



**UNIVERSIDADE FEDERAL DE SERGIPE
PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS
DOUTORADO EM LETRAS**

ANDRÉIA SILVA ARAUJO

O USO VARIÁVEL DOS PRONOMES TU, VOCÊ E CÊ NA FUNÇÃO DE SUJEITO:
um estudo do padrão de comportamento referencial

SÃO CRISTÓVÃO/SE

2022

ANDRÉIA SILVA ARAUJO

O USO VARIÁVEL DOS PRONOMES TU, VOCÊ E CÊ NA FUNÇÃO DE SUJEITO:
um estudo do padrão de comportamento referencial

Tese apresentada ao Programa de Pós-Graduação em Letras da Universidade Federal de Sergipe, como requisito parcial à obtenção do título de Doutora em Letras.

Área de concentração: Estudos Linguísticos.

Linha de Pesquisa: Descrição, Análise e Usos Linguísticos.

Orientadora: Prof.^a Dra. Raquel Meister Ko. Freitag

SÃO CRISTÓVÃO/SE

2022

**FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL
UNIVERSIDADE FEDERAL DE SERGIPE**

A663u Araújo, Andréia Silva.
O uso variável dos pronomes tu, você e cê na função de sujeito:
um estudo do padrão de comportamento referencial / Andréia Silva
Araújo; orientadora Raquel Meister Ko. Freitag. – São Cristóvão,
SE, 2022.
198f.: il.

Tese (doutorado em Letras) – Universidade Federal de Sergipe,
2022.

1. Língua portuguesa. 2. Linguística. 3. Sociolinguística I. Freitag,
Raquel Meister Ko., orient. II. Título.

CDU 81'367.626

ANDRÉIA SILVA ARAUJO

O USO VARIÁVEL DOS PRONOMES TU, VOCÊ E CÊ NA FUNÇÃO DE SUJEITO:
um estudo do padrão de comportamento referencial

Tese apresentada ao Programa de Pós-Graduação em Letras da Universidade Federal de Sergipe, como requisito parcial à obtenção do título de Doutora em Letras.

Área de concentração: Estudos Linguísticos.

Linha de Pesquisa: Descrição, Análise e Usos Linguísticos.

Aprovada em: 23/05/2022.

BANCA EXAMINADORA:

Prof.^a Dra. Raquel Meister Ko Freitag
Universidade Federal de Sergipe
Presidente (Orientadora)

Prof.^a Dra. Valéria Viana Sousa
Universidade Estadual do Sudoeste da Bahia
Examinadora (externa à instituição)

Prof. Dr. Valter de Carvalho Dias
Instituto Federal da Bahia
Examinador (externo à instituição)

Prof.^a Dra. Leilane Ramos Da Silva
Universidade Federal de Sergipe
Examinadora (externa ao programa)

Prof.^a Dra. Isabel Cristina Michelin de Azevedo
Universidade Federal de Sergipe
Examinadora (interna ao programa)

Aos meus pais, **Pedro Paes de Araujo** e **Luzia de Jesus Silva Araujo**, por serem a razão do meu existir e do meu insistir nesta jornada, dedico.

AGRADECIMENTOS

A Deus, por ter me dado força para superar os obstáculos que surgiram ao longo desta jornada.

À minha orientadora, Profa. Dra. Raquel Meister Ko. Freitag, pela orientação segura, pela dedicação, pela compreensão, pelos conselhos nos momentos mais difíceis, e, sobretudo, pela formação como pesquisadora propiciada durante uma trajetória de aproximadamente 15 anos entre graduação e doutorado.

À Prof.^a Dra. Leilane Ramos Da Silva, pelos ensinamentos, pelos incentivos e por ter aceitado participar de mais uma etapa importante do meu percurso acadêmico.

Aos membros da banca examinadora de defesa, Prof.^a Dra. Isabel Cristina Michelin de Azevedo, Prof.^a Dra. Leilane Ramos Da Silva, Prof.^a Dra. Valéria Viana Sousa e Prof. Dr. Valter de Carvalho Dias, pela leitura atenta e pelas significativas sugestões que contribuíram para o aprimoramento do presente estudo.

À minha amiga Josilene de Jesus Mendonça, pela parceria, pelos momentos especiais que passamos durante essa árdua caminhada, pelos conselhos e pelos incentivos.

À Profa. Dra. Marta Scherre, pela disponibilização de material bibliográfico.

Aos professores do Programa de Pós-Graduação em Letras, pelos ensinamentos durante o percurso.

Aos informantes da amostra Deslocamentos-UFS/ITA (2018), pela participação nas entrevistas realizadas.

À CAPES, pelo subsídio financeiro.

Aos membros do Gelins, pelos momentos compartilhados.

Aos meus pais, Pedro Paes de Araujo e Luzia de Jesus Silva Araujo, pelo carinho, pelo amor incondicional, por tudo que fizeram e fazem por mim, por sempre acreditarem no meu potencial e por serem meus incentivadores e meus exemplos de perseverança.

Às minhas irmãs, Leidi Araujo e Cleidi Araujo, pelo carinho, pela torcida, pela preocupação e pelo incentivo constantes.

À minha sobrinha, Maria Eduarda Araujo, e aos meus sobrinhos, Gabriel Barreto e Gustavo Barreto, pelo carinho, pelo amor e pelos momentos felizes partilhados.

A todos os meus familiares e amigos, que ficaram felizes com a concretização desta etapa profissional.

RESUMO

As formas variantes **tu**, **você** e **cê** são utilizadas no português brasileiro, tanto para se reportar a um referente determinado (interlocutor dêitico) quanto a um referente indeterminado (a qualquer pessoa presente ou não no contexto comunicativo, incluindo o próprio locutor/falante). Os estudos variacionistas sobre esse fenômeno analisam o tipo de referência como uma variável independente. No entanto, cada tipo de referência parece ter um comportamento de uma variável dependente, caracterizando-se como duas funções/regras de uso distintas para o fenômeno em questão, a saber: i) a referência determinada das formas variantes expressa a função de referência à segunda pessoa do singular (2ªPS); e ii) a referência indeterminada das formas variantes expressa a função de indeterminação do sujeito. Defendemos a tese de que os usos variáveis dos pronomes **tu**, **você** e **cê** estão relacionados à distinção do tipo de função referencial (determinada ou indeterminada), a qual interfere na proporção de distribuição das formas, sendo necessário que cada tipo de função referencial seja controlado como uma regra variável separada (dependente). A partir dos postulados teóricos da Sociolinguística Variacionista (WEINREICH; LABOV; HERZOG, 2006 [1968]; LABOV, 2008 [1972]) em interface com a Pragmática (BROWN; GILMAN, 1980; BROWN; LEVINSON, 2011 [1987]) e com a Teoria da Acomodação (BELL, 1984; GILES; COUPLAND; COUPLAND, 1991), objetivamos analisar o padrão de comportamento das variantes **tu**, **você** e **cê** com referência determinada e indeterminada no falar de universitários, em duas situações de coleta de dados: em entrevistas sociolinguísticas e em interações. Selecionamos como *corpus* de análise 16 entrevistas sociolinguísticas (amostra Deslocamentos-UFS/ITA (2018)) e 16 interações conduzidas (amostra Rede Social de Informantes Universitários de Itabaiana/SE (2013)). Para cada uma das regras variáveis em estudo, controlamos os efeitos das seguintes variáveis independentes: i) grau de intimidade entre os interlocutores; ii) relações (as)simétricas quanto ao sexo/gênero dos interlocutores; iii) turno de fala; iv) pergunta proposital; v) sexo/gênero do falante; vi) deslocamento social; vii) tipo de amostra; viii) tipo de sequência discursiva; ix) tópico discursivo; x) tipo de discurso; xi) paralelismo formal e xii) efeito gatilho. Identificamos 3205 ocorrências dos pronomes sujeitos **tu**, **você** e **cê** no *corpus* analisado. Os resultados gerais demonstraram que a variante **você** é a forma mais utilizada pelos informantes, tanto na função de referência determinada quanto indeterminada nas duas amostras analisadas. O uso frequente dessa variante foi motivado, principalmente, pelo grau de escolaridade dos informantes e pelo ambiente focalizado (uma universidade), os quais favorecem a atuação da dimensão de poder que propicia o uso da forma de maior prestígio na comunidade dentre as três formas variantes. Os usos da variante **tu** ocorreram apenas na regra de referência determinada em interações conduzidas (na fala de homens, em sequências injuntivas, em tópicos sobre vivências pessoais, em situações interacionais entre íntimos). Tais resultados apresentaram significância estatística, indicando que o tipo de coleta interfere na distribuição das variantes e que há a atuação de duas regras variáveis (de referência determinada e de referência indeterminada) nos usos das formas variáveis estudadas. Quando o tipo de amostra é de entrevistas, os usos das formas pronominais **você** e **cê** ocorreram indiferentemente do tipo de referência. Estes resultados contribuem para o panorama das mudanças no paradigma pronominal do português brasileiro e sugerem a importância da diversificação de tipos de amostras linguísticas em análises de processos variáveis pragmaticamente motivados.

PALAVRAS-CHAVE: formas pronominais de 2ªPS; função referencial; variação linguística.

ABSTRACT

The variant forms *tu* (you), *você*(you), and *cê*(you) are used in Brazilian Portuguese to report to both a determined referent (deictic interlocutor) and indeterminate referent (any person present or not in the communicative context, including the own locutor/speaker). The variationist studies about this phenomenon analyzed the kind of reference as an independent variable. However, each reference seems to have a behavior of a dependent variable being characterized as two distinct functions/rules of use for the phenomenon under study: i) the determined reference of the variant forms expresses the function of reference to the second person singular (2ndPS); and ii) the indeterminate reference of the variant forms expresses the function of indeterminacy of the subject. We argue for the thesis that the variable uses of the pronouns **tu**, **você**, and **cê** are related to the distinction of the kind of referential function (defined or indeterminate), which interferes with the proportion of the distribution of the forms, being necessary that each referential function be controlled as a separate variable rule (dependent). From the theoretical framework of the interface among the Variationist Sociolinguistics (WEINREICH; LABOV; HERZOG, 2006 [1968]; LABOV, 2008 [1972]), Pragmatics (BROWN, GILMAN, 1980; BROWN, LEVINSON, 2011 [1987]), and Accommodation theory (BELL, 1984, 2001; GILES, COUPLAND, COUPLAND, 1991), we aim at analyzing the behavioral pattern of the variants **tu**, **você**, and **cê** with a defined and an indeterminate reference in the speech of university students, in two situations of data collection: in sociolinguistic interviews and in interactions. We selected as the analyzed *corpus* 16 sociolinguistic interviews (Sample Displacements-UFS/ITA (2018)) and 16 interactions (sample Social Network of university informants from Itabaiana/SE (2013)). For each one of the variable rules under study, we controlled the effects of the following independent variables: i) level of intimacy among the interlocutors; ii) (a)symmetrical relationships in relation to interlocutors' gender; iii) speech turn; iv) deliberate question; v) speaker's gender; iv) social mobility; vii) kind of sample; viii) kind of discursive sequence; ix) discursive topic; x) kind of speech; xi) formal parallelism; and xii) triggering effect. We identified 3209 occurrences of the subject pronouns **tu**, **você**, and **cê** in the analyzed *corpus*. Overall, results revealed that the variant *você* is the most used variant by the informants, both in the functions of defined and indeterminate references in both samples. The frequent use of that variant was mostly triggered by the informants' level of schooling and the focused environment (a university), which favor the action of the power dimension enabling the use of the most prestigious form among the three variants in the community. The uses of the variant **tu** only occurred in the rule of defined reference in interactions (in men's speech, injunctive sequences, in topics about personal experiences, in interactional situations among intimate people). Such results presented statistical significance, indicating that the kind of sample interferes with the distribution of the variants and that there is action of two variable rules (defined and indeterminate references) in the uses of the variable forms used. When the sample is composed of interviews, the uses of the pronominal forms **você** and **cê** occurred regardless of the kind of reference. These results contribute to the overview of the changes in Brazilian Portuguese pronominal chart and suggest the importance of the diversification of kinds of linguist samples in analysis of variable processes pragmatically motivated.

KEYWORDS: Pronominal forms of 2ndPS; referential function; linguistic variation.

RESUMEN

Las formas variantes **tu**, **você** y **cê** son utilizadas en el portugués brasileño para reportarse tanto a un referente determinando (interlocutor deíctico) cuanto a un referente indeterminado (a cualquier persona presente o no en un contexto comunicativo, incluyendo el propio locutor/hablante). Las investigaciones variacionistas desarrolladas sobre ese fenómeno analizan el tipo de referencia como una variable independiente. Sin embargo, cada clase de referencia aparenta tener un comportamiento de una variable dependiente, y se caracteriza como dos funciones/reglas de uso distintas para el fenómeno en cuestión, es decir: i) la referencia determinada de las formas variantes expresa la función de referencia a la segunda persona del singular (2ªPS); y ii) la referencia indeterminada de las formas variantes expresa la función de indeterminación del sujeto. Defendemos la tesis de que los usos variables de los pronombres **tu**, **você** y **cê** están relacionados a la distinción del tipo de función referencial (determinada o indeterminada), a la cual interfiere en la proporción de distribución de las formas; lo que hace necesario que cada tipo de función referencial sea controlado con una regla variable separada (dependiente). De acuerdo con los postulados teóricos de la Sociolingüística Variacionista (WEINREICH; LABOV; HERZOG, 2006 [1968]; LABOV, 2008 [1972]) en interfaz con la Pragmática (BROWN; GILMAN, 1980; BROWN; LEVINSON, 2011 [1987]) y con la Teoría de la Acomodación (BELL, 1984; GILES; COUPLAND; COUPLAND, 1991), objetivamos analizar un padrón de comportamiento de las variantes **tu**, **você** y **cê** con referencia determinada e indeterminada en el hablar de universitarios, en dos situaciones de coleta de datos: en entrevistas sociolingüísticas y en interacciones. Seleccionamos como *corpus* de análisis 16 entrevistas sociolingüísticas (muestra Deslocamientos - UFS/ITA (2018)) y 16 interacciones conducidas (muestra Rede Social de Informantes Universitarios de Itabaiana /SE (2013)). Para cada una de las reglas variables investigadas, controlamos los efectos de las siguientes variables independientes: i) nivel de cercanía entre los interlocutores; ii) relaciones (a)simétricas cuanto al sexo de los interlocutores; iii) turno de habla; iv) pregunta intencionada; v) sexo del hablante; vi) deslocamiento social; vii) tipo de muestra; viii) tipo de secuencia discursiva; ix) tópico discursivo; x) tipo de discurso; xi) paralelismo formal y xii) efecto gatillo. Identificamos 3205 ocurrencias de pronombres sujetos **tu**, **você** y **cê** en el *corpus* analizado. Los resultados generales enseñaron que la variante **você** es la forma más utilizada por los informantes, tanto en la función de referencia determinada cuanto indeterminada en las dos muestras analizadas. El uso frecuente de esta variante fue motivado, principalmente, debido al grado de escolaridad de los informantes y por el ambiente focalizado (una universidad), los cuales favorecen la actuación de la dimensión de poder que propicia el uso de la forma de mayor prestigio en la comunidad en medio de las tres formas variantes. Los usos de la variante **tu** ocurrieron solamente en la regla de referencia determinada en interacciones conducidas (en el habla de hombres, en secuencias imperativas, en tópicos sobre vivencias personales, en situaciones interaccionales entre cercanos). Esos resultados presentan significancia estadística, de ese modo, evidencian que el tipo de coleta interfiere en la distribución de las variantes y que hay la actuación de las dos reglas variables (de referencia determinada y de referencia indeterminada) en los usos de las formas variables estudiadas. Cuando el tipo de muestra es proveniente de entrevistas, los usos de las formas pronominales **você** y **cê** ocurrieron indiferentemente del tipo de referencia. Esos resultados contribuyen para el panorama de los cambios en el paradigma pronominal del portugués brasileño, y surgen para la importancia de la diversificación de las clases de muestras lingüísticas en análisis de los procesos variables pragmáticamente motivados.

PALABRAS CLAVES: formas pronominales de 2ªPS; función referencial; variación lingüística.

LISTA DE ILUSTRAÇÕES

Figura 1 - Correlação entre as pessoas do discurso e o eixo interlocutivo	29
Figura 2 - Relação entre as funções de referência à 2ªPS e de Indeterminação do sujeito e as formas pronominais variantes.....	33
Figura 3 - Escala de traços semânticos referencial de pronomes de 2ªPS	40
Figura 4 - <i>Continuum</i> do tipo referência dos pronomes de 2ªPS	41
Figura 5 - Situações de convergência e divergência na interação	62
Figura 6 - Derivação intrafalante a partir de variação interfalante	68
Figura 7 - Pessoas e papéis na situação da fala	70
Figura 8 - Esquema de uma relação não recíproca (assimétrica)	76
Figura 9 - Esquema de uma relação recíproca (simétrica)	77
Figura 10 - A semântica/sistema bidimensional (poder e solidariedade) em equilíbrio	78
Figura 11 - O sistema bidimensional (poder e solidariedade) em conflito.....	79
Figura 12 - Díades sociais envolvendo conflito semântico	79
Figura 13 - Díades sociais em equilíbrio	80
Figura 14 - Esquema dos extremos de diferenciação T-V.....	83
Figura 15 - Esquema dos extremos de diferenciação T-V.....	84
Figura 16 - Mapa dos seis subsistemas dos pronomes de segunda pessoa do singular do português brasileiro no final do XX e início do século XXI, o Brasil de 1709 e o Brasil do século XXI.....	94
Figura 17 - Mapa de distribuição das variantes de pronomes de 2ªPS em capitais dos estados brasileiros.....	96
Figura 18 - Mapa de distribuição das variantes de pronomes de 2ªPS em áreas de não capitais dos estados brasileiros	97
Figura 19 - localização da cidade de Itabaiana, em Sergipe.....	131
Figura 20 - <i>Campus</i> Professor Alberto Carvalho da UFS	132
Figura 21 - Organograma do banco de dados Falares Sergipanos	138
Figura 22 - Rede Social dos oito informantes universitários do <i>Campus</i> Prof. Alberto Carvalho/UFS	141
Figura 23 – Regressão condicional para as ocorrências das variantes quanto ao tipo de função referencial e ao tipo de amostra.....	155

LISTA DE QUADROS

Quadro 1- Paradigmas dos pronomes pessoais no português brasileiro.....	15
Quadro 2 - Recursos semânticos de inclusão do paradigma de pessoa.....	29
Quadro 3 - Características gerais das pessoas do discurso e da não pessoa	30
Quadro 4 - Graus de indeterminação dos pronomes de 2ªPS	36
Quadro 5 - Principais objetivos do método de coleta.....	55
Quadro 6 - Perguntas sobre “Perigo de morte” e “Infância”	57
Quadro 7 - Dimensões da acomodação subjetiva e objetiva da fala	64
Quadro 8 - Hierarquia de atributos e papéis da audiência.....	70
Quadro 9 - Formas de tratamento da Língua Portuguesa no Limiar do século XX	85
Quadro 10 - Formas de tratamento da Língua Portuguesa no final do século XX.....	86
Quadro 11 - Atos Ameaçadores a Faces.....	90
Quadro 12 - Perfil social das documentadoras da amostra.....	134
Quadro 13 – Dados sociais dos informantes selecionados da amostra UFS-Itabaiana/SE (2018)	136
Quadro 14 - Escala de gradação para o controle da distância social entre os informantes da rede social.....	139
Quadro 15 - Dados sociais dos informantes em rede social pessoal	140
Quadro 16 - Representação das interações da rede social	141

LISTA DE TABELAS E GRÁFICOS

Tabela 1 - O uso das formas de 2ªPS em função do tipo de referência na região Norte	42
Tabela 2 - O uso das formas de 2ªPS em função do tipo de referência na região Nordeste	43
Tabela 3 - O uso das formas de 2ªPS em função do tipo de referência na região Centro-oeste	44
Tabela 4 - O uso das formas de 2ªPS em função do tipo de referência na região Sudeste	45
Tabela 5 - O uso das formas de 2ªPS em função do tipo de referência na região Sul	46
Tabela 6 - O uso das formas de 2ªPS em função do tipo de referência (não binário)	47
Tabela 7 - Diferentes abordagens da referência do você/cê	50
Tabela 8 - Uso de tu em função do grau de intimidade entre os interlocutores	99
Tabela 9 - Uso de CÊ específico e relação entre os interlocutores (Infs e Docs – GESOL-SP-2000)	101
Tabela 10 - Uso das formas de 2ªPS em função do sexo/gênero dos interlocutores	102
Tabela 11 - Uso das formas de 2ªPS em função do tipo de texto/sequência discursiva	105
Tabela 12 - Correlação entre o uso das formas de 2ªPS e a variável tópico discursivo	108
Tabela 13 - Uso das formas de 2ªPS em função do tipo de discurso na região Norte	110
Tabela 14 - Uso das formas de 2ªPS em função do tipo de discurso na região Nordeste	110
Tabela 15 - Uso das formas de 2ªPS em função do tipo de discurso na região Centro-Oeste	111
Tabela 16 - Uso das formas de 2ªPS em função do tipo de discurso na região Sudeste	111
Tabela 17 - Uso das formas de 2ªPS em função do tipo de discurso na região Sul	112
Tabela 18 - Função referencial e a variável tópico discursivo no uso das formas de 2ªPS	112
Tabela 19 - Uso das formas de 2ªPS em função do tipo de coleta	114
Tabela 20 - Uso das formas de 2ªPS em função do paralelismo formal na região Norte	116
Tabela 21 - Uso das formas de 2ªPS em função do paralelismo formal na região Nordeste	117
Tabela 22 - Uso das formas de 2ªPS em função do paralelismo formal na região Centro-Oeste	119
Tabela 23 - Uso das formas de 2ªPS em função do paralelismo formal	119
Tabela 24 - Uso das formas de 2ªPS em função do paralelismo formal na região Sudeste ...	120
Tabela 25 - Uso das formas de 2ªPS em função do paralelismo formal na região Sul	120
Tabela 26 - Correlação entre o uso das formas de 2ªPS e o paralelismo formal em contextos de referência específica	121
Tabela 27 - Uso das formas de 2ªPS em função do sexo/gênero do informante na região Nordeste	122
Tabela 28 - Relação entre o uso das formas de 2ªPS e o sexo/gênero do informante em outras regiões	123
Tabela 29 - Correlação entre o uso das formas de 2ªPS e o sexo/gênero do informante em contextos de referência determinada	124
Tabela 30 - Emprego de você e cê conforme a referência do pronome e o sexo/gênero dos informantes (GESOL-SP-2000)	124
Tabela 31 - Estratificação da amostra	133
Tabela 32 - Estratificação da amostra quanto à variável sexo/gênero	134
Tabela 33 - Estratificação da amostra utilizada	136
Tabela 34 – Distribuição das variantes tu , você e cê na função sintática de sujeito e o tipo de referência	154
Tabela 35 - Uso de você e cê em função do tipo de referência e da variável grau de intimidade entre os interlocutores	159

Tabela 36 - Uso de você e cê em função do tipo de referência e da variável relações (as)simétricas quanto ao sexo/gênero dos interlocutores	161
Tabela 37 - Uso de você e cê em função do tipo de referência e da variável turno de fala....	162
Tabela 38 - Uso de você e cê em função da referência e da variável pergunta proposital.....	163
Tabela 39 – Uso de você e cê em função do tipo de referência e da variável sexo/gênero do falante	165
Tabela 40 - Uso de você e cê em função do tipo de referência e da variável tipo de deslocamento	166
Tabela 41 - Uso de você e cê em função do tipo de referência e da variável tipo de sequência discursiva.....	168
Tabela 42 - Uso de você e cê em função do tipo de referência e da variável tópico discursivo	170
Tabela 43 - Uso de você e cê em função do tipo de referência e da variável tipo de discurso	171
Tabela 44 - Uso de você e cê em função do tipo de referência e da variável paralelismo formal	173
Tabela 45 - Uso de você e cê em função do tipo de referência e da variável efeito gatilho...	175
Gráfico 1 – Crença em relação ao próprio uso linguístico	157

SUMÁRIO

1	INTRODUÇÃO	15
2	A 2ªPS E SUA RELAÇÃO COM O TIPO DE REFERÊNCIA	24
2.1	PRONOMES PESSOAIS: A NOÇÃO DE PESSOA, NÚMERO E GÊNERO	24
2.2	A CATEGORIA PESSOA E A EXPRESSÃO DA 2ªPS	28
2.3	A REFERÊNCIA À 2ªPS E À INDETERMINAÇÃO SUJEITO: RELAÇÃO ENTRE FORMA E FUNÇÃO	32
2.4	O CONTROLE DA REFERÊNCIA DAS FORMAS TU, VOCÊ E CÊ NOS ESTUDOS SOCIOLINGÜÍSTICOS	41
2.5	FUNÇÃO REFERENCIAL DOS PRONOMES TU, VOCÊ E CÊ: CONSIDERAÇÕES GERAIS	51
3	ACOMODAÇÃO LINGÜÍSTICA, POLIDEZ E PRONOMES DE PODER E SOLIDARIEDADE	53
3.1	TEORIA DA ACOMODAÇÃO LINGÜÍSTICA E ENTREVISTAS SOCIOLINGÜÍSTICAS	53
3.1.1	Entrevistas sociolinguísticas: o método laboviano de coleta de dados linguísticos	53
3.1.2	Acomodação linguística	60
3.1.3	O modelo de <i>Audience Design</i>	66
3.2	ASPECTOS PRAGMÁTICOS E FORMAS PRONOMINAIS DE 2ªPS	74
3.2.1	As dimensões de poder e de solidariedade	75
3.2.2	A dimensão da neutralidade	82
3.2.3	A relação entre dêixis social, face e polidez	88
4	O USO VARIÁVEL DAS FORMAS PRONOMINAIS TU, VOCÊ E CÊ NO PORTUGUÊS BRASILEIRO	92
4.1	DISTRIBUIÇÃO DAS FORMAS PRONOMINAIS DE 2ªPS NO PORTUGUÊS BRASILEIRO	93
4.2	VARIÁVEIS INTERACIONAIS	98
4.2.1	Grau de intimidade entre os interlocutores	98
4.2.2	(As)simetria quanto ao sexo/gênero dos interlocutores	101
4.3	VARIÁVEIS ESTILÍSTICAS	104
4.3.1	Tipo de sequência discursiva	104
4.3.2	Tópico discursivo	107
4.3.3	Tipo de discurso	109
4.3.4	Tipo de coleta	114
4.4	VARIÁVEL ESTRUTURAL	115
4.5	VARIÁVEL SOCIAL	121
4.6	PROBLEMAS METODOLÓGICOS NO ESTUDO DA VARIAÇÃO DAS FORMAS DE 2ªPS: LIMITAÇÕES DOS ESTUDOS REALIZADOS	125
5	O USO VARIÁVEL DAS FORMAS PRONOMINAIS TU, VOCÊ E CÊ NO FALAR DE UNIVERISITÁRIOS ITABAIANENSES: ANÁLISE, RESULTADOS E DISCUSSÃO	127
5.1	O PONTO DE PARTIDA PARA O DESENVOLVIMENTO DO ESTUDO: ASPECTOS METODOLÓGICOS	127
5.1.1	Comunidade de prática em análise	128
5.1.1.1	Compreendendo a noção de comunidade de prática	128
5.1.1.2	<i>Lócus</i> da pesquisa: Universidade Federal de Sergipe, <i>Campus</i> Prof. Alberto Carvalho	129
5.1.2	Constituição da amostra: entrevistas sociolinguísticas de Itabaiana/SE (2018)	132
5.1.3	A amostra rede social de informantes universitários de Itabaiana/SE (2013)	137
5.1.4	As variáveis controladas	143
5.1.5	Tratamento estatístico dos dados	152
5.2	O PADRÃO DE COMPORTAMENTO REFERENCIAL DAS FORMAS TU, VOCÊ E CÊ: RESULTADOS E DISCUSSÃO	153
5.2.1	Análise das variantes tu, você e cê	153
5.2.2	Análise das variantes você e cê	158
5.2.2.1	Grau de intimidade entre os interlocutores	159
5.2.2.2	Relações (as)simétricas quanto ao sexo/gênero dos interlocutores	160
5.2.2.3	Turno de fala	161

5.2.2.4	Pergunta proposital.....	163
5.2.2.5	Sexo/gênero do falante	164
5.2.2.6	Deslocamento	165
5.2.2.7	Tipo de sequência discursiva	166
5.2.2.8	Tópico de discurso.....	169
5.2.2.9	Tipo de discurso	170
5.2.2.10	Paralelismo formal	172
5.2.2.11	Efeito gatilho	173
6	CONSIDERAÇÕES FINAIS.....	176
	REFERÊNCIAS	179
	APÊNDICES	189
	APÊNDICE A – ROTEIRO DE ENTREVISTA	190
	ANEXOS.....	195
	ANEXO A – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO.....	196
	ANEXO B – FICHA SOCIAL DO INFORMANTE.....	197
	ANEXO C – NORMAS ADOTADAS PELO GRUPO DE ESTUDOS EM LINGUAGEM INTERAÇÃO E SOCIEDADE (GELINS) PARA A REALIZAÇÃO DE TRANSCRIÇÃO ORTOGRÁFICA	198

1 INTRODUÇÃO

A língua não é um sistema estável e muito menos homogêneo como a tradição gramatical tenta transparecer, mas sim, dinâmica e heterogênea. Todas as línguas naturais humanas possuem um caráter dinâmico (cf. WEINREICH; LABOV; HERZOG, 2006 [1968]). Em decorrência de seu dinamismo e heterogeneidade, a língua sofre variação e/ou mudança no decorrer do tempo. No português, é possível constatar diversos fenômenos variáveis no âmbito fonológico ou ainda em níveis gramaticais mais altos, para além da fonologia (cf. FREITAG, 2007), tais como: sintático, morfológico, semântico, discursivo, lexical, morfossintático, entre outros.

Entre os fenômenos variáveis em nível morfossintático, no português brasileiro, estão aqueles relacionados ao paradigma pronominal (cf. CASTILHO, 2019) pelo fato de, no decorrer dos anos, outras formas terem surgido para exercer funções já existentes no paradigma tradicional, conforme se evidencia, no quadro a seguir, quanto aos pronomes pessoais:

Quadro 1- Paradigmas dos pronomes pessoais no português brasileiro

PESSOA	PB TRADICIONAL		PB EM USO	
	SUJEITO	COMPLEMENTO	SUJEITO	COMPLEMENTO
1ª singular	Eu	me, mim, comigo	eu, a gente	eu, me, mim, prep.+ eu, mim
2ª singular	tu, você, o senhor, a senhora	te, ti, contigo, prep.+ o senhor/ a senhora	tu, você, ocê, cê	Você, ocê, cê, lhe, te, ti, prep.+ tu/você/ocê/cê
3ª singular	ele, ela	o, a, lhe, se, si, consigo	Ele/ei, ela	ele, ela, lhe, prep. + ele/ela
1ª plural	Nós	nos, conosco	a gente, nós	nos, a gente, prep.+ nós/a gente
2ª plural	vós, os senhores, as senhoras, vocês	vos, convosco, prep.+ os senhores/as senhoras	Vocês, ocês, cês	vocês/ocês/cês, lhes, prep.+ vocês/ocês
3ª plural	eles, elas	os, as, lhes, se, si, consigo	Eles/eis, elas	Eles/eis, elas, lhes, prep.+ eles, elas

Fonte: Elaborado pela autora a partir de Bechara (2015, p. 173) e Castilho (2019, p. 477).

Ao compararmos o paradigma tradicional dos pronomes pessoais do caso reto (pronomes sujeitos) e o paradigma em uso do português brasileiro (PB), expostos no quadro 1, verifica-se que, neste último, há a inclusão de duas formas: **você¹(s)** (e suas reduções **ocê(s)** e **cê(s)**) e **a gente**. A primeira é utilizada para expressar a segunda pessoa do singular (2ªPS)

¹ A classificação canônica da forma **você** é como pronome de tratamento, visto que esta se originou da expressão nominal de tratamento **Vossa Mercê** (cf. LOPES, 2007; FARACO, 2017 [1996]).

(**você, ocê e cê**) e do plural (**vocês, ocês, cês**), já a segunda é usada para expressar a primeira pessoa do plural (1ªPP). A entrada dessas formas no paradigma pronominal ocasionou processos de variação/mudança linguística em níveis gramaticais distintos (cf. LOPES, 2007). Dentre os processos variacionais decorrentes dessa inserção, interessa-nos os que ocorrem no escopo da 2ªPS.

De acordo com a perspectiva tradicional, a referência à 2ªPS no paradigma pronominal do português é realizada pelo pronome pessoal **tu**. Entretanto, no PB corrente, a forma **você** tem sido utilizada para fazer essa referência, assim como as variantes **ocê** e **cê**. Muitas pesquisas sociolinguísticas, nas últimas décadas, têm focalizado o fenômeno da variação no uso das formas pronominais de 2ªPS como objeto de estudo (MARTINS, 2010; ALVES, 2010; FRANCESCHINI, 2011; NASCIMENTO, 2011; SCHERRE *et al.*, 2015; LOPES, 2017; LOPES *et al.*, 2018; entre outros), as quais evidenciam que esse fenômeno está intrinsecamente relacionado a aspectos sociais, interacionais/pragmáticos, estilísticos e dialetais.

Scherre *et al.* (2015) realizaram um mapeamento sociolinguístico sobre a variação no sistema pronominal de 2ªPS no PB, tendo como base estudos sincrônicos que utilizaram dados da oralidade e que foram realizados até 2012, a fim de evidenciar, sobretudo, a proporção distribucional do fenômeno nas diferentes regiões do país. A partir do levantamento dos resultados apresentados em diversos estudos, os autores apresentaram uma proposta de mapeamento constituída por seis subsistemas de tratamento para o português brasileiro, a saber:

1. Subsistema **só você**: uso exclusivo das formas “*você/cê/ocê*”;
2. Subsistema **mais tu com concordância baixa**: uso médio de “*tu*” acima de 60% com concordância abaixo de 10%;
3. Subsistema **mais tu com concordância alta**: uso médio de “*tu*” acima de 60% com concordância entre 40% e 60%;
4. Subsistema **tu/você com concordância baixa**: uso médio de “*tu*” abaixo de 60% com concordância abaixo de 10%;
5. Subsistema **tu/você com concordância média**: uso médio de “*tu*” abaixo de 60% com concordância entre 10% e 39%;
6. Subsistema **você/tu**: uso de 1% a 90% sem concordância (cf. SCHERRE *et al.*, 2015, p. 138-139).

O mapeamento evidenciou que o fenômeno ocorre de maneira diversificada nas regiões do Brasil, como, por exemplo, na região Sul, há o predomínio da forma **tu** no Rio Grande do Sul (subsistema **mais tu com concordância baixa**); na região Sudeste, há o uso majoritário da forma **você**, com exceção de Santos (subsistema **só você**) (SCHERRE *et al.*, 2015). Pesquisas sobre o fenômeno em dados de fala ainda não foram desenvolvidas em todos os estados do país,

como é o caso de Sergipe² (um dos fatos que nos motivou a realizar a presente pesquisa com dados de fala de sergipanos, conforme ressaltaremos adiante). Os subsistemas se aplicam, inclusive para Sergipe. Com base em resultados de testes de avaliação subjetiva (cf. ARAUJO; JESUS, 2018; ARAUJO; MENDONÇA, 2018), é possível presumir que este estado (principalmente a capital Aracaju) faz parte do sexto subsistema – **você/tu**, com ocorrências de **tu** de 1% a 90% sem concordância. Tais resultados demonstram que o emprego das formas variantes pode ser distinto a depender da área geográfica.

Uma implicação a ser considerada, ao analisar esse fenômeno linguístico, é a quem esses pronomes fazem referência. As formas variantes **tu**, **você** e **cê**³ podem ser usadas tanto para se reportarem a um referente determinado quanto a um indeterminado. Os exemplos a seguir, os quais pertencem ao banco de dados *Falares Sergipanos* (FREITAG, 2013; 2017), ilustram os usos dessas variantes de acordo com o tipo de referente do pronome.

- (1) D.S.: tenha vergonha doido tome vergonha... quem nunca viu hein? respeite o cara...
sim o sítio de quem era? que **tu** ia?
D.M.: era não... é do meu avô... quer comprar? tá à venda
(int.UFS-Itabaiana2013_D.S.cdt D.M.sdt P M_M 01)⁴
- (2) D.S.: ()... *pra você qual é o benefício da Bolsa Família? ela é boa ela ruim?*
(int.UFS-Itabaiana2013_D.S.cdt D.M.sdt P M_M 01)
- (3) (...) minha fia (hes) assim eu acho os Professor eles são muito exigente não sei se é por causa da dificuldade do curso também mas **cê** vê que tem turma que é com quatro cinco aluno né? e porque realmente a gente tem né Andréia dificuldade? [...]
(33ent.UFS-Itabaiana2018_desl. II_final_jos.fs.30)⁵

² Em testes de avaliação subjetiva, aplicados a informantes de diversas localidades do estado sergipano, os participantes da pesquisa mencionaram que as formas variantes **tu** e **você** são utilizadas pelas pessoas das cidades onde moram (cf. ARAUJO; JESUS, 2018; ARAUJO; MENDONÇA, 2018; MENDONÇA, ARAUJO, 2019; PAES, 2019). Embora existam pesquisas no âmbito da avaliação/percepção linguística, não foram constatadas pesquisas com dados linguísticos da fala de sergipanos que focalizem a variação no paradigma pronominal da 2ªPS na posição de sujeito. Na posição sintática de objeto, há um estudo realizado por Araujo e Borges (2021).

³ A forma variante **ocê** é bastante rara na fala de sergipanos, principalmente, na região da Grande Aracaju e Agreste Central; em virtude disso, esta não será focalizada na presente pesquisa. Não serão foco também, em nossa investigação, as formas de tratamento **senhor/senhora**, **variante zero** (sujeitos nulos) e nem outras formas que possam surgir no *corpus* com referência à 2ªPS. A pesquisa se restringe ao estudo das formas variantes **tu**, **você** e **cê**.

⁴ Os dados foram retirados da amostra *Rede Social de Informantes Universitários de Itabaiana/SE*. A sigla refere-se ao tipo de coleta (int. = interação), à identificação da localidade de coleta e ano (UFS-Itabaiana2013), ao falante (iniciais do nome) com o domínio do tópico (cdt), ao falante (iniciais do nome) sem o domínio do tópico (sdt), ao grau de proximidade (próximo= P; distante= D), ao número da interação (número ao final).

⁵ Os dados foram retirados da *Amostra de falantes universitários de Itabaiana/SE (2018)*. A sigla refere-se ao número da entrevista (número inicial), ao tipo de coleta (ent. = entrevista), à identificação da localidade de coleta e ano (UFS-Itabaiana2018), ao tipo de deslocamento e período curso (desl.II_final), ao falante (iniciais do nome), ao sexo/gênero do informante (f = feminino), à escolaridade (s = superior), à idade do informante (número final).

- (4) C.A.: você já chegou a avaliar as competências assim positiva ou negativa alguma vez num amigo seu?

W.S.: ah já... sempre isso aí é constante... e principalmente quando **cê**... com a convivência que **cê** vai passando com certo colega **você**... aprende sabe qual é os valores dele... e existe nele... quando **você** não conhece é mais difícil você aquela pessoa é boa pessoa tudo mas depois que **cê** passa a conviver... é que **você** vai saber quais são os pontos positivos e os negativos dele

(C.A.cdt W.S.sdt P MM 17)

- (5) C.A.: a imagem... não a geralmente assim né? quando é um conhecido eu converso bastante porque a fila demora bastante pra **cê** ser atendido né? eles também não utilizam uma forma de distrair o cliente (est) já que tem poucos funcionários questão de de enxugar... reduzir os gastos... se for uma pessoa desconhecida não não não sou de puxar muita conversa não... sou mais na minha agora se for uma pessoa conhecida eu converso bastante um pouco... presto atenção às vezes eu tô numa numa fi-... (inclusive) uma casa lotérica né? eu não fico virado assim pra frente eu geralmente eu fico viro de costa né? pra o atendente ou na parede encostado... que aí eu percebo as pessoas que entra que sai... porque **você** assim olhando pra o caixa **você** não sabe quem tá atrás de **você** (est) e geralmente eu gosto de ficar na parede... olhando dos dois lados né? já já aconteceu de pessoas inclusive uma professora lá em Frei Paulo... tava no Banco e retirou o dinheiro... e quando saiu fora e um casal acompanhou ela né? e daí a pouco ela voltou e disse que tinha sido roubada... questão de saidinha no banco aquele pessoal entra... às vezes é alguém que **você** menos... pensa que né? vai lhe roubar

(C.A.cdt W.S.sdt P MM 17)

Nos três primeiros exemplos, as formas pronominais de 2ªPS são utilizadas para fazer referência a um interlocutor dêitico, que está presente na situação interlocutiva. Já no exemplo (4) as formas destacadas foram utilizadas com traços semânticos genéricos, ou seja, com referência indeterminada. Observa-se que a referência é feita a um quantitativo de pessoas mais abrangente, podendo ser qualquer pessoa; neste caso, o interlocutor, bem como qualquer outra pessoa, pode passar pela situação retratada. Em (5), os usos destacados da forma **você** expressam também indeterminação do sujeito. Entretanto, neste caso, o falante utiliza a forma para se referir a si mesmo a fim de preservar a sua face no contexto interlocutivo. Tais exemplos evidenciam que os usos dessas formas pronominais não possuem os mesmos traços semânticos, visto que aspectos discursivos e pragmáticos/interacionais operam no tipo de indeterminação, fazendo com que haja matizes/graus distintos (SOUSA, 2008; NASCIMENTO, 2011; SILVA, 2015; COSTA, 2016; LOPES, 2017) – os usos das formas no exemplo (5) possuem um grau de indeterminação menor que os usos em (4). Além disso, as formas variantes são utilizadas para se reportarem, de fato, à 2ªPS apenas em contextos de referência determinada, uma vez que nos demais contextos a referência não é feita a uma pessoa dêitica, mas sim às pessoas de forma geral (passando a ter, de certa forma, uma interpretação plural) ou, ainda, ao próprio falante.

Os estudos sociolinguísticos sobre o fenômeno em questão tratam o tipo de referência como uma variável independente. Contudo, cada tipo de referência estaria muito mais para uma variável dependente do que para uma independente, caracterizando-se como duas funções (regras) de uso das variantes em estudo (com peculiaridades distintas), a saber: a referência determinada das formas variantes constitui-se como a função de referência à segunda pessoa do singular e a referência indeterminada das formas variantes constitui-se como a função de referência à indeterminação do sujeito. Silva, M. (2019) estudou o fenômeno em questão, no falar acreano, com o objetivo de verificar se o fenômeno se trata de um caso de variação ou de escolha funcional, separando os dados de referência determinada e indeterminada. No entanto, a autora apresenta resultados separados apenas para algumas variáveis. Além disso, o quantitativo de dados do fenômeno é pouco robusto para fazer generalizações ou ter conclusões mais assertivas (221 dados).

Diante dessas constatações, partimos da tese de que os usos variáveis dos pronomes **tu**, **você** e **cê** estão relacionados à distinção do tipo de função referencial (determinada ou indeterminada), a qual interfere na proporção de distribuição das formas e nos efeitos dos condicionamentos sociais, interacionais, estilísticos e estruturais nestas, sendo necessário que cada tipo de função referencial seja controlado como uma regra variável separada. A partir dessa tese proposta, decorrem as seguintes questões e hipóteses de pesquisa:

1. Qual das variantes em questão tem uma proporção distribucional maior em cada função referencial?

Os usos das variantes **tu** e **cê** tendem a ser mais frequente na função referencial determinada, mas não suplantariam o número de ocorrências da variante **você**, uma vez que, em testes de avaliação subjetiva (cf. ARAUJO; JESUS, 2018; ARAUJO; MENDONÇA, 2018), sergipanos afirmaram utilizar mais esta última variante do que a variante **tu**.

2. O padrão de distribuição dessas variantes em cada tipo de referência é resultado do tipo de coleta dos dados de fala?

Os usos das variantes **tu**, **você** e **cê** são sensíveis ao tipo de coleta dos dados de fala. O tipo de coleta que propicia uma maior troca de turnos de fala e motiva os dois interactantes a discorrerem sobre os tópicos abordados na situação interlocutiva, como é o caso do tipo de coleta denominada de interações conduzidas, favorece o uso da variante **tu**. Já o tipo de coleta cujo foco é propiciar que apenas um dos interactantes (o entrevistado) desenvolva os tópicos abordados, como é o caso do tipo de coleta denominado de entrevistas sociolinguísticas, favorece o uso da variante **você**.

Com base nessas questões e hipóteses aventadas, objetivamos identificar o padrão de comportamento das variantes **tu**, **você** e **cê**, na posição sintática de sujeito, em contextos de função referencial determinada e indeterminada, utilizando dados de fala de estudantes universitários sergipanos, mais especificamente, da região do Agreste Central Sergipano⁶.

Para desenvolvermos a pesquisa neste âmbito, coletamos 80 entrevistas sociolinguísticas em 2018 na comunidade de prática Universidade Federal de Sergipe, *Campus Professor Alberto Carvalho* (Itabaiana/SE), das quais selecionamos apenas 16 entrevistas⁷ para compor o *corpus* de análise. Como objetivamos verificar se o tipo de metodologia de coleta dos dados de fala interfere ou não nos usos das formas variantes estudadas, utilizamos 16 interações conduzidas (conversas entre dois estudantes, em que um deles é responsável por conduzir a interação) da amostra Rede Social de Informantes Universitários de Itabaiana/SE (2013), a qual pertence ao banco de dados *Falares Sergipanos* (FREITAG, 2017).

É importante ressaltar que a escolha dessa comunidade de prática como *lócus* para obtermos os dados de fala para o desenvolvimento desta pesquisa deve-se ao sentimento de pertencimento que possuímos em relação à comunidade e à região e pelo fato de, em nossas vivências sociais, constatarmos o uso do fenômeno aqui estudado na região do Agreste sergipano. Ademais, escolhemo-la pelo fato de existir desde 2009 um processo de documentação sociolinguística do falar de universitários dessa comunidade (coordenado pela Prof.^a Dra. Raquel Meister Ko. Freitag), permitindo-nos utilizar amostras já coletadas (mais especificamente a amostra de interações conduzidas) e contribuir para a descrição linguística da região.

Para cada uma das regras variáveis em estudo, controlamos os efeitos das seguintes variáveis independentes: i) grau de intimidade entre os interlocutores; ii) relações (as)simétricas quanto ao sexo/gênero dos interlocutores; iii) turno de fala; iv) pergunta proposital; v) sexo/gênero do falante; vi) deslocamento social; vii) tipo de amostra; viii) tipo de sequência discursiva; ix) tópico discursivo; x) tipo de discurso; xi) paralelismo formal e xii) efeito gatilho. Como respaldo teórico, a presente pesquisa é desenvolvida à luz da Sociolinguística Variacionista (WEINRICH; LABOV; HERZOG, 2006; LABOV, 2008 [1972]) em interface com a Pragmática (BROWN; GILMAN, 1980; BROWN; LEVINSON, 2011 [1987]) e com a

⁶ O Agreste Central Sergipano é constituído por 14 (quatorze) municípios, a saber: Areia Branca, Campo do Brito, Carira, Frei Paulo, Itabaiana, Macambira, Malhador, Moita Bonita, Nossa Senhora Aparecida, Pedra Mole, Pinhão, Ribeirópolis, São Domingos e São Miguel do Aleixo.

⁷ A escolha de apenas 16 entrevistas se deve ao volume de dados do fenômeno em estudo e à necessidade de equiparar o quantitativo de gravações da amostra de entrevistas à de interações conduzidas com o mesmo perfil de informante.

Teoria da Acomodação (BELL, 1984, 2001; GILES; COUPLAND; COUPLAND, 1991). A associação entre as abordagens permite a mensuração de efeitos interacionais/pragmáticos, estilísticos e sociais nos usos linguísticos e, por conseguinte, obter resultados mais abrangentes do fenômeno em questão.

A realização de um estudo sobre o padrão de comportamento referencial das variantes **tu**, **você** e **cê** justifica-se pelo fato de, apesar de já existirem estudos nesse escopo no português brasileiro, haver a necessidade de uma análise mais acurada. Assim, esta pesquisa traz contribuições quanto aos procedimentos investigativos por analisar o tipo de função referencial como uma variável dependente, o que possibilita esclarecer o funcionamento do uso variável das formas **tu**, **você** e **cê** sem sobrepor as regras de expressão da 2ªPS (referência determinada) e indeterminação do sujeito (referência indeterminada).

Cabe ressaltar que a execução desta pesquisa se torna ainda mais necessária por não existirem pesquisas de cunho sociolinguístico em Sergipe, evidenciando o ineditismo no âmbito espacial. A partir desta pesquisa, ampliamos o universo de estudo a respeito dos pronomes de 2ªPS, contribuindo, assim, para a descrição do português no Brasil, mais especificamente, para o mapeamento do uso desses pronomes. Além disso, a presente pesquisa contribui para que futuramente a 2ªPS deixe de ser sub-representada nos materiais didáticos, tendo em vista que, em parte, essa sub-representação é decorrente da falta de estudos descritivos que contemplem todas as variedades do português brasileiro. Evidencia-se, dessa maneira, a relevância do estudo proposto.

A presente pesquisa está vinculada ao Grupo de Estudos em Linguagem, Interação e Sociedade – GELINS, coordenado pela Prof.^a Dra. Raquel Meister Ko. Freitag e pela Prof.^a Dra. Leilane Ramos da Silva. Outras pesquisas sobre pronomes, vinculadas ao GELINS, foram desenvolvidas sob a orientação da Prof.^a Dra. Raquel Meister Ko. Freitag, a saber: i) “Estratégias de polidez e a variação de nós vs. a gente na fala de discentes da Universidade Federal de Sergipe” (2014), de Kelly Carine dos Santos; ii) “A literatura como fonte de dados: um olhar sociolinguístico sobre a obra história da minha infância, de Gilberto Amado” (2014), de Fernanda Bispo Correia; iii) “Variação na expressão da 1ª pessoa do plural: indeterminação do sujeito e polidez”, de Josilene de Jesus Mendonça; iv) “Desenvolvimento de objeto pedagógico para o ensino produtivo de gramática: a Balança das Relações Sociais” (2019), de Jilcicleide Augusta Paes; v) “Variação no preenchimento da posição determinante antes de possessivos pré-nominais: padrões dialetais e contatos” (2020), de José Manoel Siqueira da Silva; vi) “Traços semânticos da referência à primeira pessoa do plural no português brasileiro: um estudo em tempo real” (2022), de Josilene de Jesus Mendonça.

O presente texto de tese segue estruturado em cinco seções, além desta introdução (que se constitui como a primeira seção).

Na segunda seção, intitulada “A 2ªPS e sua relação com o tipo de referência”, discutimos aspectos conceituais a respeito da 2ªPS. Inicialmente, expomos noções gerais relacionadas às categorias morfossintáticas de pessoa, número e gênero com base, principalmente, em Benveniste (1995 [1956]), Harley e Ritter (2002) e Siewierska (2004). Em seguida, apresentamos as propriedades que cooperam para a definição da 2ªPS. Logo após, debatemos a relação existente entre forma e função no processo de variação das formas pronominais de 2ªPS. A subseção 2.4 é voltada para sistematização dos resultados das pesquisas sociolinguísticas que controlaram o tipo de referência das formas **tu**, **você** e **cê** no PB. Na subseção 2.5, apresentamos nossas considerações gerais sobre as discussões feitas na presente seção.

Na terceira seção, intitulada “Acomodação linguística, polidez e os pronomes de poder e solidariedade”, apresentamos os princípios teóricos que respaldam esta pesquisa. Na subseção 3.1, explanamos a respeito da teoria da acomodação (a partir, principalmente, de Giles, Coupland e Coupland (1991) e Alan Bell (1984; 2001)), relacionando-a com o modelo de coleta das entrevistas sociolinguísticas. Em seguida, abordamos os aspectos pragmáticos que estão relacionados ao uso das variantes **tu**, **você** e **cê**. baseando-nos, sobretudo, em Brown e Gilman (1960), Brown e Levinson (2011 [1987]) e Cook (1997).

Na quarta seção, intitulada “O uso variável das formas pronominais **tu**, **você** e **cê** no PB”, apresentamos os resultados das pesquisas sobre o fenômeno em estudo. Primeiramente, explanamos a respeito dos resultados das pesquisas referentes à distribuição das formas variantes no português brasileiro com base no mapeamento feito por Scherre, Andrade e Catão (2019). Na sequência, apresentamos os resultados concernentes às variáveis independentes de cunho interacional. Nas subseções 4.3, 4.4 e 4.5, abordamos os resultados das pesquisas sobre as variáveis estilísticas, estruturais e sociais, respectivamente. Por fim, debatemos os problemas metodológicos e as limitações das pesquisas realizadas.

Na quinta seção, intitulada “O uso variável das formas pronominais **tu**, **você** e **cê** no falar de universitários itabaianenses: análise, resultados e discussão”, expomos primeiramente os procedimentos seguidos para a identificação de padrão de comportamento referencial das formas em estudo. Em seguida, discutimos os resultados obtidos com o desenvolvimento da presente pesquisa.

Na sexta seção, intitulada “Considerações finais”, retomamos a tese proposta e discutimos de que forma ela foi contemplada a partir dos resultados obtidos. Ademais, tecemos

considerações sobre as contribuições da pesquisa e possibilidades de estudos sobre fenômeno em questão.

Feitas as considerações a respeito dos direcionais inerentes a cada uma das seções da presente tese, convidamo-los à realização da leitura.

2 A 2ªPS E SUA RELAÇÃO COM O TIPO DE REFERÊNCIA

O objetivo desta seção é apresentar uma discussão a respeito de aspectos relacionados à categoria pronomes, especificamente, àqueles que contribuem para o entendimento de como se configura a 2ªPS. A seção está estruturada em cinco subseções. Apresentamos, na primeira, as categorias morfossintáticas de pessoa, número e gênero. Na subseção 2.2, abordamos as propriedades que contribuem para a definição da 2ªPS. Em seguida, explanamos a respeito da relação existente entre forma e função. Na seção 2.4, sistematizamos os resultados das pesquisas sociolinguísticas que controlaram o tipo de referência das formas **tu**, **você** e **cê** no português brasileiro. Por fim, na subseção 2.5, fazemos nossas considerações gerais a respeito das discussões feitas na presente seção.

2.1 PRONOMES PESSOAIS: A NOÇÃO DE PESSOA, NÚMERO E GÊNERO

Nas línguas naturais, uma das formas de se fazer referência discursivamente é por meio do uso de pronomes pessoais. Estes são desprovidos de conteúdo descritivo próprio (WIESE; SIMON, 2002). Isso significa que, para identificar o referente de uma forma pronominal, é necessário ancorar-se em outros elementos discursivos (tanto do contexto linguístico como do extralinguístico), existindo, assim, uma relação de dependência contextual (WIESE; SIMON, 2002).

A ligação entre um pronome e um objeto referencial pode ser feita a partir da distinção de traços “[...] morfossemânticos, discursivo-pragmáticos, morfossintáticos e sintáticos”⁸ (WIESE; SIMON, 2002, p. 2, tradução nossa). Por meio dos aspectos morfossemânticos, é possível identificar distinções conceituais que baseiam a representação do paradigma pronominal, tais como número (singular e plural), papel no ato de fala (falante, destinatário e não participante), gênero (feminino e masculino), animacidade, +/- humano⁹, polidez (HARLEY; RITTER, 2002; WIESE; SIMON, 2002). Uma característica universal do paradigma pronominal, nas línguas naturais, aparentemente, é distinguir papéis discursivos

⁸ No original: [...] morpho-semantic, discourse-pragmatic, morpho-syntactic and syntactic means that serve to establish the link between a pronoun and an object (WIESE; SIMON, 2002, p. 2).

⁹ O Inglês, por exemplo, é uma língua que codifica a diferenciação semântica entre humano vs. não humano no uso dos pronomes pessoais de terceira pessoa: utiliza-se as formas *he* e *she* para humanos e a forma *it* para não humano (cf. WIESE; SIMON, 2002). No português, não temos diferenciação de usos de pronomes para esse traço.

(noção de pessoa) e indicar a ideia de singular/plural (noção de número) (WIESE; SIMON, 2002).

A noção de pessoa é debatida em diversos campos do conhecimento (como a filosofia, sociologia, psicologia, linguística, antropologia etc.) (SIEWIERSKA, 2004). Na linguística, esta noção é apreendida, sobretudo, “como uma categoria do sistema gramatical”¹⁰ (SIEWIERSKA, 2004, p. 1, tradução nossa). Essa categoria está relacionada à distinção entre: o falante de uma informação, o destinatário dessa informação e a parte sobre a qual se fala.

Tradicionalmente, tem-se nomeado o falante de primeira pessoa, o destinatário de segunda pessoa e a parte sobre a qual se fala de terceira pessoa. Entretanto, no que se refere à primeira e à segunda pessoas, a categoria de pessoa não exprime unicamente quem é o falante e o destinatário, respectivamente, mas também os papéis de participante ou de discurso destes em uma determinada situação interlocutiva (cf. SIEWIERSKA, 2004). A discussão sobre essa categoria será focalizada na seção 2.2.

A expressão da categoria de número, comumente, é feita, nos paradigmas pronominais, pela distinção entre singular e plural. O uso da marca de singular é realizado quando se quer fazer referência a um ser ou item dentro ou fora do eixo interlocutivo, já o uso da marca de plural é realizado quando se quer fazer referência a mais de um ser ou itens (cf. SIEWIERSKA, 2004). Entretanto, a noção de singular/plural não pode ser entendida de forma igualitária nas três pessoas do discurso, pois a terceira pessoa não compartilha os mesmos atributos da primeira e da segunda pessoas. A respeito disso, Siewierska (2004) ressalta que:

O que normalmente é entendido pela oposição singular/plural é uma distinção entre um e mais de um. Essa distinção, no entanto, não é necessariamente interpretada da mesma maneira com relação aos marcadores de pessoa, como no caso dos substantivos. Um substantivo singular refere-se a um único símbolo da entidade denotada pelo substantivo e um substantivo plural refere-se a vários símbolos da entidade relevante; portanto, livros referem-se a mais de uma instância da classe de objetos chamada livro. As formas plurais de terceira pessoa são interpretadas de maneira análoga, ou seja, referem-se a terceiros que consistem em vários indivíduos ou itens. As formas plurais de primeira pessoa, por outro lado, raramente se referem a mais de um falante. No inglês, podemos identificar vários ou mais falantes em uma cerimônia de juramento ou em alguma outra ocasião especial quando várias pessoas estão realmente falando simultaneamente. Em todos os outros casos, não identificamos mais de um falante, mas o falante e outra pessoa. Essa outra pessoa pode ser apenas o destinatário, [...] algum outro indivíduo ou grupo de indivíduos e o destinatário [...] ou algum indivíduo ou grupo de indivíduos entre os quais o destinatário não está incluído [...]¹¹ (SIEWIERSKA, 2004, p. 82, tradução nossa).

¹⁰ No original: [...] as a category of the grammatical system [...] (SIEWIERSKA, 2004, p. 1).

¹¹ No original: What is typically understood by the singular/plural opposition is a distinction between one and more than one. This distinction is not, however, necessarily interpreted in the same way with respect to person

Nesse sentido, o pronome de primeira pessoa do plural (1^aPP) não consiste no plural do pronome de primeira pessoa do singular (1^aPS), visto que a forma **nós** não possui o sentido de várias **eus**, isto é, várias pessoas falando. Quando utilizamos o pronome **nós**, este deve ser interpretado como **eu**, somado a uma pessoa ou mais, entre as quais o ouvinte pode ou não estar incluído (SIEWIERSKA, 2004).

As diferenciações de gênero nos pronomes pessoais frequentemente são baseadas em aspectos como sexo/gênero, humanidade, animado/inanimado (para algumas línguas, não é o caso do PB), os quais consistem em uma aglutinação de parâmetros semânticos e formais, que, em termos gerais, são chamados de classe (SIEWIERSKA, 2004). A maior parte da marcação do contraste de gênero é realizada com base no sexo/gênero, por exemplo, os marcadores usados para fazer referência para os homens são masculinos e para as mulheres são femininos (cf. SIEWIERSKA, 2004).

Os pronomes de primeira e segunda pessoas dificilmente são marcados por gênero¹²; as oposições de gênero são propriedades da terceira pessoa (SIEWIERSKA, 2004). Uma justificativa para isso ocorrer está relacionada ao fato de o sexo/gênero, tanto da primeira quanto da segunda pessoa, ser evidente para os membros do eixo interlocutório, tornando-se, assim, redundante qualquer outro tipo de marcação de gênero (SIEWIERSKA, 2004). Quando for necessária essa marcação em um enunciado que haja formas pronominais de primeira ou segunda pessoas, esta será feita por outro elemento, como, por exemplo, por intermédio da flexão do adjetivo. Em contrapartida, a marcação de gênero da terceira não é óbvia, visto que nem sempre a terceira pessoa está presente no contexto extralinguístico do ato comunicativo. Realizar a marcação de gênero da terceira pessoa, portanto, contribui para que os interlocutores possam identificar qual é o referente que está em foco. Há línguas que codificam oposição de gênero na terceira pessoa com formas para singular (ele/ela) e plural (eles/elas), formas distintas para os traços +/- humano e +/- animado (cf. SIEWIERSKA, 2004). No caso específico do

markers as in the case of nouns. A singular noun refers to a single token of the entity denoted by the noun and a plural noun refers to multiple tokens of the relevant entity; thus books refers to more than one instance of the class of objects called book. Plural third persons are interpreted in an analogous way, that is they refer to third parties consisting of several individuals or items. Plural first-person forms, by contrast, only very rarely refer to more than one speaker. The English we may identify several or more speakers in a swearing-in ceremony or some other special occasion when a number of people are actually speaking simultaneously. In all other instances we does not identify more than one speaker but rather the speaker and somebody else. This somebody else may be just the addressee[...] some other individual or group of individuals and the addressee [...] or some individual or group of individuals among which the addressee is not included [...] (SIEWIERSKA, 2004, p. 82).

¹² Siewierska (2004) afirma que são poucas línguas que fazem marcação de gênero na primeira e segunda pessoas. A autora ressalta que a maior parte das línguas que marcam gênero na segunda pessoa é do Norte da África e que, na primeira pessoa, são muito raras as línguas que fazem essa marcação.

português, este não apresenta formas pronominais de terceira pessoa distintas para expressar referentes com os traços +/- humano e +/- animado.

Em termos gerais, as diferenças ressaltadas entre as três pessoas do discurso decorrem principalmente:

[...] do fato de que as formas de primeira e de segunda pessoas são expressões inerentemente dêiticas, ou seja, sua interpretação depende das propriedades do contexto extralinguístico do enunciado em que ocorrem. Embora a primeira pessoa seja sempre o falante da expressão e a segunda o ouvinte, a identidade real de cada uma depende de quem pronuncia a expressão que as contém a quem, quando e onde. Eles pertencem à classe de expressões frequentemente denominadas shifters (Jakobson, 1971). As formas de terceira pessoa, por outro lado, são expressões essencialmente anafóricas. Sua interpretação depende não do contexto extralinguístico, mas do contexto linguístico da expressão¹³ (SIEWIERSKA, 2004, p. 7, tradução nossa).

