faz no seu trabalho? Conte a sua rotina, um dia de trabalho.
3. É essa sua profissão?
4. Você tem vontade de trabalhar em outra coisa ou em outro lugar? Por quê?

5. Como foi a experiência para você do primeiro emprego?
6. Se fosse para você escolher hoje uma profissão, qual você escolheria ? Por quê?
7. Você se lembra de algum fato interessante ocorrido no seu trabalho?

Escola:

1. Onde você estuda? Tem quanto tempo que estuda lá?
2. Quais foram os motivos que impediram/ que dificultaram que você estudasse?
3. Você gosta da escola? Por quê?
4. Você teve vontade de estudar?
5. Você acha o estudo interessante e importante? Por quê?
6. O que você acha da educação em Conquista?
7. Você vê diferença na educação de hoje e na educação de antes? Em que são diferentes?
8. Você se lembra de alguma história interessante que aconteceu na escola?
9. Qual é a disciplina que você mais gosta? Por quê?
10. Se você pudesse retirar uma disciplina da escola, qual você retiraria? Por quê?
11. Você pensa em fazer vestibular? Para quê?
12. Quais os motivos que contribuíram para você escolher esse curso?
13. O que você mais gosta (ou mais gostava) na escola?

Família:

1. Como é a sua família? Você tem quantos irmãos?
2. Como é que são seus pais? Fale um pouco sobre eles.
3. Como é que são seus irmãos? Onde moram? O que fazem?
4. Vocês passavam muito tempo juntos na infância? O que vocês faziam juntos?
5. Lembra de alguma história interessante vivida com a sua família na infância?
6. Lembra de alguma história interessante vivida com a sua família?
7. Lembra de alguma viagem? Conte.
8. Você tem filhos? Como é o seu relacionamento com eles?
9. O que você acha que faz por eles que seus pais não fizeram com você?
10. Você é casado (a), tem namorado (a)? Como vocês se conheceram?

Lazer:

1. O que você costuma fazer em Conquista nos finais de semana? Você costuma sair? Para onde?
2. O que costuma fazer nesse lugar?
3. Você disse que gosta de cinema/ novela. Tem algum (a) filme/novela em especial que tenha marcado você? Que você goste mais? Qual?
4. Conte um pouco a história dele (a).
5. O que você acha das opções de lazer em Conquista?
6. O que você acha que poderia melhorar?
7. Qual é a sua diversão preferida?
8. Qual é o estilo de música que você mais gosta? Por quê?
9. O que você acha dessas músicas atuais?
10. Qual é a sua religião? Fale um pouco a respeito dela.
11. Com relação à religiosidade, a sua família também pensa como você?
12. Costuma ler livros? Lembra de algum que tenha lido? Conte a história.

Pessoais:

1. Estudar e trabalhar para você são difíceis de conciliar? Por quê?
2. O que você acha da sua forma de falar? Por quê?
3. Você mudaria alguma coisa no seu jeito de falar?
4. Você acha que todos os brasileiros falam da mesma forma? Por quê?
5. Você conhece alguém que fala diferente de você? Como é essa diferença?

6. Você já teve alguma doença mais séria? Qual foi?
7. Você já esteve diante de um evento de morte de uma pessoa querida? Como foi?
8. Como você se sentiu?
9. Você já fez algo que se arrependeu depois? Conte.
10. O que mais magoa você?
11. Você tem algum sonho? Conte.
12. Se tivesse um cargo tipo presidente da república o que faria de imediato? Por quê?
13. O que você gostaria de ver publicado na manchete de um jornal?
14. Você gosta de novelas? Quantas costuma assistir diariamente? Qual é a que mais gosta na atualidade ? Por quê?
15. Ao ler revistas, o que procura nas mesmas? Qual é o seu maior interesse nas revistas?
16. E futebol? Qual é o seu time? Gosta de assistir aos jogos pela televisão ou rádios? Por quê?
17. Costuma ir a estádios?
18. Como é assistir a um jogo em um estádio?
19. Costuma viajar nas férias?
20. Há algum lugar específico que sempre vai ou escolhe lugares diferentes?
21. O que você procura fazer quando viaja? Geralmente, qual é o motivo que o faz viajar: férias, ver amigos, visitar parentes, trabalhar, participar de eventos?
22. E esse São João? O que você fez nesse São João?
23. Tem planos para as próximas férias? Quais?

