



Universidades de Brasília
Instituto de Letras
Departamento de Linguística, Português e Línguas Clássicas
Programa de Pós-Graduação em Linguística

Tu e mais quantos? – A segunda pessoa na fala brasiliense

Carolina Queiroz Andrade

Brasília

2010

Universidade de Brasília
Instituto de Letras
Departamento de Linguística, Português e Línguas Clássicas
Programa de Pós-Graduação em Linguística

Tu e mais quantos? – A segunda pessoa na fala brasiliense

Carolina Queiroz Andrade

Dissertação apresentada ao Departamento de Linguística, Português e Línguas Clássicas do Instituto de Letras da Universidade de Brasília como requisito parcial à obtenção do título de Mestre em Linguística.

Orientadora: Prof.^a Dr.^a Maria Marta Pereira Scherre.

Brasília
2010

Banca Examinadora

Prof.^a Dr.^a Maria Marta Pereira Scherre – Presidente: UnB/UFES

Prof.^a Dr.^a Loremi Loregian-Penkal – Membro efetivo externo: UNICENTRO-PR

Prof.^a Dr.^a Heloisa Maria Moreira Lima Salles – Membro efetivo interno: UnB/LIP

Prof.^a Dr.^a Daisy Bárbara Borges Cardoso – Membro Suplente: SEEDF

Para Marta Scherre, minha mãe acadêmica.

Agradecimentos

Às diversas figuras de Deus e ao meu país, por todos os percalços e oportunidades.

À minha filha e ao meu marido, pela infinita compreensão, apoio e amor.

À professora e orientadora Marta Scherre, por sua história e exemplo, pelo seu admirável amor à pesquisa e pela sua inesgotável generosidade.

Às professoras Orlene Lúcia de Sabóia Carvalho e Janaína de Aquino Ferraz, pelo grande apoio acadêmico e por tantas portas abertas.

À professora Rachel Dettoni, pela dedicação com que nos ensinou, tanto na graduação como na pós, sobre as bases sociolinguísticas. Aproveito para agradecer também pelo convite para a participação no VALCO, que é, entre outras coisas, uma importante ferramenta de registro de fala do Centro-Oeste.

À professora Adriana Chan, pelo entusiasmo com que não respondia as nossas dúvidas, uma postura colaboradora para o despertar de uma consciência investigativa.

À Universidade de Brasília, aos professores e aos funcionários, principalmente à Renata, por cada orientação, ajuda e apoio.

Às companheiras de fenômeno, Nivia Lucca e Edilene Patrícia Dias, por tantos materiais emprestados e pelas trocas. Agradeço ainda à Edilene Patrícia e à Juliana Shibata, pela leitura de última hora do *abstract*.

Aos amigos Patrícia Tavares, Cíntia Pacheco, Aline Mesquita, Dalmo Borges, Daisy Bárbara, Caroline Cardoso, Josane de Oliveira, Germano Martins, Janaína Ferraz, Lara Villas Bôas, Thalita Santos, Fernanda Gláucia e Sara Andrade pelas parcerias, pelas trocas e imenso apoio acadêmico. Agradeço imensamente à Lara Villas Bôas pela revisão desta dissertação.

Aos meus parentes e amigos, especialmente à tia Rosa, pela força e empolgação.

À Vila Planalto e à escola da Vila, pela oportunidade de contar um pouquinho dessa história, principalmente na pessoa da Andréa Seraffini dos Passos Longo, que trabalhava nessa escola e fez a ponte fundamental entre pesquisadora e informantes.

Ao CNPq, pelo apoio financeiro.

A todos que direta ou indiretamente inspiraram e apoiaram esta pesquisa, de coração, eu agradeço!

A imobilidade das coisas que nos cercam talvez lhes seja imposta pela nossa certeza de que essas coisas são elas mesmas e não outras, pela imobilidade de nosso pensamento perante elas.

(Marcel Proust)

SUMÁRIO

ÍNDICE DE TABELAS	ii
ÍNDICE DE FIGURAS	iv
ÍNDICE DE QUADROS	v
ÍNDICE DE GRÁFICOS	vi
1. INTRODUÇÃO	7
1.1 Breve contextualização sobre o meu encontro com a pesquisa linguística	10
1.2 Principais quês e porquês da presente pesquisa	10
2. O <i>LOCUS</i>	17
2.1. Vila Planalto, Brasília	17
2.2. A fala da Vila Planalto como objeto de estudo	23
3. TEORIA E METODOLOGIA	27
3.1. Metodologia desenvolvida para constituição dessa amostra	33
3.1.1. Descrição das gravações: facilidades e dificuldades	35
3.2 Descrição da metodologia de análise da amostra	38
4 UM RETRATO DE ALGUNS TRABALHOS JÁ REALIZADOS	41
4.1 Os pronomes segundo alguns gramáticos e alguns linguistas	41
4.2 Referência à 2ª pessoa e a perspectiva T/ V	43
4.3 Alguns trabalhos sobre os pronomes pelo Brasil	45
4.3.1 Região Sudeste	46
4.3.2 Região Sul	54
4.3.3 Região Nordeste	58
4.3.4 Região Norte	61
4.3.5 Distrito Federal, Região Centro-Oeste	61
4.4 Sobre o que vimos	67
5.1. Ponderações preliminares	69
5.2. Descrição do corpus	71
5.3 Como as variáveis serão analisadas: breve visão de conjunto	74
5.4 Efeitos de natureza externa	75
5.4.1 Faixa etária e tipo de relação entre os interlocutores (simetria <i>versus</i> assimetria)	76
5.4.2 Origem dos pais	86
5.4.3 Sexo	94
5.4.4 Tipo de relato	99
5.5 Efeitos de natureza interna	101
5.5.1 Função sintática	101
5.5.2 Tipo de referência, presença de vocativo e entonação	104
5.5.3 Paralelismo Linguístico	109
5.6 Considerações sobre as análises	114
6. CONSIDERAÇÕES FINAIS	119
7. REFERÊNCIAS BIBLIOGRÁFICAS	124
8. APÊNDICE	130

ÍNDICE DE TABELAS

Tabela 01 – Origem dos moradores da Vila Planalto, segundo a CODEPLAN (2009).	21
Tabela 02 – Variação <i>tu/ você</i> na fala carioca em função da idade.....	48
Tabela 03 – Variação <i>tu/ você</i> na fala carioca em função do sexo.....	49
Tabela 04 – Função sintática da variação <i>tu/ você</i> em Santos, SP.....	50
Tabela 05 – Variação <i>tu/ você</i> na fala mineira em função da idade.....	51
Tabela 06 – Função sintática da variação <i>tu/ você</i> em São João da Ponte, MG.....	52
Tabela 07 – Variação <i>você/ ocê/ cê</i> segundo a função sintática (Gonçalves, 2008).....	54
Tabela 08 – O <i>tu</i> em função da localidade e do sexo, em Santa Catarina (SC) e no Rio Grande do Sul (RS).....	56
Tabela 09 – O <i>tu</i> em função da localidade e faixa etária, em SC e RS.....	57
Tabela 10 – A variação <i>você/ ocê/ cê</i> no Distrito Federal em função do sexo (dados dos informantes de 10 a 14 anos).....	62
Tabela 11 – Função sintática da variação <i>você/ cê/ ocê</i> em Brasília (dados da entrevistadora).....	63
Tabela 12 – A variação <i>tu/ você</i> no Distrito Federal em função da idade (Dias, 2007).	65
Tabela 13 – A variação <i>tu/ você</i> no Distrito Federal em função do sexo (Dias, 2007)..	65
Tabela 14 – A variação <i>tu/ você</i> no Distrito Federal em função do sexo e idade (Dias, 2007).....	66
Tabela 15 – Detalhamento dos informantes.....	73
Tabela 16 – Variação <i>você/ cê/ tu</i> em Brasília em relação à idade (todos os dados)....	76
Tabela 17 – Variação <i>cê/ tu/ você</i> na Vila Planalto em relação à faixa etária.....	79
Tabela 18 – Variação <i>você, cê</i> e <i>tu</i> em relação ao tipo de relação e a faixa etária (com todos os dados).....	81
Tabela 19 – Variação <i>você, cê</i> e <i>tu</i> na Vila Planalto em relação ao tipo de relação e a faixa etária.....	82
Tabela 20 – Resultados de pesos relativos da variação <i>você/ cê/ tu</i> na Vila Planalto em relação à faixa etária.....	83
Tabela 21 – Tipo de relação entre os interlocutores (dados da Vila).....	85
Tabela 22 – Variação <i>você/ cê/ tu</i> em relação à origem dos pais (todos os dados).....	88
Tabela 23 – Variação <i>você/ cê/ tu</i> em relação às origens dos pais, resultados em pesos	90
Tabela 24 – Resultados gerais quanto ao sexo (estudo de 2005).....	95
Tabela 25 – Resultados do sexo em função da região administrativa (estudo de 2005)	95
Tabela 26 – Resultados gerais, com as três faixas etárias em função do sexo (estudo de 2007).....	95
Tabela 27 – Resultados quanto ao sexo, mas somente da faixa etária de 13 a 19 anos (estudo de 2007).....	95
Tabela 28 – Efeito do fator sexo em rodadas binárias dos dados da Vila Planalto.....	97
Tabela 29 – Efeito do fator sexo em percentuais e pesos (dados da Vila Planalto).....	97
Tabela 30 – Fator sexo em função das localidades, baseados em rodadas binárias.....	98
Tabela 31 – Variação <i>você/cê/tu</i> em relação ao tipo de fala (dados da Vila Planalto).	100
Tabela 32 – A variação <i>você/cê/tu</i> em função do tipo de fala, a partir de rodada eneária.....	100
Tabela 33 – Variação <i>você/ cê/ tu</i> em relação à função sintática, com todos os dados da amostra.....	102
Tabela 34 – Variação <i>você/ cê/ tu</i> em relação à função sintática, resultados em percentuais, com todos os dados da amostra.....	103

Tabela 35 – Resultados em pesos da variação <i>você/ cê/ tu</i> quanto à função sintática, com todos os dados da amostra	104
Tabela 36 – Variação <i>você/cê/tu</i> em relação ao tipo de referência, valores em percentuais com todos os dados da amostra	105
Tabela 37 – Variação <i>você/cê/tu</i> em relação ao tipo de referência, valores em pesos (todos os dados)	105
Tabela 38 – Variantes <i>você/cê/tu</i> em relação à presença ou ausência de vocativo (todos os dados)	106
Tabela 39 – Variação <i>você/ cê/ tu/</i> em relação à entonação, resultados em percentuais para toda a amostra	108
Tabela 40 – Variação <i>você/ cê/ tu/</i> em relação à entonação, resultados em pesos	108
Tabela 41 – Variação entre <i>cê/você</i> quanto ao paralelismo, resultados de 2004.....	110
Tabela 42 – Variação <i>você/cê/ tu/</i> em relação ao Paralelismo (percentuais).....	111
Tabela 43 – Resultados da variação em relação ao Paralelismo (em rodadas binárias, com todos os dados)	112
Tabela 44 – Variação <i>você/cê/tu</i> em relação ao Paralelismo (todos os dados, resultados em pesos)	113

ÍNDICE DE FIGURAS

Figura 01 – Expedição Cruls	18
Figura 02 – Vila Planalto.....	24
Figura 03 – Brasília:	24
Figura 04 – Distrito Federal.....	26
Figura 05 – Brasil	26

ÍNDICE DE QUADROS

Quadro 01 – Proposta de quadro pronominal para o Brasil	42
Quadro 02 – Variação <i>tu/você</i> na região Sudeste	53
Quadro 03 – Variação <i>tu/você</i> na região Sul.....	58
Quadro 04 – A variação <i>tu/você</i> na região Nordeste	60
Quadro 05 – A variação <i>tu/você</i> na região Norte	61
Quadro 06 – A variação <i>tu/você/cê</i> no DF, região Centro-Oeste	66

ÍNDICE DE GRÁFICOS

Gráfico 01 – Origens da população brasiliense	22
Gráfico 02 – Idade	78

Resumo

O presente trabalho trata da variação pronominal que ocorre em Brasília entre as formas *você/cê/tu*. Os pressupostos teórico-metodológicos utilizados foram baseados na Teoria da Variação e Mudança Linguística, preconizada por Weinreich, Labov e Herzog (1968). Com base nessa teoria, verificamos que a variação não é aleatória, mas ocorre de forma organizada, segundo um conjunto de fatores internos e externos à língua. Foram analisadas amostras de fala de crianças e adolescentes entre 7 a 15 anos, com foco especial para a Vila Planalto, a fim de verificar se, já nos primeiros anos, as crianças usariam a variante *tu* que, em Brasília, figura como variante inovadora. Além disso, investigamos se a variante *tu* vem assinalando uma expansão em falas femininas. Verificamos, por fim, se a variante *cê* figura como uma forma de esquiva, usada para evitar pronomes que possam soar muito formais, como “senhora”, por um lado, ou muito informais, como *tu* ou *você*, por outro. A hipótese sobre a aquisição em tenra idade da variante *tu* foi confirmada na amostra da Vila Planalto (onde foram coletados cerca de 70% dos dados), uma vez que essa variante ocorreu de forma semelhante nas faixas etárias de 7-11 e de 12 a 15 anos. Uma variável que atua paralelamente à faixa etária é a relação entre os interlocutores. Se simétrica, tende a favorecer o uso do *tu*; se assimétrica, tende a favorecer o uso de *você* ou *cê*. A hipótese sobre a expansão do pronome *tu* em falas femininas foi apenas parcialmente confirmada, pois, apesar de sua ocorrência em números absolutos e percentuais ser superior, se comparada com resultados de trabalhos realizados anteriormente, os efeitos em termos de pesos relativos ficaram apenas equivalentes nessa comparação. A hipótese sobre o *cê* ser uma forma de esquiva também se confirmou, uma vez que, quando houve preocupação quanto ao uso dos pronomes, as formas escolhidas foram *senhora* e *cê*. Na análise dos dados, a origem dos pais emergiu como uma variável importante no entendimento do surgimento do *tu* na fala brasiliense, sendo a origem nordestina um fator fundamental neste processo. Contudo na Vila Planalto, a naturalidade brasiliense dos pais já assume seu papel de expansão e fixação do *tu* no rol dos pronomes de segunda pessoa. Os resultados alcançados também indicaram que a fala de Brasília, apesar de ainda incipiente em relação a ter uma variedade linguística típica, vem assinalando um desenvolvimento de características da oralidade, e o *tu*, como uma forma marcada, faz parte dessa focalização. Além das forças sociais e interacionais analisadas, forças internas, comuns a outras pesquisas de segunda pessoa se mostraram relevantes no entendimento deste processo variável, entre as quais destacamos a função sintática, o tipo de referência e o paralelismo linguístico. No complexo jogo entre as variantes *você/cê/tu*, foi possível verificar que a abordagem estatística que dá conta de forma mais completa dos dados é a que considera individualmente as três variantes,

ou seja, é a análise eneária, tendo em vista que, a depender das forças em jogo, *você* e *cê* se opõem a *tu*; ou *você* e *tu* se opõem a *cê*; ou *você*, *cê* e *tu* apresentam comportamentos distintos.

Abstract

This dissertation deals with the variation in Brasilia on the usage of the pronominal forms “você” / “cê” / “tu”. The theory and methodology were based on the “Theory of Language Variation and Change”, developed by Weinreich, Labov and Herzog (1968). Based on this theory, we found that the variation is not random, but occurs in an organized way, according to internal and external factors of the language. We analyzed speech data of children and teenagers aged between 7-15 years old, with special focus on Vila Planalto, to determine whether, in the first years, children would use the variant “tu”, which is the innovative variant in Brasilia. Furthermore, we investigated whether the variant “tu” is expanding in the speech of women. We verified, finally, whether the variant “cê” appears as a form of avoidance, used to prevent the use of pronouns that may sound very formal, like “senhora” or very informal, like “você” or “tu”. The hypothesis about the acquisition in an early age of “tu” was confirmed in the sample of Vila Planalto (where 70% of the data were collected), since this variant was similar in the age groups from 7 to 11 and 12 to 15. A variable that acts in parallel with the age group is the relationship between the interlocutors, if symmetrical, it tends to favor the use of “tu”, if asymmetrical, it tends to favor the use of “você” or “cê”. The hypothesis of the expansion of the pronoun “tu” in female speech was only partially confirmed because, despite the fact that in frequency it is higher when compared to previous works, the effects are similar when the weights calculated by GoldVarb are compared. The hypothesis of “cê” working as a form of avoidance was also confirmed, because when there was concern about the use of pronouns, the forms chosen were either “senhora” or “cê”. In the analysis of the data, the geographic origin of the parents has emerged as an important variable in understanding the emergence of “tu” in the speech of Brasília: the Northeast region is shown as an essential factor in this process, but in Vila Planalto, where parents are natural from Brasilia, this factor has taken the role of expansion and fixation of “tu” as one of the choices of second person pronouns. The results also indicated that the speech of Brasilia, although still incipient in relation to having a typical linguistic variety, has developed characteristics that shows it has its own orality, and “tu”, as a marked form, is part of this dialect focalization. Beyond the social forces and interactional analysis, internal forces, common to other studies of second person pronouns, were found to be relevant in understanding this variation process, among which we highlight the syntactic function, the type of reference and linguistic parallelism. In the complex interplay between on the forms você/ cê/ tu, we observed that the statistical approach that better accounts for the use of these variants is the one which considers the three of them individually, that it is the n-ary analysis, because, depending on the forces in play, “você” and “cê” oppose “tu”, or “você” and “tu” are opposed to “cê”, or “você”, “tu” and “cê” present different behavior.

1. INTRODUÇÃO

1.1 Breve contextualização sobre o meu encontro com a pesquisa linguística

Entrei para a graduação em Letras Português do Brasil como Segunda Língua em 2003, na Universidade de Brasília (UnB). Foi no meu segundo semestre dentro da universidade (já em 2004), depois de cursar a disciplina Lexicografia, Semântica e Pragmática Contrastiva, com a professora Orlene Lúcia Sabóia de Carvalho, que eu tive meu primeiro contato direto com a pesquisa linguística, quando, ao final do semestre, a professora Orlene me indicou para a professora Maria Marta Pereira Scherre, que estava selecionando dois alunos de iniciação científica. A professora Marta Scherre me ligou e nos encontramos em uma sala no Departamento de Linguística, na UnB. As portas da pesquisa se abriram para mim naquele momento.

Como aluna de iniciação científica, participei, durante três anos e meio, da pesquisa sobre a variação do imperativo no português brasileiro, encabeçada pela professora Marta Scherre. Nesse período aprendi muito, a pesquisa me trouxe bons e muitos frutos. Foi natural me apaixonar pela Sociolinguística diante da oportunidade que tive. Natural, também, foi querer continuar na área, mas agora estudando características mais específicas da fala da minha cidade, Brasília¹.

1.2 Principais quês e porquês da presente pesquisa

Definir é Limitar
(Oscar Wilde)

O presente trabalho é uma análise sobre a variação que ocorre em Brasília com relação aos usos dos pronomes de segunda pessoa do singular, quais sejam: o *senhor(a)*; *você*, *cê* e *tu*. Através dessa análise, pretendemos contribuir com as seguintes questões gerais: 1) mapeamento geográfico dos usos de pronome de segunda pessoa no Brasil; 2) acompanhamento desse fenômeno em Brasília, dando continuidade às pesquisas desenvolvidas por Andrade (2004), que estudou a variação entre as formas *você/ ocê/ cê*; Lucca (2005), que estudou a variação entre *tu/ você*; e Dias (2007) que estudou a

¹ Havia 3 possibilidades de escolha para investigação de mestrado: 1) a variação no imperativo em um *corpus* do Amapá; 2) comparação do imperativo nas línguas portuguesa e espanhola; 3) variação pronominal em Brasília.

mesma variação que Lucca, mas em três faixas etárias diferentes; 3) discussão a respeito da hipótese de haver, em Brasília, um estágio de focalização dialetal, ou seja, gostaríamos de saber se já podemos dizer que existe uma variedade brasileira com traços linguísticos identificadores em formação na Capital. Sendo assim, gostaríamos de dar uma modesta contribuição para o desenvolvimento da Sociolinguística no Brasil e, em especial, em Brasília.

No ano de 1997, em uma viagem a Porto Seguro (lugar muito visitado por brasilienses), em época de pré-carnaval, eu, em meio à folia, conversei por algum tempo com um rapaz nativo (que vendia bebidas por ali) e, a certa altura da conversa, ele me perguntou: “você é de Brasília?” (até então, quando eu viajava, a pergunta corriqueira era “de onde você é?”). Eu, naquele momento, mesmo sem ser linguista, mas com a curiosidade natural do brasiliense, respondi perguntando: “sou, como você sabe?”, e ele: “pelo sotaque!”. Mais intrigada que antes, perguntei: “como é o meu sotaque?”, e ele respondeu: “é sem sotaque!”. Essa observação “sotaque sem sotaque” é bastante comum, embora o reconhecimento do sotaque brasiliense não o seja.

Assim, temos, a partir da narração anteriormente colocada, que: 1) já há o reconhecimento do “sotaque brasiliense” (situação rara, mas possível, principalmente em regiões muito visitadas por brasilienses); 2) a característica mais expressiva desse sotaque é a falta de sotaque (ou ausência de características marcadas); 3) o brasiliense se interessa pelo seu modo de falar. Tais constatações remetem a uma ideia mais geral: quando nos assumimos como comunidade distinta, necessitamos afirmar nossas diferenças, contrastando-as com as demais comunidades. E fazemos isso, principalmente, através de características culturais, como, por exemplo, por meio de uma variedade linguística com características próprias: é uma questão de identidade. Pagotto (2004, p. 89) teoriza: “(...) É ao se colocar como igual a um outro ou diferente dele, que o sujeito imprime significado social às formas linguísticas”.

Em meados de 2004, lembro-me de ter ouvido meu cunhado (brasiliense, com, na época, 17 anos) usando o *tu* em suas conversas. Naquele momento eu estranhei o fato, (acho que por isso mesmo me lembro bem!). Àquela época, mantinha meus ouvidos abertos por conta das descobertas recentes sobre variação linguística e, sem dúvida, o uso do pronome *tu* era de se estranhar, pois o brasiliense não tinha, até então (no meu conhecimento), o costume de usar essa forma pronominal de segunda pessoa. Depois disso, o que observava era que o *tu* foi começando a ficar mais presente na boca dos brasilienses, fato que era registrado em trabalhos variacionistas desenvolvidos na

universidade. Muito antes de pensar em trabalhar com esse fenômeno, eis que eu, brasileiro, comecei a observar o *tu* em minha própria fala também.

É curioso como a fala casual, espontânea, não é percebida por quem a usa. Minha cunhada, Maria Carolina, foi minha parceira em um jogo, em meados de janeiro de 2010. Durante o jogo, ela só se referiu a mim pelo pronome *tu* (nessa altura eu já estava habituada a ouvir o *tu* na boca dos brasileiros, sem comentar tal fato). Mas eis que a certa altura do jogo, surgiu o assunto a respeito do que eu estava desenvolvendo na minha pesquisa. Ao relatar a variação, minha cunhada se pronunciou revelando que não usava o *tu*. Para sua surpresa, eu lhe disse que ela não havia usado outro pronome durante aquele jogo e, depois disso, ela percebeu que o utilizava, pois o *tu* quis sair de sua boca mais algumas vezes durante aquela tarde.

É importante dizer que, apesar de ser comum em Brasília (pelo menos entre pessoas de faixa etária acima de 30 anos), pessoas que usam, mas frequentemente não se sentem usando o *tu*, ou seja, embora haja essa “negação” sobre o uso dessa forma, não parece que o *tu* seja um pronome estigmatizado, ainda que seu uso, em determinadas faixas etárias e contextos, possa indicar muitas vezes uma atitude de “casualidade, brincadeira, ironia” (cf. Dias, 2007, p. 74-90).

Lucca (2005) foi pioneira em registrar o uso do *tu* em Brasília, em dissertação de mestrado. Antes disso, porém, essa forma, na academia, era vista como pertencente a outras localidades. O *tu* era sentido como “algo de fora”, poderia ser estranhado em Brasília, pois mesmo os imigrantes originários de localidades onde o *tu* ocorre com maior frequência tendiam, morando em Brasília, a usar a forma *você*, por conta da difusão dialetal (falaremos de difusão e focalização dialetal mais adiante). Hoje, porém, é fato que tal variante está em expansão na fala dos brasileiros, ainda que alguns grupos de brasileiros não a percebam.

A seguir, apresentaremos dois trechos transcritos² de variação pronominal na fala brasileiro, extraídos de uma gravação realizada em 2008, da conversa de um brasileiro -B; 14 anos, sexo masculino, com um candango -P; 13 anos, sexo masculino, piauiense que chegou em Brasília aos dois anos, ambos moradores da Vila Planalto. No primeiro trecho, o assunto é um jogo de xadrez; no segundo, os informantes estão

² A transcrição foi feita originalmente de forma mais próxima à transcrição fonética. Ao registrarmos trechos dessa amostra, porém, buscamos apresentá-los em linguagem mais aproximada à forma escrita. Alguns trechos apresentam a palavra reduzida {init.} que significa ininteligível.

falando de namoradas e de um outro rapaz. Os pronomes analisados (*VOCÊ, CÊ, TU*) estão em letras maiúsculas.

B: VOCÊ não pó ah, tipo, tem, ah, vamos supor, a torre, ela tá na frente do seu rei, CÊ num tem pra onde saí, CÊ num pode fazer o, o roque, VOCÊ tem que encontrar um jeito de sair do cheque, tendeu? Porque senão VOCÊ se fôdis, né? Vamos jogar agora?

P: Bora!

B: E aí, véi... Quantos anos CÊ tem mesmo?

P: Treze. (...)

B: Ah, num pego ninguém da escola não, doido... Só pego de fora... tô dizeno... Ah tá, sou TU não, uô... Museu! Moleque chato, véi, esse Ulisses... /é/ véi

P: É doido, vou meter porrada nele, aí.

B: TU e mais quantos? TU num bate nem nuã formiga, véi (...)

O -P é o único informante da amostra que é candango. Os candangos são os brasileiros que não nasceram em Brasília, mas a adotaram como sua cidade. Pode-se considerar que tal informante adotou não só Brasília, mas também o modo de falar brasileiro. A esse respeito, Labov (2008, p. 352) teoriza: “Quando se mudam para uma nova área dialetal aos 3 ou 4 anos, as crianças parecem adotar os padrões básicos da área nova”. Consideramos, portanto, os exemplos colocados como falas autenticamente brasileiras.

Os dois trechos colocados anteriormente ocorreram em momentos diferentes da conversa. No primeiro, os informantes iniciavam a conversa e estavam claramente seguindo algumas orientações da pesquisadora, como ensinar a jogar xadrez. No outro momento, os informantes já estavam mais à vontade, sem demonstrar que se preocupavam com a gravação.

Os trechos anteriormente colocados fazem parte do *corpus* da presente pesquisa e são, até o momento, exemplos prototípicos dessa variação. Apontam, a exemplo de pesquisas anteriores a esta, para a importância dos contextos, da idade, do sexo, entre outros fatores que influenciam a ocorrência dos pronomes em estudo.