As três categorias não são interdependentes, pelo contrário. Uma das regras universais, que foram propostas por Greenberg (1963 apud HARLEY; RITTER, 2002, p. 483, tradução nossa), evidencia a dependência existente, por exemplo, entre a categoria número e gênero, a saber: “Universal 32: 'Sempre que um verbo concorda com um sujeito ou objeto nominal em gênero, ele também concorda em número’”¹⁴. Nesse âmbito, três situações são possíveis: i) se o sujeito da oração é singular e a concordância verbal é singular, é expresso um número singular; ii) se o sujeito é plural e a concordância verbal também é plural, é expresso um número plural; e iii) se o sujeito é plural e a concordância verbal é singular, é expresso um número duplo (HARLEY; RITTER, 2002, p. 483), como, por exemplo, em português, os usos de **nós vai, eles/elas vai**; podemos considerar o inverso como em **a gente vamos**.

Demonstramos, nesta seção, que os pronomes pessoais não codificam apenas a categoria gramatical pessoa. Dentre as possibilidades, estão a categoria número - a qual está intimamente ligada à noção de pessoa -, e gênero (cf. SIEWIERSKA, 2004). A categoria número é a mais comum, além de que, na interação com a categoria pessoa, é considerada a mais complexa (cf. SIEWIERSKA, 2004). Na próxima subseção, retomamos a noção de pessoa gramatical, a qual se caracteriza como central para os pronomes pessoais, a fim de aprofundar a discussão a respeito da expressão da 2ªPS.

¹³ No original: [...] of the fact that first- and second-person forms are inherently deictic expressions, that is their interpretation is dependent on the properties of the extralinguistic context of the utterance in which they occur. Although the first person is always the speaker of the utterance and the second the hearer, the actual identity of each depends on who utters the utterance that contains them to whom, when and where. They belong to the class of expressions often referred to as shifters (Jakobson 1971). Thirdperson forms, on the other hand, are essentially anaphoric expressions. Their interpretation depends not on the extralinguistic but on the linguistic context of the utterance (SIEWIERSKA, 2004, p. 4).

¹⁴ No original: Universal 32: 'Whenever a verb agrees with a nominal subject or object in gender it also agrees in number (1963 apud HARLEY; RITTER, 2002, p. 483).

2.2 A CATEGORIA PESSOA E A EXPRESSÃO DA 2ªPS

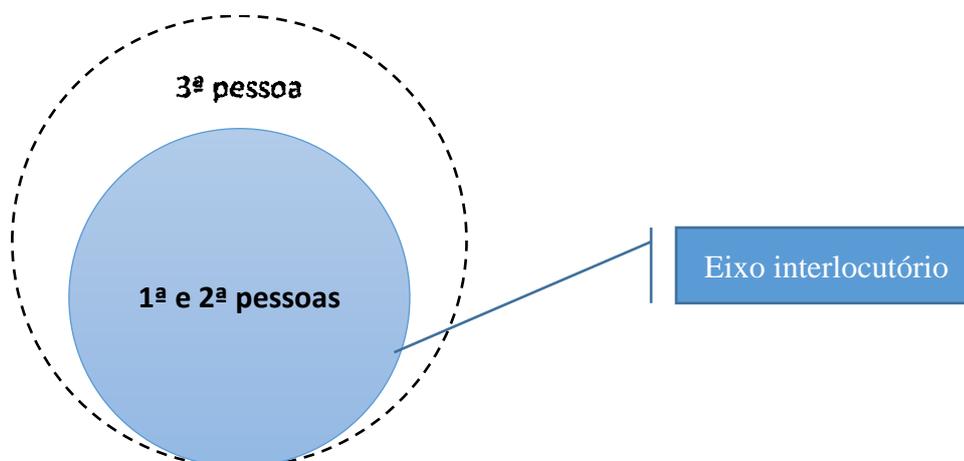
Benveniste (2006 [1989]) parte da premissa de que a língua possibilita o funcionamento subjetivo e referencial do discurso. Esse duplo funcionamento é essencial para a distinção, postulada pelo autor, entre o **eu** e o **não eu**, assim como entre a pessoa e a não pessoa. É por meio da enunciação¹⁵ que se cria um “locutor” e que, simultaneamente, concebe-se um “interlocutor/alocutário”. O **eu** indica o indivíduo locutor da enunciação, enquanto o **tu** denota o alocutário da enunciação. O referente de **eu/tu** é sempre atual (agora) e “só pode ser identificado na instância do discurso”, isso ocorre pelo fato de o **eu** ser “o centro da enunciação” e ser a partir deste (**eu**) que se instaura o **tu**, isto é, “só existe o tu em função de um eu” e vice-versa (BENVENISTE, 1995 [1956], p. 250).

Tradicionalmente, a noção de pessoa está associada às categorias dos pronomes pessoais e do verbo. A instauração da categoria de pessoa ocorre nas trocas sociais e discursivas. A distinção entre pessoa e não pessoa é operada com base na participação ou não da cena enunciativa. Logo, o **eu** e o **tu** são considerados pessoas já que estes fazem parte da situação interlocutiva. A não pessoa é o elemento que não participa da enunciação, ou seja, não é o locutor e nem o alocutário; esta é representada pelo **ele**.

A primeira pessoa (**eu**) e a segunda pessoa (**tu**) possuem uma dependência discursiva (por isso estão dentro do eixo interlocutivo representado na figura 1), pois dependem dos papéis discursivos desempenhados no momento da conversação, isto é, a referência é não fixa. Isso ocorre em virtude do caráter inversível existente entre **eu** e **tu**: o **eu** pode se tornar **tu** e o **tu** pode se tornar **eu**. Quanto à terceira pessoa, esta não se encontra dentro da situação interlocutiva e sua referência é fixa. Enquanto a marcação de primeira e de segunda pessoas só pode ser feita por formas pessoais, a referência à terceira pode ser realizada por qualquer expressão lexical (SIEWIERSKA, 2004). Dessa forma, existe uma diferenciação entre essa tríade: apenas a primeira e a segunda pessoas fazem parte do eixo pessoa/interlocutivo, a terceira pessoa está fora desse eixo - por esse motivo, nos termos de Benveniste (1995 [1956]), é chamada de não pessoa -, e pode fazer parte ou não do contexto discursivo (por isso, na figura 1, o pontilhado no círculo do qual faz parte).

¹⁵ A enunciação é definida pelo autor como o “colocar em funcionamento a língua por um ato individual de utilização”, consiste no ato de dizer, na passagem da língua para fala (BENVENISTE, 2006 [1989], p. 82).

Figura 1 - Correlação entre as pessoas do discurso e o eixo interlocutivo



Fonte: Elaboração própria.

É por meio da categoria gramatical de pessoa que os distintos papéis desempenhados pelas “pessoas” no evento de fala são demarcados, isto é, quem é o locutor (**eu**), o alocutário/interlocutor (**tu**) e o outro (**ele**) (uma ou mais terceira pessoa) (LEVINSON, 2004). A marcação de pessoa no paradigma pronominal pode ser feita a partir de dois aspectos semânticos: inclusão do falante (+S) e inclusão do destinatário (+A) (cf. quadro 2). A 1ªPS é caracterizada pelo papel (+S) e a 2ªPS pelo papel (+A) e pela ausência do papel (-S). Já a terceira pessoa não possui nenhuma das distinções semânticas mencionadas, sendo caracterizada pelo (-S), (-A) e por (+O) (“O” significa outro, ou seja, uma ou mais terceira pessoa) (LEVINSON, 2004).

Quadro 2 - Recursos semânticos de inclusão do paradigma de pessoa

1ª pessoa	(+S)
2ª pessoa	(-S, +A)
3ª pessoa	(-S, -A, +O)

Fonte: Elaboração própria com base em Levinson (2004).

Contrariamente à tradição, que considera três pessoas do discurso, Benveniste (1995 [1956]) propõe a existência da pessoa subjetiva (**eu**), da pessoa não subjetiva (**tu**) e da não pessoa (**ele**). Para o autor, a noção de “pessoa” é própria de **eu/tu** apenas, isto significa que há somente duas pessoas e uma não pessoa. A não pessoa (**ele**) é o “conteúdo representativo da enunciação”, sendo um “elemento constitutivo da relação interpessoal” (BENVENISTE, 1995 [1956], p. 250). Sem a não pessoa não se institui o **eu** e o **tu**, visto que esses se constituem em função de um **ele**, ou seja, do que é dito. A não pessoa (**ele**) abarca “realmente uma indicação

de enunciado sobre alguém ou alguma coisa, mas não referida a uma ‘pessoa’ específica” (BENVENISTE, 1995 [1956], p. 250). Nas palavras do autor:

Nas duas primeiras pessoas, há ao mesmo tempo uma pessoa implicada e um discurso sobre essa pessoa. *Eu* designa aquele que fala e implica ao mesmo tempo um enunciado sobre o “eu”: dizendo eu, não posso deixar de falar de mim. Na segunda pessoa, “tu” é necessariamente designado por *eu* e não pode ser pensado fora de uma situação proposta a partir do “eu”; e, ao mesmo tempo, eu enuncia algo como um predicado de “tu”. Da terceira pessoa, porém, um predicado é bem enunciado somente fora do “eu-tu”; essa forma é assim exceptuada da relação pela qual “eu” e “tu” se especificam. Daí, ser questionável a legitimidade dessa forma como “pessoa” (BENVENISTE, 1995 [1956], p. 250).

Constata-se que o cerne da categoria pessoa é a oposição gerada entre os elementos do discurso **eu/tu** e **ele**. Apenas a primeira e a segunda pessoas fazem parte do eixo interlocutivo. Essas têm seus papéis discursivos bem definidos desempenhados por seres que possuem o traço + humano, podendo ser inversíveis e intercomplementares com os membros da interação verbal. A terceira pessoa é caracterizada como aquela que comporta seres animados e inanimados não implicados na situação interlocutória, isto é, está fora da interação verbal, não comunga, assim, da troca de papéis no ato discursivo. A primeira e a segunda pessoas só podem ser definidas na situação discursiva, enquanto a terceira pode ser definida fora desta. Assim, a inversibilidade, atributo de **eu** e **tu**, não é possível no caso da não pessoa (**ele**). Em termos gerais, o quadro a seguir resume as principais características para essa tríade:

Quadro 3 - Características gerais das **pessoas do discurso** e da **não pessoa**

Características	1ª pessoa	2ª pessoa	Não pessoa
Faz parte do eixo interlocutório	✓	✓	
Define-se apenas na situação discursiva	✓	✓	
Pode não ser definida na situação discursiva			✓
Refere-se apenas a seres humanos	✓	✓	
Pode referir-se a seres animados (humanos, animais) e a inanimados			✓
Os papéis discursivos são inversíveis e intercomplementares	✓	✓	

Fonte: Elaboração própria a partir de Benveniste (1995 [1956]) e Wiese e Simon (2002).

A expressão da 2ªPS é realizada quando um locutor se dirige a um interlocutor que está presente na situação interlocutiva, uma vez que estes (**eu** e **tu**) só existem quando são atualizados na instância do discurso (cf. BENVENISTE, 1995 [1956]). Existe uma relação

intrínseca entre o **eu-tu**, à medida que o **eu** se constitui locutor, simetricamente, constitui-se um **tu** alocutário. Tais particularidades evidenciam que a expressão da 2ªPS possui um referente dêitico que consiste no alocutário (também chamado de receptor, ouvinte, destinatário, interlocutor). Isso significa que este referente está presente na situação interlocutiva, isto é, o referente é determinado.

Além da forma **tu**, outros pronomes podem ser utilizados para fazer referência à segunda pessoa do discurso, como a forma **você** como exemplificado em (6). Convencionalmente, a forma **você** é classificada como pronome de tratamento pelo fato de esta ter se originado da expressão nominal de tratamento **Vossa Mercê** (cf. LOPES, 2007; FARACO, 2017 [1996]). Em decorrência da dinamicidade e heterogeneidade da língua, a forma **você** passou a ser usada como pronome pessoal em referência à 2ªPS, a qual é feita, de acordo com a perspectiva tradicional, pelo pronome pessoal **tu**. Há também, para fazer referência a tal pessoa do discurso, as formas **ocê** (como em (7)) e **cê** (como em (8)), reduções do pronome **você**.

(6) D.S.: qual foi o motivo que **você** levou pra... fazer a graduação?
(D.S.cdt D.M.sdt P MM 01)

(7) Inf.: “(...) mais tarde... à noite que ele chegou... FULANO... onde que **ocê** FOI fulano?... ah... fui dá uma volta por aí (...)
(GONÇALVES, 2008, p. 112, adaptado)

(8) D.M.: **cê** já saiu chorando com as brincadeiras?
(D.S.cdt D.M.sdt P MM 01)

(9) DOC.: você considera Itabaiana uma cidade segura?
Inf.: (est)... não... por que se **você** sair na rua com seu celular falando com alguém os donos vêm tomar... e quando **você** não tem é capaz de tomar uma bala ou uns tapas na cara... uma conhecida... ela tava sem celular eles levaram o cabelo dela... por que ela tava de cabelo solto ela tinha o cabelo muito grande eles passaram o canivete e levaram... aí **cê** se pergunta... que cidade nós estamos?... a segurança daqui é péssima
(01ent.UFS-Itabaiana2018_desl. I_início_ray.fs.19)

No entanto, essas formas de 2ªPS podem ser usadas como estratégia de indeterminação (como em (9)), ou seja, com referência genérica. Em virtude disso, é necessário que se faça uma distinção entre o que é forma e função, pois, partindo da premissa de Benveniste (1995 [1956]) de que a segunda pessoa se constitui na cena enunciativa, defendemos que os usos impessoais e genéricos desses pronomes não fazem de fato referência à segunda pessoa do discurso. É justamente sobre esse aspecto que discutimos na subseção a seguir.

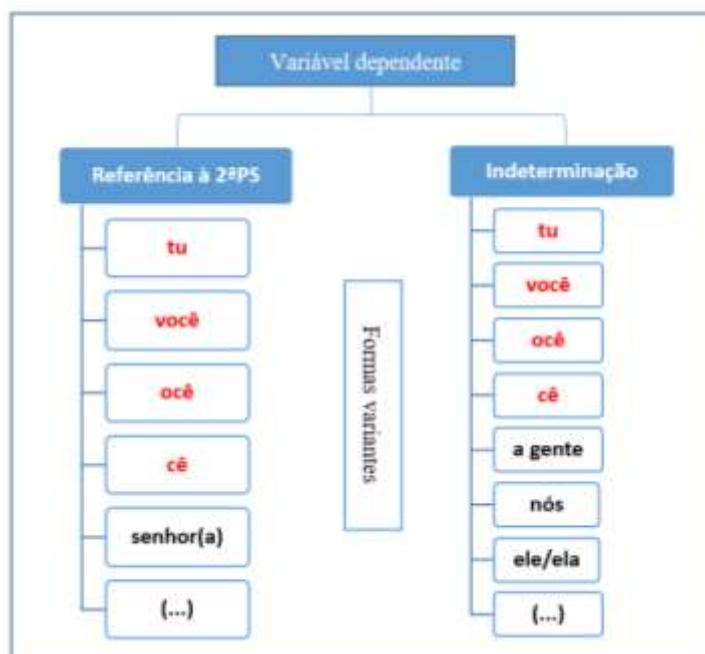
2.3 A REFERÊNCIA À 2ªPS E À INDETERMINAÇÃO SUJEITO: RELAÇÃO ENTRE FORMA E FUNÇÃO

O estudo sobre variação e mudança linguística pode ocorrer a partir de dois direcionais: partindo da forma para a função ou da função para forma¹⁶. O primeiro direcional é o escopo da Teoria Funcionalista/da Gramaticalização (HOPPER, 1991; HOPPER; TRAUGOTT, 1993), já o segundo é o escopo da Teoria da Variação e Mudança Linguística/Sociolinguística Variacionista (WEINREICH; LABOV; HERZOG, 2006 [1968]; LABOV, 2008 [1972]). A visão funcionalista tem como preocupação central o estudo de uma determinada forma que exerce diferentes funções (possui mais de um significado), a fim de identificar quais as motivações de uso em cada função exercida. A abordagem variacionista tem como foco o estudo de um significado (função) e as distintas formas usadas para codificá-lo, com intuito também de verificar as variáveis condicionantes de uso. Nesse último âmbito, estão imbricadas as noções de variável e de variantes linguísticas. Tem-se uma variável linguística quando duas ou mais formas são intercambiáveis com o mesmo valor de referência, representacional ou conteúdo informativo em um dado contexto (WEINREICH; LABOV; HERZOG, 2006 [1968]); isso significa que uma variável linguística consiste em uma função. Variantes linguísticas são os modos alternativos (formas distintas) de se dizer algo com o mesmo significado referencial, representacional ou conteúdo informativo (LABOV, 2008 [1972]).

As formas **tu**, **você** e **cê** são usadas, no PB, para desempenhar a função de referência à 2ªPS (determinada) e podem ser utilizadas como formas genéricas para expressar a função de indeterminação do sujeito (cf. figura 2). As referidas formas caracterizam-se como variantes linguísticas, enquanto as duas funções mencionadas caracterizam-se, cada uma, como variável linguística (também chamada de variável dependente). Assim, há sobreposição de formas para uma mesma função e sobreposição de funções para uma mesma forma.

¹⁶ Para uma discussão mais aprofundada sobre essa questão, ver Freitag e Gonçalves (2011).

Figura 2 - Relação entre as funções de referência à 2ªPS e de Indeterminação do sujeito e as formas pronominais variantes



Fonte: Elaboração própria.

A função de referência à 2ªPS é expressa quando as formas pronominais **tu**, **você**, **cê** e dentre outras são usadas para se referir ao interlocutor/alocutário, o qual está presente na situação de interlocução, como exemplificado em (10) e (11). Em (10), temos um trecho de uma entrevista entre uma documentadora e uma informante, as quais são amigas desde a adolescência, mas, na época da gravação, não tinham contato frequente. Observa-se que a documentadora utilizou a variante **você** para se referir à informante, enquanto a informante, dando continuidade ao tópico discursivo “curso de graduação”, usou a variante **tu** para se dirigir à documentadora. Já em (11) temos um excerto em que a documentadora utiliza a variante **cê** para se referir à informante ao abordar o tópico discursivo voltado para “lembranças da infância”.

- (10) Doc.: mas tá gostando do curso?
 Inf.: eu gosto do curso apesar das dificuldade né? (hes) mas eu gosto.
 Doc.: se arrependeu não?
 Inf.: não até tentei um tempo trocar de curso e não se identifiquei com outro até perdi mais de dois anos nisso nu-num-num tive coragem de trocar ((risos))
 Doc.: eita! **você** fez quantas disciplinas assim?
 Inf.: umas dez ((risos))
 Doc.: de qual curso?
 Inf.: de Pedagogia ... (hes) **tu** acredita Andréia? perdi minha fia dois anos (hes) quase e depois quando saiu a vaga não tive coragem
 (33ent.UFS-Itabaiana2018_desl. II_final_jos.fs.30)

- (11) Doc.: **cê** lembra assim quem era seu melhor amigo assim no-pa-na-na adolescência? ou na infância?

(33ent.UFS-Itabaiana2018_desl. II_final_jos.fs.30)

O uso das formas **tu**, **você** e **cê** com referência à 2ªPS também acontecem em discursos reportados, o qual ocorre quando o falante menciona uma conversa realizada em um momento passado, com um outro interlocutor, como se esta estivesse ocorrendo na situação interlocutiva atual. No exemplo (12), temos um trecho que apresenta o uso da forma **você** é feito com referência à 2ªPS em um discurso reportado.

- (12) Inf.: [...] o Professor falou... “você não sabe se eles têm preparo pra estar na faculdade... se **você** não... deixar eles entrarem na faculdade... (est)... é como **você** falar... que uma pessoa não tem capacidade de dirigir sem ter dado o carro pra ela dirigir”

(16ent.UFS-Itabaiana2018_desl. I_final_eri.ms.22)

Quando o uso dos pronomes **tu**, **você** e **cê** é feito para se referirem não apenas à pessoa com quem se fala, mas também, dentre outras possibilidades, às pessoas em geral, temos a expressão da função de indeterminação. Neste caso, nem sempre o interlocutor está incluso nessa referência, conforme discutiremos mais à frente. Em (13), temos um exemplo de uso das formas **cê** e **você** com referência indeterminada. Observa-se que, na situação retratada, as formas destacadas se referem a qualquer pessoa, pois as pessoas, de modo geral, só passam a identificar pontos positivos ou negativos de um amigo quando o conhecem de fato.

- (13) C.A.: você já chegou a avaliar as competências assim positiva ou negativa alguma vez num amigo seu?

W.S.: ah já... sempre isso aí é constante... e principalmente quando **cê**... com a convivência que **cê** vai passando com certo colega **você**... aprende sabe qual é os valores dele... e existe nele... quando **você** não conhece é mais difícil **você** aquela pessoa é boa pessoa tudo mas depois que **cê** passa a conviver... é que **você** vai saber quais são os pontos positivos e os negativos dele

(C.A._{cdt} W.S._{sdt} P M_M 17)

Percebe-se, por intermédio desse exemplo, que os usos desses pronomes pessoais com traços semânticos genéricos não fazem referência unicamente ao interlocutor e sim a um quantitativo de pessoas mais abrangente. Neste caso, o interlocutor é um possível exemplo, isto é, a referência não está voltada tão somente a ele, como já apontaram Corradello (1997), Sousa (2008) e Nascimento (2011) em seus estudos.

O uso das formas **tu**, **você** e **cê**, para expressar indeterminação, pode funcionar também, em algumas situações, como uma estratégia de preservação de face do locutor, ou melhor, como estratégia de polidez (discutimos sobre esse assunto na seção 3 desta tese), por exemplo, quando

o falante/locutor as utiliza para referir-se a si mesmo. No trecho a seguir, exemplo (5) deste texto rerepresentado em (14), os usos da forma **você**, os quais estão destacados, foram utilizados pelo falante para fazer referência a si mesmo, visto que este está discorrendo sobre como ele se comporta quando está em algum estabelecimento e precisa esperar para ser atendido. Trata-se de um comportamento particular do informante (ficar em uma determinada posição, por exemplo, encostado na parede do estabelecimento de forma a obter uma melhor visibilidade de quem entra no estabelecimento), pois, provavelmente, não é o mesmo seguido por muitas pessoas.

- (14) C.A.: a imagem... não a geralmente assim né? quando é um conhecido eu converso bastante porque a fila demora bastante pra **cê** ser atendido né? eles também não utilizam uma forma de distrair o cliente (est) já que tem poucos funcionários questão de de enxugar... reduzir os gastos... se for uma pessoa desconhecida não não não sou de puxar muita conversa não... sou mais na minha agora se for uma pessoa conhecida eu converso bastante um pouco... presto atenção às vezes eu tô numa numa fi-... (inclusive) uma casa lotérica né? eu não fico virado assim pra frente eu geralmente eu fico viro de costa né? pra o atendente ou na parede encostado... que aí eu percebo as pessoas que entra que sai... porque **você** assim olhando pra o caixa **você** não sabe quem tá atrás de **você** (est) e geralmente eu gosto de ficar na parede... olhando dos dois lados né? já já aconteceu de pessoas inclusive uma professora lá em Frei Paulo... tava no Banco e retirou o dinheiro... e quando saiu fora e um casal acompanhou ela né? e daí a pouco ela voltou e disse que tinha sido roubada... questão de saidinha no banco aquele pessoal entra... às vezes é alguém que você menos... pensa que né? vai lhe roubar

(D.M._{cdt} C.A._{sdt} D M_M 11)

A partir dos exemplos (13) e (14), nota-se que os usos das formas pronominais em questão com valor genérico não possuem os mesmos traços semânticos, pois aspectos discursivos e pragmáticos atuam no tipo de indeterminação, fazendo com que haja matizes/ graus distintos de indeterminação (cf. CORRDELLO, 1997; NASCIMENTO, 2011; SILVA, 2015; COSTA, 2016). Com base em pistas inferenciais, tanto discursivas quanto pragmáticas, a função de indeterminação expressa pelos pronomes **tu**, **você** e **cê** pode ser dividida em dois ou mais subtipos (NASCIMENTO, 2011). Corradello (1997) apresentou quatro usos da forma **você** com matizes diferentes de indeterminação: i) referência a interlocutor + todo mundo; ii) referência a locutor + grupo específico – interlocutor; iii) referência ao próprio locutor; iv) referência a uma terceira pessoa. Nascimento (2011), ao analisar a variação entre **você** e **cê**, reelaborou a proposta de graus de indeterminação do pronome **você** feita por Corradello (1997), expandido-a de quatro para oito tipos: i) 1ª pessoa do singular (oculta); ii) 1ª pessoa do singular (expandido); iii) 1ª pessoa do plural exclusiva; iv) 1ª pessoa do plural inclusiva (+definida); v) 1ª pessoa do plural inclusiva (-definida); vi) 3ª

pessoa do singular ou do plural; vii) Híbrida; e viii) Expletivo (cf. quadro 4). Tal proposta não se aplica apenas as formas variantes **você** e **cê**, mas também a forma **tu**.

Quadro 4 - Graus de indeterminação dos pronomes de 2ªPS

Referência	traço semântico referencial/ Pessoas	Descrição
1ª pessoa do singular (oculta)	+1ª p, - 2ª p, - 3ª p	referente ao próprio locutor
1ª pessoa do singular (expandida)	+1ª p, [+ 2ª p?], + 3ª p	referente ao locutor e a qualquer pessoa/locutor é um exemplo em meio as pessoas em geral
1ª pessoa do plural exclusiva	+1ª p, - 2ª p, + 3ª p	referente ao locutor + grupo específico de indivíduos, no qual o interlocutor não está incluso
1ª pessoa do plural inclusiva (+definida)	+1ª p, + 2ª p, + 3ª p	referente às pessoas em geral
1ª pessoa do plural inclusiva (-definida)	+1ª p, + 2ª p, + 3ª p	referente às pessoas em geral
3ª pessoa do singular ou do plural	- 1ª p, - 2ª p, + 3ª p	referente à terceira pessoa
Híbrida	+1ª p, [+ 2ª p?], + 3ª p	referente híbrido por indicar simultaneamente duas leituras: interlocutor (referência específica) e qualquer pessoa (referência genérica)
Expletivo	-----	sem referência

Fonte: Elaboração própria a partir de Corradello (1997) e Nascimento (2011).

Os pronomes **tu**, **você** e **cê** podem ser utilizados para fazer referência à primeira pessoa do singular como forma de generalização desta. A identificação desse tipo de uso é feita por meio da presença de algum elemento no discurso que referencie o próprio falante, como, por exemplo, pronomes de primeira pessoa (eu, me, mim, meu, etc.), verbos conjugados na primeira pessoa do singular ou alguma evidência referencial que esteja implícita na situação (tais como julgamentos, avaliações, sentimentos, ponto de vista do próprio falante etc.) (NASCIMENTO, 2011).

Nascimento (2011) classificou a referência à primeira pessoa do singular, por meio do uso das formas pronominais de 2ªPS, em duas categorias: primeira pessoa “oculta” (em que há “ocultamento do eu” (CORRADELLO, 1997)) e primeira pessoa “expandida” (cf. quadro 4). A autora ressalta que a distinção entre essas duas categorias está no fato de apenas a referência à primeira pessoa oculta ser utilizada como estratégia de preservação de face do locutor. Neste caso, algo particular do locutor é generalizado com intuito de proteger a face ou de não se comprometer diretamente com o que está sendo dito. O “ocultamento do eu” ocorre quando o locutor não quer expressar diretamente com marcas linguísticas algo particular dele, que o envolve, como ele faz/pensa sobre algo, por isso opta por utilizar uma referência indeterminada. Portanto, esse tipo de referência genérica possui apenas o traço semântico de primeira pessoa

do singular (+1ª p, - 2ª p, - 3ª p) (cf. quadro 4). Em (15), temos um exemplo de uso da variante **cê** como uma estratégia de indeterminação para realizar “ocultamento do eu”, visto que o falante está relatando algo que ele fez. É possível verificar o uso da forma **eu** na primeira oração e a conjugação do verbo na 1ªPS na última oração, elementos estes que funcionam como pistas de que o uso da forma **cê**, na segunda oração, foi feita no lugar da forma de 1ªPS. Em (16), também temos o uso das formas **você** e **cê** também como estratégia de generalização para preservar a face do locutor, fato este constatado a partir de pistas implícitas (consistem no contexto retratado na avaliação e na percepção expressas pelo falante).

(15) INF.: Às vezes **EU** até parece que tirei um cochilo. Parece que **cê** encostou e tirou um cochilo! Tô desse jeito, Van. (Jéssica-F1FI-2007)
(Exemplo (87) de Nascimento (2011, p. 110, grifo da autora))

(16) INF.: Tinha um cara... o cara era chinês, mas **cê** falando/ **cê** falando/[hes.] **você** falando com ele... ah **você** tinha vontade de mata(r) o rapaz. (José-M3SI-2006)
(Exemplo (89) de Nascimento (2011, p. 110, grifos da autora))

A referência à primeira pessoa do singular expandida é expressa em momentos que o falante utiliza as formas pronominais de 2ªPS para fazer remissão a si mesmo, colocando-se como “um exemplo em meio às pessoas em geral (daí a denominação P1 ‘expandido’)” (NASCIMENTO, 2011, p. 111), ou seja, algo de cunho geral é particularizado. Isso ocorre quando temos uma situação que acontece com as pessoas de forma geral, mas que, no contexto interlocutivo, está retratando algo que ocorreu com o locutor, como podemos observar no exemplo (17). Neste caso, a pessoa referida, por meio do uso de um pronome de 2ªPS, está presente na situação de interlocução, no entanto, não é a segunda pessoa/o interlocutor, mas sim a 1ªPS; o próprio falante consiste em um representante dentre outras pessoas que poderiam, por exemplo, vivenciar uma determinada situação. Os traços semânticos de pessoa envolvidos neste tipo de indeterminação são a primeira pessoa e a terceira pessoa, a segunda pessoa pode ou não estar incluída na referência (+1ª p, [+ 2ª p?], [+ 3ª p]).

(17) INF.: ... é bom também... pra **você** estuda(r) ... porque aparece palavra que **cê** nunca viu na tua vida e (isso) vai te forçar ao quê? pra **você** forma(r) outra **cê** vai te que pegar o dicionário é o que **EU** acostumei. (Silmará-F3FI-2007)
(Exemplo (90) de Nascimento (2011, p. 110, grifos da autora))

As formas pronominais de 2ªPS também podem ser utilizadas para fazer referência à primeira pessoa do plural (NASCIMENTO, 2011). Esse uso tem matizes referenciais distintas e, em virtude disso, pode ser dividido em dois subtipos: com referência exclusiva e com referência inclusiva (NASCIMENTO, 2011). A referência à primeira pessoa do plural exclusiva acontece quando as formas pronominais em questão são utilizadas para se referirem ao próprio

locutor/falante e a um grupo específico de pessoas do qual o interlocutor não faz parte, apresentando, dessa forma, os traços semânticos [+1ª, -2ªp, +3ªp]. Em (18), observa-se que o locutor utilizou a forma **você** para fazer referência a si mesmo e a um grupo de pessoas que vivenciaram a situação retratada, da qual o interlocutor não faz parte.

- (18) INF.: ANTIGAMENTE tinha-se educação... **VOCE** dava a benção pro teu pai dava benção pra suas tias... né?... **VOCE** frequentava a igreja ia na missa. [VOCE = EU + as pessoas de antigamente] (Marco-M3FI-2007)

(Exemplo (93) de Nascimento (2011, p. 111, grifos da autora))

A referência à primeira pessoa do plural inclusiva é expressa quando as formas pronominais de 2ªPS são usadas para fazer menção às pessoas de forma geral, ou seja, ao próprio locutor, ao interlocutor e a todo mundo [+1ª, +2ªp, +3ªp]. Nascimento (2011) subdividiu esse tipo de referência considerando os traços [+/- definido]. A referência à primeira pessoa do plural inclusiva com o traço [+ definido] ocorre em situações em que a referência do pronome de 2ªPS utilizado se restringe a “um determinado conjunto de pessoas”, como, por exemplo, a uma categoria profissional (professores, porteiro, médicos, feirante etc), “dentro de um universo mais amplo” do qual o interlocutor está incluso (NASCIMENTO, 2011, p. 113). Observa-se, em (19), que “a presença do sintagma nominal ‘porteiro’” delimita o escopo da referência da forma pronominal **cê** – refere-se às pessoas que desejam ou têm a possibilidade de “exercer essa profissão” (NASCIMENTO, 2011, p. 113). Já o uso das formas de 2ªPS para fazer a referência à primeira pessoa do plural inclusiva com o traço [- definido] é feito nos contextos em que o locutor quer se dirigir às pessoas de forma geral, a toda humanidade. Em (20), o uso da forma **você** refere-se a toda humanidade, pois ter saúde é uma necessidade de todos.

- (19) INF.: Tá ficando cada vez pior eu acho né? pra__ se(r) PORTEIRO hoje **CÊ** tem que te(r) o segundo grau... né? (Fabiana-F1SI-2003)

(Exemplo (94) de Nascimento (2011, p. 113, grifos da autora))

- (20) DOC.: o que você acha que hoje seria prioridade na prefeitura?... O que você sente que **VOCE** precisa e que **TODO MUNDO** precisa?

INF.: [que **TODO MUNDO**...A saúde. A saúde eu acho que tem que se(r) principal.Porque sem saúde **VOCE** não vive bem, sem médico, sem... a saúde pra mim é o principal. (Marta-F2FI-2008)

(Exemplo (95) de Nascimento (2011, p. 113, grifos da autora))

A referência à terceira pessoa ocorre quando as formas são empregadas pelo locutor não para se referirem a si mesmo ou ao interlocutor, mas sim a uma terceira pessoa ou mais de uma (singular ou plural) (NASCIMENTO, 2011). A identificação é feita a partir de indícios referenciais, como podemos observar que o falante, ao usar a forma **você**, no exemplo (21), está se referindo aos empresários, os quais não fazem parte da situação interlocutiva.

- (21) Então **O EMPRESÁRIO** que pensa que eh, aumentando... havendo um aumento coletivo dos salários ele vai diminuir o seu lucro..., eu acho que O EMPRESARIADO BRASILEIRO é inteligente, ele não pensa dessa forma. É claro que eh, **TODO BOM, BOM NEGOCIANTE** sabe que se o seu eh, cliente tiver eh, boas condições pra adquiri(r) o seu produto, mais pessoas vão adquirir o seu produto e aí **você** passa a ganha(r) ... te(r) um lucro mais saudável, maior, né? com/ com maior quantidade. (Mauro-M1SI-2003)

(Exemplo (96) de Nascimento (2011, p. 114, grifos da autora))

A referência híbrida é expressa nos casos em que as formas são utilizadas para referir-se, concomitantemente, “ao interlocutor (referência específica) e a ‘qualquer pessoa’ (referência genérica)” (NASCIMENTO, 2011, p. 114). Geralmente esse tipo de referência ocorre em contextos hipotéticos e de exemplificações. Em (22), temos um exemplo de uso das formas **você** e **cê** com referência híbrida, visto que a situação retratada inclui o interlocutor e qualquer outra pessoa.

- (22) Por exemplo, SE você mora em frente uma pracinha, **cê** vê a pracinha detonada, ninguém tem a coragem de pega(r) um enxada e i(r) limpa(r). (Marta-F2F-2008)

(Exemplo (97) de Nascimento (2011, p. 114, grifos da autora))

Os usos expletivos das formas ocorrem quando não há referente. Isso ocorre nos contextos em que as formas funcionam como sujeito de verbos que possuem uma interpretação existencial quando o pronome é omitido (cf. NASCIMENTO, 2011), como exemplificado em (23) por intermédio de sentenças com e sem o uso do pronome e troca de verbos para evidenciar a possibilidade de interpretação existencial.

- (23) a) hoje **você** já **tem** outras instituições aí tão caminhando, tão indo bem, tão evoluindo né? (Fabiana-F1S-2003)

b) hoje ___ **já tem** outras instituições aí tão caminhando, tão indo bem, tão evoluindo né?

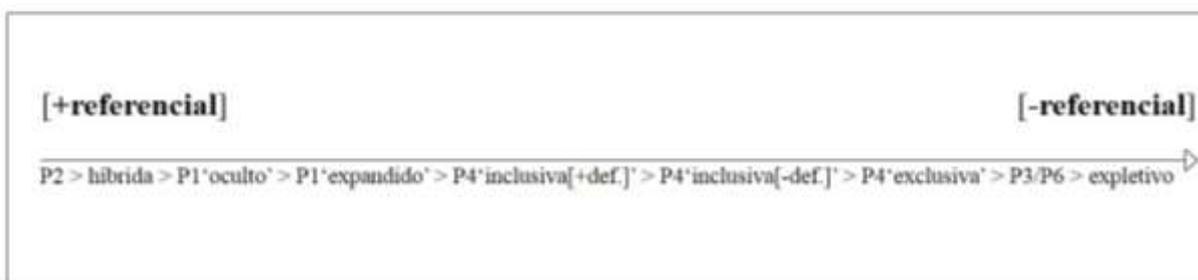
c) hoje ___ **já existem** outras instituições aí tão caminhando, tão indo bem, tão evoluindo né?

d) hoje ___ **já há** outras instituições aí tão caminhando, tão indo bem, tão evoluindo né?

(Exemplo (99) de Nascimento (2011, p. 115, grifos da autora))

Nascimento (2011), para justificar a necessidade de se considerar subtipos de referência genérica, argumenta que a oposição binária da referência (determinada *versus* indeterminada) não é suficiente em uma análise de traços semânticos referenciais de pronomes de 2ºPS por não dar conta da complexidade envolvida no processo. Além disso, a autora defende que se deve controlar a referência das formas de 2ªPS em uma escala, ou seja, em uma gradação referencial: da mais referencial (+específica, -genérica) a menos referencial (-específica, +genérica) (cf. figura 3).

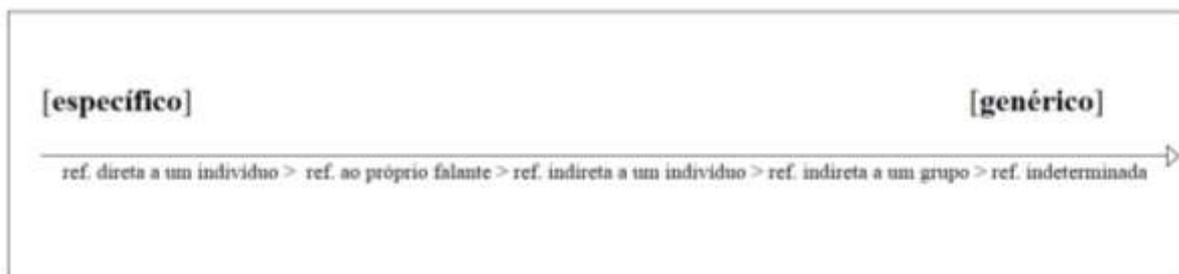
Figura 3 - Escala de traços semânticos referencial de pronomes de 2ªPS



Fonte: Nascimento (2011, p. 117, adaptado).

A identificação desses subtipos de referência genérica não é simples de ser compreendida, pois há sobreposição de traços semânticos referenciais, o que dificulta classificar com precisão cada contexto de uso das formas. A própria pesquisadora deixa transparecer ao mencionar, por exemplo, que há subtipos que se aproximam um do outro e que por isso poderiam ser analisados conjuntamente, como é o caso da referência à primeira pessoa do singular expansiva e da referência à primeira pessoa do plural. Nascimento (2011) realizou quatro análises a partir do controle do tipo de referência considerando os dados de referência específica e de referência genérica. Na primeira análise, os dados de referência genérica foram categorizados nos oito subtipos descritos e nas demais foram amalgamados gradativamente, conforme apresentamos na seção 2.4.

Costa (2016) também propõe, em seu estudo, o controle do tipo de referência dos pronomes de 2ªPS a partir de um *continuum* referencial (cf. figura 4). A autora propõe cinco subtipos de referência (incluindo a referência específica): i) direta a um indivíduo (específica); ii) indireta a um indivíduo; iii) ao próprio falante; iv) indireta a um grupo; e v) genérica/indeterminada. Diferentemente de Nascimento (2011), Costa (2016) controlou separadamente as ocorrências de referência específica em discurso direto das que ocorrem em discurso indireto/reportado. A autora denominou esse tipo de uso de: referência indireta a um indivíduo. Outros subtipos de referência controlados pela pesquisadora foram a referência específica ao próprio falante e a referência indireta a um grupo, os quais correspondem, respectivamente, à referência P1 'oculto' e à referência P4 inclusiva [+definido], denominações propostas por Nascimento (2011). O último subtipo de referência controlado pela pesquisadora foi a indeterminada/genérica, a qual é expressa quando o falante utiliza os pronomes de 2ªPS para referir-se a interlocutor indeterminado (qualquer pessoa), enquadrando-o em uma situação hipotética (COSTA, 2016). Tal tipo de referência se aproxima, aparentemente, da referência híbrida e da referência P4 inclusiva [-definida], controladas por Nascimento (2011).

Figura 4 - *Continuum* do tipo referência dos pronomes de 2ªPS

Fonte: Costa (2016, p. 137, adaptado).

No *continuum* do tipo de referência dos pronomes de 2ªPS, Costa (2016) considerou que a referência indireta a um indivíduo é menos específica que a referência ao próprio falante, conforme apresentado na figura 4. Discordamos dessa proposta da autora pelo fato de esse tipo de referência ser feito ao interlocutor direto do locutor e, mesmo sendo um discurso reportado, consiste em um uso dêitico das formas pronominais de 2ªPS. Portanto, esse subtipo de referência deve ocupar uma posição anterior, dentro do *continuum*, a da referência ao próprio falante.

Essas propostas de controle do traço semântico referencial dos pronomes em estudo são bastante significativas, no entanto, ainda precisam de um refinamento para que se possa realizar análise mais acuradas. No entanto, este não é o propósito da presente tese, uma vez que focalizamos o tipo de referência de forma binária – determinada e indeterminada - e como duas regras variáveis, buscando identificar os efeitos das variáveis independentes nos usos dos pronomes **tu**, **você** e **cê** em cada uma das regras.

2.4 O CONTROLE DA REFERÊNCIA DAS FORMAS TU, VOCÊ E CÊ NOS ESTUDOS SOCIOLINGUÍSTICOS

Os estudos sociolinguísticos têm, em sua maioria, investigado o tipo de referência das formas **tu**, **você** e **cê** como uma variável independente. O controle dessa variável ocorre majoritariamente por meio da oposição binária: traço semântico determinado/específico (expresso quando o referente do pronome faz parte da situação interlocutiva, isto é, o pronome codifica a referência ao interlocutor definido no contexto interacional) vs. indeterminada/genérica (expresso quando há uma abrangência referencial), como pode ser verificado nas pesquisas de Martins (2010), Alves (2010), Nogueira (2013), Silva, Suziane (2017), entre outras. Alguns pesquisadores, que também analisaram a referência como variável

independente, controlaram-na a partir de um *continuum* de indeterminação, conforme discorremos na seção anterior. Temos como exemplo deste caso os trabalhos de Nascimento (2011), Silva, I. (2015) e Costa (2016).

Discutimos, a seguir, os resultados obtidos pelos pesquisadores que controlaram a referência das formas **tu**, **você** e **cê** como uma variável independente, com intuito de evidenciar se as taxas de frequências de tais formas possuem relação com o traço semântico do referente. Para tanto, realizamos o teste de qui-quadrado de Pearson¹⁷, por intermédio da função *chisq.test*, para as análises univariadas. Apresentamos, primeiramente, os resultados dos estudos que controlaram o tipo de referência de forma binária, agrupando-os por distribuição regional. Cabe ressaltar que foram selecionadas as pesquisas realizadas a partir dos anos 2000 em virtude do quantitativo expressivo de estudos desenvolvidos.

Tabela 1 - O uso das formas de 2ªPS em função do tipo de referência na região Norte

Estudos	Tipo de referência	Tu		Você		T	Qui-quadrado
		A	%	A	%		
Martins (2010)	Específico	334	74,7	113	25,3	447	$X^2 (1, N= 806) = 44.65,$ $p < 0.00001$
Tefé/AM	Genérico	186	51,8	173	48,2	359	
Brito, Castro e Furtado (2020)	Referencial	249	67	124	33	373	$X^2 (1, N= 388) = 0.07,$ $p = 0.7913$
Cametá/PA	Genérico	9	60	6	40	15	

Fonte: Elaboração própria.

Identificamos dois estudos realizados na região Norte que controlaram o traço semântico do referente de forma binária: Martins (2010), desenvolvido na cidade de Tefé/AM, e Brito, Castro e Furtado (2020), desenvolvido na cidade de Cametá/PA. Os resultados demonstraram que a frequência de uso da forma **tu** é alta nos dois estudos referidos, havendo taxa de uso maior nos contextos de referência específica, principalmente, no estudo de Martins (2010) (334/447 - 74,7%). Contudo, as análises univariadas realizadas revelam significância estatística distinta. Os resultados, presentes na tabela 1, evidenciam que o efeito dessa variável nos usos das formas **tu** e **você** é estatisticamente significativo ($X^2 = 44.65$, $p < 0,00001$) na pesquisa de Martins (2010), ocorrendo o contrário no estudo de Brito, Castro e Furtado (2020) ($X^2 = 0.07$, $p < 0.7913$). É importante ressaltar que os resultados apontaram que há uma distribuição mais equitativa das formas em contextos de referência genérica, sobretudo, no estudo de Martins (2010), o que confirma a necessidade de se analisar os usos de referência específica separados dos usos de referência genérica.

¹⁷ O nível de significância considerado para os testes de qui-quadrado, neste trabalho, foi de 0.05. É importante ressaltar que realizamos os testes de qui-quadrado pelo fato de muitas das pesquisas consultadas não apresentarem o nível de significância das variáveis.

Tabela 2 - O uso das formas de 2ªPS em função do tipo de referência na região Nordeste

Estudos	Tipo de referência	Tu		Você		Cê		T	Qui-quadrado
		A	%	A	%	A	%		
Alves (2010)* Cidad. do Maranhão	Específico	100	43,1	132	56,9	—	—	232	$X^2(1, N= 328) = 6.70,$ $p = 0.009617$
	Genérico	26	27,1	70	72,9	—	—	96	
Carneiro (2011) São Luís/MA	Determinado	192	71,9	75	28,1	—	—	267	$X^2(1, N= 277) = 20.17,$ $p < 0.00001$
	Indeterminado	0	0	10	100	—	—	10	
Guimarães (2014) Fortaleza/CE	Específico	785	53,3	687	46,7	—	—	147 2	$X^2(1, N= 1555) = 59.57,$ $p < 0.00001$
	Genérico	7	8,4	74	91,6	—	—	83	
Nogueiro (2013) F. de Santana/BA	Específico	74	26	211	74	—	—	285	$X^2(1, N= 822) = 150.07,$ $p < 0.00001$
	Genérico	0	0	537	100	—	—	537	
Nogueiro (2013) Salvador/BA	Específico	9	2,9	301	97,1	—	—	310	$X^2(1, N= 891) = 14.26,$ $p = 0.0001592$
	Genérico	0	0	581	100	—	—	581	
Alves (2015)** São Luís/MA	Específico	761	93,3	55	6,7	—	—	816	$X^2(1, N= 1028) = 243.76,$ $p < 0.00001$
	Genérico	110	51,9	102	48,1	—	—	212	
Silva, Suziane (2017) D. Gouveia/AL	Específico	—	—	204	92	18	8	222	$X^2(1, N= 508) = 1.24,$ $p = 0.2638$
	Genérico	—	—	271	95	15	5	286	

Fonte: Elaboração própria.

Notas: *Dados de **você**, **ocê** e **cê** apresentados juntos pelo(s) autor(es).

Os dados de **tu com concordância e de **tu sem concordância** foram apresentados pelo(a) pesquisador(a) de forma separada, no entanto para fins comparabilísticos, amalgamamos esses dados.

Identificamos seis estudos (Nogueira (2013) analisa duas amostras) sobre a variação das formas pronominais **tu**, **você** e **cê** (somente o estudo de Silva, Suziane (2017) apresenta dados separados dessa última variante), realizados na região Nordeste, que controlaram o traço semântico do referente (tabela 2). Nos contextos de referência específica, os resultados demonstraram que apenas nos estudos de Guimarães (2014) (o qual utilizou dados de fala da cidade de Fortaleza/CE), Carneiro (2011) e Alves (2015) (ambos usaram dados de fala de São Luís/MA) houve o favorecimento do uso da forma **tu** em detrimento da forma **você**. Silva, Suziane (2017), ao analisar dados de fala da cidade de Delmiro Gouveia/AL, constatou o uso apenas das variantes **você** e **cê**, sendo que as ocorrências da variante **você**, tanto no contexto de referência específica quanto no de referência genérica, foram predominantes (92% e 95%, respectivamente).

Quanto aos contextos de traço semântico genérico, a distribuição das frequências evidenciou que houve, majoritariamente, o favorecimento do uso da forma **você** nos estudos consultados, com exceção apenas da pesquisa de Alves (2015), realizada com dados de fala de São Luís/MA, que apresentou taxas de uso das formas **tu** e **você** muito próximas (51,9% e

48,1%, respectivamente). Tal resultado obtido por Alves (2015) diverge do obtido por Carneiro (2011), que foi também desenvolvido com dados de fala de São Luís/MA. As motivações para a obtenção de resultados distintos podem estar pautadas no fato de que o método de coleta utilizado para a obtenção dos dados foi diferente: Alves (2015) analisou dados coletados por meio de gravações de conversas semiespontâneas e Carneiro (2011) analisou dados coletados por intermédio de gravações de entrevistas sociolinguísticas. Entretanto, esperava-se que, em virtude do método utilizado, houvesse uma maior frequência de uso da forma **você**, tanto nos contextos de referência genérica quanto específica, no estudo de Carneiro (2011). Os resultados do estudo desenvolvido por Alves em 2015 também são diferentes daqueles obtidos pela autora em 2010. Embora as duas pesquisas tenham sido realizadas com dados de fala do Maranhão, há distinções quanto às cidades coletadas. No primeiro estudo, os dados foram coletados em diferentes cidades do estado, os quais foram amalgamados nas análises realizadas, o que pode ter gerado resultados enviesados.

Observa-se, portanto, que a forma **você** tem comportamentos diferentes a depender do contexto de uso. Os resultados do teste de associação de qui-quadrado foram estatisticamente significativos para a maioria dos estudos do Nordeste dispostos na tabela 2 (a única exceção é o estudo de Silva, Suziane (2017)), mostrando que há uma relação de dependência entre os usos das formas em questão e o traço semântico do referente.

Tabela 3 - O uso das formas de 2ªPS em função do tipo de referência na região Centro-oeste

Estudos	Tipo de referência	Tu		Você		Cê		T	Qui-quadrado
		A	%	A	%	A	%		
Lucca (2005) Brasília/DF	Específico	291	72	112	28	—	—	403	$X^2(1, N=453) = 2.2689e-31$, $p = 1$
	Genérico	36	72	14	28	—	—	50	
Dias (2007)* Brasília/DF	Específico	109	14,8	628	85,2	—	—	737	$X^2(1, N=900) = 13.79$ $p = 0.0002034$
	Genérico	6	3,7	157	96,3	—	—	163	
Andrade (2010) Brasília/DF	Específico	283	37,5	266	35,3	205	27,2	754	$X^2(1, N=835) = 37.07$ $p < 0.00001$
	Genérico	5	6,2	52	64,2	24	29,6	81	

Fonte: Elaboração própria.

Nota: *Dados de **você** e **cê** rodados juntos.

Constatamos, na região Centro-oeste, que apenas três estudos sobre a variação dos pronomes **tu**, **você** e **cê** controlaram o tipo de referência: Lucca (2005), Dias (2007) e Andrade (2010) (os três utilizaram dados de fala da cidade de Brasília/DF). Os resultados da análise univariada de distribuição das formas variantes em função do traço semântico referencial foram estatisticamente significativos para os dois últimos estudos, no entanto, as taxas de uso apresentaram um padrão de distribuição distinto. Os resultados de Lucca (2005) demonstraram

que, na amostra de conversas espontâneas ocultas analisadas, o uso da forma **tu** é favorecido, tanto em contexto de referência específica (291/403 - 72%) quanto no contexto de referência genérica (36/50 - 72%), evidenciando um efeito “neutro” desta variável nos usos das formas em questão na amostra analisada.

Contrariamente, os resultados de Dias (2007), os quais foram obtidos a partir da análise de dados de conversas espontâneas não ocultas, evidenciaram um uso mais frequente da forma **você** independentemente do tipo de referência (específica (628/737 - 85,2%) e genérica (157/163 - 96,3%)). Andrade (2010) obteve resultados distintos das duas pesquisas mencionadas ao analisar dados de fala coletados a partir de entrevistas sociolinguísticas. A autora controlou as variantes **tu**, **você** e **cê** e constatou que há uma diferenciação no uso das formas a depender do tipo de referência: em contextos específicos, a forma **tu** foi favorecida (283/754 - 37,5%) e, em contextos genéricos, houve o favorecimento da forma **você** (52/81 - 64,2%). Por outro lado, se considerarmos os dados de **você** e **cê** juntos, tem-se um resultado semelhante ao de Dias (2007). Os resultados obtidos nessas pesquisas evidenciam um comportamento de uso das variantes bastante diverso, apontando a necessidade de novos estudos e de um refinamento da análise de associação do fenômeno ao tipo de referência.

Tabela 4 - O uso das formas de 2ªPS em função do tipo de referência na região Sudeste

Estudos	Tipo de referência	Você		Ocê		Cê		T	Qui-quadrado
		A	%	A	%	A	%		
Peres (2006) Belo Horizonte/MG	Definido	108	15,1	36	5,1	569	79,8	713	X ² (1, N= 1453) = 57.04 p < 0.00001
	Indefinido	234	31,6	20	2,7	486	65,7	740	
Gonçalves (2008) Arcas/MG	Definido	55	14	97	26	226	60	378	X ² (1, N= 510) = 49.01 p < 0.00001
	Indefinido	58	44	23	17	51	39	132	
Calmon (2010) Vitória/ES	Específico	302	72,6	—	—	114	27,4	416	X ² (1, N= 1819) = 0.63 p = 0.4257
	Genérico	1048	74,7	—	—	355	25,3	1403	
		Tu		Você		cê		T	
		A	%	A	%	A	%		
Silva, Suelen (2017) Ressaquinha/MG	Específico	199	45,4	239	54,6	—	—	438	X ² (1, N= 579) = 35.17 p < 0.00001
	Genérico	24	17	117	83	—	—	141	

Fonte: Elaboração própria.

Na região Sudeste, os estudos realizados no estado de Minas Gerais, a partir de amostras de entrevistas sociolinguísticas, apresentaram uma distribuição não uniforme das variantes em função do traço semântico do referente (cf. tabela 4). A pesquisa desenvolvida por Peres (2006) identificou o uso das formas **você**, **ocê** e **cê** na fala da capital Belo Horizonte/MG, sendo o uso desta última variante o mais frequente em contextos de referência específica (569/713 - 79,8%)

como também de referência genérica (486/740 - 65,7%). Os resultados obtidos por Gonçalves (2008), com a análise de dados de fala da cidade de Arcas/MG, diferem dos obtidos por Peres (2006) para os contextos de referência genérica, nos quais houve uma taxa de frequência de uso maior da variante **você** (58/132 - 44%). O estudo realizado por Silva, Suelen (2017) com dados de fala da cidade de Ressaquinha/MG, diferentemente das duas pesquisas mencionadas, identificou o uso das variantes **tu** e **você**. Os resultados evidenciaram que, tanto os contextos de referência específica quanto de referência genérica, o uso da forma **você** recorrentes, apresentando, respectivamente, as frequências 54,6% (239/438) e 83% (117/141).

Calmon (2010) desenvolveu sua pesquisa utilizando dados de fala da cidade de Vitória/ES coletados por meio de gravações de entrevistas sociolinguísticas. A autora identificou na amostra o uso das variantes **você** e **cê** e constatou que a primeira variante é mais frequente nos dois tipos de referência controlados (específico (302/416 - 72,6%) e genérico (1049/1403 - 74,7%)). A análise univariada de qui-quadrado evidenciou que apenas os resultados do estudo de Calmon (2010) não são estatisticamente significativos para a escolha das variantes em função do tipo de referência.

Tabela 5 - O uso das formas de 2ªPS em função do tipo de referência na região Sul

Estudos	Tipo de referência	Tu		Você		T	Qui-quadrado
		A	%	A	%		
Loregian-Penkal (2004) cidades do Sul	Determinado	804	91	81	9	885	$X^2 (1, N= 989) = 27.65,$ $p < 0.00001$
	Indeterminado	862	82	183	18	1045	
Loregian-Penkal (2004) cidades do R. G. do Sul	Determinado	460	90	52	10	512	$X^2 (1, N= 1864) = 1.82,$ $p = 0.1773$
	Indeterminado	1191	88	171	12	1352	
Franceschini (2011) Concórdia/SC	Determinado	162	79	42	21	204	$X^2 (1, N= 926) = 1.56,$ $p = 0.2107$
	Indeterminado	350	48	372	52	722	

Fonte: Elaboração própria.

Identificamos dois estudos realizados na região Sul sobre a variação entre as formas **tu** e **você** que controlaram o traço semântico do referente de forma binária: Loregian-Penkal (2004) e Franceschini (2011). As autoras utilizam os termos determinado (referente recuperável) e indeterminado (referente genérico), os quais equivalem, de forma geral, aos termos específico e genérico, respectivamente.

A partir de dados de fala de entrevistas sociolinguísticas, Loregian-Penkal (2004) realizou um estudo a respeito desse fenômeno em algumas cidades da região Sul, as quais foram agrupadas da seguinte forma: cidades do Sul (Porto Alegre/RS, Florianópolis/SC e Ribeirão/SC) e cidades do interior do Rio Grande do Sul (Flores da Cunha, Panambi e São Borja). Os resultados evidenciaram, nos dois grupos de cidade analisados, o favorecimento da

variante **tu**, tanto em contextos determinados quanto indeterminados. Tais resultados devem ser considerados com muita cautela, pois a amalgamação dos dados de falantes provenientes de localidades distintas pode ter provocado enviesamentos, principalmente, pelo fato de o fenômeno estudado ter se mostrado sensível ao aspecto dialetal em muitas pesquisas.

Os resultados da pesquisa desenvolvida por Franceschini (2011), com base na análise de dados decorrentes de entrevistas sociolinguísticas realizadas com falantes de Concórdia/SC, diferenciam-se dos obtidos por Loregian-Penkall (2004) para os contextos de referência indeterminada, visto que o houve um favorecimento da forma **você** (372/722 - 52%). A análise univariada realizada por intermédio do teste de qui-quadrado mostra que a distribuição das formas em função do traço semântico do referente é significativa somente para o primeiro estudo de Loregian-Penkall (2004) exposto na tabela 5.

Tabela 6 - O uso das formas de 2ªPS em função do tipo de referência (não binário)

Estudos	Tipo de referência	Tu		Você		T	Qui-quadrado
		A	%	A	%		
Modesto (2006) Santos/SP	Direta	154	42	209	58	363	X ² (2, N= 708) = 40.40, p < 0.00001
	Indireta	39	33	79	67	118	
	Indeterminada	39	17	188	83	227	
Silva, I. (2015) São José dos Campos/SP	Determinado	14	7	180	93	194	X ² (3, N= 530) = 2.66, p = 0.4461
	Reportado determinado	11	10	93	90	104	
	Eu/self	6	10	53	90	59	
	Genérico	10	5	163	94	173	
Costa (2016) Cametá/PA	Indireta/específica a um indiv.	66	70,2	28	29,8	94	X ² (4, N= 389) = 10.20, p = 0.03715
	Direta/específica a um indiv.	59	72,8	22	27,2	81	
	Ao próprio falante	25	65,8	13	34,2	38	
	Genérico	17	53,1	15	46,9	32	
	Indireta a um grupo	140	57,4	104	42,6	144	
Lopes (2017) Chapecó/SC	Interlocutor	41	57,7	30	42,3	71	X ² (3, N= 268) = 12.76, p = 0.00517
	Particular	33	54,1	28	45,9	61	
	Grupo	28	32,2	59	67,8	87	
	Genérico	20	40,8	29	59,2	49	

Fonte: Elaboração própria.

Modesto (2006), Silva, I. (2015), Costa (2016) e Lopes (2017) optaram por não realizar o controle da referência de modo binário em seus estudos, como se pode observar nos dados dispostos na tabela 6. Modesto (2006) realizou uma pesquisa sobre a variação entre **tu** e **você**, utilizando dados de falantes provenientes da cidade de Santos/SP, os quais foram obtidos por meio de gravação de conversas espontâneas ocultas e conscientes. O autor controlou a variável

tipo de referência a partir de três níveis: referência direta, referência indireta e referência indeterminada. A referência classificada como direta é expressa quando o falante utiliza as formas variantes de 2ªPS para se referir ao seu interlocutor, o qual faz parte da situação interlocutiva. A referência indireta é feita quando o falante reproduz uma situação de interlocução sucedida em uma ocasião antecedente a atual (ou seja, vale-se de um discurso reportado), na qual utiliza as formas variantes de segunda pessoa para se referir ao seu interlocutor presente na situação de conversação. Já a referência indeterminada ocorre nos contextos em que as formas variantes são utilizadas para se remeter a qualquer pessoa, consistindo em uma referência generalizada em que não se é possível identificar quem é o interlocutor de fato, trata-se de um interlocutor “imaginário”. A situação retratada não consiste na situação interlocutiva atual, o que ocorre é “um enquadramento de um interlocutor indeterminado em uma situação contextual hipotética” (MODESTO, 2006, p. 77). Os resultados obtidos por Modesto (2006) evidenciaram que a forma **você** foi mais frequente em todos os tipos de referência controlados, principalmente, em contexto de referência indeterminada (188/227 - 83%).

Silva, I. (2015), ao analisar a amostra de entrevistas sociolinguísticas realizadas com falantes de São José dos Campos/SP, controlou a variável tipo de referência de forma semelhante a Modesto (2006), embora nomeie o grupo de variáveis de forma distinta: referência determinada equivale à referência direta, referência reportada determinada equivale à referência indireta, referência genérica equivale à referência indeterminada. A diferença de controle entre os autores está no fato de Silva, I. (2015) considerar a estratégia referencial eu/self (que é expressa quando o falante utiliza as formas variantes de segunda pessoa para se referir a si mesmo) como um nível à parte, enquanto Modesto (2006) possivelmente (já que o autor não faz remissão a esse uso) considera esse tipo de uso como referência indeterminada. Silva, I. (2015) constatou que a forma **você** é predominantemente favorecida em todos os tipos de referência controlada. Isso se deve ao baixo uso da forma **tu** pelos falantes na amostra analisada. Considerando apenas os dados da forma **tu**, constata-se que ela ocorreu com maior frequência em contextos determinados. Embora tenham usado métodos de coleta distintos, os resultados de Silva, I. (2015), em termos gerais, aproximam-se dos obtidos por Modesto (2006), evidenciando que esse pode ser um padrão de uso nas cidades de grande porte do estado de São Paulo.