Anexo 2 – Ficha social do informante



Universidade Estadual do Sudoeste da Bahia – UESB
 Departamento de Estudos Linguísticos e Literários – DELL
 Grupo de Pesquisa em Sociofuncionalismo
 &
 Grupo de Pesquisa em Linguística Histórica

FICHA SOCIAL DO INFORMANTE

Número: _____ Data : ____/____/20____
 Entrevistador: _____ Local / Bairro : _____

1. Nome: _____

2. Endereço: _____

3. Telefone(s) para contato: _____

4. Data de nascimento: _____

5. Naturalidade: _____

6. Oriundo da : () Zona urbana () Zona rural

7. Há quanto tempo você mora nesse bairro?

8. Qual é a naturalidade dos seus pais?

PAI: _____

MÃE: _____

9.Há quanto tempo seus pais moram em Vitória da Conquista?

10. Você estuda?

() sim

() não

() nunca estudou

11. Qual a série? _____

12. Até que série estudou? _____

13. Por que não continuou os estudos? _____

14. Você trabalha? () sim () não

15.Em quê? _____

16.É essa a sua profissão? () sim () não

17. Você tem uma outra profissão? () sim () não

18.Qual é a sua profissão? _____

16. Você se sustenta sozinho(a)? () sim () não

17.Você recebe ajuda financeira de quem?

() família

() outros (quem? _____)

18.Qual a sua renda mensal aproximada (ou renda familiar)

Renda individual : _____

Renda familiar : _____

19. Além de você, quantas pessoas moram em casa?

20. Qual é a relação de parentesco que há entre vocês?

() parente (s) (indicar) : _____

() não parente (indicar): _____

21. Você costuma ver TV? () sim () não

22. Qual (is) programa(s) assistidos por você?

() novela Quais ? _____

() notícias Quais? _____

() esportes Quais? _____

() outros Quais? _____

23. Você acompanha alguma novela ? Qual (quais)? _____

23. Você costuma ouvir rádio? () sim () não

24. Em que horário você ouve? _____

25. Qual(is) é/são o(s) programa(s) ouvido(s) por você? _____

26. Em média , quanto tempo do seu dia você passa:

a) assistindo TV _____

b) ouvindo rádio _____

26. Você lê jornal?

() sim, diariamente () não () de vez em quando

27. Qual (is) jornal (is)? _____

28. Quais são as partes do jornal que você mais tem interesse?

29. Você gosta de ler revistas? () sim () não

30. Qual (is) revista (s)? _____

31. Você costuma usar internet? () sim () não

32. Onde? _____

33. Quanto tempo você costuma usar? _____

34. O que você costuma acessar (ver) na internet?

31. Você costuma ir ao cinema?

☐ sempre

☐ não

☐ de vez em quando

32. Qual tipo de filme você prefere?

☐ romance

☐ comédia

☐ drama

☐ suspense

☐ ação

☐ outros

33. Você se lembra de algum filme em especial?

33. Qual a sua diversão favorita? _____

34. Você gosta de micareta/carnaval? ☐ sim ☐ não

35. Você gosta de São João? ☐ sim ☐ não

36. Você gosta do Natal ☐ sim ☐ não

37. Você gosta de futebol? ☐ sim ☐ não

38. Qual time? _____

39. Você tem alguma religião? ☐ sim ☐ não

40. Qual é a sua religião? _____

41. Você é uma pessoa que

☐ nunca sai de Vitória da Conquista

☐ só sai a negócio

☐ sempre sai para passear

42. Passa muito tempo fora da cidade?

☐ menos de um mês

☐ mais de um mês (especificar)

Atitude: Receptivo/Extrovertido ☐

Refratário/Introvertido ☐

A IMPORTÂNCIA DESSA FICHA É:

- CONTRIBUIR PARA A SELEÇÃO DOS INFORMANTES DESEJADOS NA PESQUISA;
- CONTRIBUIR PARA A ELABORAÇÃO DE PERGUNTAS PARA A ENTREVISTA.