Convém, no entanto, deixar claro que, apesar dessa influência anteriormente mencionada ocorrer em função de fatores linguísticos e sociais, as formas variantes devem ocorrer com o mesmo valor de verdade nos mesmos contextos de fala. Ao

abordar esse assunto e sobre o que vem a ser uma possível variação linguística, Oliveira (1987, p. 22) afirma: “Admitimos, então, que as restrições em termos de contexto e valor de verdade são indispensáveis”. Em outras palavras, para ser um caso de variação linguística, as variantes têm de poder ocorrer de forma variável nos mesmos contextos linguísticos e sociais, mantendo o mesmo significado referencial. Temos que, excetuando a restrição da forma *cê* não ocorrer em função sintática de objeto, as três formas *cê/ você/ tu* ocorrem de forma variável, nos mesmos contextos e com o mesmo valor de verdade ao longo de nossa amostra. A seguir, apresentaremos exemplos de trechos da amostra com as três variáveis. Esses exemplos são de quatro diferentes gravações. O primeiro é um trecho de um menino de 14 anos; o segundo é de uma menina de 13 anos; o terceiro e o quarto são trechos de duas meninas de 14 anos. Cada dado em caixa alta deve poder ser substituído pelas demais formas variantes sem alterar o significado para que se possa dizer que são formas linguisticamente variáveis nos termos da sociolinguística laboviana.

B: ô mas o Pablo, TU é burro demais, num sei nem como CÊ conseguiu chegar sexta-feira, doido...

R: desse jeito a <Claudia> vai me encarar a Daniela, e eu também encaro, eu não quero nem saber! Aí {init.} eu perguntei pra Geovana, e a Geovana falou bem assim, acho que é inveja e ciúmes de VOCÊ tirar nota mais alta {init.} e porque TU fala mais com o Dimas e ela não fala, eu acho que é por isso que ela tem ciúmes de TU.

J: {init.} um exemplo, se TU, vou até {init.} pra dizer que VOCÊ, TU é eu e ele é TU, não, TU é eu, aí nós tava brincando e ele pede pra ficar com com, comigo, no caso que é TU, TU vai fazer o quê?

I: não a pequena, a, a Crislaine. Aí tá, aí eu entrei lá pra dentro aí eu falei “Rosana, eu vou ficar ali dentro um pouco”. Aí a Rosana “tá bom”. Aí a Crislaine “mãe, eu vou com a Ingrid”. Aí a Rosana “não, CÊ vai ficar aqui”. Aí tá, aí eu peguei, eu fiquei lá dentro conversando com o Alison, se rolar/ e a Rosana?/ lá fora, ela tava já chapada/ ela num viu o Alison não?/ Viu, mas ela nem pensou que tava nós dois aqui não. Porque o portão, né, tipo desses portão fechadão?! Aí tava com a porta fechada o tempo todinho, aí nem, aí teve uma hora que a porta tava aberta e nós entramos, aí tá, a gente fomos lá pro canto, bem no cantinho mesmo, num dava nem pra vê a gente. Aí toda hora, quando alguém vinha, aí minha prima avisava. Aí tá, aí minha prima o tempo todinho “Ingrid, CÊ num vai vim mais pra aqui pra fora não?” Eu falei “oxi, eu to lá dentro fazendo nada”.

Aí tá, ia depois tava eu e, a, aí eu fui ajudar minha, a minha tia a dar banho na minha prima pequenininha, aquela bebezinha?! Aí tá, aí eu demorei lá dentro, né, porque até dar banho, até trocar de roupa e tudo, aí minha prima chegou “Ingrid que demora é essa!” eu “minha filha, eu tava dando banho na, na Isabele, num vem reclamar de nada não, VOCÊ num é minha mãe nem nada, VOCÊ não é minha mãe nem nada e tal”. Aí, aí ela bem assim “ah, mas pra que demora isso tanto?” Eu falei “minha filha, quem demorou fui eu, oxi. TU num tá tomando conta de mim, eu vim sozinha num vim num foi contigo”.

Decidimos estudar a fala de crianças e adolescentes nas idades entre 7 e 15 anos, de ambos os sexos, para verificarmos, principalmente, se o uso da variante *tu* ocorre de forma semelhante aos resultados encontrados por Lucca (2005). Lucca pesquisou a faixa etária imediatamente posterior, de 15 a 19 anos, e encontrou elevado uso da variante *tu* em sua amostra. Gostaríamos de medir se há diferença da ocorrência desse pronome em nossos resultados. Ou seja, gostaríamos de investigar se é em tenra idade, ou na faixa etária entre 7 a 15 anos, que a variante *tu* entra no dialeto brasiliense, fase esta que se caracteriza por ser de pleno desenvolvimento linguístico. Caso não encontremos uma elevada ocorrência dessa variante em nossa amostra, tal resultado endossará a hipótese de forma localizada na idade (pelo menos nos limites de nossa investigação). Gostaríamos também de investigar se o *tu* também está em expansão na fala das pessoas do sexo feminino. Caso a hipótese de aquisição do pronome *tu* se confirme, essa tendência será naturalmente esperada.

Por último, gostaríamos de verificar a seguinte questão específica: se o *tu* e o *cê* concorrem entre si, em função do *tu*, por ser tônico, ter maior capacidade de chamar atenção do interlocutor, nos termos de Paredes (2003, p. 164-165):

Como se sabe, a forma **você**, originária da forma de tratamento Vossa Mercê, tem passado por sucessivas reduções, que culminam nas formas **ocê** no dialeto mineiro (cf. Ramos, 1997) e **cê**, na fala carioca, sendo que esta última já demonstrou até ultrapassar em frequência de uso a forma plena (cf. Paredes Silva, 1998). Através do pronome **tu** o falante estaria resgatando um monossílabo tônico para competir, com vantagem, com o clítico **cê** na função de atrair a atenção do interlocutor e compensar a perda de corpo fônico que se vem assinalando.

Se confirmada essa questão, nesse universo de dados, poderíamos inferir que o *cê*, nessa relação, pode servir à fala como um tipo de pronome de esquiva (ou coringa), já que sua forma é mais opaca em oposição ao *tu*, sendo este mais marcado (e semelhantemente informal). Mas ficaria a dúvida: será que o *cê* poderia ocorrer para substituir (nas posições que ele pode ocupar) qualquer forma pronominal de segunda pessoa? Como forma opaca ou não marcada, o *cê* poderia surgir no lugar de: *senhor(a)* (forma extremamente formal); *você* (que é a substituição mais natural e de maior ocorrência, uma vez que o *cê* é uma redução do *você*) e *tu*, principalmente em contextos em que essa variante não ocorre – em função de características diatópicas, por um lado, e, por outro lado, em circunstâncias tão formais que restrinjam o surgimento dessa variante mais marcada?

Apresentamos até aqui as vias pelas quais a presente pesquisa surgiu e também os principais questionamentos que orientam a presente investigação. A seguir, no segundo capítulo, discorreremos sobre a localidade estudada: Vila Planalto – Brasília (aspectos geográficos, históricos e sociais); no terceiro capítulo, faremos a apresentação da teoria e da metodologia; no quarto, teremos a revisão bibliográfica; no quinto, haverá a análise dos resultados e, por último, teremos as considerações finais.

2. O LOCUS

*“We are one, but not the same”
(U2)*

Antes que seja esclarecido o porquê de focalizar a localidade que estudamos, a saber, a Vila Planalto, inserida na região administrativa de Brasília, Distrito Federal, precisamos contextualizar a cidade em foco, dizendo, por exemplo, como surgiu a Vila Planalto, como são seus habitantes, como é a sua socioeconomia, ou seja, é necessário que nosso olhar enxergue essa cidade dentro da grande Brasília³, visualizando os meandros de seus aspectos sociais. Assim, a explanação a seguir trata de aspectos históricos, geográficos que moldaram a sociedade pertencente à Vila Planalto, em Brasília. E, na sequência, explicaremos o porquê da escolha dessa cidade como foco da presente investigação.

2.1. Vila Planalto, Brasília

Marquês de Pombal, no século XVIII, foi o primeiro ilustre a mencionar a necessidade de interiorização da capital do Brasil. Muitos visionários, entre eles o padre Dom Bosco, tiveram, no século XIX, o mesmo “sonho”: o de ter a capital no centro do país. A realização desse desejo tinha inúmeras vantagens, entre as quais se destacam: retirar o Distrito Federal da zona litorânea, por ser este lugar teoricamente muito vulnerável; trazer o poder político para o centro do país, colocando-o, de maneira até mesmo simbólica, como “igualmente acessível” frente às demais regiões; e propiciar o desenvolvimento do interior do país.

Porém, o que foi efetivamente realizado de importante na direção desse desejo comum durante aproximadamente cem anos foi: 1) a Expedição Cruls, que tinha como objetivo a elaboração de um relatório sobre o Planalto Central, onde se estabeleceria a futura capital; 2) a criação da lei que estabelecia a mudança da capital para o interior do Brasil expressa em duas Constituições, a de 1891 e 1934. A seguir, temos um mapa elaborado pela Expedição Cruls.

³ Temos duas maneiras de nos referirmos a Brasília: 1) Brasília, que se confunde com a região administrativa Brasília, que é o Plano Piloto em forma de avião e seus arredores; 2) grande Brasília, uma visão mais abrangente de Brasília, que extrapola os limites do Plano Piloto e arredores, pois abarca também as demais regiões administrativas e que se confunde com a ideia de Distrito Federal. No presente estudo usaremos simplesmente Brasília para a primeira definição e grande Brasília para a segunda.

Figura 01 – Expedição Cruls



(fonte: <<http://www.gdf.df.gov.br/045/04501020.asp>>)

Finalmente, no ano de 1956 do século XX, o então Presidente da República Juscelino Kubitschek iniciou seu mandato realizando esse sonho da construção da nova capital no Planalto Central do país, que ficaria localizada em uma região geográfica inserida no Estado de Goiás. Sobre o destino da nova capital, Kubitschek (extraído do site: <<http://www.gdf.gov.br/045/04501021.asp>>) relatou:

Coube a Brasília uma tarefa bem mais profunda e de muito maior alcance: a de puxar para o Oeste a massa a povoar o Brasil igualmente e, através desse empuxo migratório interno, realizar, quando muito no período de duas décadas, a verdadeira integração nacional.

A construção de Brasília foi uma grande empreitada arquitetônica, social e política. Os desafios eram faraônicos: construir uma nova cidade no período de três anos e meio, moderna, que pudesse atender aos anseios de uma nação tão mesclada como a nossa e que ainda pudesse se tornar a nova capital para um “novo” Brasil. Havia, como podemos notar na citação anterior, uma ideologia por detrás da construção da cidade

que era relativa ao progresso, ao desenvolvimento econômico e industrial de todo um país⁴.

Para a construção de Brasília, muitas providências foram evidentemente necessárias. Houve contratação de engenheiros, arquitetos e urbanistas, dentre os quais o nome mais famoso, em torno da construção da cidade, é o do arquiteto Oscar Niemeyer. Além das pessoas envolvidas na construção, houve também a contratação de empresas de engenharia que construiriam a cidade. O Governo Federal contratou várias dessas empresas para a execução da grande obra. Entre as mais famosas estavam: Pacheco Fernandes, Tamboril, Rabelo, e DFL (Departamento de Força e Luz). Para abrigar os engenheiros, técnicos e peões dessas empresas, foram criados vários acampamentos bem no centro da futura cidade (até mesmo para agilizar o andamento da obra). Esse aglomerado de barracões ficou conhecido mais tarde como Vila Planalto.

Os acampamentos tinham o nome das construtoras, os barracos eram de madeira, e havia ali alguma infraestrutura, como um cinema, o clube do Departamento de Força e Luz, o campo de futebol da construtora Rabelo e pequenas lojinhas de comércio. Os barracos e essa infraestrutura seriam demolidos depois da inauguração de Brasília. Mas, como muitos operários passaram anos na construção e vivendo nos acampamentos, eles começaram a constituir família ali, e isso veio modificar o que viria a seguir.

Depois da inauguração de Brasília, durante os anos 70 e 80, o governo local tentou inúmeras vezes remover o acampamento, pois a Vila Planalto não estava no projeto original da cidade. Diante das dificuldades em regularizar essa situação, muitos moradores da Vila aceitaram propostas de se assentarem em outras regiões administrativas, as chamadas cidades satélites (que atualmente são chamadas de regiões administrativas). É importante frisar que a Vila está localizada em uma área valorizada da cidade, próxima à Esplanada dos Ministérios, no meio do Setor de Clubes, à beira do Lago Paranoá e perto das embaixadas (cf. mapas no final do capítulo). Assim, havia muitos interesses para que essa área fosse destinada a construções de clubes ou que fosse vendida para compradores que pudessem pagar o preço de uma área nobre (durante o governo de José Ornellas houve um projeto de vendas dessa área).

Por outro lado, no fim da década de 80, alguns moradores da Vila se organizavam na tentativa de fixar a Vila Planalto como uma região administrativa e de ter a posse da terra. Um dos grupos fundamentais na tentativa de fixação da Vila foi o

⁴ É importante lembrar que a construção de Brasília se deu num período posterior à grande depressão mundial (relativa à economia).

grupo das 10. Esse grupo se reunia na tentativa de encontrar meios de lutar pela área. Em um desses encontros, a filha de uma das integrantes do grupo, Leiliane Rebouças (com 10 anos na época), teve a ideia de escrever uma carta para o então Presidente da República José Sarney. E foi isso o que aconteceu: em agosto de 1987 a menina conseguiu entregar a carta ao presidente. No dia seguinte, Leiliane recebeu resposta a sua carta: Sarney encarregara o então Governador de Brasília, José Aparecido, a resolver o problema. No dia 27/08/1987, o governador se reuniu com a menina e outros interessados, e se comprometeu a lutar pelo tombamento de Brasília como patrimônio histórico e a incluir a Vila Planalto, pois este era o único local que ainda mantinha características da época da construção. No dia 21 de abril de 1988, o governador assinou o documento do tombamento de Brasília e da fixação da Vila Planalto.

Nasce assim a Vila Planalto, ou melhor, ganha “certidão de nascimento” com direito de existir. Porém, esse “bairro” mais pertence ao governo que aos seus moradores, pois os moradores da Vila receberam, em 1988, somente o direito de ali viver. A posse funciona apenas como uma autorização para usufruir o local. Eles, até os dias atuais, não receberam a escritura de suas casas. Mesmo assim, a especulação imobiliária cresceu bastante nas últimas décadas. A CODEPLAN (Companhia de Planejamento do Distrito Federal) estipula que, nos dias atuais, apenas 56% dos moradores receberam a posse diretamente do governo. Por outro lado, na mesma pesquisa, a informação é de que 25,5% dos residentes da Vila Planalto têm entre 40 e 59 anos e somente 11,7% estão acima de 60 anos de idade. Ou seja, uma parte desses moradores herdou a posse de seus pais, os pioneiros, e outra parte dessa população comprou a posse através de “contratos de gaveta” (expressão usada para contratos de compra e venda que têm valor legal, mas não são documentos definitivos dos imóveis, nem poderiam ser, já que ali não há imóveis escriturados) de outros candangos/brasilienses.

Os moradores da Vila Planalto ou, pelo menos, os moradores que frequentam a escola pública da Vila Planalto nos dias atuais são, em boa parte, filhos ou netos das gerações de pioneiros da cidade. Foram esses alunos da escola pública da Vila Planalto que participaram da presente pesquisa. A origem dos seus avós ou pais pode ser mais bem definida socialmente que regionalmente: poucos engenheiros se estabeleceram ali. A maior parte da população da Vila, originalmente, foi composta por operários e seus familiares, seguido das pessoas que viabilizavam a estrutura local, como pequenos comerciantes ou pessoas que trabalhavam com gêneros alimentícios, por exemplo.

A seguir, a tabela apresenta as origens detalhadas dos moradores da Vila Planalto. Esta tabela foi extraída de um relatório de pesquisa domiciliar e socioeconômico realizada na Vila Planalto em 2009, pela CODEPLAN:

Tabela 01 – Origem dos moradores da Vila Planalto, segundo a CODEPLAN (2009)

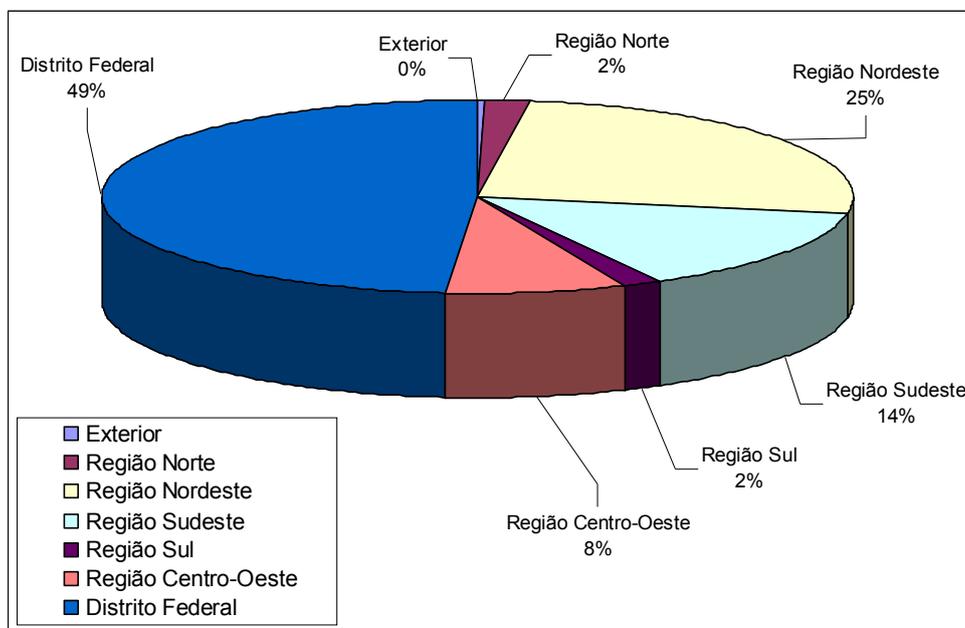
Unidade da Federação	Número	Percentual	% / Região
Distrito Federal	4284	56,1%	DF 56,1%
Mato Grosso	10	0,1%	Centro-Oeste 4,7%
Mato Grosso do Sul	24	0,3%	
Goiás	331	4,3%	
Acre e Amapá	-	-	Norte 1,4%
Amazonas	5	0,1%	
Roraima	5	0,1%	
Rondônia	10	0,1%	
Pará	29	0,4%	
Tocantins	54	0,7%	
Maranhão	214	2,8%	Nordeste 24,9%
Piauí	336	4,4%	
Ceará	395	5,2%	
Rio Grande do Norte	102	1,3%	
Paraíba	263	3,4%	
Pernambuco	297	3,9%	
Alagoas	5	0,1%	
Sergipe	29	0,4%	
Bahia	258	3,4%	Sudeste 10,9%
Minas Gerais	618	8,1%	
Espírito Santo	29	0,4%	
Rio de Janeiro	131	1,7%	
São Paulo	54	0,7%	Sul 1,9%
Paraná	44	0,6%	
Santa Catarina	24	0,3%	
Rio Grande do Sul	73	1,0%	
Total	7.634	100%	100%

Fonte: CODEPLAN: Levantamento Domiciliar Socioeconômico, Vila Planalto, 2009 – com adaptações.

A grande Brasília, nos dias atuais, é uma cidade que vive entre a sua segunda e terceira geração (a terceira geração é constituída por crianças ou adolescentes) já que

uma considerável parte dos seus habitantes nasceu aqui⁵. Segundo informações de 2007 (contidas no *site* <www.gdf.gov.br/sites/200/216/00000327.pdf>) da CODEPLAN, a população de Brasília é composta das origens segundo o gráfico 02 a seguir:

Gráfico 01 – Origens da população brasiliense



A Vila Planalto segue a mesma tendência de origens apresentada pela grande Brasília. Como aponta a tabela anteriormente colocada, 56,1% da população da Vila é composta de brasilienses. Já a grande Brasília, como apontou o gráfico acima, tem em sua população 49% dos habitantes composta por brasilienses.

Pensando ainda na primeira geração, os pioneiros, como demonstram a tabela 01 e o gráfico 01 (e a exemplo da história de Brasília), estes vieram de todas as partes do país. Ao perguntar a alguns moradores tão antigos ali quanto a própria Vila, a sensação generalizada entre eles é de que sua população foi formada em maioria por nordestinos. Podemos inferir que eles estão certos. Embora o estado de Minas Gerais seja o que mais tenha mandado gente para a Vila, 8,1%, se considerarmos em bloco, a região de maior representatividade na Vila Planalto realmente é a Nordeste, com uma população de 24,9% (apenas 10,9% das pessoas ali são originárias da região Sudeste). Essa distribuição é bastante semelhante ao que ocorre com o restante do Distrito Federal, que tem 25% da população com origem nordestina.

⁵ Informação inicialmente coletada em palestra de Bortoni-Ricardo, em setembro de 2009.

Segundo a pesquisa da CODEPLAN, vivem na Vila Planalto cerca de 7.634 habitantes. A sensação dos moradores quando questionados é de que a Vila cresceu na última década, mas não muito se comparada ao restante do Distrito Federal.

Vale dizer que a grande Brasília, até os dias atuais, é uma cidade receptora de imigrantes do Brasil inteiro. Diversas características cooperam para essa realidade, entre elas: 1) ser a capital da República; 2) ter uma demanda crescente de empregos por conta da expansão da cidade; 3) possuir políticas locais que trazem (já traziam desde meados da década de 90) alguma qualidade de vida aos menos favorecidos, tais como: bolsas escola e família, entrega de mantimentos (como leite e pão), distribuição de lotes, entre outros. Vale dizer que tais políticas se tornaram, naturalmente, um atrativo para brasileiros socialmente menos favorecidos⁶. Porém, mesmo com essa realidade, a Vila Planalto, apesar de ser parte de Brasília, é uma das poucas localidades em que não se observa uma expansão da cidade, nem mesmo uma intensa mudança de moradores. A Vila, na realidade de Brasília, é uma das localidades mais tradicionais, no sentido de se manter como original.

2.2. A fala da Vila Planalto como objeto de estudo

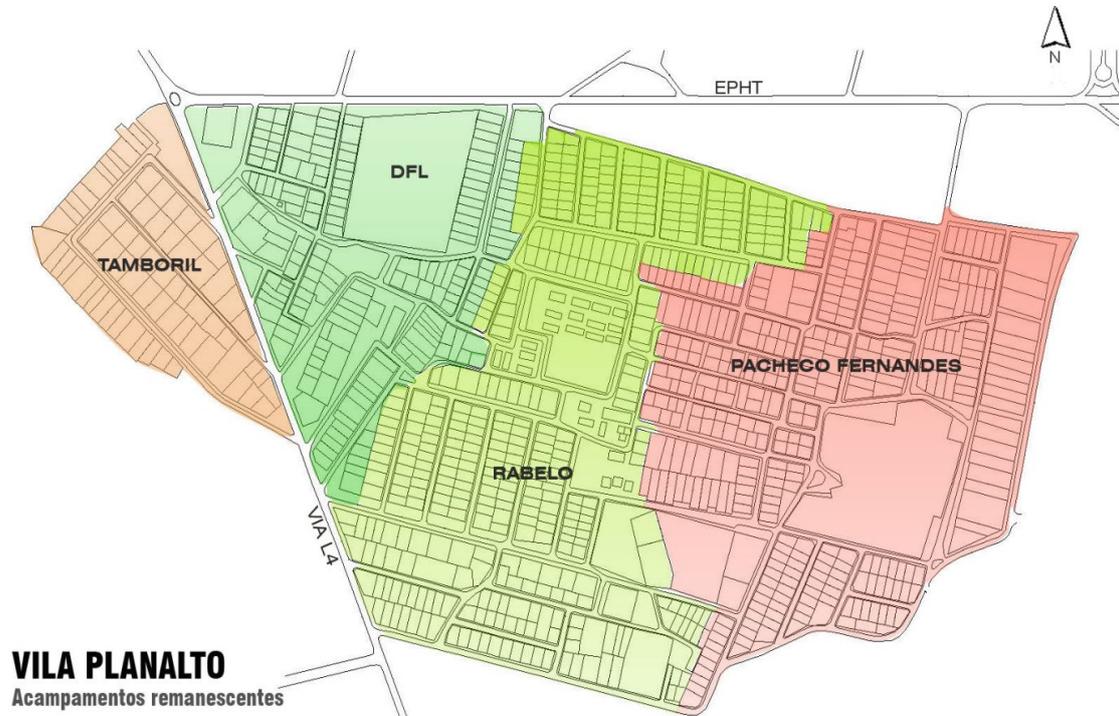
A fala dos brasilienses da Vila Planalto se caracteriza por ser uma excelente fonte de estudos, porque ali existe um substrato da variedade que está se formando em Brasília. A Vila é, sem dúvida, uma das cidades mais antigas de Brasília, pois surgiu em 1958. E, como houve relativamente uma considerável manutenção de sua população, podemos entrever na fala da população da Vila Planalto os comportamentos linguísticos de difusão e focalização dialetal que podemos encontrar em maior ou menor grau em toda a grande Brasília.