Costa (2016) desenvolveu uma pesquisa, a partir da análise de uma amostra de interação face a face de grupo focal de falante da cidade de Cametá/PA, na qual controlou a variável tipo de referência nos moldes de Silva, I. (2015), fazendo o acréscimo do nível referência a um

grupo. Esse tipo de referência é expresso quando o falante utiliza uma das formas de segunda pessoa dirigida “a um interlocutor, que não é singular, único, em termos de especificidade pessoal, mas sim um grupo, que encapsula os interlocutores, tornando-os comuns, por meio de alguns traços sociais”; sendo que esse grupo referenciado é “reconhecido pelos interlocutores, por meio do contexto discursivo, como sendo o interlocutor/ouvinte em particular a quem o falante se direciona” (COSTA, 2016, p. 137-138). Os traços sociais, por exemplo, podem ser: exercer “a mesma profissão”, ser morador “do mesmo bairro, cidade ou país”, possuir “o mesmo nível social ou de escolaridade etc.” (COSTA, 2016, p. 138). Os resultados obtidos pela autora demonstraram o favorecimento da forma **tu** em todos os tipos de referência. É importante ressaltar que o uso da forma **tu** foi fortemente condicionado pela referência direta/específica a um falante (59/81 - 72,8%) e pela referência indireta/específica a um indivíduo (66/94 - 70,2%), ou seja, a forma é favorecida em contextos de referência específica (cf. tabela 6).

Lopes (2017) analisou dados de entrevistas sociolinguísticas realizadas com falantes da cidade Chapecó/SC. O controle da variável tipo de referência foi feito de modo semelhante à Costa (2016): i) referência ao interlocutor – ocorre quando o falante utiliza as formas pronominais para se dirigir à pessoa com quem está falando na situação interlocutiva; ii) referência particular – ocorre quando o falante utiliza as formas para referir a si mesmo; iii) referência a um grupo definido – ocorre quando o falante utiliza as formas para se referir a algumas pessoas do discurso ou a um grupo de pessoas determinado; iv) referência genérica – ocorre quando as formas são utilizadas para se referir a todas as pessoas, a qualquer pessoa. Os resultados obtidos pela autora evidenciaram que a forma **tu** tende a ser favorecida em contextos de referência ao interlocutor (41/71 - 57,7%) e de referência particular (33/61 - 54,1%), sendo desfavorecida, principalmente, em contextos de referência a um grupo específico (28/78 - 32,2%). Tais resultados são, em termos gerais, semelhantes aos obtidos por Franceschini (2011), os quais foram apresentados na tabela 5, e indicam que o uso do **tu** tende a se manter na região Sul para contextos de referência específica. Além disso, sinalizam que há, possivelmente, uma mudança em curso para os contextos de referência genérica, uma vez que a forma **você** aparece nos estudos mais recentes como a mais utilizada pelos falantes, contextos nos quais estes utilizavam, majoritariamente, a forma **tu**, como demonstrou o estudo de Loregian-Penkall (2004).

A análise univariada realizada demonstrou que, para os estudos de Modesto (2006), Costa (2016) e Lopes (2017), os testes de qui-quadrado obtidos equivalem a uma significância estatística maior ou igual a 0,05, assim, os traços semânticos do referente condicionam os usos

das formas variantes. Já para o estudo de Silva, I. (2015) esse controle não se mostrou estatisticamente significativo.

Tabela 7 - Diferentes abordagens da referência do **você/cê**

CÊ	Referência	A/T	%	Qui-quadrado
9 níveis	P2 (Específica)	150/227	66,1	$X^2(8, N= 1859) = 67.42,$ $p < 0.00001$
	P1 ‘oculto’	63/90	70,0	
	P1 ‘expandido’	140/242	57,9	
	P4 ‘exclusivo’	89/195	45,6	
	Híbrido	70/136	51,5	
	P4 ‘inclusivo’ [+ definido]	265/486	54,5	
	P4 ‘inclusivo’ [- definido]	100/217	46,1	
	P3 /P6 (3ª p. sg./pl.)	73/201	36,3	
	Expletivos	13/65	20,0	
	Subtotal	963/1859	51,8	
6 níveis	P2 (Específica)	150/227	66,1	$X^2(5, N= 1859) = 52.11,$ $p < 0.00001$
	P1 + P4 ‘exclusivo’	292/527	55,4	
	Híbrido	70/136	51,5	
	P4 ‘inclusivo’ [± definido]	365/703	51,9	
	P3 /P6 (3ª p. sg./pl.)	73/201	36,3	
	Expletivos	13/65	20,0	
	Subtotal	963/1859	51,8	
3 níveis	Específica	150/227	66,1	$X^2(2, N= 1859) = 25.77,$ $p < 0.00001$
	Genérica	800/1567	51,1	
	Expletivos	13/65	20,0	
	Subtotal	963/1859	51,8	
2 níveis	Específica	150/227	66,1	$X^2(1, N= 1859) = 1.56,$ $p = 0.2107$
	Genérica (+ Expletivos?)	813/1632	49,8	
	Subtotal	963/1859	51,8	

Fonte: Nascimento (2011, p. 154, adaptado).

Nascimento (2011) analisou a variação entre **você** e **cê** em uma amostra de entrevistas sociolinguísticas de falantes da cidade de São Paulo/SP. A pesquisadora subdividiu o tipo de referência em nove níveis (conforme já discutimos na seção anterior): i) P2 específica; ii) P1 ‘ocultamento do eu’; iii) P1 ‘expandido’; iv) P4 ‘exclusivo’; v) P4 ‘inclusivo’ [+definido]; vi) P4 ‘inclusivo’ [-definido]; vii) híbrido; viii) P3/P6 (3ª pessoa); ix) expletivos. Nascimento (2011) realizou 4 análises com controle da variável tipo de referência: uma com os 9 níveis e as demais com amalgamações (com 6 níveis, 3 níveis e 2 níveis).

Na análise da referência com base em nove níveis, os resultados evidenciaram que os contextos de uso com referência específica ao interlocutor favorecem o uso da variante **cê** (150/227 - 66,1%); em contrapartida, os contextos considerados expletivos desfavoreceram-na (13/65 - 20%). No que concerne aos diferentes usos genéricos das formas **você** e **cê**, os resultados obtidos por Nascimento (2011) demonstraram que a referência à P1 ‘oculto’, à P1

‘expandido’ e à P4 exclusivo condicionaram o uso da variante **cê**. Os contextos de P4 inclusiva [+definida], P4 inclusiva [-definida], P3/P6 e híbridos desfavoreceram o uso da variante reduzida.

Para a análise com seis níveis, a autora amalgamou os níveis semelhantes: P1 ‘oculto’ + P1 ‘expandido’ + P4 exclusivo e P4 inclusiva [+definida] + P4 inclusiva [-definida]. Os resultados revelaram que a variante **cê** tende a ser utilizada nos contextos de referência às pessoas presentes na interlocução - primeira e segunda pessoas -, ocorrendo o contrário nos contextos de referência a não pessoa (nos termos de Benveniste (1995 [1956])). Os contextos de referência híbrida e de referência P4 ‘inclusiva’ pouco favorecem o uso da variante **cê**.

Na análise com três níveis, Nascimento (2011) amalgamou os níveis considerados de cunho genérico e manteve os níveis específicos e expletivos separados. Mesmo com a amalgamação feita, os resultados apontaram o favorecimento da variante reduzida em detrimento da variante plena apenas nos contextos de referência específica. O mesmo ocorre na análise binária - referência específica vs. genérica + expletivos. Os testes de qui-quadrado demonstraram que apenas a análise de referência com oposição binária não apresenta diferenças estatisticamente significativas. Além disso, o teste aponta que a análise da referência com controle de seis níveis é a que apresenta melhor resultado estatístico dentre as outras análises feitas quanto ao tipo de referência.

2.5 FUNÇÃO REFERENCIAL DOS PRONOMES TU, VOCÊ E CÊ: CONSIDERAÇÕES GERAIS

As discussões feitas nas seções anteriores evidenciam a necessidade de se elaborar uma proposta com mais acuracidade para o controle do traço semântico referencial dos pronomes de 2ªPS. Partindo da noção de pessoa promulgada por Benveniste (1995 [1956]), é notório que nem todos os usos das formas pronominais **tu**, **você** e **cê** fazem referência à 2ªPS, ou seja, ao interlocutor determinado, o qual faz parte da situação interlocutiva. Tais formas podem ser utilizadas também para se reportarem a qualquer pessoa do discurso, e, nesses casos, estas passam a ter abrangência referencial, deixando, assim, de ser uma referência específica para ser, em algum grau, indeterminada/genérica. Defendemos, portanto, que o uso das formas **tu**, **você** e **cê** com traços indeterminados não pode ser considerado como de referência à 2ªPS, pois já não se refere a uma pessoa dêitica, mas sim as pessoas de forma geral (passando a ter, de certa forma, uma interpretação plural) ou, ainda, ao próprio falante.

É importante ressaltar que os resultados dos estudos sistematizados demonstraram que as taxas de frequências das variantes quanto ao tipo de referência possuem um comportamento diversificado em cada região e, muitas vezes, dentro de uma mesma região. Apenas nos estudos da região Norte e da região Sul os resultados evidenciaram o favorecimento (em termos percentuais) de uma mesma forma – a variante **tu**.

Em virtude dessa diversidade de resultados e pelos demais motivos já apresentados, acreditamos ser necessário realizar as análises dos usos dessas formas separando os dados com base na função desempenhada por estas. Consideramos que cada função consiste em um fenômeno diferente e que, conseqüentemente, possui condicionamentos sociais, interacionais/pragmáticos, estilísticos e estruturais/formais também distintos.

3 ACOMODAÇÃO LINGUÍSTICA, POLIDEZ E PRONOMES DE PODER E SOLIDARIEDADE

Nesta seção, apresentamos as noções essenciais para entender os aspectos que influenciam o uso das variantes **tu**, **você** e **cê**, isto é, o que tende a influenciar a escolha de uma das formas em uma situação interlocutiva. Para tanto, dividimo-lo em duas subseções. Apresentamos, na primeira subseção, a teoria da acomodação, correlacionando-a com o modelo de coleta das entrevistas sociolinguísticas. Na subseção seguinte, discorremos sobre os aspectos concernentes à polidez linguística e às dimensões de poder, solidariedade e neutralidade que estão relacionados ao uso das variantes **tu**, **você** e **cê**.

3.1 TEORIA DA ACOMODAÇÃO LINGUÍSTICA E ENTREVISTAS SOCIOLINGUÍSTICAS

Nesta subseção, objetivamos demonstrar a importância de analisar a variação entre as formas de 2ªPS na fala de universitários sergipanos à luz da teoria da acomodação linguística, utilizando dados de fala coletados por meio de entrevistas sociolinguísticas. Assim, descrevemos, primeiramente, o método de coleta de entrevistas sociolinguísticas desenvolvido por William Labov (1966; 1984; 2008 [1972]). Na sequência, discorremos sobre a teoria da acomodação nos respaldando, principalmente, no estudo de Giles, Coupland e Coupland (1991). Em seguida, ainda no escopo da teoria acomodação, apresentamos o modelo de *Audience Design* proposto por Alan Bell (1984; 2001) pela aderência que este possui ao objeto de estudo da presente tese.

3.1.1 Entrevistas sociolinguísticas: o método laboviano de coleta de dados linguísticos

Os estudos sociolinguísticos, na vertente variacionista, têm como objetivo identificar padrões regulares de usos linguísticos em comunidades de fala/prática. Pesquisadores dessa área utilizam, frequentemente, como fonte de dados, a entrevista sociolinguística. Trata-se de um método de coleta de dados tradicionalmente usado entre os pesquisadores da Sociolinguística Variacionista desde os estudos iniciais de Labov (1966; 1984; 2008 [1972]),

caracterizando-se, assim, como o principal método de coleta de dados linguísticos da área (cf. MILROY; GORDON, 2003) e o mais difundido no Brasil (FREITAG, 2014).

A entrevista é uma situação comunicativa dialógica entre pelo menos duas pessoas, as quais alternam os turnos conversacionais, sendo que uma destas figura o papel de entrevistador e a outra de entrevistado. Há vários tipos de entrevistas (definidos de acordo com os seus propósitos, temas abordados, público-alvo, formalidade etc.), tais como: entrevista jornalística, entrevista médica, entrevista científica, entrevista de emprego, entre outras (cf. VALLE; GÖRSKI, 2014). Os tipos de entrevistas, em termos gerais, compartilham algumas características, a saber: i) é estruturada no par pergunta-resposta; ii) envolve pelo menos dois interlocutores – o entrevistador e o entrevistado; iii) o entrevistador é responsável por abrir e fechar a entrevista, fazer perguntas, estimular a palavra do outro, instigar “a transmissão de informações” solicitadas, introduzir novos tópicos, orientar e reorientar a interação; iv) o entrevistado tem o papel de “responder e fornecer as informações pedidas” (SCHNEUWLY; DOLZ, 2004, p. 86). Nesses termos, podemos considerar a entrevista sociolinguística um subgênero da entrevista. Por outro lado, a entrevista sociolinguística pode ser considerada um macrogênero textual pelo fato de poder apresentar diferentes gêneros textuais em sua composição, como, por exemplo, uma receita culinária, uma narrativa de experiência pessoal, um relato de opinião, entre outros (cf. TAVARES, 2015).

Labov (1984) ressalta que a entrevista sociolinguística consiste em um método de coleta bem desenvolvido com dez objetivos principais:

Quadro 5 - Principais objetivos do método de coleta

- (1) gravar com razoável fidelidade de 1 a 2 horas de fala de cada entrevistado/falante.
- (2) obter toda a gama de dados demográficos necessários para a análise dos padrões sociolinguísticos (idade; história residencial, escolar, ocupacional e de linguagem; localização e relações familiares; renda, renda ou valores da casa; participação em grupos e associações).
- (3) obter respostas comparáveis a perguntas polêmicas e de interesse em várias culturas (experiência do perigo de morte; destino; premonições; luta e regras para uma luta justa; atitudes em relação a outros grupos raciais e étnicos; aspirações educacionais).
- (4) promover narrativas da experiência pessoal, onde as normas e estilos de interação pessoal da comunidade são mais claramente revelados, e onde o estilo é mudado regularmente para o vernáculo.
- (5) estimular a interação entre o grupo de pessoas presentes e grave também a conversa não endereçada ao entrevistador.
- (6) separar tópicos de maior interesse para o falante e permitir que ele ou ela conduza a explicação do tópico da conversa.
- (7) traçar os padrões de comunicação entre os membros da comunidade e estabelecer a posição do falante na rede de comunicação.
- (8) obter um registro de atitudes evidentes em relação à linguagem, características linguísticas e estereótipos linguísticos.
- (9) obter informações específicas sobre estruturas linguísticas por meio de eliciações formais: leitura de textos e listas de palavras.
- (10) realizar experimentos de campo sobre reações subjetivas às percepções de formas linguísticas (testes pares mínimos e comutação; testes de autorrelato; testes de reação subjetiva; testes de antecedentes familiares).

Fonte: Labov (1984, p. 8, tradução nossa)¹⁸.

A entrevista é um gênero formalmente estruturado (conforme evidenciam as informações presentes no quadro 5), em decorrência disso, consideram-se os dados de fala decorrentes desta como monitorados, isto é, como uma fala cuidada. No entanto, em uma interação, podemos monitorar os nossos usos linguísticos ou não. Ao não monitorar, estamos fazendo uso que se aproxima do vernáculo, o qual, é definido como “o estilo em que se presta o mínimo de atenção ao monitoramento da fala” (LABOV, 2008 [1972], p. 244). É nesse estilo que os dados linguísticos mais casuais e espontâneos emergem; e, em virtude disso, os

¹⁸ No original: (1) to record with reasonable fidelity from 1 to 2 hours of speech from each speaker.

- (2) to obtain the full range of demographic data necessary for the analysis of sociolinguistic patterns (age; residential, school, occupational, and language history; family location and relations; income, rent or house values; group memberships and associations).
- (3) to obtain comparable responses to questions that define contrasting attitudes and experiences among various sub-cultures (experience of the danger of death; fate; premonitions; fight and rules for a fair fight; attitudes towards other racial and ethnic groups; educational aspirations).
- (4) to elicit narratives of personal experience, where community norms and styles of personal interaction are most plainly revealed, and where style is regularly shifted towards the vernacular.
- (5) to stimulate group interaction among the people present. and So record conversation not addressed to the interviewee.
- (6) to isolate from a range of topics those of greatest interest to the speaker, and allow him or her to lead in defining the topic of conversation.
- (7) to trace the patterns of communication among members of the neighborhood, and establish the position of the speaker in the communication network.
- (8) to obtain a record of overt attitudes towards language, linguistic features and linguistic stereotypes.
- (9) to obtain specific information on linguistic structures through formal elicitation: reading texts and word lists.
- (10) to carry out field experiments on subjective reactions towards perceptions of linguistic forms (minimal pair and commutation tests; self-report tests; subjective reaction tests; family background tests).

sociolinguistas variacionistas o têm como alvo basilar de seus estudos (cf. TAVARES, 2015). Desse modo, o grande objetivo de utilizar o método de coleta das entrevistas sociolinguísticas é coletar dados de fala espontânea ou mais natural, o vernáculo de uma comunidade de fala/prática.

No contexto de gravação de uma entrevista, há alguns empecilhos para se obter dados linguísticos vernaculares, tais como: o pesquisador/entrevistador e o gravador, gerando o que Labov (2008 [1972]) chama de “paradoxo do observador”, uma vez que a presença destes interfere na naturalidade da situação comunicativa. Contudo, a entrevista sociolinguística foi pensada de forma a fazer emergir o vernáculo por intermédio de estratégias que minimizam os efeitos do paradoxo do observador.

Uma das alternativas para tentar minimizar os efeitos do paradoxo do observador seria o entrevistador/pesquisador fazer um contato prévio com os membros da comunidade, a fim de que, no momento da entrevista, o entrevistado fique mais confortável em interagir com o entrevistador. Ademais, a realização de entrevistas sociolinguísticas é feita a partir de um roteiro de perguntas/temas estabelecido previamente pelo pesquisador, também com o intuito de minimizar os efeitos do paradoxo do observador (LABOV, 1984; 2008 [1972]) e fazer, assim emergir usos linguísticos o mais próximo possível do vernáculo do falante. Labov (1984; 2006 [1966]) afirma que as perguntas devem ser feitas em blocos/módulos temáticos e ressalta que a ordem, tanto das perguntas quanto dos módulos, deve partir do mais geral e impessoal para o mais específico e pessoal. O autor propõe os seguintes temas para serem abordados nas entrevistas: religião, economia, política, trabalho, brigas, família, infância, perigo de morte, medos, entre outros. Ao responder perguntas relacionadas a temas sobre experiências pessoais ou de cunho mais íntimos, tais como situações de risco de morte e recordações da infância (como as perguntas presentes no quadro 6), possivelmente o falante irá se envolver emocionalmente com o tema decorrido e esquecer que está sendo gravado, deixando de monitorar a fala.

Quadro 6 - Perguntas sobre “Perigo de morte” e “Infância”

<i>Perigo de morte</i>	<i>Infância</i>
1. Você já esteve em uma situação em que pensou que havia um sério perigo de você ser morto? Que você pensou consigo mesmo: "É o fim?" 2. O que aconteceu? 3. Como você se sentiu depois? (LABOV, 2006 [1966], p. 415, Tradução nossa) ¹⁹	1. Conte um fato que marcou sua infância. 2. Quando você era criança, você costumava brincar de quê? 3. Como eram seus pais? Eles eram muito rígidos? 4. Caso a pessoa tenha dito que tem irmãos, pergunte como era a relação dela com os irmãos na infância: vocês eram companheiros? Brigavam? Você lembra de alguma travessura que fizeram juntos? (Ver Apêndice A)

Fonte: Elaboração própria.

A diversidade de assuntos abordados no desenrolar da entrevista sociolinguística propicia o surgimento de diferentes tipos de sequências discursivas, visto que o entrevistador pode solicitar ao entrevistado, por exemplo: i) indicações sobre como chegar em um determinado local, como fazer uma receita (sequência injuntiva); ii) que relate algo que aconteceu em sua vida ou na comunidade (sequência narrativa); iii) opinião pessoal a respeito de um assunto (sequência argumentativa/opinativa); iv) que descreva um fato, objeto ou pessoa (sequência descritiva) etc. Dessa forma, as perguntas do entrevistador funcionam como um gatilho para o surgimento na resposta do entrevistado de um determinado tipo de sequência discursiva (FREITAG *et al.*, 2009; FREITAG, 2014). Assim como pode ocorrer mudança de estilo na fala do entrevistado de acordo com o tópico abordado, o tipo de sequência discursiva também propicia alteração estilística: sequências narrativas e descritivas tendem a propiciar uma fala menos monitorada, ocorrendo o contrário em sequências expositivas, argumentativas/opinativas.

Os tópicos discursivos abordados durante a entrevista são introduzidos pelo entrevistador por meio da pergunta e o desenvolvimento destes é feito pelo entrevistado/informante por meio da resposta. No entanto, o entrevistado também pode fazer a inclusão de outros tópicos (participante ativo na entrevista) e falar sobre eles. Ao discorrer sobre cada tópico da entrevista, espera-se que o entrevistado se prolongue em cada turno que detém a palavra para que, desse modo, tenha-se um quantitativo extenso de dados linguísticos. Assim, o fato de o entrevistador ter o papel de conduzir a entrevista não significa que este o faça em todo o transcorrer da gravação. De certa forma, a condução da entrevista é compartilhada entre o entrevistador e entrevistado, sendo que quanto mais este último participa dessa condução mais

¹⁹ No original: Danger of Death.

1. Have you ever been in a situation where you thought there was a serious danger of your being killed? That you thought to yourself, "This is it?"
2. What happened?
3. How did you feel afterwards?

os dados de fala gravados tendem a se aproximar do vernáculo ou de uma conversa entre conhecidos/amigos (cf. LABOV, 1984; 2001).

Além dos empecilhos mencionados, outros aspectos podem surgir durante a realização da entrevista e interferir no desenvolvimento desta de modo a barrar o surgimento do vernáculo do falante. Dentre os aspectos estão: “a desenvoltura e loquacidade do informante, o preparo do entrevistador, o grau de empatia estabelecida entre os interlocutores, o interesse pelo assunto, o nível de conhecimento sobre determinado tema e assim por diante” (VALLE; GÖRSKI, 2014, p. 104-105)²⁰. A junção desses aspectos faz com que uma entrevista se diferencie, em maior ou menor grau, da outra (cf. VALLE; GÖRSKI, 2014).

É notório, então, que todo o protocolo de coleta da metodologia laboviana visa modificar a situação de controle envolvida na entrevista sociolinguística de forma a propiciar uma situação comunicativa mais próxima de uma conversação espontânea entre pessoas conhecidas/amigas e, dessa forma, obter dados linguísticos mais próximos do vernáculo do falante (LABOV, 2001). Seguindo a proposta de metodologia de coleta laboviana, é possível reverter o quadro de monitoramento de estilos de fala, por exemplo, ao abordar tópicos discursivos que promovam um maior envolvimento emocional do falante, fazendo com que este preste menos atenção ao monitoramento da fala. Assim, na entrevista sociolinguística, é possível captar estilos de fala casual (que se aproximam do vernáculo), os quais se alternam com estilos de fala monitorada. Portanto, “a entrevista se configura, assim, como uma forma bastante sistemática [...] para se captar em momentos bem determinados registros próximos ao vernáculo” (VALLE; GORSKI, 2014, p. 77). Além disso, o protocolo de coleta da entrevista sociolinguística foi delineado com o objetivo de obter um quantitativo de dados exponencial que permita realizar tratamento estatístico, por meio do qual seja possível identificar tendências distribucionais, tanto sociais quanto linguísticas e estilísticas, a respeito de qualquer fenômeno da língua em variação (cf. TAVARES, 2015).

É importante ressaltar, ainda, que o roteiro é elaborado de forma a não direcionar a entrevista para que surjam usos de um fenômeno variável em específico, mas sim que o *corpus* constituído possa ser utilizado para a análise de qualquer fenômeno linguístico em variação. Isso se deve ao fato de a constituição de bancos de dados linguísticos, seguindo a metodologia da Sociolinguística Variacionista, consistir em uma atividade onerosa, que exige muito tempo

²⁰ “O grau de empatia estabelecida entre os interlocutores” é um aspecto que está relacionado ao tipo de relação e ao grau de proximidade existentes entre o entrevistador e o entrevistado, e, conseqüentemente, às dimensões de poder e de solidariedade envolvidas na situação comunicativa, bem como à polidez. Na seção 3.2, abordamos sobre essas dimensões (e sua relação com os pronomes de segunda pessoa) e a noção de polidez.

do pesquisador, tanto para coleta quanto para o processo de transcrição subsequente. Portanto, a aplicação dessa metodologia “precisa ser otimizada, subsidiando a pesquisa de diversos fenômenos de variação linguística, tanto na dimensão social como também na dimensão estilística” (FREITAG; SNICHELOTTO, 2015, p. 167).

Dentre as vantagens de se utilizar as entrevistas sociolinguísticas como *corpus* de extração de dados para a análise da variação linguística, está o fato de estas viabilizarem a obtenção de resultados consistentes e reaplicáveis, os quais são, em tese, comparáveis entre si (cf. LABOV, 2001). É justamente com a intenção de elevar o nível/o poder de comparabilidade entre amostras de fala que o método de coleta da entrevista sociolinguística foi elaborado (cf. ECKERT, 2001) pautando-se em critérios. Portanto, o protocolo de coleta seguido tenciona garantir confiabilidade quanto aos resultados das pesquisas feitas, isso significa que, se um pesquisador realizar um estudo sobre o mesmo fenômeno que um outro, seguindo a mesma metodologia, ele provavelmente obterá resultados semelhantes. Tal fato possibilitará que os resultados sejam generalizáveis.

É fundamental que se continue com a tradição de constituir bancos de dados a partir da coleta de entrevistas sociolinguísticas, aos moldes variacionistas, por esta ser uma metodologia convencionalizada e disseminada na área, propiciando a comparabilidade dos resultados entre estudos que possuem o mesmo/semelhante objeto de análise e assim identificar tendências de usos (cf. FREITAG; SNICHELOTTO, 2015; FREITAG, 2016). Isso não significa que novos aspectos não possam ser incorporados, pelo contrário, é necessário, por exemplo, que se incorpore “novas categorias sociodemográficas nas estratificações de bancos de dados sociolinguísticos e revisão das já existentes, a fim de acompanhar a dinâmica da sociedade e captar os perfis identitários locais de cada comunidade de fala” (FREITAG; SNICHELOTTO, 2015, p. 167), como também de cada comunidade de prática. Além disso, constituir amostras de bancos de dados sociolinguísticos que dão continuidade a outros já existente é de suma relevância para que se possa “manter a série histórica”, permitindo, assim, a captação “tendências amplas de variação/mudança” (FREITAG; SNICHELOTTO, 2015, p. 167), bem como mapear semelhanças e diferenças de usos linguísticos existentes no falar em diversas localidades do território brasileiro. Dessa forma, poder-se-á analisar os dados linguísticos, tanto em tempo real quanto em tempo aparente (a partir de amostras constituídas com estratificação da idade), sendo “esta combinação de observações [...] o método básico para o estudo da mudança em progresso” (LABOV, 1994, p. 63 apud FREITAG; SNICHELOTTO, 2015, p. 167).

Apesar de todas essas vantagens, o método de coleta da entrevista sociolinguística foi, e ainda é, alvo de críticas. Muitos pesquisadores não consideram viável utilizar dados de fala decorrentes de amostras de entrevistas sociolinguísticas em pesquisas que focalizam fenômenos linguísticos que são sensíveis a aspectos pragmáticos (como o tipo de relação interpessoal) e estilísticos. Embora haja limitações, não podemos descartá-lo em virtude dos motivos já elencados. Uma das possibilidades de utilizar a entrevista de modo mais produtivo, nas pesquisas sobre esses tipos de fenômenos, é não só focalizar nas variáveis relacionadas ao entrevistado/falante, mas também nas relacionadas ao entrevistador (o qual figura o papel de destinatário), visto que os aspectos relacionados a este podem interferir nos usos linguísticos do falante. Identificar esses aspectos é a proposta da teoria da acomodação, e é a respeito dela que tratamos na subseção a seguir.

3.1.2 Acomodação linguística

A teoria da acomodação da fala foi elaborada a partir da fundamentação da psicologia social, tendo sido desenvolvida, principalmente, por Howard Giles. O foco dessa teoria recai sobre a explicação de processos contextuais (os quais estão envolvidos na interação comunicativa) que influenciam as escolhas do falante, seja em relação ao código, ao estilo ou às estratégias sociolinguísticas (GILES; COUPLAND; COUPLAND, 1991).

A noção de acomodação pode ser entendida, em termos gerais, como a atenuação das diferenças linguísticas existentes entre os interactantes em uma situação comunicativa. Isso ocorre quando o falante alinha seu comportamento linguístico, durante a interação, aproximando-o ao do ouvinte, a fim de que este o avalie de forma positiva (preservação da face positiva do falante, nos termos de Brown e Levinson (2011 [1987])), ou seja, o falante busca aprovação social e, conseqüentemente, que a comunicação seja bem-sucedida. Desse modo, a acomodação “pode funcionar para indexar e obter solidariedade ou dissociação de um parceiro de conversação de forma recíproca e dinâmica”²¹ (GILES; COUPLAND; COUPLAND, 1991, p. 2, tradução nossa).

A acomodação é um processo que envolve um conjunto de alternativas que o falante dispõe, no decurso de uma conversação face a face, para atingir as suas necessidades comunicativas (de integração social ou de preservação de identidade grupal, por exemplo),

²¹ No original: It can function to index and achieve solidarity with or dissociation from a conversational partner reciprocally and dynamically.

sejam elas (alternativas) no escopo da fala e dos padrões discursivos, como também do comportamento não verbal durante a interação social. Para Giles, Coupland e Coupland (1991), a teoria da acomodação consiste em um paradigma bastante desenvolvido, a partir do qual é possível explicar vários aspectos que condicionam os usos linguísticos, tais como:

(1) consequências sociais (atitudinais, atribucionais, comportamentais e comunicativas), (2) fatores ideológicos e macrosociais, (3) variáveis e processos intergrupais, (4) práticas discursivas em contextos naturalistas e (5) expectativa de vida do indivíduo e mudanças de linguagem de grupo²² (GILES; COUPLAND; COUPLAND, 1991, p. 4, tradução nossa).

Essa teoria tem dois pressupostos básicos, a saber: i) os interactantes de uma situação comunicativa possuem “motivações para adaptar sua comunicação em relação às suas percepções dos estilos comunicativos dos interlocutores”²³; ii) os interactantes elaboram “impressões e avaliações de parceiros com base nas suas expectativas em relação ao estilo comunicativo de um parceiro quanto ao seu próprio estilo”²⁴ (STREET JR, 1991, p. 134, tradução nossa). Destarte, “estratégias de acomodação podem caracterizar realinhamentos indiscriminados de padrões de código ou seleção da linguagem, [...] relacionados a constelações de crenças, atitudes e condições socioestruturais subjacentes”²⁵ (GILES; COUPLAND; COUPLAND, 1991, p. 2, tradução nossa).

Em uma situação comunicativa face a face, os interactantes podem ter comportamentos linguísticos convergentes e divergentes (cf. figura 5, situações 1 e 2, respectivamente). Os comportamentos linguísticos convergentes são aqueles em que há uma aproximação entre o modo de falar do falante e do ouvinte/interlocutor. Isso significa que os interlocutores (ou pelo menos um deles) de uma situação comunicativa ajustam o seu modo de falar ao do outro, reduzindo, dessa forma, as diferenças de estilos (cf. GILES; COUPLAND; COUPLAND, 1991). Desse modo, os interactantes que tencionam lograr aprovação, estabelecer relação com o outro, comunicar-se produtivamente etc., geralmente, modificam seu estilo linguístico de forma a convergir com o estilo do interlocutor, promovendo-se, assim, trocas comunicativas suaves e estáveis, sem atritos/desentendimentos (cf. STREET JR, 1991).

²² No original: (1) social consequences (attitudinal, attributional, behavioral, and communicative), (2) ideological and macro societal factors, (3) intergroup variables and processes, (4) discursive practices in naturalistic settings, and (5) individual life span and group-language shifts.

²³ No original: Motivations for adapting their communication relative to their perceptions of the communicative styles of interlocutors.

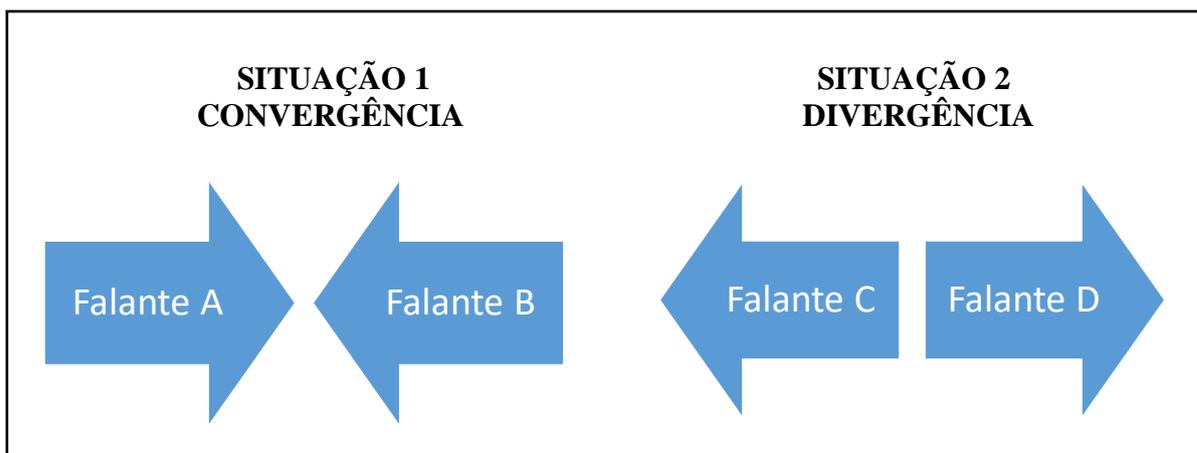
²⁴ No original: Impressions and evaluations of partners with respect to their expectations for a partner's communicative style relative to their own style.

²⁵ No original: accommodation strategies can characterize wholesale realignments of patterns of code or language selection, [...] related to constellations of underlying beliefs, attitudes, and sociostructural conditions.

Quanto aos comportamentos linguísticos divergentes, estes acontecem quando o falante se distancia ou evita uma aproximação do estilo comunicativo do interlocutor na interação, constituindo-se como uma forma de dissociação. Para tanto, o falante realça as diferenças dialetais existentes entre ele e o interlocutor, mantendo, portanto, a integridade, distância ou identidade no nível pessoal ou grupal (cf. GILES; COUPLAND; COUPLAND, 1991). Neste caso, temos uma interação instável, na medida em que o falante quer mostrar dissociação e desaprovação do parceiro, controlar a interação por meio da mudança de estilo ou ainda motivar o interlocutor a utilizar um outro estilo, por exemplo, um estilo mais reservado e formal. Conseqüentemente, isso pode promover o fim da interação, mudanças de comportamento do interlocutor ou ocasionar impressões negativas (PATTERSON, 1983 apud STREET JR, 1991). Tanto os comportamentos linguísticos convergentes quanto os divergentes, podem ocorrer por intermédio de recursos verbais (fala, extensão vocal, prosódia etc.) e não verbais (expressões faciais, movimentos corporais etc.) (cf. GILES; COUPLAND; COUPLAND, 1991).

Na figura 5, temos a ilustração de situações comunicativas em que há convergência (situação 1) e divergência (situação 2). A situação de convergência, evidenciada pelo uso de setas que estão indo uma ao encontro da outra, demonstra que o falante A e o falante B estão adaptando seus comportamentos linguísticos de modo que estes passem a ser semelhantes. Em contrapartida, a situação de divergência é demarcada pelo uso de setas em direção contrária, o que indica que o falante C e o falante D optaram por ter um comportamento linguístico que se distancia um do outro.

Figura 5 - Situações de convergência e divergência na interação



Fonte: Elaboração própria.

Em uma situação interlocutiva, é possível que os interactantes tenham papéis sociais distintos ou outras diferenças sociais e, em decorrência disso, um deles possua poder sobre o outro ou um status social maior, como, por exemplo, nas relações pai-filho, médico-paciente, professor-aluno, entrevistador-entrevistado (STREET JR, 1991). A demarcação de poder/status

social pode ser realizada por meio da linguagem, isto é, das formas linguísticas utilizadas pelo falante durante uma interação. Um exemplo disso é o uso das formas pronominais em estudo - **tu** e **você** (e suas reduções) -, que, dependendo da cultura do lugar, podem ser usadas para manifestar poder (na subseção 3.5, abordamos essa questão). Nessas situações de diferenciação de poder, os falantes podem manter cada um (ou apenas um deles) o seu comportamento linguístico, a fim de ressaltar as diferenças sociais existentes entre eles (cf. GILES, 1977 apud STREET JR, 1991). Nesse âmbito, é importante destacar que dois padrões de acomodação podem emergir na situação interlocutiva retratada: complementariedade e divergência. A complementariedade emerge quando os interactantes desejam mutuamente manter as diferenças de comportamento linguístico, propiciando uma troca interativa suave e tipicamente estável. Já a divergência linguística só ocorrerá de fato quando um dos interlocutores quiser manter a diferenciação existente entre eles, externando-a por meio dos seus usos linguísticos. Portanto, os ajustes acomodativos são decorrentes dos interesses pessoais dos interactantes, das percepções dos interactantes, dos estilos dos interlocutores e das exigências situacionais-relacionais (STREET JR, 1991).

As mudanças de comportamento linguístico, em uma interação face a face, tanto no âmbito da convergência quanto no da divergência, podem acontecer de forma *upward* ou *downward*. A primeira forma ocorre quando a mudança parte de uma variedade menos prestigiada para uma consensualmente mais prestigiada, tendo-se, portanto, um movimento ascendente (de baixo para cima). Já a segunda, *downward*, ocorre quando a mudança é direcionada para uma variedade estigmatizada ou socialmente menos valorizada no contexto, isto é, parte-se de uma variedade mais prestigiada para uma menos prestigiada, havendo, assim, um movimento descendente (de cima para baixo) (cf. GILES, COUPLAND, COUPLAND, 1991). A título de exemplificação podemos mencionar o fato de que, em um determinado momento de uma entrevista sociolinguística, o entrevistador pode adotar um estilo mais prestigiado e o entrevistado pode realinhar o seu estilo de modo a se aproximar do estilo do entrevistador. Neste caso temos um exemplo de convergência ascendente. Um exemplo de convergência descendente é a mudança de estilo dos apresentadores do Jornal Nacional da rede televisiva Globo para um estilo que se aproxime mais do utilizado pela maioria dos telespectadores, ou seja, eles realinham o comportamento linguístico para um padrão de menor prestígio.

A mudança de comportamento linguístico, em uma situação comunicativa, não ocorre de forma total. Isso quer dizer que o falante não realiza mudanças em todas as variáveis e níveis linguísticos disponíveis. As mudanças podem ser unimodais (quando ocorre em uma única

dimensão variável ou nível) ou multimodais (quando implica mudanças em várias dimensões). Ademais é importante ressaltar que a convergência e a divergência não são estratégias mutualmente exclusivas, visto que a convergência de algumas características linguísticas pode ser “igualada à divergência simultânea de outras”²⁶ (GILES; COUPLAND; COUPLAND, 1991, p. 11-12). Portanto, em uma interação face a face, é possível que em um momento os falantes façam usos linguísticos convergentes, e, em um outro momento, divergentes. Em decorrência disso, o processo de (não) acomodação pode ser parcial ou total.

Os processos de convergência e de divergência, em qualquer situação comunicativa, podem ser usados de forma simétrica ou assimétrica. O uso de forma simétrica ocorre quando o falante e o interlocutor fazem mudanças com o intuito de aproximar as características linguísticas um do outro. Quanto ao uso assimétrico, este acontece quando a convergência ou a divergência de um dos interactantes não é correspondida pela do outro (GILES; COUPLAND; COUPLAND, 1991).

A teoria da acomodação linguística está relacionada também a aspectos cognitivos, tais como as dimensões objetivas e subjetivas. A dimensão objetiva se refere às mudanças que o falante faz no seu estilo de forma a se aproximar (convergência) ou se distanciar (divergência) do estilo dos outros interlocutores. No que se refere à dimensão subjetiva, esta diz respeito à crença do falante quanto ao fato de ele ou os outros estarem convergindo ou divergindo (GILES; COUPLAND; COUPLAND, 1991). O quadro a seguir apresenta as estratégias de acomodação e as dimensões objetivas e subjetivas inter-relacionadas, resultando em quatro níveis possíveis de correlação entre estas.

Quadro 7 - Dimensões da acomodação subjetiva e objetiva da fala

		Acomodação subjetiva	
		Convergência	Divergência
Acomodação objetiva	Convergência	A	B
	Divergência	C	D

Fonte: Thakerer *et al.* (1982 apud GILES, COUPLAND, COUPLAND, 1991, p. 15, tradução nossa).

Observa-se que nas células A e D existe uma correspondência entre o que ocorre na dimensão objetiva e na dimensão subjetiva: a estratégia utilizada pelo falante é a mesma que ele acredita estar usando. Na célula A, o falante está convergindo com a sua audiência

²⁶ No original: [...] matched by simultaneous divergence of others.

(dimensão objetiva) e possui a crença de que está convergindo com o outro (dimensão subjetiva); e, na célula D, o falante está divergindo do seu interlocutor (dimensão objetiva) e acredita que de fato isso está acontecendo (dimensão subjetiva). Dessa forma, nas células A e D, a dimensão objetiva coincide com a subjetiva. Entretanto, nas células B e C, ocorre justamente o contrário, isto é, há uma dissonância entre o que o falante emprega e o que ele acha que está empregando. Em B, o falante está convergindo com o ouvinte (dimensão objetiva), mas acredita que está divergindo com este (dimensão subjetiva) e, em C, temos o oposto: o falante diverge do ouvinte (dimensão objetiva) e acredita que está convergindo (dimensão subjetiva) (GILES; COUPLAND; COUPLAND, 1991).

O grau de saliência de um traço linguístico pode interferir nos processos de acomodação linguística, pelo fato de que quanto mais perceptíveis pelo falante ele for, mais passíveis de valorização social estará. Desse modo, a acomodação possui uma relação intrínseca com os diferentes níveis de percepção e de avaliação social das variantes linguísticas (CAMPBELL *et al.*, 2014). Isso se deve, em termos gerais, ao fato de a acomodação ser cognitivamente mediada pela preconceção que o falante possui em relação ao modo como os outros, categorizados socialmente, falarão (HEWSTONE; GILES, 1986 apud GILES; COUPLAND; COUPLAND, 1991).

Com base na avaliação social que um fenômeno linguístico recebe, é possível classificá-los em: indicadores, marcadores e estereótipos (LABOV, 2008[1972]). Os indicadores dizem respeito a traços linguísticos que operam no nível do inconsciente e que possuem distinção social (conforme a idade e o grupo social) e pouca força avaliativa na sociedade; em decorrência disso, não possuem padrão de variação estilística, sendo utilizados pelos indivíduos mais ou menos do mesmo modo em qualquer contexto. Os marcadores são traços linguísticos que exibem estratificação social e estilística e podem estar abaixo do nível de consciência. Os estereótipos, por fim, são traços linguísticos representativos de indivíduos ou grupo de indivíduos e reconhecidos pelo falante; os quais estão acima do nível de consciência do falante e são fortemente avaliados pela sociedade (LABOV, 2008[1972]). Correlacionando essas noções à acomodação linguística, “[...] Podemos esperar que algumas variáveis - os indicadores - não se acomodem, mas a maioria mostrará algum grau de mudança em direção ao destinatário”²⁷ (BELL, 1984, p. 166, tradução nossa). Portanto, as variáveis que se caracterizam como indicadores, nos termos de Labov (2008 [1972]), são menos (ou não) afetadas pela

²⁷ No original: “[...] We can expect that a few variables - the indicators - will not accommodate, but most will show some degree of shift towards the addressee.

acomodação, ou seja, o falante não mudará seu estilo ao conversar com um interlocutor que tenha alguma distinção, como, por exemplo, ser de outra classe social; ocorrendo o contrário com aquelas que se caracterizam como marcadores e estereótipos em uma determinada comunidade.

Percebe-se, desse modo, que a visão estereotipada das pessoas leva em consideração características sociais que os membros da interação possuem, tais como idade, escolaridade, classe etc. Por exemplo, em uma entrevista sociolinguística, tanto o entrevistador quanto o entrevistado, podem mudar seu estilo linguístico para acomodar-se ao do outro. No entanto, essa mudança estará condicionada às características sociais de um e de outro, ao jogo de poder envolvido e ao custo dessas mudanças a face positiva ou negativa dos interactantes. Street (1991) ressalta que algumas pesquisas a respeito da relação médico-paciente evidenciam que os médicos, às vezes, potencializam o poder que possuem sobre os pacientes por meio de estratégias divergentes, como, por exemplo, mudanças de tópico e de estilo na interação.

Correlacionada à noção de estereótipos, está a noção de prototipicidade. Gallois e Callan (1988 apud GILES; COUPLAND; COUPLAND, 1991) descobriram que, quanto mais um novo membro de uma comunidade se aproximar do falar prototípico do grupo, mais bem avaliado socialmente este será, ocorrendo o oposto com a pessoa que se distanciar do falar prototípico. Assim, tais noções são relevantes no processo de convergência ou de divergência linguística (GILES; COUPLAND; COUPLAND, 1991).

A (não) atuação da acomodação linguística envolve aspectos relacionados, por exemplo, às normas sociais, à manutenção da identidade da comunicação, à atitude dos falantes em relação ao ouvinte/audiência. Alan Bell (1984) desenvolveu, com base na teoria da acomodação, um modelo voltado para a análise dos efeitos da audiência nos usos linguísticos, o qual denominou de *Audience Design*. Na subseção seguir, apresentamos esse modelo.

3.1.3 O modelo de *Audience Design*

Os sociolinguistas analisam a variação linguística, em termos gerais, a partir de duas categorias iniciais de variáveis que podem favorecer ou barrar a aplicação de uma regra variável: linguísticas e extralinguísticas. As variáveis linguísticas são aqueles de cunho fonológico, sintático, morfológico etc. As variáveis extralinguísticas possuem duas categorias/dimensões, as quais foram propostas desde os primeiros estudos de William Labov: social e estilística. A dimensão social, chamada também de variação interfalante, está voltada

para a identificação das diferenças linguísticas entre vários falantes, a qual é feita por meio do controle da classe social, gênero, idade etc. A dimensão estilística, ou variação intrafalante, está voltada para a identificação das diferenças linguísticas de um único falante (BELL, 1984; 2001). Essas duas dimensões são a base para compreender o modelo de *Audience Design* proposto por Alan Bell (1984; 2001).

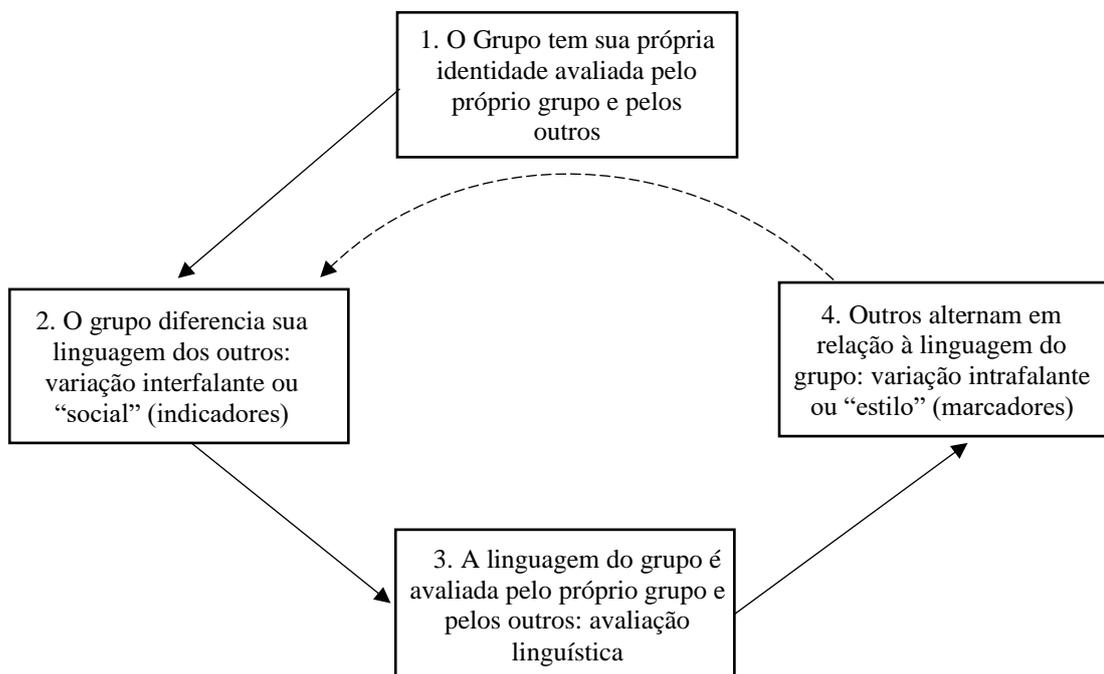
O modelo de *Audience Design* parte do princípio de que a variação intrafalante deriva da variação interfalante (BELL, 1984; 2001). Existe uma inter-relação, entre variação interfalante, variação intrafalante e avaliação linguística, o que consiste em uma evidência fulcral “sobre a derivação e a natureza da mudança de estilo”²⁸ (BELL, 1984, p. 150, tradução nossa). É possível que haja variação interfalante sem que haja variação intrafalante, no entanto, não há variação intrafalante sem que tenha havido primeiro variação interfalante, isto é, a variação social antecede/pressupõe a estilística (cf. BELL, 1984). Isso ocorre porque “a variação na dimensão do estilo no discurso de um único falante deriva e ecoa a variação existente entre os falantes na dimensão ‘social’”²⁹ (BELL, 1984, p. 151, tradução nossa).

Na figura 6, temos a representação das mudanças intrafalantes a partir de variação interfalantes. Observa-se, na primeira célula da figura, que a avaliação social da identidade dos falantes consiste na mola propulsora para o significado das diferenças entre os interlocutores. A partir dela, no fluxo derivacional, é possível surgir variação interfalante ou social, entendida como variação na fala entre indivíduos e grupos de individuais. A próxima etapa do fluxo da derivação é a avaliação linguística, ou seja, os usos linguísticos dos falantes são avaliados pelo próprio grupo e por outras pessoas. E, por fim, na última célula, temos a variação intrafalante ou estilística que ocorre quando os falantes alternam seus usos linguísticos em relação ao do grupo (BELL, 1984).

²⁸ No original: [...] on the derivation and nature of style shift.

²⁹ No original: Variation on the style dimension within the speech of a single speaker derives from and echoes the variation which exists between speakers on the "social" dimension.

Figura 6 - Derivação intrafalante a partir de variação interfalante



Fonte: Bell (1984, p. 152, tradução nossa)

A variação interfalante e a variação intrafalante estão relacionadas, respectivamente, às noções de indicadores e de marcadores definidas por Labov (2008 [1972]). Os fenômenos linguísticos que possuem traços de indicadores estão abaixo do nível da consciência, sendo assim, têm pouca força avaliativa e não possuem padrão de variação estilística, apenas com distinção social, ficando presos na célula 2 de desenvolvimento do fluxo derivacional (cf. figura 6) (BELL, 1984). A maioria dos fenômenos linguísticos é caracterizado como marcadores, possuindo variação nas duas dimensões, isto é, tanto social quanto estilística (LABOV, 2008 [1972]).

Os fenômenos linguísticos passíveis de mudança de estilo são aqueles que passam pela avaliação dos falantes da comunidade. Isso significa que “a mudança de estilo e a avaliação de uma variável sempre co-ocorrem e pressupõem que a variável seja socialmente diferenciada”³⁰ (BELL, 1984, p. 157, tradução nossa). Conforme demonstra a figura 6, a diferenciação de estilo de uma determinada variável linguística é decorrente “da diferenciação social por meio de avaliação social”³¹ (BELL, 1984, p. 157, tradução nossa). Nesse âmbito, Bell (1984) ressalta, ainda, que:

³⁰ No original: Style shift and evaluation of a variable always co-occur and presuppose that the variable is socially differentiated.

³¹ No original: [...] style differentiation of a variable is derived from social differentiation by way of social evaluation.

[...] a diferenciação social não precisa levar à avaliação e mudança de estilo. Os indicadores não estão sujeitos a avaliação. Uma variável pode ser diferenciada entre os falantes sem ser avaliada, mas assim que os falantes começam (inconscientemente) a avaliá-la, eles aparentemente também começam a mudar de estilo. A avaliação e a mudança de estilo são recíprocas³² (BELL, 1984, p. 157, tradução nossa).

Portanto, uma variável linguística terá variação intrafalante no discurso de um falante somente se houver variação interfalante (BELL, 1984, p. 158). Isso está relacionado ao fato de o falante ter ou não contato com uma variedade de estilos, a qual é decorrente do contato com outros falantes de outras comunidades. Além disso, uma variável linguística precisa ser utilizada de forma distinta entre os membros de uma comunidade, para que assim possa ser objeto de avaliação social e, conseqüentemente, passe a fazer parte do repertório linguístico dos falantes como uma opção de mudança de estilo. Caso o falante passe a não ter contato com pessoas que utilizam determinadas estilos, ele provavelmente deixará de utilizá-los, ou seja, perdem os estilos que foram adquiridos em outras comunidades ou por meio de pessoas que não fazem parte do círculo íntimo de pessoas com as quais o falante tem contato frequentemente (cf. BELL, 1984).

Nos estudos sociolinguísticos, os pesquisadores relacionam, comumente, a dimensão social da variação linguística aos atributos sociais de uma pessoa (neste caso do falante), tais como sexo/gênero, idade, classe social etc. Partindo do princípio da teoria da acomodação de que a variação intrafalante (estilística) é decorrente da variação interfalante (social), faz-se necessário relacionar tais atributos de uma pessoa à dimensão do estilo. No entanto, deve-se considerar as características sociais do ouvinte e não apenas as do falante. Dessa forma, para o estudo do estilo, é necessário que haja uma mudança de foco do falante para o ouvinte no controle dos aspectos envolvidos na dimensão social (BELL, 1984).

Com base nessas considerações, o modelo de *Audience Design* parte do pressuposto de que os falantes, ao elaborarem suas respostas, levam em consideração a audiência presente no contexto comunicativo, ou seja, “[...] os falantes criam seu estilo para o público”³³ (BELL, 1984, p. 159, tradução nossa). Em termos gerais, os falantes acomodam seu estilo para torná-lo semelhante ao da audiência, objetivando obter aprovação desta.

³² No original: [...] social differentiation need not lead to evaluation and style shift. Indicators are not subject to evaluation. A variable may be differentiated between speakers without being evaluated, but as soon as speakers begin (unconsciously) to evaluate it, they apparently also begin to style-shift. Evaluation and style shift are reciprocal.

³³ No original: [...] speakers design their style for their audience.

Em uma situação comunicativa, temos o falante (primeira pessoa do discurso) e sua audiência, a qual pode ser definida como a plateia ou as pessoas que ouvem o falante. No modelo proposto por Bell (1984), a audiência é constituída por quatro pessoas, que podem desempenhar papéis de: destinatário (addressee), auditor (auditor), ouvinte (overhearer) e “bisbilhoteira” (eavesdropper) (cf. figura 7).

Figura 7 - Pessoas e papéis na situação da fala



Fonte: Bell (1984, p. 159).

O papel de cada pessoa que faz parte da audiência é definido com base em quatro atributos: ser conhecida, ser ratificada ou ser abordada pelo falante (cf. quadro 8). O destinatário (segunda pessoa do discurso) é o principal membro da audiência, pois possui os três atributos. O auditor é o membro da audiência que tem dois atributos: ser conhecido e ser ratificado no grupo. No que concerne ao ouvinte e ao “bisbilhoteiro”, o primeiro possui o atributo de ser conhecido pelo grupo, já o segundo é desconhecido pelo grupo, não possuindo, conseqüentemente, nenhum atributo. Tanto o auditor como o ouvinte e o bisbilhoteiro fazem parte da terceira pessoa do discurso (BELL, 1984).

Quadro 8 - Hierarquia de atributos e papéis da audiência

	Conhecida	Ratificada	Abordada pelo falante
Destinatário	+	+	+
Auditor	+	+	-
Ouvinte	+	-	-
“Bisbilhoteiro”	-	-	-

Fonte: Bell (1984, p. 161, tradução nossa).

É importante ressaltar que os papéis que a audiência desempenha são, grosso modo, ordenados implicitamente de acordo com a quantidade de atributos: quanto menos atributos

possuir o membro da audiência mais distante estará do falante, havendo, assim, uma hierarquia de papéis da audiência, como pode ser observado na figura 7. Seguindo essa escala implícita, do mais próximo ao mais distante do falante, uma variável linguística pode variar estilisticamente conforme os papéis da audiência: destinatário > auditor > ouvinte (BELL, 1984). O membro da audiência que desempenha o papel de “bisbilhoteiro” não pode afetar o estilo de um falante justamente pelo fato dele ser desconhecido (BELL, 1984), isto é, os participantes da situação comunicativa não sabem que o “bisbilhoteiro” está presente (este seria uma pessoa que escuta a conversa às escondidas). No caso do *corpus* analisado nesta tese, a audiência é constituída apenas pelo destinatário e, em virtude disso, focamos, principalmente, nos aspectos teóricos relacionados a este.

Na maioria das situações comunicativas, a variação de estilo é afetada, dentre os papéis que os membros da audiência podem assumir, pelo destinatário. Vários estudos já foram desenvolvidos com o intuito de identificar “a força do efeito do destinatário no estilo de um falante”³⁴ (BELL, 1984, p. 163, tradução nossa). Bickerton (1980) e Thelander (1982), por exemplo, gravaram dados de fala dos mesmos informantes em duas situações de comunicação: em entrevistas e em interações entre amigos íntimos/colegas (cf. BELL, 1984). Os resultados desses estudos evidenciaram que os falantes mudaram o estilo a depender do tipo de coleta e do perfil do destinatário. As mudanças para um estilo mais padrão da variável eram maiores “em uma entrevista do que em grupo de colegas”³⁵ (THELANDER, 1982, p. 71 apud BELL, 1984, p. 163, tradução nossa), evidenciando, dessa forma, o efeito do entrevistador nos usos linguísticos do entrevistado. Em nosso *corpus* de análise, temos dados de fala coletados de forma semelhante - em uma das amostras, os informantes foram gravados em interação com amigos, e, na outra, temos gravações de entrevistas em que entrevistador e entrevistado não se conhecem, o que nos permite verificar se o fenômeno focalizado na presente tese possui taxas diferentes à medida que se modificam o tipo de coleta de dados e o perfil do destinatário.

Os falantes, ao responderem seus destinatários, avaliam três aspectos possíveis: i) “as características pessoais de seus destinatários e projetam seu estilo para se adequarem”³⁶; ii) “o nível geral de estilo do discurso de seus destinatários e mudam em relação a ele”³⁷; e iii) “os níveis de seus destinatários para variáveis linguísticas específicas e mudam em relação a esses

³⁴ No original: [...] the strength of the addressee's effect on a speaker's style [...].

³⁵ No original: [...] in an interview than in a peer group.

³⁶ No original: Speakers assess the personal characteristics of their addressees, and design their style to suit.

³⁷ No original: Speakers assess the general style level of their addressees' speech, and shift relative to it.

níveis”³⁸ (BELL, 1984, p. 167, tradução nossa). Algumas vezes, o estilo projetado pelo falante não condiz com o do destinatário, ocorrendo convergência subjetiva, mas que não se realiza de forma objetiva, conforme descrevemos na seção anterior.

Além da *Audience Design*, duas outras variáveis não pessoais principais podem influenciar no design do estilo de um falante: tópico e cenário. Alguns estudos, como de Blon e Gumperz (1972) e Giles e Poesland (1975), constataram que a mudança de tópico discursivo produz mudança de dialeto/sotaque (cf. BELL, 1984). Douglas-Cowie (1978) e Coupland (1981) também identificaram que a mudança de tópico produz mudança de estilo em algumas variáveis linguísticas (cf. BELL, 1984). Nesse âmbito, Bell (1984) destaca que:

É sabido que determinados tópicos ou cenários levam os falantes a mudar de estilo em uma determinada direção. [...] A única explicação adequada da direção da mudança de estilo não pessoal reside na associação e derivação da mudança projetada pelo público. Ou seja, os falantes associam classes de tópicos ou cenários a classes de pessoas. Eles, portanto, mudam de estilo ao falar sobre esses tópicos ou ambientes, como se estivessem conversando com destinatários a quem associam ao tópico ou ambiente. Tópicos como ocupação ou educação, e cenários como escritório ou escola, mudam para um estilo adequado para abordar um empregador ou professor. Da mesma forma, tópicos íntimos ou um ambiente doméstico suscitam discursos apropriados para destinatários íntimos - familiares ou amigos. A base de toda mudança de estilo de acordo com fatores não pessoais reside então na mudança projetada pelo público [...].³⁹ (BELL, 1984, p. 181, tradução nossa)

A organização dos tópicos discursivos agrupados em temas, que podem ser abordados em qualquer interação, faz parte das técnicas de coleta de entrevistas sociolinguísticas. Toda a agenda de coleta de uma entrevista sociolinguística é elaborada de forma a lograr um discurso casual. Para obter esse tipo de discurso, a manipulação dos tópicos pelo entrevistador é feita, deliberadamente, conforme discutimos na seção 3.1.1. A partir dessa manipulação, é possível “[...] simular as condições que produzem a fala para um destinatário íntimo”⁴⁰ (BELL, 1984, p. 182, tradução nossa).

Outra variável que interfere na projeção do estilo do falante para os seus interlocutores é a dimensão do relacionamento/o tipo de relação existente entre eles – se são íntimos ou não

³⁸ No original: Speakers assess their addressees' levels for specific linguistic variables, and shift relative to those levels.

³⁹ No original: It is well known that given topics or settings provoke speakers to style-shift in a certain direction. [...]The only adequate explanation of the direction of nonpersonal style shift lies in its association with, and derivation from, audience-designed shift. That is, speakers associate classes of topics or settings with classes of persons. They therefore shift style when talking on those topics or in those settings as if they were talking to addressees whom they associate with the topic or setting. Topics such as occupation or education, and settings such as office or school, cause shifts to a style suitable to address an employer or teacher. Similarly, intimate topics or a home setting elicit speech appropriate for intimate addressees - family or friends. The basis of all style shift according to nonpersonal factors lies then in audience-designed shift.

⁴⁰ No original: [...] simulate the conditions which produce speech to an intimate addressee.

íntimos, se são conhecidos ou estranhos. As pesquisas sobre acomodação evidenciam que há uma diferenciação na forma como os interactantes ajustam seu estilo linguístico em função desse aspecto. Em uma situação comunicativa em que os interlocutores não se conheciam antes de interagirem (ou seja, eram estranhos), como muitas vezes ocorre em entrevista sociolinguística, o grau de relacionamento é baixo. Quando isso acontece outra variável atua de forma mais proeminente nos usos linguísticos: o status social dos interlocutores. No entanto, quando os interlocutores possuem um grau de proximidade forte, o relacionamento se fortalece, tornando, assim, a influência do status social dos interlocutores nos usos linguísticos menos importante (BELL, 1984). Dessa forma, “existe uma distribuição recíproca, então, de relacionamento e de status, de modo que, se a força de uma variável é maior, a outra é menor”⁴¹ (BELL, 1984, p. 169, tradução nossa).