Não só a Vila Planalto, mas Planaltina, Núcleo Bandeirante e Taguatinga são localidades, ou regiões administrativas, mais antigas que a região administrativa de Brasília (ou Plano Piloto). Mas o grande diferencial da Vila Planalto é mesmo sua incapacidade de expansão, por estar no centro de Brasília (nos dois níveis: grande Brasília e região administrativa) e a sua já mencionada manutenção da população. A seguir, teremos a apresentação de alguns mapas: da Vila Planalto, de Brasília (que traz

⁶ E, por ser esse tipo de política altamente recorrente durante aproximadamente duas décadas, é possível notar uma cultura relativa ao paternalismo (e até mesmo ao coronelismo) que se criou em Brasília em torno de tais benefícios.

circulada a Vila Planalto), do Distrito Federal ou o que temos chamado de grande Brasília, e do Distrito Federal dentro do Brasil:

Figura 02 – Vila Planalto



(fonte das figuras: <codeplan@codeplan.df.gov.br>)

Figura 03 – Brasília:

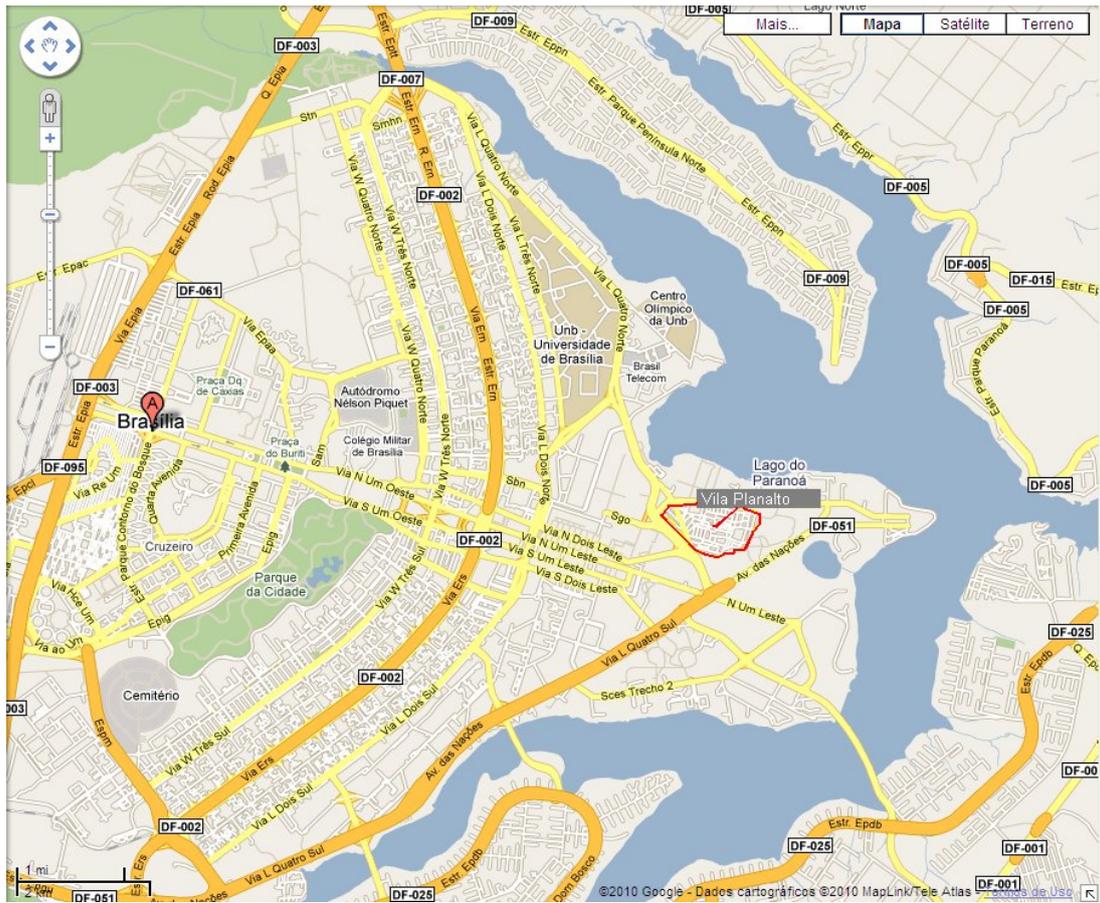


Figura 04 – Distrito Federal

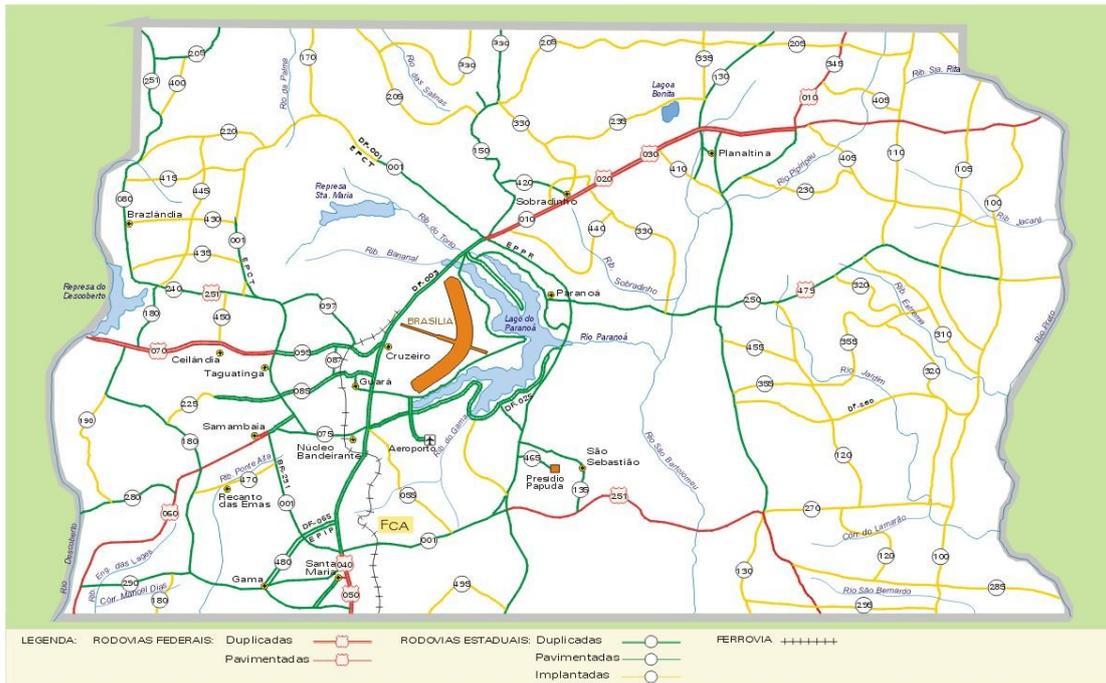
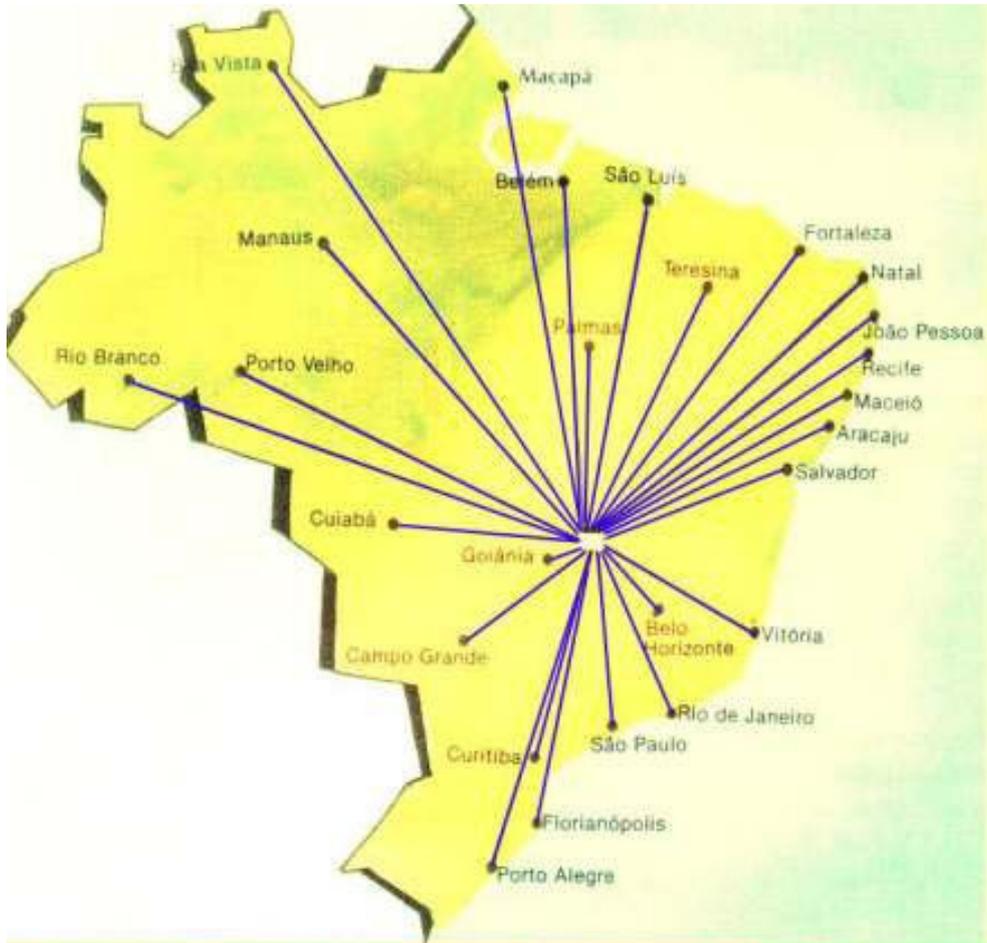


Figura 05 – Brasil



3. TEORIA E METODOLOGIA

A teoria adotada no presente trabalho é conhecida por Teoria da Variação e da Mudança Linguística ou Sociolinguística Quantitativa. Segundo Faraco (2006, p. 9-29), a teoria nasceu em um momento sócio-histórico em que a Linguística havia sido inicialmente dominada pelas ideias de Saussure (início do séc. XX) e, mais tarde, por Chomsky (por volta dos anos 1960 do séc. XX). Ambos privilegiavam os estudos linguísticos focados em fatos internos das línguas. Foi no Simpósio “Direções para Linguística Histórica”, realizado em 1966, nos EUA, que houve uma renovação dos estudos linguísticos, em que a diacronia e a linguística histórica puderam ser reintroduzidas na pauta dos estudiosos. Na década de 50 já havia surgido as primeiras pesquisas sociolinguísticas, como a tese de doutorado de Weinreich (1951). Mas foi a partir das pesquisas de Labov (1966) que a sociolinguística se estabeleceu como teoria e metodologia para a investigação da língua em uso, privilegiando não somente fatos internos, mas também fatos externos à língua.

A presente análise se baseia na teoria desenvolvida por Weinreich, Labov & Herzog (1968) de que as línguas são heterogêneas, e esta heterogeneidade é passível de ser descrita. Significa dizer que não só as línguas variam e mudam, como variam e mudam de acordo com motivações linguísticas e sociais. Temos assim que uma vocação natural da área seja, justamente, estudar como, quando, onde e por que as línguas variam e mudam. Nesse sentido Tagliamonte (2007, p. 3) teoriza:

A Sociolinguística tende a colocar ênfase na linguagem no contexto social. A análise da variação está incorporada à sociolinguística, a área da linguística que tem como ponto de partida as regras gramaticais e, em seguida, estuda os pontos que essas regras fazem contato com a sociedade.⁷”

⁷ Extraído do original: “Sociolinguistics tends to put emphasis on language in social context (...). Variations analysis is embedded in sociolinguistics, the area of linguistics which takes as a starting point the rules of grammar and then studies the points at which these rules make contact with society” (Tradução minha).

No mesmo sentido, Labov (2008, p. 21) teoriza:

não se pode entender o desenvolvimento de uma mudança lingüística sem levar em conta a vida social da comunidade em que ela ocorre. Ou, dizendo de outro modo, as pressões sociais estão operando continuamente sobre a língua, não de algum ponto remoto no passado, mas como uma força social imanente agindo no presente vivo.

Observemos o trecho do nosso *corpus* a seguir:

Rosana, primeiro TU não vai impedir a boca dela, TU não vai conseguir segurar ela nem nada, então de todo jeito ela vai ficar escondido com alguém, mesmo VOCÊ não sabendo, ela vai ficar (...)

A variação que ocorre no trecho acima (entre as variantes *tu/ você*) é condicionada principalmente por **fatores sociais** como sexo, idade, estilo de fala etc. É difícil hoje imaginar que a ocorrência do pronome *tu* na variedade brasileira, que é caracterizado pela realização com verbos (em grande parte das ocorrências) sem marcas morfológicas de segunda pessoa, fosse produto de erro dos falantes e que a variação entre *tu* e *você* ocorresse de forma caótica ou, ainda, que tal variação fosse produto de um movimento somente interno da língua. A esse respeito, Weinrich, Labov & Herzog (1968, p. 36) afirmam: “Um dos corolários de nossa abordagem é que, numa língua que serve a uma comunidade complexa (i.e., real), a ausência de heterogeneidade estruturada é que seria disfuncional”. E, ainda a esse respeito, Labov (2008, p. 20) afirma “(...) uma abordagem que considera apenas as pressões estruturais dificilmente pode contar a história toda. Nem todas as mudanças são altamente estruturadas, e nenhuma mudança acontece num vácuo social (...)”.

Sturttvant (*apud* Labov, 2008, p. 21) já indicava um caminho para compreensão sobre a **propagação** de um fenômeno de variação lingüística. Em suas palavras: “Antes que um fenômeno possa se difundir de palavra para palavra [...] é necessário que uma das formas rivais adquira algum tipo de prestígio”. Por isso, estudamos a fala real e

analisamos nessas falas o que ocorre tanto internamente como externamente para que elas sejam elas mesmas e não outras.

É importante observar que, embora a sociolinguística tenha colocado os fatos sociais no escopo dos condicionamentos da variação linguística, os fatos internos permanecem também como campo de investigação sociolinguística. Na presente análise, por exemplo, teremos o fator função sintática que medirá as pressões internas que influenciam a ocorrência das variantes *cê/ você/ tu*. Mais adiante retomaremos esse assunto.

Tendo como força norteadora do presente estudo a Teoria da Variação e Mudança Linguística, gostaríamos ainda de encaixar a presente pesquisa em alguns direcionamentos teóricos que se inserem dentro da teoria da variação. Assim, uma teoria que se aplica em nossa maneira de perceber o fenômeno de variação em estudo é o que a sociolinguística interacional concebe por **enquadres** (Goffman, 1998, p. 70-97) ou os **estilos contextuais** de fala, concebido pela sociolinguística laboviana. Infere-se de Labov (2008, p. 101-138) que os estilos estão imbricados em uma complexa rede de características, tais como: contexto de fala, assunto, gênero discursivo, entre outros. Segundo Labov, “o problema é controlar o contexto e definir os estilos de fala que ocorrem dentro de cada contexto (...)”. Labov (2008, p. 313) define a diferença do que vem a ser fatores sociais e estilísticos e, sobre isso, o autor teoriza:

Por social entendo aqueles traços da língua que caracterizam vários subgrupos numa sociedade heterogênea; e, por “estilística”, as alternâncias pelas quais um falante adapta sua linguagem ao contexto imediato de fala. Ambas estão incluídas no comportamento “expressivo” – o modo como o falante diz ao ouvinte algo sobre si mesmo e seu estado mental, além de dar informação representacional sobre o mundo. A variação social e estilística pressupõe a opção de dizer “a mesma coisa” de várias maneiras diferentes, isto é, as variantes são idênticas em valor de verdade ou referencial, mas se opõem em sua significação social/ estilística.

Goffman (1998) dá um exemplo bastante esclarecedor sobre os estilos (ou enquadres, segundo a sociolinguística interacional), quando narra uma história em que um famoso presidente, durante uma entrevista, assume a postura linguística de um chefe de estado respondendo às questões sobre governabilidade, mas, ao final da entrevista, por já conhecer uma jornalista ali presente, o presidente tece comentários jocosos a respeito das roupas que sua conhecida usava, abandonando a postura linguística anterior. No exemplo dado por Goffman, o personagem modifica seu estilo de fala (ou enquadre), pois este sempre dependerá do contexto de fala que os falantes estão vivenciando no momento da enunciação. Temos, dessa forma, que em todo ato comunicativo está presente um estilo, e este influencia os diversos usos que podemos fazer da língua, inclusive a seleção entre os fenômenos que estão em variação.

Como nossa análise se encaixa na faixa etária de 7 a 15 anos, outra base teórica que norteia nossa discussão é a semente lançada por Labov (2001, p. 447) **estendendo o período de aquisição da língua materna**, em que, até os cinco anos, a aquisição da língua materna ocorre sob a influência do dialeto dos pais, mas no estágio seguinte da infância, na pré-adolescência e na adolescência, essa influência vem principalmente dos pares. Labov (2008, p. 349) afirma que “as crianças não falam ‘somente’ como os pais. Na grande maioria dos casos que temos estudado ou encontrado, as crianças seguem o padrão de seus pares” (com adaptações). Na infância e adolescência o falante começa a perceber as diferenças entre as variedades a que tem acesso e se posiciona em relação a eles. Significa dizer que, em nosso *corpus* (que se insere em uma comunidade de fala), as crianças devem compartilhar valores sociais para a língua e para as variações linguísticas que utilizam, uma vez que suas influências avaliativas e linguísticas estão difundidas dentro da comunidade. Como exemplo dessa percepção pode-se apontar um resultado de um incipiente teste de percepção que realizamos entre alguns informantes da presente pesquisa. Perguntado sobre o uso da variante *cê*, um dos informantes respondeu da seguinte forma:

Não uso. Porque é coisa de roceiro.

Estudamos a língua falada em um local em que o contato linguístico entre variedades regionais é uma realidade. Esse contato remete aos conceitos de difusão e focalização dialetal. Segundo Milroy e Le Page (*apud* Hanna, 1986, e Corrêa, 1998), o dialeto difuso é um dialeto sem uma característica particular que não pode ser

reconhecido ou identificado, pois se caracteriza por uma mistura dialetal. A difusão dialetal ocorre, principalmente, quando há grande movimentação social ou geográfica, em que o falante abre mão de características muito específicas de seu dialeto original em favor de um dialeto convergente ou mais próximo do novo grupo ao qual pertence ou pretende pertencer. Vale dizer que fatores como prestígio ou identidade incidem no quanto os grupos mantêm ou perdem de seus dialetos originais. Já a focalização dialetal é o movimento inverso, ou seja, quando um dialeto pode ser identificado ou especificado uma vez que é divergente de outros dialetos. Temos, como exemplos de focalização no Brasil, as variedades carioca, gaúcha, mineira, entre outros.

Como uma de nossas hipóteses é medir se as pessoas do sexo feminino em Brasília (ou, especificamente, na Vila Planalto) estão expandindo a ocorrência da variante *tu* na fala, é necessário saber o que a teoria da variação sugere como interpretações possíveis sobre essa **diferença entre os sexos em relação à fala**. Labov (2008, p. 346-347) especula que “as mulheres usam as formas mais ‘avançadas’ em sua própria fala informal e se corrigem mais nitidamente no outro extremo da fala monitorada” (grifo nosso), ou seja, elas tendem a aderir (até mais que os homens) às mudanças quando o contexto de fala é mais informal, ao passo que são mais conservadoras quando o contexto é mais formal. Isso ocorre, segundo Labov, porque as mulheres “são mais sensíveis ao padrão de prestígio”. É importante frisar, porém, que o próprio Labov recomenda que essas sugestões de interpretação ainda são especulações. Assim, por serem especulações, necessitam de mais análises para que se chegue a um posicionamento mais consistente acerca da interpretação das diferenças que pode haver entre os sexos no uso linguístico. Vale dizer que Borges (2009, p. 32-34) estudou a questão do gênero em sua tese de doutorado. Nela, a autora demonstra que essa questão constitui um fator externo extremamente relevante no que concerne à variação linguística e que não há consenso entre os linguistas sobre as tendências de fala de um determinado sexo, se será mais favorecido em fala prestigiada ou menos, se algum dos sexos será mais favorecido em questões ligadas à identidade ou menos. Porém, há consenso sobre o fato de homens e mulheres terem a suas falas diferenciadas.

Hanna (1986) e Corrêa (1998) estudaram a fala de Brasília. As autoras realizaram pesquisas sociolinguísticas sobre características fonéticas na oralidade (ou realizações de pronúncias) e, a partir de seus estudos, concluíram que Brasília passa por uma fase de **focalização dialetal** em que sua principal característica é a falta de marcas regionais. Ou seja, segundo Hanna (1986, p. 151) e Corrêa (1998, p. 87), os brasilienses

tendem a focalizar um dialeto em direção a um dialeto suprarregional. A partir das análises aqui explicitadas, podemos inferir que a forma *tu*, por ser uma forma marcada no português brasileiro para referência da segunda pessoa e, ainda, por ser uma forma inovadora no “dialeto brasiliense”, poderia ser compreendida então como um aspecto da eminente focalização dialetal que ocorre em Brasília.

Outra teoria ainda mais específica e fundamental em nossa análise diz respeito às características de duas das variantes estudadas, *você* e *cê*: a gramaticalização da expressão *Vossa Mercê*. Essa expressão mudou ao passar do tempo, por motivos históricos e sociais que não abordaremos no momento. Mas, segundo Antenor Nascentes (1956), citando também Leite de Vasconcelos em seu “Dialeto Brasileiro”, tal expressão passou por várias formas. Algumas delas estão a seguir:

Vossa Mercê > *vossemecê* > *voss'micê* > *vossuncê* > *vassum'cê* > *vom'cê* > até chegar às formas de hoje: *você*, *ocê* (forma pertencente principalmente ao dialeto mineiro) e *cê*.

Temos que as principais características de gramaticalização são: **perda de massa fônica e esvaziamento de sentido**. Ou seja, a expressão cerimoniosa *Vossa Mercê* passou, nos dias atuais, a ser *você/ ocê/ cê* deixando de ter *status* honorífico, figurando, assim, como pronomes.

As teorias anteriormente colocadas servem de base para as reflexões que faremos sobre a amostra utilizada que se pretende como representativa da comunidade estudada. Em última análise, as teorias revisadas nos auxiliarão na tarefa de compreender e discorrer a respeito do que linguisticamente ocorre em Brasília (ou mais especificamente, com a amostra analisada).

Para a análise dos dados, os passos metodológicos foram: 1) constituição da amostra; 2) estudo sobre a comunidade de fala da qual a amostra faz parte; 3) levantamento dos dados em análise; 4) classificação dos dados segundo critérios sociais e linguísticos a fim de operacionalizar as hipóteses levantadas; 4) testes estatísticos dos dados através do programa *GoldvarbX*; 5) análise dos resultados.

3.1. Metodologia desenvolvida para constituição da amostra

Mesmo havendo diretrizes tradicionais para a constituição da amostra, optamos por constituí-la de forma alternativa, pois o fenômeno de variação pronominal que ocorre em Brasília dificilmente se mostraria em uma entrevista laboviana típica, uma vez que a variação entre as formas *você/ cê/ tu* é regida, principalmente, por fatores sociais que envolvem as diversas características e inter-relações entre os interlocutores. Assim, seria difícil fazer emergir os dados que buscávamos (que são mais típicos de conversas casuais ou informais) em uma entrevista do tipo laboviana, realizada com pessoas que não se relacionavam ou não se conheciam previamente e, mais ainda, que pertenciam a mundos sociais completamente díspares, como no caso desse estudo, em que o grupo de colaboradores ou informantes é composto de crianças e adolescentes de 7 a 15 anos. Sobre isso, Tagliamonte (2007) pondera:

Dependendo da natureza dos dados e da natureza da pesquisa sob investigação, diversos fatores serão importantes. Quem será estudado e por quê? Quais são os limites do grupo ou comunidade? Se a localização no espaço é importante, onde será? Se a localização no tempo é importante, quando isso será realizado?⁸

Continuando nessa mesma linha de raciocínio: se a forma de gravar os dados é importante, como realizar essas gravações? Decidiu-se realizar as gravações da seguinte forma: os colaboradores da pesquisa ficariam com o aparelho de MP4 para realizar gravações de conversas casuais com amigos. Para que fosse possível a gravação de conversas casuais (sem atenção para a forma com que se fala), a gravação seria feita sem o conhecimento prévio dos interlocutores. Ao término da gravação, porém, o informante (ou colaborador) revelaria aos interlocutores a gravação da conversa e pediria autorização para utilizá-la em pesquisa científica (respeitando o anonimato de todos os envolvidos). Porém, com as crianças de 7 a 11 anos, encontramos dificuldades

⁸ Do original: “depending on nature of the data and nature of the research question under investigation, different factors will be important. Who are you going to study, and why? What are the boundaries of the group or community? If location in space is important, where is it? If location in time is important, how will this be accomplished?”

em realizar as gravações nesse formato. Por isso, a metodologia desenvolvida para as gravações dessa faixa etária foi diferenciada. Mais adiante voltaremos a essa questão.

Seria ideal também que os informantes fossem escolhidos de forma aleatória, sorteados, de preferência, em uma lista telefônica, por exemplo. Mas nisso também buscamos uma forma alternativa, pois seria difícil contactar idades tão específicas de um modo aleatório. Como este estudo tem por base a coleta e a análise de dados de fala casual de crianças de 7 a 15 anos, foi natural optar por buscar parceria em uma escola. Entramos, assim, em contato com a escola pública da Vila Planalto para conseguir ter acesso às crianças que comporiam o nosso corpo de colaboradores (ou informantes).

Para contactar os alunos, houve a necessidade de autorização por parte do Conselho Regional de Educação e da escola. Uma vez vencida as etapas burocráticas, pudemos contactar a escola, mas é importante dizer que o acesso sempre foi condicionado aos interesses da escola como um todo.

Assim, os colaboradores de 12 a 15 anos foram escolhidos por uma única professora da escola Centro de Ensino Fundamental 01 do Planalto, Andréa Longo. O único critério que ela utilizou para a escolha dos informantes foi o grau de responsabilidade que os alunos já demonstravam em sala. A professora então, de posse do MP4, emprestava-o a um aluno, que, por sua vez, ficava com o aparelho (em média de três a quatro dias) até fazer uma gravação de conversa casual (sem – ou com o mínimo de – preocupação com a linguagem utilizada) com os amigos. Em seguida, o aluno devolvia o MP4 para a professora e ela o entregava à pesquisadora. Somente então a pesquisadora conferia a gravação e julgava se entraria ou não para a amostra.

Ao tentar realizar gravações com crianças de 7 a 11 anos nos moldes das gravações de crianças de 12 a 15 anos, tivemos problemas de ordens técnicas. Constatamos que as idades dos informantes e os graus de responsabilidade que estes apresentavam não eram compatíveis com a empreitada proposta. Por isso, buscamos outras formas de gravação para esta faixa etária. Para conseguir realizar essas gravações, muitas ideias surgiram. A primeira foi tentar que os professores gravassem esses alunos em sala, na hora da aula das crianças, para evitar ao máximo as consequências do paradoxo do observador⁹ e para evitar que fosse imposto aos informantes algo diferente do que eles já conheciam (pois, como já mencionado, os dados tendem a surgir mais em

⁹ O paradoxo do observador consiste na presença do pesquisador ou observador que pode alterar a fala do informante, no sentido desta ficar mais monitorada (cf., principalmente, Labov, 1966, 1972a e 1972b; e Tarallo, 1986).

ambientes linguísticos descontraídos), mas, neste contexto, assumiríamos que o estilo dos alunos poderia ser mais formal.

Depois de algumas tentativas frustradas, em que havia somente a fala da professora ou falas muito formais entre alunos e professora, decidimos que eu mesma faria as gravações, embora não fosse o ideal para esta pesquisa. Assim, foi permitido que eu me reunisse com um grupo de alunos, por um tempo limitado, cerca de uma hora, em uma sala de aula dentro da escola e em alguns dias definidos. Para diminuir o possível enviesamento dos dados, causado pela minha presença, decidi levar minha filha de 10 anos para que os alunos pudessem se sentir mais à vontade. Os critérios dos professores para a seleção dos alunos que se reuniriam com a pesquisadora foram sempre critérios de acomodação à aula (exemplo: crianças que já haviam terminado a atividade proposta em sala). Assim, geralmente o horário que eu tinha acesso às crianças era próximo ao final da aula, normalmente quatro crianças eram escolhidas pelas suas professoras, sendo, na maioria das vezes, duas meninas e dois meninos, e então eu e as crianças seguíamos para outra sala, onde a gravação seria feita. As gravações ocorreram da seguinte maneira: uma vez reunidos na sala designada para nosso encontro, eu (ou minha filha) lia um livro com os alunos e, na sequência, discutíamos tal livro. Depois disso podíamos conversar livremente sobre assuntos diversos. Apesar de sobrar pouco tempo para conversas informais, esse método se mostrou satisfatório, como teremos oportunidade de ver mais adiante.

Houve também, em nossa amostra, gravações realizadas com crianças oriundas de outras localidades. Essas crianças são do meu ciclo de convívio e, apesar de ser um grupo socialmente diferente do grupo da Vila Planalto, consideramos adequado ter registro de fala de outras localidades para servir como parâmetros de comparações.

3.1.1. Descrição das gravações: facilidades e dificuldades

Labov (2008, p. 63) argumenta que “os meios empregados para coletar dados interferem nos dados a serem coletados”. Como as crianças de 12 a 15 anos ficaram com o aparelho de MP4 para realizar as gravações, tais gravações se mostraram adequadamente casuais.

A instrução da pesquisadora para com os informantes foi de fazer gravações ocultas. Assim, o informante faria a gravação de uma conversa com os amigos e depois

revelaria a eles que a conversa havia sido gravada. Em seguida, pediria a autorização de seus interlocutores para a utilização dos dados ali coletados. Caso algum amigo não autorizasse, a instrução era de apagar a gravação. Nenhum informante, porém, seguiu essas instruções. Ao longo da escuta das gravações, ficou patente que todos os colaboradores, em todas as gravações, estavam cientes de estarem sendo gravados desde o início das gravações. Não houve, também, indícios de que algum colaborador tenha solicitado a destruição da gravação por motivos de manutenção de privacidade. Vejamos, a seguir, um trecho de transcrição em que o informante demonstra que a gravação não era oculta:

B: {init.} Pô, valeu, de... Dezesseis minutos estragado agora por causa que vocês falaram “é um mp4”
P: Não, não vou fazer tudo de novo não, véi
B: dezesseis minutos enrolando, sem assunto, e eu fazendo um assunto, dezesseis minutos de trabalho
*: deleta!
B: eu não sei deletar
*: {init.}
B: vou tentar disfarçar aqui

Outra instrução era de que os amigos procurassem brincar, jogar ou ensinar algo uns para os outros durante as gravações. O intuito dessa instrução era de que houvesse maior descontração e, portanto, mais probabilidade de a gravação ficar boa, em estilo casual, o que aumentaria a probabilidade dos dados relevantes para a análise surgirem. Essa instrução foi parcialmente seguida. É perceptível que, na maioria das gravações utilizadas, houve a tentativa de seguir tais instruções. Constatamos, porém, que a gravação ficava menos casual quando os informantes claramente tentavam seguir as instruções para que ficasse casual, entretanto, a gravação ficava mais casual à medida que os interlocutores esqueciam as instruções.