Alguns estudiosos defendem que os interactantes que conversam pela primeira vez tendem convergir, principalmente, em relação à “velocidade da fala, latência de resposta, sotaque, intimidade de autodivulgações e comportamentos faciais e gestuais”⁴² (STREET; GILES, 1982 apud STREET JR, 1991, p. 141, tradução nossa). Bell (1984) ressalta, tomando como base os resultados do estudo de Douglas-Cowie (1978), que após um certo tempo de gravação, geralmente na segunda hora de gravação, os interlocutores passam a se familiarizar e, conseqüentemente, passam a utilizar variantes menos padrão. Os usos dos pronomes de segunda pessoa são afetados por essas duas variáveis: as mesmas formas usadas entre pessoas íntimas são utilizadas para se referir aos inferiores socialmente, e as formas usadas entre não íntimos são as mesmas usadas para se dirigir a superiores (BELL, 1984). As variáveis status social e relacionamento são equivalentes, respectivamente, ao “poder” e à “solidariedade” propostos por Brown e Gilman (1966). Bell (1984, p. 169) faz uma ressalva de que “a diferença de código é mais nítida do que na mudança de estilo e as decisões são mais evidentes, mas os determinantes são os mesmos”⁴³. Retornamos a discussão da influência dessas variáveis na subseção 3.2.

Grosso modo, podemos assumir que as pressões para acomodar o discurso estão presentes em qualquer interação. No entanto, há situações de interação, que possuem um grau maior de pressão do que outras, como as situações públicas. Uma dessas situações é falar em público, visto que “quanto maior o público de um falante, maiores as pressões a serem

⁴¹ There is a reciprocal distribution, then, of relationship and status, so that if the strength of one variable is greater, the other is less.

⁴² [...] speech rate, response latency, accent, intimacy of self-disclosures, and facial and gestural behaviors.

⁴³ The code difference is sharper than in style shift and the decisions more overt, but the determinants are the same.

entendidas e a obter aprovação”⁴⁴ (BELL, 1984, p. 170, tradução nossa). Outra situação pública de grande pressão é a instituição de contato público, tais como instituições de serviço – lojas, restaurantes, empresas etc. – que precisam da aprovação de seus clientes (BELL, 1984). Geralmente os funcionários de instituições desse tipo acomodam a fala aos seus destinatários, neste caso, clientes. Esses são os dois tipos de situações públicas destacadas por Bell (1984) como as que mais reforçam o efeito, na fala, das variáveis relacionamento e status.

A discussão feita nesta subseção demonstra que o falante projeta seu estilo de forma a adequar-se ao destinatário, em termos gerais, à audiência (BELL, 1984). O modelo de *Audience Design* é aderente para a análise da variação de formas pronominais de 2ªPS. Uma análise nesse âmbito deve considerar as variáveis de cunho pessoal (tanto do falante quanto do destinatário), tais como: idade, sexo/gênero, status social, escolaridade, tipo de relacionamento entre os interactantes etc. As variáveis não pessoais, tais como tópico discursivo, cenário, sequência discursiva entre outras, devem ser consideradas na análise, no entanto, estas possuem, em tese, efeito secundário na mudança de estilo (BELL, 1984). Todas essas variáveis podem ser controladas a partir de amostras de dados de fala obtidos por meio de entrevistas sociolinguísticas, o que evidencia a validade do método em estudos de fenômenos linguísticos sensíveis à dimensão do estilo.

3.2 ASPECTOS PRAGMÁTICOS E FORMAS PRONOMINAIS DE 2ªPS

Os nossos usos linguísticos, em qualquer situação interlocutiva, são intrinsecamente regidos por aspectos sociais e culturais característicos da comunidade da qual fazemos parte. Isso significa que as normas linguísticas de uma determinada comunidade não são, necessariamente, as mesmas de outra (cf. MEYERHOFF, 2006), pois os fatores socioculturais de cada sociedade/comunidade se diferenciam, ou seja, cada grupo social elabora seu próprio sistema linguístico. Destarte, as mudanças sociais e culturais que ocorrem em uma comunidade fazem com que as normas de usos linguísticos também se modifiquem. Nas subseções a seguir, apresentamos as dimensões de poder, solidariedade e neutralidade, bem como a teoria da polidez, as quais estão intrinsecamente relacionadas às normas socioculturais e, conseqüentemente, linguísticas de cada sociedade.

⁴⁴ No original: The larger a speaker's audience, the greater the pressures to be understood and to win approval.

3.2.1 As dimensões de poder e de solidariedade

A forma organizacional da sociedade e as relações de poder existentes entre as pessoas se manifestam também linguisticamente. Uma das formas de materialização linguística dessas relações ocorre por intermédio dos usos de pronomes pessoais de segunda pessoa do discurso. Em muitas línguas há a disposição do falante dois pronomes singulares de tratamento à segunda pessoa do discurso. O uso de dois pronomes singulares de segunda pessoa/tratamento teve início no Latim com **tu** e **vos** e, posteriormente, expandiu-se para outras línguas (cf. BROWN; GILMAN, 1960).

No entanto, nem sempre foi assim no Latim. Na antiguidade, o Latim tinha apenas a forma **tu** para se referir à 2ªPS. Posteriormente, por volta do século IV, a forma plural **vos** passou a ser usada para se dirigir ao imperador (cf. BROWN; GILMAN, 1960). Nessa época existiam dois imperadores: um responsável pelo império Oriental, com sede em Constantinopla, e outro responsável pelo do Oeste, com sede em Roma. No entanto, em virtude das reformas de Diocleciano, o escritório imperial foi administrativamente unificado, mantendo-se os dois imperadores. Assim, palavras dirigidas a um deles, foram, por implicação, dirigidas a ambos (cf. BROWN; GILMAN, 1960). Decorre desse contexto, possivelmente, o uso de **vos** como forma de tratamento, em resposta a essa pluralidade implícita. Uma outra possível explicação é o fato de o imperador ser representante de seu povo e falar como seu representante, podendo assim ser considerado plural (cf. BROWN; GILMAN, 1960). Esse uso pode ainda ter sido inspirado pelo poder que um imperador possui, se levarmos em consideração que “a pluralidade é uma metáfora muito antiga e onipresente do poder”⁴⁵ (BROWN; GILMAN, 1960, 254, tradução nossa).

Brown e Gilman (1960) utilizaram os símbolos T e V (do latim **tu** e **vos**) como designadores genéricos em qualquer língua. O primeiro como designador genérico do pronome singular familiar e o segundo como designador do pronome singular respeitoso (cf. BROWN; GILMAN, 1960). A escolha de um ou de outro pronome de segunda pessoa está estreitamente relacionada à semântica deste. O termo semântica é entendido aqui como a “covariação entre o pronome utilizado e a relação objetiva que há entre falante e interlocutor”⁴⁶ (BROWN; GILMAN, 1960, p. 252, tradução nossa). No processo interacional, há duas noções semânticas

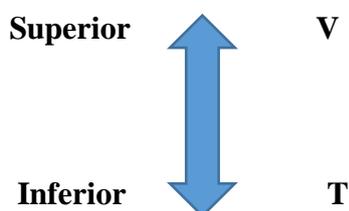
⁴⁵ No original: [...] for plurality is a very old and ubiquitous metaphor of power.

⁴⁶ No original: [...] covariation between the pronoun used and the objective relationship that exists between speaker and address.

aplicadas para a investigação de toda a vida social, a saber: dimensão do poder e dimensão da solidariedade (cf. BROWN; GILMAN, 1960).

A dimensão do poder atua quando um dos membros de uma situação discursiva – em que há, pelo menos, duas pessoas - possui uma posição superior em relação ao(s) outro(s) membro(s), sendo capaz de reger o comportamento do outro. Nessa dimensão, não é possível que duas pessoas tenham poder uma sobre a outra na mesma área do comportamento, consistindo, assim, em uma relação não recíproca, uma vez que necessariamente deve existir no contexto de interlocução uma pessoa na posição superior e outra na posição inferior. Trata-se, portanto, de uma relação vertical, assimétrica e não recíproca (cf. BROWN; GILMAN, 1960). Dessa forma, em uma interação em que a semântica do poder⁴⁷ está em jogo, o superior diz T e recebe V do inferior (cf. figura 8). Essa posição de superioridade, ou seja, de poder, manifesta-se de formas variadas socialmente, como, por exemplo, por meio de: “força física, riqueza, sexo/gênero, idade, papel institucionalizado na igreja, no Estado, no exército ou no seio família”⁴⁸ (BROWN; GILMAN, 1960, p. 255, tradução nossa), status social etc. Por exemplo, se um falante A é o empregador/chefe de um falante B, B não é o empregador/chefe de A, configurando-se como relação assimétrica.

Figura 8 - Esquema de uma relação não recíproca (assimétrica)



Fonte: Elaborado pela autora.

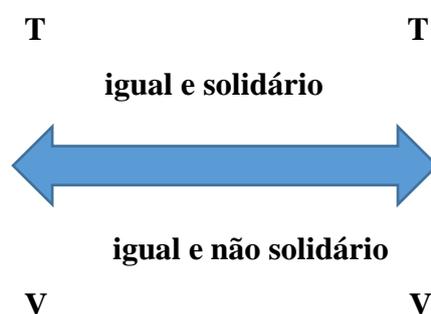
No entanto, nas relações sociais, as interações também ocorrem entre pessoas que possuem o mesmo poder de comportamento uma sobre a outra, havendo, deste modo, uma relação entre iguais (recíproca), sem que um dos membros seja inferior ou superior, isto é, pessoas/membros da interação possuem poder aproximadamente equivalentes (simetria). A

⁴⁷ Os autores ressaltam que “A semântica de poder não-recíproca é associada com uma sociedade relativamente estática, em que o poder é distribuído por direito de nascença e não está sujeito a muita redistribuição. O poder semântico estava intimamente ligado com os sistemas feudais e senhoriais” (BROWN; GILMAN, 1960, p. 265, tradução nossa). No original: The non-reciprocal power semantic is associated with a relatively static society, in which power is distributed by birthright and is not subject to much redistribution. The power semantic was closely tied with feudal and manorial systems.

⁴⁸ No original: [...] physical strength, wealth, age, sex, institutionalized role in the church, the state, the army or within the family.

título de exemplificação, se em uma situação discursiva o falante A tem a mesma profissão que B, B tem a mesma profissão de A; existe, desse modo, uma relação de simetria entre eles. Quando a interação ocorre nesses termos, não há, em tese, diferenciação de tratamento. Desse modo, o falante diz T e recebe T ou diz V e recebe V (cf. BROWN; GILMAN, 1960). A esse tipo de relação, Brown e Gilman (1960) chamaram de semântica/dimensão da solidariedade. Destarte, contrariamente à dimensão do poder, em uma interação em que a dimensão da solidariedade está em evidência, temos uma relação horizontal, simétrica e recíproca entre os falantes, conforme ilustra o esquema a seguir.

Figura 9 - Esquema de uma relação recíproca (simétrica)



Fonte: Elaborado pela autora.

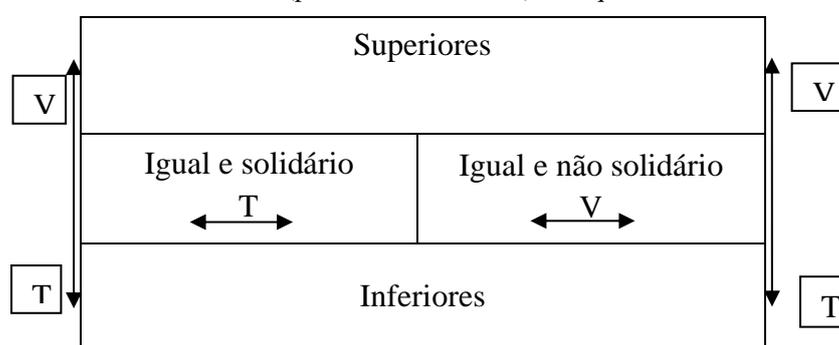
Brown e Gilman (1960) ressaltam que inicialmente não existia nas línguas regras que distinguíssem o tratamento usado entre iguais. Entretanto, de forma muito gradual, foi desenvolvida uma diferenciação de tratamento, nesse âmbito, gerando a dicotomia T de intimidade e V de formalidade. O uso da forma V estaria condicionado aos contextos comunicativos em que há diferenças entre as pessoas. Contudo, “nem sempre as diferenças entre pessoas implicam uma diferença de poder”⁴⁹ (BROWN; GILMAN, 1960, p. 257, tradução nossa). Isso significa dizer que, quando há uma distinção de poder na situação discursiva, emerge-se o uso de V apenas em uma direção de tratamento (do inferior para o superior), já quando há diferenças entre os interlocutores que não estão relacionadas com poder, emerge-se o uso de V em ambas as direções (cf. BROWN; GILMAN, 1960). Muitas línguas europeias modernas, atualmente, possuem essa distinção, tais como: **tu** e **vous** no francês, **du** e **Sie** no alemão, **tu** e **lei** no italiano, **tu** e **usted** no espanhol e em outras línguas (cf. BROWN; GILMAN,

⁴⁹ No original: Not all differences between persons imply a difference of power.

1960). No inglês antigo, existia essa distinção por meio do uso de **thou** e **you**, mas atualmente apenas esse último pronome é utilizado para qualquer situação⁵⁰.

Nas sociedades em que a solidariedade se mantém como forma de distinguir o tratamento entre pessoas de poderes iguais, o sistema bidimensional da semântica dos pronomes encontra-se/encontrava-se em equilíbrio. Isso significa que esse sistema possui pesos que seguem dois direcionais: enquanto a dimensão da solidariedade pesaria no sistema como forma de tratamento entre pessoas iguais em poder, a dimensão do poder pesaria no tratamento entre pessoas de diferentes poderes. A figura a seguir ilustra essa situação.

Figura 10 - A semântica/sistema bidimensional (poder e solidariedade) em equilíbrio



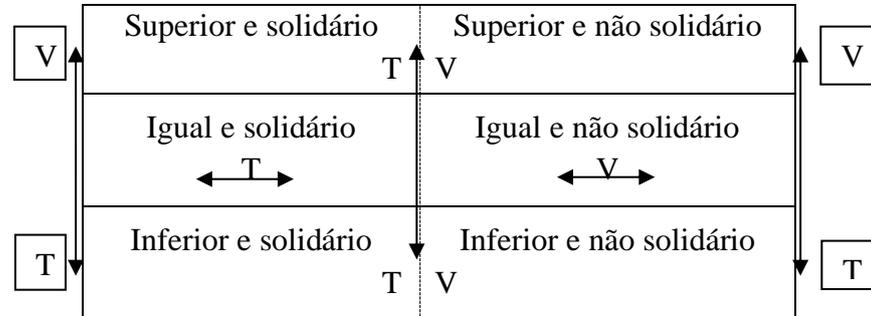
Fonte: Adaptado de Brown e Gilman (1960, p. 259, tradução nossa).

Brown e Gilman (1960) ressaltam que, durante um tempo considerável, o sistema semântico bidimensional parece ter permanecido em equilíbrio nas línguas e pode estar ainda em sociedades mais fechadas. Contudo, as sociedades se transformaram e se transformam ao longo do tempo e o surgimento de ideologias igualitárias fizeram com que muitas se democratizassem. Conseqüentemente, com as mudanças sociais, as normas de usos linguísticos também se modificam, como é o caso dos usos dos pronomes. Em outras palavras, o desenvolvimento de sociedades mais abertas agiu, sobretudo, contra a semântica de não reciprocidade do poder e em prol da semântica da solidariedade. Em virtude disso, o sistema semântico dos pronomes se desestabilizou em muitas comunidades. Tal desestabilização decorre do fato de que a dimensão da solidariedade passou a ser, potencialmente, aplicável a qualquer pessoa abordada. Isso significa que uma pessoa superior pode ser solidária (por exemplo, um empregador pode usar V em direção ao funcionário) ou não em uma situação

⁵⁰ Conforme os autores, “Em inglês geral, é claro, ‘**tu**’ não é mais usado. A explicação de seu desaparecimento não é de forma alguma certa; no entanto, as forças em ação parecem ter incluído uma reação popular contra o radicalismo dos Quakers e dos Levelers e também uma tendência geral em inglês em direção à inflexão verbal simplificada” (BROWN; GILMAN, 1960, p. 268, tradução nossa). No original: In English at large, of course, ‘**thou**’ is no longer used. The explanation of its disappearance is by no means certain; however, the forces at work seem to have included a popular reaction against the radicalism of quakers and Levelers and also a general trend in English toward simplified verbal inflection.

interlocutiva, assim como uma pessoa inferior, em termos de poder, pode, do mesmo modo, ser ou não solidária (cf. BROWN; GILMAN, 1960). A figura abaixo evidencia essa situação.

Figura 11 - O sistema bidimensional (poder e solidariedade) em conflito



Fonte: Adaptado de Brown e Gilman (1960, p. 259, tradução nossa).

Observa-se, na figura 11, que essa extensão da dimensão da solidariedade cria seis categorias de pessoas, as quais são definidas com base nas suas relações com um falante (cf. BROWN; GILMAN, 1960). O conflito das regras de tratamento (evidenciado pelo pontilhado nas linhas da figura) surge em dois pontos do sistema bidimensional: na parte superior esquerda da figura e na parte inferior direita. O conflito de tratamento surge, para as pessoas que estariam na parte superior esquerda da figura, pelo fato de que estas seriam influenciadas pela dimensão do poder por estarem em uma posição de superioridade, mas também pela dimensão da solidariedade uma vez que são iguais solidárias. Portanto, para essa parte do sistema, o poder indica V e a solidariedade T. Já em relação à parte inferior direita as pessoas sentiriam o peso da dimensão da solidariedade por estarem em uma posição de inferioridade, assim como pela dimensão do poder por serem não solidárias. Para essa parte do sistema bidimensional, o poder indica T e a solidariedade V (cf. BROWN; GILMAN, 1960). Na figura 12, são apresentadas díades sociais nas quais o conflito abstrato retratado na figura 11 seria sentido.

Figura 12 - Díades sociais envolvendo conflito semântico

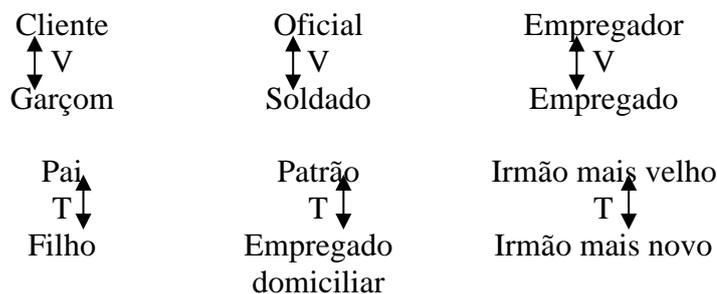


Fonte: Adaptado de Brown e Gilman (1960, p. 260, tradução nossa).

Na figura 12, há dois direcionais nas díades: em um a direção da força semântica é inequívoca, no outro as duas forças semânticas se opõem. A primeira fileira de díades (no sentido horizontal) envolve conflito no tratamento para os inferiores que não são solidários, e a segunda fileira de díades envolve conflito no tratamento aos superiores que são solidários (cf. BROWN; GILMAN, 1960).

Brown e Gilman (1960) destacam que o poder semântico prevaleceu durante o século XIX, fazendo com que, por exemplo, garçons, soldados comuns e empregados fossem chamados T, pais, patrões e irmão mais velhos fossem chamados V. Os autores ressaltam que, no século XX, a semântica da solidariedade ganhou preeminência. Isso fez com que díades, como as que estão presentes na figura 13, passassem a usar mutuamente o mesmo tratamento, isto é, passassem a retribuir o pronome de solidariedade ou o pronome de não solidariedade. Assim, o sistema da semântica dos pronomes passa de bidimensional para um sistema unidimensional com o recíproco T para o solidário e o recíproco V para o não solidário (cf. BROWN; GILMAN, 1960).

Figura 13 - Díades sociais em equilíbrio



Fonte: Adaptado de Brown e Gilman (1960, p. 260, tradução nossa).

Segundo os autores, a prática de meados do século XX era reinterpretar atributos de poder como atributos de relações simétricas de solidariedade (cf. BROWN; GILMAN, 1960). Dessa forma, as relações como *pai de* seriam reinterpretadas como pertencer à *mesma família que*. Outras situações desse tipo, são ilustradas a seguir:

mais rico do que	➡	mesma renda que
mais nobre que	➡	mesmo tipo de ascendência
mais velho que	➡	mesma idade que
(...)		(...)

Esse novo direcionamento da semântica dos pronomes pode demorar muito tempo para se tornar regra nas línguas ou até mesmo não se tornar, pois é possível que o processo de

mudança semântica⁵¹ regrida a depender dos preceitos sociais e culturais presentes em uma sociedade. Assim, determinados usos linguísticos que estavam em desuso ou pouco frequentes podem retornar com maior frequência de uso na língua. Além disso, cada língua pode ter um comportamento distinto em decorrência das diferenças de poder que emanam em cada sociedade, isto é, cada uma tem normas de comportamento linguísticos e sociais demarcados pela estruturação da sociedade, conforme já ressaltamos.

Normas de comportamento podem ser entendidas como as práticas consistentes que um grupo possui. As implicações da participação em um determinado grupo são muito relevantes; por exemplo, fazer parte uma determinada classe social “sugere um tipo de vida da família, um nível de educação, um conjunto de pontos de vista político etc.”⁵² (BROWN; GILMAN, 1960, p. 277, tradução nossa), o que ajudará a determinar as ações dessa pessoa. Nem sempre o falante faz a escolha certa de um pronome para um determinado contexto, e acaba violando a norma de grupo. Essa escolha inesperada de um pronome possivelmente está relacionada ao fato de o falante considerar a sua relação com o outro como aquela que exige, para o momento, o pronome utilizado. Quebrar as normas de poder pode significar que o falante considera, temporariamente o interlocutor como um estranho ou um íntimo, isso denota que a simpatia é estendida ou retirada. Logo, T pode ser usado como uma forma de manifestar raiva ou desprezo e o V admiração ou respeito (BROWN; GILMAN, 1960). O uso de T com tal valor é comumente inserido “entre pessoas que normalmente trocam V, mas ele pode, é claro, também ser usados por um subordinado a um superior. À medida que a distância social é maior, a derrubada da norma é mais surpreendente e, geralmente, representa uma extremidade maior de sentimento”⁵³ (BROWN; GILMAN, 1960, p. 278, tradução nossa).

O estudo pioneiro de Brown e Gilman (1960) sobre a semântica dos pronomes com base num sistema bidimensional é de suma importância para compreender o funcionamento dos pronomes de segunda pessoa. No entanto, com os processos de globalização do mundo, essa proposta não dá conta de todos os aspectos envolvidos nos usos dos pronomes de tratamento, apresentando, assim, limitações.

⁵¹ Um exemplo é o caso da França em que durante a revolução francesa a semântica da solidariedade foi impulsionada, mas “nos anos posteriores a solidariedade diminuiu e as diferenças de poder que sempre existem em todos os lugares foram expressas mais uma vez” (BROWN; GILMAN, 1960, p. 266, tradução nossa). No original: in later years solidarity declined and the differences of power which always exist everywhere were expressed once more.

⁵² No original: [...] suggests a kind of family life, a level of education, a set of political views and much besides.

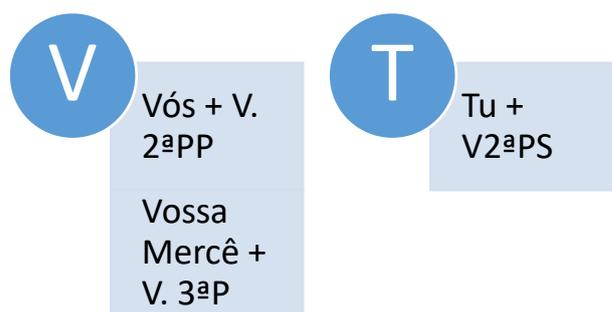
⁵³ No original: [...] introduced between persons who normally exchange V but it can, of course, also be used by a subordinate to a superior. As the social distance is greater, the overturned of the norm is more shocking and generally represents a greater extremity of passion.

3.2.2 A dimensão da neutralidade

Como explanamos, Brown e Gilman (1960) apresentaram uma proposta de interpretação das formas pronominais de tratamento, com base na premissa de que, tanto nas línguas quanto em qualquer sociedade, as dimensões de poder e de solidariedade são atuantes. Partindo dessa proposta, Cook (1997) defende que essas duas dimensões atuam conjuntamente, sem que se excluam uma à outra.

Entretanto, o sistema binário proposto por Brown e Gilman (1960) não é suficiente para línguas em que um sujeito pronominal não é a forma generalizada em paradigmas de segunda pessoa, como é o caso do português (COOK, 1997). O português, assim como outras línguas românicas, herdou do Latim os pronomes **tu** e **vós**, acompanhados de suas desinências verbais, para expressar a segunda pessoa do singular e do plural, respectivamente (cf. BROWN; GILMAN, 1960; FARACO, 2017 [1996]). A escolha da forma plural era feita para designar um único interlocutor em contextos de tratamento cerimonioso, enquanto a escolha da forma singular era para um tratamento não cerimonioso. Para situações também cerimoniosas, surgiu a forma **Vossa Mercê**, por volta do século XIV, usada para se dirigir ao rei (cf. FARACO, 2017 [1996]). O uso desse sujeito nominal era feito com a desinência verbal de segunda pessoa. Essa inserção de um sujeito nominal como forma de tratamento conduziu o verbo para a terceira pessoa, tendo a sua consolidação; ao passo que o uso de honoríficos foi estendido do tratamento ao rei para o tratamento de outras pessoas das partes mais elevadas da pirâmide social, como membros da aristocracia, do clero e casualmente da burguesia (cf. COOK, 1997). Por conseguinte, a esfera V do sistema bidimensional proposto por Brown e Gilman (1960) passou a ser ocupada por duas fórmulas distintas de tratamento: o paradigma tradicional com o sujeito pronominal **vós** acompanhado da desinência de 2ªPP e o sintagma sujeito nominal (**Vossa Mercê**) acompanhado desinência de terceira pessoa do singular. Quanto à esfera T, não houve mudança; prosseguiu a ser ocupada pelo paradigma tradicional com o sujeito pronominal **tu** acompanhado da desinência de 2ªPS (cf. COOK, 1997). A figura a seguir ilustra essa diferenciação entre T e V.

Figura 14 - Esquema dos extremos de diferenciação T-V



Fonte: Elaboração própria a partir de Cook (1997).

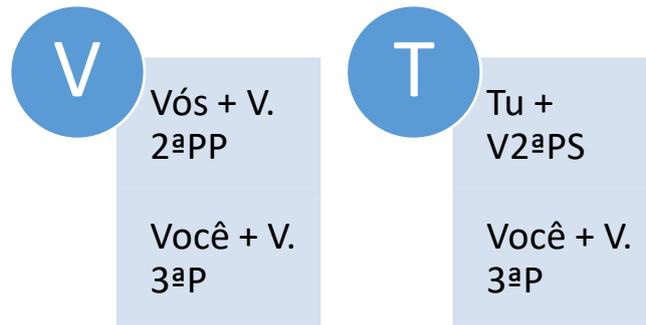
O sintagma **Vossa Mercê** passou por um processo de evolução e transformou-se na forma reduzida **voce** (cf. FARACO, 2017 [1996]). Como esse sintagma nominal passou a transitar nas mais diversas esferas sociais, perdeu o caráter honorífico (outras formas surgiram com esse valor, a exemplo, Vossa Alteza e Vossa Majestade), deixou, assim, de ocupar a esfera V e passou a ocupar a esfera T, e em alguns contextos foi percebido como um tratamento ofensivo⁵⁴ (cf. COOK, 1997). Em vista disso,

Existem a partir de agora dois formatos para a dualidade V-T nas formas de tratamento da língua portuguesa. A esfera V é servida pelo paradigma tradicional como sujeito pronominal *vós* e a desinência verbal da segunda pessoa do plural, e também pelo novo sintagma sujeito nominal +verbo com a desinência da terceira pessoa. A esfera T é servida pelo paradigma tradicional com o sujeito pronominal *tu* e a desinência verbal da segunda pessoa do singular, e também pelo novo sintagma sujeito nominal +verbo com a desinência da terceira pessoa. Temos, assim, tanto para V como para T um formato com um duplo apoio significativo, no pronome denotador do sujeito e na inflexão verbal, e um formato com um único apoio significativo, o do substantivo denotador de sujeito, e uma mesma inflexão verbal usada em comum (COOK, 1997, p. 452).

Dessa maneira, o formato *sujeito sintagma nominal + verbo na terceira pessoa* passou a estar presente nos dois extremos da diferenciação T-V (cf. figura 15). Em consequência, a coexistência dos dois formatos resultou “em certos desvios de ênfase e certas deslocções de uso” (COOK, 1997, p. 452). Entre as mudanças está o fato de que o paradigma tradicional da 2ªPP começou a se tornar obsoleto no PB, pois este foi gradativamente deslocado pelo *sintagma sujeito nominal + verbo na terceira pessoa* (cf. COOK, 1997). Enquanto a 2ªPP canônica tornou-se arcaizante, o seu equivalente para o singular persistiu (cf. COOK, 1997). Vale ressaltar que no português falado a Leste do oceano atlântico, o paradigma tradicional se manteve.

⁵⁴ No português brasileiro atual, essa situação ocorre não por meio do uso da forma **voce**, mas sim com o uso da forma **tu**, pois em alguns lugares esta forma é considerada desrespeito quando dirigida a pessoas mais velhas/idosos.

Figura 15 - Esquema dos extremos de diferenciação T-V



Fonte: Elaboração própria a partir de Cook (1997).

Essas alterações no paradigma pronominal da segunda pessoa fizeram com que a língua portuguesa utilizasse “um modo de tratamento que se subtrai à dualidade V-T” (COOK, 1997, p. 451). Cook (1997) destaca, a esse respeito, que com:

A generalização aos pólos opostos da dualidade V-T da forma de tratamento em formato sujeito nominal + verbo na terceira pessoa resultou, morfológica e sintacticamente, na abertura da possibilidade para a evolução da forma de neutralidade, a qual se tornou obtível pela simples omissão do sujeito pronominal. Esta omissão corresponde a uma transferência para o sintagma de composição nominal e verbal do artifício de economia de esforço sintático que já era prática no latim e continuar a em uso nos paradigmas tradicionais do português (por exemplo Mattoso Câmara 81). O denotador do sujeito - pronominal neste caso - era com frequência dispensado, ficando a indicação da pessoa gramatical ao cuidado exclusivo da desinência verbal (COOK, 1997, p. 451).

Então, no português, surgiu uma maneira de se referir ao outro que não está relacionada nem a dimensão do poder nem da solidariedade, mas sim a neutralidade (representada por N), à qual é gramaticalmente obtida pelo uso do verbo na terceira pessoa sem sujeito expresso. Essa possibilidade de omissão do sujeito com valoração de neutralidade somente veio a ser cultivada quando as mudanças na estrutura social propiciaram uma menor rigidez na hierarquização social (cf. COOK, 1997). O quadro 9, elaborado por Cook (1997), apresenta o cenário geral dos usos das formas de tratamento em português no limiar do século XX. Neste mostram-se as distintas modalidades de abrangimento da hierarquia social da época.

Quadro 9 - Formas de tratamento da Língua Portuguesa no Limiar do século XX

Forma de Tratamento	Modo de Neutralidade (potencial/latente)
Dirigida a um receptor individual:	
Esfera V vós + verbo 2ª p.pl. (arcaizante) ou suj. nom. sing. + verbo 3ª p. sing. =>	Verbo 3ª p. sing.
Esfera T Tu + verbo 2ª p. sing. ou suj. nom. sing. + verbo 3ª p. sing. =>	
Dirigida a um receptor coletivo:	
Esfera V vós + verbo 2ª p.pl. (prática reduzida) ou suj. nom. pl. + verbo 3ª p. pl. =>	Verbo 3ª p. pl.
Esfera T vós + verbo 2ª p. pl. (prática reduzida) ou suj. nom. pl. + verbo 3ª p. pl. =>	

Fonte: Cook (1997, p. 255).

O quadro 9 apresenta duas colunas, uma para as formas de tratamento e a outra para o modo de neutralidade latente/potencial. A primeira coluna expõe as formas de tratamento a partir da focalização de dois aspectos: formas dirigidas a um receptor individual e formas dirigidas a um receptor coletivo. Quanto às formas dirigidas a um receptor individual, a autora demonstra que na esfera V nesse tempo o paradigma herdado do Latim *vós + verbo na 2ªPP* estava em uso (destacando o processo arcaizante deste), e inclui o *sujeito nominal + verbo na terceira pessoa* e relaciona a este a possibilidade de uso de apenas o verbo 3ª PS no modo de neutralidade. A única diferença da esfera T em relação a V é que esta apresenta *tu + verbo na 2ª PS* no lugar do tratamento cerimonioso. De forma semelhante ocorre com as formas dirigidas a um receptor coletivo, tanto na esfera V quanto na T o uso era da forma *vós + verbo na 2ªPP* (com uma prática reduzida) ou o *sujeito nominal plural + verbo na 3ªPP*, a este último está correlacionado o uso do verbo na 3ªPP no modo neutralidade potencial (cf. COOK, 1997).

Quadro 10 - Formas de tratamento da Língua Portuguesa no final do século XX

Forma de Tratamento	Modo de Neutralidade (efectivo)
Dirigida a um receptor individual:	
Esfera V suj. nom. sing. + verbo 3ª p. sing. =>	Verbo 3ª p. sing.
Esfera T Tu + verbo 2ª p. sing. (enfraquecido) ou suj. nom. sing. + verbo 3ª p. sing. =>	
Dirigida a um receptor coletivo:	
Esfera V suj. nom. pl. + verbo 3ª p. pl. =>	Verbo 3ª p. pl.
Esfera T suj. nom. pl. + verbo 3ª p. pl. =>	

Fonte: Cook (1997, p. 255).

O quadro 10 apresenta as formas de tratamento do português utilizadas no final do século XX em que se tem efetivamente um sistema trimensional. Nesse sistema, constata-se uma redução de fórmulas, em comparação ao quadro anterior, “que faculta a efetivação da forma de neutralidade” (COOK, 1997, p. 456). Segundo Cook (1997) o uso da 2ªPP não é mais feito na língua portuguesa (está praticamente extinto) e que a utilização de *tu + verbo 2ªPS* está enfraquecido. A autora destaca que “a opção **tu** subsiste, porém, como única exceção à tendência geral para a prática do sintagma *sujeito nominal + verbo na terceira pessoa*, este com a permanente possibilidade de converter-se em modo de neutralidade, socialmente, de aplicação genérica” (COOK, 1997, p. 457). Nessa perspectiva, é importante ressaltar que:

Ao procurar-se um esquema de interpretação para as formas de tratamento em português, um aspecto fundamental a ter em conta é o facto de ser possível subtrair a produção social de sentido V-T através da omissão do denotador de sujeito. Trata-se de um fenómeno que excede os limites da teorização de Brown e Gilman. Focalizando um formato binário de escolhas V e T, o modelo é insuficiente para o caso português, onde também está em causa uma dimensão de neutralidade. O modelo de Brown e Gilman é um útil instrumento de análise em relação ao primeiro período que considera, até fins do século passado [século XIX], em que o contexto social de realização das formas de tratamento era de rígida estratificação e hierarquização social, mas encontra os seus limites quando se procuram analisar as formas de tratamento num contexto social com uma mais ampla margem para negociação de modo de inter-relacionamento (COOK, 1997, p. 458).

Em virtude disso, Cook (1997) defende que o conflito do sistema bidimensional de tratamento seja solucionado por intermédio dessa outra dimensão semântica atuante, a da neutralidade. Em termos gerais, a dimensão da neutralidade emerge quando se quer “evitar uma tomada de posição dentro do contraste formal-informal” (COOK, 1997, p. 451) e permite que os interlocutores negociem posições de relacionamento dentro das esferas/dimensões. Essa dimensão não se opõe as demais, mas opera em coexistência com o sistema T-V instituído. No entanto, Cook (1997) pondera que:

Um aspecto importante a considerar é a ambivalência das mensagens de sentido social que as formas de tratamento podem transmitir e a dificuldade em conciliar essa ambivalência como postulado de Brown e Gilman, de que em sociedades mais igualitárias se verifica uma redução do factor poder e a tendência é para T ser usado no círculo da solidariedade e V com estranhos. Como vimos para *você*, uma forma T pode ser portadora tanto de solidariedade positiva como negativa. Também *tu*, geralmente associado à solidariedade positiva, pode ser veículo de uma intenção patronizante de solidariedade negativa. De facto, em muitas sociedades se pratica o uso de uma mesma forma de tratamento tanto para sinal de intimidade como para instrumento de controlo em afirmação de ascendência (por ex. Chaika). Do mesmo modo, estratégias V podem ter gradações várias no campo da solidariedade, pois a cortesia é um eixo de dois pólos, um positivo e outro negativo, como demonstrado por Brown e Levinson (1978, 1987). Em consequência, o estreitamento de distância social das opções T pode expressar afeição ou intromissão insultuosa; enquanto que o distanciamento social em V pode expressar respeito ou cortesia mas pode também insinuar exclusão, e o uso de uma opção V inapropriada pode servir a intenção de ofender. Numa língua com a riqueza semântica de um sujeito nominal, como se verifica em português, haverá ainda uma maior variedade de matizes possíveis tanto em T como em V para realizações simétricas e assimétricas de relacionamento (COOK, 1997, p. 458).

Portanto, provavelmente, a diminuição de poder em favorecimento da solidariedade, que foi identificada por Brown e Gilman (1960), será interpretada de maneira mais produtiva “como redução de aceitabilidade de meios de expressão gramaticalmente estabelecidos para um relacionamento de disreciprocidade, pois o exercício de poder é uma constante em qualquer sociedade” (COOK, 1997, p. 458).

A análise de Cook (1997) é voltada principalmente para o português de Portugal. Torna-se essencial verificar no português brasileiro atual como se configura esse sistema tridimensional. É preciso realizar uma pesquisa para verificar se a forma *você* opera em apenas uma das esferas ou em ambas. Se for nas duas esferas, essa forma pronominal poderá ser considerada como pertencente a dimensão da neutralidade, como propõe Souza (2011). Além disso, é importante averiguar ainda em qual esfera a redução *cê* opera no português brasileiro. A presente pesquisa busca, dentre seus objetivos, investigar esses aspectos.

3.2.3 A relação entre dêixis social, face e polidez

Vários estudos foram realizados após a publicação da proposta de Brown e Gilman (1960) sobre os pronomes de poder e de solidariedade. Essas investigações subsequentes evidenciaram que as dimensões de poder e de solidariedade não são suficientes para caracterizar o uso social adequado dos pronomes (cf. SIEWIERSKA, 2004), como demonstramos na subseção anterior. Outras dimensões são necessárias para que isso ocorra, como, por exemplo: “posição social/classe, status, cargo, geração, formalidade, informalidade, discurso público, discurso privado, intimidade, distância social e alto grau de excitação emocional”⁵⁵ (MÜHLHÄUSLER; HARRÉ, 1990, p. 132 apud SIEWIERSKA, 2004, p. 214, tradução nossa). Todas essas dimensões estão intrinsicamente relacionadas às dêixis sociais, as quais podem ser expressas de diferentes maneiras no sistema individual de uma língua, sendo, possivelmente, a forma mais comum por meio “da manipulação das distintas semânticas refletidas em um dado paradigma pessoal, como pessoa, número, inclusão e gênero”⁵⁶ (SIEWIERSKA, 2004, p. 214, tradução nossa).

A maioria dos dêiticos sociais de uma língua é marcador de polidez (há um subconjunto menor que é usado para codificar uma atitude de denegrição/impolidez), dentre estes estão os pronomes pessoais de 2ªPS (cf. TROUGOTT; DASHER, 2002). A polidez é um mecanismo linguístico utilizado para a preservação de face em qualquer situação interlocutiva, visto que todo ato linguístico, em uma interação face a face, é uma atividade inerentemente ameaçadora pelo fato de que, ao entrarem em contato um com outro, os participantes ocasionam um desequilíbrio das faces (BROWN; LEVINSON, 2011[1987]). Decorre disso, o fato de que, em toda atividade verbal, na qual não se cumprem as máximas conversacionais de Grice, o uso de recursos de proteção à face pode ocorrer, ou seja, os falantes teriam que utilizar estratégias de polidez para reparar o dano causado a face. Portanto, em todo o processo da interação há o monitoramento pelos falantes, mesmo que inconscientemente, das faces envolvidas e estes recorrem a estratégias de polidez para evitar possíveis conflitos. A polidez, nesse âmbito, é compreendida como uma estratégia linguística empregada com o intuito de evitar conflitos na interação verbal. Em termos gerais, consiste em uma estratégia para preservarmos a nossa face

⁵⁵ No original: rank, status, office, generation, formality, informality, public discourse, private discourse, intimacy, social distance and high degree of emotional excitement.

⁵⁶ No original: the manipulation of the semantic distinctions reflected in a given person paradigm, such as person, number, inclusivity and gender.

e a face do outro a fim de propiciar uma comunicação econômica e eficaz, sem atritos (BROWN; LEVINSON, 2011[1987]).

A visão de polidez de Brown e Levinson (2011 [1987]) parte da premissa de que os participantes de uma situação interlocutiva são dotados de duas características: a face e a racionalidade. Esta é entendida como a execução de uma forma específica “de raciocínio [...] que garante inferências de fins ou objetivos para meios que satisfaçam esses fins, [...] um sistema de raciocínio prático deve permitir que se passe dos fins aos meios e meios adicionais, preservando a "satisfação" desses meios”⁵⁷ (BROWN; LEVINSON, 2011[1987], p. 48, tradução nossa). Aquela é concebida pelos autores⁵⁸ como a autoimagem pública que cada falante constrói de si e que quer proteger dos possíveis danos no decurso de uma interação face a face. O termo face refere-se a “algo que está emocionalmente revestido, que pode ser perdida, mantida ou reforçada, e deve ser constantemente cuidada durante a interação”⁵⁹ (BROWN; LEVINSON, 2011[1987], p. 61, tradução nossa). Em termos gerais, os participantes de uma situação interlocutiva cooperam para que a face seja mantida durante a interação face a face e pressupõem a cooperação do interlocutor, baseando-se no fato de que, no processo de interação, a vulnerabilidade da face é mútua (BROWN; LEVINSON, 2011[1987]).

A noção de face está atrelada a dois tipos de desejos que os falantes concedem uns aos outros: o desejo de ser aprovado, estimado, apreciado (face positiva) e o desejo de não sofrer imposição, de preservar seu espaço pessoal, de ter as suas ações livres/de não ser impedido de agir (face negativa) (BROWN; LEVINSON, 2011[1987]). Em qualquer situação de interação social, conforme ressaltamos, as faces dos interactantes estão em risco: tanto o ouvinte quanto o falante produzem, em sua maioria, atos (verbais e não verbais) que ameaçam a face de ambos, visto que nem sempre há compatibilidade de interesse entre eles. Logo, as faces dos participantes de uma interação são “ao mesmo tempo e contraditoriamente, o alvo de ameaças permanentes e o objeto de um desejo de preservação” (KERBRAT-ORECCHIONI, 2006, p. 80). Tais atos verbais/não verbais são chamados de Atos Ameaçadores à Face (face Threatening Acts – FTAs) por Brown e Levinson (2011[1987]).

⁵⁷ No original: of reasoning [...] which guarantees inferences from ends or goals to means that will satisfy those ends [...] a system of practical reasoning must allow one to pass from ends to means and further means while preserving the ‘satisfactoriness’ of those means.

⁵⁸ Os autores inspiram-se no conceito de face proposto por Goffman (1967).

⁵⁹ No original: is something that is emotionally invested, and that can be lost, maintained, or enhanced, and must be constantly attended to in interaction.

Um FTA pode atingir, simultaneamente, quatro faces em uma situação de interlocutiva entre dois participantes. Em consequência disso, existem quatro categorias de atos de fala, conforme o tipo de face que ameaça, os quais são descritos no quadro a seguir:

Quadro 11 - Atos Ameaçadores a Faces

	Atos Ameaçadores a Face Negativa	Atos Ameaçadores a Face Positiva
<i>Afetam ao Falante</i>	1. Atos que violam o seu território, como: expressar agradecimentos, aceitar ofertas, promessas involuntárias e ofertas, assumir a gafe do outro.	2. Atos autodegradantes, como: pedir desculpas, aceitar um elogio, confessar culpa, autocriticar-se, falta de cooperação.
<i>Afetam ao Ouvinte</i>	3. Atos que violam o seu território, como, por exemplo, as perguntas indiscretas; Atos diretivos, como: ordenar, pedir, sugerir, aconselhar, lembrar (o falante indica que o ouvinte deve lembrar-se de fazer algo), ameaçar, advertir; Fazer ofertas, promessas, elogios.	4. Atos que ameaçam o narcisismo do outro, tais como: desaprovar, criticar, queixar-se, repreender, acusar, insultar, desafiar; Menção de temas tabus, expressão de violência.

Fonte: Elaborado a partir de Brown e Levinson (2011[1987], p. 65-68).

Observa-se, a partir das informações do quadro acima, que os FTAs podem acontecer em dois direcionais: do falante para o falante ou do falante para o ouvinte. As categorias de atos de fala 1 e 2 são autoameaçadores, uma vez que se referem à própria face do falante; já as categorias 3 e 4 referem-se à atitude do falante para com o ouvinte. Enquanto as categorias 1 e 3 são atos ameaçadores da face negativa, as categorias 2 e 4 são atos ameaçadores da face positiva (cf. BROWN; LEVINSON, 2011[1987]).

A possibilidade de ocorrer atos de fala ameaçadores durante uma interação faz emergir o uso de estratégias de polidez para minimizar esse poder ofensivo em virtude do fato de o falante ter três desejos na situação interlocutiva: a) comunicar o conteúdo do FTA, b) ser eficiente ou urgente e c) manter as faces envolvidas em qualquer grau (cf. BROWN; LEVINSON, 2011[1987]). Brown e Levinson (2011[1987]) afirmam que, a depender das circunstâncias para executar um FTA, um determinado tipo de estratégia de polidez é utilizado a saber: i) positiva - age como “uma ação reparadora” direcionada à preservação da face positiva do interlocutor, isto é, a ameaça do FTA à face positiva do ouvinte é evitada ou minimizada; ii) negativa – é uma ação reparadora direcionada à preservação da face negativa do interlocutor, exercendo, assim, a função de minimizar ou anular os efeitos da imposição de um FTA; ou iii) encoberta – é aquela que possibilita o falante realizar o FTA de forma implícita, sem clareza; de modo que não se possa constatar qual a intenção real deste (falante) ao comunicar o ato (BROWN; LEVINSON, 2011[1987]).

É importante ressaltar que três variáveis contextuais condicional o tipo de estratégia a ser usada em uma situação interlocutiva: i) a distância social existente entre os interlocutores –

trata-se de uma dimensão simétrica de semelhança/diferença e se refere ao grau de familiaridade e solidariedade entre o falante e o ouvinte; ii) o poder relativo existente entre os interlocutores – consiste em uma dimensão assimétrica, está relacionado ao poder que o falante exerce sobre o ouvinte e vice-versa; e iii) o grau do custo da imposição de um ato comunicativo – é definido cultural e situacionalmente e se refere aos riscos intrínsecos ao ato que irá realizar, ou seja, pode ser aprovado ou não pelo interlocutor (BROWN; LEVINSON, 2011[1987], p. 76-78). Portanto, as estratégias são usadas de acordo com o objetivo que se quer atingir na situação comunicativa e estão relacionadas com os tipos de relações existentes entre os interlocutores, evidenciando a relação existente entre convenções socioculturais e princípios de usos da língua. Nesse âmbito, vale ressaltar que os usos das formas de 2ªPS estão fortemente correlacionadas a essa questão.

Qualquer processo comunicativo é regulado por normas sociais e estas determinam o comportamento linguístico dos participantes na condução da interação verbal entre pares, sendo a atuação da polidez um fator essencial para manter o equilíbrio na interação. As variáveis idade, sexo/gênero, classe social, bastantes difundidas nos estudos sociolinguísticos, também interferem no modo como a polidez se manifesta. Ademais, a manifestação da polidez varia de cultura para cultura - o que é polido em uma determinada cultura pode não ser em outra em virtude das mudanças de costumes e de valores – isto é, está fortemente atrelado ao contexto sociocultural de um lugar (cf. ARAUJO, 2014). Existem diferenças na construção social da polidez (cf. TROUGOTT; DASHER, 2002).

É a partir do cruzamento dessas informações socialmente demarcadas e dos objetivos a serem alcançados no processo interativo que a escolha da estratégia é feita. A título de exemplificação, em uma interação entre um jovem e um idoso a tendência é que emergjam estratégias de polidez, como o uso da forma pronominal **Senhor**, em decorrência das normas sociais que induzem seu uso como forma de deferência, em respeito aos mais velhos. Em contrapartida, o uso dessa forma pronominal pode ser interpretado de forma distinta entre os interlocutores: na tentativa de demonstrar respeito pelo interlocutor com a intenção de preservá-lo a face negativa, o falante pode estar, na verdade, ameaçando a face positiva do seu interlocutor ao usar a formalidade, contrariando as expectativas deste de ser reconhecido como parte integrante do grupo social; ou ao contrário, se o falante utilizar uma forma que denota maior intimidade, como o pronome **tu**, ao tentar diminuir a distância social entre os interlocutores, a atitude do falante pode ser interpretada como uma ameaça à face negativa do ouvinte, como uma excessiva familiaridade que não é aceita pelo interlocutor. Destarte, o valor de cada estratégia é definido no momento da interlocução e não pode estar dissociado deste, o que demonstra ainda mais a associação entre esta e a dêixis social.

4 O USO VARIÁVEL DAS FORMAS PRONOMINAIS TU, VOCÊ E CÊ NO PORTUGUÊS BRASILEIRO

Estudar fenômenos linguísticos variáveis é importante não só para descrever como estes ocorrem, mas também para compreender o próprio funcionamento da língua. Quando temos um fenômeno linguístico que já foi amplamente pesquisado em uma determinada língua (como o português brasileiro), é de suma relevância fazer um levantamento dos estudos desenvolvidos em diversas localidades do país, a fim de realizar uma sistematização dos resultados e, a partir desta, identificar tendências/padrões de uso e fazer generalizações (com as ressalvas pertinentes) – conforme apregoa a Sociolinguística Variacionista.

A realização de sistematizações dos resultados de pesquisas brasileiras tem sido foco de muitos estudiosos. Scherre *et al.* (2015) realizaram a sistematização de estudos sincrônicos sobre a variação dos pronomes ‘tu’ e ‘você’ que utilizaram dados da oralidade e que foram realizados até 2012. Scherre, Andrade e Catão (2019) realizaram um novo mapeamento para agregar as pesquisas desenvolvidas a partir 2013. O primeiro levantamento focaliza os resultados de distribuição das frequências de usos das formas em todas as regiões do país, sem abordar os resultados referentes às variáveis condicionantes do fenômeno em questão.

Diante desse quadro, constata-se a necessidade de se realizar um estudo de sistematização que abarque também as variáveis condicionantes do fenômeno, a fim de discutir o que foi aferido até então pelas pesquisas científicas. Deste modo, o objetivo central desta seção é realizar uma revisão sistemática dos estudos sobre a variação entre as formas **tu**, **você** e **cê** com enfoque nas principais variáveis condicionantes de uso destas. Essa sistematização nos permite comparar os resultados dos estudos já realizados com os obtidos em nossa pesquisa e, assim, verificar se há diferenças.

Realizamos um levantamento dos estudos feitos, a partir dos anos 2000, a respeito das formas variantes **tu**, **você** e **cê** no português brasileiro. É importante ressaltar que nem sempre as discussões apresentadas aqui refletem necessariamente as conclusões dos estudos consultados. A partir dos dados absolutos das ocorrências das formas, alguns resultados foram recalculados de forma a atender aos propósitos da presente pesquisa de comparabilidade e de generalização. Além disso, realizamos análises inferenciais univariadas na plataforma R

(CORE TEAM, 2020) para cada variável independente, por meio do teste de qui-quadrado Pearson⁶⁰ utilizando a função *chisq.test*.

A seção está estruturada em seis subseções. Na primeira, apresentamos os resultados dos estudos referente à distribuição das formas variantes no português brasileiro com base no mapeamento feito por Scherre, Andrade e Catão (2019). Em seguida, abordamos os resultados dos estudos a respeito das variáveis independentes de cunho interacional. Nas subseções 4.3, 4.4 e 4.5, expomos os resultados das pesquisas acerca das variáveis estilísticas, estrutural/formal e sociais, respectivamente. Na subseção 4.6, discutimos os problemas metodológicos e as limitações dos estudos realizados. Passemos, então, a discussão desses resultados.

4.1 DISTRIBUIÇÃO DAS FORMAS PRONOMINAIS DE 2ªPS NO PORTUGUÊS BRASILEIRO

Scherre *et al.* (2015) realizaram um mapeamento das pesquisas desenvolvidas sobre a variação das formas de 2ªPS no português brasileiro com dados de fala do final do século XX e início do século XXI (até 2012). Os autores, a partir do levantamento dos resultados apresentados em diversos estudos, propõem seis subsistemas de tratamento para o português brasileiro, considerando a realização da (não) concordância verbal, a saber: i) **só você**, com as variantes **você/cê/ocê**; ii) **mais tu** (> 60%) e concordância explícita com **tu baixa** (< 10%); iii) **mais tu** (> 60%) e concordância explícita com **tu alta** (de 40% a 60%); iv) **tu/você** (**tu** < 60%) e concordância explícita com **tu baixa** (< 10%); v) **tu/você** (**tu** < 60%) e concordância explícita com **tu média** (de 10% a 39%); e vi) **você/tu - tu** de 1% a 90% sem concordância explícita com o **tu**. A figura abaixo evidencia a distribuição das formas variáveis no território nacional.

⁶⁰ Assim como nas análises realizadas na seção 2.4 desta tese o nível de significância considerado para os testes de qui-quadrado foi de 0.05.

Figura 16 - Mapa dos seis subsistemas dos pronomes de segunda pessoa do singular do português brasileiro no final do XX e início do século XXI, o Brasil de 1709 e o Brasil do século XXI



Fonte: Scherre *et al.* (2015, p. 142), com adaptações de Scherre, Andrade e Catão (2019).

O panorama geral dos usos dos pronomes de 2ªPS do PB, representado no mapa 1 da figura 16, demonstra que o subsistema “só **você**, com as variantes **você/cê/ocê**” é predominante nas amostras de fala analisadas pelos pesquisadores. Os resultados dispostos no mapa evidenciam que o fenômeno ocorre de maneira bastante diversa em muitas regiões do Brasil: na região Sul, há o predomínio da forma **tu** no Rio Grande do Sul; na região Sudeste, o uso predominante da variante **você**; e, em outras regiões, o uso das duas formas. A distribuição dos subsistemas em questão (mapa 1) assemelha-se mais ao mapa do Brasil de 1709 (mapa 2) do que ao mapa das regiões do Brasil do XXI (mapa 3) (SCHERRE; ANDRADE; CATÃO, 2019), visto que o quantitativo de estados era menor no século XVIII. Com base no mapa 1, presente na figura 16, os subsistemas estão distribuídos da seguinte forma:

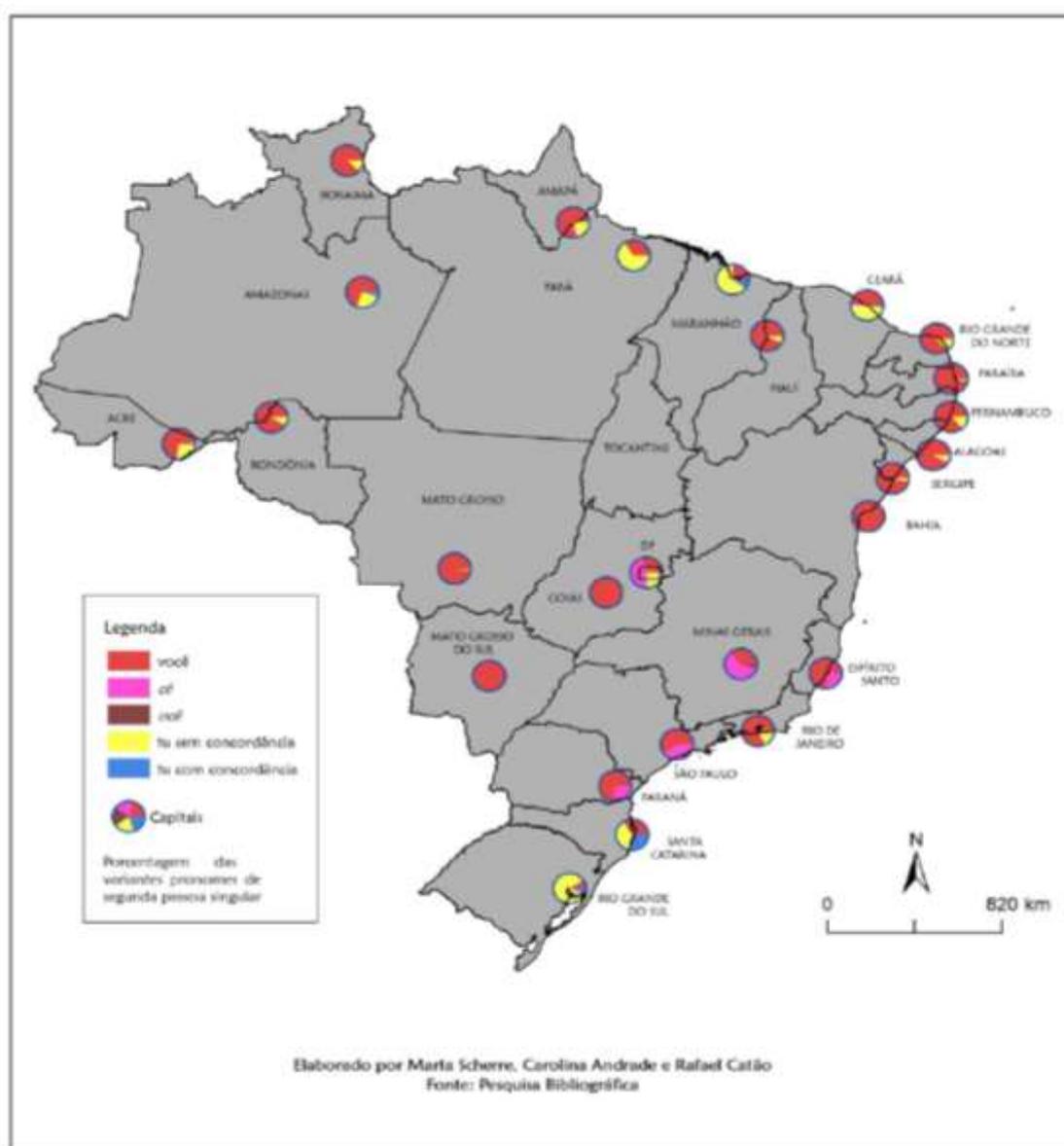
1. O subsistema **só você** é suprarregional, mas se concentra na área central do país com bastante uniformidade, envolvendo a antiga capitania de São Paulo, segundo o mapa que mostra a divisão administrativa do Brasil em 1709 (cf. Mattos, 2013, p. 36). São representantes da região Centro-Oeste os estados de Goiás, Mato Grosso e Mato Grosso do Sul. Neste caso, fica excluído apenas o Distrito Federal ou a Grande Brasília. A região Sudeste é representada pelos do Espírito Santo, Minas Gerais (exceto a cidade de São João da Ponte) e São Paulo (exceto Santos). A região Nordeste é representada pela Bahia, basicamente pela sua capital, Salvador; a região Norte, pelo Tocantins; e a região Sul, pelo estado do Paraná.
2. O subsistema **tu com concordância baixa** é encontrado em regiões dos extremos, Norte e Sul. O estado do Amazonas é o representante da região Norte. Na região Sul, é o Rio Grande do Sul.
3. O subsistema **tu com concordância alta** tem representação na região Norte, com o estado do Pará; e, na região Sul, é o estado de Santa Catarina.
4. O subsistema **tu/você com concordância baixa** é encontrado nas regiões Nordeste e Sul. Os estados representantes da região Nordeste são Maranhão e Tocantins; e o da região Sul é Santa Catarina.
5. O subsistema **tu/você com concordância média** é encontrado nas significativo de estados: Maranhão, Piauí, Ceará, Paraíba, Pernambuco. Na região Norte, o estado representativo é o Amazonas; e na região Sul, Santa Catarina.
6. O subsistema **você/tu sem concordância** só não tem representante na região Sul, pelo menos por ora, pela nossa interpretação. Na região Centro-Oeste, é representado, nos dias atuais, pelo Distrito Federal; na região Sudeste, pelos estados do Rio de Janeiro, São Paulo e Minas Gerais; na região Nordeste, pelos estados do Maranhão e Bahia; e, na região Norte, pelos estados de Roraima e Acre. (SCHERRE *et al.*, 2015, p. 142-143)

Tais resultados demonstram que as formas variantes podem ter um emprego distinto a depender da área geográfica. Com o desenvolvimento de várias pesquisas sobre o fenômeno em questão, Scherre, Andrade e Catão ((2019) fizeram um mapeamento mais refinado dos resultados de distribuição das formas, levando em consideração cinco variantes de pronomes de 2ªPS: **você, cê, ocê, tu sem concordância** e **tu com concordância**.

A partir dessa análise, os autores elaboraram os mapas contrapondo os estudos realizados em capitais com aqueles de não capitais. Na figura 17, que, na maioria das capitais do Nordeste, prevalece o uso da variante **você**. Os poucos dados da forma **tu** ocorrem sem concordância verbal, apenas a capital do Ceará e a capital do Maranhão possuem um uso pouco mais expressivo da forma **tu sem concordância**; sendo que, na capital São Luís, há alguns dados de **tu com concordância**. Semelhantemente ocorre nas capitais das regiões Norte e Centro-Oeste: uso majoritário da forma **você** nas amostras analisadas, com exceção da capital do Pará (predomínio da forma **tu sem concordância**) e do Distrito Federal (uso das formas **você, tu sem concordância verbal** e, principalmente, **cê**). Os usos das variantes nas regiões Sudeste e Sul são mais diversificados. Na região Sudeste, temos os seguintes resultados: a capital de Minas Gerais apresenta dados de **você** e **cê**, sendo o uso desta última forma mais

frequente; a capital do Espírito Santo e a capital de São Paulo apresentam usos das formas **você** e **cê**, principalmente da primeira variante; e a capital do Rio de Janeiro apresenta dados de uso da forma **você** (majoritariamente) e da forma **tu sem concordância**. Na região Sul, a capital do Paraná apresenta usos das variantes **você** e **cê**, sobretudo, da primeira forma; já, na capital de Santa Catarina, temos o uso das variantes **você**, **tu com concordância** e **sem concordância verbal**. Quanto à capital do Rio Grande do Sul, esta apresenta o uso predominantemente da forma **tu sem concordância verbal** e apenas alguns dados das variantes **você** e **tu com concordância**.

Figura 17 - Mapa de distribuição das variantes de pronomes de 2ªPS em capitais dos estados brasileiros



Fonte: Scherre, Andrade e Catão (2019).

A figura 18 relaciona os usos das variantes às áreas de não capitais que foram objetos de descrição nas pesquisas científicas sobre o fenômeno em questão. Constata-se que a

distribuição das variantes nas áreas de não capitais se modifica um pouco da que ocorre nas capitais: nas regiões Norte e Nordeste, há um aumento do uso da variante **tu sem concordância**; na região Sudeste, diferentemente dos resultados para as capitais, a forma **ocê** aparece como uma das variantes utilizadas pelos falantes de Minas Gerais e o uso do **tu sem concordância** aparece como uma das formas utilizadas em São Paulo e no Paraná. Já na região Sul as áreas de não capital em Santa Catarina há um uso mais expressivo da variante **você** e no Rio Grande do Sul mantém-se o uso majoritário de **tu sem concordância**. Em relação às áreas de não capitais da região Centro-Oeste, não foram constatadas pesquisas científicas nesse âmbito.