É importante informar que nem sempre as gravações ficavam boas. De fato, a maioria das gravações realizadas foi descartada. Os problemas mais comuns que levaram a esse descarte foram: 1) inibição intensa – muitas crianças ficavam mudas ou monossilábicas durante longos períodos nas gravações: houve uma gravação, por exemplo, em que só se escutava “vai”, “é sua vez” e o barulho do arrastar das pedras de um jogo de xadrez; 2) crianças brincando de entrevistar e/ou representando personagens; 3) nível alto de ruídos, inclusive de televisão. Só foram utilizadas, portanto, as

gravações que passavam por um controle de qualidade em que faziam parte os critérios: boa ou razoável possibilidade de audição; razoável nível de espontaneidade e diálogos.

As gravações realizadas com a presença da pesquisadora com crianças de 7 a 12 anos presumivelmente não são tão casuais quanto as anteriormente explicadas. Como já dissemos, os grupos (pesquisadora + crianças ou pesquisadora + filha – menina de 10 anos que participou da maioria das gravações + demais crianças) iam para uma sala reservada, dentro da escola, onde a pesquisadora (ou a filha) lia um livro. O livro era discutido pelo grupo e depois as crianças ficavam à vontade para conversar sobre outros assuntos. Assim, em boa parte do tempo as conversas não assumem o estilo casual. Na última fase do encontro que surgia a conversa no estilo casual e, conseqüentemente, os dados. Mesmo percebendo que sobrava pouco tempo para conversas casuais (os alunos podiam ficar com a pesquisadora em torno de 40 minutos) não poderíamos deixar de ler a historinha que, em um só tempo, tornava os alunos mais próximos à pesquisadora e colaborava para o crescimento intelectual deles. Nessas gravações, também foram enfrentados problemas, tais como: difícil acesso às crianças; alto nível de ruídos (já que as gravações eram realizadas dentro da escola) e, algumas vezes, não chegava a haver conversas casuais. Quando ocorria algum desses problemas, a gravação também era descartada.

Já no terceiro tipo de gravação, com as crianças não pertencentes à comunidade da Vila Planalto, não houve a pretensão de que as gravações fossem ocultas. Era colocado um gravador perto das crianças ou elas mesmas ligavam o gravador enquanto conversavam. Foram quatro gravações dessa natureza e todas elas foram realizadas nas casas dos informantes: em duas os interlocutores eram irmãos e nas outras duas os interlocutores eram amigos (nesses casos, a casa era de apenas um dos informantes, naturalmente).

É importante lembrar que, embora a amostra seja representativa das comunidades de fala, esta não foi aleatória num sentido amplo; as gravações não foram realizadas todas do mesmo jeito e nem na mesma comunidade de fala. Por um lado, há um grupo de gravações que são mais espontâneas (no sentido lato), de crianças de 12 a 15 anos e, por outro lado, com crianças mais novas, há gravações maximamente espontâneas (no sentido lato) no contexto de pesquisador presente (entre outras limitações): essas duas na localidade Vila Planalto. E, ainda, há gravações de conversas espontâneas realizadas nas casas dos colaboradores, em um ambiente familiar, mas em outras localidades (Asa Norte, Sudoeste e Jardim Botânico – Lago Sul). Ao total, foram

11 horas e 24 minutos de gravação, sendo 43 informantes. Destes, 25 são do sexo feminino e 18 do masculino, sendo 9 informantes não pertencentes à comunidade da Vila Planalto. Houve, nas gravações, os informantes que, apesar de não terem o perfil descrito, fazem parte da amostra. São eles: duas mães, um rapaz de vinte anos e uma menina de cinco anos. As gravações desses informantes foram, grosso modo, incidentais. No entanto, a única fala incidental que interessa à presente investigação é a fala da menina de 5 anos, que entrou para a análise dos dados.

Depois das gravações e do controle de qualidade, essas gravações foram transcritas e, em seguida, houve a identificação dos dados relevantes à presente pesquisa. Cada dado identificado foi codificado a partir de um grupo de fatores ou variável independente que, em princípio, pode influenciar a variação. Esse arquivo se encontra no apêndice do presente trabalho. Depois de codificados, os dados foram submetidos ao pacote de programas *GoldvarbX* 2001 e *Varbrul* 1988/1989. A rodada final também se encontra em apêndice. Os resultados alcançados são evidências estatísticas do que realmente influencia e o que não influencia a variação. A partir daí, podemos fazer inferências, tirar conclusões, enfim, interpretar os resultados. As possibilidades de aprofundamento da análise são diretamente proporcionais ao amadurecimento científico do pesquisador, uma vez que as ferramentas utilizadas possibilitam infinitas combinações e refinamentos de análises.

3.2 Descrição da metodologia de análise da amostra

Ao utilizar os programas *GoldvarbX* (Sankoff; Tagliamonte; Smith, 2005) e *Varbrul* 1988/1992 (Pintzuk, 1988), utilizamos programas estatísticos, baseados em modelos matemáticos, para medir as influências que cada fator exerce sobre a variação em análise (cf. Sankoff, 1988). Quando realizamos rodadas, os primeiros resultados que dispomos são percentuais simples em função de cada variante *versus* cada fator, como, por exemplo: percentuais de *tu* para o sexo masculino. É comum que, nessas primeiras rodadas, encontremos fatores categóricos, que são um conjunto de fatores que se comportaram 100% com apenas uma única variante e, portanto, não apresentaram variação dentro da amostra. Assim, os passos seguintes são os de resolver os efeitos categóricos, eliminando dados de algum fator ou amalgamando fatores com efeitos semelhantes. São muitos os motivos para tomar as decisões de isolar ou amalgamar

dados, como: isolar (ou simplesmente retirar) – quando tais dados não fazem parte da análise variável ou necessitam de outro tipo de investigação; juntar – quando são poucos dados e se assemelham. A presente análise, como será possível verificar mais adiante, contou com algumas retiradas e amalgamações.

Depois das decisões tomadas quanto a isolar ou amalgamar, e dos resultados em percentuais se tornarem sem efeito categórico, ou seja, somente quando os resultados em percentuais apresentarem variação em função de todas as variantes e variáveis é que passamos ao próximo passo, que é a realização de rodadas de pesos relativos. Tais rodadas, além de atribuir pesos relativos a todos os fatores que influenciam a variação, também fazem a seleção das variáveis estatisticamente significativas nas etapas binárias, a saber, as que envolvem duas variantes. Os pesos são basicamente valores numéricos corrigidos, mais consistentes que frequências ou percentuais. Na presente pesquisa, a rodada de pesos relativos se tornou complexa porque, inicialmente, usou-se o programa *GoldvarbX*, que faz somente rodadas binárias. A realização de rodadas binárias ocorreu em três fases: *tu* versus *você*, *tu* versus *cê* e *cê* versus *você*. Mais adiante essa questão será retomada, mas é oportuno informar que as rodadas binárias, apesar de não apresentarem números mais adequados quanto aos pesos relativos, justamente porque a variação do fenômeno aqui analisado não é binária e sim eneária (com três variantes), foram fundamentais à presente análise por apresentar a possibilidade de seleção de fatores, ou seja, foi possível, por meio das rodadas binárias, que se chegasse aos fatores que apresentam relevância estatística para a presente variação. E, assim, tais rodadas servem, como veremos mais adiante, para a investigação de como os fatores se inter-relacionam quando da seleção ou não seleção das variáveis independentes.

Como a presente variação é eneária, entre *você*, *cê* e *tu*, o objetivo é fazer correlações entre elas, em conjunto e não somente duas a duas. Por isso, o passo seguinte foi rodar os dados no programa *Varbrul* 1988/1992, que possibilitou a rodada eneária de pesos relativos (mesmo não fazendo seleção). Tal rodada, por medir os pesos relativos para cada variante em contraste com as demais, tornou o resultado em pesos mais adequados para esse tipo de análise. Assim, apresentaremos no capítulo 5 do presente estudo a análise dos resultados em função de rodadas eneárias e binárias.

Naro (2003, p. 15-17) explica que, no jogo das variações estudadas, um fator não influencia uma variante de forma isolada. Em suas palavras “a operação de uma regra variável é sempre o efeito da atuação simultânea de muitos fatores”, ou, em outras palavras, “é, em princípio, impossível medir diretamente, nos dados de uso real, a

influência de uma dada categoria, sem medir simultaneamente o efeito das outras categorias”. Temos assim que, quando rodamos os dados no programa, este apresenta os pesos relativos em função de cada fator, com todos os fatores analisados conjuntamente. Assim, é possível que ocorra, quando os fatores estão sendo analisados, um resultado apontando a não convergência ou a sobreposição de fatores, que quer dizer basicamente que um fator se sobrepôs ao outro, que suas codificações estão de certa maneira cruzadas, como se o mesmo fenômeno num dado fosse marcado duas vezes, pelos dois fatores. Esse tipo de resultado também ocorreu na presente pesquisa, quando da investigação dos fatores “tipo de relação” e “faixa etária”, como teremos oportunidade de apresentar. Para a verificação de tais sobreposições, a exemplo do que sugere Guy (2007, p. 57-61), utilizamos o cruzamento dos dados para analisar esses resultados: mais adiante retomaremos esta questão.

É oportuno informar, por ora, que a análise dos resultados se deu de forma complexa, pois a cada nova rodada houve a necessidade de aprofundamento da investigação, não somente das análises quanto à variação em foco, mas também em relação ao método a ser utilizado, para que os resultados chegassem a um nível máximo de controle, contundência e robustez. Nas análises dos resultados, no capítulo 5 do presente trabalho, desenvolveremos as explicações das situações aqui mencionadas.

4 UM RETRATO DE ALGUNS TRABALHOS JÁ REALIZADOS

“O caminho de um milhão de milhas começa por um passo”

(Lao Tsé)

Muitos trabalhos já foram feitos acerca dos pronomes no português brasileiro e alguns foram realizados sobre os pronomes na variedade brasiliense. Os estudos acerca desse assunto têm se mostrado bastante frutíferos e não estão esgotados. Assim, para que se entenda melhor os comportamentos dos pronomes, apresentaremos a seguir uma síntese do quadro pronominal que encontramos em algumas gramáticas do português, bem como uma revisão dos principais trabalhos realizados por linguistas que norteiam a presente discussão.

4.1 Os pronomes segundo alguns gramáticos e alguns linguistas

Temos em nossa tradição gramatical o seguinte quadro de pronomes pessoais do caso reto: 1ª, 2ª e 3ª pessoa do singular: *eu, tu, ele(a)*, e do plural: *nós, vós, ele(as)*. A segunda pessoa do singular corresponde a somente o *tu* nas gramáticas dos seguintes autores: Napoleão Mendes de Almeida (2005); Said Ali (2001); Evanildo Bechara (2004), Rocha Lima (2006); Cunha & Cintra (2001). Salvo as gramáticas de Said Ali e a de Napoleão Mendes de Almeida, as demais gramáticas analisadas apontam o *você* como pronome de tratamento de 2ª pessoa com verbo de 3ª pessoa. Napoleão Mendes de Almeida, gramático reconhecidamente tradicionalista, coloca o *você* como pronome de terceira pessoa. Por seu turno, Said Ali comenta o *você* como “um pronome nascido do uso e abuso da fórmula *vossa mercê*, o qual perdeu o brilho e o significado original” (cf. p. 75 – com adaptações nossas). O autor, porém, não tece mais comentários a respeito da classificação gramatical da forma *você*. Por outro lado, temos em Mira Mateus et al. (2003) um quadro pronominal (a nosso ver) mais contemporâneo, em que as autoras colocam o *você* no quadro de pronomes pessoais de formas fortes, dentro dos nominativos (junto ao *tu*).

Nenhuma das gramáticas registra o clítico *cê*, forma reduzida de *você*, largamente usada no português brasileiro. Porém, essa variante já tem sido discutida em alguns trabalhos de linguistas desde a metade do século XX. Vitral (*apud* Loregian-

Penkal, 2004) o descreve como um pronome que está em estágio avançado de gramaticalização, chegando ao estágio de cliticização.

Em geral, o *senhor* é registrado como forma de tratamento de 2ª pessoa (em Napoleão Mendes de Almeida é registrado como de 3ª pessoa) com verbos na terceira pessoa (a exemplo do que ocorre com o *você*, por ser forma indireta). É consenso entre os gramáticos e linguistas que sua ocorrência denota respeito e tratamento cerimonioso.

Vale dizer que, em muitos materiais didáticos de ensino de português do Brasil como segunda língua, o quadro pronominal é apresentado de forma menos tradicional, somente com os pronomes: eu, você, ele, ela, nós, a gente, vocês, o senhor, a senhora, eles e elas. Ou seja, apresentam os pronomes com a supressão do *vós* e até mesmo do *tu* (cf. Lima et al., 1991, p. 10, e Henriques & Grannier, 2001, p. 4). Esse fato pode ter significado uma tentativa dos autores de aproximação entre o ensino da língua portuguesa e a realidade linguística que o aluno encontraria no uso da língua falada. Ao suprimir o pronome *tu*, porém, tais materiais didáticos corroboraram a ideia errônea sobre o desaparecimento deste pronome no escopo linguístico do brasileiro, como teremos oportunidade de demonstrar mais adiante. Loregian-Penkal (2004) sugere o seguinte quadro pronominal que, ao que tudo indica, se aproxima mais à realidade linguística observada no Brasil.

Quadro 01 – Proposta de quadro pronominal para o Brasil

Pronomes Pessoais Sujeito	
Pessoa Gramatical	Pronomes
1ª pessoa do singular	eu, a gente
2ª pessoa do singular	tu, você
3ª pessoa do singular	ele, ela
1ª pessoa do plural	nós, a gente
2ª pessoa do plural	(vós), vocês
3ª pessoa do plural	eles, elas

(Extraído de Loregian- Penkal (2004, p. 48) – com adaptações)

Quando se trata das formas variantes *tu* ou *você* (ou os pronomes de segunda pessoa no Brasil), esse assunto costuma ser bastante controverso entre linguistas. Illari (cf. “O português da gente”), entre outros autores, aponta o *você* como verdadeiro

pronome de 2ª pessoa do português brasileiro, sendo o uso do *tu* no Brasil restrito a alguns pontos das regiões Sul e Norte. Faraco (1996, p. 77) ao explicar o porquê da queda do s final dos verbos ligados à segunda pessoa em orações afirmativas ou o uso da desinência de 3ª pessoa (você) estar ocorrendo na segunda pessoa (*tu*), afirma que isso ocorre “como consequência do predomínio de você sobre *tu* no Brasil” (com adaptações). Outros autores, por outro lado, têm registrado a ocorrência de *tu* em diferentes localidades no Brasil, até mesmo com sinais de expansão desse pronome, como, por exemplo, em Paredes (2003) “O retorno do pronome *tu* a fala carioca” ou em Lucca (2005) “A variação *tu/ você* na fala brasiliense”. Há ainda em andamento trabalhos no norte do país que ajudarão a quantificar esse predomínio do *você* apontado por Faraco e outros autores. Mas é fato que o *você* não está isolado na referência à segunda pessoa no Brasil. Como já mencionado, há diversos trabalhos científicos contemporâneos que registraram e analisaram o pronome *tu* em uso no português brasileiro.

Outro fato a ser considerado é acerca da teorização de Duarte (1996, p. 109-125) em que a autora registra que “há, no português do Brasil, uma crescente simplificação nos paradigmas flexionais” e isso, entre outras consequências, levaria a uma tendência de maior preenchimento do sujeito (ambos os fenômenos se influenciam mutuamente). O fato de estarmos preenchendo mais o sujeito, além de influenciar o paradigma flexional na direção de uma simplificação, pode significar também uma influência direta na variação dos pronomes que utilizamos. É razoável pensar que, quando uma forma é muito explicitada, ela pode ficar mais suscetível à variação linguística.

Por todo o exposto, acreditamos que a investigação sobre os pronomes de segunda pessoa seja um terreno fértil para a pesquisa linguística.

4.2 Referência à 2ª pessoa e a perspectiva T/V

Em um célebre texto de Brown & Gilman (1960) “The pronouns of Power and Solidarity”, os autores teorizam sobre a existência nas línguas (de forma geral) de dois tipos de pronomes de segunda pessoa: são os pronomes T/V. Eles os explicam segundo uma semântica de poder e solidariedade, e fazem isso sob uma perspectiva diacrônica. Antigamente, a semântica do poder regia o uso dos pronomes, uma vez que, se os interlocutores fossem socialmente estratificados, os interlocutores das classes inferiores

tratariam seus superiores por V (pronome com traço de maior reverência) ao passo que geralmente seriam tratados por T (pronome de uso íntimo ou sem reverência). Por sua vez, os pronomes, se usados de forma mútua por seus interlocutores, geralmente indicavam que tais interlocutores se tratavam como iguais, pois pertenciam à mesma classe ou tinham simetria (social, econômica, intelectual etc.). Hoje, porém, a semântica do poder não é mais a “predominante”, pois os usos mútuos (em que a tendência é, com o uso de V, o interlocutor receber o mesmo V, e o mesmo ocorrer com o T) são os mais difundidos (pelo menos na cultura ocidental). É a chamada semântica da solidariedade. É importante salientar que, teoricamente, uma semântica não exclui a outra e que a semântica do poder continua a ocorrer, ainda que em menor escala. Os autores também teorizam sobre um pronome ser mais íntimo e o outro mais formal, T/V respectivamente. Faraco (1996, p. 51), ao fazer referência ao que ocorre no alemão entre *du* versus *Sie*, aponta um exemplo prototípico dessa relação T/V, sendo o *du* equivalente ao T (informal) e o *Sie* equivalente ao V (formal).

Lucca (2005, p. 87) faz uma reflexão sobre os usos dos pronomes *tu* e *você*. Tal estudo se encaixa na semântica da solidariedade, de Brown & Gilman (1960), em que o uso elevado de *tu* (72%) nos grupos de jovens brasileiros que a autora estudou, na faixa etária de 15 a 19 anos (sendo 92% dos informantes do sexo masculino), indicou, entre outras coisas, que a forma *tu* ocorria em maior escala quando exprimia uma relação de maior solidariedade.

Dias (2007, p. 7-23) faz uma revisão bastante detalhada da relação de poder e solidariedade e sintetiza esta teoria em diversas línguas. Em sua análise, que se caracterizou por uma investigação sobre a variação *tu* e *você* em três faixas etárias diferentes e com mais dados de mulheres, o uso da variante *tu* (em média 13%) não atingiu o mesmo valor atingido por Lucca. Na faixa etária equivalente a de Lucca, de 13 a 19 anos, o *tu* atingiu 29,8%. Isso ocorreu justamente porque ambas as pesquisas analisaram grupos socialmente distintos e, ainda, a forma de gravar os dados foram diferentes. No caso dos grupos estudados por Dias (2007), o *tu* não se manteve como indicador de solidariedade (mais adiante discutiremos mais detalhadamente esse estudo).

O fato é que, ao comparar os resultados de Dias e Lucca, fica evidente que, embora possamos classificar o *tu* como pronome do tipo T e o *você* como do tipo V, em muitas ocasiões, as nuances dessa classificação são bastante complexas, pois envolvem outros fatores de ordem social. Entendemos que essa classificação de pronomes T/V vale como auxiliar no entendimento da variação dos nossos pronomes, mas não são

estanques no sentido de, por exemplo, o *tu* ser o pronome informal e o *você* o pronome formal no português brasileiro. Essa observação foi discutida em Dias (2007, p. 10) quando a autora aponta para o fato de diferentes pronomes no português poderem ser encaixados nos papéis T/V, como, por exemplo, o *você* para T e o *senhor* para V. A esse respeito, Lucca (2005, p. 62) também afirma que “no português do Brasil podemos lidar, respectivamente, com as formas *tu* ou *você*, *cê* ou *você*, *você* ou *senhor*, conforme o dialeto”. E, ainda sobre essa relação entre pronomes e intimidade *versus* formalidade, nas palavras de Faraco (1996, p. 67) “(...) a evolução do Vossa Mercê afetou também o sistema do tratamento íntimo, já que você ou substitui o tu (como em geral aconteceu no português brasileiro) ou criou um nível de relativa formalidade entre tu e outras expressões mais formais”.

As considerações aqui colocadas sobre os pronomes T/V, embora nos auxiliem no entendimento da variação, nos levam a outra consideração acerca dessa teoria, pois, na variedade em estudo, há dados, por exemplo, de um menino que utiliza 4 variantes para se referir à segunda pessoa: *tu*, *você*, *cê* e o *senhor*. Sendo esse tipo de variação comum no repertório brasiliense (e, de forma geral, na variedade do português brasileiro), consideramos que as nuances entre formalidade e informalidade ou distanciamento e proximidade, por exemplo, não são somente dicotômicas, como sugere a aplicação dos pronomes T/V, nem estão em distribuição complementar, segundo determinadas aplicações semânticas e suas respectivas variedades. Podemos observar, assim, gradações de formalidade ou de distanciamento que se comportam como num *continuum* e estão submetidas a uma complexa teia de fatores sociais e interacionais, como teremos oportunidade de ver na seção 4.3, a seguir.

4.3 Alguns trabalhos sobre os pronomes pelo Brasil

Apresentaremos, a seguir, uma revisão de alguns trabalhos realizados pelo Brasil sobre a variação dos pronomes de segunda pessoa do singular na oralidade. Esses trabalhos, apesar de nem sempre serem teoricamente variacionistas, são norteadores sobre o que vem acontecendo no Brasil em função da variação entre os pronomes. Têm em comum, além do objeto de estudo, a observação de fatores de natureza externa à língua, como sexo e faixa etária.

A maioria dos estudos que serão revisados foram feitos em tempo aparente. O estudo em tempo aparente se caracteriza por analisar gravações ou entrevistas de pessoas de diferentes faixas etárias a fim de verificar e projetar os fenômenos de variação na linha do tempo, ou seja, é feito para que se possa observar a evolução da variação na língua. Os estudos aqui revisados também podem ser em tempo real, que se caracterizam por gravar e analisar o mesmo tipo de fala em diferentes épocas, ou seja, verificam os fenômenos da fala através do tempo, como foi o caso do estudo de Paredes (2003), com os dados do Programa de Estudos de Uso da Língua (PEUL).

Desejamos com isso ter, principalmente, um retrato da variação pronominal no Brasil e ainda saber como essa variação tem sido interpretada pelos autores.

4.3.1 Região Sudeste

Rio de Janeiro, capital

Paredes Silva (2003, p. 160-169) realizou pesquisa de linha variacionista em que analisa os dados dos pronomes *tu* e *você*, coletados em amostras do PEUL/UFRJ de dois períodos distintos, com o nome de Censo no primeiro período, que retrata a fala carioca em meados dos anos de 1980, e com o nome de Tendência no segundo período, que retrata a fala carioca dos anos 2000. Ambas as amostras se caracterizam por serem entrevistas labovianas típicas, gravadas com o conhecimento dos falantes.

Ao estudar os pronomes através da amostra Censo, do PEUL, Paredes percebeu que tais resultados não condiziam com a realidade que ela observava no dia a dia. A pesquisadora, apesar de observar a crescente ocorrência do pronome *tu* na fala carioca, não conseguia captar esses dados nas amostras do PEUL, inclusive na amostra BDI¹⁰. Paredes (2003, p. 167) pondera que os dados coletados no PEUL não refletiram a realidade dos pronomes em análise, pois, em função do gênero discursivo, o pronome *tu* apresentou baixíssima ocorrência.

A fim de buscar resultados que refletissem a fala carioca quanto ao uso dos pronomes, a pesquisadora formou uma amostra com dados coletados em gravações

¹⁰ Amostra Banco de Dados Interacionais (Roncarati et al., 1996), também pertencente ao PEUL, que tinha a finalidade de registrar conversas casuais dos falantes. Esta amostra foi feita com o conhecimento dos falantes.

ocultas, que se caracterizassem por conversas casuais. Surgiu assim a amostra Paredes 96.

A pesquisadora concluiu que o fato de não ter colhido de forma satisfatória a variação dos pronomes nas amostras do PEUL se deve, principalmente, ao gênero da amostra, ao tipo de interação que estava sendo registrado e se havia ou não o conhecimento dos informantes em relação à gravação, pois tudo isso influenciou os resultados quanto ao uso dos pronomes. Através de sua pesquisa, Paredes (2003) chega a outras conclusões relevantes sobre os usos dos pronomes na cidade do Rio de Janeiro: 1) há o pronome *tu* na fala carioca com concordância de 3ª pessoa; 2) há uma mudança em curso: os mais jovens usam mais o *tu* enquanto os mais velhos usam menos (tabela 02, a seguir, detalhará tal consideração); 3) há a ocorrência maior da forma *tu* na fala dos homens. A autora atribui essa tendência ao estigma que pode existir em torno do uso do pronome de 2ª pessoa com flexão de 3ª pessoa, fato que vai de encontro à prescrição gramatical e é condizente com a tendência feminina de preferir formas mais prestigiadas (a tabela 03 adiante detalhará tal consideração).

A seguir, a tabela 02 apresenta os números do pronome *tu* alcançados por Paredes na amostra Paredes 96 (cf. Paredes, 2003, p. 167) em comparação com as amostras do PEUL, em que analisa tipo de amostra *versus* faixa etária¹¹:

¹¹ Os valores das tabelas foram retirados das tabelas de Paredes (2003, p. 160-169). Os números totais variam em função de adaptações feitas por mim, como, por exemplo, a não explicitação de alguma faixa etária.

Tabela 02 – Variação *tu/ você* na fala carioca em função da idade

Idades/ amostras	Número/ total	Percentual	Peso relativo
Paredes 96			
10-19 anos	69/ 106	65%	.50
20-29 anos	130/ 185	70%	.60
30-39 anos	36/ 77	47%	.28
Totais	235/ 368	64%	
Corpus Censo (1980)			
7-14 anos	14/ 65	22%	.61
15-25 anos	21/ 213	10%	.61
26-49 anos	7/ 211	3%	.36
Totais	42/ 489	9%	
Corpus Tendência (2000)			
7-14 anos	1/ 61	2%	.16
15-25 anos	26/ 195	13%	.64
26-49	13/ 167	8%	.50
Totais	40/ 423	9%	

A tabela acima demonstra, de forma geral, que a idade é fator relevante para a análise dos dados. Tais resultados apontam, relativamente, que os mais novos tendem a usar mais o *tu* (apesar de que no *corpus* Tendência há mais dados de *tu* na idade intermediária). A tabela 03 a seguir apresenta os números encontrados por Paredes 96 (cf. Paredes, 2003) em função do sexo.

Tabela 03 – Variação *tu/ você* na fala carioca em função do sexo

Sexo/ amostras	Número/ total (tu)	Percentual (tu)	Peso relativo (tu)
Paredes (96)			
Homens	132*/ 192	69%	.57
Mulheres	103/ 176	59%	.43
Totais	235/ 368	68%	
Corpus Censo			
Homens	37/ 339	11%	.72
Mulheres	5/ 347	1%	.29
Totais	42/ 686	6%	
Corpus Tendência			
Homens	34/ 267	13%	.72
Mulheres	7/ 312	2%	.31
Totais	41/ 579	7%	

(* com adaptações¹²)

Em termos de comprovação do retorno do *tu* à fala carioca, as amostras do PEUL deixam a desejar, mas, a partir dessas amostras, já podemos perceber as tendências da variante *tu* quanto aos fatores de natureza externa: idade e sexo e tipo de registro da amostra. Ao comparar os pronomes nas amostras, Paredes encontra cerca de 95% de *você* nas amostras do PEUL contra 35% na amostra Paredes 96, sugerindo não só o retorno do pronome *tu* como uma hegemonia dessa forma em conversas casuais entre pessoas jovens.