Figura 18 - Mapa de distribuição das variantes de pronomes de 2ªPS em áreas de não capitais dos estados brasileiros



Fonte: Scherre, Andrade e Catão (2019).

No mapa dos seis subsistemas dos pronomes de 2ªPS do português brasileiro, presente na figura 16, não há menção de pesquisas com esse escopo em Sergipe. O mapa de distribuição

das formas na fala de informantes de capitais do Brasil (figura 17) evidencia que houve algum estudo com amostras de falantes sergipanos após 2012. Trata-se do estudo de Cardoso *et al.* (no prelo apud SCHERRE, ANDRADE, CATÃO, 2019) realizado com dados do Atlas Linguístico do Brasil (ALIB), o qual focaliza o tratamento do interlocutor em capitais. Todavia, não consiste em um estudo que busca identificar os condicionamentos linguísticos, estruturais/formais, interacionais, estilísticos e sociais nos usos dos pronomes de 2ªPS, como é a proposta da presente tese. A pesquisa desenvolvida por Cardoso *et al.* (no prelo apud SCHERRE, ANDRADE, CATÃO, 2019) aponta que Sergipe faz parte do sexto subsistema proposto por Scherre *et al.* (2015) – **você/tu**, com ocorrências de **tu** de 1% a 90% sem concordância.

Embora a proposta de mapeamento dos estudos desenvolvidos sobre os pronomes de 2ªPS seja válida e relevante para que se tenha um panorama geral dos usos das formas de tratamento, não se pode ignorar que há problemas em se fazer generalizações e comparações a partir de um material de natureza tão heterogênea, como os próprios autores do texto reconhecem. Isso é decorrente do fato de que as pesquisas não foram, de certo modo, feitas para serem comparadas, já que não houve uma “agenda de pesquisa” para isso e cada pesquisador desenvolveu o estudo conforme suas possibilidades. Na última subseção desta seção, retomamos essa problemática.

4.2 VARIÁVEIS INTERACIONAIS

As variáveis interacionais estão intrinsecamente relacionadas ao contexto discursivo, às relações estabelecidas/existentes entre o locutor/falante e o interlocutor/ouvintes. Apresentamos, nas duas subseções a seguir, os resultados das pesquisas que controlaram as seguintes variáveis: grau de intimidade entre os interlocutores e (as)simetria quanto ao sexo/gênero dos interactantes.

4.2.1 Grau de intimidade entre os interlocutores

Brown e Gilman (2003 [1960]) ressaltam que o uso dos pronomes pessoais de segunda pessoa do discurso está relacionado ao grau de intimidade entre os interlocutores. Alguns pesquisadores brasileiros controlaram essa variável independente com o intuito de verificar o

efeito desta nos usos das formas pronominais **tu**, **você** e **cê** na fala. Na tabela a seguir, estão dispostos os resultados obtidos por três estudos (Mota (2008); Martins (2010) e Guimarães (2014)) a partir do controle do grau de intimidade entre os interlocutores.

Tabela 8 - Uso de **tu** em função do grau de intimidade entre os interlocutores

Estudos	Íntimo		Não íntimo		Qui-quadrado
	A/T	%	A/T	%	
Mota (2008) São João da Ponte/MG	42/156	26	5/313	1	$X^2 (1, N= 469) = 71.27,$ $p < 0.00001$
Martins (2010) Tefé/AM	348/430	80,9	172/376	45,7	$X^2 (1, N= 806) = 106.95,$ $p < 0.00001$
Guimarães (2014) Fortaleza/CE	767/1409	54,4	25/146	17,1	$X^2 (1, N= 1555) = 72.21,$ $p < 0.00001$

Fonte: Elaborada própria.

Os resultados obtidos por Mota (2008) evidenciaram que o efeito da variável independente grau de intimidade entre os interlocutores é estatisticamente significativo ($x^2 = 71.27, p < 0.00001$) nos usos das formas **tu** e **você** em São João da Ponte/MG. Na amostra de entrevistas analisada pela autora, tanto em contextos íntimos quanto em não íntimos, o uso mais frequente foi da forma **você**. Entretanto, ao analisarmos os resultados considerando o total de ocorrências das formas, constatamos que o uso da forma **tu** foi favorecido quando os interlocutores eram íntimos (42 dados da forma **tu**, de um total de 47, ocorreram no contexto de maior intimidade), já a forma **você** foi usada com maior frequência quando os interlocutores não eram íntimos.

No estudo de Martins (2010), o qual foi realizado com dados de fala decorrentes entrevistas com informantes da cidade de Tefé/AM, essa variável independente também se mostrou significativa ($X^2 = 106.95, p < 0.00001$). Os resultados demonstraram que, nas interações em que os interlocutores eram íntimos, a forma **tu** foi favorecida, apresentando um percentual de 80,9% (348/430) das ocorrências; ao passo que os contextos interlocutivos de não intimidade entre os falantes apresentaram um percentual de 45,7% (172/376) dessa variante. É importante ressaltar que, das 19 entrevistas analisadas, três e meia foram feitas sem o conhecimento dos informantes que participaram da pesquisa. Essa análise conjunta de dados coletados por métodos distintos possivelmente interferiu nos resultados.

Guimarães (2014) controlou o efeito do grau de intimidade entre os interlocutores da cidade de Fortaleza/CE. Assim como ocorreu nos dois estudos anteriores, o efeito da variável foi estatisticamente significativo ($X^2 = 72.21, p < 0.00001$). Os resultados obtidos pela autora demonstraram que, nas situações interlocutivas em que os participantes eram íntimos, houve,

sutilmente, um maior uso da forma **tu** (767/1409 - 54,4%) e, naquelas em que não havia intimidade entre os interlocutores, essa variante foi desfavorecida (25/146 - 17,1%) ou seja, a forma **você** foi favorecida com um percentual de 82,9%. É relevante ressaltar que, diferentemente das outras duas pesquisas, os dados de fala analisados por Guimarães (2014) não são decorrentes de gravação de entrevistas, mas sim de diálogos naturais.

Constata-se, portanto, que os resultados desses três estudos apresentam uma direcionalidade semelhante entre si ao evidenciarem que o traço íntimo favorece o uso do pronome **tu** e o traço não íntimo favorece o uso do pronome **você**. Além disso, verificamos, a partir da análise inferencial univariada (qui-quadrado), que a variável independente grau de intimidade entre os interlocutores condiciona o uso das formas pronominais em estudo, fazendo com que o falante acomode seu estilo tendo como parâmetro a relação existente com o interlocutor. Tais resultados comprovam a premissa clássica de Brown e Gilman (1960) de que essa variável é um fator condicionante nos usos das formas pronominais em questão.

Cabe fazermos uma ressalva quanto ao modo como a variável independente grau de intimidade entre os interlocutores foi controlada em cada um dos estudos. Mota (2008) e Martins (2010) fizeram o controle dessa variável utilizando como parâmetro o tipo de díade, ou seja, o tipo de relação existente entre os interlocutores, como, por exemplo, se a interação ocorreu entre: pai e filho, amigos, desconhecidos, entrevistador e entrevistado, irmãos, chefe e subordinado etc. Quanto mais os interlocutores convivem no dia a dia, interagem um com o outro, são mais próximos um do outro, significa que eles são mais íntimos; quando isso não ocorre, os interlocutores são menos íntimos.

Contrariamente a estes pesquisadores, Guimarães (2014) controlou o grau de intimidade a partir do índice de troca de turnos entre os informantes, da proporção da duração da fala e do tipo de assunto. Com base nisso, a autora considerou como alto grau de intimidade entre os falantes as conversações “em que havia um alto índice de troca de turno entre os informantes e eles falavam quase na mesma proporção” e como baixo grau de intimidade as conversações “impessoais que tratavam do passado ou apenas do trabalho dos informantes, com textos longos e com pouca troca de turno entre os interlocutores” (GUIMARÃES, 2014, p. 126). O controle realizado dessa forma é bastante controverso, pois os critérios estabelecidos não permitem mensurar de fato o grau de intimidade entre os falantes; por exemplo, o falante pode conversar sobre o passado ou trabalho com qualquer pessoa, falar menos por não ser íntima do interlocutor ou por ser tímida, ser mais objetiva. Logo, a comparação dos resultados das pesquisas deve ser considerada de forma cautelosa, tanto pelas diferenças de tipo de coleta quanto pela forma como a variável grau de intimidade foi controlada.

Tabela 9 - Uso de CÊ específico e relação entre os interlocutores (Infs e Docs – GESOL-SP-2000)

Grau de intimidades entre os interlocutores	Cê (REF Específica)					
	Informantes			Documentadores		
	Nº	%	p.r.	Nº	%	p.r.
+ próximo	92/162	56,8	.56	354/592	59,8	.63
- próximo	58/124	46,8	.43	113/409	27,6	.32
Total (T)	150/286	52,4	Ranger: .13	467/1001	46,7	Range: .31
	<i>Input: 0.554 l.l. = -319.398 p < 0.02</i>			<i>Input: 0.458 l.l. = -619.506 p = 0.000</i>		

Fonte: Nascimento (2011, p. 142, adaptado).

Nascimento (2011) constatou, na amostra de dados de fala de São Paulo/SP, que as diferenças de usos das variantes **você** e **cê** em relação ao grau de intimidade entre os interlocutores é estatisticamente significativa. Os resultados da autora evidenciaram que a variante **cê** é favorecida entre interlocutores mais próximos na expressão de referência determinada e desfavorecida entre os menos próximos, tanto na fala dos informantes quanto na fala dos documentadores. Embora as diferenças percentuais não sejam altas, temos indícios de que de fato o grau de intimidade atua na escolha das variantes.

4.2.2 (As)simetria quanto ao sexo/gênero dos interlocutores

A determinação do tipo de relação existente entre os interlocutores está correlacionada às dimensões de poder e de solidariedade. Quando os interlocutores possuem, por exemplo, o mesmo sexo/gênero, o mesmo grau de escolaridade, a mesma faixa etária, fazem parte do mesmo grupo social, exercem a mesma profissão ou o mesmo papel dentro da hierarquia institucional, possivelmente, estes terão um comportamento linguístico mais ou menos solidário em uma situação interlocutiva. Nesse caso, temos uma relação simétrica, pois a interação ocorre entre iguais (BROWN; GILMAN, 1960).

Nas situações conversacionais em que os interlocutores possuem sexos/gêneros diferentes, faixas etárias diferentes, fazem parte de grupos sociais diferentes ou ocupam posições distintas na hierarquia institucional, é provável que haja a atuação da dimensão do poder, ou seja, um dos interlocutores possui poder sobre o outro, fazendo com que o outro tenha, em tese, um comportamento linguístico em conformidade com essa diferença. Nessas situações, o tipo de relação existente entre os interlocutores é assimétrico (BROWN; GILMAN, 1960). Deste modo, a escolha por uma ou outra forma variante é feita, dentre as outras variáveis, mediante a projeção que o falante faz de sua audiência e, no caso específico do fenômeno em

questão, representa a marcação de uma atitude do falante (BROWN; GILMAN, 1960). Focalizamos na discussão da (as)simetria estabelecida a partir do sexo/gênero dos interactantes.

Em uma interação, o sexo/gênero dos interlocutores é uma variável que pode influenciar os usos linguísticos do falante. O controle dessa variável independente em pesquisas sobre fenômenos que são sensíveis ao perfil social da audiência, como é o caso da variação entre as formas **tu**, **você** e **cê**, torna-se primordial. Identificamos quatro estudos sobre variação de formas de 2ªPS que controlaram essa variável, a saber: Lucca (2005), Dias (2007), Silva, S. (2019) e Nascimento (2011). Este último estudo focaliza a variação entre **você** e **cê** e os dois primeiros são voltados para a variação entre **tu** e **você**. Na tabela a seguir, expomos os resultados obtidos pelos autores.

Tabela 10 - Uso das formas de 2ªPS em função do sexo/gênero dos interlocutores

Estudos	Sexo/gênero dos interloc.	Tu		Você		Cê		T	Qui-quadrado
		A	%	%	A	%			
Lucca (2005) Brasília/DF	Mesmo gênero	290	79	77	21	—	—	367	X ² (1, N= 450) = 38.70 p < 0.00001
	Gênero oposto	37	43	46	57	—	—	83	
Dias (2007) Brasília/DF	Mesmo sexo	39	7,7	470	82,3	—	—	509	X ² (2, N= 900) = 27. 64 p < 0.00001
	Sexo oposto	71	19,6	291	80,4	—	—	362	
	Todos	5	17,2	24	82,8	—	—	29	
Silva, S. (2019) Coité do Nóia/AL	H-H	24	16	126	84	—	—	150	X ² (2, N= 520) = 14.50 p = .00071
	M-M	27	13	175	87	—	—	202	
	H-M	6	4	162	96	—	—	168	
Nascimento (2011) São Paulo/SP	F-F	—	—	47	49	492	51	965	X ² (3, N= 2475) = 6.51 p = .089294
	M-M	—	—	397	55,2	322	44,8	719	
	F-M	—	—	206	52,6	186	47,4	392	
	M-F	—	—	204	51,1	195	48,9	399	
Silva, S. (2019) Coité do Nóia/AL	H-H	—	—	93	79,5	24	20,5	117	X ² (2, N= 463) = 0.58, p = 0.7477
	M-M	—	—	151	82	33	18	184	
	H-M	—	—	128	79	34	21	162	

Fonte: Elaboração própria.

Lucca (2005), a partir de um *corpus* constituído por gravações de conversas espontâneas ocultas de informantes brasileiros, controlou a variável independente sexo/gênero dos interlocutores considerando dois aspectos: se os interactantes eram do mesmo sexo/gênero (homem-homem, mulher-mulher) ou se eram de sexo/gênero opostos (mulher-homem). A autora constatou que o emprego da forma **tu** foi favorecido quando os interlocutores possuíam o mesmo sexo/gênero, apresentando um percentual de 79% das ocorrências. Nas interações em

que os interlocutores eram do sexo/gênero oposto, houve uma tendência sutil para o uso da variante **você**, com um percentual de 57% das ocorrências. A análise inferencial univariada evidenciou que esses resultados são estatisticamente significativos ($X^2 = 38,70$, $p < 0.00001$).

Dias (2007) também analisou dados de conversas espontâneas de informantes da cidade de Brasília/DF. No entanto, diferentemente da pesquisa de Lucca (2005), os informantes selecionados por Dias (2007) tinham consciência de que estavam sendo gravados. A codificação dessa variável independente foi feita considerando a interação entre informantes do mesmo sexo/gênero, do sexo/gênero oposto ou de todos (“quando o informante se dirige a um grupo de pessoas dos dois sexos” (DIAS, 2007, p. 85)). Na amostra analisada, a forma **você** foi mais frequente e, por isso, foi favorecida em todos os tipos de controle da variável sexo/gênero dos interlocutores, principalmente, entre informantes do mesmo sexo/gênero. Ademais, os resultados obtidos demonstraram que a forma **tu** é mais recorrente entre interlocutores de sexo/gênero oposto. A realização da análise univariada evidenciou que essa variável é estatisticamente significativa ($X^2 = 27.64$, $p < 0.00001$).

Silva, S. (2019), diferentemente de Lucca (2005) e Dias (2007), analisou separadamente os dados decorrentes de diálogos entre mulheres daqueles decorrentes de diálogos entre homens. Ao analisar diálogos de informantes da cidade Coité do Nória/AL, a autora verificou que, embora os percentuais mostrem que em qualquer tipo de relação entre sexos/gêneros a forma **você** seja mais frequente, o uso da variante **tu** foi mais elevado entre informante do mesmo sexo/gênero, principalmente, nas relações mulher-mulher (27 ocorrências). Nesse estudo, a variável independente sexo/gênero dos interlocutores também se mostrou significativa ($X^2 = 14.50$, $p = 0.00071$). A autora também pesquisou a variação entre as formas **você** e **cê** nessa mesma amostra. Os resultados demonstraram que a maioria das ocorrências é da forma **você**, a qual teve alta porcentagem nos três tipos de controle da variável sexo/gênero dos interlocutores. A forma variante **cê** teve um quantitativo maior de ocorrências nas interações entre homem-mulher (34 ocorrências). Os resultados do teste de qui-quadrado evidenciaram que o efeito dessa variável não é significativo para a variação entre as formas **você** e **cê** na amostra analisada.

O estudo de Nascimento (2011) focalizou a variação entre as formas **você** e **cê** na fala de informantes de São Paulo/SP a partir de gravações de entrevistas. Para analisar o efeito da variável sexo/gênero dos interlocutores, a autora realizou o controle de quatro formas: feminino-feminino, masculino-masculino, feminino-masculino, masculino-feminino. O controle foi feito desse modo, com intuito de verificar se a simetria ou a assimetria da relação entre outros interlocutores condicionava o uso das variantes. Os resultados evidenciaram leve

favorecimento da forma reduzida **cê** nas interações entre mulheres. Nos demais tipos de controle da variável, houve um leve favorecimento da forma **você**. Assim como no estudo de Silva (2011) sobre as variantes **você** e **cê**, o efeito dessa variável também não foi estatisticamente significativo na amostra analisada por Nascimento (2011).

Em termos gerais, constata-se que há um mecanismo de acomodação linguística envolvido na medida em que o falante muda seu estilo de acordo com o perfil do interlocutor.

4.3 VARIÁVEIS ESTILÍSTICAS

Nesta seção, são apresentados os resultados referentes ao controle de variáveis estilísticas. Constatamos quatro variáveis de cunho estilístico controladas nas pesquisas sobre a variação entre os pronomes **tu**, **você** e **cê**: tipo de sequência discursiva, tópico discursivo, tipo de discurso e tipo de coleta.

4.3.1 Tipo de sequência discursiva

O tipo de sequência discursiva é uma variável independente que pode atuar como um fator condicionante nos usos linguísticos dos falantes e conseqüentemente determinar o estilo a ser utilizado, tendo também relação com os processos acomodativos da dinâmica da língua. Na tabela a seguir, apresentamos os resultados de estudos que controlaram essa variável.

Tabela 11 - Uso das formas de 2ªPS em função do tipo de texto/sequência discursiva

Estudos	Tipo de Texto/ G. discursivo	Tu		Você		Cê		T	Qui-quadrado
		A	%	A	%	A	%		
Loregian-Penkal (2004) Florianópolis/SC, Porto Alegre/RS, Ribeirão/SC	Argumentativo	823	90	87	10	—	—	910	X ² (3, N= 1973) = 23.07, p = 0.00000039
	Narrativo	712	87	103	13	—	—	815	
	Explicações	79	85	14	15	—	—	93	
	Receitas	155	100	0	0	—	—	155	
Loregian-Penkal (2004) Chapecó/SC Blumenau/SC Lages/SC	Argumentativo	187	23	633	77	—	—	820	X ² (3, N= 2144) = 15.70, p < 0.00001
	Narrativo	216	21	831	79	—	—	1047	
	Explicações	63	40	91	60	—	—	154	
	Receitas	85	69	38	31	—	—	123	
Loregian-Penkal (2004) Flores da Cunha/RS Panambi/RS São Borja/RS	Argumentativo	420	87	59	13	—	—	479	X ² (3, N= 1864) = 76.09, p < 0,00001
	Narrativo	994	94	56	6	—	—	1050	
	Explicações	121	80	29	20	—	—	150	
	Receitas	143	77	42	23	—	—	185	
Traesel (2016) Florianópolis/SC	Diálogo	77	65	41	35	—	—	118	X ² (4, N= 544) = 41.53, p < 0,00001
	Narrativa	79	49	83	51	—	—	162	
	Receita	142	78	39	22	—	—	181	
	Conselho	38	80	53	20	—	—	80	
	Outros	1	33	2	67	—	—	3	
Lopes (2017) Chapecó/SC	Dissertativa	5	48	64	52	—	—	123	X ² (2, N= 268) = 0.54, p = 0.760106
	Narrativa	37	43,5	48	56,5	—	—	85	
	Descritiva	26	43,3	34	56,7	—	—	60	
Loregian-Penkal e Menon (2012) Curitiba/PR	Narrativos	—	—	1053	67	519	33	1572	X ² (3, N= 2200) = 28.89, p < 0,00001
	Argumentativos	—	—	407	79	109	21	516	
	Explicações	—	—	70	67	35	33	105	
	Receitas	—	—	3	70	4	30	7	

Fonte: Elaboração própria.

Loregian-Penkal (2004) realizou um estudo sobre a alternância **tu** e **você** em amostras de fala advindas de localidades diversas. Os dados codificados foram rodados no programa estatístico VARBRUL em função das localidades selecionadas, as quais foram divididas em três grupos: i) cidades de Florianópolis/SC, Porto Alegre/RS e Ribeirão da Ilha (distrito de Florianópolis); ii) três cidades do interior de Santa Catarina - Chapecó, Blumenau e Lages; e iii) três cidades do interior do Rio Grande do Sul - Flores da Cunha, Panambi e São Borja. A variável independente, denominada pela autora de gênero discurso, foi controlada a partir de

quatro níveis: segmentos predominantemente narrativos, segmentos predominantemente argumentativos, explicações e receitas; para os casos que não se encaixavam nestes tipos estabelecidos, a autora codificou-os como “não se aplica” (mas não apresenta os resultados desse nível de controle).

Os resultados do primeiro e do terceiro grupos de cidades evidenciaram que o aparecimento da variante **tu** foi favorecido em todos os tipos de gêneros controlados. No que concerne ao segundo grupo, apenas o gênero discursivo receitas obteve um maior percentual da forma **tu** do que da forma **você**. Embora os resultados dessa variável tenham se mostrado significativos nos três grupos de localidades investigados (conforme apresentado na coluna referente ao qui-quadrado da tabela 11), estes são contestáveis no sentido de que a pesquisadora juntou dados de fala de informantes provenientes de estados distintos (Florianópolis/SC e Porto Alegre/RS), o que envia os dados. Nas explicações referentes a esses resultados, Loregian-Penkal (2004) faz algumas ressalvas a respeito das rodadas estatísticas realizadas com cada localidades em separado, demonstrando como os usos das variantes foram condicionados conforme o tipo de gênero. Entretanto, autora não apresentou os dados estatísticos de forma completa de cada cidade separadamente. Além disso, nessas explicações, fica evidente que essa variável independente tem um condicionamento distinto, o que demonstra ainda mais a impossibilidade de analisar o efeito desta com a junção dos dados por grupo de cidades.

Traesel (2016), analisando a variação entre **tu** e **você** na fala de informantes da cidade de Florianópolis/SC, controlou a variável tipo de texto com base nos seguintes níveis: diálogo com o entrevistador, narrativas com envolvimento, receita, conselho e outros. Os resultados demonstraram que quando o tipo de texto é diálogo com o entrevistador e receita a forma pronominal **tu** foi mais frequente. Nos contextos de narrativa com envolvimento e conselho houve uma leve tendência de uso da forma **você**. Ao realizarmos o teste de qui-quadrado, constatamos que essa variável possui um efeito estatisticamente significativo na amostra analisada ($X^2 = 41.53$, $p < 0.00001$).

Lopes (2017), ao analisar uma amostra de fala de informantes da cidade de Chapecó/SC, controlou a variável tipo de sequência discursiva a partir de três níveis: sequência dissertativa, sequência narrativa e sequência descritiva. Os resultados percentuais evidenciaram que há um sutil favorecimento do uso da forma **você** em todos os tipos de sequências discursivas controlados. Ademais, a atuação dessa variável sobre o uso dos pronomes **tu** e **você** na amostra analisada não foi considerada estatisticamente significativa no teste de qui-quadrado que realizamos ($X^2 = 0.54$, $p = 0.760106$).

Loregian-Penkall e Menon (2012) realizaram um estudo sobre a variação entre as formas **você** e **cê** na fala de informantes da cidade de Curitiba/PR. As autoras controlaram a variável independente gênero discursivo da mesma forma que Loregian-Penkall (2004). Seus resultados demonstraram que a forma **você** é favorecida nos quatro tipos de controle da variável (cf. tabela 14). O teste de qui-quadrado evidenciou que essa variável é estatisticamente significativa para o estudo realizado ($X^2 = 28.89$, $p < 0,00001$).

A partir dos resultados apresentados, constata-se que os estudos controlam a variável de forma distinta o que inviabiliza a possibilidade de se fazer generalizações a respeito do efeito da variável no processo de variação das formas das formas em questão. Ademais, autores controlam a variável de forma inadequada ao misturar gêneros textuais/discursivos com tipos textuais/sequências discursivas.

4.3.2 Tópico discursivo

Bell (1984), ao desenvolver o modelo de *Audience Design* (conforme explanamos na seção 2), ressalta que o tópico discursivo pode ser uma das variáveis situacionais que influencia o estilo utilizado pelo falante, ou seja, que atua na mudança de estilo. O autor toma como base para a elaboração dessa hipótese os estudos, dentre outros, desenvolvidos por Blom e Gumperz (1972) - os quais constataram que os falantes de uma determinada comunidade mudavam do dialeto local para o padrão em decorrência de uma mudança de tópico - e Giles e Powesland (1975) - estes descobriram, na pesquisa a respeito do modelo de acomodação, que, ao propiciar uma mudança de tópico, pode-se levar a uma mudança de sotaque. Identificamos três estudos brasileiros sobre as variantes **tu**, **você** e **cê** que controlaram a variável independente tópico discursivo. Na tabela a seguir dispomos os resultados obtidos por esses estudos.

Tabela 12 - Correlação entre o uso das formas de 2ªPS e a variável tópico discursivo

Estudos	Tópico discursivo	Tu		Você		T	Qui-quadrado
		A	%	A	%		
Lucca (2005) Brasília/DF	Mais familiar	323	74	113	26	436	$X^2 (1, N= 453) = 18.38,$ $p < 0.00001$
	Menos familiar	4	23	13	77	17	
Dias (2007) Brasília/DF	Ironia/brincadeira	40	31,7	86	67,3	126	$X^2 (3, N= 900) = 66.37,$ $p < 0.00001$
	Conversa casual	72	12,9	486	82,1	558	
	Repreensão	2	4,9	39	95,1	41	
	Conversa profissional	1	0,6	174	99,4	175	
Guimarães (2014) Fortaleza/CE	Obs. irônicas/brincadeiras	48	90,6	5	9,4	53	$X^2 (7, N= 1555) = 69.04,$ $p < 0.00001$
	Fofoca	25	62,5	15	37,5	40	
	Conv. sobre relac. amoroso	59	63,5	34	36,5	93	
	Recordações	40	41,5	59	58,5	99	
	Conversa casual	541	51,2	516	48,8	1057	
	Conversa sobre trabalho	68	42,2	93	57,8	161	
	Religião	10	20,8	38	79,2	48	
Repreensão	1	25	3	75	4		

Fonte: Elaboração própria.

Lucca (2005) controlou o tópico discursivo considerando o grau de familiaridade que o falante possui com este: mais familiar ou menos familiar. Os tópicos mais familiares envolvem conversas banais a respeito do dia a dia do falante, já aqueles menos familiares possuem maior grau de complexidade que os tópicos voltados para a experiência de vida do falante. A pesquisadora analisou um *corpus* constituído por conversas espontâneas ocultas. A hipótese da autora era que os jovens, ao abordarem temas mais familiares, tenderiam a utilizar um estilo mais próximo do vernáculo e fariam uso do pronome **tu** por esta ser a forma pronominal de 2ªPS mais frequente entre os membros do grupo. Os resultados evidenciaram que o uso do **tu** foi mais recorrente quando o tópico discursivo era de cunho mais familiar (323/436 - 74%) do que em tópico menos familiar (113/436 - 26%), confirmando assim a hipótese aventada. O teste de qui-quadrado apontou esta variável como significativa ($X^2 = 18.38, p < .00001$). Tais resultados evidenciaram que houve uma mudança de estilo de fala.

A variável tópico discursivo foi denominada, no estudo de Dias (2007), de tipo de fala. O seu controle foi feito a partir da seguinte divisão de temas relacionados à: ironia/brincadeiras, conversa casual, repreensão, conversa profissional. A hipótese de Dias (2007) era de que houvesse um favorecimento da forma **tu** em contextos de observações irônicas, brincadeiras e conversa casual, enquanto os contextos de conversa profissional e repreensão favoreceriam o

uso da forma **você**, o que demarcaria o distanciamento entre o falante e seu interlocutor. Os resultados demonstraram que, em termos percentuais, todos os tipos de tópicos controlados favoreceram o uso da forma **você**. No entanto, ao analisarmos a distribuição das ocorrências da forma **tu**, constatou-se que esta foi mais frequente no *corpus* analisado quando os tópicos discursivos eram sobre ironias/brincadeiras e conversas casuais. O efeito dessa variável mostrou-se estatisticamente significativo ($X^2 = 66.37, p < 0.00001$).

Guimarães (2014) controlou a variável tópico discursivo de modo semelhante à Dias (2007), diferenciando-se apenas pelo acréscimo de quatro níveis, a saber: conversa sobre relacionamento amoroso, recordações, fofoca e religião. A pesquisadora tinha a hipótese de que os tópicos de cunho mais íntimos favoreceriam o uso da forma **tu**, no entanto não informou quais se encaixariam nesse escopo. Os resultados evidenciaram que os tópicos discursivos observados irônicas/brincadeiras, fofoca, conversa sobre relacionamento amoroso e conversa casual propiciaram o uso do pronome **tu** na amostra analisada. Já os tópicos recordações, conversa sobre trabalho, religião e repreensão favoreceram a forma **você**. Tais resultados sugerem que o efeito dessa variável foi significativo ($X^2 = 69.04, p < 0.00001$).

Os resultados expostos evidenciam que o falante adapta seus usos linguísticos ao tipo de assunto, conforme o grau de formalidade envolvido, havendo, assim, uma mudança de postura linguística. Novamente, a forma diversificada com que cada estudo controla a variável tópico discursivo inviabiliza a realização generalizações de seu efeito nos usos das formas pronominais de 2ªPS.

4.3.3 Tipo de discurso

O discurso pode ocorrer de forma direta ou indireta. O discurso direto, também denominado de relato próprio, não relatado ou original, ocorre quando o falante, no ato da interlocução, discorre sobre algum assunto/tema, dirigindo-se diretamente ao seu interlocutor, ou seja, sem recorrer a um discurso feito em um momento anterior. O discurso indireto, também chamado de relato de terceiros, reportado, é expresso quando o falante retoma um discurso falado em um momento anterior ao momento da enunciação, podendo ser um discurso reportado do próprio falante ou de terceiros.

Os seguintes pesquisadores controlaram do tipo de discurso com o intuito de verificar se este interfere na escolha das formas de 2ªPS: i) região Norte - Martins (2010), Silva, M. (2019), Brito Castro e Costa (2020); ii) região Nordeste – Alves (2010), Nogueira (2013), Guimarães (2014), Silva, F. (2015), Alves (2015), Silva, Suziane (2017), Silva, S. (2019); iii)

região Centro-Oeste – Andrade (2010); iv) região Sudeste – Nascimento (2011), Silva, Suelen (2017); v) região Sul – Hausen (2000), Traesel (2016), Lopes (2017). Os resultados obtidos por estes estudos são discutidos a seguir, agrupados por região.

Tabela 13 - Uso das formas de 2ªPS em função do tipo de discurso na região Norte

Estudo	Tipo de discurso	Tu		Você		T	Qui-quadrado
		A	%	A	%		
Martins (2010) Tefé/AM	Direto	431	63,7	246	36,3	677	$X^2 (1, N= 806) = 1.12, p = 0.2896$
	Relatado	89	69	40	31	129	
Brito, Castro e Costa (2020) Cameté/PA	Direto	224	63	127	37	352	$X^2 (1, N= 389) = 10.61, p = 0.001122$
	Indireto	34	92	3	8	37	

Fonte: Elaboração própria.

Na tabela 13, estão dispostos os resultados dos estudos de Martins (2010) e Brito, Castro e Costa (2020), os quais foram realizados, respectivamente, nas cidades de Tefé/AM e Cameté/PA, região Norte do Brasil. Os autores controlaram essa variável de forma binária (direto x indireto/relatado) e constataram que os dois tipos de discurso favorecem os usos da forma **tu**, principalmente, o discurso indireto/relatado. No entanto, apenas os resultados obtidos por Brito, Castro e Costa (2020) se mostraram estatisticamente significativos na análise univariada ($X^2 = 10.61, p = 0.001122$).

Tabela 14 - Uso das formas de 2ªPS em função do tipo de discurso na região Nordeste

Estudos	Tipo de discurso	Tu		Você		Cê		T	Qui-quadrado
		A	%	A	%	A	%		
Alves (2010) Cidad. do Maranhão	Próprio	93	33,9	182	66,1	—	—	275	$X^2 (1, N= 329) = 13.09, p = 0.0002958$
	Terceiro	33	61,1	21	61,1	—	—	54	
Nogueira (2013) Feira de Santana/BA	Direto	52	7,5	644	92,5	—	—	696	$X^2 (1, N= 823) = 11.53, p = 0.0006728$
	Relatado	22	17,3	105	82,7	—	—	127	
Nogueira (2013) Salvador/BA	Direto	3	0,6	654	99,5	—	—	657	$X^2 (1, N= 891) = 5.70, p = 0.01695$
	Relatado	6	2,6	228	97,4	—	—	234	
Guimarães (2014) Fortaleza/CE	Original	755	52,8	676	47,2	—	—	1431	$X^2 (1, N= 1555) = 23.08, p < 0.00001$
	Reportada	37	29,8	87	70,2	—	—	124	
Silva, F. (2015) Natal/RN	Não relatado	51	18	229	82	—	—	280	$X^2 (1, N= 378) = 2.10, p = 0.1471$
	Relatado	11	11	87	89	—	—	98	
Alves (2015) São Luís/MA*	Não relatado	818	85	143	15	—	—	961	$X^2 (1, N= 1028) = 1.31, p = 0.2511$
	Relatado	53	85	14	15	—	—	67	
Silva, Suziane (2017) Delmiro Gouveia/AL	Direto	—	—	463	94	31	6	495	$X^2 (2, N= 508) = 5.6807e-31, p = 1$
	Indireto	—	—	12	92	1	8	13	
Silva, S. (2019) * C. do Nória/AL	Relato próprio	—	—	369	81	89	19	458	$X^2 (1, N= 463) = 0.34, p = 0.5583$
	Relato de terceiros	—	—	3	60	2	40	5	
Silva, S. (2019) * C. do Nória/AL	Relato próprio	51	10	458	90	—	—	509	$X^2 (1, N= 514) = 2.8244e-29, p = 1$
	Relato de terceiros	0	0	5	100	—	—	5	

Fonte: Elaboração própria.

Notas: *Os resultados foram recalculados para fins comparabilísticos.

Na região Nordeste, foram constatadas nove pesquisas sobre a variação entre pronomes de 2ªPS que controlaram o tipo de discurso. Os resultados dos estudos feitos nessa região se distinguem dos obtidos na região Norte, visto que se constatou, em termos gerais, que, tanto o discurso direto quanto o indireto, favoreceram o uso da variante **você** nas amostras analisadas. Temos duas exceções nesse âmbito: os resultados obtidos por Guimarães (2014), a partir da análise de dados de fala de Fortaleza/CE, apontaram que, em contextos de discurso direto/original, há uma sutil tendência para o uso da variante **tu** (755/1431 - 52,8%) e os resultados obtidos por Alves (2015) demonstraram que nos dois tipos de discurso a forma **tu** é majoritariamente favorecida na amostra de dados de fala da cidade de São Luís/MA. Apenas os resultados obtidos nos estudos de Alves (2010), Nogueira (2013) e Guimarães (2014) foram estatisticamente significativos na análise inferencial univariada.

Tabela 15 - Uso das formas de 2ªPS em função do tipo de discurso na região Centro-Oeste

Estudo	Tipo de discurso	Tu		Você		Cê		T	Qui-quadrado
		A	%	A	%	A	%		
Andrade (2010) Brasília/DF	Original	254	48,7	134	25,7	134	25,7	522	X ² (1, N= 571) = 2.99, p = 0.2232
	Reportado	19	38,9	12	24,5	18	36,7	49	

Fonte: Elaborado a partir de Andrade (2010).

Na região Centro-Oeste, Andrade (2010) constatou que, dentre as formas **tu**, **você** e **cê**, a variante **tu** foi mais recorrente, tanto no discurso original/direto (254/522 - 48,7%) quanto no reportado (19/49, 38,9%), em dados de fala da cidade de Brasília/DF (cf. tabela 15). Entretanto, o teste de qui-quadrado evidenciou que tais resultados não são estatisticamente significativos.

Tabela 16 - Uso das formas de 2ªPS em função do tipo de discurso na região Sudeste

Estudo	Tipo de discurso	Tu		Você		T	Qui-quadrado
		A	%	A	%		
Silva, Suelen (2017) Ressaquinha/MG	Direto	172	40,1	257	59,1	429	X ² (1, N= 579) = 1.49, p = 0.2215
	Relatado	51	34	99	66	150	

Fonte: Elaborado a partir de Silva, Suelen (2017).

Diferentemente de Andrade (2010), os resultados de Silva, Suelen (2017) evidenciaram que, na amostra de fala da cidade de Ressaquinha/MG (região Sudeste), a forma **você** é a mais frequente nos dois tipos de discurso controlado (cf. tabela 16). O efeito dessa variável também não é significativo.

Tabela 17 - Uso das formas de 2ªPS em função do tipo de discurso na região Sul

Estudos	Tipo de discurso	Tu		Você		T	Qui-quadrado
		A	%	A	%		
Traesel (2016) Florianópolis/ SC	Não reportado	213	73	77	26	290	X ² (2, N= 544) = 38.51, p < 0.00001
	Repor. p. próx.	122	50	122	50	244	
	Repor. p. não próx.	2	20	8	80	10	
Lopes (2017) Chapecó/SC	D. p/ o entrevistador	114	47,1	128	52,9	242	X ² (2, N= 268) = 2.85, p = 0.2405
	D. relatado de 3ªP	7	29,2	17	70,8	24	
	D. rel. do próp. fal.	1	50	1	50	2	

Fonte: Elaboração própria.

Traesel (2016), conforme tabela 17, de modo distinto dos demais pesquisadores mencionados, controlou o tipo de discurso reportado, correlacionando-o ao grau de proximidade entre os interlocutores. Os resultados obtidos pelo autor evidenciaram que o discurso não reportado favorece o uso da variante **tu** (213/290 - 73%), enquanto o discurso reportado com uma pessoa não próxima favorece o uso da variante **você** (8/10 - 80%). Os dados referentes ao discurso reportado com pessoa próxima se distribuíram de modo igualitário entre as variantes **tu** e **você**. Tais resultados obtiveram associação estatística significativa (X² = 38.51, p < 0.00001). Os resultados obtidos por Lopes (2017) apontaram o favorecimento da forma **você** nos contextos de discurso para o entrevistador (sutilmente – 128/242 - 52,9) e em discurso relatado de terceiros (17/24 - 70,8%). Esses resultados obtidos pela pesquisadora não foram estatisticamente significativos.

Tabela 18 - Função referencial e a variável tópico discursivo no uso das formas de 2ªPS

Estudos	Tipo de discurso	Tu		Você		Cê		T	Qui-quadrado
		N	%	N	%	N	%		
Hausen (2000) Blumenau/SC, Chapecó/SC, Lages/SC	Indeterminado	317	22	1132	78	—	—	1449	X ² (5, N= 2148) = 72.87, p < 0.00001
	Dirige-se ao entrevistador	103	29	247	71	—	—	350	
	Função fática	43	59	30	41	—	—	73	
	Discurso relatado	71	33	146	67	—	—	217	
	Discurso relatado próprio	26	46	31	53	—	—	57	
	Dirige-se a um interveniente	1	50	1	50	—	—	02	
Nascimento (2011) São Paulo/SP (Ref. específica)	Não reportado	—	—	111	39	171	61	282	X ² (1, N= 537) = 2.62, p = 0.105
	Reportado	—	—	119	47	136	53	255	
Nascimento (2011) São Paulo/SP (Ref. genérica)	Não reportado	—	—	972	53	847	47	1819	X ² (1, N= 1838) = 3.94, p = 0.04688
	Reportado	—	—	15	79	4	21	19	
Silva, M. (2019) Rio Branco/AC (Ref. específica)	Direto	21	32,8	43	67,2	—	—	64	X ² (2, N= 128) = 2.71, p = 0.2575
	Rep. por terceiros	23	46	27	54	—	—	50	
	Rep. pelo entrevistado	7	50	7	50	—	—	14	

Fonte: Elaboração própria.

Na tabela 18, estão dispostos resultados dos estudos que separaram os dados de contextos de referência determinada (específica) e/ou indeterminada (genérica), relacionando-os à variável tópico discurso no uso das formas de 2ªPS. Hausen (2000) controlou o tipo de discurso da seguinte forma: indeterminado (“o falante se dirige a um interlocutor genérico” (HAUSEN, 2000, p. 63)), dirige-se ao entrevistador, discurso relatado, discurso relatado próprio, dirige-se a um interveniente e função fática (utilizado “para chamar a atenção do interlocutor e manter o canal de comunicação aberto” (HAUSEN, 2000, p. 37), como em “tu vê, eu comecei ir na aula com treze anos” (HAUSEN, 2000, p. 65)). Os resultados evidenciaram que apenas em contextos de função fática o houve o uso mais frequente da variante **tu** (43/73 - 59%).

Nascimento (2011) constatou que, em contextos de referência específica, a variante **cê** é favorecida, principalmente, no discurso não reportado (171/282 - 61%). Em contextos de referência genérica, os resultados evidenciaram o favorecimento da variante **você**, sobretudo, no discurso reportado (15/19 - 79%). Somente em contextos de referência genérica os resultados se mostraram significativos no teste de qui-quadrado ($X^2 = 3.94$, $p = 0.04688$).

Silva, M. (2019) apresentou os resultados obtidos sobre a variação das formas **tu** e **você**, a partir do controle do tipo de discurso, apenas em contextos de referência específica. Os resultados evidenciaram que, no discurso direto e reportado por terceiros, a variante **você** é mais frequente, já no discurso reportado pelo entrevistado as variantes tiveram a mesma frequência de uso. Tais resultados não foram estatisticamente significativos.

Em termos gerais, os efeitos da variável independente tipo de discurso nos usos das variantes não se distinguem muito na maioria das pesquisas. A diferença de uso de formas quanto ao tipo de discurso somente foi constatada nos estudos de Guimarães (2014), com dados de fala de Fortaleza/CE, e de Traesel (2016), com dados de fala de Santa Catarina/SC – nestes a variante **tu** foi favorecida. Tais resultados, menos expressivos, podem ter relação com o tipo de coleta utilizado, poucos dados do fenômeno em alguns estudos, predominância de uma das variantes na região. Os resultados obtidos por Nascimento (2011) evidenciam a necessidade de analisar os usos das variantes de 2ªPS separando os dados de referência específica daqueles de referência genérica.

4.3.4 Tipo de coleta

O tipo de coleta utilizado para a constituição do *corpus* de uma pesquisa pode interferir nos usos linguísticos do falante. Fenômenos linguísticos sensíveis a variáveis socio-pragmáticos e estilísticos são, geralmente, condicionados pelo tipo de coleta de dados. Scherre *et al.* (2015) destacam, assim como outros pesquisadores, a exemplo de Lopes *et al.* (2018), que esse seria o caso do uso variável dos pronomes de 2ªPS no português brasileiro. Constatamos um estudo na região Norte (Babilônia e Martins (2011)) e outro na região Sudeste (Calmon (2010)) que controlaram o efeito dessa variável nos usos das variantes **tu**, **você** ou **cê**. Na tabela abaixo, dispomos os resultados obtidos por tais pesquisas.

Tabela 19 - Uso das formas de 2ªPS em função do tipo de coleta

Estudos	Situação discursiva	Tu		Você		Cê		T	Qui-quadrado
		A	%	A	%	A	%		
Babilônia e Martins (2011) Manaus/AM	D2	161	70,5	65	29,5	—	—	228	$X^2 (2, N= 494) = 235.77, p < 0.00001$
	DID	9	4	214	96	—	—	223	
	EF	4	9,3	39	90,7	—	—	43	
Calmon (2010) Vitória/ES	PORTVIX- F	—	—	1350	74,2	469	25,8	1819	$X^2 (5, N= 4355) = 110.91, p < 0.00001$
	PORTVIX-E	—	—	1793	76,0	566	24	2359	
	F. Casual 1	—	—	32	31,4	70	68,6	102	
	F. Casual 2	—	—	79	63,2	46	36,8	125	
	Obs. Casual 1	—	—	14	63,6	8	36,4	22	
	Obs. Casual 1	—	—	11	61,1	7	38,9	18	

Fonte: Elaboração própria.

Babilônia e Martins (2011) desenvolveram uma pesquisa sobre o fenômeno em questão com dados de fala da cidade de Manaus/AM, região Norte do país. Os autores analisaram um *corpus* constituído por três tipos de coleta/situação discursiva, os quais somam trinta gravações e, aproximadamente, doze horas: Diálogos entre Dois Informantes (D2), Diálogos entre Informante e Documentador (DID) e Elocuções Formais (EF) (BABILÔNIA; MARTINS, 2011). Os resultados evidenciaram que o uso da forma **tu** foi favorecido em situações discursivas de D2 (161/228 – 70,5%). Quanto ao pronome **você**, este foi predominantemente mais recorrente em situações discursivas de DID (214/223 – 96%) e de EF (39/43 – 90,7%). Tais resultados possuem um efeito estatisticamente significativo ($X^2 = 235.77, p < 0.00001$). É importante ressaltar que tais resultados devem ser considerados com ponderações, pois, além do quantitativo de gravações não ser distribuído uniformemente entre os tipos de situação

discursiva analisados, os perfis sociais dos informantes são diversificados tanto pela idade quanto pelo grau de escolaridade.

Calmon (2010) utilizou dados em sua pesquisa decorrentes de dois bancos de dados: Projeto O Português Falado na Cidade de Vitória (PORTVIX) e amostra Fala Casual. Do PORTVIX foram analisadas 24 entrevistas sociolinguísticas e da amostra Fala Casual duas gravações ocultas de conversas feitas com dois grupos de pessoas, as quais possuíam perfis sociais distintos e eram íntimas. A autora separou os dados de fala que eram dos participantes da pesquisa (PORTVIX F, Fala Casual 1, Fala Casual 2) daqueles que eram dos entrevistadores/pesquisadores (PORTVIX E, observador Casual 1, observador Casual 2). Os resultados obtidos apontam que os participantes da pesquisa usam as variantes **você** e **cê**, sendo que a primeira foi mais frequente (com exceção da Fala Casual 1), principalmente, na amostra de entrevistas sociolinguísticas. A análise inferencial univariada realizada aponta que esses resultados são estatisticamente significativos ($X^2 = 110.91$, $p < 0.00001$).

Observa-se que os resultados desses estudos evidenciaram que há distinções de usos entre as duas localidades: enquanto no *corpus* de Manaus/AM as frequências de usos das variantes se modificam de acordo com o tipo de situação discursiva/tipo de coleta, no *corpus* de Vitória/ES essa distinção não se mostrou proeminente. Tal fato pode estar relacionado ao baixo quantitativo de gravações de Fala Casual utilizado no estudo de Camon (2010), o que evidencia a necessidade de pesquisas que considerem tipo de coleta/situação discursiva no controle de variáveis condicionantes.

4.4 VARIÁVEL ESTRUTURAL

Nesta subseção discutimos os resultados dos estudos sobre pronomes de 2ªPS que controlaram a variável estrutural paralelismo formal. O paralelismo linguístico consiste na tendência à repetição de uma mesma estrutura seja nos níveis fonológico, morfológico, lexical, sintático ou semântico. Trata-se da repetição de uma mesma forma linguística configurando ocorrências em cadeia, isto é, “marcas levam a marcas e zeros levam a zeros” (SCHERRE, 1998, p. 54). O controle do paralelismo tem se mostrado uma variável significativa em investigações variacionistas (FREITAG; ARAUJO, 2011; NASCIMENTO, 2011), assim, é uma variável de suma importância na investigação de qualquer fenômeno linguístico.

O paralelismo formal é uma das variáveis controladas pelos pesquisadores a fim de identificar as motivações que podem condicionar a implementação de uma variante em

detrimento de outra. Constatamos vários estudos sobre a variação entre as formas de 2ªPS que controlaram essa variável. Na sequência, apresentamos os resultados dessas pesquisas, agrupando-os por região.

Tabela 20 - Uso das formas de 2ºPS em função do paralelismo formal na região Norte

Estudos	Paralelismo	Tu		Você		T	Qui-quadrado
		A	%	A	%		
Martins (2010) Tefé/AM	Isolada não prec.	152	76,8	46	23,2	198	$X^2(4, N= 750) = 225.85, p < 0.00001$
	1ª da série não prec.	70	50	70	50	140	
	Precedida de tu	229	89,5	27	10,5	256	
	Precedida de você	30	20	120	80	150	
	Precedida de senhor	3	50	3	50	6	
Costa (2016) Cametá/PA	Não 1ª da série, precedida por tu	192	92	17	8	212	$X^2(4, N= 484) = 192.13, p < 0.00001$
	Não prec. de forma pron., isol. na oração	42	60,9	27	39,1	69	
	1º item da série, não prec. de forma pron.	51	59,3	35	40,7	86	
	Não 1º item da série, prec. por você	18	15,4	99	74,6	117	
	Não 1º item da série, prec. por senhor(a)	1	20	4	80	5	
Brito, Castro e Costa (2020) Cametá/PA	1º item da série e prec. da f. pron. tu	36	81	8	19	44	$X^2(2, N= 300) = 17.71, p = 0.0001422$
	Isolado na oração	132	54	114	46	246	
	Precedido de você	2	20	8	80	10	

Fonte: Elaboração própria.

Martins (2010), Costa (2016), Brito, Castro e Costa (2020) foram os pesquisadores da região Norte que controlaram a variável paralelismo formal. Os resultados obtidos por Martins (2010) evidenciaram que os contextos de ocorrência isolada e precedida de **tu** favorecem o uso da variante **tu** (cf. tabela 20). Os contextos em que a forma pronominal é a primeira da série não há favorecimento de nenhuma das variantes, visto que a distribuição destas ocorreu de forma equitativa. Quando a ocorrência da variante é precedida da forma **você**, há condicionamento do uso da variante **você** (120/150 – 80%).

A pesquisa de Costa (2016), em Cametá/PA, e de Brito, Castro e Costa (2020), em Cametá/PA apresentam resultados semelhantes aos obtidos por Martins (2010), em Tefé/AM. Conforme tabela 20, os contextos de ocorrência isolada na oração, primeiro item da série e precedido de **tu** favorecem o uso da forma **tu**. Nas situações em que a ocorrência é precedida pelo pronome **você**, o uso da variante **você** é favorecido. A análise univariada apontou os resultados dessas três pesquisas como estatisticamente significativos.

Tabela 21 - Uso das formas de 2ºPS em função do paralelismo formal na região Nordeste

Estudos	Paralelismo	Tu		Você		Cê		T	Qui-quadrado
		A	%	A	%	A	%		
Guimarães (2014) Fortaleza/CE	Isolado	479	61,3	297	38,7	—	—	782	X ² (2, N= 1555) = 73.00, p < 0.00001
	Primeira da série	126	44,7	156	55,3	—	—	282	
	Não primeira da série	187	38,1	304	61,9	—	—	491	
Alves (2015) * São Luís/MA	Isolada	495	89	61	11	—	—	554	X ² (4, N= 1021) = 194.22, p < 0.00001
	Primeira da série	141	82	31	18	—	—	172	
	Não 1ª da série, prec. de tu sem concord.	186	92	16	8	—	—	202	
	Não 1ª da série, prec. de tu com concord.	36	92	3	8	—	—	39	
	Não 1ª da série, precedida de você	11	20	43	80	—	—	54	
Guimarães (2019)* Fortaleza/CE	Isolado	79	22,3	275	77,7	—	—	354	X ² (4, N= 601) = 57.48, p < 0.00001
	Primeiro da série	9	10,6	76	89,4	—	—	85	
	Precedido de tu	10	76,9	3	23,1	—	—	13	
	Precedido de cê	2	7,7	24	92,3	—	—	26	
	Precedido de você	5	4,1	118	95,9	—	—	123	
Silva, S. (2019) Coité do Nóia/AL	Realização isolada	47	14	306	87	—	—	353	X ² (3, N= 520) = 36.70, p < 0.00001
	Primeiro da série	5	7	64	93	—	—	69	
	Antecedido por tu	4	80	1	20	—	—	5	
	Antecedido por você/cê	1	1	92	99	—	—	93	
Rocha et al. (2016) V. da Conquista/BA	Antecedido pelo pronome você	—	—	108	77	33	23	141	X ² (1, N= 167) = 54.19, p < 0.00001
	Antecedido pela variante cê	—	—	39	31	87	69	126	
Silva, Suziane (2017) Delmiro Gouveia/AL	Realização isolada	—	—	64	91	6	9	70	X ² (3, N= 508) = 21.81, p < 0.00001
	Primeira da série	—	—	99	94	6	6	105	
	Antecedido por você	—	—	299	95	15	5	314	
	Antecedido por cê	—	—	13	68	6	32	19	
Silva, S. (2019) * Coité do Nóia/AL	Realização isolada	—	—	52	81	12	19	64	X ² (4, N= 463) = 0.66, p < 0.00001
	Primeiro da série	—	—	245	80	61	20	306	
	Antecedido por tu	—	—	1	100	0	0	1	
	Antecedido por você	—	—	72	81	17	19	89	
	Antecedido por cê	—	—	2	66,7	1	33,3	3	

Fonte: Elaboração própria.

Notas: *Os resultados foram recalculados para fins comparabilísticos.

Na região Nordeste, conforme tabela 21, constatamos sete estudos que controlaram a variável paralelismo formal, destes quatro foram sobre a variação **tu** e **você** (Guimarães (2014),

em Fortaleza/CE; Alves (2015), São Luís/MA; Guimarães (2019), em Fortaleza/CE; Silva, S. (2019), em Coité do Nória/AL) e três sobre a variação **você** e **cê** (Rocha, Santos e Sousa (2016), em Vitória da Conquista/BA; Silva, Suziane (2017), em Delmiro Gouveia/AL; Silva, S. (2019), em Coité do Nória/AL).

As pesquisas realizadas a respeito da variação **tu** e **você** obtiveram resultados relativamente distintos. Guimarães (2014) constatou, na amostra analisada, que nos contextos de uso isolado a variante **tu** é mais recorrente (479/782 – 61,3%), já os contextos de primeira da série e de não primeira da série favorecem o uso da variante **você**. Alves (2015) controlou a variável paralelismo a partir de cinco níveis: isolada; não primeira da série, precedida de **tu** sem concordância; não primeira da série, precedida de **tu** com concordância; não primeira da série, precedida de **você**. A variante **tu** só não foi favorecida em contextos de não primeira da série, precedida de **você** (11/54 – 20%). Diferentemente dos resultados obtidos por Alves (2015), Guimarães (2019) e Silva, S. (2019) constataram em seus estudos que a variante **tu** foi favorecida apenas em contextos precedido de **tu**, os demais contextos (isolado, primeira da série, precedido de **você**, precedido de **cê**) condicionaram o uso da forma **você**. Observa-se que os resultados de Guimarães (2019) se diferenciam dos obtidos pela pesquisadora em 2014 com dados também de Fortaleza/CE quanto aos contextos de uso isolado das variantes: no primeiro estudo realizado a forma **tu** foi favorecida, já no segundo a forma **você**. A análise inferencial univariada realizada mostra que esses resultados são estatisticamente significativos para todos os estudos.

Os estudos sobre a variação **você** e **cê** obtiveram praticamente os mesmos resultados. Todos os níveis de paralelismo controlados favoreceram o uso da variante **você**, com exceção apenas do nível antecedido pelo pronome **cê** que no estudo de Rocha, Santos e Sousa (2016) favoreceu o uso da variante **cê** (87/126 – 69%). A análise univariada de qui-quadrado evidencia que há associação estatisticamente significativa entre o uso das variantes **você** e **cê** e a variável paralelismo formal em todos os estudos.

Tabela 22 - Uso das formas de 2ºPS em função do paralelismo formal na região Centro-Oeste

Estudos	Paralelismo	Tu		Você		Cê		T	Qui-quadrado
		A	%	A	%	A	%		
Andrade (2010) Brasília/DF	Isolada na oração	176	36,6	164	34,1	141	29,3	481	$X^2(4, N= 835)$ = 201.32, $p < 0.00001$
	1º item da série	44	31,9	57	41,3	37	26,8	138	
	Precedido por tu	60	84,5	5	7	6	8,5	71	
	Precedido por você	5	5,2	79	81,4	13	13,4	97	
	Precedido por cê	3	6,2	13	27,1	32	66,7	48	
Andrade (2015) Brasília/DF	Isolado	254	34,2	317	42,7	172	23,1	743	$X^2(4, N= 1319)$ = 309.78, $p < 0.00001$
	1º da série	73	29,8	113	46,1	59	24,1	245	
	Precedido de tu	74	84,1	7	8	7	8	88	
	Precedido de você	6	3,6	135	81,8	24	14,5	165	
	Precedido de cê	4	5,1	19	24,4	55	70,5	78	

Fonte: Elaboração própria.

Foram realizados três estudos sobre a variação entre as formas de 2ªPS na região Centro-Oeste, mais especificamente em Brasília/DF, que controlaram a variável paralelismo formal: Andrade (2004), que analisou a variação entre **você** e **cê**, Andrade (2010) e Andrade (2015), que analisaram a variação entre **tu**, **você** e **cê** (cf. tabela 22). Os resultados obtidos por Andrade (2010) e Andrade (2015) apontaram que a precedência de uma variante condiciona, na sequência, o uso da mesma variante, o que evidencia o efeito do paralelismo. Os resultados desses dois estudos diferem apenas para os contextos de uso isolado das variantes: no primeiro estudo houve, de modo sutil, o favorecimento da variante **tu** (176/481 – 36,6%), já no segundo houve o favorecimento da variante **você** (113/245 – 46,1%). Tais resultados se mostraram estatisticamente significativos na análise univariada.

Tabela 23 - Uso das formas de 2ºPS em função do paralelismo formal

Andrade (2004) Brasília	Paralelismo	Você		Cê		T	Qui-quadrado	
		N	%	N	%			
Entrevistadora	Forma prec. de cê	32	27	84	72	116	$X^2(2, N= 707) = 156.29, p < 0.00001$	
	Forma prec. de você	333	85	57	15			390
	1º de uma série	156	68	45	32			201
Entrevistado	Forma prec. de cê	7	44	9	56	16	$X^2(2, N= 115) = 7.26, p = 0.02648$	
	Forma prec. de você	29	78	8	22			37
	1º de uma série	35	66	27	44			62

Fonte: Elaboração própria.

Andrade (2004) analisou os dados do fenômeno em questão que ocorreram na fala da entrevistadora separado daqueles que foram utilizados pelo entrevistado. Os resultados obtidos demonstraram que, independentemente de quem é o falante, a precedência de uma variante favoreceu, na sequência, o uso da mesma variante, ou seja, manteve-se o paralelismo (cf. tabela 23). A autora constatou que nos contextos de uso isolado a variante *voê* foi favorecida. Esses resultados se mostram estatisticamente significativo na análise univariada de qui-quadrado.

Tabela 24 - Uso das formas de 2ºPS em função do paralelismo formal na região Sudeste

Estudos	Paralelismo	Tu		Você		Cê		T	Qui-quadrado
		A	%	A	%	A	%		
Nascimento (2011) São Paulo/SP	Sem antecedente	—	—	716	53,4	625	46,6	1341	$X^2(2, N= 2475) = 274.35, p < 0.00001$
	Antecedido por <i>voê</i>	—	—	458	71	187	29	645	
	Antecedido por <i>cê</i>	—	—	106	21,7	383	78,3	489	
Santos (2012) Rio de Janeiro/RJ	Isolado	7	22,6	24	77,4	—	—	31	$X^2(4, N= 297) = 36.26, p < 0.00001$
	Forma empreg. após a 2ª interv. da pesq.	37	19,2	156	80,6	—	—	93	
	Precedido de nulo/zero	12	22,2	42	78,8	—	—	54	
	Precedido de <i>voê</i>	9	9,2	89	90,8	—	—	98	
	Precedido de <i>tu</i>	14	66,7	7	43,3	—	—	21	

Fonte: Elaboração própria.

Dois pesquisas desenvolvidas na região Sudeste controlaram a variável paralelismo formal: Nascimento (2011), em São Paulo/SP, e Santos (2012), no Rio de Janeiro/RJ. Nascimento (2011) constatou que os contextos antecedidos por *cê* favoreceram o uso da variante *cê* (383/489 – 78,3%). Os contextos sem antecedentes ou antecedidos por *voê* condicionaram o uso da variante *voê*. Quanto ao estudo de Santos (2012) sobre a variação entre *tu* e *voê*, os resultados evidenciaram que os contextos precedidos de *tu* favoreceram o uso da variante *tu* (14/21 - 66,7%). Os demais contextos controlados condicionaram o uso da variante *voê*. Os resultados destas pesquisas são estatisticamente significativos (cf. tabela 24).

Tabela 25 - Uso das formas de 2ºPS em função do paralelismo formal na região Sul

Paralelismo	Tu		Você		T	Qui-quadrado
	N	%	N	%		
Isolada	245	67	122	33	367	$X^2(2, N= 926) = 32.78, p < 0.00001$
Formas diferentes	40	51	38	49	78	
Formas iguais	227	47	254	53	481	

Fonte: Elaboração própria a partir de Franceschini (2011).

Na região Sul, apenas Franceschini (2011) apresentou resultados referentes ao controle da variável paralelismo formal ao analisar dados de fala da cidade de Concórdia/SC. A pesquisadora constatou que os contextos de uso isolado favoreceram a utilização da variante *tu*

(245/367 – 67%). Nos contextos de formas diferentes (não paralelismo) e de formas iguais (paralelismo) as taxas de frequências são praticamente as mesmas para as variantes **tu** e **você** (cf. Tabela 25). A análise univariada de qui-quadrado mostra que estes resultados são estatisticamente significativos ($X^2 = 32.78$, $p < 0.00001$).

Tabela 26 - Correlação entre o uso das formas de 2ºPS e o paralelismo formal em contextos de referência específica

Paralelismo formal	Tu		Você		T	Qui-quadrado
	N	%	N	%		
Ocor. isoladas	42	44,2	53	55,8	95	$X^2 (3, N= 125) = 4.23, p = 0.2367$
Paral. binário	5	25	15	75	20	
Paral. ternário	1	20	4	80	5	
Paral. enário	1	20	4	80	5	

Fonte: Elaboração própria a partir de Silva, M. (2019).

Silva, M. (2019) analisou dados de fala da cidade Rio Branco/AC, região Norte. Diferentemente dos estudos anteriores, a pesquisadora analisou apenas os dados de 2ºPS com referência específica. Silva, M. (2019), conforme tabela 26, controlou a variável paralelismo formal da seguinte forma: ocorrências isoladas, paralelismo binário (quando o dado em análise é precedido por apenas um pronome, é a segunda ocorrência da sequência), paralelismo ternário (quando o dado em análise é precedido por dois pronomes, é a terceira ocorrência da sequência), paralelismo eneário (quando o dado em análise é precedido por três pronomes, é a quarta ocorrência da sequência). Os resultados mostraram que em todos os níveis controlados a variante **você** é favorecida. Tais resultados não são estatisticamente significativos. Cabe ressaltar que a quantidade ocorrências do fenômeno na amostra analisada foi muito baixa, o que evidencia a necessidade de um estudo com maior poder exploratório.