Santos, São Paulo

Modesto (2007) registra a variação entre *tu* e *você* na cidade de Santos. Sua pesquisa segue pressupostos teóricos variacionistas, funcionalistas e interacionais. O autor sugere que a variação entre as formas de tratamento *tu* e *você* ocorre principalmente em contextos informais de fala (Modesto, 2007, p. 5). Por esse motivo, o pesquisador coletou seus dados a partir de gravações secretas e não secretas, sendo as secretas realizadas com intuito de se chegar a contextos menos formais de fala. Modesto

¹² Do original 32, mas que, pelos outros valores, pelos totais e pela interpretação, inferimos ser 132.

(2007, p. 6) analisou, entre outras, as variáveis sexo, gênero, escolaridade, referência, função sintática subjetiva e objetiva, incluindo nas análises objetivas de *tu* a forma *te* (além das formas *ti*, *a [para]* e *contigo*), e o *você* precedido ou não de preposição.

Quanto aos resultados, nas palavras do autor (Modesto, 2007, p. 9): “Entre todos os grupos de fatores controlados, apenas dois – gênero e faixa etária – foram considerados estatisticamente irrelevantes para a seleção das formas de tratamento (...)”. Embora haja, nessa pesquisa, a não seleção desses dois fatores sociais (que são fundamentais para maioria das pesquisas sobre essa variação no Brasil até o momento) outros grupos de fatores selecionados nos chamaram atenção, como, por exemplo, o monitoramento e a expressividade.

Os resultados apontam para 46% de ocorrência do pronome *tu* em falas não monitoradas, com o peso de 0,61, contra 17% de pronome *tu* em falas monitoradas, com peso relativo de 0,37. Em relação à expressividade (emoção e força apelativa aplicadas na enunciação), os resultados apontaram que, em enunciados com maior expressividade, a ocorrência da forma *tu* foi favorecida com o peso de 0,65, ao passo que, em enunciados com menor expressividade, a forma *tu* foi desfavorecida com o peso de 0,39. Sobre ambos os fatores, Modesto (2007, p. 13) afirma que “a forma *tu* é usada na maioria das vezes em contextos de maior expressividade e menor monitoramento”. Quanto aos resultados em relação à função sintática, vejamos a tabela a seguir.

Tabela 04 – Função sintática da variação *tu/você* em Santos, SP

Função sintática	Você	Tu
Subjetiva (sujeito)	440/ 617 = 71%	177/ 617 = 28%
Objetiva (objeto)	36/ 91 = 39%	55/ 91 = 60%
Total	476/ 708 = 67%	232/ 708 = 33%

(extraída de Modesto, 2006, p. 85 – com adaptações)

Os resultados da tabela acima demonstram um relativo favorecimento da forma *tu* em função objetiva. Modesto (2006, p. 98-100) afirma que esse resultado é favorecido principalmente pela alta ocorrência da forma *te*. Modesto (2007, p. 16-17) afirma ainda que os falantes alternam as formas *te*, *tu* e *você*, o que demonstra que eles não estão conscientes da associação do *te* ao *tu*, pois há em seus dados uma considerável ocorrência do *te* referindo-se ao *você*. Em outras palavras, podemos supor que pode

haver na pragmática uma dissociação quanto ao uso da forma oblíqua *te* à forma reta *tu*, uma vez que o *te* tem servido ao *você* da mesma maneira que ao *tu*.

Sobre a variação do *tu/ você* na fala de Santos, o autor conclui: 1) não há, nos dados encontrados, concordância canônica do pronome *tu* com a flexão verbal de 2ª pessoa; 2) as formas *tu* e *você* continuam vivas na região e 3) não há indícios de mudança em progresso, “(...) pois as variáveis gênero e faixa etária, grande indicadores de mudanças lingüísticas, mostraram-se estatisticamente irrelevantes pelo programa de análise” (Modesto, 2007, p. 24). Ou seja, essa variação seria estável.

São João da Ponte, norte de Minas Gerais

Mota (2008) registrou a variação *tu/ você* em uma cidade do Município de Montes Claros, norte de Minas Gerais. Em sua pesquisa, a autora realizou gravações labovianas de conversas ou narrações de diversos temas, em diferentes faixas etárias e de ambos os sexos. Os resultados indicaram que, no total, 10% dos dados são de ocorrência do pronome *tu*. O percentual, apesar de relativamente pequeno, é surpreendente, já que a ideia comumente aceita até então seria de que em Minas Gerais, de maneira geral, não existe a variante *tu* e, além disso, tais dados foram coletados em entrevistas labovianas que, ao que tudo indica, desfavorecem esse registro.

Os dados levantados por Mota (2008), quando submetidos ao tratamento estatístico, não foram selecionados quanto ao fator gênero (ou sexo). Mas, por outro lado, o fator faixa etária se mostrou significativo e foi selecionado pelo programa. A seguir, a tabela traz os resultados alcançados por Mota (2008, p. 69) da ocorrência de *tu* em função da faixa etária:

Tabela 05 – Variação *tu/ você* na fala mineira em função da idade

Faixa etária	Ocorrência	Percentual	Peso relativo
7-14 anos	4/31	12%	0,30
15- 25 anos	10/274	3%	0,72
26- 49 anos	32/131	24%	0,19
Acima de 50 anos	1/33	3%	0,14
Total	47/ 469	10%	

(extraído de Mota, 2008, p. 69 – com adaptações)

Reparamos que os percentuais e as ocorrências das faixas etárias de 15 a 25 anos e de 26 a 49 anos não condizem com os pesos relativos nem com a interpretação dada pela pesquisadora. Assim, assumimos que os resultados em pesos, bem como a interpretação da autora, é que valem como verdade¹³. Estudos semelhantes, a saber, Lucca (2005), que estudou a faixa etária entre 15 e 19 anos; Dias (2007), que estudou três faixas etárias diferentes, entre elas, a de 13 a 19 anos, e Paredes (2003), que estudou as faixas de 10 a 19 anos e 20 a 29 anos (entre outras), apontam para resultados equivalentes aos de Mota (2008), em que a tendência observada pela autora é de que a faixa etária de 15 a 25 usa mais a variante *tu* que as demais faixas.

Sobre a função sintática, Mota realizou sua investigação de forma semelhante a Modesto (2006 e 2007), vejamos seus resultados na tabela a seguir.

Tabela 06 – Função sintática da variação *tu/ você* em São João da Ponte, MG

Função Sintática	Você	Tu
Sujeito	392/ 411 = 95%	19/ 411 = 5%
Objeto	11/ 13 = 85%	2/ 13 = 15%
Objeto de preposição	19 = 100%	0%
Total	422/ 443 = 95%	21/ 443 = 5%

(extraído de Mota, 2008, p. 61 a 63 – com adaptações e retirados os casos de *tua* e dos *pronomes possessivos* da análise)

A exemplo da pesquisa realizada por Modesto (2006 e 2007), a maioria dos casos de objeto computados para a variante *tu* eram da forma *te*, mais precisamente 25 das 49 ocorrências correspondem à forma *te*, sendo 1 ocorrência da forma *ti* e somente 2 ocorrências de *tu* explícito. O que nos leva a reafirmar a ideia de associação pragmática da forma *te* com o *você* da mesma forma que é associada ao *tu*. E, por outro lado, podemos inferir também que o *tu* explícito ocorre mais como sujeito que como objeto.

Outra observação relevante da pesquisadora é de que o *tu*, em sua amostra, “é um fenômeno da zona rural” (cf. Mota, 2008, p. 83); ela ressalta que são fatores sociais que regem a variação, como, por exemplo, a relação entre interlocutores de intimidade *versus* não intimidade. Outra consideração relevante é de que, apesar do baixo

¹³ Acreditamos que deve ter havido apenas um erro de digitação, em que se trocaram alguns números das duas faixas etárias dentro da tabela.

percentual, as pessoas de faixas etárias mais elevadas mantêm em seu repertório o uso do *tu*. A autora interpreta esse fato como uma forma de manutenção do dialeto (ou como uma possível variação localizada na idade) e não como mudança em progresso (a exemplo da interpretação dada por Modesto, 2007), nas palavras de Mota (2008, p. 83 – com adaptações): “O fato do ‘tu’ estar presente na fala dos informantes da faixa etária de 26-49 anos e acima de 50 anos constitui uma evidência de que não se trata de uma inovação na comunidade”.

A seguir, vejamos um quadro que engloba os resultados da variação *tu* e *você* na Região Sudeste.

Quadro 02 – Variação *tu/ você* na Região Sudeste

Rio de Janeiro, RJ (Paredes Silva, 2003)	9% a 64%
Rio de Janeiro, RJ (Lopes & Marcutulio, 2009)	34%
Santos, SP (Modesto, 2007)	28%
São João da Ponte, MG (Mota, 2008)	5%
Belo Horizonte, MG (Ramos, 1997)	Não há registro
Uberlândia, MG (Herênio, 2006)	Não há registro
Vitória, ES (Calmon, 2009 e 2010)	Não há registro

Resta dizer que há, nesta região, alguns estudos (boa parte estudos diacrônicos) sobre os pronomes *você*, *ocê* e *cê*. A título de exemplificação, faremos um breve registro sobre a pesquisa de Gonçalves (2008), que estudou essa variação a partir de dados do dialeto mineiro de Arcos, centro-oeste de Minas Gerais, em comparação com outras pesquisas do mesmo tipo para realizar o contraste do português brasileiro com o português europeu. Gonçalves encontrou em seus dados o seguinte resultado para a função sintática:

Tabela 07 – Variação *você/ ocê/ cê* segundo a função sintática (Gonçalves, 2008)

	Você	Ocê	Cê
Sujeito	21%	19%	60%
Objeto direto	12%	88%	-
Objeto indireto	35%	50%	15%
Complemento nominal	36%	46%	18%

(extraído de Gonçalves, 2008, p. 188 – com adaptações)

O autor explica que o ambiente favorável às três formas é a função de sujeito. Gonçalves também encontra em sua amostra, ainda que de forma escassa, casos de *cê* ocupando posição de complemento nominal e de objeto indireto, fenômeno que vai de encontro a algumas pesquisas realizadas com as três variantes.

4.3.2 Região Sul

Loregian-Penkal (2004), entre outros autores, pesquisou os três estados da Região Sul do país¹⁴ quanto ao uso dos pronomes *tu/ você* e encontrou resultados interessantes. A ideia mais difundida no Brasil é de que no Sul se usa o *tu*. Essa generalização, embora em parte verdadeira, necessita de relativização. A depender da cidade, o *tu* é mais ou menos corriqueiro, ou, ainda, pode não fazer parte do repertório local. No Paraná, por exemplo, a pesquisadora não encontrou ocorrência de *tu*, sendo o pronome *você* de uso categórico nas áreas monolíngues. Nos outros dois estados da Região Sul, a pesquisadora encontrou uma alta ocorrência do *tu*. Apresentaremos, a seguir, uma compilação (detalhados em Loregian-Penkal, 2004) dos resultados encontrados em função do pronome *tu*, são eles:

Porto Alegre, RS – ocorrência na ordem de 80% em falas masculinas e 99% em falas femininas.

Flores da Cunha, RS – ocorrência na ordem de 68% nas falas masculinas e 96% nas falas femininas.

Panambi, RS – ocorrência na ordem de 79% das falas masculinas e 90% nas falas masculinas.

¹⁴ Através do banco de dados do VARSUL (Projeto Variação Linguística Urbana da Região Sul).

São Borja, RS – ocorrência na ordem de 89% nas falas masculinas e 99% nas falas femininas.

Florianópolis, SC – ocorrência na ordem de 59% em falas masculinas e 91% em falas femininas.

Ribeirão da Ilha¹⁵, em Florianópolis, SC – ocorrência na ordem de 96% em falas masculinas e 97% em falas femininas.

Chapecó, SC – ocorrência na ordem de 41% em falas masculinas e 59% em falas femininas.

Blumenau, SC – ocorrência na ordem de 14% em falas masculinas e 53% em falas femininas.

Lages, SC – ocorrência na ordem de 10% em falas masculinas e 23% em falas femininas.

Podemos inferir dos resultados anteriormente colocados que o comportamento da variação entre os pronomes é regido por fatores sociais também nas regiões estudadas pela autora. Nessas regiões, os fatores como a localidade específica e o sexo (e também a idade, como teremos oportunidade de ver mais adiante) influenciam o uso das variantes. Vejamos, a seguir, uma tabela dos resultados da variante *tu* em função das localidades e do sexo.

¹⁵ Este bairro de Florianópolis está separado no trabalho de Loregian-Penkal (2004) porque é um bairro peculiar se comparado aos demais, pois mantém características linguísticas (inclusive com mais concordância canônica para a variante *tu*) dos açorianos, imigrantes que se estabeleceram nesta região.

Tabela 08 – O *tu* em função da localidade e do sexo, em Santa Catarina (SC) e no Rio Grande do Sul (RS)

Florianópolis, Porto Alegre e Ribeirão da Ilha (<i>input</i> : 0,88)			
Masculino	663/ 881	75%	0,20
Feminino	1131/ 1167	96%	0,74
Total	1794/ 2048	87%	
Flores da Cunha, Panambi e São Borja (<i>input</i> : 0,89)			
Masculino	574/ 739	78%	0,23
Feminino	1138/ 1213	96%	0,74
Total	1712/ 1952	88%	
Chapecó, Blumenau e Lages (<i>input</i> : 0,27)			
Masculino	231/ 1315	18%	0,42
Feminino	353/ 919	38%	0,61
Total	584/ 2234	26%	

(extraído de Penkal, 2004, p.135 – com adaptações)

No Rio Grande do Sul, há a tendência de maior uso do *tu* se comparado com os resultados de Santa Catarina. Em ambos os estados, o uso dessa variante entre as mulheres é maior que o uso entre os homens. Resultados de outros fatores indicam que a idade também influencia fortemente a variação, apesar do fator faixa etária não ter sido selecionado com todas as localidades. Nas palavras de Loregian-Penkal (2004, p. 140), “a faixa etária de 25 a 49 é a que lidera o uso do pronome *tu*”. Vejamos, a seguir, os resultados do uso do *tu* em relação à localidade e faixa etária.

Tabela 09 – O *tu* em função da localidade e faixa etária, em SC e RS

Florianópolis, Porto Alegre e Ribeirão da Ilha (<i>input</i> : 0,88)			
25 a 49 anos	1020/ 1157	88%	0,55
Mais de 50 anos	774/ 891	86%	0,44
Total	1794/ 2048	87%	
Flores da Cunha, Panambi e São Borja (<i>input</i> : 89)			
25 a 49 anos	1023/1156	90%	0,60
Mais de 50 anos	689/796	87%	0,36
Total	1712/1952	88%	
Chapecó, Blumenau e Lages (<i>input</i> : 0,27)			
25 a 49 anos	459/ 1483	31%	0,62
Mais de 50 anos	125/ 751	16%	0,27
Total	584/ 2234	26%	

(extraído de Loregian-Penkal, 2004, p.140 – com adaptações)

Pelo fato de a faixa etária de 25 a 49 anos liderar o uso do *tu*, Loregian-Penkal (2004, p. 141) teoriza: “o uso de *tu* talvez esteja associado a uma menor formalidade, ou a uma maior intimidade (...)”. Essa ideia se coaduna com as demais interpretações de outros autores acerca do uso do pronome *tu* em faixas etárias semelhantes a essa.

A pesquisadora conclui, a respeito do *tu* nos estados do Rio Grande do Sul e Santa Catarina, que, apesar de haver localidades em que o *você* é bastante usado (como em Lages, por exemplo), segundo a autora (cf. Loregian-Penkal, 2004, p. 230), “o pronome *tu* permanece sendo uma forma bastante produtiva na linguagem oral”, funcionando, muitas vezes, como uma marca de identidade local.

Por fim, o quadro 03, a seguir, traz uma relação dos usos dos pronomes em função das quantidades de informantes que usaram tais pronomes na amostra analisada. Assumimos, assim, na ilustração a seguir, que as capitais dos estados da Região Sul são exemplificativas (sabemos que não há como seguramente generalizar) dos usos dos pronomes nessa região.

Quadro 03 – Variação *tu/ você* na Região Sul

Capital	Uso categórico do <i>tu</i>	Uso categórico de <i>você</i>	Uso variável
Curitiba, PR	Não houve o uso	24 pessoas	Não houve o uso
Florianópolis, SC	13 pessoas	01 pessoa	10 pessoas
Porto Alegre, RS	14 pessoas	01 pessoa	9 pessoas
Total	27 pessoas	26 pessoas	19 pessoas

(extraído de Menon e Loregian-Penkal, *apud* Loregian-Penkal, 2004, p. 67 – com adaptações e retirando algumas cidades)

4.3.3 Região Nordeste

A seguir, apresentaremos uma compilação de resultados encontrados sobre a variação desta região, sendo que muitos desses resultados foram retirados dos trabalhos de Dias (2007); Lucca (2005) e Scherre et al. (2009).

Paraíba

Bezerra (*apud* Lucca, 2005) registra a variação *tu/ você* na fala espontânea de crianças nascidas em João Pessoa, Paraíba. Seus resultados apontam que as crianças usam mais o *tu* que o *você* em seu repertório linguístico, chegando o *tu* à marca de 69% das ocorrências no total. As meninas, que usam os pronomes com cerca de 71% de *tu*, lideram, relativamente, esse uso. Os meninos usaram o pronome *tu* em 65% dos casos. A autora explica que o fenômeno de variação em seus dados “está relacionado ao tipo de ato comunicativo envolvido na interação – pedido, repreensão, insulto, desculpa, consentimento, entre outros” (cf. Lucca, 2005, p. 22).

Pedrosa (1999) também estuda o pronome *tu* através do banco de dados do VALPB¹⁶, mas tendo por foco a concordância. Sobre a ocorrência do *tu*, nas palavras da autora (Pedrosa, 1999, p. 44):

o nosso objeto de estudo não foi encontrado em abundância no *corpus* por nós analisado, mas a nível de percepção ousamos dizer que a

¹⁶ Projeto Variação Linguística no Estado da Paraíba, que tem por uma de suas finalidades registrar a fala na Paraíba.

utilização desse pronome é muito vasta em nossa comunidade, principalmente entre pessoas que possuem uma intimidade razoável, levando-nos a concluir que a falta de intimidade entre o falante e o entrevistador possa ter sido o fator inibidor da pouca ocorrência com o “tu”.

Vale dizer que a maior parte dos dados sem concordância de Pedrosa (1999, p. 38) está justamente na faixa etária de 26 a 49 anos. A partir dos estudos realizados por Bezerra (1994) e Pedrosa (1999), podemos dizer que há a alternância das formas *tu/você* em João Pessoa, Paraíba.

Fortaleza, Ceará.

Soares (*apud* Dias, 2007, p. 38-39, e Lucca, 2005, p. 25) estudou a alternância dos pronomes em Fortaleza e concluiu que essa variação é ternária, entre os pronomes *tu*, *você* e *senhor(a)*. Ao contabilizar os dados de Soares, Dias encontrou os seguintes percentuais totais: 28% de *tu*, 66% de *você* e 6% de *senhor* (ou o nome). Soares considera que os pronomes estejam condicionados pelos seguintes fatores: situação discursiva, papel social, faixa etária e grau de intimidade entre os falantes. Vale dizer que a investigação realizada por Soares (1980) não é variacionista. E, ainda, a autora usou diferentes tipos de *corpora*, porque, segundo ela, os dados colhidos em questionários não refletiram a realidade da fala em Fortaleza. Por isso ela buscou registrar outros *corpora* que comprovassem o uso da variante *tu* (que, segundo a autora, ocorre com concordância verbal variável).

Recife, Pernambuco

Sette (1980) pesquisou a variação *tu*, *você* e *senhor* na fala recifense. A pesquisadora aplicou testes de percepção em seus informantes. Os resultados a que chegou foram de que o *tu* é usado, embora em menor escala, se comparado ao *você*, na variedade recifense do português e o *senhor(a)* geralmente é usado para tratamento cerimonioso ou denotando respeito em função da idade e em função do cargo ocupado (com média de 85% dessa ocorrência quando se trata de pai ou mãe, em ambiente familiar; já no ambiente de trabalho os mais velhos tendem a receber esse tratamento

independente de ocuparem cargos inferiores; os superiores tendem a receber somente o *senhor(a)* independentemente da idade). Segundo os informantes, o que rege a variação dos pronomes *tu* e *você* na fala são fatores de ordem social, como intimidade e proximidade. Alguns informantes, porém, afirmam utilizar ambas as formas sem distinção (cerca de 12,5%). A autora comparou os dados quantitativos dos testes com outras observações assimétricas da fala para concluir que o *você* é mais usado em Recife que o *tu* (65% dos informantes afirmaram usar mais o *você* que o *tu*) e, através disso, a autora pôde observar as situações que favorecem um ou outro uso. Assumimos que o trabalho de Sette configura uma valiosa contribuição de como os recifenses percebem essa variação em sua fala.

Compilamos alguns resultados de trabalhos realizados sobre a variação *tu/ você* que ocorre na Região Nordeste. É necessário, porém, que se faça a seguinte ressalva: sabemos da complexidade de variedades culturais e linguísticas que compõem as regiões estudadas e que cada estado goza de características peculiares que o diferencia dos demais. O intuito aqui é apenas que tenhamos uma ideia, ainda que incipiente e não muito detalhada, do comportamento da variação pronominal que ocorre nas regiões, incluindo a Região Nordeste. Não pretendemos com isso generalizar, mas exemplificar a variação em foco.

Quadro 04 – A variação *tu/ você* na região Nordeste

Cidades/ Estados	Em função de Tu
Helvécia, Rio de Contas, Cinzento e Sapé, BA (Oliveira, 2005)	12%
Imperatriz, MA (Herênio, 2006)	27%
Fortaleza, CE (Soares, 1980)	28%
João Pessoa, PB (Bezerra, 1994)	69%

4.3.4 Região Norte

Soares & Leal (*apud* Lucca, 2005, p. 25-26) fizeram uma pesquisa em Belém do Pará sobre a variação *tu/ você* e *senhor* na interação entre pais e filhos e chegaram às seguintes conclusões: 1) os filhos tratam os pais usando as três variantes, na ordem de 49,1% de *tu*, 12,3% de *você* e 38,6% de *senhor*; 2) os pais, por sua vez, tratam seus filhos de forma também variável entre *você* e *tu*, na ordem de 76,8% de *tu* contra 23,2% de *você*. Como se tratam de dados colhidos na relação entre pais e filhos, pode-se inferir dos resultados encontrados que o *tu* é o pronome mais usado nas interações cujos falantes têm mais proximidade e intimidade.

Tomando como exemplo representativo (e levando em conta a ressalva feita anteriormente sobre a não generalização) da variação pronominal na Região Norte, apresentamos, a seguir, resultados sobre a ocorrência variável dos pronomes em função da forma *tu*.

Quadro 05 – A variação *tu/ você* na Região Norte

Cidade/ Estado	Em função de Tu
Belém, PA (Soares & Leal, 1993)	62,8%
Tefé, AM (Martins, 2010 ¹⁷)	Acima de 80%

4.3.5 Distrito Federal, Região Centro-Oeste

A seguir, teremos a revisão de alguns trabalhos de variação entre pronomes no Distrito Federal.

O Distrito Federal

Andrade (2004) realizou uma pesquisa com as variantes *você*, *ocê* e *cê* em Brasília, com uma amostra de fala da faixa etária entre 10 e 14 anos e também com os dados da pesquisadora, com na época 25 anos. A amostra foi constituída em 1991, feita a partir de gravações labovianas típicas e constam em Malvar (1992). A localidade

¹⁷ Resultados ainda não publicados. Dissertação de mestrado em andamento (cf. Scherre et al., 2009).

específica de Brasília que a amostra representa fica no entorno, mais precisamente aos arredores de Sobradinho, considerada como zona rural. Os resultados que a autora alcançou da variação dos pronomes em relação ao sexo estão descritos na tabela a seguir:

Tabela 10 – A variação *você/ ocê/ cê* no Distrito Federal em função do sexo (dados dos informantes de 10 a 14 anos)

Sexo	Você	Cê	Ocê
Masculino	31/ 63 = 49%	24/ 63 = 38%	5/ 63 = 5%
Feminino	40/ 68 = 59%	20/ 68 = 29%	4/ 68 = 6%
Totais	71/ 131 = 54%	44/ 131 = 34%	7/ 131 = 5%

(tabela extraída de Andrade, 1994, p. 73 – com adaptações e retirados os casos de pronomes nulos ou zeros)

A ocorrência da variante *cê*, de forma geral, foi bastante significativa, embora tenha sido menor que a ocorrência da variante *você*. Há que se considerar, como dito anteriormente, que se tratou de uma coleta de dados tradicional, do gênero discursivo entrevista, em que os informantes sabiam previamente da gravação, fato que influencia a fala a ficar mais monitorada. Outro resultado interessante é o fato de as pessoas do sexo masculino favorecerem o uso do *cê*, com o peso de 0,61, enquanto as pessoas do sexo feminino desfavoreceram esse uso, com peso de 0,40 (cf. p. 73).

Andrade (2004, p. 49-50) também analisa a função sintática, e um resultado relevante para a nossa análise é a verificação da fala da entrevistadora (com 25 anos na época da gravação – que apresentou mais dados, os quais se assemelham aos resultados das crianças estudadas). Vejamos a tabela a seguir:

Tabela 11 – Função sintática da variação *você/ cê/ ocê* em Brasília (dados da entrevistadora)

	Você	Cê	Ocê
Sujeito	602/ 996 = 60%	226/ 996 = 23%	15/ 996 = 2%
Objeto de verbo	8/ 8 = 100%	0/ 8 = 0%	0/ 8 = 0%
Objeto de preposição	38/ 38 = 100%	0/ 38 = 0%	0/ 38 = 0%
Dupla função	5/ 6 = 83%	0/ 6 = 0%	1/ 6 = 17%
Total	674/ 1069 = 63%	226/ 1069 = 21%	16/ 1069 = 1%

(extraído de Andrade, 2004, p. 50 – com adaptações e retirando os casos de: zeros, 153 casos, que correspondem a 15% e demais casos que somam 21 ocorrências)

Os dados da tabela anteriormente colocada sugerem que, quando o pronome serve à dupla função (sujeito e objeto) ele até pode ser um *ocê*, mas categoricamente não poderá ser um *cê*, pois, nos dados de Andrade (2004) não há casos desta forma como objeto. O *ocê* (como verificado também em Gonçalves, 2008) parece não ter restrição sintática apesar de essa forma ser escassa nas amostras analisadas. O *cê* é sintaticamente restrito à função de sujeito, fato que vai de encontro aos resultados de Gonçalves (2008), mas que se coaduna com a maioria das pesquisas variacionistas acerca dessa variação. A função de sujeito, por seu turno, é a mais produtiva para todas as variantes em estudo (uma vez que forma a ordem natural do português em SVO) sendo, portanto, essa a função irrestrita para todas as variantes.

Por fim é importante ressaltar que Andrade (2004) não encontra em sua amostra da fala de Brasília a ocorrência da forma *tu*.

Lucca (2005) realizou pesquisa sobre a variação dos pronomes *tu* e *você* (em que os casos de *cê* foram acoplados ao *você*) na fala de jovens brasilienses em três regiões administrativas diferentes do Distrito Federal, a saber: Brasília, Taguatinga e Ceilândia. A autora revela que sua coleta de dados ocorreu em circunstâncias prototípicas para o uso da variante *tu*, uma vez que a maioria dos dados foi coletada em gravações de conversas espontâneas e ocultas entre jovens rapazes que eram amigos. Essa pesquisa não teve a idade como fator social de investigação, uma vez que a pesquisadora considerou as pequenas diferenças de idade entre os informantes como indicador de que tais informantes pertenciam à mesma faixa etária, de 15 a 19 anos.