Em termos gerais, os resultados obtidos com o controle da variável paralelismo formal nesses estudos evidenciaram que a probabilidade de usar **tu**, ao invés de **você**, é significativamente maior quando o falante utiliza também a variante **tu** em oração antecedente. As variantes **cê** e **você** possuem comportamento análogo: há maior probabilidade para o uso de **cê** quando o antecedente formal é **cê**, e de **você** quando o antecedente for **você**. Tais constatações corroboram com a premissa defendida por Scherre (1998) de que marcas levam a marcas.

4.5 VARIÁVEL SOCIAL

Entre as variáveis clássicas controladas nos estudos sociolinguísticos está a variável sexo/gênero, a qual tem se mostrado estatisticamente significativa para o estudo de muitos

fenômenos linguísticos variáveis. Em virtude do quantitativo de estudos que controlaram essa variável, apresentamos nesta seção apenas os resultados das pesquisas desenvolvidas na região Nordeste. Constatamos 12 pesquisas realizadas com dados de fala dessa região: Oliveira (2007), Santana (2008), Divino (2008), Alves (2010), Assunção e Almeida (2011), Carneiro (2011), Nogueira (2013) (analisou duas amostras: uma de Salvador/BA e outra de Feira de Santana/BA), Guimarães (2014), Alves (2015), Rocha, Santos e Sousa (2016), Silva, Suziane (2017) e Silva, S. (2019). Retomamos também os resultados de Andrade (2010; 2015), Calmon (2010) e Nascimento (2011) por estes focalizarem a variação **você** e **cê**. Os resultados dessas pesquisas são apresentados na tabela a seguir.

Tabela 27 - Uso das formas de 2ºPS em função do sexo/gênero do informante na região Nordeste

Estudos	Sexo/ gênero	Tu		Você		Cê		T	Qui-quadrado
		A	%	A	%	A	%		
Oliveira (2007) Poções/BA e S. A de Jesus/BA	Feminino	61	14	380	86	—	—	441	$X^2(1, N= 1128) = 1.48, p = 0.2227$
	Masculino	77	11	610	89	—	—	687	
Santana (2008) F. de Santana/BA	Feminino	75	54	63	46	—	—	138	$X^2(1, N= 243) = 50.72, p < 0.00001$
	Masculino	10	9	95	91	—	—	105	
Divino (2008) S. A. de Jesus/BA	Feminino	104	22,9	350	77,1	—	—	454	$X^2(1, N= 875) = 1.33, p = 0.2475$
	Masculino	82	19,5	339	80,5	—	—	421	
Alves (2010) São Luís/MA	Feminino	49	40,8	71*	59,2	—	—	120	$X^2(1, N= 328) = 0.32, p = 0.5712$
	Masculino	77	37	131	63	—	—	208	
Assunção e Almeida (2011) F. de Santana/BA	Feminino	5	7	63	93	—	—	68	$X^2(1, N= 155) = 0.47, p = 0.4925$
	Masculino	6	13	40	87	—	—	87	
Carneiro (2011) São Luís/MA	Feminino	116	70,7	48	29,3	—	—	164	$X^2(1, N= 277) = 0.23, p = 0.6285$
	Masculino	76	67,2	37	32,8	—	—	113	
Nogueira (2013) F. de Santana/BA	Feminino	28	6,3	415	93,7	—	—	443	$X^2(1, N= 822) = 7.74, p = 0.005399$
	Masculino	46	12,1	333	87,9	—	—	379	
Nogueira (2013) Salvador/BA	Feminino	6	1,2	489	98,8	—	—	495	$X^2(1, N= 891) = 0.11, p = 0.736$
	Masculino	3	0,8	393	99,2	—	—	396	
Guimarães (2014) Fortaleza/CE	Feminino	470	58,6	332	41,4	—	—	802	$X^2(1, N= 1555) = 38.36, p < 0.00001$
	Masculino	322	42,8	431	57,2	—	—	753	
Alves (2015) São Luís/MA	Feminino	376	77	111	23	—	—	487	$X^2(1, N= 1028) = 39.34, p < 0.00001$
	Masculino	495	91	46	9	—	—	541	
Rocha, Santos e Sousa (2016) V. Conquista/BA	Feminino	—	—	165	68	79	32	244	$X^2(1, N= 405) = 23.38, p < 0.00001$
	Masculino	—	—	69	43	92	57	161	
Silva, Suziane (2017) Delmiro Gouveia/AL	Feminino	—	—	200	94	12	6	206	$X^2(1, N= 502) = 0.21, p = 0.6425$
	Masculino	—	—	275	93	21	7	296	
Silva, S. (2019) Coité do Nóia/AL	Feminino	31	10	277	90	—	—	308	$X^2(1, N= 520) = 0.41, p = 0.5183$
	Masculino	26	12	186	88	—	—	212	

Fonte: Elaboração própria.

Os resultados demonstraram que, na maioria das pesquisas desenvolvidas no Nordeste, os informantes de ambos os sexos/gêneros utilizam com mais frequência a variante **você** (cf. tabela 27). Apenas nos estudos de Carneiro (2011) e Alves (2015) o pronome **tu** foi a variante mais usada pelos informantes tanto do sexo/gênero masculino quanto pelo feminino. Santana (2008) e Guimarães (2014) obtiveram resultados distintos dos demais estudos: informantes do sexo/gênero feminino utilizaram mais a forma **tu**, enquanto os informantes do sexo/gênero masculinos utilizaram com mais frequência a forma **você**. Rocha, Santos e Sousa (2016) analisaram a variação entre as formas **você** e **cê** e alcançaram também resultados distintos em função da variável independente sexo/gênero: informantes do sexo/gênero feminino utilizaram com maior frequência a variante **você** e os do sexo/gênero masculino utilizaram a variante **cê**. O teste de qui-quadrado evidenciou que os resultados obtidos nas pesquisas de Santana (2008), Nogueira (2013) (amostra Feira de Santana/BA), Guimarães (2014) e Rocha, Santos e Sousa (2016) são estatisticamente significativos.

Tabela 28 - Relação entre o uso das formas de 2ºPS e o sexo/gênero do informante em outras regiões

Estudos	Sexo/gênero	Tu		Você		Cê		T	Qui-quadrado
		A	%	A	%	A	%		
Andrade (2010) Brasília	Feminino	145	44	113	34	72	22	330	X ² (2, N= 574) = 29.038, p < 0.00001
	Masculino	128	52	36	15	80	33		
Calmon (2010) – amostra PORTVIX Brasília	Feminino	—	—	511	79,5	132	20,5	643	X ² (1, N= 1819) = 13.929, p = 0.0001898
	Masculino	—	—	839	71	337	29		
Calmon (2010) – amostra Fala Casual Brasília	Feminino	—	—	97	63,4	56	36,6	153	X ² (1, N= 269) = 19.677, p < 0.00001
	Masculino	—	—	41	35	75	65		
Andrade (2015) Brasília	Feminino	176	27	353	53	135	20	654	X ² (2, N= 1318) = 23.284, p < 0.00001
	Masculino	235	36	238	36	181	28		

Fonte: Elaboração própria.

Os resultados dispostos na tabela 28 referem-se a dados de fala da cidade Brasília, os quais foram selecionados por apresentarem resultados também para a variante **cê**. O estudo de Andrade (2010) constatou que, dentre as três variantes, a forma **tu** foi a mais frequente na amostra na fala de mulheres (145/330 – 44%) e de homens (128/244 – 52%). Já no estudo de Andrade (2015) a forma mais frequente passa a ser a variante **você** para ambos os sexos/gêneros. Calmon (2010) não constatou dados de **tu** em nenhuma das amostras analisadas. Seus resultados evidenciaram que a variante **cê** foi a forma preterida na fala dos homens apenas

na amostra Fala Casual de Brasília (75/116 – 65%). Em todos os estudos os resultados foram estatisticamente significativos.

Tabela 29 - Correlação entre o uso das formas de 2ºPS e o sexo/gênero do informante em contextos de referência determinada

Estudo	Sexo/gênero	Tu		Você		T	Qui-quadrado
		A	%	A	%		
Silva, M. (2019) Rio branco/AC	Feminino	29	44,6	36	55,4	65	$X^2 (1, N= 128) = 0.88, p = 0.3475$
	Masculino	22	34,9	41	65,1	63	

Fonte: Elaboração própria a partir de dados de Silva, M. (2019).

Silva, M. (2019) analisou apenas os dados das variantes **tu** e **você** em contexto de referência determinada. Os resultados evidenciaram que os informantes independentemente do sexo/gênero tendem a utilizar mais a variante **você**, sobretudo, os do sexo/gênero masculino (cf. tabela 29). Esses resultados não apresentaram significância estatística no teste de qui-quadrado, fato este que pode estar relacionado ao baixo quantitativos de dados analisados.

Tabela 30 - Emprego de **você** e **cê** conforme a referência do pronome e o sexo/gênero dos informantes (GESOL-SP-2000)

Sexo do falante	CÊ					
	REF Específica			REF Genérica		
	Nº/Total	%	p.r.	Nº/Total	%	p.r.
Mulheres	69/89	77,5	.64	480/913	52,6	.52
Homens	81/138	58,7	.40	333/719	46,3	.47
Total	150/227	66,1	Range: .12	813/1632	49,8	Range: .07
<i>Input: 0.692 l.l. = -121.374 p < 0.02</i>			<i>Input: 0.465 l.l. = -909.954 p < 0.05</i>			

Fonte: Nascimento (2011, p. 157).

Para a variante sexo/gênero do falante, Nascimento (2011) realizou o cruzamento desta com a variável tipo de referência (cf. tabela 30). Os resultados evidenciaram que, quando a referência é específica, a variante **cê** foi favorecida tanto na fala de mulheres (69/89 – 77,5%) quanto de homens (81/138 – 58,7%). Quando a referência é genérica, a variante **cê** foi a preterida na fala das mulheres, no entanto as taxas percentuais diminuem consideravelmente (480/913 – 52,6%) em comparação as obtidas em contextos de referência específica. Os usos **cê** na fala de homens em contextos de referência genérica também foram distintos daqueles obtidos para a referência específica, tendo em vista que a variante **cê** foi sutilmente desfavorecida (333/719 – 46,3%). Tais resultados foram estatisticamente significativos.

Os resultados obtidos por Nascimento (2011) confirmam a importância de analisar dados de referência específica separados daqueles de referência genérica, tendo em vista que usos das variantes **você** e **cê**, no estudo da referida pesquisadora, tiveram um padrão de comportamento distinto. Destarte, esse comportamento pode se repetir em outras realidades linguísticas.

4.6 PROBLEMAS METODOLÓGICOS NO ESTUDO DA VARIAÇÃO DAS FORMAS DE 2ªPS: LIMITAÇÕES DOS ESTUDOS REALIZADOS

Assim como o mapeamento feito por Scherre *et al.* (2015), a presente sistematização dos resultados das pesquisas desenvolvidas a respeito da variação das formas **tu**, **você** e **cê** deve ser considerada com cautela, uma vez que essas pesquisas se diferenciam uma da outra em maior ou menor grau. Como ressaltamos na subseção 4.1, não houve uma “agenda de pesquisa” para estabelecer um padrão metodológico a ser seguido com o intuito de permitir maior comparabilidade entre os estudos, porém nem tinham a obrigação de fazê-la, pois tal fato não era o objetivo destes.

A sistematização feita evidencia que algumas decisões metodológicas adotadas pelos pesquisadores precisam ser reconsideradas para possibilitar um melhor esclarecimento sobre o padrão de comportamento das variantes **tu**, **você** e **cê**, como, por exemplo, analisar conjuntamente dados: das variantes em contextos de referência determinada e indeterminada; das variantes **você**, **ocê** e **cê**; oriundos de comunidades de fala distintas; das variantes em funções sintáticas distintas. Ainda temos a questão de que, em termos gerais, os estudos controlam as variáveis independentes de forma distinta, fazendo com que não haja uma colinearidade entre estas de modo a propiciar comparabilidade efetiva dos resultados das pesquisas realizadas. Tais ações podem fazer emergir maior ou menor frequência de usos de uma ou de outra variante, principalmente, no caso do pronome **tu** em localidades que este não é a primeira forma da comunidade. Além disso, outra problemática refere-se à estratificação das amostras dos estudos, visto que esta interfere nos resultados das frequências das referidas formas de pronominais, uma vez que não somente aspectos sociais, mas também interacionais, estilísticos e estruturais influenciam nos usos, sendo fundamentais para o entendimento do fenômeno.

Diante da discussão realizada, fica evidente que, para realizar um estudo mais acurado a respeito do fenômeno, é difícil e complexo. Na presente tese, realizamos as análises

considerando os seguintes aspectos: i) análise do tipo de referência como uma variável dependente a partir da perspectiva de que consistem em duas funções distintas; ii) análise apenas dos pronomes retos explícitos na posição de sujeito, visto que o uso de formas de 2ªPS em outras posições possui condicionamentos distintos; iii) análise separadamente dados obtidos a partir de comunidade de fala e da utilização de coleta distintas para verificar os efeitos das variáveis controladas, correlacionando-os; iv) controle variáveis relacionadas tanto ao perfil social do entrevistado quanto ao do entrevistador. A partir dessas decisões metodológicas, esperamos obter resultados mais fidedignos às peculiaridades do fenômeno em estudo.

5 O USO VARIÁVEL DAS FORMAS PRONOMINAIS TU, VOCÊ E CÊ NO FALAR DE UNIVERISITÁRIOS ITABAIANENSES: ANÁLISE, RESULTADOS E DISCUSSÃO

O objetivo da presente seção é apresentar e discutir o padrão de comportamento referencial das variantes **tu**, **você** e **cê** no falar de universitários itabaianenses. Para tanto, apresentamos, na subseção a seguir, os procedimentos seguidos para a identificação de tal padrão comportamental na fala dos estudantes. Na subseção seguintes, discutimos os resultados obtidos com o desenvolvimento da presente pesquisa.

5.1 O PONTO DE PARTIDA PARA O DESENVOLVIMENTO DO ESTUDO: ASPECTOS METODOLÓGICOS

As decisões metodológicas constituem uma das principais ações do desenvolvimento de uma pesquisa científica a respeito de qualquer objeto. Portanto, elas devem ser adequadas ao problema de pesquisa e aos objetivos traçados, assim como serem claramente delineadas. Esta seção é dedicada à exposição de todo o percurso metodológico que seguimos para o estudo do uso variável dos pronomes sujeito **tu**, **você** e **cê** em contextos de função referencial determinada (expressão da 2ªPS) e indeterminada (indeterminação do sujeito) no falar itabaianense, mais especificamente, de estudantes da Universidade Federal de Sergipe, *Campus* Prof. Alberto Carvalho. Compõem o nosso *corpus* de análise entrevistas sociolinguísticas (amostra Deslocamentos-UFS/ITA (2018)) e interações conduzidas (amostra Rede Social de Informantes Universitários de Itabaiana/SE (2013)).

Antes de apresentarmos os procedimentos metodológicos propriamente ditos, discorreremos sobre a noção de comunidade de prática e apresentamos a que foi escolhida como *locus* da pesquisa. Na sequência, subseção 5.1.2, expomos os procedimentos seguidos para constituir a Amostra de falantes universitários de Itabaiana/SE em 2018. Na subseção 5.1.3, apresentamos a amostra *Rede Social de Informantes Universitários de Itabaiana/SE*. Posteriormente, apresentamos as variáveis que controlamos para realização deste estudo. E, por fim, descrevemos a natureza da análise e o tipo de tratamento estatístico aplicado aos dados.

5.1.1 Comunidade de prática em análise

Para desenvolver a presente pesquisa, focalizamos na comunidade de prática da Universidade Federal de Sergipe, *Campus* Prof. Alberto Carvalho, situado no município de Itabaiana/SE. Antes de apresentarmos as características dessa comunidade de prática (na subseção 5.1.1.2), é necessário discutirmos o conceito de comunidade de práticas. É o que fazemos na seção a seguir.

5.1.1.1 Compreendendo a noção de comunidade de prática

Tradicionalmente, seguindo os preceitos de Labov (2008 [1972]), as pesquisas sociolinguísticas variacionistas tomam o estudo da variação/mudança linguística centrado em comunidades de fala. Uma comunidade de fala é considerada, na perspectiva laboviana, como um agrupamento de indivíduos que compartilham normas e atitudes sociais perante uma língua ou uma variedade linguística (LABOV, 2008 [1972]). No entanto, Eckert (2000) propõe uma mudança de foco: que os pesquisadores passem a centrar também o estudo da variação linguística em comunidade de prática. Esta pode ser entendida como uma junção de pessoas, de forma contínua/frequente, em prol de um empreendimento comum (ECKERT; MCCONNEL-GINET, 2010 [1992]). É no decurso de uma atividade que as pessoas realizam conjuntamente, em torno de um objetivo comum, que há o surgimento das práticas – “o modo de fazer as coisas, modos de falar, crenças, valores, relações de poder” (ECKERT; MCCONNEL-GINET, 2010 [1992], p. 102). Uma comunidade de prática pode ser concebida “por pessoas trabalhando juntas em uma fábrica, habitués de um bar, companheiros de brincadeira em uma vizinhança, a família nuclear, parceiros policiais e seu etnógrafo, a Suprema Corte etc.” (ECKERT; MCCONNEL-GINET, 2010 [1992], p. 102). Além disso, Eckert e McConnel-Ginet (2010 [1992]) advertem que:

Comunidades de prática podem ser grandes ou pequenas, intensas ou difusas; elas nascem e morrem, podem sobreviver a muitas mudanças de membros e podem estar intimamente articuladas a outras comunidades. As pessoas participam de múltiplas comunidades de prática, e a identidade individual é baseada nesta participação. Em lugar de conceber o indivíduo como uma entidade à parte, pairando sobre o espaço social, ou como um ponto em uma rede, ou como membro de um conjunto específico ou de um conjunto de grupos, ou como um amontoado de características sociais, precisamos enfocar as comunidades de prática. Tal foco possibilita-nos ver o indivíduo como agente articulador de uma variedade de formas de participação em múltiplas comunidades de prática (ECKERT; MCCONNEL-GINET, 2010 [1992], p. 102-103, grifos nossos).

Assim, uma comunidade de prática consiste em uma construção social, estando relacionada às práticas cotidianas dos indivíduos, os quais participam de múltiplas comunidades e moldam seus usos linguísticos de acordo com a sua participação, com o papel desempenhado na situação interlocutiva (ECKERT, 2006; ECKERT; MCCONNELL-GINET, 2010 [1992]). Nessa perspectiva, a variação linguística é vista “como um recurso para a construção do significado social” (FREITAG; MARTINS; TAVARES, 2012, p. 923). A realização de estudos dos padrões de fala com foco em comunidade de prática permite a observação da atuação de valores sociopessoais, os quais podem interferir em uma deriva ou não. Esse é o caso da variação de usos das formas pronominais de 2ªPS que estão relacionados, a depender da comunidade, à atuação de valores sociopessoais.

Tal perspectiva de comunidade é adotada por sociolinguistas que buscam estudar a variação a partir de uma dimensão estilística, a fim de apreender, detalhadamente, a dinâmica do valor social das variáveis linguísticas (cf. FREITAG; MARTINS; TAVARES, 2012). Conseqüentemente, analisa-se como a construção da identidade do indivíduo e a construção do significado social são efetivadas (ARAÚJO, 2014). A observação de comunidades de prática, portanto:

(...) permite ao pesquisador identificar como as variantes linguísticas assumem significado social, possibilitando estabelecer relação mais direta entre língua e significado do que um estudo baseado em uma comunidade de fala, que, dado o seu delineamento, não permite controlar as relações estabelecidas entre os falantes e suas implicações na dinâmica linguística (FREITAG, 2013, p. 158-159).

Nos termos que propõem Eckert e McConnell-Ginet (2010 [1992]), é possível considerar que a Universidade Federal de Sergipe se constitui como uma comunidade de prática, uma vez que há um conjunto de pessoas agregadas para aprender, construir e fazer a gestão do conhecimento (cf. FREITAG; MARTINS; TAVARES, 2012). É justamente sobre essa comunidade de prática que se debruça a presente pesquisa. Na subseção a seguir discorreremos sobre ela.

5.1.1.2 *Lócus* da pesquisa: Universidade Federal de Sergipe, *Campus* Prof. Alberto Carvalho

A Universidade Federal de Sergipe é constituída por seis *Campi*, a maioria decorrente do Programa de Expansão com Interiorização da Universidade Pública (EXPANDIR). O *Campus* Cidade Universitária Prof. José Aloísio de Campos, situado no município de São Cristóvão, região metropolitana da capital do estado, é a sede da UFS, sendo este responsável,

a nível macroinstitucional, pelo gerenciamento dos outros *Campi*. Os demais *Campi* da instituição, conforme resumo executivo, nº 1, 2017, são:

- i) *Campus* Prof. João Cardoso Nascimento, situado em Aracaju/SE (instalado em 1984);
- ii) *Campus* Prof. Alberto Carvalho, situado no município de Itabaiana/SE (instalado em 2006);
- iii) *Campus* de Laranjeiras/SE (instalado em 2007);
- iv) *Campus* Universitário Prof. Antônio Garcia Filho, situado na cidade de Lagarto/SE (instalado em 2011);
- v) *Campus* do Sertão, situado em Nossa Senhora da Glória/SE (instalado em 2014).

Dentre esses *Campi*, a presente pesquisa focaliza o estudo de dados de fala de universitários provenientes do *Campus* Prof. Alberto Carvalho (Itabaiana/SE), o qual é considerado aqui como uma comunidade de prática, visto que os estudantes permanecem no ambiente universitário pelo menos quatro horas diárias, desenvolvendo atividades, compartilhando conhecimentos, valores/crenças, estilo/variedade linguística etc. Utilizamos dados de falas de universitários desse *Campus*, a partir de dois tipos de coletas: entrevistas sociolinguísticas e interações conduzidas.

O *Campus* Professor Alberto Carvalho está situado no município de Itabaiana/SE. A cidade de Itabaiana localiza-se no Agreste Central Sergipano, a, aproximadamente, 57 km da capital Aracaju⁶¹. O município possui uma área de 336.693 km² e uma estimativa populacional, em 2019, de 95.427 habitantes⁶². A figura a seguir apresenta a localização da cidade de Itabaiana, em Sergipe.

⁶¹ Quilometragem obtida a partir do Google Maps.

⁶² Estimativa do Instituto Brasileiro de Geografia e Estatística (IBGE). Informação disponível em: <https://cidades.ibge.gov.br/brasil/se/itabaiana/panorama>.

Figura 19 - localização da cidade de Itabaiana, em Sergipe



Fonte: Google Maps (adaptado).

É importante ressaltar que o município de Itabaiana é uma microrregião que se sobressai economicamente por possuir o maior comércio do interior do Estado. O setor comercial é caracterizado como “forte”. Dentre os motivos para que isso aconteça está o fato de que a cidade funciona como entreposto comercial na medida em que a população das cidades circunvizinhas vem consumir os produtos do comércio de Itabaiana (FREITAG; MARTINS; TAVARES, 2012). Portanto, é uma cidade receptora que gera grande mobilidade social.

O grande desenvolvimento econômico da cidade de Itabaiana contribuiu para que o Governo Federal a escolhesse para receber o primeiro *Campus* universitário do programa de Expansão com Interiorização da Universidade Pública (EXPANDIR). A criação do *Campus* Professor Alberto Carvalho (figura 20) ocorreu em novembro de 2005 por meio da Resolução nº 19 do Conselho Superior (CONSU) da UFS (JESUS, 2019). As atividades do *Campus* foram iniciadas no dia 14 de agosto de 2006 e há, atualmente, dez cursos em funcionamento, dos quais sete são de licenciatura (Biologia, Física, Letras-Português, Pedagogia e Química) e três são de bacharelados (Administração, Ciências Contábeis e Sistema de Informação). Diariamente, o *Campus* recebe cerca de 2000 estudantes (conforme resumo executivo, nº 1, 2017).

Figura 20 - *Campus* Professor Alberto Carvalho da UFS



Fonte: Portal UFS Itabaiana/SE.⁶³

Em decorrência da instalação do *Campus* universitário, a probabilidade de ter acesso ao nível superior aumentou para grande parte da população itabaianense e das cidades circunvizinhas. Para muitos estudantes do *Campus* Professor Alberto Carvalho, ser universitário representa uma conquista familiar e uma possibilidade de mudança da realidade vivenciada. Além disso, ter acesso ao nível superior consiste em uma grande responsabilidade para muitos estudantes, pois pesa o fato de estes serem “o primeiro universitário em uma família de pais que não tiveram a oportunidade de ter acesso à escolarização” (FREITAG; MARTINS; TAVARES, 2012, p. 932).

É relevante destacar que o perfil do estudante desse *Campus* é diferente do estudante do *Campus* São Cristóvão/SE, visto que a maioria dos estudantes é proveniente, tanto da zona urbana quanto da zona rural, da cidade de Itabaiana/SE e, principalmente, das cidades circunvizinhas. Poucos estudantes são da capital ou de fora do estado (geralmente são de cidades da Bahia ou de Alagoas que são próximas da cidade de Itabaiana/SE, como, por exemplo, a cidade baiana Coronel João Sá). Por ser um *Campus* menor (o quantitativo de estudantes é cerca de 11% do quantitativo de São Cristóvão/SE), os estudantes possuem um maior contato entre si, o que possivelmente faz com que tenham um perfil linguístico mais homogêneo (menos divergente, nos termos da teoria da acomodação linguística).

5.1.2 Constituição da amostra: entrevistas sociolinguísticas de Itabaiana/SE (2018)

Para realizar análise da fala de estudantes da Universidade Federal de Sergipe - *Campus* Prof. Alberto Carvalho foi necessário constituir uma amostra com gravações de entrevistas sociolinguísticas. Essa necessidade de coleta decorre do fato de o banco de dados *Falares*

⁶³ Disponível em: <http://itabaiana.ufs.br/conteudo/64526-Campus-prof-alberto-carvalho-ufs-itabaiana-tem-seu-primeiro-programa-de-mestrado-academico-aprovado-pela-capes>.

Sergipanos não possuir uma amostra recente de entrevistas desse *Campus*. O banco possui apenas a amostra Falantes Cultos de Itabaiana/SE que foi constituída entre 2009 e 2011.

Os estudantes que participaram da amostra são de Itabaiana/SE e das cidades circunvizinhas. Com base nessas características de mobilidade/deslocamentos sociais dos estudantes, estratificamos a amostra em três células sociais, a saber: sexo/gênero (masculino e feminino); deslocamento (D. I – estudantes que moram em Itabaiana; D. II – estudantes que vêm de outra cidade (vai e volta); D. III – estudantes que passaram a morar em Itabaiana por causa da universidade)⁶⁴; e período de curso (Inicial – até o 2º período; Meio – 4º, 5º e 6º períodos; Final - a partir do sétimo período). Assim, constituímos uma amostra de 18 células, para cada uma foram selecionados três estudantes (tabela 31), perfazendo um quantitativo de 54 entrevistas. A estratificação, a partir do controle de deslocamentos sociais e período de curso, foi delineada pela Prof.^a Dr.^a Raquel Meister Ko. Freitag, coordenadora do Projeto Falares Sergipanos – etapa 2, ao qual a presente amostra está vinculada. Essa estratificação foi seguida também na constituição da amostra de entrevistas sociolinguísticas do *Campus* São Cristóvão/SE, realizada por Ribeiro (2019) e por Côrrea (2019). A coleta das entrevistas para compor as duas amostras foi feita em 2018 concomitantemente.

Tabela 31 - Estratificação da amostra

Sexo/gênero	Período de curso					
	Início de curso		Meio do curso		Final de curso	
	M	F	M	F	M	F
Deslocamento I	3	3	3	3	3	3
Deslocamento II	3	3	3	3	3	3
Deslocamento III	3	3	3	3	3	3
Total parcial	18		18		18	
Total	54					

Fonte: Elaboração própria.

No total, foram coletadas 80 entrevistas sociolinguísticas⁶⁵, aos moldes laboviano, durante o período de abril a agosto de 2018. A maioria das entrevistas teve uma duração média

⁶⁴ Haveria um outro tipo de deslocamento com apenas os estudantes que não são sergipanos, mas que vieram morar no estado por motivos de estudos ou outro qualquer (conforme a estratificação delineada pela coordenadora do Projeto Falares Sergipanos – etapa 2), assim como fizeram Ribeiro (2019) e Côrrea (2019). Como não conseguimos encontrar informantes desse perfil para preencher as células, não controlamos esse tipo de deslocamento.

⁶⁵ A constituição e todo o planejamento de coleta dessa amostra foram realizados conjuntamente com Josilene de Jesus Mendonça, que se vale da mesma amostra para desenvolver sua pesquisa de doutorado intitulada em

de 60 minutos (algumas gravações possuem uma maior duração outras um pouco menos), totalizando aproximadamente 98 horas de gravação. Para considerar as 80 entrevistas coletadas (nesta amostra há a inclusão das entrevistas realizadas com estudantes de outros estados), a estratificação da amostra pode ser realizada quanto ao sexo/gênero dos participantes, conforme a tabela abaixo:

Tabela 32 - Estratificação da amostra quanto à variável sexo/gênero

SEXO/GÊNERO	
Masculino	Feminino
40	40
Total	80

Fonte: Elaboração própria.

Quanto à seleção desses informantes, esta ocorreu, em sua maioria, a partir da abordagem “bola de neve” (cf. MILROY; MILROY, 1992). Nessa abordagem, um informante, que possui contato com os documentadores (entrevistadores), indica outro para fazer parte da pesquisa. Como as documentadoras possuíam contato com estudantes da instituição, foi possível utilizar esse método para realizar as entrevistas. A existência de contato com estudantes da instituição decorre do fato de duas documentadoras serem estudantes do *Campus* e uma das pesquisadoras ter atuado como professora contratada do Departamento de Letras do *Campus*, além disso, as documentadoras são da própria cidade de Itabaiana/SE ou de cidades circunvizinhas. Quando não conseguíamos colocar essa abordagem em prática, abordávamos estudantes no *Campus* e os convidávamos para participar da pesquisa. Constituir a amostra foi bastante dispendioso, principalmente, pelo *Campus* possuir um quantitativo menor de estudantes se compararmos com o *Campus* de São Cristóvão/SE. No quadro a seguir, apresentamos o perfil social das documentadoras da presente amostra.

Quadro 12 - Perfil social das documentadoras da amostra

Documentadores	Escolaridade	Idade	Local de criação
DOCA	Doutoranda	30	Moita Bonita/SE
DOJ	Doutoranda	28	Itabaiana/SE
DOCC	Graduanda	21	Campo do Brito/SE
DOCD	Graduanda	21	Campo do Brito/SE

Fonte: Elaborado pela autora.

Traços semânticos da referência à primeira pessoa do plural no português brasileiro: um estudo em tempo real (MENDONÇA, 2022). As bolsistas do Programa de Iniciação Científica, Cósia Karine Vieira Borges e Damiana Karina Vieira Borges, colaboraram no processo de coleta das entrevistas.

É importante ressaltar que o roteiro⁶⁶ de perguntas (cf. APÊNDICE A), utilizado durante a realização das entrevistas, foi elaborado minuciosamente para que, por meio deste, possa-se controlar as seguintes variáveis independentes que podem barrar ou não uma deriva na língua: nível sociocultural, rede de relacionamento ampla ou restrita (MILROY, 1980), grau de mobilidade social (no sentido do quanto a pessoa se desloca da sua cidade para outros lugares) (MARSHALL, 2014). Além disso, é relevante destacar que a coleta seguiu todos os preceitos do comitê de ética científica. Antes de iniciar a gravação de cada entrevista, informávamos ao participante os objetivos da pesquisa, os possíveis riscos envolvidos na sua participação e solicitávamos que este lesse o termo de Consentimento Livre e Esclarecido (ANEXO A) e, caso concordasse com o que constava no termo, assinasse-o. Também solicitamos que os participantes preenchessem uma ficha social (ANEXO B) para que pudéssemos verificar se eles se encaixavam na estratificação apresentada na tabela 31. A amostra de entrevistas sociolinguísticas constituída está vinculada ao projeto *Banco de Dados Falares Sergipanos – etapa 2*, sob a coordenação da Profa. Dra. Raquel, o qual foi aprovado pelo CEP/CONEP com o número CAAE 68318317.0.0000.5546.

Finalizada a coleta das entrevistas sociolinguísticas, procedeu-se ao processo de transcrição destas. Vale ressaltar que não se trata de uma tarefa fácil, visto que é uma etapa que exige muito tempo e atenção para sua realização, ainda mais que são quase 100 horas de gravação. As entrevistas foram transcritas por meio do *software* Elan (BRUGMAN; RUSSEL, 2004), o qual permite o alinhamento entre áudio e texto da transcrição. Para tanto, adotamos as normas de transcrição utilizadas pelo Grupo de Estudos em Linguagem Interação e Sociedade (GELINS), as quais tomam como ponto de referência os princípios ortográficos da escrita do português com algumas adaptações (ANEXO C). Feito isso, organizamos o *corpus* e criamos um código de identificação para cada um dos informantes, resguardando assim, o anonimato. Posteriormente, é necessário realizar a revisão das transcrições para verificar se estas foram feitas de forma adequada e, conseqüentemente, tentar diminuir o número de segmentos incompreendidos pelos transcritores. Concluído esse processo, a amostra passará a compor o *Banco de Dados Falares Sergipanos*. Portanto, as 80 entrevistas coletadas consistem em um produto técnico decorrente da presente tese que irá compor o referido banco de dados e, conseqüentemente, contribuirá para o desenvolvimento de novas pesquisas sobre o falar Sergipano.

⁶⁶ O roteiro de entrevista foi elaborado pelas pesquisadoras Andréia Silva Araujo e Josilene de Jesus Mendonça.

Tabela 33 - Estratificação da amostra utilizada

Deslocamento	Sexo/gênero	
	M	F
Deslocamento I	4	4
Deslocamento II	4	4
Total	16	

Fonte: Elaborado pela autora.

Em virtude do volume de dados das formas pronominais nas entrevistas coletadas e do tempo escasso para a extração/análise deste, optamos pela análise de 16 entrevistas (cf. tabela 33) da amostra que constituímos a partir da comunidade de prática UFS-Campus-ITA. No quadro a seguir, apresentamos os dados sociais dos informantes selecionados.

Quadro 13 – Dados sociais dos informantes selecionados da amostra UFS-Itabaiana/SE (2018)

Informante	Sexo/gênero	Deslocamento	Idade	Curso/período	Cidade
Ana.P	Feminino	I	24	Pedagogia/8º	Itabaiana/SE
Mid	Feminino	I	22	Química/2º	Itabaiana/SE
Ray	Feminino	I	19	Geografia/2º	Itabaiana/SE
She	Feminino	I	25	Pedagogia/8º	Itabaiana/SE
Dav	Masculino	I	19	Física/2º	Itabaiana/SE
Eri	Masculino	I	22	Física/8º	Itabaiana/SE
Reg	Masculino	I	24	Biologia/10º	Itabaiana/SE
Thi	Masculino	I	19	Física/2º	Itabaiana/SE
Ana	Feminino	II	22	Pedagogia/8º	Moita Bonita/SE
Ani	Feminino	II	19	Geografia/1º	Moita Bonita/SE
Cla	Feminino	II	22	Administração/10º	Frei Paulo/SE
Ian	Feminino	II	22	Pedagogia/2º	Macambira/SE
Car	Masculino	II	26	Matemática/12º	Moita Bonita/SE
Eve	Masculino	II	30	Administração/10º	Frei Paulo/SE
Jos	Masculino	II	21	Geografia/2º	Malhador/SE
Mar	Masculino	II	19	Geografia/2º	Malhador/SE

Fonte: Elaborado pela autora.

Para selecionarmos os 16 informantes (quadro 13) dentre os 80 contactados para constituir a amostra de entrevistas, utilizamos como parâmetro o perfil dos informantes que fazem parte da amostra de interações conduzidas, quais sejam: um grupo de informantes que é

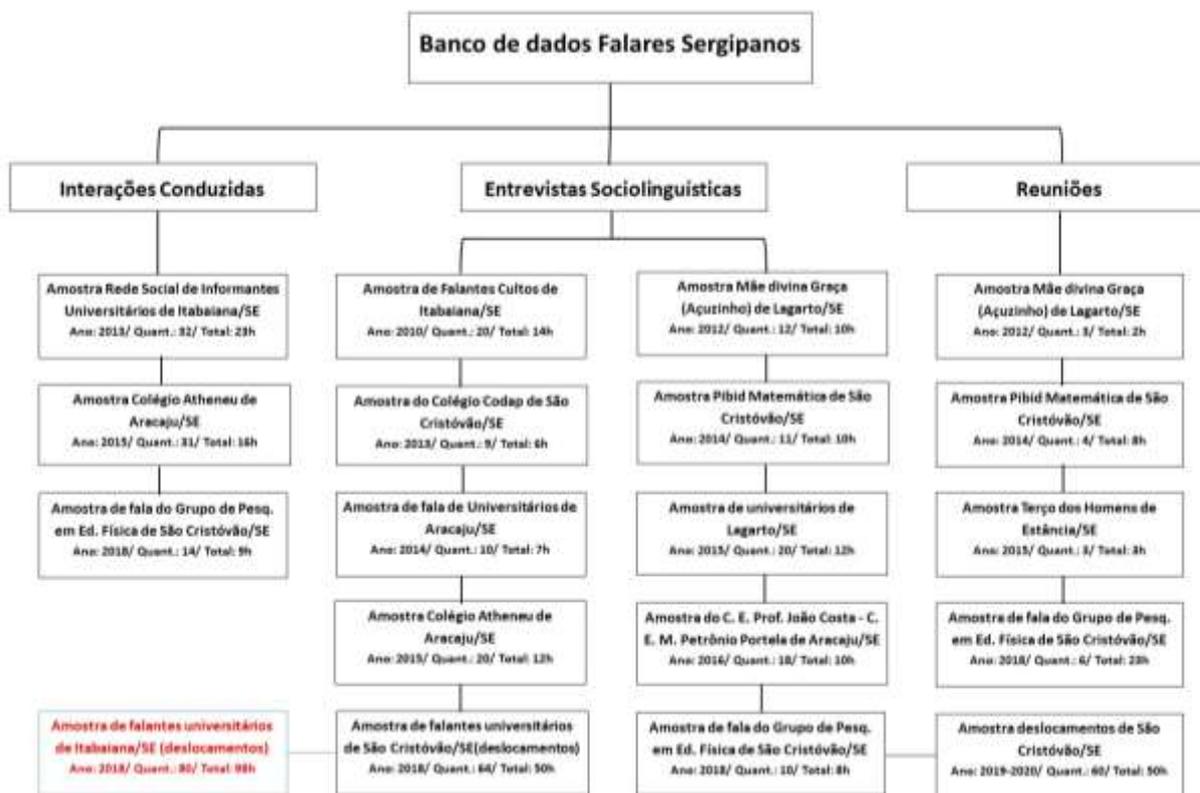
morador de Itabaiana/SE e um grupo de informantes que é morador de uma cidade do agreste sergipano - Frei Paulo. Como não há na amostra oito entrevistas de informantes de uma mesma cidade circunvizinha a Itabaiana/SE, selecionamos aqueles que fossem das cidades mais próximas. Além disso, um outro critério utilizado foi o entrevistador e o entrevistado não serem íntimos, a fim de compararmos dados de fala de informantes não íntimos com os dados de fala de informantes íntimos.

5.1.3 A amostra rede social de informantes universitários de Itabaiana/SE (2013)

O *Banco de Dados Falares Sergipanos*, vinculado ao Grupo de Estudos em Linguagem, Interação e Sociedade (GELINS), “é uma base de documentação linguística ampla da variedade de português falado em Sergipe [...]. Seu propósito é dar subsídios à investigação de variedades linguísticas do português, em seus diferentes níveis (do morfofonológico ao discursivo) [...]” (FREITAG, 2013, p. 156). Esse banco de dados é composto por dados de fala de informantes das seguintes cidades sergipanas: Aracaju (capital), São Cristóvão, Lagarto, Estância e Itabaiana.

O termo amostra “refere-se ao grupo de indivíduos [...] selecionados para representar, no estudo, a população ou o universo do qual fazem parte e que o pesquisador deseja estudar” (GUY; ZILLES, 2007, p. 109). As amostras desse banco de dados são constituídas por gravações de entrevistas sociolinguísticas, interações conduzidas e reuniões. A figura a seguir ilustra as amostras que compõem o banco *Falares Sergipanos*.

Figura 21 - Organograma do banco de dados Falares Sergipanos



Fonte: Elaborado pela autora.

Além da amostra que constituímos, utilizamos a seguinte amostra do *Banco de Dados Falares Sergipanos: Rede Social de Informantes Universitários de Itabaiana/SE (2013)*, a qual possui CAAE de número 0386.0.107.000-11. Dessa forma, usamos duas amostras constituídas a partir de métodos distintos. Tal diferenciação possibilita identificar se o tipo de metodologia de coleta dos dados de fala interfere ou não nos usos das formas variantes estudadas.

A amostra *Rede Social de Informantes Universitários de Itabaiana/SE* foi constituída em 2013 a partir da aplicação do modelo metodológico de interações conduzidas (não há presença do entrevistador, mas sim de dois informantes que interagem e conduzem a interação, isto é, os próprios informantes conduzem o tópico discursivo), desenvolvido por Araujo (2014) e Santos (2014) para atender às necessidades de análises de seus objetos de estudo do mestrado. Esse modelo metodológico foi elaborado com o intuito de controlar variáveis pragmáticas (distância social, custo da imposição e poder relativo), relacionados à polidez linguística, que podem influenciar os usos linguísticos dos falantes (ARAUJO, 2014).

Araujo, Santos e Freitag (2014) tomaram a Universidade Federal de Sergipe, *Campus Prof. Alberto Carvalho*, como a comunidade de prática a ser investigada. Trata-se de uma

comunidade de prática composta por várias redes sociais⁶⁷, as quais são formadas conforme “o tipo de relação existente entre os membros, tais como: aluno-aluno, aluno-professor, aluno-funcionário, professor-professor, professor-funcionário, funcionário-funcionário” (ARAUJO, 2014, p. 56). Dentre esses tipos de relações, as autoras definiram a relação aluno-aluno como foco para pesquisa.

Para controlar a variável distância social, Araujo, Santos e Freitag (2014), basearam-se na proposta de controle de Blake e Josey (2003) e Oushiro (2011). Estas estabeleceram uma escala de gradação para controlar essa variável com base na frequência com que os informantes interagem. O controle dessa variável na constituição da amostra ocorreu por meio de uma escala de gradação “que vai do grau máximo de proximidade (grau 1) ao grau mínimo de proximidade (grau 5) entre os informantes” (ARAUJO; SANTOS; FREITAG, 2014, p. 105). O quadro a seguir descreve cada um dos graus da escala.

Quadro 14 - Escala de gradação para o controle da distância social entre os informantes da rede social

<i>Grau 1</i> – Bastante próximo. Os informantes possuem laços fortes por fazerem parte do círculo imediato de familiares/amigos e interagem diariamente;
<i>Grau 2</i> – Próximo. Os informantes não possuem laços fortes por não fazerem parte do círculo imediato de familiares/amigos, mas interagem frequentemente;
<i>Grau 3</i> – Próximo. Os informantes não possuem laços fortes por não fazerem parte do círculo imediato de familiares/amigos e não interagem frequentemente;
<i>Grau 4</i> – Neutro. Os informantes são conhecidos um do outro e interagem raramente;
<i>Grau 5</i> – Distante. Os interlocutores não se conhecem anteriormente e só interagem no momento da gravação da interação.

Fonte: Adaptado de Araujo, Santos e Freitag (2014, p. 105).

As pesquisadoras focalizaram no grau 1 e no grau 5 da escala de gradação para o controle da distância social entre os informantes. Foram selecionados 8 informantes para participar das gravações de interação, os quais foram distribuídos em dois grupos - cada um dos grupos possui duas mulheres e dois homens. Um dos requisitos era que os membros de cada grupo possuíssem relações de proximidade entre si, mas não com os informantes pertencentes ao outro grupo (ARAUJO; SANTOS; FREITAG, 2014). No quadro 15, são apresentadas algumas informações sociais desses informantes.

⁶⁷ A noção de rede social é entendida nessa perspectiva como um “conjunto de atores/pessoas que têm relações entre si, sejam elas por laços fortes (grau de proximidade alto), sejam por laços fracos (grau de proximidade baixo)” (ARAUJO; SANTOS; FREITAG, 2014, p. 103).

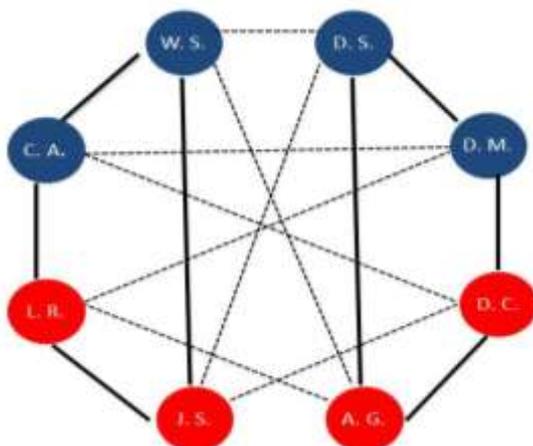
Quadro 15 - Dados sociais dos informantes em rede social pessoal

Informante	Sexo/gênero	Deslocamento	Idade	Curso/período	Cidade
D. C.	Feminino	I	28	Geografia/8º	Itabaiana/SE
A. G.	Feminino	I	25	Geografia/8º	Campo do Brito/SE
D. M.	Masculino	I	24	Geografia/8º	Itabaiana/SE
D. S.	Masculino	I	21	Geografia/8º	Itabaiana/SE
L. R.	Feminino	II	21	Pedagogia/8º	Frei Paulo/SE
J. S.	Feminino	II	19	Pedagogia/2º	Frei Paulo/SE
W. S.	Masculino	II	19	Ciências Cont./2º	Frei Paulo/SE
C. A.	Masculino	II	30	Administração/9º	Frei Paulo/SE

Fonte: Adaptado de Araujo, Santos e Freitag (2014, p. 104).

O primeiro grupo é formado pelos informantes D.S., D.M., D.C. e A.G. Estes estavam no último período do curso de Geografia e apenas a informante A.G. não morava em Itabaiana/SE. O segundo grupo é formado por W.S., C.A., L.R. e J.S. Diferentemente do primeiro grupo, os membros do segundo grupo são de três cursos: as informantes cursavam Pedagogia (uma estava no segundo período e a outra no penúltimo); os informantes masculinos eram de curso diferentes, um cursava Ciências Contábeis (segundo período) e o outro Administração (penúltimo período) (cf. exposto no quadro 15). Todos os membros deste grupo são provenientes da cidade de Frei Paulo, a qual está localizada no agreste sergipano, a cerca de 20 km de Itabaiana/SE. Apesar de os membros do segundo grupo não cursarem, na época da coleta, a mesma graduação, estes possuíam grau 1 de proximidade por manterem contato diariamente ao utilizarem o mesmo transporte de sua cidade para a universidade (ARAÚJO, 2014). Além disso, os membros dos dois grupos tinham idades próximas (de 21 a 30 anos). Na figura 22, são apresentadas as relações/conexões formadas pelos informantes considerando os laços fracos ou fortes existentes entre eles (MILROY, 1980).

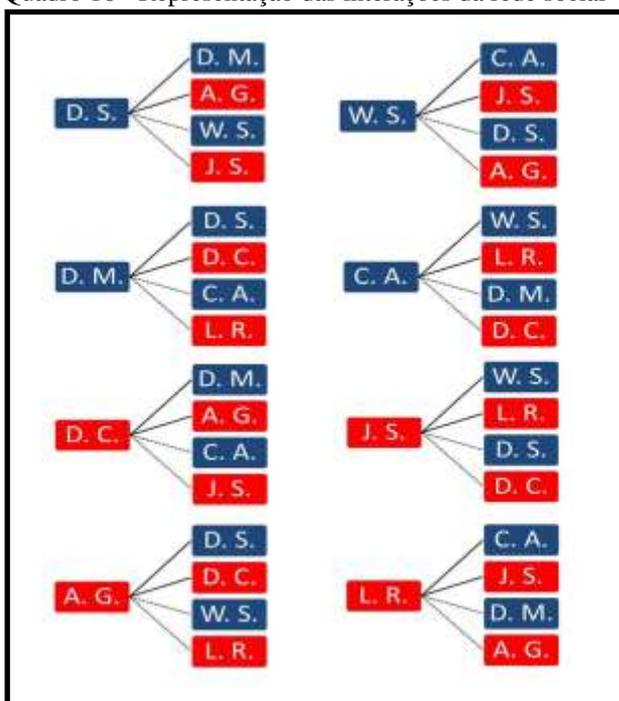
Figura 22 - Rede Social dos oito informantes universitários do *Campus Prof. Alberto Carvalho/UFS*



Fonte: Araujo, Santos e Freitag (2014, p. 108).

A partir dessas relações/conexões expostas na figura 22, pode-se constatar que os membros da rede estão conectados por linhas contínuas e tracejadas e, nos termos de Milroy (2002), representam respectivamente a existência de laços fortes e de laços fracos. Os membros de cada grupo são conectados por laços fortes entre si, enquanto os membros de grupos diferentes possuem laços fracos. No quadro 16, estão ilustradas as relações constituídas entre cada membro do grupo “(laço forte com linha contínua e laço fraco com linha tracejada) e com quem cada um interagiu (masculino em azul e feminino em vermelho)” (ARAUJO; 2014, p. 57).

Quadro 16 - Representação das interações da rede social



Fonte: Santos (2014, p. 36).

Conforme apresentado no quadro acima, cada informante interagiu com 4 pessoas diferentes (um homem e uma mulher, próximos dele; um homem e uma mulher, distantes dele). Dessa forma, é possível controlar se o sexo/gênero do interlocutor interfere na variação dos fenômenos linguísticos que ocorrerem na amostra. Logo, o controle da variável sexo/gênero pode ser realizado da seguinte forma (ARAUJO; SANTOS; FREITAG, 2014, p. 106):

- i) Feminino → masculino;
- ii) Feminino → feminino;
- iii) Masculino → feminino;
- iv) Masculino → masculino.

Diferentemente das entrevistas sociolinguísticas, nas interações conduzidas não há a presença do entrevistador e nem do roteiro de perguntas. Nestas interações, os próprios participantes conduzem a conversa a partir da seleção do tópico retratado nas microssituações presentes em cartões. Antes de iniciar a gravação, as pesquisadoras explicavam todo o procedimento que deveria ser seguido e disponibilizavam 50 cartões, nos quais constavam temas diversos por meio de microssituações (cf. ARAUJO; SANTOS; FREITAG, 2014), como, por exemplo, sobre questões sociais (ex. (24)) e vivências pessoais (ex. (25)). Cada participante escolhia aleatoriamente 10 cartões (dois de cada tipo de situação) para utilizar durante a interação. Com base no tipo de assunto introduzido na interação é possível identificar o custo da imposição “vão da aparente neutralidade a situações que envolvem a preservação das faces negativa e positiva” (ARAUJO; SANTOS; FREITAG, 2014, p. 106).

(24) Maltratar animais é crime e prevê pena de 3 meses a um ano de detenção. Bruna e Letícia presenciaram o vizinho espancando um cachorro. Bruna pensou de imediato em acionar a polícia militar ambiental, já Letícia pensou em prestar atendimento ao animal. (ARAUJO; SANTOS; FREITAG, 2014, p. 113)

(25) As intrigas sempre estão presentes entre os irmãos. Sempre há aquele momento em que um olha para o outro e diz: “nunca mais fale comigo”. Mas poucas horas depois eles já estão juntos novamente. (ARAUJO; SANTOS; FREITAG, 2014, p. 109)

Quanto à variável pragmática relações de poder, esta foi controlada a partir de quem estava com o domínio do tópico na interação, ou seja, de quem estava conduzindo o tópico, inserindo um novo tópico na interação. Após seguir todos esses procedimentos, Araujo (2014) e Santos (2014) realizaram as gravações das interações, as quais possuem em média 40 minutos. No total, foram gravadas 32 interações conduzidas (contabilizando aproximadamente 23 horas de gravação). Na presente pesquisa, utilizamos apenas as 16 interações realizadas entre íntimos, isso porque as gravações de entrevistas sociolinguísticas selecionadas o entrevistador e o

entrevistados possuem laços fracos, o que nos permitiu controlar o fator grau de intimidade entre os interlocutores.

5.1.4 As variáveis controladas

Os pronomes **tu**, **você** e **cê** possuem uma natureza referencial dupla: determinada e indeterminada. Para descrever o padrão de comportamento referencial dessas variantes, consideramos cada tipo de referência como uma variável dependente. Temos a expressão referência determinada quando o locutor utiliza uma das variantes para se referir a um interlocutor dêitico (ex. (26)) e a expressão da referência indeterminada quando o locutor utiliza uma das variantes para se referir a si mesmo ou às pessoas de forma geral (ex. (27)).

(26) A.G.: então nesse caso ou até mesmo se eu tivesse eu pre- eu poderia até dar entendeu? era melhor dizer assim “<<tame>> eu tô dando” nesse caso de como **cê** tá me falando de quê? de saúde... do que “ah tô emprestando sem saber se vou receber” e ficar “ah não me pagou não me pagou” e sabendo que foi pra algo de saúde então se eu tivesse eu preferia dar... do que emprestar

(D.S.cdt A.G.sdt P MF 02)

(27) A.G.: não não é hierarquia assim é ques- eu defendo essa questão assim que é o seu pai sua mãe eles querem o seu bem então **você** tem que tá ali pra... realmente... como não é dever obrigação a ele de ser assim mas de respeitar de tá ali não ir se voltar contra eles

(D.S.cdt A.G.sdt P MF 02)

Controlamos variáveis independentes de cunho interacional, social, estilístico e estrutural/formal. Variáveis interacionais são aquelas que possuem relação direta com a situação interlocutiva estabelecida entre o locutor e o alocutário/interlocutor. É a partir das relações estabelecidas entre estes que se pode constatar a atuação das dimensões da solidariedade e do poder propostas por Brown e Gilman (1960) e da dimensão da neutralidade aventada por Cook (1997), bem como a atuação da acomodação linguística proposta por Bell (1984), nos usos das formas pronominais **tu**, **você** e **cê**. No presente trabalho, controlamos as seguintes variáveis interacionais: i) grau de intimidade entre os interlocutores; ii) relações (as)simétricas quanto ao sexo/gênero dos interlocutores; iii) turno de fala; iv) pergunta proposital. As variáveis sociais controladas foram: v) sexo/gênero do falante e vi) deslocamento social. No que concerne às variáveis estilísticas, controlamos as seguintes: vii) tipo de sequência discursiva, viii) tópico discursivo, ix) tipo de discurso e x) tipo de amostra. As variáveis estruturais/formais selecionadas foram: xi) paralelismo formal e xii) efeito gatilho. Na sequência, discorreremos sobre tais variáveis e apresentamos suas respectivas hipóteses.

i) Grau de intimidade entre os interlocutores

A variável grau de intimidade entre os interlocutores refere-se ao quão próximo os interlocutores são um do outro: íntimos ou não íntimos. Aqueles que possuem relações de laço forte (parentes, amigos, namorado(a), colega de trabalho/estudo etc.) e que interagem diariamente/frequentemente são considerados como íntimos, ou seja, são próximos um do outro; portanto, possuem uma relação interpessoal de solidariedade, a qual favorece o uso de T. Já os interlocutores que possuem relações de laço fraco e que não/raramente interagem são considerados como não íntimos. Mesmo em uma interação em que não há diferenças de poderes é possível que os interlocutores utilizem a forma V por não haver intimidade suficiente entre eles (cf. BROWN; GILMAN, 1960). Assim, a proximidade/a distância social, em termos de grau de intimidade, pode condicionar o comportamento linguístico dos interactantes como bem destacam Brown e Levinson (2011 [1987]) na discussão a respeito da teoria de Polidez.

Estudos sociolinguísticos têm evidenciado que, quando a relação é de intimidade entre os interlocutores, há o favorecimento do uso da forma **tu** (cf. MARTINS, 2010; GUIMARÃES, 2014) e para as relações de não intimidade há o favorecimento da forma **você** (cf. MOTA, 2008; MARTINS, 2010; GUIMARÃES, 2014). Com base em tais resultados, para a regra variável de referência determinada, a nossa hipótese é de que haja o favorecimento das formas pronominais **tu** e **cê** quando os interlocutores são íntimos por acreditarmos que, neste tipo de interação, exista, durante a conversação, maior grau de informalidade; e entre os não íntimos, há o condicionamento da forma **você**. Para a regra variável de indeterminação do sujeito, a nossa hipótese é de a forma **você** seja favorecida independentemente do grau de intimidade entre os interlocutores por, possivelmente, este ser o pronome de uso de livre trânsito em Sergipe, ou seja, essa forma pronominal pode pertencer a dimensão da neutralidade, conforme propõe Souza (2011). Como selecionamos apenas as interações conduzidas nas quais os interlocutores são íntimos e as entrevistas nas quais entrevistador e entrevistado não são íntimos, é necessário realizar o cruzamento das amostras para verificar a influência dessa variável na escolha dos pronomes **tu**, **você** e **cê** pelos informantes.

ii) Relações (as)simétricas quanto ao sexo/gênero dos interlocutores

Uma das formas de se analisar a (as)simetria na relação entre os interactantes é a partir do controle do sexo/gênero destes. Temos o estabelecimento da simetria quando os

interlocutores são do mesmo sexo/gênero (masculino-masculino ou feminino-feminino) e da assimetria quando estes são de sexo/gênero diferentes (masculino-feminino ou feminino-masculino). De acordo com Brown e Gilman (1960), as relações de simetria estão relacionadas à dimensão da solidariedade e as de assimetria à dimensão de poder. Portanto, os falantes acomodam seus usos linguísticos a seus interlocutores, conforme discute Bell (1984) em sua teoria da acomodação.

Os resultados da pesquisa de Lucca (2005) evidenciam que relações simétricas quanto ao sexo/gênero dos interlocutores favorecem o uso da forma pronominal de solidariedade **tu**. Os demais estudos sobre o uso de **tu** *versus* **você** constataram que esta última forma é favorecida em todos os tipos de (as)simetria (DIAS, 2007; SILVA, S., 2019); nestes o quantitativo de ocorrência da forma **tu** é baixo, o que justifica tal resultado. O estudo de Nascimento (2011) sobre o uso de **você** *versus* **cê** evidenciou um sutil favorecimento da forma **cê** na relação feminino-feminino e um sutil favorecimento da forma **você** nos demais tipos de relação. A nossa hipótese é de que as formas **tu** e **cê** sejam mais frequentes em relações simétricas, principalmente, em contextos de regra variável de referência determinada.

iii) Turno de fala

Em uma conversação, há, no mínimo, a presença de dois participantes. Nesta pode ser estabelecida uma assimetria entre os participantes a partir do papel interacional desempenhado por estes, fazendo emergir a dimensão do poder (BROWN; GILMAN, 1960; BROWN; LEVINSON, 2011 [1987]). Na amostra de interações conduzidas, um dos participantes é o responsável por incluir e mudar o tópico discursivo da conversação (dominador do tópico discursivo – falante I), tendo, portanto, maior poder na situação interativa. Já o outro participante tem a função de desenvolver o tópico discursivo, possuindo, assim, menor poder na situação interativa (não dominador do tópico discursivo – falante II).

A variável turno de fala é controlada apenas nas interações conduzidas, pois, nas entrevistas sociolinguísticas, o entrevistador utiliza um roteiro prévio de perguntas, no qual contém formas pronominais em estudo. Aventamos a hipótese de que os usos dos pronomes em estudo, nos contextos de referência determinada, são mais frequentes na fala do participante dominador do tópico e que as taxas de frequência das formas **tu** e **cê** são mais altas, visto que, em situações assimétricas, aquele que está em uma posição hierárquica mais alta tende a utilizar

T e receber V (BRONW; GILMAN, 1960). Nos contextos de referência indeterminada, temos a hipótese de que a forma **você** é mais frequente tanto na fala do falante I quanto do falante II.

iv) Pergunta proposital

No roteiro de perguntas, utilizado nas gravações das entrevistas sociolinguísticas, foram incluídas perguntas que motivassem o uso dos pronomes **tu**, **você** e **cê** com referência determinada pelo entrevistado, como em (28) e (29). Isso porque as entrevistas sociolinguísticas não propiciam a troca de turnos tendo em vista que o foco desse tipo de coleta é que o entrevistado discorra sobre o tópico discursivo e não o entrevistado; e estas favorecem o uso dessas formas pronominais como estratégia de indeterminação do sujeito, como em (30).

(28) Doc.: Caso eu fosse um político, o que diria para mim?
(06ent.UFS-Itabaiana2018_desl. I_início_eve.ms.24)

(29) Doc.: O que eu tenho que fazer para deixar de ser consumista? Que dicas **tu** me daria?
(06ent.UFS-Itabaiana2018_desl. I_início_eve.ms.24)

(30) Doc.: o que você acha da coleta seletiva?
Inf.: a coleta seletiva eu acho que ela é importante... porque... ((BARULHO)).. **cê** tem que fazer... essa coleta... s- com frequência... e você... e também assim a questão... de **você**... s- fazer a separação... (est)... do lixo orgânico do lixo inorgânico eu acho que é um... eu acho que é assim que é de extrema importância fazer isso... ((BARULHO))
(06ent.UFS-Itabaiana2018_desl. I_início_eve.ms.24)

Assumimos a hipótese de que, nos contextos de perguntas propositais, os entrevistados utilizariam as formas **tu**, **você** e **cê** com valor referencial determinado em suas respostas, propiciando, sobretudo, o uso da primeira e da terceira formas referidas.

v) Sexo/gênero do falante

A variável sexo/gênero é tradicionalmente controlada em estudos sociolinguísticos e tem apresentado significância estatística para a ocorrência de fenômenos linguísticos variáveis em diferentes níveis linguísticos. A tendência apontada pelos estudos prévios realizados com dados da região nordeste (CALMON, 2010; ANDRADE, 2015; NOGUEIRA, 2013; SILVA, Suziane, 2017; SILVA, S., 2019), é de que, tanto homens quanto mulheres, usam com mais

frequência a forma **você**. Entretanto, ao separar os dados de referência determinada daqueles de referência indeterminada, Nascimento (2011), ao analisar dados de fala de São Paulo/SP, constatou que a variante **cê** é a favorecida na expressão de referência determinada na fala das mulheres e dos homens; e apenas na fala das mulheres quando temos a expressão de indeterminada. Com base dos resultados obtidos por Nascimento (2011), temos a hipótese de que, em contextos de função referencial determinada, os usos de **cê** apresentam taxas de frequência maiores na fala das mulheres, entretanto, tendo em vista os resultados dos demais pesquisadores, não suplantaria o uso da forma **você**. Em contextos de referência indeterminada, a variante **você** é a preterida na fala de ambos.

vi) Deslocamento social

O controle da variável deslocamento social visa identificar se o fator geográfico influencia no comportamento linguístico dos informantes. Controlamos dois níveis para essa variável: deslocamento I – estudantes da UFS-*Campus*-ITA que são moradores da cidade de Itabaiana/SE, local onde fica a comunidade de prática focalizada; deslocamento II – estudantes que moram em outra cidade da região Agreste e que se deslocam diariamente para UFS-*Campus*-ITA. Ao interagir, os estudantes da comunidade de prática UFS-*Campus*-ITA, assim como qualquer falante, realizam trocas linguísticas, passam a utilizar formas linguísticas de seus interlocutores a fim de se acomodarem linguisticamente à comunidade ou à norma acadêmica tendo em vista que “a faculdade é uma experiência de realocação comum”⁶⁸ (CAMPBELL-KIBLER *et al.*, 2014, p. 21). Espera-se que os estudantes tenham comportamento linguístico convergente, tanto para regra de referência determinada quanto para a indeterminada, utilizando a forma **você**, pronome que apresenta indícios de possuir livre trânsito no estado de Sergipe e relacionado ao estilo formal, pois apresenta avaliação positiva pela comunidade acadêmica (ARAUJO; MENDONÇA, 2018).

vii) Tipo de sequência discursiva

As sequências discursivas podem ser definidas como estruturas caracterizadas por marcas linguísticas de tempos verbais, aspectos sintáticos, relações lógicas etc. (PAREDES

⁶⁸ No original: “college is a common relocation experience” (CAMPBELL-KIBLER *et al.*, 2014, p. 21).