Os fatores analisados e selecionados pelo programa foram – por um lado, os linguísticos: 1) paralelismo formal (tendência de formas semelhantes se agruparem) e 2)

tipo de estrutura (se afirmativa, interrogativa ou exclamativa); por outro lado, os extralinguísticos: 1) gênero do falante, 2) tipo de relação entre os pares, 3) região administrativa e 4) familiaridade com o tema.

O fator gênero do falante, na pesquisa de Lucca (2005), foi o primeiro a ser selecionado pelo programa, e os resultados apontaram para o peso relativo da variante *tu* na ordem de 0,55 para o gênero masculino, enquanto o gênero feminino ficou na ordem de 0,09. A diferença entre os pesos revelou que o gênero masculino favorece de forma indubitável o uso do *tu*, ao passo que o gênero feminino o desfavorece (é importante salientar que cerca de 90% dos dados de Lucca são de falas masculinas).

Em função da região administrativa, a cidade de Ceilândia foi a que mais favoreceu o uso do *tu*, com o peso de 0,68, enquanto Taguatinga e Brasília ficaram, cada uma, com 0,43 de peso relativo. Segundo a pesquisadora, a variante *tu* teria aparecido no repertório brasileiro em função do grande contingente de imigrantes nordestinos que hoje vivem no Distrito Federal. O fato de Ceilândia ter se destacado com mais ocorrências da variante *tu* corrobora esse teoria, já que, entre as três regiões administrativas estudadas, Ceilândia conta com o maior contingente de imigrantes nordestinos.

A autora conclui que o *tu* é mais recorrente nas seguintes situações conversacionais: “falantes jovens do gênero masculino que interagem com pares solidários tratando de temas cotidianos” (cf. p. 113) e que essa variação ocorre principalmente segundo fatores sociais.

Dias (2007) estudou a mesma variação no Distrito Federal, com dados apenas da região administrativa de Brasília, porém, com dados de ambos os sexos equilibradamente (51,6% dos dados de falantes do sexo feminino e 48,4% dos dados de falantes do sexo masculino) e em três faixas etárias diferentes, a saber: 13 a 19 anos (19% dos dados); 20 a 29 anos (47,1% dos dados) e mais de 30 anos (33,9% dos dados).

Os fatores analisados pela autora (Dias, 2007, p. 64-94) e selecionados pelo programa foram: por um lado, o único linguístico selecionado: 1) tipo de fala (*frames* ou enquadres) – a) conversa casual, com o peso de 0,56; b) conversa profissional ou acadêmica, com o peso de 0,17; c) repreensões, com o peso de 0,34 e d) observações irônicas, deboches e brincadeiras, com o peso de 0,80. Por outro lado, os fatores extralinguísticos (ou sociais): 1) sexo; 2) faixa etária; 3) estilo do falante; 4) tipo de relacionamento com o interlocutor e 5) faixa etária do interlocutor.

Os resultados alcançados pela autora em função da faixa etária estão na tabela a seguir:

Tabela 12 – A variação *tu/ você* no Distrito Federal em função da idade (Dias, 2007)

Faixa Etária	Frequência de <i>tu</i>	Peso relativo
13 a 19 anos	51/ 171 = 29,8%	0,76
20 a 29 anos	53/ 424 = 12,5%	0,56
Mais de 30 anos	11/ 305 = 3,6%	0,28
Totais	115/ 900 = 12,8%	-

(extraído de Dias, 2007, p. 70 – com adaptações)

Os pesos indicaram o favorecimento do *tu* por parte da faixa etária de 13 a 19 anos. Outra observação da autora é que os informantes com mais de trinta anos tenderam a usar a variante *tu* em situações mais específicas, enquanto as outras faixas etárias tenderam a usar esta variante em diversos tipos de fala. Nas palavras de Dias (2007, p. 74), "as frequências de uso de *tu* nas brincadeiras/ ironias cresce à medida que elevamos a faixa etária, e o inverso acontece com outros tipos de fala". E, ainda, "o *tu* está passando de um uso altamente específico para uso de contextos variados".

Quanto ao fator sexo, a autora chegou aos seguintes resultados:

Tabela 13 – A variação *tu/ você* no Distrito Federal em função do sexo (Dias, 2007)

Sexo	Frequência de <i>Tu</i>	Peso relativo
Feminino	50/ 464 = 10,8%	0,41
Masculino	65/ 436 = 14,9%	0,60
Totais	115/ 900 = 12,8%	-

(extraído de Dias, 2007, p. 75 – com adaptações)

Ao analisar ambos os fatores – sexo e faixa etária – a autora chegou aos seguintes percentuais em função da variante *tu*:

Tabela 14 – A variação *tu/ você* no Distrito Federal em função do sexo e idade (Dias, 2007)

Sexo	13 – 19 anos	20 – 29 anos	Mais de 30	Total
feminino	24/ 106 = 22,6%	26/ 269 = 9,7%	0/ 89 = 0%	50/464 = 10,8%
masculino	27/ 65 = 41,5%%	27/ 155 = 17,4%	11/ 216 = 5,1%	65/ 436 = 14,9%
Totais	51/ 171 = 29,8%	53/ 424 = 12,5%	11/ 305 = 3,6%	115/ 900 = 12,8%

(extraído de Dias, 2007, p. 76 – com adaptações)

Dias (2007) analisa as demais variáveis (selecionadas e não selecionadas), mas, para este recorte, optamos por salientar somente as variáveis já apontadas. A autora confirmou suas hipóteses de que, quanto mais novo o falante, mais *tu* usaria e também a tendência de *tu* ocorrer mais na fala de pessoas do sexo masculino. Lucca (2005), quando optou por estudar o grupo prototípico dessa variante, demonstrou que tinha exatamente essa hipótese. Podemos perceber não os mesmos resultados numéricos, mas, de forma geral, as mesmas tendências nos dados de Dias (2007) e Lucca (2005), em que a maior tendência de ocorrência de *tu* é na fala de falantes do sexo masculino na faixa etária entre 13 e 19 anos, em relações solidárias ou íntimas. Na comparação, porém, os resultados de Dias lançam a semente do possível aumento da ocorrência de *tu* na fala das meninas.

Apresentaremos, a seguir, um quadro com os resultados da variação *tu/você/cê* no Distrito Federal (DF), Região Centro-Oeste:

Quadro 06 – A variação *tu/você/cê* no DF, Região Centro-Oeste

	Você	Cê	Tu
Adriana (2004) ¹⁸	108/ 189 = 57%	48/ 189 = 25%	Não há registro
Lucca (2005)	125/ 453 = 23%		327/453 = 72%
Dias (2007)	785/ 900 = 87,2%		115/ 900 = 12,8%

(extraído de Andrade, 2004, p. 45; Lucca, 2005, p. 78 e Dias, 2007, p. 64 – com adaptações. Retirando os “nulos e ocês” de Andrade)

¹⁸ É importante lembrar que a pesquisa foi realizada em 2004, mas os dados são de 1991.

4.4 Sobre o que vimos

Assumimos nesta revisão que o *você* e o *tu* são os pronomes de segunda pessoa do singular mais “generalizados” no português. Porém, eles podem ser sentidos como pronomes de uso não local em algumas localidades do país, como, por exemplo, o *tu* em Curitiba ou o *você* em Porto Alegre. Assim, tais pronomes mais “generalizados” obedecem a características diatópicas. Talvez por serem os mais difundidos, os pronomes *tu* e *você* foram, até o momento, os mais estudados por pesquisadores no Brasil.

De maneira geral, os resultados para o uso do *tu* em relação às faixas etárias foram semelhantes, apontando, em todas as pesquisas revisadas, para um aumento de *tu* nas faixas etárias mais novas bem como a diminuição de sua ocorrência em faixas etárias mais elevadas. Porém, houve algumas formas diferentes de interpretação desse fenômeno: por um lado, esse fato foi interpretado como caracterizador de mudança em curso, sugerindo, de maneira geral, o aumento da ocorrência da variante *tu*. Por outro lado, esse fato também foi interpretado como sendo variação estável e localizada na idade, em que os mais novos tendem a usar mais a variante *tu*, mas que, mais tarde, quando os falantes estão mais maduros, essa variante tende a ocorrer menos em suas falas e, em alguns locais, tal ocorrência pode até mesmo se especializar, se transformando em uma variante de uso de contextos específicos como em brincadeiras ou ironias, por exemplo.

Quanto ao fator sexo, essa variação (entre *tu* e *você*) também apresentou características diferentes em função, principalmente, da localização dos grupos de falantes. Um fato curioso, porém, é que, em localidades em que o *tu* é mais usado por pessoas do sexo masculino, a interpretação dada pelos pesquisadores é de que, em geral, as mulheres tendem a preferir formas prestigiadas ou mais próximas à prescrição gramatical (afinal, a ocorrência de *tu* com concordância canônica é de média para inexistente em todo o país). Por esse motivo, as mulheres usariam menos a variante *tu*, por esta variante não ser prestigiada já que viria sem concordância ou com concordância variável. Essa interpretação, de mulheres tenderem às formas mais prestigiadas, se encaixa na especulação sobre a diferença da fala entre os sexos, feita por Labov, mas com a ressalva de se aplicarem especificamente quando em contexto de fala mais formal.

Em outras palavras, Labov sugere que as mulheres tendem à fala mais formal em contextos mais formais de fala.

Entretanto, em localidades em que o uso de *tu* por mulheres é mais alto, a saber, no Rio Grande do Sul e em Santa Catarina, a interpretação dada é de que, nessas localidades, as mulheres tendem ao conservadorismo quanto ao uso de *tu* (cf. Loregian-Penkal, 2004, p. 139). Portanto, essa tendência existiria mais por uma questão ligada à identidade local. Vale lembrar que em tais localidades, em que o índice de *tu* é mais alto na fala de mulheres que na de homens, o *tu* é largamente usado sem concordância canônica, há, porém, localidades nessas regiões em que a concordância é considerável, cf. Loregian-Penkal (2004). Segundo Labov, as mulheres, especificamente em fala informal, tendem ao maior uso da variante “inovadora”. Mas não podemos dizer que essa teoria se aplica a esse caso concreto porque o *tu* não foi considerado inovação nem em Santa Catarina nem no Rio Grande do Sul.

O clítico *cê*, como redução do *você*, não ocorre em todas as localidades onde há o uso do *você*, principalmente quando o uso do *você* é sentido como mais formal e está em desvantagem acentuada na variação com o *tu* (por exemplo, como ocorre em Tefê, AM). Nos trabalhos revisados sobre a variação em Santa Catarina e Rio Grande do Sul, por exemplo, não houve menção a essa variante. Acreditamos que, se lá há a ocorrência do *cê*, esta deve ser pequena. Nessas localidades, onde o *cê* é de raro a inexistente, o *tu* tende a ocorrer mais e de forma mais generalizada (e não somente em determinadas situações conversacionais). Podemos, nesse sentido, pensar em atribuir a falta de ocorrência do *cê* à baixa ocorrência de sua forma plena, o *você*.

Nas localidades em que o *cê* ocorre, porém, seu uso é relevantemente alto e também indica (a exemplo do que ocorre com o *tu*) uma tendência de maior informalidade em oposição à sua forma plena (cf. Gonçalves, 2008). A variante *cê*, na maioria das pesquisas que vimos sobre a variação pronominal, é acoplada ao *você* (em Lucca e Dias, por exemplo). Outro fato relevante é que essa forma é a única das variantes que é sintaticamente restrita.

O *senhor(a)*, por seu turno, é o único opositor que torna a variante *você* um pronome acentuadamente informal. Tal forma também é pouco discutida nos trabalhos revisados e os resultados sempre apontam que o *senhor(a)* é um pronome que denota tratamento cerimonioso ou respeitoso (usos mais comuns no português brasileiro, e, também, é a forma com que é apresentada em manuais de português, por gramáticos e linguistas) e, por ser assim, sua ocorrência, ao que tudo indica, tende a ser muito baixa.

5. ANÁLISE DOS RESULTADOS

“Esmiuçar as entranhas das formas linguísticas e sentir a sistematicidade que envolve as línguas, dialetos e variedades, sem julgamento de valor (...)Partilhar esse bem constitui mais do que um dever, é uma responsabilidade social, é uma questão de cidadania.”
(Marta Scherre)

5.1. Ponderações preliminares:

Antes de entrar na análise propriamente dita, é interessante informar que há outras formas amplamente utilizadas por brasilienses (e altamente recorrentes em nossa amostra) para se referir a segunda pessoa. Em conversas informais, as mais utilizadas atualmente são os vocativos: “tia(o)” para se referir a professores, mães ou pais de amigos ou aos amigos dos pais; “véi” ou “doido” para amigos. Podemos nos referir a segunda pessoa, também, somente a olhando ou tocando, falando diretamente com ela. Nesses casos, se não houver o uso dos pronomes, nomes ou vocativos, será o chamado sujeito nulo. Na presente pesquisa não analisaremos esses tipos de dados.

Percebemos que os pronomes em análise dentro da nossa amostra, o *tu*, o *você* e o *cê* são pronomes que têm uma diferença de simetria/assimetria; respeito/não respeito, ou seja, estão no já mencionado *continuum*, sendo possível decifrar tais intenções somente no momento da enunciação, dentro de seu contexto. Assim, é fato que o *tu* pode indicar, muitas vezes, proximidade/intimidade, brincadeiras ou até mesmo não respeito (cf. Dias, 2007), mas pode também ser dito sem preocupação a respeito disso, principalmente em conversas casuais entre pessoas jovens (ou que têm uma atitude jovem). As crianças pesquisadas, de maneira geral, se referiram à pesquisadora (com a qual não tinham simetria, mas demonstravam respeito) por *tu*, *você* e *cê* e raramente demonstraram preocupação com o pronome a ser usado. É importante relatar, ainda, que houve apenas dois informantes que perceptivelmente demonstraram preocupação com a forma pela qual se refeririam à pesquisadora – foram os únicos momentos de preocupação com relação a como se referir aos interlocutores em toda a amostra analisada. Apenas um dos informantes usou, algumas vezes, a forma nominal a *senhora* (que não entrou para a análise estatística dos dados por estar na fala de um único informante, mais adiante retomaremos essa questão) e outra informante evitou o uso dos

pronomes, usando-os apenas 4 vezes ao longo do registro de sua fala, utilizando também o recurso do sujeito nulo.

O pronome *você* transita em dois mundos: da fala formal à informal. Permito-me testemunhar que, nos únicos momentos em que não trato naturalmente (e de forma inconsciente) meus interlocutores por *você*, é quando estou vivenciando o papel de aluna ou quando estou em uma situação de extrema formalidade em que necessito tratar meu interlocutor por *senhor(a)* (ocasiões bastante hierárquicas e raras). Penso que chamar um professor (ou uma autoridade) por *você* é muita intimidade. Mas, por outro lado, chamá-lo(s) de *senhor(a)* pode marcar uma distância maior do que a que eu gostaria no momento da enunciação (e muitas vezes não agrada ao interlocutor(a)). Uso, frequentemente, o *cê* para tentar passar despercebida. O *você* é adequado em situações formais, quando usado em reuniões, discussões técnicas, no trabalho de forma geral, em palestras etc; ou seja, o *você* terá *status* de pronome formal a partir do momento em que surgir em um contexto formal de fala. Assim, podemos dizer que o *você* pertence de forma irrestrita (sendo sutilmente diferente segundo algumas intenções no momento e no contexto da enunciação) a todo o *continuum* informalidade/formalidade.

Por seu turno, o *cê* também transita entre dois mundos, sendo basicamente informal, mas podendo surgir até mesmo em contextos formais de fala, como na narração anteriormente explicitada. Contudo, se notado, certamente será sentido como um pronome mais informal. Uma característica interessante do *cê* é que este parece sempre manter, mesmo dentro da formalidade, o grau de respeito que pode ser atribuído ao *você* (mesmo em situações em que o *você* seria sentido como demasiadamente informal, o *cê* não parece aderir a esse sentido do *você*).

Talvez a alta ocorrência do *cê* (bem como dos demais pronomes) em vários *corpora*, também se deva ao fato, já anteriormente mencionado, de estarmos preenchendo mais o sujeito (ou perdendo característica de língua *pro-drop* – cf. Duarte).

O *tu* é, a rigor, o pronome considerado típico da intimidade, como já dito anteriormente. Nos lugares onde este pronome é nativo, como no Rio Grande do Sul e em Santa Catarina, ou em localidades do Pará ou do Amazonas, seu uso parece se evidenciar bastando para os interlocutores que se reconheçam como pertencentes ao mesmo lugar (e resguardada alguma eventual formalidade, que pode provocar a conjugação verbal na segunda pessoa – cf. Loregian-Penkal, 2004). O linguista Germano Martins, amazonense, residente em Tefé, relatou em comunicação pessoal que, uma vez que se sinta mais íntimo ou amigo de alguém, já se sente autorizado a utilizar o

tu com tal interlocutor (independente de onde esse interlocutor seja), mas se este interlocutor for alguma autoridade (como pais ou professores) ou tiver em idade elevada, o linguista não o tratará por *tu* nem que se torne muito íntimo. Pelas frequências apresentadas nos trabalhos já realizados, onde o *tu* é nativo e pelo fato de o *tu* ter sido registrado com alta recorrência em entrevistas labovianas nessas localidades, acreditamos que esse pronome ocorra naturalmente até mesmo em contextos formais de fala.

Por outro lado, em lugares onde o *tu* não é nativo, ou melhor, não marca identidade local, como no Rio de Janeiro, por exemplo, mas que são lugares onde o *tu* é bastante recorrente, a tendência desse pronome é surgir em contextos casuais de fala. Ou seja, ao que tudo indica, esse *tu* não entra em contextos formais de fala, sendo relativamente e socialmente mais restrito que o *tu* de ambientes nativos. Podemos dizer, de forma geral, que o *tu*, nesses lugares, é um pronome da casualidade, intimidade e contextos desse tipo. Podemos dizer, assim, que este pronome, à semelhança dos demais, obedece a fortes características diatópicas e interacionais, como já tivemos oportunidade de demonstrar.

5.2. Descrição do *corpus*

Como já mencionado no item 3.1.1 do presente trabalho, foram ao todo utilizadas 14 transcrições (sendo 16 gravações ao total, mas duas delas sendo continuações), o que resultou em um total de 11 horas e 24 minutos de gravação. Os resultados globais contam com 953 dados, sendo cerca de 70% dos dados de pessoas moradores da Vila Planalto e 30% dos dados de moradores de outras localidades, como: Jardim Botânico, no Lago Sul (3 meninas e 1 menino); Asa Norte (2 meninos); Sudoeste (2 meninas e 1 menino). Excetuando a Vila Planalto, que é socialmente híbrida, as demais regiões administrativas não são muito diferentes entre si em termos de classe social. Há, na amostra estudada, 24 meninas e 18 meninos que compõem a faixa etária escolhida, de 7 a 15 anos. Além deles, há uma menina de 5 anos (com 68 dados), uma mulher em torno de 50 anos (com apenas 6 dados) e um rapaz de 20 anos (com 23 dados). Usaremos nos testes estatísticos apenas os dados da menina de 5 anos, pois: a) essa idade se aproxima significativamente da faixa etária estudada; b) essa informante é moradora de fora da Vila; c) seu comportamento linguístico se assemelha

ao comportamento linguístico de sua localidade. Os dados dos outros dois informantes não comporão os refinamentos estatísticos do presente estudo, já que estes informantes têm idades distintas das faixas etárias aqui analisadas.

Além desses dados de faixas etárias diferentes, serão retirados da maioria dos testes estatísticos também os casos de função de objeto (direto e indireto), já que uma das variantes em estudo, o *cê*, é sintaticamente restrita e não surge nesta função. É interessante ressaltar que essa categoricidade do não surgimento do pronome *cê* em função de objeto em nossa amostra, por si só, já configura um importante resultado.

Quanto ao sexo, 413 dados são de meninos, o que corresponde a 43,3% e 540 dados são de meninas, 56,7% do total. Acreditamos que dentro da faixa etária escolhida para o presente estudo contamos com uma amostra equilibrada acerca do sexo dos informantes. Informações mais detalhadas sobre os informantes quanto a sexo, idade, origem dos pais e localidade onde moram em Brasília encontram-se na tabela a seguir (retirando os dois informantes já mencionados que não comporão as análises estatísticas):

Tabela 15 – Detalhamento dos informantes

Informantes	Idade	Sexo	Origem mãe	Origem pai	Localidade de residência
(L)	5	F	RJ	AM	Lago Sul
(T)	7	F	RJ	AM	Lago Sul
@	7	F	BSB	CE	Vila Planalto
(z)	7	F	BSB	BSB	Vila Planalto
(a)	7	F	RN	RN	Vila Planalto
(s)	7	M	BSB	MA	Vila Planalto
(l)	7	M	BSB	BSB	Vila Planalto
(t)	8	F	BSB	PB	Vila Planalto
(K)	8	F	BSB	BSB	Vila Planalto
(F)	8	M	BSB	BSB	Vila Planalto
(C)	8	M	BSB	BSB	Vila Planalto
(q)	9	F	PE	PE	Vila Planalto
(n)	9	F	PE	BSB	Vila Planalto
(D) ¹⁹	9 e 10	F	BSB	RJ	Lago Sul
(U)	9	M	BSB	BSB	Vila Planalto
(B)	10	F	RJ	BSB	Sudoeste
(o)	10	M	ES	RJ	Asa Norte
(i)	10	M	BSB	BSB	Asa Norte
(M)	10	M	BSB	BSB	Vila Planalto
(m)	10	F	RN	RN	Vila Planalto
(A)	11	F	TO	Maranhão	Vila Planalto
(h)	11	M	BSB	Bahia	Vila Planalto
(y)	12	M	MG	MG	Lago Sul
(N)	12	F	BSB	RN	Vila Planalto
(V)	12	F	PB	PB	Vila Planalto
(G)	12	F	PI	CE	Vila Planalto
(f)	12	M	MG	MG	Vila Planalto
(Y)	12	M	RJ	BSB	Sudoeste
(E)	13	F	PE	PE	Vila Planalto
(R)	13	F	MG	RN	Vila Planalto
(k)	13	F	BSB	MT	Vila Planalto
(Q)	13	F	PI	CE	Vila Planalto
(p)	13	M	GO	RS	Vila Planalto
(b)	14	F	RJ	BSB	Sudoeste
(l)	14	F	BSB	BSB	Vila Planalto
(J)	14	F	BSB	BSB	Vila Planalto
(u)	14	M	PB	MG	Vila Planalto
(X)	15	F	BA	BA	Vila Planalto
(W)	15	F	BSB	PB	Vila Planalto
(1)	15	M	-	-	Vila Planalto
(2)	15	M	-	-	Vila Planalto
(?)	15	M	PE	PE	Vila Planalto
(P)	15	M	PE	PE	Vila Planalto
43 informantes		25 fem./ 18 masc.			

¹⁹ Essa informante é a filha da pesquisadora e foi gravada em várias ocasiões diferentes ao longo de mais de um ano, portanto, com idades diferentes. Os informantes 1 e 2 não apresentaram dados, portanto não foram contactados para a descoberta sobre as origens de seus pais.

5.3 Como as variáveis serão analisadas: breve visão de conjunto

A análise dos dados foi baseada nos seguintes fatores:

Variável dependente: *você, cê e tu*.

Variáveis independentes:

Extralinguísticas (ou externas): Faixa etária

Tipo de relação (simétrica *versus* assimétrica)

Localidade (região administrativa)

Origem dos pais

Tipo de fala (relatada *versus* original)

Linguísticas (ou internas):

Função sintática

Presença ou ausência de vocativo

Entonação (frases interrogativas *versus* não interrogativas)

Polaridade da sentença (afirmativa *versus* negativa)

Referencialidade (genérica *versus* específica)

Paralelismo formal

A maioria dos fatores aqui analisados foram fatores analisados e relevantes em outras pesquisas do mesmo tipo. A presença ou não de vocativo, a polaridade da sentença e a origem dos pais foram fatores inovadores, não discutidos nos trabalhos aqui revisados. É importante dizer, porém, que Lucca (2005) lança em seu trabalho uma semente para que a análise sobre a origem dos pais fosse feita, já que seus resultados apontaram para a importância dessa origem em lugares onde havia a maior ocorrência do pronome *tu* em Brasília, a saber, em Ceilândia. Os fatores polaridade da sentença e presença de vocativo foram fatores de controle. Tipo de relato, referencialidade, tipo de relacionamento entre os interlocutores e paralelismo formal foram fatores estudados na maioria das pesquisas realizadas sobre os pronomes que aqui revisamos. Tais fatores, em maior ou menor grau, influenciaram de forma contundente a variação dos pronomes. Por fim, a função sintática, sexo, faixa etária e localidade são fatores que, além de sempre recorrentes nas demais pesquisas, influenciam de forma definitiva a interpretação que podemos atribuir aos fenômenos em análise. O controle da função sintática é fundamental para a análise gramatical do fenômeno linguístico; o controle da faixa etária ajuda a indicar o percurso da variação no tempo, e o sexo e a localidade ajudam a entender a variação no espaço, além de auxiliarem acerca do entendimento sobre o significado social por detrás da variação.

Cada dado de nossa amostra passou pela codificação segundo as características apresentadas anteriormente e, na sequência, os dados foram rodados pelos programas. Fizemos vários testes, e nossa análise passará, como já mencionado, por rodadas eneárias, sem seleção e rodadas binárias, com seleção de fatores. Lembramos que as rodadas eneárias, apesar de não selecionarem variáveis, são fundamentais em nosso estudo uma vez que medem os pesos relativos considerando as três variantes juntas, tornando tais pesos contundentes para cada variante. As rodadas binárias serão basicamente em três fases: *tu* versus *você*, *tu* versus *cê* e *cê* versus *você*. Nessas rodadas haverá a seleção de variáveis, que significa, basicamente, que tal variável apresenta significância estatística (cf. Guy, 2007; Naro, 2003; Scherre & Naro, 2003). Através das rodadas binárias podemos analisar o comportamento de cada fator com os demais e podemos também observar a influência de cada fator estudado em função de cada uma das variantes, como já tivemos oportunidade de explicar.

A seguir, teremos a apreciação dos resultados em função dos fatores que influenciam no comportamento das variáveis em nossa amostra, ou seja, que foram selecionados pelo programa *GoldvarbX*. Teremos também algumas considerações sobre fatores que não foram selecionados pelo programa, mas que apresentaram características cuja apreciação também servirá à análise, ou seja, discutiremos alguns desses fatores por entender que, mesmo descartados, são importantes para a compreensão geral tanto da análise quanto da variação em estudo.

Relembramos que o procedimento tomado em casos de fatores categóricos foi sempre o de retirá-los ou amalgamá-los, pois só é possível fazer tais testes com fenômenos variáveis.

5.4 Efeitos de natureza externa

“Sabe-se muito bem que todo e qualquer falante tende, em maior ou menor grau, a assimilar falas do grupo em que está inserido, consciente ou inconscientemente”

(Marta Scherre)

Inicialmente falaremos dos resultados em função dos fatores externos (ou sociais). Segundo a revisão bibliográfica feita no capítulo anterior, esses fatores são os que mais influenciam a variação em estudo.

5.4.1 Faixa etária e tipo de relação entre os interlocutores (simetria *versus* assimetria)

Controlamos a faixa etária por ser esse um fator que influenciou o fenômeno em análise em outros estudos do mesmo tipo, como tivemos oportunidade de comentar em tópicos anteriores. Além disso, gostaríamos de saber em que idade a variante *tu* entra para o repertório linguístico do brasileiro. Por isso escolhemos estudar a faixa etária específica de 7 a 15 anos, que se caracteriza por ser diretamente anterior, e às vezes interseccionada, às faixas etárias estudadas por Andrade (2004), Lucca (2005) e Dias (2007).