SILVA, 1999 apud FREITAG *et al.*, 2019). Controlamos quatro tipos de sequências discursivas: narrativa, argumentativa, expositiva-explicativa e injuntiva. A sequência narrativa, em termos gerais, apresenta relatos de fatos/acontecimentos de forma a estabelecer progressão temática e temporal (ex. (31)). A sequência argumentativa é caracterizada pela demonstração ou refutação de uma ideia/tese (ADAM, 1987), a fim de convencer o outro (leitor/ouvinte), com base em fatos, evidências etc., de que tal ideia tem fundamento, lógica de ser (ex. (32)). Quanto a expositiva-explicativa, esta é caracterizada pela apresentação de um determinado assunto/tema, expondo informações a respeito deste e as explicando (ex. (33)). Já a injuntiva trata-se de uma sequência que apresenta atos diretivos sobre como fazer algo, instruções de como realizar determinadas ações (ADAM, 1987), isto é, objetiva apresentar orientações/instruções de como o interlocutor deve proceder em determinada situação (ex. (34)).

(31) L.R.: Washington chegou "não vamos fazer o seguinte... eu fico com a sua turma que eu gosto de criança... e **você** vai pra o ensino fundamental e eu fico com o pessoal do maternal"... lá em Carira isso...

(C.A.cdt L.R.sdt P MF 18)

(32) D.M.: sim... sim sim sim é essa questão... mas **você** num disse se fosse ofertado pra poucos... uma quantidade menor... isso já vinha... deveria ser... oferecido já da base

(D.S.cdt D.M.sdt P MM 01)

(33) W.S.: na realidade gosto... mas queria tá cursando outro... esse assim se **cê** me perguntar se eu penso... em parar ele não de jeito nenhum... agora vontade de cursar outro eu tenho

(D.M.cdt D.S.sdt P MM 09)

(34) D.S.: mas é... **tu** brincava <<ni>> sítio mais quem? com seus irmãos era?

(D.S.cdt D.M.sdt P MM 01)

Por meio do tipo de sequência discursiva é possível identificar o grau de atenção prestada à fala (SNICHELOTTO, 2014). Nesse ínterim, sequências discursivas de cunho narrativo e injuntivo estão relacionadas ao estilo de menor grau de monitoramento (informal) e as sequências discursivas de cunho argumentativo e expositivo-explicativo estão relacionadas ao estilo de maior grau de monitoramento (formal). As pesquisas que apresentam resultados referentes a essa variável são sobre o falar da região Sul, onde predomina o uso da variante **tu** (exceção da cidade de Curitiba/PR (LOREGIAN-PENKAL; MENON, 2012)); já na região Nordeste, na maioria das localidades, o uso desta variante não é tão frequente (SCHERRE; ANDRADE; CATÃO, 2019). Partimos da hipótese de que as sequências discursivas argumentativas e expositiva-explicativas favorecem o uso da forma pronominal **você** por ser um contexto mais formal, tanto para a expressão de referência determinada quanto de referência

indeterminada. Enquanto sequências discursivas narrativas e injuntivas favorecem o uso da forma **tu** e **cê**, pois tendem a ser contextos mais informais e propiciarem maior envolvimento do informante, principalmente, em sequências narrativas. Estas duas sequências discursivas estão relacionadas à dimensão da solidariedade (BROWN; GILMAN, 1960), as quais podem favorecer o uso dos pronomes em estudo, sobretudo, **tu** e **cê** com referência determinada.

viii) Tópico discursivo

A variável tópico discursivo refere-se ao assunto abordado na situação interlocutiva, isto é, sobre o que os interlocutores estão falando. Assim como a variável tipo de sequência discursiva, o tópico discursivo é uma variável que permite identificar a mudança de estilo (mais formal ou menos formal), isso significa que essa variável influencia o grau de monitoramento da fala feito pelos interlocutores (BELL, 1984). Categorizamos a variável tópico discursivo em três níveis: educação (ex. (35)), questões sociais (ex. (36)) e vivências pessoais (ex. (37)).

(35) Inf.: por isso que eu acho que... se **você** privatizar as escolas... priva- privatizar pelo menos uma parte da educação... e transformar... porque se uma escola é privada ela pode ser regional... porque ela vai ser controlada pela aquela por pessoas daquela região... ou seja ela vai se adaptar

(04ent.UFS-Itabaiana2018_desl. I_início_thi.ms.19)

(36) Inf.: ah... assim... muita gente critica... mas... é uma (hes) é um processo... tá acontecendo a inflação que tá hoje vem de um processo... num sei se **você** lembra tem uma parte do governo da Dilma antes das... pegar o segundo mandato... que a energia tava um baratinha tava uma delícia... na verdade ela tava segurando os preços... da energia... aí quando a outra fui impeachemada ela perd- ela

(04ent.UFS-Itabaiana2018_desl. I_início_thi.ms.19)

(37) Inf.: eu perguntaria o porquê da fofoca... se **você** tem algo contra mim... e eu tentaria resolver isso da melhor forma possível... caso não resolvesse isso aí meu filho o bicho ia pegar... por que... eu não sou de levar desaforo pra casa...

(01ent.UFS-Itabaiana2018_desl. I_início_ray.fs.19)

Tomando como base os resultados de Lucca (2005) e Guimarães (2014), a nossa hipótese é de que assuntos de maior complexidade - educação e questões sociais - propiciam o uso de estilo de fala mais monitorada, o que favorece o uso da forma pronominal **você** em contextos de referência determinada e indeterminada. Já os assuntos de menor complexidade, como os classificados como vivências pessoais, propiciam o uso de estilo de fala menos monitorada, favorecendo o uso das formas **tu** e **cê** (conforme constatado nos estudos de Lucca

(2005) e Guimarães (2014) para a primeira forma referida) em contexto de referência determinada.

ix) Tipo de discurso

A variável tipo de discurso foi controlada a partir de dois níveis: discurso direto e discurso indireto. O discurso direto refere-se a fala elaborada no momento da enunciação (ex. (38)). O discurso indireto diz respeito à remissão de uma fala própria ou de terceiros realizada em um momento anterior a situação interlocutiva atual/enunciação (ex. (39)).

- (38) Inf.: o livre arbítrio é a sua escolha eu tenho que respeitar ... posso não concordar com você mas eu vou lutar até o fim pelo seu direito de dizer o que **cê** pensa o que **você** é

(01ent.UFS-Itabaiana2018_desl. I_início_ray.fs.19)

- (39) Inf.: porque assim Administração como a filosofia da Administração é "ah **você** pode administrar sua casa **você** pode administrar uma casa um- um- uma Multinacional" realmente ah você pode aplicar as coisas do da Administração no seu dia a dia de uma não sei você vai adaptando as coisa entendeu

(34ent.UFS-Itabaiana2018_desl. II_final_eve.ms.30)

Temos a hipótese de que, em contextos de referência determinada, há um favorecimento maior ao uso das formas **tu** e **cê** no discurso direto (conforme resultados de Nascimento (2011) para a última forma pronominal mencionada). Em contexto de referência indeterminada, a forma **você** é favorecida nos dois tipos de discurso (cf. NASCIMENTO, 2011).

x) Tipo de amostra

O tipo de amostra é uma variável que pode determinar o nosso comportamento linguístico nas situações interlocutivas. Em nosso estudo, utilizamos duas amostras que possuem procedimentos de constituição distintos: entrevistas sociolinguísticas e interações conduzidas. A nossa hipótese é de que as interações conduzidas, por seu caráter mais dialógico (propiciar mais trocas de turnos, com objetivo de que os dois participantes discorram sobre os tópicos discursivos), favoreçam a ocorrência de referência determinada e o uso das formas **tu** e **cê**. Quanto às entrevistas sociolinguísticas, estas favorecem a ocorrência da regra de indeterminação do sujeito. Além disso, temos a hipótese de que o uso da forma **você** é frequente nos dois tipos de amostras, principalmente, em entrevistas sociolinguísticas.

xi) Paralelismo formal

De acordo com Scherre (1998), há uma tendência na língua de que uma vez utilizada uma marca linguística, esta tende a se repetir em cadeia, princípio denominado de paralelismo formal. Controlamos essa variável a partir dos seguintes níveis: isolada (ex. (40)) primeira menção (ex. (41)); antecedida por **tu** (ex. (42)); antecedida por **você** (ex. (43)); e antecedida por **cê** (44)).

- (40) Doc.: as pessoas no local onde você mora costumam se ajudar?
 Inf.: não (est) ... **você** precisar de uma xícara de sal... eles dizem assim o mercadinho tá na esquina
 (01ent.UFS-Itabaiana2018_desl. I_início_ray.fs.19)
- (41) L.R: é mas chegar no teu estabelecimento al- **você** perceber alguém alguém pensa que tá fazendo você não tá vendo... se você vê você vai fazer o quê?...
 (J.S.cdt L.R.sdt P F_F 22)
- (42) Inf.: **tu** acha que **tu** gasta fora do normal é?
 (33ent.UFS-Itabaiana2018_desl. II_final_jos.fs.30)
- (43) Doc.: você considera Itabaiana uma cidade segura?
 Inf.: (est)... não... por que se **você** sair na rua com seu celular falando com alguém os donos vem tomar... e quando **você** não tem é capaz de tomar uma bala ou uns tapas na cara... uma conhecida.. ela tava sem celular eles levaram o cabelo dela
 (01ent.UFS-Itabaiana2018_desl. I_início_ray.fs.19)
- (44) D.S.: ((RISOS)) é pra falar a verdade e eu não tô falando a verdade não?
 D.M.: e onde **cê** aprendeu tudo isso? **cê** já trouxe de casa foi?
 (D.S.cdt D.M.sdt P M_M 01)

Assumimos a hipótese de que o princípio do paralelismo formal atua tanto na regra de referência determinada quanto na indeterminada, isto é, a precedência de uma forma pronominal favorece, em seguida, o uso desta mesma forma. Nas ocorrências das formas em estudo isolada ou primeira menção, pressupomos que haja o uso mais frequente da forma **você**, como já constataram os estudos prévios que possuem baixa ocorrência da variante **tu** (NASCIMENTO, 2011; SILVA, SUZIANE, 2017; GUIMARÃES, 2019; SILVA, S., 2019).

xii) Efeito gatilho

O efeito gatilho ocorre quando o locutor de uma conversação utiliza uma marca linguística e o interlocutor a utiliza na sequência, havendo, assim, um “engatilhamento” de usos

linguísticos (cf. OLIVEIRA, 2006; SANTOS; ARAUJO, FREITAG, 2011). Controlamos essa variável a partir da identificação de uso de forma pronominal (**tu**, **você** ou **cê**) na fala do locutor e a manutenção desta no turno de fala do interlocutor (presença de efeito gatilho, conforme exemplo (45)) ou da ausência de uso da forma pelo interlocutor (ausência de efeito gatilho, conforme exemplo (46)).

(45) D.S.: **cê** consegue ler?

D.M.: **cê** só escuta o zum zum zum mas você não escuta ali que a menina tá sem namorado não... você só escuta o zum zum zum

(D.S._{cdt} D.M._{sdt} P M_M 01)

(46) D.S.: **cê** não é filhinho de papai não?

D.M.: infelizmente não

(D.S._{cdt} D.M._{sdt} P M_M 01)

Temos a hipótese de que a última forma pronominal utilizada pelo locutor/falante engatilhe o uso subsequente desta no turno de fala do interlocutor, principalmente, quando as formas pronominais em uso pelo locutor (responsável por introduzir o tópico, realizar a pergunta) forem **tu** e **cê** e propiciarem a expressão da **referência determinada** na fala do interlocutor (falante II) (responsável por responder as perguntas nas entrevistas e interações conduzidas). Para análise, considerou-se o limite de até 100 palavras do início do turno do interlocutor (atual locutor), a fim de verificar se a primeira forma pronominal utilizada por este é a mesma utilizada pelo locutor do turno anterior.

5.1.5 Tratamento estatístico dos dados

A análise dos dados seguiu uma abordagem de cunho quantitativo, uma vez que esse tipo de método é necessário em investigações linguísticas “para que se possa compreender melhor a estrutura da língua, bem como sua função” (LABOV, 2008 [1972], p. 73). Todas as ocorrências das formas pronominais **tu**, **você** e **cê** na posição de sujeito foram codificadas de acordo com as variáveis descritas na subseção anterior. Feito isso, os 3205 dados codificados foram submetidos à plataforma R (CORE TEAM, 2021), em interface com o RStudio, para a realização das análises estatísticas. Neste, realizamos análises inferenciais univariadas para o teste de qui-quadrado de *Pearson*, o teste exato de *fisher* e o teste de V^2 Crammer, a fim de identificar o nível de significância das análises. O nível de significância considerado foi o valor alfa de $p < 0.05$, o que significa que os resultados com p-valor menor ou igual a 0.05 são

estatisticamente significativos. Esses testes foram feitos a partir das funções *chiq.test* e *fisher.test*. O teste de V^2 Crammer consiste em uma medida de associação baseada Qui-quadrado, cujo objetivo é verificar a força de associação entre as variáveis dependentes e independentes. Tal força varia entre 0 e 1: “o valor 1 indica a máxima relação entre as variáveis e 0 a ausência de relação” (CORMELATO *et al.*, 2020, p. 110-111). Assim, quanto mais o teste de V^2 Crammer se aproximar de 1 mais forte é a força de associação entre as variáveis e vice-versa.

Realizamos também uma análise multivariada das variáveis função referencial e tipo de amostra, com intuito de verificar simultaneamente os efeitos destas nos usos das variantes. Para tanto, fizemos uma árvore inferencial de regressão condicional por meio do pacote *partykit* (HOTHORN; HORNIK; ZEILEIS, 2015), com a função *crtree*. Informações mais detalhadas sobre esse tipo de análise são apresentadas em Freitag e Pinheiro (2020).

Posteriormente a essa etapa de manipulação dos dados, procedemos à interpretação dos resultados, a qual é apresentada na subseção a seguir.

5.2 O PADRÃO DE COMPORTAMENTO REFERENCIAL DAS FORMAS TU, VOCÊ E CÊ: RESULTADOS E DISCUSSÃO

Nesta subseção, descrevemos e discutimos os resultados estatísticos a respeito do uso variável dos pronomes **tu**, **você** e **cê**, a fim de identificar o padrão de comportamento referencial destes por meio da análise dos efeitos das variáveis independentes estabelecidas (e de suas respectivas hipóteses) já delineadas nas subseções anteriores. A seção está dividida em duas subseções. Na primeira, apresentamos os resultados de distribuição geral das três variantes estudadas e realizamos uma análise de cunho qualitativo para os dados da forma pronominal **tu**. Na subseção seguinte, debatemos os resultados referentes apenas as variantes **você** e **cê** e sua relação com as variáveis independentes.

5.2.1 Análise das variantes **tu**, **você** e **cê**

Analizamos as ocorrências das variantes **tu**, **você** e **cê** em duas amostras de fala de universitários da comunidade de prática UFS-Itabaiana/SE (16 entrevistas sociolinguísticas e 16 interações conduzidas). Consideramos apenas as ocorrências em contextos de função

sintática de sujeito, conforme descrito na seção antecedente. No conjunto de 3205 ocorrências das formas, foram identificadas 22 (0,7%) ocorrências de **tu**, 759 (24%) ocorrências de **cê** e 2424 (75,3%) ocorrências de **você**. Na tabela a seguir, apresentamos os resultados de distribuição geral dessas variantes em função das duas regras variáveis estudadas: expressão da 2ªPS (referência determinada) e expressão da indeterminação do sujeito (referência indeterminada).

Tabela 34 – Distribuição das variantes **tu**, **você** e **cê** na função sintática de sujeito e o tipo de referência

Tipo de referência	Tu	Você	Cê	Total
Determinada	22 1,3%	1214 74%	395 24%	1631
Indeterminada	0 0%	1210 77%	364 23%	1574
Total	22	2424	759	3205

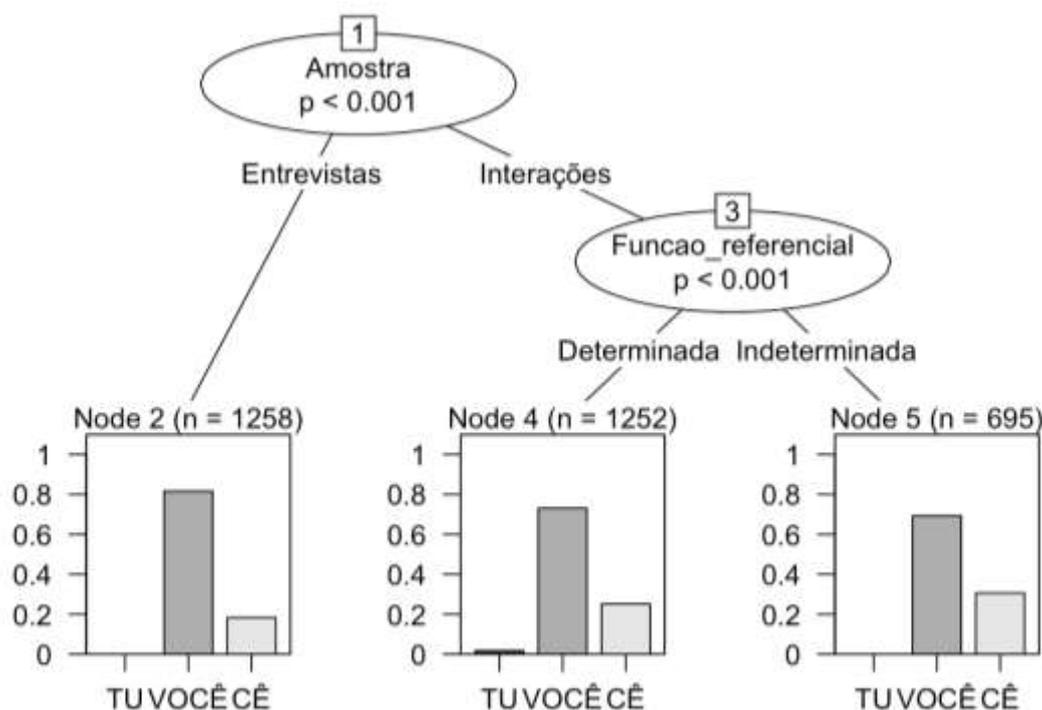
$\chi^2=22.266 \cdot df=2 \cdot Cramer's V=0.083 \cdot p=0.000$

Fonte: Elaboração própria.

Os resultados de distribuição geral das variantes **tu**, **você** e **cê** em função do tipo de referência apresentaram diferenças de frequências estatisticamente significativas ($\chi^2=22.266$, $p=0.000$), porém com associação fraca ($V=0.083$). Constatamos que a variante **você** foi a mais frequente, tanto na referência determinada (1214/1631 – 74%) quanto na referência indeterminada (1210/1574 – 77%). A segunda variante mais frequente foi o pronome **cê**, com taxas percentuais muito próximas para as duas regras referenciais analisadas (**determinada**: 395/1631 – 24%; **indeterminada**: 364/1574 – 23%). A variante **tu** foi a menos frequente na fala dos universitários da comunidade de prática analisada, e esta ocorreu apenas em contextos de referência determinada (22/1631 – 1,3%). Relacionando esses resultados a proposta de Scherre *et al.* (2015) de subsistemas de tratamento para o português brasileiro, temos um indício de que o falar itabaianense faz parte do subsistema **você/tu** (com **tu** de 1% a 90% sem concordância).

Para responder o questionamento sobre a interferência da variável tipo de coleta na distribuição dos resultados do fenômeno estudado, realizamos uma análise de árvore inferencial condicional. Esse tipo de análise possibilita identificar a associação entre a variável dependente e as variáveis independentes a partir do agrupamento destas em um diagrama arbóreo (FREITAG; PINHEIRO, 2020).

Figura 23 – Regressão condicional para as ocorrências das variantes quanto ao tipo de função referencial e ao tipo de amostra



Fonte: Elaboração própria.

O primeiro nó da árvore condicional (figura 23) é relativo ao tipo de amostra, o qual separa as ocorrências no nível entrevistas e no nível interações. Para o nível de entrevistas foram computadas 1258 ocorrências, sendo 1027 (82%) da variante **você**, 231 (18%) da variante **cê** e nenhuma da variante **tu**. Na sequência, ligada ao nível de interações, a função referencial passa a ser o fator segmentador que separa o padrão de comportamento das variantes em estudo nos níveis de referência determinada e indeterminada. Esta última apresentou um total de 695 ocorrências distribuídas do seguinte modo: 482 (69%) da forma **você**; 213 (21%) da forma **cê**; e nenhuma da forma **tu**. Já a referência determinada, em interações, apresentou 1252 ocorrências, das quais 915 (73,1%) foram da variante **você**, 315 (25,2%) da variante **cê** e 22 (1,7%) da variante **tu**. Observa-se que a variante **tu**, além de ocorrer apenas em contextos de referência determinada, também só foi constatada na amostra de interações conduzidas.

Os dados da variante **tu** ocorreram na fala de apenas dois estudantes que sempre foram moradores da cidade de Itabaiana/SE. O estudante D.S. (de 23 anos) utilizou 21 vezes a variante **tu** nas seguintes interações: quando era responsável por conduzir a interação (incluir e mudar de tópico discursivo) ao interagir com o estudante D.M. (de 24 anos), como exemplificado em (47) e (48), e com a estudante A.G. (de 25 anos), como exemplificado em (49); quando estava fazendo uma pergunta (sequência injuntiva); e quando estava em uma interação conduzida pelo estudante D.M., como exemplificado em (50). O estudante D.M., quando estava responsável

por conduzir o tópico, utilizou uma vez a variante **tu** ao interagir com o estudante D.S., como exemplificado em (51).

- (47) D.S.: *nas férias tu foi pra onde?*
 D.M.: na feira?
 D.S.: *nas férias à toa... na feira nós não tem lugar pra ir... é que nem*
 D.M.: eu fiquei em casa... eu não tenho pra onde ir
 (D.S.cdt D.M.sdt P M_M 01)
- (48) D.S.: *sua inimiga... mas olhe... na tua infância tu fazia o quê?*
 D.M.: o quê?
 D.S.: *na tua infância você fazia o quê? ()*
 D.M.: eu não tô escutando não meu fio... ((SUSSURRO)) eu estudava... pouco
 D.S.: *pouco?*
 D.M.: trabalhava muito... e brincava mais ou menos
 D.S.: *e desenho tu aprendeu quando?*
 D.M.: na escola com minha Professora... foi ela que me ensinou porque eu riscava as cadeira
 (D.S.cdt D.M.sdt P M_M 01)
- (49) D.S.: vamo lá... (hes) você conhece alguém velhaco?
 [...]
 D.S.: e você conhecia bem a pessoa?
 A.G.: conhecia
 D.S.: e ela fez um negócio desse?
 A.G.: fez
 D.S.: e ainda fala com ela?
 A.G.: falo... fala comigo
 D.S.: vixe e **tu** é uma tola enganada que gosta de ser enganada é?
 (D.S.cdt A.G.sdt P MF 02)
- (50) D.M.: *não teve foi na quinta da paralização... dos Professores dia onze... lembrei... teve... só foram umas cinquenta pessoas pra rua*
 D.S.: **tu** veio?
 D.M.: *hã? eu não eu tava em casa ruim com minha alergia matando*
 D.S.: também você é igual à doida da sala né? quando tem prova dá alergia
 D.M.: *não mas não tinha nem aula como é que*
 D.S.: é quando tem manifestação você tem alergia
 (D.M.cdt D.S.sdt P M_M 09)
- (51) D.M.: *então... vou falar com ela pra lhe descer uns cacete de novo pra ver se você lembra*
 D.S.: ô castigo
 D.M.: *e também quando tu bate teu irmão?*

D.S.: *hã?*

D.M.: *quando cê bate seu irmão?*

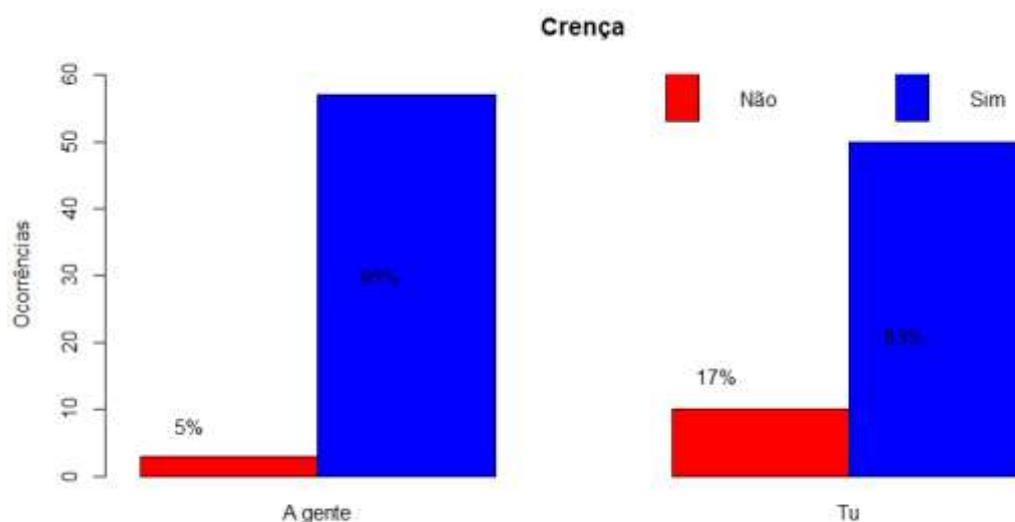
D.S.: *o quê?*

D.M.: *né pra lembrar ele não que tem que fazer as coisa certo?*

(D.M._{cdt} D.S._{sdt} P M_M 09)

Os exemplos de 47 a 51 demonstram que os contextos de usos da variante **tu** na amostra de interações possuíam um estilo mais informal e os tópicos discursivos foram sobre vivências pessoais (19 dados) e questões sociais (3 dados). A maioria dos dados da variante **tu** ocorreu na interação entre D.S. e D.M., principalmente no momento que D.S. era irônico/sarcástico ou “fazia brincadeiras” ao abordar o tópico. É importante ressaltar que os dois estudantes possuem uma relação de alta intimidade, visto que um frequenta a casa do outro, conhece ou possui informações a respeito da família, o que contribui para emergir formas pronominais da dimensão de solidariedade, conforme ressaltam Brown e Gilman (1960).

Gráfico 1 – Crença em relação ao próprio uso linguístico



Fonte: Araujo e Mendonça (2018, p. 135).

Embora as taxas de frequência da variante **tu** sejam baixas nas amostras analisadas, é possível que o uso desta seja mais frequente na comunidade e na região do Agreste sergipano. Tal ressalva baseia-se em Araujo e Mendonça (2018), as quais realizaram um estudo de atitude linguística a respeito das formas pronominais **a gente** e **tu** com 60 estudantes da comunidade de prática focalizada na presente pesquisa. Os resultados obtidos pelas autoras demonstraram que 83% dos estudantes afirmaram utilizar a variante **tu** (gráfico 1). Além disso, constataram que 21% estudantes julgaram esse pronome como errado. Já Araujo e Jesus (2018), ao realizar um estudo de atitude linguística com estudantes do ensino fundamental e médio, identificaram

que a maioria destes avalia a variante **tu** de forma negativa, associando-a ao adjetivo “feio”. Esses resultados evidenciam que é necessário expandir a amostra para outros tipos de relações hierárquicas, grau de intimidade, escolaridade, ambientes menos formais etc. para que se possa ter resultados mais acurados do padrão de comportamento linguístico quanto ao fenômeno em questão.

A partir dessa análise exploratória da variação **tu**, **você** e **cê** por meio da construção de uma árvore de inferência condicional, podemos concluir que há indícios de distinção no padrão de comportamento referencial do fenômeno. Os resultados apresentados indicam que o tipo de amostra interfere na distribuição das formas analisadas em cada função referencial: a variante **tu** foi utilizada apenas em contextos de expressão de referência determinada em interações conduzidas (na fala de homens, em sequências injuntivas, em tópicos sobre vivências pessoais, em situações interacionais entre íntimos), evidenciando, assim, que há a atuação de duas regras variáveis (de referência determinada e de referência indeterminada) nos usos das formas variáveis estudadas. Quando o tipo de amostra é de entrevistas, os usos das formas pronominais ocorreram indiferentemente do tipo de referência.

Em virtude do baixo número de ocorrências, excluímos os dados referentes à variante **tu** para análise estatística com as demais variáveis independentes. Nas subseções a seguir, discutimos os resultados nesse âmbito.

5.2.2 Análise das variantes **você** e **cê**

Apresentamos nesta subseção os resultados referentes à análise das variantes **você** e **cê** para cada função referencial (determinada e indeterminada). Para identificar os condicionamentos dos usos das variantes **você** e **cê** em cada uma das regras referenciais, controlamos as seguintes variáveis: i) grau de intimidade entre os interlocutores; ii) relações (as)simétricas quanto ao sexo/gênero dos interlocutores; iii) turno de fala; e iv) pergunta proposital; v) sexo/gênero do falante; vi) deslocamento social; vii) tipo de sequência discursiva; viii) tópico discursivo; e ix) tipo de discurso; x) paralelismo formal; e xi) efeito gatilho. Discutimos os resultados obtidos nas subseções a seguir.

5.2.2.1 Grau de intimidade entre os interlocutores

De acordo com a teoria de Poder e Solidariedade proposta por Brown e Gilman (1960) e com a teoria de Polidez Linguística proposta por Brown e Levinson (2011 [1987]), o grau de intimidade entre os interlocutores de uma interação interfere no padrão de comportamento linguístico destes. Nesse sentido, situações interacionais de maior grau intimidade entre os interlocutores (relações de laço forte (parentes, amigos, namorado(a), colega de trabalho/estudo etc.) e com interação diariamente/frequentemente) favorecem o uso das formas pronominais de 2ªPS relacionadas à dimensão da solidariedade, por exemplo, o uso do pronome **cê**. Em situações interações de menor grau de intimidade entre os interlocutores (relações de laço fraco (desconhecidos) e com interação raramente/ausente) favorecem o uso das formas pronominais da dimensão do poder ou da neutralidade, tal como o uso do pronome **você**.

Para a regra variável de referência determinada, a nossa hipótese era de que, nas situações interacionais entre interlocutores íntimos, haveria o favorecimento do uso da forma **cê** (cf. NASCIMENTO, 2011) e nas situações interacionais entre interlocutores não íntimos, haveria o favorecimento da forma **você** (cf. MOTA, 2008; MARTINS, 2010; GUIMARÃES, 2014). Quanto à regra variável de referência indeterminada, a nossa hipótese era de haveria o favorecimento da forma pronominal **você** independentemente do grau de intimidade entre os interlocutores por possivelmente essa variante ser o pronome de uso de livre trânsito em Sergipe, pertencendo a dimensão da neutralidade. Apresentamos na sequência os resultados obtidos.

Tabela 35 - Uso de **você** e **cê** em função do tipo de referência e da variável grau de intimidade entre os interlocutores

Grau de intimidade entre os interlocutores	Referência determinada			Referência indeterminada		
	VOCÊ	CÊ	T	VOCÊ	CÊ	T
Íntimo	925 74%	315 26%	1240	482 69%	213 31%	695
Não íntimo	299 79%	80 21%	379	728 83%	151 17%	879
Total (T)	1214	395	1609	1210	364	1574
	$\chi^2=2.931 \cdot df=1 \cdot \text{Cramer's}$ $V=0.044 \cdot p=0.087$			$\chi^2=38.851 \cdot df=1 \cdot \text{Cramer's}$ $V=0.159 \cdot p=0.000$		

Fonte: Elaboração própria.

Os resultados concernentes à referência determinada demonstraram que a forma **você** foi mais frequente, tanto em situações interlocutivas entre *íntimos* (925/1204 – 74%) quanto entre *não íntimos* (299/379 – 79%), refutando a nossa hipótese inicial. Tais resultados não foram

estatisticamente significativos. Esses resultados não corroboram os obtidos por Nascimento (2011) que constatou, no falar de São Paulo/SP, o favorecimento da forma **cê** com referência determinada na fala dos informantes (92/162 – 57%) e dos documentadores (354/595 – 60%) em situações interlocutivas entre *mais próximos*. Entretanto se considerarmos o número de ocorrências da variante **cê** em termos absolutos, constatamos que esta é mais frequente entre íntimos para a regra de referência determinada.

Nos contextos de referência indeterminada, obtivemos um resultado semelhante ao obtido para referência determinada, tendo em vista que a variante **você** foi a forma preterida pelos informantes nos dois níveis de grau de intimidade entre os interlocutores controlados e o efeito dessa variável foi estatisticamente significativo ($\chi^2=38.851$, $p=0.000$), porém com associação fraca ($V=0.159$). Tais resultados confirmam a nossa hipótese de que essa variante apresenta indícios de pertencer a dimensão da neutralidade por possuir baixa interferência no padrão de usos das variantes em situação de maior ou menor grau de intimidade entre os interlocutores. Esses resultados assemelham-se aos constatados por Guimarães (2014).

5.2.2.2 Relações (as)simétricas quanto ao sexo/gênero dos interlocutores

A interrelação de interlocutores quanto ao sexo/gênero destes permite identificar a atuação da (as)simetria nas relações pessoais. O controle dessa variável possibilita verificar os efeitos da dimensão da solidariedade com base na simetria (interações entre informantes do mesmo sexo/gênero), da dimensão do poder (interações entre informantes de sexo/gênero diferentes); e conseqüentemente observar se há um comportamento linguístico acomodativo de acordo com a relação estabelecida.

Para essa variável, controlamos os seguintes níveis: i) masculino-masculino e ii) feminino-feminino, configurando relações de simetria; iii) masculino-feminino e iv) feminino-masculino, configurando relações de assimetria. Assumimos a hipótese de que em relações simétricas a forma **cê** seja frequente, sobretudo, para a regra variável de referência determinada. Nas situações interlocutivas de assimetria, a variante **você** seria a preterida pelos informantes da comunidade de prática analisada para os dois tipos de referência. Na seqüência, dispomos os resultados obtidos.

Tabela 36 - Uso de **você** e **cê** em função do tipo de referência e da variável relações (as)simétricas quanto ao sexo/gênero dos interlocutores

Relações (as)simétricas quanto ao sexo/gênero dos interlocutores	Referência determinada			Referência indeterminada		
	VOCÊ	CÊ	T	VOCÊ	CÊ	T
Feminino-Feminino	380 92%	34 8%	414	335 89%	43 11%	378
Masculino-Masculino	220 59%	151 41%	371	92 48%	98 52%	190
Feminino-Masculino	446 79%	121 21%	567	683 79%	182 21%	865
Masculino-Feminino	168 65%	89 35%	257	100 71%	41 29%	141
Total	1214	395	1609	1210	364	1574
	$\chi^2=129.156 \cdot df=3 \cdot \text{Cramer's } V=0.283 \cdot p=0.000$			$\chi^2=120.806 \cdot df=3 \cdot \text{Cramer's } V=0.277 \cdot p=0.000$		

Fonte: Elaboração própria.

Os resultados obtidos demonstraram que a variante **você** é mais frequente para todos os níveis da variável relações (as)simétricas quanto ao sexo/gênero dos interlocutores para as duas regras variáveis estudadas; exceto para o nível masculino-masculino em contexto de referência indeterminada, no qual não houve favorecimento de nenhuma das variantes. Este último resultado é semelhante ao obtido por Nascimento (2011) e por Silva, S. (2019), no entanto, as autoras apresentaram os resultados dos dois tipos de referência conjuntamente. Tais resultados apresentaram significância estatística ($\chi^2=129.156$, $p=0.000$) e com força associativa fraca ($V=0.283$).

É importante ressaltar que, nos níveis da variável em que há a presença de um interlocutor do sexo/gênero masculino com o domínio do tópico, as taxas percentuais da variante **cê** são mais altas do que nos níveis em que isso não ocorre. Esse resultado indicia que houve um comportamento linguístico distinto, mesmo que sutil, conforme o tipo de relação estabelecida. Em termos gerais, temos a confirmação parcial da hipótese aventada, haja vista que os informantes tendem a utilizar a variante **você** tanto nas situações interacionais de simetria quanto assimetria.

5.2.2.3 Turno de fala

A variável turno de fala foi controlada apenas na amostra de dados de interações conduzidas (os próprios participantes da pesquisa conduzem a interação) a partir de dois níveis:

falante I e falante II. No processo de constituição dessa amostra, o falante I possuía o papel de incluir e mudar o tópico (por meio da realização de perguntas para o falante II) discursivo na situação interlocutiva. Quanto ao falante II, apesar de poder também incluir/mudar o tópico, tinha como principal papel desenvolver o tópico discursivo proposto pelo falante I. O controle dessa variável nos permitiu identificar a interferência da dimensão do poder, discutida por Brown e Gilman (1960) Brown Levinson (2011 [1987]), nos usos das variantes **você** e **cê** na expressão da referência determinada e indeterminada. Nesse âmbito, temos uma situação interlocutiva assimétrica em que o falante I possuía mais poder do que o falante II.

Partimos da hipótese de que, em contextos de referência determinada, a variante **cê** seria a mais frequente no turno de fala do falante I, tendo em vista que este está em uma posição hierárquica mais alta por ser o dominador do tópico (ao fazer uma pergunta para o interlocutor, geralmente, utiliza-se pronomes de 2ªPS). Para a expressão da referência indeterminada, pressupomos que a forma **você** é mais frequente, tanto no turno do falante I quanto do falante II. Na tabela a seguir, dispomos os resultados obtidos.

Tabela 37 - Uso de **você** e **cê** em função do tipo de referência e da variável turno de fala

Turno de fala	Referência determinada			Referência indeterminada		
	VOCÊ	CÊ	T	VOCÊ	CÊ	T
Falante I	769 77%	231 23%	1000	171 74%	61 26%	232
Falante II	146 63,5%	84 36,5%	230	311 67%	152 33%	463
Total	915	315	1230	482	213	695
			$\chi^2=16.984 \cdot df=1 \cdot \text{Cramer's } V=0.120 \cdot p=0.000$	$\chi^2=2.807 \cdot df=1 \cdot \text{Cramer's } V=0.067 \cdot p=0.094$		

Fonte: Elaboração própria.

Quando a referência é determinada, os resultados demonstraram que os falantes I utilizam mais a forma **você** (769/1000 – 77%), confirmando parcialmente a hipótese aventada. No que concerne ao falante II, os resultados evidenciaram que este utilizou com mais frequência a forma pronominal **você** (146/230 – 63,5%), o que confirma a nossa hipótese. A análise univariada evidenciou que a relação entre as variantes e a variável turno de fala na expressão da referência determinada foi estatisticamente significativa ($\chi^2=16.984$, $p=0.000$), mas com associação fraca ($V=0.120$).

Para a referência indeterminada, os resultados obtidos evidenciaram que a variante **você** é favorecida, tanto no turno de fala do falante I (171/232 – 74%) quanto do falante II (311/463 – 67%), confirmando a nossa hipótese. A análise univariada apontou que a relação entre as

variantes focalizadas e a variável turno de fala não apresenta significância estatística na expressão da referência indeterminada.

5.2.2.4 Pergunta proposital

A partir do controle da variável pergunta proposital, objetivamos verificar a interferência desta no falar dos entrevistados. Essa variável foi controlada apenas na amostra de dados de entrevistas sociolinguísticas e teve como objetivo propiciar o uso das variantes em questão com referência determinada haja vista que esse tipo de amostra favorece a ocorrência da referência indeterminada. A hipótese aventada era de que, quando houvesse a presença de pergunta proposital, as formas **você** e **cê**⁶⁹ com valor referencial determinado seriam mais frequentes nas respostas dos entrevistados, principalmente, a última forma referida. Os resultados são apresentados na tabela abaixo.

Tabela 38 - Uso de **você** e **cê** em função da referência e da variável pergunta proposital

Pergunta proposital	Referência determinada			Referência indeterminada		
	VOCÊ	CÊ	T	VOCÊ	CÊ	T
Presença	237 81%	56 19%	293	195 78,6%	53 21,4%	248
Ausência	62 72%	24 28%	86	533 84,5%	98 15,5%	631
Total	299	80	379	728	151	879
	$\chi^2=2.582 \cdot df=1 \cdot \text{Cramer's}$ $V=0.090 \cdot p=0.108$			$\chi^2=3.867 \cdot df=1 \cdot \text{Cramer's}$ $V=0.070 \cdot p=0.049$		

Fonte: Elaboração própria.

No nível presença de pergunta proposital, a expressão da referência determinada foi mais frequente (293 ocorrências) do que a referência indeterminada (248 ocorrências) (cf. tabela 38). Embora a pergunta direcionasse o falante a referir-se, por meio do uso das formas **você** ou **cê**, ao seu interlocutor dêitico (neste caso, o entrevistador), ele as utilizava como forma de indeterminação do sujeito. Observa-se que esse uso ocorre no exemplo (52) no qual o falante utiliza a forma para se referir às pessoas/aos estudantes que desejam obter uma bolsa de Iniciação Científica.

(52) DOC.: se eu quiser ter uma bolsa cientí- se eu quiser ter uma bolsa científica **cê** saberia me dizer como eu posso conseguir uma?

⁶⁹ A hipótese englobava também o uso da variante **tu**, entretanto, como já discutimos, essa forma pronominal só foi constatada na amostra de interações conduzidas que selecionamos.

INF.: geralmente é por editais né?... que as pessoas se inscrevem se eu não me engano é por editais ou **você** procura algum professor... e ele te orienta... como é que **você** consegue entrar eu num sei bem a bolsa né? que a bolsa é quando tem remuneração... mas quando **você** quer entrar em algum projeto de pesquisa que é o PIBIC né? que chama?... geralmente **cê** procura um professor ou eles mesmo chega na sala e comenta olhe nós temos tantas vagas pra PIBIC

(20ent.UFS-Itabaiana2018_desl. II_início_ian.fs.22)

Constamos alta taxa percentual da forma **você** nesse nível para a expressão das duas referências: 237/203 - 81% em contextos de expressão de referência determinada e 195/248 – 78,6% em contextos de expressão de referência indeterminada. Os resultados obtidos demonstraram que a presença de pergunta proposital não favoreceu um uso mais frequente da variante **cê** na fala dos entrevistados, visto que as taxas percentuais dessa variante no nível presença de pergunta proposital (referência determinada = 56/293 – 19% de **cê**; referência indeterminada = 53/195 – 21,4% de **cê**) são semelhantes às taxas obtidas no nível ausência de pergunta proposital (referência determinada = 56/293 – 19% de **cê**; referência indeterminada = 53/195 – 21,4% de **cê**) para os dois tipos de referência (cf. tabela 38). Os resultados referentes à variável pergunta proposital não foram significativos estatisticamente para o uso das variantes com expressão indeterminada ($\chi^2=2.582$, $p=0.108$). Portanto, não confirmamos a nossa hipótese para esta variável.

5.2.2.5 Sexo/gênero do falante

A variável sexo/gênero do falante foi controlada com o objetivo de verificar a diferenciação do comportamento linguístico, quanto aos usos das formas pronominais **você** e **cê**, entre mulheres e homens para a expressão da 2ªPS e para a expressão da indeterminação do sujeito. Tomando como base os estudos consultados que focalizaram a variação das formas **você** e **cê** (CALMON, 2010; ANDRADE, 2015; NOGUEIRA, 2013; SILVA, Suziane, 2017; SILVA, S., 2019), a nossa hipótese inicial era de que as mulheres e os homens utilizassem mais a variante **você**, tanto para a referência determinada quanto para indeterminada. Além disso, pressupomos que a variante **cê** seja mais frequente em contextos de referência determinada e na fala das mulheres, assim como constatou Nascimento (2011), mas esta não suplantaria o uso da variante **você**. Na tabela a seguir, apresentamos os resultados obtidos com o controle dessa variável.

Tabela 39 – Uso de **você** e **cê** em função do tipo de referência e da variável sexo/gênero do falante

Sexo/gênero do falante	Referência determinada			Referência indeterminada		
	VOCEÊ	CÊ	T	VOCEÊ	CÊ	T
Feminino	651 89%	84 11%	735	446 87%	69 13%	515
Masculino	563 65%	311 36%	874	764 72%	295 28%	1059
Total	1214	395	1609	1210	364	1574
	$\chi^2=124.462 \cdot df=1 \cdot \text{Cramer's}$ $V=0.280 \cdot p=0.000$			$\chi^2=39.935 \cdot df=1 \cdot \text{Cramer's}$ $V=0.161 \cdot p=0.000$		

Fonte: Elaboração própria.

Os resultados evidenciaram que o uso da variante **você** na fala das mulheres e dos homens é favorecido (cf. tabela 39) nos dois tipos de referência pronominal, o que confirma a hipótese aventada. No entanto, observamos que os homens utilizaram mais a variante **cê** (referência determinada = 311/874 – 36%; referência indeterminada = 295/1059 – 28%), embora não tenha suplantado os usos da variante **você**. Esses resultados são estatisticamente significativos tanto para a referência determinada ($\chi^2=124.462$, $p=0.000$) quanto para a referência indeterminada ($\chi^2=39.935$, $p=0.000$), mas com força associativa fraca.

Tais resultados vão de encontro aos obtidos por Nascimento (2011), a qual analisou dados de fala de São Paulo/SP e constatou que as mulheres utilizam com maior frequência a variante **cê** para a referência determinada (principalmente para esta) e indeterminada, já os homens utilizam essa forma com mais frequência apenas para a referência determinada. Observa-se, portanto, diferenças dialetais visto que os resultados concernentes à análise de dados de fala provenientes da região Nordeste (nosso estudo) se diferem daqueles referentes à região Sudeste (estudo de Nascimento (2011)).

5.2.2.6 Deslocamento

A variável deslocamento possibilita analisar o comportamento linguístico dos estudantes à luz da teoria da acomodação linguística. Essa variável foi controlada por meio de dois níveis: deslocamento I e deslocamento II. Fazem parte do primeiro os estudantes da UFS-Campus-ITA que são moradores da cidade de Itabaiana/SE, enquanto do segundo tipo, fazem parte os estudantes que moram em outra cidade da região Agreste Sergipano e que se deslocam diariamente para UFS-Campus-ITA. A nossa hipótese, para essa variável, era de que os estudantes do *deslocamento II* teriam um comportamento linguístico convergente com aqueles

que fazem parte do deslocamento I. Desse modo, os estudantes dos dois deslocamentos utilizariam a variante **você**, tanto para regra de referência determinada quanto para a indeterminada, pelo fato dessa forma pronominal possuir avaliação positiva na comunidade acadêmica estudada (ARAÚJO; MENDONÇA, 2018), os informantes possuírem alto grau de escolaridade e o ambiente (a Universidade) de realização ser formal. Todos esses fatores contribuem para que haja convergência de uso. Apresentamos, na sequência, os resultados alcançados com o controle dessa variável.

Tabela 40 - Uso de **você** e **cê** em função do tipo de referência e da variável tipo de deslocamento

Deslocamento	Referência determinada			Referência indeterminada		
	VOCEÊ	CÊ	T	VOCEÊ	CÊ	T
Deslocamento I	439 68,6%	201 31,4%	640	654 79,6%	168 20%	822
Deslocamento II	775 80%	194 20%	969	556 74%	196 26%	752
Total	1214	395	1609	1210	364	1574
	$\chi^2=26.364 \cdot df=1 \cdot \text{Cramer's}$ $V=0.129 \cdot p=0.000$			$\chi^2=6.679 \cdot df=1 \cdot \text{Cramer's}$ $V=0.067 \cdot p=0.010$		

Fonte: Elaboração própria.

Os resultados evidenciaram que os estudantes do deslocamento I e do deslocamento II apresentam comportamento linguístico convergentes. Isso porque as taxas percentuais de uso da variante **você** e da variante **cê** para os dois níveis da variável foram semelhantes independentemente do tipo de referência. Esses resultados podem ser justificados pelo fato de todos estarem realizando um curso de nível superior e a realização da entrevista ter ocorrido em um ambiente de formalidade, o que faz emergir o uso da variante de estilo formal (atuação da dimensão do poder). Destarte, os estudantes podem ter convergido linguisticamente, em relação ao fenômeno aqui estudado, com o intuito de obter aceitação dos membros da comunidade acadêmica em questão e não sofrerem estigma quanto ao seu modo de falar.

5.2.2.7 Tipo de sequência discursiva

A variável tipo de sequência discursiva possibilita verificar o estilo de fala dos interlocutores (grau de monitoramento – mais ou menos formal), o qual pode influenciar a escolha de uma ou de outra forma pronominal em estudo. Os níveis controlados para essa variável foram: sequências discursivas de cunho narrativo (ex. (53)) e injuntivo (ex. (54)) estão relacionadas ao estilo de menor grau de monitoramento (informal) e as sequências discursivas

de cunho argumentativo (ex. (55)) e expositivo-explicativo (ex. (56)) estão relacionadas ao estilo de maior grau de monitoramento (formal).

(53) DOC.: e hoje o que eles falam sobre os seus estudos?

INF.: eu sempre ouço que ... eu tenho que estudar pra ter um emprego melhor ... que eu tenho que estudar pra ter futuro melhor pra não acabar como eles ... a minha mãe dona de casa que ela não trabalha ... e o meu pai ... descarregando carros caminhão ... aí ele fala “olha eu tô me acordando duas e meia da manhã ... **você** tem que estudar pra **você** não precisar acordar esse horário”... sempre dizendo pra eu conseguir algo melhor do que aquilo que ele tem

(01ent.UFS-Itabaiana2018_desl. I_início_ray.fs.19)

(54) DOC.: e como faço pra chegar nesses locais você ... saberia me dizer contaria o passo a passo?

INF.: olhe primeiro... primeiro você entra pela entradinha da UFS que é uma só tem a dos fundos mas é pra quem entra com carros e motos... (est)... **cê** vai pela entrada principal... aí **você** vira a direita... vai ver que tem um um caminhezinho aí vai até o final tem duas escadarias... duas não uma né?... é tem uma escadaria dessa escadaria **você** desce... segue direto aí onde **você** vai ver um aglomerado de pessoas e tem lá o nome DCE

(20ent.UFS-Itabaiana2018_desl. II_início_ian.fs.22)

(55) DOC.: se eu fosse morar aqui em Itabaiana que tipo de trabalho eu poderia fazer para ganhar meu próprio sustento?

INF.: dependendo da sua criatividade... **você** pode vender coisas tipo... se **você** for uma muito boa cozinheira... o que não falta é coisas pra você vend- tipo... ambientes pra **você** vender **você** for... criativa também... na feira tem bastante emprego... agora assim não é coisa leve é coisa bastante pesada que vai te desgastar... por que diz que Itabaianense eu conheci um rapaz um caminhoneiro que ele disse que se encantou... pela garra da mulher itabaianense que ele disse que a gente... vê ele viu aqui mulheres se acordando cedo na feira batalhando tudo e ele disse que na cidade dele num ver isso... é raro **você** ver uma mulher buscando seu próprio dinheiro

(01ent.UFS-Itabaiana2018_desl. I_início_ray.fs.19)

(56) INF.: os meus alunos me consideram uma tia boa... calminha... mas nem sempre isso vai dizer se **você** vai ser uma boa professora ou não... Carolina se **você** for calma se **você** for cuidadosa... claro que tem que ser né? com criança tem que ser... pra tratar com pessoas tem que ser... mas não consigo dizer um perfil assim (est)... pra ser professor

(31ent.UFS-ItabaiAna2018_desl. II_final_Ana.fs.22)

Com base nos resultados obtidos por Loregian-Penkall e Menon (2012) para o falar da cidade de Curitiba/PR, assumimos a hipótese de que as sequências discursivas argumentativas e expositiva-explicativas favoreceriam o uso da variante **você**, por ser um contexto mais formal, tanto para a expressão de referência determinada quanto de referência indeterminada. Já em contextos de sequências discursivas narrativas e injuntivas o uso da forma **cê** com referência determinada seria favorecido pelo fato destas sequências promoverem maior envolvimento do

informante, principalmente, em sequências narrativas, e assim, emergir um estilo mais informal. Na tabela a seguir, expomos os resultados constatados.

Tabela 41 - Uso de **você** e **cê** em função do tipo de referência e da variável tipo de sequência discursiva

Tipo de sequência discursiva	Referência determinada			Referência indeterminada		
	VOCÊ	CÊ	T	VOCÊ	CÊ	T
Argumentativa	55 66%	28 34%	83	611 84%	118 16%	729
Expositiva-explicativa	236 72%	91 28%	327	321 73%	<u>120</u> 27%	441
Injuntiva	817 78%	<u>237</u> 22%	1054	249 68%	116 32%	365
Narrativa	106 73%	39 27%	145	29 74%	10 26%	39
Total	1214	395	1609	1210	364	1574
	$\chi^2=8.534 \cdot df=3 \cdot Cramer's$ $V=0.073 \cdot p=0.036$			$\chi^2=39.404 \cdot df=3 \cdot Cramer's$ $V=0.158 \cdot p=0.000$		

Fonte: Elaboração própria.

Os resultados obtidos com o controle da variável tipo de sequência discursiva confirmaram, parcialmente, as nossas hipóteses iniciais. Constatamos que, para a regra de referência determinada, todos os tipos de sequência favoreceram o uso da variante **você**, sobretudo, a sequência injuntiva (817/1054 – 72%). Esse tipo de sequência foi também o que mais propiciou o uso da variante **cê**, em termos absolutos - 237 ocorrências. Para a regra de referência indeterminada, tivemos o favorecimento da variante **você** em todos os tipos de sequências, principalmente, quando a sequência discursiva era argumentativa (611/729 – 84%). Já em valores absolutos a forma **cê** foi favorecida nas sequências expositivo-explicativas – 120 ocorrências. Os resultados obtidos para o tipo de sequência narrativa se distanciaram dos obtidos por Loregian-Penkall e Menon (2012), uma vez que, em nosso estudo, o número de ocorrências nesse nível da variável foi bastante baixo. Uma possível justificativa para essa situação pode estar relacionada ao fato de a maioria dos tópicos abordados nas entrevistas e interações conduzidas não propiciava a narração. Por fim, é importante ressaltar que as diferenças de frequências entre as variáveis foram estatisticamente significativas para as duas regras referencias, entretanto com força associativa fraca.

5.2.2.8 Tópico de discurso

A variável tópico discursivo foi selecionada no presente estudo com o intuito de verificar se o assunto abordado durante a conversação dos universitários itabaianenses interfere nos usos das variantes **você** e **cê** na expressão da referência determinada e da indeterminada. Os assuntos abordados nas entrevistas e interações conduzidas foram categorizados em três níveis: educação (ex. (57)), questões sociais (ex. (58)) e vivências pessoais (ex. (59)).

(57) A.G.: *é e o principal... ponto assim que deveria ser essas manifestações era a questão a Educação uma Educação de realmente de qualidade que seja prioridade do governo né? então... um dos pontos que a população realmente tinha que... protestar e cobrar do governo era a questão da Educação porque eu acho que a Educação de qualidade é a base de tudo... não é verdade?*

D.S.: *é né? depende... depende... (hes) tem autores que discutem que a escola é pro é um espaço onde se constrói... (hes) pe- (hes) sujeitos... pensantes sujeitos reflexivos sobre suas ações... mas há autores que defendem que a escola é um é um é... é um é o puxadinho da fábrica... é o puxadinho da fábrica são lá onde é os é um espaço onde se há o treinamento da sociedade pra que se su- pra que se substitua o operariado... do do futuro... quer dizer... se **você** pensa por essa questão de que a a escola é é essa é esse caminho a- é esse espaço alienatório... da sociedade... a Educação não é o futuro de nada... a Educação apenas é o futuro... que vai perpetuar essa sociedade desigual... mas se **você** a... (hes) se **você** constrói uma Educação que faz com que o sujeito reflita sobre sua prática... aí sim a Educação seria algo que transformaria a sociedade... porque pro porque pro... (hes) reproduziria cidadãos... (hes) críticos... sobre a sobre a socie- sobre a sua realidade... e não apenas... (hes) reproduz reprodutores de de conceitos*

(05 A.G.cdt D.S.sdt P FM 05)

(58) D.S.: **você** acha **você** acha você como é que **você** vê essa questão da Bolsa Família?
A.G.: *mas é verdade... é uma visão como se ficasse pra sempre depender do que os governantes de alguma forma não de solucionar os problemas daquele... real... família brasileira que está precisando né? de realmente de condições que eles mesmos teriam que dar como **você** mesmo colocou... como a questão do... das manifestação né? que tá ocorrendo poderia ser até um dos pontos que poderia se manifestar poderia ser isso né? que realmente eles dessem e não que ele pudesse resolver com a questão do com medidas paliativas com a questão de Bolsa Família... o que **você** tem a dizer sobre a questão das manifestações porque assim... alguns vão pra lutar mesmo outros vão pra... bagunçar pra não dar nenhuma importância né?*

(05 A.G.cdt D.S.sdt P FM 05)

(59) D.S.: *mas é... tu brincava <<ni>> sítio mais quem? com seus irmãos era?*

D.M.: meus irmãos

D.S.: *antes de ficar de mal?*

D.M.: *é... com meus irmãos e meu primo*

(D.S.cdt D.M.sdt P MM 01)

Com base nos resultados de Lucca (2005) e Guimarães (2014), defendemos a hipótese de que os tópicos sobre educação e questões sociais favoreceriam o uso da variante **você** em contextos de referência determinada e indeterminada pelo fato destes tópicos, geralmente, propiciarem a utilização de um estilo de fala monitorada. No que concerne ao tópico vivências pessoais, pressupomos que este favoreceria o uso da variante **cê**, principalmente, para expressar a referência determinada, propiciar o uso de estilo de fala menos monitorada. Na sequência, expomos os resultados relativos a essa variável.

Tabela 42 - Uso de **você** e **cê** em função do tipo de referência e da variável tópico discursivo

Tópico discursivo	Referência determinada			Referência indeterminada		
	VOCE	CÊ	T	VOCE	CÊ	T
Educação	134 73%	49 27%	183	287 81%	67 19%	354
Questões sociais	459 80%	113 20%	572	550 76%	175 24%	725
Vivências pessoais	621 73%	233 27%	854	373 75%	122 25%	495
Total	1214	395	1609	1210	364	1574
	$\chi^2=11.034 \cdot df=2 \cdot \text{Cramer's}$ $V=0.083 \cdot p=0.004$			$\chi^2=4.573 \cdot df=2 \cdot \text{Cramer's}$ $V=0.054 \cdot p=0.102$		

Fonte: Elaboração própria.

Os resultados obtidos confirmaram a nossa hipótese inicial de que os tópicos sobre educação e questões sociais condicionariam o uso da forma **você** para as duas regras de referência (tabela 42). Tais resultados assemelha-se aos obtidos por Lucca (2005) e Guimarães (2014). No que se refere ao nível vivências pessoais, a nossa hipótese foi refutada pelo fato de este tópico não ter favorecido a utilização da variante **cê** (233/854 – 27%) para as duas regras referencias analisadas. Ressaltamos que, em termos absolutos, a variante **cê** foi mais frequente no nível em questão na regra de referência determinada (233 ocorrências). A análise univariada evidenciou que o efeito dessa variável foi significativo apenas para a regra de referência determinada e apresentou força de associação baixa.

5.2.2.9 Tipo de discurso

A variável tipo de discurso foi analisada a partir do nível discurso direto, elaborado no momento da situação interlocutiva como em (60), e do nível discurso indireto, citação de discurso próprio ou de terceiro, o qual foi elaborado em momento anterior e reproduzido na atual situação interlocutiva como em (61).

(60) INF.: é basicamente eu acho que o mesmo eu acho que o curso de letras e geografia são parecidos em alguns quesitos... no caso paciência acho que **você** já tem... trabalho em grupo também

(01ent.UFS-Itabaiana2018_desl. I_início_ray.fs.19)

(61) DOC.: e pra eu conseguir uma vaga no mercado de trabalho e me destacar na área de geografia?

INF.: aí você me pegou ainda não sei ((RISOS))... não sei te responder por que uma vez um antigo professor meu me perguntou... se eu conhecia alguém formado... aí eu falei que conhecia então vou recapitular a pergunta “ **você** conhece alguém muito bom no que faz formado desempregado?”... aí eu “não”... justamente aí a questão quando você é muito bom no que você faz independente do trabalho você vai se destacar e consequentemente não vai ficar desempregado

(01ent.UFS-Itabaiana2018_desl. I_início_ray.fs.19)

A hipótese inicial para essa variável era de que o discurso direto favoreceria o uso da variante **cê**, em contextos de referência determinada, assim como constatou Nascimento (2011) no falar de São Paulo/SP. Para a regra variável referência indeterminada, a forma **você** seria a preterida nos dois tipos de discurso (cf. NASCIMENTO, 2011). Os resultados obtidos estão dispostos na tabela a seguir.

Tabela 43 - Uso de **você** e **cê** em função do tipo de referência e da variável tipo de discurso

Tipo de discurso	Referência determinada			Referência indeterminada		
	VOCÊ	CÊ	T	VOCÊ	CÊ	T
Direto	1112 75%	368 25%	1480	1161 69%	354 31%	1515
Indireto	102 79%	27 21%	129	49 83%	10 17%	59
Total	1214	395	1609	1210	364	1574
	$\chi^2=0.791 \cdot df=1 \cdot Cramer's$ $V=0.025 \cdot p=0.374$			$\chi^2=0.979 \cdot df=1 \cdot Cramer's$ $V=0.029 \cdot p=0.322$		

Fonte: Elaboração própria.

As hipóteses iniciais foram confirmadas parcialmente para a variável tipo de discurso. Os resultados constatados evidenciaram que, tanto para a regra variável de referência determinada quanto para a indeterminada, o discurso direto favoreceu o uso da variante **você**. De modo análogo, o discurso indireto também propiciou o uso da variante **você** nos dois tipos de referência. Apesar disso, observamos que as taxas de frequência da variante **cê** são mais altas no discurso direto também as duas regras. As diferenças de frequências de usos das formas variantes não foram estatisticamente significativas para as duas regras variáveis estudadas. É

importante ressaltar que esses resultados confirmam parcialmente a hipótese aventada, uma vez que a variante **cê** não é a preterida no discurso direto.

5.2.2.10 Paralelismo formal

O paralelismo formal consiste na tendência a repetição de uma forma linguística em cadeia (SCHERRE, 1998). No presente estudo, o controle dessa variável foi feito por meio dos seguintes níveis: isolada (ex. (62)), primeira menção (ex. (63)); antecedida por **você** (ex. (64)); e antecedida por **cê** (65)).

(62) DOC.: em relação assim à violência já aconteceu alguma coisa que te deixou assustado?... lá em Frei Paulo

INF.: não lá em Frei Paulo nunca passei por nenhuma situação assim não de violência não a chegar ver não graças a Deus é coisa simples briga de vizinho essas coisa assim de família mas uma situação (que **cê** disse) de violência assim eu num nunca presenciei graças a Deus na minha cidade não

(34ent.UFS-Itabaiana2018_desl. II_final_eve.ms.30)

(63) INF.: **você** teria que... ter em mente uma coisa... se você realmente... a área que você tá... você não gosta de jeito nenhum não tem outra coisa porque... eu penso assim... sei que você já construiu a sua história até aqui... e... e tipo... você... sair da área que você já... de certa forma foi construindo

(13ent.UFS-Itabaiana2018_desl. I_final_she.fs.25)

(64) DOC.: você gosta de morar lá em Frei Paulo?