Os resultados da tabela 16, a seguir, foram extraídos das primeiras rodadas realizadas com todos os dados. Ela mostra os resultados gerais somente em percentuais, com as três variantes que entraram para os testes estatísticos, *você*, *cê* e *tu*, em relação às idades.

Tabela 16 – Variação *você/ cê/ tu* em Brasília em relação à idade (todos os dados)

Idade	<i>Você</i>	<i>Cê</i>	<i>Tu</i>
5 (1 menina)	59/ 64 = 92,2%	5/ 64 = 7,8%	0/ 64 = 0,0%
7 (6 informantes)	28/ 47 = 59,6%	13/ 47 = 27,7%	6/ 47 = 12,8%
8 (4 informantes)	7/ 33 = 21,2%	3/ 33 = 9,1%	23/ 33 = 69,7%
9 (4 informantes)	36/ 61 = 59,0%	18/ 61 = 29,5%	7/ 61 = 11,5%
10 (6 informantes)	92/169 = 54,4%	68/ 169 = 40,2%	9/ 169 = 5,3%
11 (2 informantes)	5/ 51 = 9,8%	35/ 51 = 68,6%	11/ 51 = 21,6%
12 (6 informantes)	29/ 98 = 29,6%	16/ 98 = 16,3%	53/ 98 = 54,1%
13 (5 informantes)	30/ 53 = 56,6%	5/ 53 = 9,4%	18/ 53 = 34,0%
14 (4 informantes)	22/139 = 15,8%	50/ 139 = 36,0%	67/ 139 = 48,2%
15 (6 informantes)	10/120 = 8,3%	16/ 120 = 13,3%	94/ 120 = 78,3%
Totais	318/ 835 = 38,1%	229/ 835 = 27,4%	288/ 835 = 34,5%

Antes, porém, de discorrermos sobre as faixas etárias, gostaríamos de fazer um adendo sobre a primeira idade que vimos na tabela acima, de 5 anos. Embora essa idade não pertença à faixa etária escolhida para o presente estudo, analisaremos tais ocorrências a fim de exemplificar a fala de uma criança com idade diretamente anterior à faixa etária em foco. Observa-se que na fala da informante não há registro do *tu* quando da gravação da amostra, porém, encontramos dados de *tu* dessa informante em

situações de observação participante. Apresentaremos, a seguir, alguns exemplos de dados de observação participante:

Dados do dia 27/07/2008 (falando com a pesquisadora).

L. (feminino, 4 anos): “tia, TU pega água pra mim?”

L.(feminino, 4 anos): “TU faz suco pra mim, tia, de morango?”

Dado do dia 23/12/2008 (falando com amiguinha)

L. (feminino, 5 anos): “TU tá doida, Duda?”

Dados do dia 19/07/2009 (falando com a mãe)

L.: (feminino, 5 anos) “foi TU, mãe!”

m.: “quem?”

L: “TU, TU, por quê? Não tá certo?”

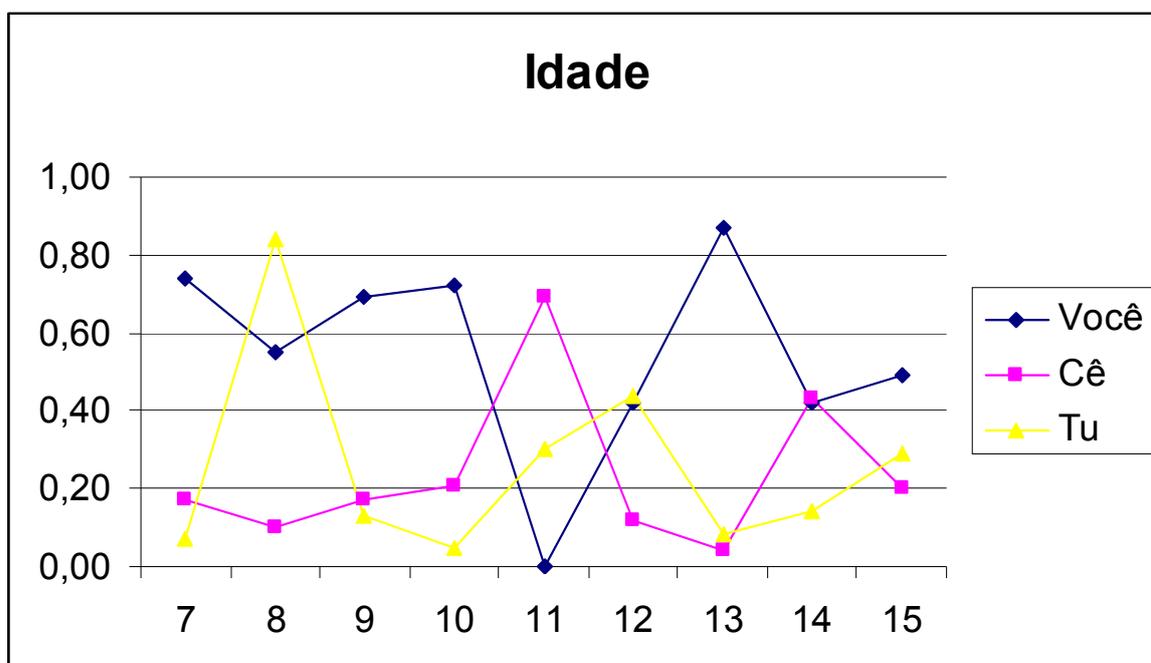
m.: “não, eu só achei engraçado você falando”

Na gravação, ocasião em que os dados de *tu* não emergiram, o diálogo ocorreu entre as irmãs L e T, a primeira com então 5 anos e a outra com 7 anos, em um ambiente familiar, na casa delas. Acreditamos que o *tu* não tenha surgido na gravação, de 40 minutos aproximadamente, por ser este um pronome ainda raro na fala das informantes e por ser mais raro também na localidade em que as informantes residem, fora da Vila, ou na fala dos pais das informantes, como verificaremos mais adiante. Foi possível coletar dados de *tu* das duas informantes, no formato de observação participante, porque o convívio entre elas e a pesquisadora é intenso.

Vale dizer que houve outra gravação em que não foi registrada a ocorrência do pronome *tu*. Tal registro foi semelhante à gravação das irmãs, pois foi feita entre três irmãos, na casa deles: duas meninas, uma de 10 e outra de 14 anos, e um menino de 12 anos, todos moradores do Sudoeste – a gravação foi feita por eles. Assim, temos que as duas gravações sem registro do *tu* foram feitas fora da Vila Planalto, com o conhecimento dos informantes e em ambiente familiar. Como havia a ausência de *tu* nessas gravações, que tinham informantes de diferentes idades, descartamos uma explicação dessa ausência acontecer em função da idade ou, em outras palavras, pelo fator faixa etária. As outras duas gravações de fora da Vila Planalto registraram a ocorrência do pronome *tu*. Nas demais gravações que compõem a amostra houve sempre a variação entre as três variantes, em maior ou menor grau.

Temos que a idade detalhada com todos os informantes foi sempre selecionada em primeiro lugar pelo programa, nas rodadas binárias. Isso significaria que a idade seria um fator importante para essa amostra. Observemos o gráfico a seguir, baseado em pesos:

Gráfico 02 – Idade



Ao analisar o gráfico 02 (idade), percebemos que as diferenças entre as idades são significativas quanto ao uso dos pronomes, mas essas diferenças não obedecem a uma lógica cronológica. Por exemplo, vimos na tabela anteriormente colocada, que apresenta resultados em percentuais, que as idades 8 e 15, em relação a variante *tu*, são mais parecidas entre si que as idades 14 e 15. Isso, como podemos verificar, não se repetiu nos resultados em pesos demonstrado no gráfico anterior, porém, mesmo em pesos, os resultados continuam não apresentando resultados sem ordem cronológica.

Assim, embora esse fator tenha sido selecionado pelo programa, um resultado também relevante é que não houve convergência nos testes realizados com as idades assim detalhadas. Significa dizer que os resultados poderiam estar enviesados ou, ainda, poderia indicar que alguns desses fatores em análise estivessem sobrepostos (cf. Guy & Zilles, 2007). Antes de buscar solução para as sobreposições, visualizamos a variação em comportamentos binários. Mesmo assim, os resultados continuaram a não indicar tendências cronológicas em relação às idades detalhadas.

Como já havíamos notado diferenças na fala das crianças pertencentes à Vila e das crianças de fora, principalmente no que tange ao tipo de gravação realizada e seus contextos de fala, e também pelo fato da amostra feita na Vila Planalto ter sido mais ampla e homogênea quanto às idades, o passo seguinte foi separar os dados desses dois grupos em diferentes localidades: dados da Vila Planalto, por um lado, e os de fora da Vila Planalto, por outro (através do fator de controle localidade). Dessa forma, passamos a rodar os dados em dois blocos assim distintos. Essa separação em nossa análise também foi coerente com o foco acerca das diferentes localidades dentro da grande Brasília. Porém, ao visualizar os resultados, pôde-se perceber uma relativa manutenção das tendências acerca das idades detalhadas nos dados somente da Vila Planalto. Novamente a idade foi selecionada como significativa, embora não tenha, novamente, apresentado nem convergência, nem uma ordem cronológica nos valores relativos. A falta de convergência, como dito anteriormente, novamente foi atribuída a algum tipo de enviesamento ainda pendente ou de sobreposição de grupos de fatores. Resolvemos, assim, estabelecer dois cortes etários relativamente naturais: 7 a 11 anos por um lado, crianças, e 12 a 15 anos por outro, adolescentes. Essa junção também se mostrou coerente com as diferenças nos tipos de gravações mencionadas anteriormente.

Tabela 17 – Variação *cê/ tu/ você* na Vila Planalto em relação à faixa etária

Faixa etária	Você	Cê	Tu
Crianças – de 7 a 11 anos	62/ 77 = 32%	77/ 191 = 40%	52/ 191 = 27%
Adolescentes – de 12 a 15 anos	87/ 383 = 23%	75/ 383 = 20%	221/ 383 = 58%
Totais	149/ 574 = 26%	152/ 574 = 26%	273/ 574 = 48%

Em termos dos resultados em percentuais, os números da tabela acima indicariam que haveria um aumento de ocorrência de *tu* na faixa etária de adolescentes, e, por outro lado, haveria a diminuição da ocorrência de *tu* na faixa etária de crianças. Porém, ao rodarmos os dados com pesos relativos, a faixa etária, com os dois cortes etários, não foi selecionada.

Esperávamos dois tipos de resultados em relação à faixa etária: 1) por um lado, as crianças estariam adquirindo cada vez mais cedo a variante *tu*. Nessa perspectiva, o *tu* seria produtivo já nas primeiras idades, ou na faixa etária de 7 a 11 anos, e seguiria crescendo em ocorrências nas idades seguintes, na faixa etária de 12 a 15 anos; 2) ou, por outro lado, o *tu* seria uma variante típica da adolescência. Assim, o *tu* seria raro nas

primeiras idades e sua ocorrência aumentaria consideravelmente na adolescência, até chegar a maior ocorrência na idade mais alta em nossa amostra, ou, se visto pela ótica das faixas etárias, o *tu* simplesmente seria bastante produtivo somente na faixa etária de 12 a 15 anos. Mas, nas duas hipóteses, a idade ou faixa etária sempre seria fator influenciador. A diferença entre as hipóteses 1 e 2 seria mais relacionada à quantidade de ocorrências nas primeiras idades ou faixa, porque já tínhamos resultados das dissertações de Lucca (2005) e Dias (2007), e ambas comprovaram a produtividade no *tu* na faixa etária de adolescentes, inclusive indicando que a adolescência seria a faixa etária prototípica para o uso de *tu*. Porém, esses resultados em que não houve a seleção pelo programa indicaram que a idade (ou a faixa etária) não tinha a importância que esperávamos para essa amostra, com os dois cortes estabelecidos.

Buscamos entender o que levou o programa a descartar a faixa etária. Assim, analisamos a rodada em que a faixa etária havia sido descartada e percebemos que, quando o tipo de relação entre os interlocutores foi selecionado pelo programa, a faixa etária deixava de ter relevância estatística. Um teste fundamental para esse entendimento dos resultados foi o cruzamento dos dados. Cruzamos a faixa etária em função do tipo de relação e não encontramos dados assimétricos na faixa etária de adolescentes (nem mesmo com todos os dados em jogo). Percebemos que a junção das diversas idades em duas faixas etárias contribuiu para os resultados quanto ao tipo de relação. Esse resultado indicou que o tipo de relação (simetria *versus* assimetria) e a faixa etária estavam sobrepostos e que o tipo de relação era mais importante para a análise dos dados da Vila Planalto que a faixa etária e, por isso, esse último fator havia sido descartado pelo programa. Vejamos a seguir o resultado do cruzamento dos fatores faixa etária e tipo de relação com todos os dados (de dentro e fora da Vila Planalto):

Tabela 18 – Variação *você*, *cê* e *tu* em relação ao tipo de relação e a faixa etária (com todos os dados)

Faixa etária	7 – 11 anos	12 – 15 anos
Tipo de relação		
Simétrica:		
<i>Você</i>	155/ 245 = 63%	91/ 410 = 22%
<i>Cê</i>	68/ 245 = 28%	87/ 410 = 21%
<i>Tu</i>	22/ 245 = 9%	232/ 410 = 57%
Assimétrica:		
<i>Você</i>	72/ 180 = 40%	Não há dados
<i>Cê</i>	74/ 180 = 41%	Não há dados
<i>Tu</i>	34/ 180 = 19%	Não há dados

A decisão de controlar o tipo de relação entre os informantes surgiu, como explicado anteriormente, da necessidade de fazer diferentes levantamentos para a composição do *corpus*. A ideia inicial era gravar somente uma localidade e somente um tipo de relação: de fala casual, entre amigos. Mas, como já dito anteriormente, acabamos realizando gravações fora da Vila também, a título de experiência, e gravamos também em diferentes contextos, por causa da gravação da faixa etária de 7 a 11 anos, em que os informantes eram imaturos para realizar de forma autônoma as gravações para a pesquisa. Assim, esse fator acabou se tornando também um controle dos diferentes contextos de gravação que, por sua vez, também polarizou as diferentes faixas etárias. A hipótese inicial para essa variável seria de que em relações simétricas a variante *tu* seria favorecida.

Fizemos testes cruzando os tipos de relação e as faixas etárias somente com os dados da Vila Planalto. Vejamos a seguir:

Tabela 19 – Variação *você*, *cê* e *tu* na Vila Planalto em relação ao tipo de relação e a faixa etária

Faixa etária	7 – 11 anos	12 – 15 anos
Tipo de relação		
Simétrica:		
<i>Você</i>	5/ 33 = 15%	87/ 383 = 23%
<i>Cê</i>	9/ 33 = 27%	75/ 383 = 20%
<i>Tu</i>	19/ 33 = 58%	221/ 383 = 58%
Assimétrica:		
<i>Você</i>	57/ 158 = 36%	Não há dados
<i>Cê</i>	68/ 158 = 43%	Não há dados
<i>Tu</i>	33/ 158 = 21%	Não há dados

No teste que gerou a tabela 19, como já esperávamos, a variante *tu* ficou ainda mais consistente nos dados só da Vila Planalto. Os números da tabela indicaram também uma tendência de, na Vila Planalto, a variante *tu* ocorrer de forma independente da faixa etária, ou seja, é como se os falantes dessa localidade já estivessem adquirindo, em tenra idade, em relações simétricas, a variante *tu*, pois os percentuais de *tu*, tanto na faixa de 7 a 11 anos como na faixa de 12 a 15 anos, foram de 58%. Esse aumento significativo do percentual em favor da variante *tu* nos dados somente da Vila pode significar que, nesta região administrativa, haja o maior uso da variante *tu* que nas demais regiões estudadas na presente pesquisa, mas, embora os resultados apontem para isto, há algumas ressalvas a essa conclusão, uma vez que: a) os dados de fora da Vila não foram colhidos no mesmo tipo de contexto dos dados da Vila; b) não há equilíbrio na quantidade de dados de fora da Vila para a realização de uma comparação direta.

Fizemos também um teste retirando o tipo de relação da rodada com os dados da Vila e a faixa etária não voltou a ser selecionada. Podemos inferir disto que a faixa etária não influencia a ocorrência de *tu* para a amostra da Vila Planalto, justamente porque ambas as faixas estudadas produziram *tu* de forma bastante semelhante, como já tivemos oportunidade de apontar. Vale ressaltar que nossos resultados apontaram para essa tendência somente quando os interlocutores estão em relações simétricas. É fato que esses resultados confirmam a hipótese 1 anteriormente colocada sobre as faixas etárias, em que os mais novos estão adquirindo cada vez mais cedo a variante *tu*. Os

resultados confirmam também a hipótese de falas simétricas favorecerem a ocorrência da variante *tu*.

Por fim, fizemos uma rodada eneária. Lembramos que este tipo de rodada tem a seguinte ressalva: não faz seleção. Alcançamos o seguinte resultado em função da faixa etária:

Tabela 20 – Resultados de pesos relativos da variação *você/ cê/ tu* na Vila Planalto em relação à faixa etária

Faixa etária	Você	Cê	Tu
7 a 11anos	62/ 191 = 32% Peso = 0,33	77/ 191 = 40% Peso = 0,37	52/ 191 = 27% Peso = 0,30
12 a 15 anos	87/ 383 = 23% Peso = 0,33	75/ 383 = 20% Peso = 0,30	221/ 383 = 58% Peso = 0,37
Total	149/ 574 = 26%	152/ 574 = 26%	273/ 574 = 48%

Observamos na tabela 20 que *o você* está equilibrado nas duas faixas etárias, com pesos iguais, sobre as demais variantes, porém, podemos observar um comportamento inverso entre uma e outra: na faixa de crianças, há o relativo aumento da variante *cê*, enquanto na faixa de adolescentes há o relativo aumento da variante *tu*. Mesmo assim, o que fica evidenciado a partir da tabela 20 é que as três variantes apresentam um comportamento de equilíbrio ou não influência pelas diferentes faixas etárias, pois seus pesos ficaram muito próximos ao valor de referência, que é 0,33.

Conclui-se em relação às faixas etárias que, mesmo sem seleção para a amostra da Vila Planalto, essa análise contribuiu de forma expressiva para o entendimento dos comportamentos das variantes em análise. O fator tipo de relação influenciou de forma indubitável os resultados; esta influência também refletiu no fator faixa etária. Ou seja, nesta amostra, tais fatores estavam sobrepostos e o tipo de relação foi mais relevante que a faixa etária nesse conjunto de dados. Os resultados da rodada eneária para o tipo de relação serão vistos mais adiante.

Podemos inferir também que, com os dados de fora da Vila, os resultados indicaram (a exemplo de quando toda a amostra esteve em análise, com 835 dados) um favorecimento do *tu* por parte da faixa etária de 12 a 15 anos. Mas, ao analisar os dados, percebemos que se tratava de um enviesamento da fala de um único informante (y) de 12 anos, que usou o *tu* de forma quase isolada nos dados de fora da Vila, pois, além dos

dados dele (13 ocorrências de *tu*), houve apenas mais três informantes que usaram o *tu*, todos da faixa etária de 7 a 11 anos: uma menina que usou a variante apenas duas vezes e mais dois meninos que usaram o *tu* cada um uma única vez, totalizando apenas mais 4 ocorrências dessa variante para a amostra de fora da Vila. Fizemos um teste para verificar se retirando os dados deste informante a faixa etária ou o tipo de relação seria selecionado. O resultado foi que o programa selecionou e depois descartou o fator faixa etária e o tipo de relação não foi nem selecionado nem descartado. Esse resultado significa que o programa não pôde ter certezas quanto aos fatores analisados, indicando que, apesar da importância desses fatores para a análise, os dados não chegaram a revelar as suas tendências. Como os dados desse informante enviesaram os resultados, e os dados sem esse informante não puderam apontar as tendências quanto aos fatores em jogo, precisaremos de investigações futuras de dados de fora da Vila para avaliarmos se há ou não a influência do fator faixa etária e do tipo de relação no uso dos pronomes nessas localidades (fora da Vila).

Explicamos, até aqui, o fator faixa etária em contraste com o fator tipo de relação justamente porque ambos os fatores estavam sobrepostos, como tivemos oportunidade de demonstrar. Quanto ao tipo de relação, gostaríamos ainda de ponderar sobre alguns resultados de nossa análise.

Considerando somente os dados da Vila Planalto, o fator tipo de relação foi selecionado em duas das três rodadas binárias e somente nas rodadas em que o *tu* estava em foco, com um *range*²⁰ de 36 (diferença de pesos entre o *tu* – com 0,58 e 0,59; *você* – com 0,22, e *cê* – 0,23, respectivamente) em ambas as rodadas.

As variantes *cê* e *você* tiveram um comportamento semelhante quando a variante *tu* estava em contraste. Mas vale lembrar que tais resultados não separaram os casos categóricos (entre simetria e assimetria) que correspondem às idades de 12 a 15 anos, portanto, é necessário que haja cautela para realizar inferências (cf. tabelas 18 e 19 do presente estudo). Através do cruzamento dos dados entre as variáveis faixa etária e tipo de relação, percebe-se que o percentual da variante *tu* na simetria é equivalente nas duas faixas etárias, de 7 a 11 anos e de 12 a 15 anos, ao passo que em relações assimétricas percebemos uma acentuada queda na ocorrência da variante *tu*. Vejamos, a seguir, os resultados da rodada eneária:

²⁰ *Range* é basicamente a diferença entre os pesos apresentados pelo programa em função de suas variantes.

Tabela 21 – Tipo de relação entre os interlocutores (dados da Vila)

Tipo de relação	Você	Cê	Tu
Simétrica	0,27	0,26	0,47
Assimétrica	0,38	0,40	0,22

A rodada eneária, além de confirmar a hipótese sobre o favorecimento do *tu* em relações simétricas, também revela o favorecimento da variante *cê* em relações assimétricas, com uma diferença mais significativa para esta rodada, que tem como número de referência o 0,33 (nas rodadas binárias esse número é 0,50). Percebemos também um favorecimento da variante *você* nas relações assimétricas, semelhante ao *cê*.

Finalmente, gostaríamos de comparar nossos resultados sobre o tipo de relação com as três dissertações realizadas em Brasília. Andrade (2004) não estudou esse tipo de variável, uma vez que o *corpus* utilizado foi constituído por entrevistas labovianas. Podemos inferir que o estudo de Lucca (2005) tenha sido somente com interlocutores simétricos, uma vez que suas gravações foram ocultas e entre amigos. A autora controlou, porém, se entre tais amigos havia uma relação entre pares ou não pares. Seus resultados indicaram que, entre os pares, a ocorrência de *tu* era favorecida. Por fim, infere-se do estudo de Dias (2007) que este foi mesclado, com registros simétricos e assimétricos de interação entre interlocutores. A autora não controlou a simetria entre falantes, mas o grau de intimidade e/ou amizade. Seus resultados indicaram que, quanto mais íntimo, mais o *tu* tende a ocorrer. Os resultados que encontramos no presente estudo para o tipo de relação se coadunam, de certa maneira, com os resultados de Lucca e Dias, pois a nossa ideia de simetria/assimetria nas relações se assemelham às ideias de pares (em contraste com não pares) e íntimos (em contraste com não íntimos) (guardadas algumas diferenças) estudados nas pesquisas das autoras citadas. Podemos dizer, pelo menos, que essa comparação cabe na presente análise.

5.4.2 Origem dos pais

Uma investigação necessária para o entendimento da variação é a influência da fala dos pais sobre a fala das crianças e adolescentes, para sabermos, entre outras coisas, as origens em que se baseia a variedade que está se formando em Brasília. Lucca (2005) realizou pesquisa em três regiões administrativas da grande Brasília, a saber, Ceilândia, Taguatinga e Brasília. A pesquisadora relacionou a entrada e a elevada ocorrência de *tu* no repertório linguístico dos falantes pertencentes a Ceilândia, os que mais favoreceram a ocorrência da variante *tu*, ao fato de que nessa região administrativa há um grande contingente de imigrantes nordestinos, ou seja, a autora sugeriu que a variante *tu* se deve à influência da fala dos pais (nordestinos) das crianças estudadas. Assim, seria necessário medir até que ponto os pais²¹ influenciam a fala das crianças e adolescentes e até que ponto os pais influenciam essa fala.

Inicialmente, fizemos a codificação de cada estado de origem do pai e cada estado de origem da mãe de todos os informantes. Porém, como o resultado ficou bastante diversificado, o passo seguinte foi unir maximamente as semelhanças e separar as diferenças em alguns blocos que nos permitissem perceber tendências. Ressaltamos que houve uma análise muito detalhada de diferentes testes para que chegássemos à composição de tais blocos, e que o resultado não poderia deixar de ser complexo, mesmo que a explicação também tivesse de ser complexa. Assim, depois de vários testes, a mais coerente combinação a que chegamos foi a seguinte:

a) mãe e pai do nordeste, sendo ambos do mesmo estado – Vila Planalto, 8 falantes

b) mãe e pai do nordeste, mas de estados diferentes – Vila Planalto, 3 falantes mais 1 falante de “sentimento e origens nordestinas – família piauiense”

c) um dos pais do nordeste – Vila Planalto, 7 falantes (com predominância de mãe de Brasília e pai do nordeste, com somente um caso de pai de Brasília e mãe do nordeste – assim, poderíamos chamar também de “um dos pais de Brasília”)

²¹ Os pais (ou avós) das crianças estudadas passaram ou passam por uma fase de difusão dialetal. Sabemos que questões ligadas à identidade que se tem com o lugar de origem, o que a sociolinguística interacional (Giles, 1980) chama de acomodação, influencia a fala das pessoas em fase de difusão dialetal (cf. Bortoni-Ricardo, 1985 e Borges, 2009). No presente estudo não investigaremos essas influências identitárias. Mas pretendemos, em trabalhos futuros, analisar esse aspecto.

d) mãe e pai brasilienses – Vila Planalto, 8 falantes

e) mãe e pais de Minas Gerais – Vila Planalto, 1 só falante

f) um dos pais de Minas Gerais (mesmo quando um deles era do nordeste) – Vila Planalto, 2 falantes

g) um pai de Brasília e outro do Mato Grosso – Vila Planalto, 1 só falante

h) sem origem nordestina – de fora da Vila, 8 falantes (porém, há um falante (y) que não tem pais nordestinos mas têm traços de família nordestina, já que seus demais familiares residem em Ceilândia)

Apesar da diversidade de origens dos pais das crianças da Vila Planalto e de fora da Vila, só houve 4 casos de influências de origem de Minas Gerais, em toda a Região Sudeste; apenas um caso de influência de origem do Rio Grande do Sul, no Sul (em que o outro pai era de Goiás, mas esse informante ficou codificado como sendo influência nordestina de estados diferentes, porque o restante da família era do Piauí, ele mesmo havia nascido no Piauí, mas veio para Brasília com 1 ano; e ainda, o informante demonstrou fortes características dialetais nordestina em sua fala); um caso de influência de origem de Tocantins, no Norte; apenas um caso de um dos pais ser de Mato Grosso, na Região Centro-Oeste. Sendo assim, a grande maioria das origens é a nordestina seguida pela brasiliense. Os resultados gerais das origens dos pais estão descritos na tabela a seguir:

Tabela 22 – Variação *você/ cê/ tu* em relação à origem dos pais (todos os dados)

Origem dos pais	Você	Cê	Tu
A* – Ambos os pais nordestinos, do mesmo estado (8 falantes)	41/ 204 = 20%	35/ 204 = 17%	128/ 204 = 63%
B – Ambos os pais nordestinos, mas de estados diferentes (4 falantes)	15/ 38 = 39%	12/ 38 = 32%	11/ 38 = 29%
C – mãe de Brasília e pai do nordeste (1 dos 7 casos é o inverso)	24/ 58 = 41%	9/ 58 = 16%	25/ 58 = 43%
D – Ambos os pais brasilienses (8 falantes)	38/ 148 = 26%	50/ 148 = 34%	60/ 148 = 41%
E – Ambos os pais mineiros (1 falante)	5/ 37 = 14%	26/ 37 = 70%	6/ 37 = 16%
F – um dos pais de Minas Gerais e o outro do Nordeste (2 falantes)	15/ 74 = 20%	18/ 74 = 24%	41/ 74 = 55%
G – mãe de Brasília e pai de Mato Grosso (1 falante)	11/ 15 = 73%	2/ 15 = 13%	2/ 15 = 13%
H – fora da Vila sem origens nordestinas (9 falantes)	169/ 261 = 65%	77/ 261 = 30%	15/ 261 = 6%
Totais	149/574 = 26%	152/574 = 26%	273/574 = 48%

(*mantivemos a relação das origens representadas por letras para auxiliar a comparação dos resultados apresentados nas páginas vindouras)

A tabela 22 apresenta resultados em percentuais para a origem dos pais das crianças estudadas. Seus números relativos sugerem que a origem que mais favoreceria a ocorrência de *tu* é a origem nordestina de ambos os pais, com 63%, portanto acima da média 48%. A ocorrência de *tu* também ficou acima da média, de forma surpreendente, quando os pais eram originários de Minas Gerais e do Nordeste. Abaixo da média, mas sendo a terceira maior ocorrência dessa variante, estão as crianças cuja origem dos pais é nordestina e brasiliense, seguidos pela origem somente brasiliense. Curiosamente, a ocorrência do *tu* não foi favorecida pelas crianças cujos pais são nordestinos, mas de estados diferentes.