INF.: morar sim porque é uma cidade tranquila mas assim você perde você na- na questão econômica você fica muito limitado né **você** num tem uma visibilidade como é uma chance de crescimento maior entendeu de que uma capital tenha ou então uma cidade maior tenha uma cidade muito pacata uma cidade pequena com comércio pouco desenvolvido a indústria só tem uma indústria então é limite muito sua a- sua a chance de crescimento principalmente na minha área administrativa

(34ent.UFS-Itabaiana2018_desl. II_final_eve.ms.30)

(65) D.S.: ((*RISOS*)) *é pra falar a verdade e eu não tô falando a verdade não?*

D.M.: e onde cê aprendeu tudo isso? **cê** já trouxe de casa foi?

(D.S._{cdt} D.M._{sdt} P M_M 01)

A hipótese formulada para a presente variável era de que o emprego de uma variante pronominal levaria ao uso da mesma variante na sequência, tanto na regra de referência determinada quanto na indeterminada. Ademais, pressupomos que, no nível isolada e no nível primeira menção, a variante **você** seria a preterida, conforme constatado por outros estudos,

como, por exemplo, o de Nascimento (2011) e o de Silva, S. (2019). Na tabela 44, estão dispostos resultados obtidos.

Tabela 44 - Uso de **você** e **cê** em função do tipo de referência e da variável paralelismo formal

Paralelismo formal	Referência determinada			Referência indeterminada		
	VOCÊ	CÊ	T	VOCÊ	CÊ	T
Isolada	381 67%	189 33%	570	204 70,6%	85 29,4%	289
Primeira menção	276 81%	65 19%	341	271 75%	91 25%	362
Antecedida por você	501 86%	81 14%	582	642 86,5%	100 13,5%	742
Antecedida de cê	56 48%	60 52%	116	93 51,4%	88 48,6%	181
Total	1214	395	1609	1210	364	1574
	$\chi^2=110.113 \cdot df=3 \cdot \text{Cramer's}$ $V=0.262 \cdot p=0.000$			$\chi^2=112.272 \cdot df=3 \cdot \text{Cramer's}$ $V=0.267 \cdot p=0.000$		

Fonte: Elaboração própria.

Os resultados obtidos foram semelhantes para as duas regras de referência analisada (cf. tabela 44). Constatamos que, nos níveis isolada e primeira menção, a variante **você** foi de fato a mais frequente para ambas as referências, o que confirma a hipótese aventada. O nível antecedida por você favoreceu, majoritariamente, a ocorrência da variante **você**, tanto para a referência determinada (501/642 – 86%) como para a indeterminada (694/742 – 86,5%), o que ratifica a nossa hipótese inicial. O nível antecedida de cê não condicionou o uso da variante **cê** em situações de referência determinada, conforme esperávamos. Igualmente, não confirmamos a nossa hipótese para o nível antecedida de cê para a regra de referência indeterminada, visto que as taxas de uso das variantes foram semelhantes, destoando, portanto, dos resultados obtidos por Nascimento (2011). A análise univariada realizada para a variável paralelismo formal foi estatisticamente significativa, porém apresentou baixa força associativa.

5.2.2.11 Efeito gatilho

O controle da variável efeito gatilho objetivou identificar se a presença de uma das variantes em estudo na fala do locutor motiva, logo na sequência, o interlocutor a usar a mesma forma quando estiver com o turno de fala, promovendo, dessa forma, um “engatilhamento” de usos linguísticos (cf. OLIVEIRA, 2006; SANTOS; ARAUJO, FREITAG, 2011). Essa variável foi controlada da seguinte maneira: i) presença de efeito gatilho - quando havia o uso de uma

das formas pronominais em estudo na fala do locutor e esta era utilizada no turno de fala do interlocutor (primeira forma pronominal utilizada na sequência após a troca de turno de fala, como em (66); ii) ausência de efeito gatilho - quando havia o uso de uma das formas pronominais (a última utilizada) em estudo na fala do locutor e esta não era mantida no turno de fala do interlocutor, como em (67).

(66) DOC.: **você** considera Itabaiana uma cidade segura?

INF.: (est)... não... por que se **você** sair na rua com seu celular falando com alguém os donos vem tomar... e quando **você** não tem é capaz de tomar uma bala ou uns tapas na cara... uma conhecida... ela tava sem celular eles levaram o cabelo dela porque ela tava de cabelo solto ela tinha o cabelo muito grande eles passaram o canivete e levaram... aí cê se pergunta... que cidade nós estamos?... a segurança daqui é péssima
(01ent.UFS-Itabaiana2018_desl. I_início_ray.fs.19)

(67) DOC.: **cê** sabe- saberia me informar dizer o passo a passo

INF.: bom... ao chegar em Malhador quando **você** entra... eh... **você** pode... porque Malhador... eh... ah... as entradas a entrada de Malhador é como se fosse uma... uma meia lua (est)... porque **você** entra... **você** faz isso aqui... **você** desse no caso... assim (est)... então quando **você** entra... **você** vai... direto a uns... ah meu Deus... poderia botar um quilômetro... **você** vai chegar na praça (est)... né na praça de eventos... e aí **você**... poderia virar a direita... e **você**... ia a mais ou menos uns trezentos a quinhentos metros... uhum... **você** viraria à esquerda e já estaria nesse restaurante (est) ... de frente a ele no caso

(22ent.UFS-Itabaiana2018_desl. II_início_jos.ms.21)

Para a presente variável, tínhamos a hipótese de que a última forma pronominal usada pelo locutor funcionasse como um gatilho para que, na sequência de troca de turnos de fala, o interlocutor utilizasse a mesma forma pronominal. Isso ocorreria, sobretudo, nas situações em que a variante utilizada pelo locutor (responsável por introduzir o tópico, realizar a pergunta) fosse a forma pronominal **cê** e propiciasse a expressão da referência determinada na fala do interlocutor.

(68) DOC.: e o custo pra se manter no curso de peda- de pedagogia é caro?

INF.: não... não num é muito não comparado a outros cursos que tem não o que a gente tem mais de... gasto assim porque eles num tem as exigência né?... com material o negócio deles é **você** ter as apostilas fazer as leituras essas coisas... mas não exige que **você** tenha computador se **você** tiver é bom... num é como outros curso que é imprescindível **você** tá com o computador ali direto todo as aula (né?) pra prática

(20ent.UFS-Itabaiana2018_desl. II_início_ian.fs.22)

(69) DOC.: aí **você** falando desses prós e contras quais seria do curso de Geografia?

INF.: eh... vamos lá... com- eu falei nestante sobre a questão de **você** ter uma outra visão de mundo (est)... é **você** não ficar preso ali a um conhecimento... eh... vamos supor leigo(est)... eh... **você** passa a conhecer vári- várias coisas **você** passa a entender o mundo esse- esse o- o pró **você** passa a conhecer o mundo eh... eh... **você** pas- coisas que **você** ver... e **você** olhava assim tipo "ah por que que isso acontece" quando **você**... passa entender aquilo **você** fica bem

(22ent.UFS-Itabaiana2018_desl. II_início_jos.ms.21)

Das 1609 ocorrências das variantes **você** e **cê**, concernentes à regra de referência determinada, 1435 aconteceram em contextos nos quais não havia a possibilidade de ocorrer o efeito gatilho. De igual modo, ocorreu para a referência indeterminada: das 1574 ocorrências das variantes, 1326 ocorreram nesse tipo de contexto referido. Esse fato é decorrente de duas situações: i) não havia nenhuma das formas pronominais em estudo no turno do locutor/entrevistador (ex. (68); ou ii) a forma em análise no turno de fala subsequente não era a primeira da série (ex. (69)). Portanto, esses dados não foram considerados na análise dessa variável. Observemos os resultados expostos na tabela subsequente.

Tabela 45 - Uso de **você** e **cê** em função do tipo de referência e da variável efeito gatilho

Efeito gatilho	Referência determinada			Referência indeterminada		
	VOCÊ	CÊ	T	VOCÊ	CÊ	T
Presença	90 86%	15 14%	105	170 92%	15 8%	185
Ausência	37 53,6%	32 46,4%	69	28 44,4%	35 55,6%	63
Total	127	47	174	198	50	248
	$\chi^2=23.789 \cdot df=1 \cdot \text{Cramer's}$ $V=0.122 \cdot p=0.000$			$\chi^2=44.992 \cdot df=1 \cdot \text{Cramer's}$ $V=0.226 \cdot p=0.000$		

Fonte: Elaboração própria.

Constatamos que, quando há a presença de efeito gatilho, a variante **você** é a mais frequente, tanto para a referência determinada (90/105 – 86%) quanto para a referência indeterminada (170/185 – 92%). Quando houve ausência de efeito gatilho, os resultados evidenciaram que o pronome mais frequente foi a variante **você** para a regra de referência determinada (37/69 – 53,6%). Em contextos de referência indeterminada, a ausência de efeito gatilho não favoreceu o uso de nenhuma das variantes. Esses resultados não confirmaram a nossa hipótese inicial de que a variante **cê** seria favorecida em contextos de efeito gatilho. Os resultados obtidos possuem diferenças estatisticamente significativas para as duas regras referenciais analisadas, mas apresentaram associação baixa.

Em síntese, os resultados referentes à variação entre as formas **você** e **cê** evidenciaram que não há um padrão comportamental referencial distinto nas amostras de fala analisadas. Na seção a seguir, apresentamos as considerações finais a respeito dos resultados constatados.

6 CONSIDERAÇÕES FINAIS

Esta pesquisa teve como objetivo identificar o padrão de comportamento das variantes **tu**, **você** e **cê** com referência determinada e indeterminada no falar de universitários da UFS-Campus/ITA. Para tanto, compuseram o nosso *corpus* de análise 16 entrevistas sociolinguísticas (amostra Deslocamentos-UFS/ITA (2018)) e 16 interações conduzidas (amostra Rede Social de Informantes Universitários de Itabaiana/SE (2013)). A pesquisa foi desenvolvida à luz da Sociolinguística Variacionista (WEINREICH; LABOV; HERZOG, 2006 [1968]; LABOV, 2008 [1972]) em interface com a Pragmática (BROWN; GILMAN, 1980; BROWN; LEVINSON, 2011 [1987]) e com a Teoria da Acomodação (BELL, 1984, 2001; GILES; COUPLAND; COUPLAND, 1991).

Cada uma das regras referenciais em estudo (determinada e indeterminada) foi relacionada às seguintes variáveis independentes que podem condicionar os usos das variantes **tu**, **você** e **cê**: i) grau de intimidade entre os interlocutores; ii) relações (as)simétricas quanto ao sexo/gênero dos interlocutores; iii) turno de fala; iv) pergunta proposital; v) sexo/gênero do falante; vi) deslocamento social; vii) tipo de amostra; viii) tipo de sequência discursiva; ix) tópico discursivo; x) tipo de discurso; xi) paralelismo formal e xii) efeito gatilho.

Os dados foram codificados e tratados estatisticamente no R (CORE TEAM, 2021) em interface com o *Rstudio*. Constatamos 3205 ocorrências, das quais 22 foram de **tu**, 2424 foram de **você** e 759 de **cê**. Nota-se, portanto, que a variante **tu** é pouco frequente na comunidade prática UFS-Itabaiana/SE. Tal fato evidencia que a comunidade em questão faz parte, dentre os subsistemas propostos por Scherre *et al.* (2015), do subsistema **você/tu** (com **tu** de 1% a 90% sem concordância), o que nos dá indícios de que Itabaiana/SE e cidades circunvizinhas fazem também parte deste.

Ademais, os resultados indiciam que os usos das formas **tu**, **você** e **cê** na comunidade de prática em estudo estão relacionados à escolha de função referencial: **tu** ocorre apenas para expressar referência determinada, **você** e **cê** são utilizadas para expressar as duas referências (determinada e indeterminada). Os resultados também evidenciaram que o tipo de amostra interfere na distribuição das formas, pois os dados de **tu** ocorreram somente em interações conduzidas. Na análise qualitativa dos dados concernentes ao pronome **tu**, constatamos que essa variante ocorreu majoritariamente em contextos de: sequências injuntivas (perguntas), tópicos sobre vivências pessoais, interlocutores íntimos, relações simétricas no nível masculino-masculino, turno do falante I, falante masculino, isolada, discurso direto,

deslocamento (mora na cidade de Itabaiana/SE). Portanto, para o nosso questionamento “O padrão de distribuição dessas variantes em cada tipo de referência é resultado do tipo de coleta dos dados de fala?”, os resultados confirmaram a hipótese inicial de que o uso da variante **tu** é condicionado ao tipo de coleta dos dados de fala.

Em virtude do baixo número de ocorrências, os dados de **tu** foram excluídos das análises estatísticas. Para a regra variável de referência determinada, apenas as variáveis tipo de discurso, grau de intimidade e pergunta proposital não foram estatisticamente significativas. A variante **você** foi favorecida em todas as variáveis, com exceção do nível antecedido de **cê** pertencente a variável paralelismo formal em que não houve favorecimento das variantes. Já para a regra variável de referência indeterminada somente as variáveis independentes tópico discursivo, turno de fala e tipo de discurso não foram estatisticamente significativas. Novamente a variante **você** foi a preterida pelos estudantes, mas nos seguintes níveis nenhuma das formas variantes foi favorecida: (as)simetria entre os interlocutores quanto ao sexo/gênero - nível masculino-masculino, efeito gatilho – nível ausente).

É importante ressaltar que o favorecimento da variante **você** ocorreu, sobretudo, pelo fato de os informantes da pesquisa possuírem alto grau de escolaridade (tendo em vista que são universitários) e o ambiente de coleta de entrevistas e interações ter sido a universidade, propiciando, assim, uma situação interlocutiva mais formal. Consequentemente, essa situação fez emergir o uso da variante que possui prestígio na comunidade que fora lócus da pesquisa – a forma **você**. Em termos gerais, para o questionamento “Qual das variantes em questão terá uma proporção distribucional maior em cada função referencial?”, os resultados evidenciaram que a hipótese de que haveria distinção quanto à distribuição das variantes e aos efeitos das variáveis controladas não foi confirmada.

O comportamento referencial das formas **tu**, **você** e **cê** não ficou tão evidente pelo fato de o *corpus* analisado, de modo geral, não favorecer expressivamente a atuação da dimensão da solidariedade. Assim, as constatações da presente tese não são conclusivas, isto é, não podemos afirmar que os usos das variantes **você** e **cê** não estão relacionados à distinção do tipo de função referencial, haja vista todo resultado, a priori, ser indício, porque cada ato de fala é um evento circunscrito a um momento.

Faz-se necessário que futuras pesquisas ampliem a estratificação das amostras de dados para outros níveis de escolaridade, faixas etárias, grau de intimidade, relações assimétricas etc., o que propiciará uma análise mais acurada dos efeitos das variáveis no fenômeno em questão. É preciso, também, que haja a interrelação entre vários fatores para criar uma situação interlocutiva que propicie o uso de **tu** em localidades em que essa forma não é a de livre trânsito

na comunidade. Além disso, realizar uma pesquisa que considere o traço semântico referencial dos pronomes de 2ªPS em uma perspectiva de gradação e que verifique os efeitos de cada variável independente em cada grau, é de suma importância.

REFERÊNCIAS

- ALVES, C. C. B. **O uso do tu e do você no português falado no Maranhão.** 2010. Dissertação (Mestrado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade Federal do Ceará, Fortaleza, 2010.
- ALVES, C. C. B. **Pronomes de segunda pessoa no espaço maranhense.** 2015. Dissertação (Mestrado em Linguística) – Programa de Pós-Graduação em Linguística Universidade de Brasília, Brasília, 2015.
- ANDRADE, A. L. V. S. de. **A variação você, cê e ocê no português brasileiro falado.** 2004. Dissertação (Mestrado em Linguística) Universidade de Brasília, Brasília, 2004
- ANDRADE, C. Q. **A fala brasiliense origem e expansão do uso do pronome tu.** Tese (Doutorado em Linguística) - Programa de Pós-Graduação em Linguística, Universidade de Brasília, Brasília, 2015.
- ANDRADE, C. Q. **Tu e mais quantos?** A segunda pessoa na fala brasiliense. 2010. Dissertação (Mestrado em Linguística) – Universidade de Brasília, Brasília, 2010.
- ARAUJO, A. S. **“Você me faria um favor?”** O futuro do pretérito e a expressão de polidez. 2014. Dissertação (Mestrado em Letras) – Programa de Pós-Graduação em Letras, Universidade Federal de Sergipe, São Cristóvão, 2014.
- ARAUJO, A. S. *et al.* **O controle da mobilidade na constituição de amostras sociolinguísticas.** *In:* ABRALIN 50, Maceió (Apresentação de Trabalho/Comunicação), 2019.
- ARAUJO, A. S.; BARRETO, E. A.; FREITAG, R. M. **Banco de dados de falantes cultos de Itabaiana.** Aracaju, 2012.
- ARAUJO, A. S.; BORGES, D. K. V. Variação no uso de pronomes-objeto de segunda pessoa na fala de estudantes Itabaianenses. **Paraguacu**, v. 1, n. 1, p. 146-167, 2021.
- ARAUJO, A. S.; JESUS, E. A. B. Sociolinguística e ensino: avaliação e atitude linguística no contexto escolar. **Interdisciplinar**, São Cristóvão/SE, v. 29, p. 87-107, 2018.
- ARAUJO, A. S.; MENDONÇA, J. J. Atitudes linguísticas de universitários em relação às formas pronominais a gente e tu. **Tabuleiro de letras**, Salvador/BA, v. 12, p. 128-144, 2018.
- ARAUJO, A. S.; MENDONÇA, Y. S.; BARRETO, V. S. O uso da segunda pessoa do singular nas redes sociais. **Revista Feira de Ciência & Cultura**, São Cristóvão/SE, v. 6, n. 11, p. 32-34, 2019.
- ARAUJO, A. S.; SANTOS, K. C. dos; FREITAG, R. M. Ko. Redes sociais, variação linguística e polidez: procedimentos de coleta de dados. *In:* FREITAG, R. M. Ko. (org.). **Metodologia de Coleta e Manipulação de Dados em Sociolinguística.** São Paulo: Editora Edgard Blücher, 2014. p. 99-116.

ASSUNÇÃO, J. da S.; ALMEIDA, N. da S. F. de. **A realização de tu e você na variante linguística de falantes feirenses**. 2011. Relatório de Pesquisa (Iniciação Científica) - Universidade Estadual de Feira de Santana, Feira de Santana, 2011.

BABILÔNIA, L.; MARTINS, S. A. A influência dos fatores sociais na alternância dos pronomes tu/você na fala manauara. **Guavira Letras**, Três Lagoas, v. 13, n. 1, ago./dez., p. 49-60, 2011.

BATTISTI, E. O estudo sociolinguístico da variação. *In*: CELSUL, **Anais** [...]. 2008. p. 1-13.

BECHARA, E. **Moderna Gramática Portuguesa**. 38. ed. Rio de Janeiro: Nova Fronteira, 2015.

BELL, A. Back in style: reworking audience design. *In*: ECKERT, P.; RICKFORD, J. R. (eds.). **Style and sociolinguistic variation**. Cambridge: Cambridge University Press, 2001. p.139-169.

BELL, A. Language style as audience design. **Language in Society**. Cambridge University Press, v. 13, n. 2, p. 145-204, 1984.

BENVENISTE, É. A natureza dos pronomes pessoais. *In*: BENVENISTE, É. **Problemas de Linguística Geral I**. Campinas: Pontes, [1956] 1995. p. 277-283.

BENVENISTE, É. Estrutura da língua e estrutura da sociedade. *In*: BENVENISTE, É. **Problemas de Linguística Geral II**. Campinas: Pontes [1989] 2006. p. 93-104.

BENVENISTE, É. O aparelho formal da enunciação. *In*: BENVENISTE, É. **Problemas de Linguística Geral II**. Campinas: Pontes [1989] 2006. p. 81-90.

BISPO, F. C. **A literatura como fonte de dados: um olhar sociolinguístico sobre a obra história da minha infância, de Gilberto Amado**. 2014 Dissertação (Mestrado em Letras) - Programa de Pós-graduação em Letras, Universidade Federal de Sergipe, São Cristóvão, 2014.

BLAKE, R.; JOSEY, M. The /ay/ diphthong in Martha's Vineyard community: what can we say 40 years after Labov? **Language in Society**, n. 32, v. 4, p. 451-485, 2003.

BORTONI-RICARDO, S. M. **Do campo para a cidade: estudo sociolinguístico de migração e redes sociais**. São Paulo: Parábola Editorial, 2011.

BORTONNI-RICARDO, S. M. **Educação em língua materna: a sociolinguística em sala de aula**. São Paulo: Parábola, 2004.

BRASIL. Lei nº 12.711, de 29 de agosto de 2012. **Dispõe sobre o ingresso nas Universidades Federais e nas instituições Federais de ensino técnico de nível médio e outras providências**. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato20112014/2012/Lei/L12711.htm. Acesso em: 08 fev. 2020.

BRITO, K. M. T.; CASTRO, K. P.; FURTADO, R. M. da S. C. A alternância no uso das formas de tratamento pronominais 'tu' e 'você' no português falado entre os estudantes do ensino médio no município de Cametá-PA. **Revista Falas Breves**, n. 8, maio, 2020.

BROWN, P.; LEVINSON, S. C. **Politeness: some universals in language usage**. Cambridge: Cambridge University Press, 2011[1987].

BROWN, R.; GILMAN, A. The pronouns of power and solidarity. *In: SEBEOK, T. A. (ed.). Style in language*. Massachusetts: MIT Press, 1960. p. 253-76.

BRUGMAN, H.; RUSSEL, A. Annotating Multimedia/Multi-modal resources with ELAN, *In: Proceedings of LREC 2004 - Fourth International Conference on Language Resources and Evaluation*. Lisboa, 2004.

CALMON, E. N. **Ponte da passagem: você e cê transitando na fala de Vitória (ES)**. 2010. Dissertação (Mestrado em Linguística) – Universidade Federal do Espírito Santo, Vitória, 2010.

CAMPBELL-KIBLER, K.; WALKER, A.; ELWARD, S.; CARMICHAEL, K. Apparent time and network effects on long-term cross-dialect accommodation among college students. *University of Pennsylvania Working Papers in Linguistics*, v. 20, n. 2, p. 21-29, 2014.

CARNEIRO; H. M. S. **As formas de tratamento tu/você no português falado ludovicense**. Tese (Doutorado em Linguística e Língua Portuguesa) - Programa de Pós-Graduação em Linguística e Língua Portuguesa, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Araraquara, 2011.

CASTILHO, A. T. **Nova gramática do português brasileiro**. São Paulo: Contexto, 2010.

COOK, M. Uma teoria de interpretação das formas de tratamento na língua portuguesa. *In: Hispania*, v. 80, n. 3, Set. 1997. p. 451-464.

CORMELATO, P. H. *et al.* Métodos estatísticos para desfechos qualitativos. *In: CAPP, E.; OTTO, H. N. (orgs.). Bioestatística Quantitativa Aplicada*. Porto Alegre: UFRGS, 2020, p.101-124.

CORRADELLO, E. de F. A. **Quem é você?** – análise de um pronome pessoal. 1997. Dissertação (Mestrado) - Instituto de Estudos da Linguagem (IEL – UNICAMP), Campinas, 1997.

CORRÊA, T. R. de A. **A variação na realização de /t/ e /d/ na comunidade de práticas da UFS: mobilidade e integração**. 2019. Dissertação (Mestrado em Letras) - Programa de Pós-graduação em Linguística, Universidade Federal de Sergipe, São Cristóvão, 2019.

COSTA, R. M. da S. **A alternância das formas pronominais tu, você e o(a) senhor(a) na função de sujeito no português falado em Cameté PA**. 2016. Dissertação (Mestrado em Linguística) - Programa de Pós-Graduação em Linguística, Universidade Federal do Ceará, Fortaleza, 2016.

DIAS, E. P. **O uso do tu no português brasileiro falado**. Brasília, 2007. Dissertação (Mestrado em Linguística) – Programa de Pós-Graduação em Linguística. Universidade de Brasília, Brasília, 2007.

DIVINO, L. S. do A. **Como trato o meu receptor?** A propósito do uso de tu/você em Santo Antônio de Jesus/BA. 2008. Dissertação (Mestrado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade Federal da Bahia, Salvador, 2008.

ECKERT, P. Communities of Practice. *In*: BROWN, K. (ed.). **Encyclopedia of Language and Linguistics**, 2. ed., Elsevier: Amsterdam, 2006. p. 683-685.

ECKERT, P. **Linguistic variation a social practice**. Oxford: Blackwell, 2000.

ECKERT, P. Style and social meaning. *In*: ECKERT, P.; RICKFORD, J. R. (eds.). **Style and Sociolinguistic Variation**. Cambridge: Cambridge University Press, 2001. p. 119-126.

ECKERT, P.; MCCONNELL-GINET, S. Comunidades de práticas: lugar onde co-habitam linguagem, gênero e poder (1992). *In*: OSTERMANN, A. C; FONTANA, B. F. **Linguagem. Gênero. Sexualidade**. São Paulo: Parábola Editorial, 2010. p. 93-108.

FARACO, C. A. O tratamento você em português: uma abordagem histórica (1996). **Labor Histórico**, Rio de Janeiro, v. 3, n. 2, p. 1-134, jul./dez. 2017.

FRANCESCHNI, L. **Variação pronominal nós/a gente e tu/você em Concórdia – SC**. 2011. Tese (Doutorado em Letras) - Programa de Pós-Graduação em Letras da Universidade Federal do Paraná, Curitiba, 2011.

FREITAG, R. M. K. Falares Sergipanos. *In*: ATAÍDE, C. *et al.* **Gelne 40 anos: experiências teóricas e práticas nas pesquisas em Linguística e Literatura**. São Paulo: Editora Edgard Blucher, 2017. p. 119-129.

FREITAG, R. M. K. Sociolinguística no/do Brasil. **Cadernos de Estudos Lingüísticos**, v. 58, n. 3, p. 445-460, 2016.

FREITAG, R. M. K.; PINHEIRO, B. F. M. Modelo de árvore de inferência condicional para explicar usos linguísticos variáveis. *In*: CARVALHO, C.; LOPES, N. S.; RODRIGUES, A. T. **Sociolinguística e Funcionalismo: vertentes e interfaces**. Salvador: EDUNEB, 2020.

FREITAG, R. M. K.; REIS, M.; BACK, A. C. D. P.; ROST-SNICHELOTTO, C. A. O controle do gênero textual/sequências discursivas na motivação da variação sociolinguística: apontamentos metodológicos. **Odisseia**, Natal, n. 3, p. 1-23, 2009.

FREITAG, R. M. Ko. **A expressão do passado imperfectivo no português: variação/gramaticalização e mudança**. 2007. Tese (Doutorado em Linguística) - Programa de Pós-graduação em Linguística, Universidade Federal de Santa Catarina, Florianópolis, 2007.

FREITAG, R. M. Ko. Banco de dados falares sergipanos. **Working Papers em Linguística**, v. 14, p. 156-164, 2013.

FREITAG, R. M. Ko. Dissecando a entrevista sociolinguística: estilo, sequência discursiva e tópico discursivo. *In*: GÖRSKI, E. M; COELHO, I. L.; SOUZA, C. M. N. de. (orgs.). **Variação estilística: reflexões teórico-metodológicas e propostas de análise**. Florianópolis: Insular, 2014. v.3. p.123-139.

FREITAG, R. M. Ko.; ARAUJO, A. S. Passado condicional no português: formas e contextos de uso. **Caligrama** – Revista de Estudos Românicos da Faculdade de Letras da UFMG, v. 16, p. 199-228, 2011.

FREITAG, R. M. Ko.; MARTINS, M. A.; TAVARES, M. A. Bancos de dados sociolinguísticos do português brasileiro e os estudos de terceira onda: potencialidades e limitações. **Alfa: Revista de Linguística**, São Paulo, v. 56, n. 3, p. 917-944, 2012.

FREITAG, R. M. Ko; GONÇALVES, S. C. L. Da forma para função ou da função para forma? **Guavira Letras**, Minas Gerais, n. 3, p. 89-108, 2011.

FREITAG, R. M. Ko; SNICHELOTTO, C. A. R. Análises contrastivas: estabilidade, variedade ou metodologia? **Working Papers em Linguística**, Florianópolis, v. 16, n. 1, p. 157-169, 2015.

GILES, H.; COUPLAND, J.; COUPLAND, H. Accommodation theory: communication, context and consequence. *In*: GILES, H.; COUPLAND, J.; COUPLAND, H. (orgs.). **Contexts of Accommodation Developments in Applied Sociolinguistics**. Cambridge: Cambridge University Press, 1991. p. 1-68

GONÇALVES, C. **Uma abordagem sociolinguística do uso das formas você, ocê e cê no português**. 2008. Tese (Doutorado em Linguística), Programa de Pós-Graduação em Linguística, Universidade de São Paulo, São Paulo, 2008.

GUIMARÃES, T. de A. A. S. **TU É DOIDO, MACHO!** A variação das formas de tratamento no falar de Fortaleza. 2014. Dissertação (Mestrado em Linguística Aplicada) - Programa de Pós-Graduação em Letras, Universidade Estadual do Ceará, Fortaleza, 2014.

GUIMARÃES, T. de A. A. S. **Tu e você no falar de fortaleza-CE**: variação e avaliações linguísticas. 2019. Tese (Doutorado em Linguística Aplicada) - Programa de Pós-Graduação em Letras, Universidade Estadual do Ceará, Fortaleza, 2019.

GUY, G. R.; ZILLES, A. **Sociolinguística Quantitativa** – Instrumental de Análise. São Paulo: Parábola, 2007.

HARLEY, H.; RITTER, E. Person and number in pronouns: A feature-geometric analysis. **Language**, p. 482-526, 2002.

HAUSEN, T. A. P. **Concordância verbal do pronome “tu” no interior do Estado de Santa Catarina**. 2000. Dissertação (Mestrado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade Federal do Paraná, Curitiba, 2000.

HERÊNIO, K. P. **“Tu” e “você” em uma perspectiva intra-linguística**. 2006. Dissertação (Mestrado) - Instituto de Letras e Linguística, Universidade Federal de Uberlândia, Uberlândia, 2006.

HOPPER, P. On some principles in the grammaticalization. *In*: TRAUGOTT, E. C.; HEINE, B. (eds.). **Approaches to grammaticalization**. Amsterdam /Philadelphia: John Benjamins Publishing Company, v. 1- 2, 1991. p. 7-35.

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

HOTHORN, T.; HORNIK, K.; ZEILEIS, A. Ctree: conditional inference trees. **The Comprehensive R Archive Network**, [s. l.], p. 1-34, 2015. Disponível em: <https://bit.ly/2PRvGYP>. Acesso em: 30 abr. 2022.

JESUS, J. M. **Efeitos do Pibid nos cursos de licenciatura do Campus professor Alberto Carvalho/UFS**: estudo comparativo entre egressos participantes e não participantes do programa durante e depois da formação inicial. 2018. Tese (Doutorado em Educação) – Universidade Federal de Sergipe, São Cristóvão, 2018.

KERBRART-ORECCHIONI, C. **Análise da conversação**: princípios e métodos. Tradução: Carlos Piovezani Filho. São Paulo: Parábola, 2006.

LABOV, W. **Padrões sociolinguísticos**. São Paulo: Parábola Editorial, 2008 [1972].

LABOV, W. Field methods of the project on linguistic change and variation. *In*: BAUGH, J.; SHERZER, J. (eds.). **Language in use**. Englewood Cliffs: Prentice Hall, 1984. p. 28-53.

LABOV, W. **Principles of Linguistic Change**: Internal Factors. Oxford UK & Cambridge, EUA: Blackwell, 1994.

LABOV, W. **Principles of Linguistic Change**: Social Factors. Oxford UK & Cambridge, EUA: Blackwell, 2001.

LABOV, W. **The Social Stratification of English in New York City**. Cambridge University Press, 2006 [1966].

LEVINSON, S. C. Deixis. *In*: HORN, L. R.; WARD, G. **The Handbook of Pragmatics**. Blackwell publishing, 2004. p. 97-121.

LOPES, C. R. S. *et al.* A reorganização do sistema pronominal de 2ª pessoa na história do Português Brasileiro: a posição de sujeito. *In*: LOPES, C. R. S. (coord.). **Mudança sintática das classes de palavra**: perspectiva funcionalista. 1. ed. São Paulo: Contexto, 2018. p. 24-141.

LOPES, C. R. S. Pronomes Pessoais. *In*: BRANDÃO, S. F.; VIEIRA, S. R. (orgs.). **Ensino de gramática**: descrição e uso. 1. ed. São Paulo: Contexto, 2007. p. 103-114.

LOPES, J. B. **Percepções e atitudes linguísticas dos chapecoenses frente à referência a segunda pessoa do singular (tu/você) no português brasileiro**. 2017. Dissertação (Mestrado em Linguística) - Programa de Pós-Graduação em Estudos Linguísticos da Universidade Federal da Fronteira Sul – UFFS, Chapecó, 2017.

LOREGIAN-PENKAL, L. **Re(análise) da referência de segunda pessoa na fala da Região sul**. Tese (Doutorado em linguística) – Programa de Pós-graduação em Linguística, Universidade Federal do Paraná, Curitiba, 2004.

LOREGIAN-PENKAL, L.; MENON, O P. da S. Você, ocê e cê em Curitiba. **Signum-Estudos de Linguagem**, v. 15, p. 201-221, 2012.

LUCCA, N. N. G. **A variação tu/ você na fala brasiliense**. 2005. Dissertação (Mestrado em Linguística) – Programa de Pós-Graduação em Linguística. Universidade de Brasília, Brasília, 2005.

MARSHALL, J. **Language Change and Sociolinguistics**: Rethinking Social Networks. 1. ed., Edinburgh, 2004.

MARTINS, G. F. **A alternância tu/você/senhor no município de Tefé:** Estado do Amazonas. 2010. Dissertação (Mestrado) – Programa de Pós-Graduação em Linguística, Universidade de Brasília, Brasília, 2010.

MARTINS, M. A. **Entre estrutura, variação e mudança:** uma análise sincrônica das construções com SE indeterminador no Português do Brasil. 2005. Dissertação (Mestrado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina, Florianópolis, 2005.

MENDONÇA, J. de J. **Traços semânticos da referência à primeira pessoa do plural no português brasileiro:** um estudo em tempo real. 2022. Tese (Doutorado em Letras) – Programa de Pós-Graduação em Letras, Universidade Federal de Sergipe, São Cristóvão, 2022.

MENDONÇA, J. de J. **Variação na expressão da 1ª pessoa do plural:** indeterminação do sujeito e polidez. 2016. Dissertação (Mestrado em Letras) – Programa de Pós-Graduação em Letras, Universidade Federal de Sergipe, São Cristóvão, 2016.

MENDONÇA, J. de J.; ARAUJO, A. S. Evaluation of the pronouns `a gente? and `tu? and of the grammatical patterns of agreement. **Revista de Estudos da Linguagem**, v. 27, p. 1613-1648, 2019.

MEYERHOFF, Miriam. **Introducing sociolinguistics.** New York: Routledge, 2006.

MILROY, L. **Language and social networks.** Oxford: Blackwell, 1980.

MILROY, L.; GORDON, M. **Sociolinguistics: method and interpretation.** Oxford: Blackwell, 2003.

MILROY, L.; MILROY, J. Social network and social class: toward an integrated sociolinguistic model. **Language in Society**, v. 21, n. 1, p. 1-26, 1992.

MODESTO, A. T. T. **Formas de tratamento no português brasileiro a alternância tu você na cidade de Santos/SP.** 2006. Dissertação (Mestrado em Letras) – Programa de Pós-Graduação em Linguística, Universidade de São Paulo, São Paulo, 2006.

MOTA, M. A. **A variação dos pronomes ‘tu’ e ‘você’ no português oral de São João da Ponte.** 2008. Dissertação (Mestrado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade Federal de Minas Gerais, Belo Horizonte, 2008.

NASCIMENTO, I. B. do. **O uso variável do pronome de segunda pessoa você(s)/cê(s) na cidade de São Paulo.** 2011. Dissertação (Mestrado) – Programa de Pós-Graduação em Semiótica e Linguística Geral, Universidade de São Paulo, São Paulo, 2011.

NOGUEIRA, F. M. da S. B. **Como os falantes de Feira de Santana e Salvador tratam seu interlocutor?** 2013. Dissertação (Mestrado em Letras) - Programa de Pós-Graduação em Língua e Cultura, Universidade Federal da Bahia, Salvador, 2013.

OLIVEIRA, L. A. F. de. Tu e Você no português popular do Estado da Bahia. *In:* VIII SEMINÁRIO DE PESQUISA E PÓS-GRADUAÇÃO DA UNIVERSIDADE FEDERAL DA BAHIA SALVADOR/UFBA (Apresentação de Trabalho/Comunicação), Salvador, 2007.

OUSHIRO, L. **Identidade na pluralidade: avaliação, produção e percepção linguística na cidade de São Paulo**. 2015. Tese (Doutorado em Linguística) - Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2015.

OUSHIRO, L. **Uma análise variacionista para as interrogativas-Q**. 2011. Dissertação (Mestrado em Linguística). Programa de Pós-graduação em Linguística, Universidade de São Paulo, São Paulo, 2011.

PAES, J. A. **Desenvolvimento de objeto pedagógico para o ensino produtivo de gramática: a balança das relações sociais**. Dissertação (Mestrado Profissional) - Programa de Pós-Graduação em Letras, Universidade Federal de Sergipe, São Cristóvão, 2019.

PERES, E. P. **O uso do você ocê cê em Belo Horizonte um estudo em tempo aparente e em tempo real**. 2006. Tese (Doutorado em Letras/Linguística) – Programa de Pós-Graduação em Letras/Linguística, Universidade Federal de Minas Gerais, Belo Horizonte, 2006.

R CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria, 2021. Disponível em: <https://www.r-project.org/>. Acesso em: 06 jan. 2022.

RIBEIRO, C. C. de S. **Deslocamento geográfico e padrões de uso linguístico: a variação entre as preposições locativas em ~ ni na comunidade de práticas da Universidade Federal de Sergipe**. 2019. Dissertação (Mestrado em Letras) - Programa de Pós-graduação em Linguística, Universidade Federal de Sergipe, São Cristóvão, 2019.

ROCHA, P. G. da. A. **O sistema de tratamento do português de Florianópolis: um estudo sincrônico**. 2012. Tese (Doutorado em Linguística) - Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina, Florianópolis, 2012.

ROCHA, W. J. C.; SANTOS, L. O. dos; SOUSA, V. V. O pronome você e sua variante cê: um estudo (socio)funcional. **Interdisciplinar- Revista de Estudos em Língua e Literatura**. São Cristóvão: UFS, v. 24, p. 143-158, 2016.

SANTANA, R. R. de. **Tipos de tipo em uma comunidade de práticas universitária**. 2019. Dissertação (Mestrado em Letras) - Programa de Pós-graduação em Letras, Universidade Federal de Sergipe, São Cristóvão, 2019.

SANTOS, A. M.; ARAUJO A. S.; FREITAG, R. M. Ko. **O “efeito gatilho” e a continuidade tópica: atuação do domínio tempo, aspecto e modalidade**. Signótica, Goiânia, v. 23, n. 2, p. 247–265, 2012. DOI: 10.5216/sig.v23i2.15221. Disponível em: <https://www.revistas.ufg.br/sig/article/view/15221>. Acesso em: 7 maio. 2020.

SANTOS, K. C. dos. **Estratégias de polidez e a variação de nós x a gente na fala de discentes da Universidade Federal de Sergipe**. 2014. Dissertação (Mestrado em Letras) - Programa de Pós-graduação em Letras, Universidade Federal de Sergipe, São Cristóvão, 2014.

SANTOS, V. M. **“Tu vai para onde?... Você vai para onde?": manifestações da segunda pessoa na fala carioca**. 2012. Dissertação (Mestrado em Letras Vernáculas) – Pós-Graduação em Letras Vernáculas, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2012.

SATHLER, E. H. B. **Estratégias de polidez utilizada por brasileiros em situações de elogios**: um estudo sociointeracionista. 2011. Dissertação (Mestrado em Linguística). Programa de Pós-Graduação em Linguística da Universidade Federal de Brasília, Brasília, 2011.

SCHERRE, M. M. P. Paralelismo linguístico. **Revista Estudos Linguísticos**, Belo Horizonte, v. 7, n. 2, p. 29-59, jul./dez., 1998. Disponível em: <http://www.periodicos.letras.ufmg.br/index.php/relin/article/view/2293>. Acesso em: 20 jan. 2018.

SCHERRE, M. M. P.; ANDRADE, C. Q.; CATÃO, R. de C. Redesenhando o mapa dos pronomes tu/você/ cê ocê no português brasileiro falado. *In*: I CONGRESSO NACIONAL DE ESTUDOS LINGUÍSTICOS. (Apresentação de Trabalho/Comunicação). Universidade Federal do Espírito Santo, Vitória, 2019. p. 1-30.

SCHERRE, M. M. P.; DIAS, E. P.; ANDRADE, C.; MARTINS, G. F. Variação dos pronomes “tu” e “você”. *In*: MARTINS, M. A.; ABRAÇADO, J. **Mapeamento sociolinguístico do português brasileiro**. São Paulo: Contexto, 2015. p. 133-172.

SCHNEUWLY, B.; DOLZ, J. **Gêneros orais e escritos na escola**. Campinas/SP: Mercado de Letras, 2004.

SIEWIERSKA, A. **Person**. Cambridge University Press, 2004.

SILVA, F. C. **Variação entre os pronomes tu e você na função de sujeito na conversação em Natal**: uma abordagem sociofuncionalista. 2015. Dissertação (Mestrado em Estudos da Linguagem) - Programa de Pós-Graduação em Estudos da Linguagem, Universidade Federal do Rio Grande do Norte, Natal, 2015.

SILVA, I. da. **Em terras de você o natural é misturar pronomes de segunda pessoa do singular estudos dos pronomes TU e VOCÊ no português popular do Brasil**. 2015. Tese (Doutorado em Letras) - Programa de Pós-Graduação em Filologia e Língua Portuguesa, Universidade de São Paulo, São Paulo, 2015.

SILVA, J. M. S. da. **Variação no preenchimento da posição determinante antes de possessivos pré-nominais**: padrões dialetais e contatos. 2020. Dissertação (Mestrado em Letras) - Programa de Pós Graduação em Letras, Universidade Federal de Sergipe, São Cristóvão, 2020.

SILVA, M. R. da. **Tu e você na variedade rio branquense**: Um caso de variação ou de escolha funcional? 2019. Tese (Doutorado em Estudos Linguísticos) - Programa de Pós-Graduação em Estudos Linguísticos, Universidade Estadual Paulista “Júlio de Mesquita Filho”, São José do Rio Preto, 2019.

SILVA, S. de O. P. **A variação pronominal de segunda pessoa do singular em Coité do Nóia**. 2019. Dissertação (Mestrado em Linguística) - Programa de Pós-Graduação em Letras e Linguística, Universidade Federal de Alagoas, Maceió, 2019.

SILVA, Suelen C. **A variação dos pronomes tu e você na fala mineira de Ressaquinha**. 2017. Dissertação (Mestrado em Linguística) - Instituto de Ciências Humanas e Sociais, Universidade Federal de Ouro Preto, Mariana, 2017.

SILVA, Suziane. O. P. **A variação você/cê na fala dos sertanejos alagoanos**. 2017. Monografia (Graduação em Letras) - Departamento de Letras Delmiro Gouveia, Universidade Federal de Alagoas, 2017.

SNICHELOTTO, C. A. R. Correlação entre sequências discursivas e marcadores discursivos de base verbal: um caso de variação estilística ou de motivação semântico-pragmática *In*: GÖRSKI, E. M.; COELHO, I. L.; NUNES DE SOUZA, C. M. (orgs.). **Variação estilística: reflexões teórico-metodológicas e propostas de análise**. Florianópolis: Insular, 2014. p. 227-246.

SOUSA, V. V. **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 em Linguística). Programa de Pós-Graduação em Letras, Universidade Federal da Paraíba, João Pessoa, Paraíba, 2008.

STREET JR, R. L. Accommodation in medical consultations. *In*: GILES, H.; COUPLAND, J.; COUPLAND, H. (orgs.). **Contexts of Accommodation Developments in Applied Sociolinguistics**. Cambridge: Cambridge University Press, 1991. p. 131-157.

TAVARES, M. A. Textos de diferentes gêneros produzidos em entrevistas sociolinguísticas: o caso do banco de dados Varsul. **Veredas on-line**, n. 2, p. 176-194, 2015.

TRAESEL, R. **As formas variáveis de tratamento ao interlocutor na ilha de Santa Catarina**: estudo da fala de adolescentes dos ingleses. 2016. Dissertação (Mestrado em linguística) Programa de Pós-Graduação em linguística, Universidade Federal de Santa Catarina, Florianópolis, 2016.

TRAUGOTT, E. C.; DRASHER, R. B. **Regularity in Semantic Change**. Cambridge: Cambridge University Press, 2002.

UNIVERSIDADE FEDERAL DE SERGIPE. UFS em números 2017. Edição especial http://indicadores.ufs.br/uploads/page_attach/path/1367/ufs_em_numeros_2014.pdf

VALLE, C. R. M.; GÖRSKI, E. M. Por um tratamento multidimensional da variação estilística na entrevista sociolinguística. *In*: GÖRSKI, E. M.; COELHO, I. L.; NUNES DE SOUZA, C. M. (orgs.). **Variação estilística: reflexões teórico-metodológicas e propostas de análise**. Florianópolis: Insular, 2014. p. 93-121.

WEINREICH, U.; LABOV, W.; HERZOG, M. **Fundamentos empíricos para uma teoria da mudança linguística**. Tradução: Marcos Bagno. São Paulo: Parábola Editorial, 2006

WIESE, H.; SIMON, H. J. Grammatical properties of pronouns and their representation: na exposition. *In*: SIMON, H. J.; WIESE, H. **Pronouns: grammar and representation**. Linguistik Aktuell/Linguistics Today. Amsterdam/Philadelphia, v. 52, 2002. p. 1-21.

APÊNDICES

APÊNDICE A – ROTEIRO DE ENTREVISTA⁷⁰

As primeiras perguntas são mais de cunho pessoal

- Qual o seu nome?
- Quantos anos você tem?
- Para você o que é ser universitário?
- Que curso você faz?
- Por que escolheu esse curso?
- Quais os pontos positivos e negativos de seu curso?
- O custo para se manter em seu curso é caro? Quanto você já gastou? Com que gasta?
- Pretende atuar nessa profissão?
- O que eu devo fazer para seguir essa profissão?
- Que perfil eu devo ter para poder fazer o seu curso?
- O que eu devo fazer para conseguir uma vaga no mercado de trabalho e me destacar na sua área?
- Em quais turnos você faz disciplinas? Pergunte como a pessoa costuma se alimentar na universidade. Onde? O que costuma comer? Como faço para chegar nesses locais que você costuma comer? (Solicitar o passo a passo para chegar lá)
- Você sente dificuldade na universidade?
- Você recebe alguma bolsa da universidade? Caso tenha uma bolsa de pesquisa, pergunte o que ela pesquisa, qual a relação com o orientador(a). A quanto tempo ganha a bolsa? Diga que você queria ter uma bolsa de pesquisa, pergunte quais são os requisitos que você teria que ter para ganhar a bolsa.
- Pergunte se a pessoa já fez pesquisa científica. Depois diga para a pessoa que você ainda não fez uma pesquisa científica e peça para ela explicar o passo a passo para você fazer uma pesquisa científica.
- Qual meio de transporte você utiliza para vir para a universidade?
- Você trabalha? Em que área? Se sim, quais as dificuldades que você encontra em seu trabalho? Qual meio de transporte utiliza para chegar ao seu trabalho?
- Quais as suas atividades diárias ou semanais?
- Nos finais de semana, o que você mais gosta de fazer? Para onde você costuma sair?
- Você costuma ir ao cinema? Com que frequência? Com quem você mais vai ao cinema?
- E ao teatro, você já foi alguma vez? Como foi essa experiência? Você assistiu que peça?

*Dependendo do entrevistado pergunta-se mais coisas.

Em seguida, elaborara-se perguntas a respeito do local de moradia, opções de lazer e meio ambiente.

⁷⁰ Roteiro elaborado por Andréia Silva Araujo e Josilene de Jesus Mendonça.

- Onde você mora? Em que parte da cidade?
- Há quanto tempo mora nesse lugar?
- Sempre morou nesse lugar? E seus pais também? Se não, onde moraram?
- Gosta de morar nesse lugar?
- Se tivesse oportunidade moraria em outro lugar?
- Você se desloca da sua comunidade para outros lugares? Para quais?
- O que você faz em cada um deles?
- Você fez amizades nesses ambientes? Com qual frequência vocês interagem/conversam? E com as outras pessoas que fazem parte desse ambiente, você tem contato? Interage com essas pessoas?
- Você se encontra com essas pessoas em outros ambientes (finais de semana ou feriados)? Com qual frequência vocês se encontram/conversam?
- Você participa de alguma atividade em grupo na sua cidade (religião/oração, estudos, dança, esporte, academia etc.)? Com frequência vocês se reúnem? Com quem você costuma interagir nesse grupo? Participa de outros grupos?
- Se eu fosse morar na sua cidade, quais tipos de trabalho eu poderia fazer para ganhar meu próprio sustento?
- Quantas pessoas moram com você? Quem são? Se a pessoa for casada, pergunte de onde é a esposa/o e a profissão.
- A casa onde vocês moram é própria ou alugada?
- O que você mais gosta de fazer no local onde mora?
- O que é atrativo para os moradores desse lugar?
- Estou pensando em visitar a sua cidade próximo final de semana, que lugar você me indica para comer? Como eu faço para chegar a esse local? (Peça o passo a passo para chegar nesse lugar)
- Você conhece seus vizinhos? Qual a sua relação com eles? (Caso a pessoa diga que não tem muito contato com os vizinhos pergunte por quê?)
- As pessoas do local onde você mora costumam se ajudar? (Leilão, Bingo, pedir alimento emprestado como açúcar, café etc.)
- Digamos que eu seja sua vizinha e espalhei uma fofoca sobre você. E, agora que estamos frente a frente, o que você falaria para mim?
- Há coleta de lixo regularmente nessa localidade?
- Você acha que a quantidade de vezes que o serviço é oferecido é suficiente para atender a comunidade?
- Há separação de tipos de lixo?
- O que você acha da coleta letiva?
- Você pratica alguma ação para preservar o meio ambiente? Se sim, quais? Se não, quais ações poderia fazer?

Aproveitando a oportunidade do assunto relacione a questão da segurança, fazendo perguntas que englobem questões locais e globais.

- Você considera sua cidade segura?

- Já aconteceu alguma coisa que te deixou assustado(a)? (Do tipo estupro, assassinato, assalto, violência contra mulher e outras situações do tipo). Se aconteceu pergunte: o que? Como foi? Com quem foi?)
- O que você acha que deveria mudar em relação à segurança?
- Quais as medidas que devem ser tomadas para solucionar ou diminuir tais problemas?
- Pergunte como ele via a questão da segurança há dez anos, pergunte se mudou muita coisa, o que mudou.
- O que você acha que aumenta a criminalidade?

Insira questões relacionadas à educação.

- Pergunte o que ele acha da educação no Brasil
- Se ele acha que existem diferenças em relação ao ensino público e o ensino privado
- Pergunte em qual rede de ensino ele estudou durante boa parte da vida
- Seus pais estudaram? Eles estudaram até que nível? O que seus pais falam sobre seus estudos?
- Seus pais trabalham em quê?
- Você tem irmãos? Quantos? Eles também fazem/fizeram faculdade? Quantos anos eles têm?
- Qual a sua relação com seus irmãos?
- Você acha que com o passar dos anos houve uma melhoria no acesso à educação?
- Como era a didática dos professores quando você cursava o ensino básico?
- Sua visão sobre a escola mudou depois que você entrou na universidade?
- Pergunte sobre a política de cotas no ensino público

*as perguntas mudam de acordo com cada entrevistado, pois tem uns que falam mais sobre determinado assunto e outros menos.

Insira perguntas relativas a experiências pessoais.

- Você ou alguém próximo já passou por algum risco de morte?
- Conte um fato que marcou sua infância.
- Quando você era criança, você costumava brincar de quê?
- Como faço para brincar disso? (Pedir para ele/ela ensinar a brincadeira que citou)
- Quem era seu melhor amigo(a)?
- Você ainda tem contato com seus amigos de infância?
- Como eram seus pais? Eles eram muito rígidos?
- Caso a pessoa tenha dito que tem irmãos, pergunte como era a relação dela com os irmãos na infância: Vocês eram companheiros? Brigavam? Você lembra de alguma travessura que fizeram juntos?
- Você já passou por algum acontecimento constrangedor na escola? Já presenciou alguma cena constrangedora com algum colega?
- Conte um momento engraçado que aconteceu com você e seu melhor amigo.
- Digamos que eu seja a sua melhor amiga e eu contei para outra pessoa um segredo seu. O que falaria para mim?

- Conte um momento muito triste que você viveu com seus pais.
- Em sua casa, o que você faz nas suas horas vagas?
- Gosta de assistir? O que você mais gosta de assistir? Na sua casa, tem TV por assinatura? Você tem Netflix em casa? Você gosta mais de filmes ou séries? Quais filmes ou séries (a depender da resposta à pergunta anterior) você costuma assistir? Poderia nos contar um pouco do enredo da sua série (ou filme) preferida?
- Com quem você costuma assistir? Vocês têm o mesmo gosto?
- Na TV aberta, o que você costuma assistir? Tem alguma emissora ou programação preferida?
- Você ou sua família costuma ouvir rádio? Quais emissoras ou programas vocês preferem?
- Quando você quer buscar mais informações sobre uma notícia específica, qual meio de comunicação você costuma utilizar? (Se a pessoa responder internet, tv ou rádio, pergunte o site, o canal, a emissora ou programa específico)
- Você tem computador? Você tem acesso à internet em casa? (Caso isso já tenha ficado claro nas respostas anteriores, não faça essa pergunta) Você prefere acessar a internet pelo celular ou pelo computador? Há alguma atividade que você prefira ou ache mais prático fazer pelo celular? Quais sites você acessa com frequência?
- Você gosta de ler? O que você costuma ler? Você lê regularmente? (Se a pessoa responder que só lê os livros do curso, pergunte se ela lê outros tipos de texto)
- Qual a sua opinião sobre o transporte público em sua cidade? E o transporte público de Sergipe? E do Brasil?
- Como você faz para ir para o comércio?
- Em qual cidade você costuma fazer compras? (Caso a pessoa responda uma cidade diferente da dela, pergunte por quê?) Você se considera consumista?
- Diga que é consumista e pergunte: o que eu tenho que fazer para deixar de ser consumista? Que dicas tu me daria?
- Para onde você já viajou? Como foi essa experiência? Com que frequência você viaja? Que lugares gostaria de conhecer?

Elabore perguntas relacionadas à cultura local, religião, culinária e a diversidade cultural que existe no Brasil.

- O que você acha que é cultura?
- Tem alguma manifestação cultural aqui na sua cidade que é considerada tradição?
- Qual o papel da religião na sua cidade?
- Você acredita e faz parte de alguma religião? Em que sua religião se diferencia das outras?
- Tem algum ponto turístico na cidade que você mora? (se sim, pergunte: como eu faço para chegar lá caso eu queira visitar esse lugar? Peça para dizer o passo a passo que você tem que fazer para chegar nesse lugar)
- Você sabe cozinhar? (caso a pessoa diga que não, insista perguntando se pelo menos algo simples como miojo, sobremesa, arroz).

- Se sabe peça para ele te dizer uma receita e peça para ele te ensinar o passo a passo como é que faz determinada comida.
- Conhece alguma simpatia (para tirar verruga, para arranjar marido/esposa, para tirar olho grosso etc.)? Explique como que eu posso fazer tal simpatia.

Elabore questões a respeito da atualidade social e política

- Você acha que a mulher já conseguiu igualdade social? (Se o entrevistado for mulher, pergunte se ela considera que seu papel social seja diferente do que a mãe dela exerce ou exercia?)
- Qual a sua opinião em relação à diversidade sexual e aos direitos adquiridos pelos LGBTs? (Lésbicas, gays, bissexuais e transexuais)
- Você concorda com casamento homoafetivo (entre pessoas do mesmo sexo/gênero)? Qual a sua opinião em relação à adoção por casais homoafetivos?
- Você acha que o mercado de trabalho respeita as minorias sociais? (exemplos: mulher, gays, negros, pobres etc.)
- Qual a sua opinião sobre o cenário político atual do Brasil?
- Caso eu fosse um político, o que diria para mim?
- Que ações sugeriria que eu tomasse já que sou um político?
- E se eu fosse o atual presidente da república, o que diria para mim?

ANEXOS

ANEXO A – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Caro informante,

Estamos convidando-o a participar como voluntário de uma pesquisa de campo a ser realizada por meio da gravação de entrevista sobre temas relacionados a questões sociais.

A coleta será realizada com o objetivo de desenvolvermos um trabalho acadêmico vinculado ao Programa de Pós-graduação em Letras.

A entrevista coletada ficará disponível no bando de dados do Grupo de Estudos em Linguagem, Interação e Sociedade – GELINS; para ser utilizada em pesquisas futuras. Serão resguardadas todas as informações de identificação de forma que se mantenha o anonimato.

Não será cobrado nada, não haverá gastos e não estão previstos ressarcimentos ou indenizações.

Você poderá solicitar esclarecimento sobre a pesquisa em qualquer etapa do estudo. Você é livre para recusar-se a participar, retirar seu consentimento ou interromper a participação na pesquisa a qualquer momento, seja por motivo de constrangimento e/ou outros motivos. A sua participação é voluntária e a recusa em participar não irá acarretar qualquer penalidade ou perda de benefícios.

Desde já, agradecemos sua atenção e participação e colocamo-nos à disposição para maiores informações.

Consentimento para participação

Eu, _____, idade: _____, estado civil: _____, RG: _____, estou de acordo com a participação no estudo descrito acima. Eu fui devidamente esclarecido quanto aos objetivos da pesquisa, aos procedimentos aos quais serei submetido e os possíveis riscos envolvidos na minha participação. Os pesquisadores me garantiram disponibilizar qualquer esclarecimento adicional a que eu venha a solicitar durante o curso da pesquisa e o direito de desistir da participação em qualquer momento, sem que a minha desistência implique qualquer prejuízo à minha pessoa ou à minha família, sendo garantido anonimato e o sigilo dos dados referentes à minha identificação, bem como de que a minha participação neste estudo não trará nenhum benefício econômico. Ao mesmo tempo, libero a utilização de minha entrevista para fins científicos e de estudos (livros, artigos, slides e transparências), em favor dos pesquisadores, obedecendo ao que está previsto na Resolução do CNS nº 196/96. Autorizo também que a minha interação fique disponível no banco de dados acima referido para ser utilizada em pesquisas futuras.

_____, ____ de _____ de _____

Assinatura do (a) participante: _____

Assinatura do (a) pesquisador (a): _____

Assinatura do (a) coordenador (a)/orientador(a): _____

Fonte: Modelo de termo utilizado pelo Grupo de Estudos em Linguagem Interação e Sociedade (GELINS) na constituição do Banco de dados Falares Sergipanos (2017) (adaptado).

ANEXO B – FICHA SOCIAL DO INFORMANTE

Data: ___/___/____ Local da gravação: _____ Pesquisador: _____ Grau de relação entre informante e pesquisador: () bastante próximo () próximo () neutro () Distante INFORMANTE: Nome completo: _____ Apelido (se tiver): _____ Estado civil: _____ Sexo: _____ Gênero: _____ Zona de residência: () rural () urbana Qual bairro? _____ Cidade: _____ CEP: _____ Telefone para contato: _____ E-mail: _____ Local de nascimento: _____ data de nascimento: _____ Profissão/ocupação: _____ Outras atividades: _____ Instrução: () Ensino Superior em andamento Curso: _____ Período: _____ <i>Campus:</i> _____ Outra formação: () Superior completo Curso: _____ () Outro nível: _____ Escolas em que estudou: Fundamental 1 () pública () privada Fundamental 2 () pública () privada Ensino Médio () pública () privada Mora com a família: () Sim () Não Morou por mais de um ano em outro município? Sim () Não () Nome do(s) lugar(es) em que morou e tempo aproximado: _____ _____ Prestou Serviço Militar? Sim () Não () Cidade em que prestou Serviço Militar: _____ DADOS RELATIVOS AOS PAIS DO INFORMANTE Em que município nasceu e morou por mais tempo? a) o pai nasceu: _____ morou: _____ Idade: ___ Ocupação: _____ Grau de escolaridade: _____ b) a mãe nasceu: _____ morou: _____ Idade: ___ Ocupação: _____ Grau de escolaridade: _____
--

Fonte: Modelo de ficha social utilizado pelo Grupo de Estudos em Linguagem Interação e Sociedade (GELINS) na constituição do Banco de dados Falares Sergipanos (2017) (adaptado).

**ANEXO C – NORMAS ADOTADAS PELO GRUPO DE ESTUDOS EM LINGUAGEM
INTERAÇÃO E SOCIEDADE (GELINS) PARA A REALIZAÇÃO DE
TRANSCRIÇÃO ORTOGRÁFICA**

OCORRÊNCIA	SINAIS	EXEMPLIFICAÇÃO
Interrogação	?	sabe o é?
Comentário do transcritor sobre o que está acontecendo no ambiente	(())	((RISOS)) ((PIGARRO))
Estímulo do interlocutor Marcar na trilha do próprio informante	(est)	(est)
Hesitação do locutor	(hes)	foi (hes) uma brincadeira bem interessante
Truncamento de palavra	-	come- começou
Nomes próprios, profissões, nomes de cursos, filmes	iniciais maiúsculas	... fui à Petrópolis uma ...
Discurso direto	“ ”	eu saio pra apresentar trabalho fora eles têm orgulho “ah ela saiu pra outro estado tá apresentando trabalho da universidade” então de certa forma isso é um apoio...
Números	por extenso	eu tenho vinte e oito anos...
Incompreensão do que ouviu	()	
Hipótese do que ouviu	(hipótese)	ter que estudar lá no no <i>Campus</i> de São Cristóvão ia re- ia reque- requerer da minha (como a associação) que eu teria que pagar todos os meses...
Onomatopeias e siglas	caixa alta	a questão do incentivo de participação de eventos porque assim de eventos por exemplo a OCMEA ela é incentivado por todos os professores

Observações:

- Estamos realizando a revisão ortográfica (de acordo com o novo acordo ortográfico);
- Não corrigir erro de concordância do falante;
- As variações morfossintáticas devem ser transcritas;
- As variações fonológicas não devem ser consideradas;
- Se o falante disser: num, pra, tava, tou, ocê, cê e demais reduções - conserva-se como ele falou, desde que não seja fonológica;
- Marcar o infinitivo;
- Marcar a pausa preenchida com: ah, eh, uh
- Início de frase não deve ser utilizado letra maiúscula.

Fonte: Modelo de ficha social utilizado pelo Grupo de Estudos em Linguagem Interação e Sociedade (GELINS) na constituição do Banco de dados Falares Sergipanos (2017) (adaptado).