A variante *você* na tabela acima foi fortemente favorecida pela origem Brasília e Mato Grosso, alcançando o percentual de 73%, e pela origem “sem origens nordestinas”,

com 65%. O *você* também foi relativamente favorecido pelas origens: mãe brasiliense e pai nordestino e por ambos os pais nordestinos, mas de estados diferentes.

O *cê*, por sua vez, foi fortemente favorecido pelo informante cujos pais são mineiros, alcançando esta variante o percentual de 70%. Esse resultado não surpreende, mas, podemos notar, na tabela acima, que as origens: ambos os pais brasilienses e ambos os pais do nordeste (mas de estados diferentes) favoreceram relativamente a ocorrência de *cê*, ao passo que a origem de apenas um dos pais de Minas Gerais não favoreceu esta ocorrência, ficando seu percentual aproximado da média. Andrade (2004) realizou pesquisa semelhante, mas entre as variantes *cê*, *ocê* e *você*, e comparou resultados de Brasília com os resultados de Minas Gerais. A autora concluiu que a variante *cê* é mais recorrente em Minas Gerais que em Brasília. Nessa linha de raciocínio, quando observamos a origem Minas e Nordeste, surpreende que os resultados apontem para o não favorecimento de *cê* e para o robusto favorecimento de *tu*.

Ao realizarmos as rodadas binárias, a fim de medir a importância de cada fator para a variação em estudo, observamos que, em todas elas, a origem dos pais foi selecionada. Isto significa dizer que a origem dos pais influencia de forma contundente a variação entre os pronomes em estudo.

Apesar das rodadas binárias serem fundamentais para a seleção dos fatores que incidem sobre a variação e, ainda, trazerem luz sobre os comportamentos de uma variante em relação à outra, é fundamental para a presente análise a apreciação de rodadas eneárias, pois estas apresentam pesos contundentes para cada um dos pronomes em estudo, já que medem as três variantes em conjunto. Vejamos, a seguir, os resultados referentes à rodada eneária.

Tabela 23 – Variação *você/ cê/ tu* em relação às origens dos pais, resultados em pesos

Origens dos pais	Você	Cê	Tu
a) ambos os pais nordestinos, do mesmo estado (8 falantes)	0,20	0,23	0,57
b) ambos os pais nordestinos, mas de estados diferentes (4 falantes)	0,33	0,44	0,23
c) mãe de Brasília e pai do nordeste (1 dos 7 casos é o inverso)	0,25	0,15	0,60
d) ambos os pais brasilienses (8 falantes)	0,22	0,33	0,45
e) Ambos os pais mineiros (1 falante)	0,15	0,70	0,15
f) um dos pais de Minas Gerais e o outro do Nordeste (2 falantes)	0,30	0,35	0,35
g) mãe de Brasília e pai de Mato Grosso (1 falante)	0,82	0,11	0,07
h) fora da Vila Planalto – sem origens nordestinas (9 falantes)	0,50	0,39	0,09

Ao analisar os pesos da tabela anterior, pode-se inferir das seguintes origens:

- a) Ambos os pais nordestinos, do mesmo estado (8 falantes) – o *range* entre os pesos de *tu*, variante favorecida, e o *cê*, variante em nível intermediário, é de 34 pontos, o que configura uma diferença significativa em favorecimento da variante *tu*. Essa origem também desfavorece a variante *você*. Tal resultado é o esperado para esta origem quanto ao uso da variante *tu*. É interessante, porém, o fato de a variante *cê* ter sido tão (des)favorecida quanto à variante *você* nessa amostra, como já tivemos oportunidade de comentar, por não termos conhecimento de estudos que apontem para o uso da variante *cê* no nordeste. Temos duas maneiras de explicar esse resultado: 1) talvez sejam as forças de

difusão dialetal e a própria formação dialetal em Brasília, atuando para que os informantes usem a variante *cê*; 2) Talvez esse resultado também ocorra pela própria tipicidade da variante *cê*, que exprime, a exemplo da variante *tu*, uma maior informalidade ante a variante *você* (lembramos que a amostra foi composta em uma linguagem informal).

- b) Ambos os pais nordestinos, mas de estados diferentes (4 falantes) – os resultados para esta origem sugerem que a variante mais favorecida é o *cê*, seguida do *você*, sendo o *tu* desfavorecido por essa configuração de origem. É curioso, como já mencionado anteriormente, o fato de a origem nordestina de estados diferentes influenciar as variantes em estudo de forma tão diversa da origem nordestina de mesmo estado. Talvez isso funcione como se houvesse uma espécie de difusão dialetal interna, dentro de uma microlocalidade (como o ambiente familiar, por exemplo), que incide nas variantes quando a origem dos pais não é a mesma. Ao pensarmos isso, estaríamos apenas verificando em microambientes o que é largamente verificável em macroambientes.
- c) Mãe de Brasília e pai do nordeste (apenas 1 dos 7 casos é o inverso, com mãe do nordeste e pai de Brasília) – essa origem, segundo os resultados da tabela anterior, favorece fortemente a ocorrência do pronome *tu* e, por outro lado, a ocorrência do *cê* ficou desfavorecida. Acreditamos que o desfavorecimento do *cê* seja decorrente do forte favorecimento de *tu*, em função de ambos os pronomes serem de contextos informais. Talvez o resultado elevado da ocorrência do pronome *tu* por essa origem seja um forte indício que essa variante esteja em expansão e se estabelecendo em Brasília. Aqui, parece que temos duas forças propulsoras da variante *tu* agindo em conjunto.
- d) Ambos os pais brasilienses (8 falantes) – essa origem favorece a ocorrência do pronome *tu*, com o peso de 0,45. Por sua vez, o pronome *cê* fica em nível intermediário, sua ocorrência pode ser considerada como neutra, pois seu peso de 0,32 fica, praticamente, na média (0,33). Por fim, tal origem desfavorece a ocorrência do *você*, com o peso de 0,22. Esse resultado corrobora a hipótese levantada anteriormente de que em Brasília a variante *tu* está se estabelecendo, pelo menos em contextos de falas casuais.

- e) Ambos os pais mineiros (1 falante) – essa origem indicou, de forma prevista, um forte favorecimento da forma *cê*, com o peso de 0,70. As demais variantes, *você* e *tu* foram igualmente desfavorecidas, ficando ambas com o mesmo peso de 0,15. Esse resultado, mesmo sendo de um só falante, corrobora a teoria colocada anteriormente de que, quando os pais são originários do mesmo lugar, parece haver uma facilidade maior de manter e passar suas influências linguísticas, mesmo quando se vive em um lugar de intenso contato entre variedades linguísticas, como é o caso de Brasília. O inverso também parece proceder, de que, quando de origens diversas, há uma microdifusão dialetal no seio familiar.
- f) Um dos pais de Minas Gerais e o outro do Nordeste (2 falantes) – essa origem parece estar lidando com influências antagônicas acerca da variação dos pronomes. As três formas encontram-se neutralizadas, muito próximas à média. O *você*, com o peso de 0,30 e as demais, o *tu* e o *cê*, com o peso de 0,35 cada uma. Tais resultados também corroboram a teoria colocada anteriormente acerca da microdifusão dialetal em seio familiar.
- g) Mãe de Brasília e pai de Mato Grosso (1 falante) – essa origem foi separada justamente porque destoava das demais influências tanto em termos geográficos como em termos de tendências quanto aos usos dos pronomes. Tal origem, apesar de ter apenas uma informante, indicou um forte favorecimento para a variante *você*, com o peso de 0,82, desfavorecendo veementemente as demais variantes. Como não encontramos estudos acerca dos usos dos pronomes em Mato Grosso, e os resultados dessa informante destoam dos demais resultados encontrados na Vila Planalto, acreditamos que seria necessário mais estudo acerca desse tipo de origem para que se chegue a conclusões. Por ora, é suficiente lembrar que o *você* parece ser a variante mais utilizada na fala das pessoas que passam por estágio de difusão dialetal.
- h) Sem origens nordestinas diretas (9 informantes), dados de fora da Vila – essa origem demonstrou tendências de usar mais a variante *você*. É curioso que, nas origens dos pais desses informantes, há somente um casal de pais provenientes do mesmo lugar, o restante das origens são todos de pais de diferentes lugares,

como mãe do Rio de Janeiro e pai do Amazonas, por exemplo. Outro fato que chama a atenção, e que já foi mencionado em tópico anterior, foi o caso do informante y, que apresentou 13 das 17 ocorrências de *tu* na amostra de fora da Vila. Portanto, trata-se de um informante que envia os dados em favor da ocorrência da variante *tu*. O restante da família do informante y que reside em Brasília mora em Ceilândia, localidade formada por um forte contingente nordestino.

Temos que as origens estudadas de ambas as localidades, sejam elas de dentro e de fora da Vila Planalto, apresentaram, de maneira geral, um forte indício de que o *tu* na fala brasiliense seja proveniente do nordeste, corroborando a hipótese levantada por Lucca (2005). Esse grupo de fatores, como já informado, foi selecionado em todas as rodadas binárias de pesos relativos. Significa dizer que a origem dos pais é fator contundente nas tendências dialetais das crianças estudadas. O fato de a ocorrência de *cê* ser favorecida quando a origem dos pais é mineira é mais uma prova dessa relevância.

Ao analisar os resultados, percebemos que a origem dos pais, apesar de relevante, não é unicamente determinante, uma vez que as tendências apontadas pelos resultados já se distanciam em alguns casos, ainda que sutilmente, das tendências linguísticas correspondentes aos locais de origem dos pais. Por exemplo, as origens mineira e nordestina fizeram com que as tendências para as formas *cê* e *tu* ficassem igualmente neutralizadas, como foi apontado pelos resultados da rodada eneária. Temos de considerar também que a origem brasiliense já é uma origem contundente em termos de análises, que aponta resultados diversos de outros estados. Uma possível influência para que a origem brasiliense tenha despontado como força propulsora em favor de uma ou outra variante talvez seja uma possível aquisição dialetal também oriunda dos pais, em que as bases para essa aquisição seja a formação de uma variedade local. Na mesma linha de raciocínio, a origem nordestina, quando favorece o *cê* em segundo lugar, antes do *você*, pode significar que essa tendência tenha sido adquirida entre os pais, já que desconhecemos o *cê* como fazendo parte de alguma variedade nordestina. E, ainda, essas tendências podem representar que essa aquisição (diferente das características dialetais do lugar de origem dos pais), além de ser pelos pais, pode representar um aspecto da focalização dialetal que ocorre em Brasília.

Ao encontro da ideia de focalização, há o fato de termos considerado a origem dos pais brasilienses por dois motivos: a) por esse tipo de origem já ser bastante comum

em Brasília, chegando à aproximadamente metade da população atual; b) por ter (relativamente) senão características indubitavelmente próprias, pelo menos características diferenciadas das demais origens. Vale dizer, ainda, que a própria origem dos pais, em nosso *corpus*, não pode ser considerada taxativa em características dialetais, uma vez que estes pais estão passando (ou já passaram) por um período de difusão dialetal.

É importante afirmar que, embora haja evidências em nossa pesquisa que confirmam a influência dos dialetos de origem dos pais na fala das crianças e adolescentes em estudo, ao que tudo indica, essa influência está de certa maneira diluída em um conjunto complexo de influências que levam as crianças a adquirir e usar uma variedade linguística ainda aberta e em fase de formação, não somente em termos de pronomes.

5.4.3 Sexo

Como já mencionado, diferentes pesquisas na área da sociolinguística (mas não somente nessa área) têm demonstrado que há diferenças entre a fala dos homens e a fala das mulheres. Essas diferenças refletem, entre outras coisas, as diferenças culturais existentes entre os sexos. Nesse sentido, os linguistas são consensuais em abraçar tais diferenças, porém, não há consenso ainda sobre as tendências linguísticas que um dos sexos teria em detrimento do outro. Sabemos, porém, que questões de prestígio e identidade influenciam os comportamentos de ambos os sexos com relação à fala, às vezes, em diferentes direções (cf. Borges, 2009, p. 32).

A hipótese para a investigação desse fator no presente estudo é de que, apesar do sexo masculino favorecer a ocorrência da variante *tu*, como comprovado em Dias (2007) e em Lucca (2005), acreditamos que o sexo feminino vem assinalando um crescimento em favor dessa variante, uma vez que: 1) tal variante, em nossa localidade, não sofre estigma e 2) há a comprovação da hipótese de aquisição dessa variante na primeira faixa etária estudada. Observemos as tabelas a seguir, baseadas nos resultados obtidos em 2005 e 2007, respectivamente:

Tabela 24 – Resultados gerais quanto ao sexo (estudo de 2005)

Sexo	Tu (<i>versus</i> você e cê)	
Feminino	4/ 17 = 23%	Peso = 0,09
Masculino	300/ 380 = 78%	Peso = 0,55

(Lucca, 2005, p. 83 – com adaptações)

Tabela 25 – Resultados do sexo em função da região administrativa (estudo de 2005)

Sexo	Tu em Ceilândia	Tu em Taguatinga	Tu em Brasília
Feminino	2/ 4 = 50%	2/ 19 = 11%	3/ 12 = 25%
Masculino	103/ 117 = 88%	160/ 225 = 71%	57/ 76 = 75%

(Lucca, 2005, p. 100 – com adaptações)

Tabela 26 – Resultados gerais, com as três faixas etárias em função do sexo (estudo de 2007)

Sexo	Tu (<i>versus</i> você e cê)
Feminino	50/ 464 = 10,8% Peso = 0,41
Masculino	65/ 436 = 14,9% Peso = 0,60
Total	115/ 900 = 12,8%

(Dias, 2007, p. 75 – com adaptações)

Tabela 27 – Resultados quanto ao sexo, mas somente da faixa etária de 13 a 19 anos (estudo de 2007)

Sexo	Tu (<i>versus</i> você e cê)
Feminino	24/ 106 = 22,6%
Masculino	27/ 65 = 41,5%

(Dias, 2007, p. 76 – com adaptações)

Os resultados gerais de Lucca e Dias, em faixas etárias similares, foram equivalentes a respeito das frequências de *tu* em falantes do sexo feminino, pois ficaram em torno de 22 a 23%.

Lucca, que estudou somente falantes adolescentes, conclui que o sexo masculino favorece de maneira determinante a ocorrência de *tu*, com um percentual de 78%, ao passo que o sexo feminino a desfavorece, com um percentual de 23%. Chama a atenção, apesar da baixa quantidade de dados, o resultado do cruzamento do fator sexo com as regiões administrativas estudadas por Lucca e apresentadas na tabela 25. Os resultados obtidos em Ceilândia sugeriram que, nesta localidade, o sexo feminino desfavorece menos esta variante do que nas demais localidades, pois seu percentual de *tu* em falas femininas foi de 50%.

Os resultados de Dias (2007, p. 76-77) para a variante *tu* ficam equivalentes aos resultados de Lucca quanto ao sexo feminino, como já mencionado. Porém, tais resultados foram diferentes quanto ao sexo masculino. Dias, que estudou a fala somente na região administrativa de Brasília, e de três faixas etárias diferentes, explica que as diferenças quanto à recorrência dos dados de *tu* nos resultados de 2005 e 2007, podem ter ocorrido, também, devido às diferenças existentes acerca da composição das amostras. Mesmo que o *tu* tenha ocorrido menos na amostra de 2007, os pesos encontrados por Dias confirmaram o favorecimento do *tu* em falantes masculinos bem como o desfavorecimento do *tu* em falantes femininos.

Embora os resultados das amostras de 2005 e 2007 tenham sido percentualmente similares para o sexo feminino, nos resultados de 2007 há uma diferença menor entre os sexos que nos resultados de 2005. Significaria dizer que houve uma tendência maior de mulheres usarem a variante *tu* na amostra de 2007 que na amostra de 2005.

Para a presente análise, realizamos, como explicado anteriormente, testes estatísticos rodando todos os informantes juntos, com dados da Vila e de fora da Vila, e testes com tais localidades separadas. Realizamos também rodadas binárias, a fim de testar a influência dos fatores através da seleção e, também, rodada eneária a fim de medir o real peso de cada variante frente às demais.

Os resultados a que chegamos sugeriram, de forma geral, que a variante *tu* é favorecida pelo sexo masculino. Nas rodadas binárias que realizamos com todos os informantes o fator sexo foi sempre selecionado pelo programa. Ao rodarmos, porém, os dados da Vila Planalto, o sexo deixou de ser selecionado em uma das rodadas binárias, na rodada *tu/ cê*.

Vejamos, inicialmente, os resultados alcançados através dos dados da Vila Planalto. Ao rodarmos os dados desta localidade, o fator sexo foi selecionado nas rodadas em que o *você* estava em contraste, sejam as rodadas *tu/ você* e *cê/ você*.

Observemos, a seguir, uma tabela que apresenta o resultado das rodadas binárias em função do sexo, com percentuais e pesos relativos.

Tabela 28 – Efeito do fator sexo em rodadas binárias dos dados da Vila Planalto

Sexo	<i>Tu</i> (versus <i>você</i>)	<i>Tu</i> (versus <i>cê</i>)	<i>Cê</i> (versus <i>você</i>)
Feminino	145/ 258 = 56,2% Peso = 0,38	145/ 217 = 66,8% (0,46)*	72/ 185 = 38,9% Peso = 0,41
Masculino	128/ 164 = 78% Peso = 0,68	128/ 208 = 61,5% (0,54)*	80/ 116 = 69% Peso = 0,64
Total	273/ 422 = 64,7%	273/ 425 = 64,2%	152/ 301 = 50,5%

(*Pesos relativos retirados do nível 6 da rodada relevante, nível este que representa a última oportunidade desta variável ser selecionada)

Os testes binários indicaram que o sexo masculino favorece de forma semelhante as variantes *tu* e *cê*, ao passo que desfavorecem a variante *você*. Por seu turno, o sexo feminino tende a desfavorecer, de forma semelhante, o *tu* e o *cê*, ao passo que tende a favorecer o *você*. Vejamos, a seguir, os resultados que obtivemos em rodada eneária.

Tabela 29 – Efeito do fator sexo em percentuais e pesos (dados da Vila Planalto)

Sexo	<i>Você</i>	<i>Cê</i>	<i>Tu</i>
Feminino	113/ 330 = 34% Peso = 0,44	72/ 330 = 22% Peso = 0,30	145/ 330 = 44% Peso = 0,25
Masculino	36/ 244 = 15% Peso = 0,24	80/ 244 = 33% Peso = 0,34	128/ 244 = 52% Peso = 0,41

Os resultados alcançados pelo fator sexo em rodada eneária indicaram que ambos os sexos tendem a usar a variante *cê* de forma neutra, com pesos de 0,30 e 0,34 sendo o peso de referência 0,33. Já os comportamentos das variantes *você* e *tu* ficaram inversamente favorecidos pelos sexos; por um lado, o sexo feminino favorecendo o *você*, com o peso de 0,44 e, por outro, o sexo masculino favorecendo o *tu*, com o peso de 0,41.

Os resultados a que chegamos também indicaram que, se comparados aos resultados de Lucca, o sexo feminino vem diminuindo as diferenças entre o sexo masculino quanto ao uso da variante *tu*. Mas esse resultado também pode decorrer devido a diferenças quanto à coleta de dados. Embora nossos resultados tenham

alcançado os percentuais de 44% de uso de *tu* pelo sexo feminino, ficando bem acima dos percentuais alcançados por Dias, nossos pesos ficaram desfavorecidos por este sexo de forma semelhante aos resultados de Dias, com pesos um pouco abaixo do peso de referência. Assim, a diminuição entre as diferenças de uso do *tu* por ambos os sexos não foram reveladas em nossa amostra pelos resultados em pesos, mas pelos números de ocorrências e pelos percentuais apresentados. Para uma melhor comparação a esse respeito, decidimos realizar uma rodada binária de *tu* versus *você* e *cê*, e os resultados estão na tabela a seguir.

Tabela 30 – Fator sexo em função das localidades, baseados em rodadas binárias

Sexo/resultados por localidade	Tu (<i>versus</i> você e <i>cê</i>)
Todos os dados	
Feminino	0,36
Masculino	0,67
Dados só da Vila	
Feminino	0,41
Masculino	0,61
Dados só de fora da Vila	
Feminino	0,35
Masculino	0,78

A partir dos resultados que obtivemos na rodada binária realizada aos moldes das análises de Lucca (2005) e Dias (2007), do *tu* versus as variantes *você* e *cê* juntas, fica novamente evidenciado que o *tu* é relativamente desfavorecido pelas falas femininas da Vila Planalto, mas esse desfavorecimento diminuiu em relação aos resultados obtidos por Lucca em 2005. Por outro lado, nossos resultados são semelhantes aos resultados encontrados por Dias (2007) segundo a comparação entre os pesos (0,41 e 0,40, respectivamente). É importante notar, porém, que esse desfavorecimento ocorre também em função do forte favorecimento do *tu* por parte das falas masculinas, assim, é necessário olhar os números também em função das relações que estão sendo medidas.

No caminho inverso, os pesos das meninas de fora da Vila indicaram um acentuado desfavorecimento da ocorrência de *tu*. Tais resultados demonstram a

necessidade de mais investigações a respeito da influência das localidades e do tipo de coleta para a investigação sobre a ocorrência dos pronomes em estudo.

O fato de o sexo masculino favorecer o uso de *tu* já era esperado e foi confirmado. O fato de a variante *você* ter sido favorecida pelo sexo feminino talvez se encaixe na explicação de que as mulheres tendem a preferir as estruturas reconhecidamente mais formais da língua. É importante, todavia, lembrar que o *corpus* do presente estudo foi, na medida do possível, um *corpus* de fala casual, em que se evitou os contextos formais de interação, a exemplo de outros estudos da mesma natureza.

5.4.4 Tipo de fala

Medimos esse fator por acreditar ser possível que o tipo de fala, se original ou reportada, pudesse interferir no uso dos pronomes. A fala reportada constitui uma parte pequena dos nossos dados, ocorrendo em apenas 6% dos casos. Lucca (2005, p. 83) controlou este fator, mas ele só foi utilizado para fazer cruzamentos de dados, não tendo sido estatisticamente relevante para o programa, ou seja, não houve seleção. Dias (2007, p. 65) também o controlou, mas, novamente, ele não se mostrou estatisticamente significativo. Vejamos a seguir um exemplo dessa ocorrência em nossa amostra: (exemplo de apenas um turno de fala, em que, fora das aspas, representa a fala original da informante e, entre aspas, representa a fala reportada da informante)

Por causa que ela num vai falar, como é que eu vou puxar o assunto? Ia falar: “e aí, Michele, como é que TU acha, TU acha o Rafael ainda uma criança? /é/. Ela vai dizer “não, eu mudei” TU vai ver se ela não vai dizer isso. Ela nunca mais que vai dizer pra mim “não, ele é gente boa, eu pensava que ele era uma criança” {init.} “e aê Michele, tá gostando do criança?” TU vai ver, eu vou gravar pra VOCÊ no meu celular. “e aê Michele, tá gostando da criança? Gente boa, né, véi?” E ontem mesmo ela falou de novo isso.

Ao realizarmos as rodadas binárias, com todos os dados e depois somente com os dados da Vila, esse fator foi selecionado nas rodadas, mas apenas na relação *tu/cê*. Esses resultados se coadunam com os de Lucca e Dias, no sentido de só haver tido seleção em uma relação que as pesquisas de 2005 e 2007 não investigaram (a relação investigada em 2005 e 2007 foi somente a *tu/você*, e esta não foi selecionada em nenhuma das três pesquisas). Os resultados com todos os dados e somente da Vila ficaram equivalentes em termos de tendências. Vejamos, a seguir, uma tabela sobre o tipo de relato, mas com os dados somente da Vila.

Tabela 31 – Variação *você/cê/tu* em relação ao tipo de fala (dados da Vila Planalto)

Tipo de fala	Você	Cê	Tu
Original	134/ 522 = 25,7%	134/ 522= 25,7%	254/ 522 = 48,7%
Reportada	12/ 49 = 24,5%	18/ 49 = 36,7%	19/ 49 = 38,9%
Total	146/ 571 = 25,6%	152/ 571 = 26,6%	273/ 571 = 47,8%

Os resultados da fala original ficaram próximos à média. Significaria dizer que esse tipo de relato não influencia a variação em estudo. Em relação à fala reportada, os resultados apontam para a variante *cê* como a que mais favorece esse tipo de relato, pois sua ocorrência é de 36,7%, ficando 10 pontos percentuais acima da média.

Como já mencionado, a relação *tu/ cê* foi a única a ser selecionada pelo programa e o peso atribuído ao *tu*, nessa relação, foi de 0,22 para a fala reportada, ou seja, a fala reportada desfavorece o uso de *tu*, fato que comprova o favorecimento de *cê* nesta relação. Vejamos, a seguir, o resultado do tipo de fala em rodada eneária:

Tabela 32 – A variação *você/cê/tu* em função do tipo de fala, a partir de rodada eneária

Tipo de fala	Você	Cê	Tu
Original	0,28	0,24	0,47
Reportada	0,36	0,42	0,21

Através dos resultados da rodada eneária, pode-se comprovar o favorecimento da variante *cê* pela fala reportada. Pode-se inferir também que a variante *tu* é

desfavorecida por esse tipo de relato. Já a variante *você* fica neutra ou levemente favorecida pela fala relatada, mas essa diferença não foi considerada estatisticamente significativa.

Ao compararmos os resultados da rodada eneária realizada com os dados da Vila e a rodada eneária com todos os informantes, não se pôde perceber alterações nos números alcançados, o que significa que as diferentes localidades não alteram os resultados para esse fator.

5.5 Efeitos de natureza interna

Passemos, a seguir, para a apreciação dos fatores que influenciam internamente ou linguisticamente a variação em análise. Para a apreciação desses resultados, não analisaremos as duas localidades em separado por acreditar que as tendências linguísticas podem ser apreendidas de qualquer amostra, independentemente da localidade. Assim, daqui para frente usaremos toda a amostra, com dados da Vila Planalto e de fora da Vila em conjunto.

5.5.1 Função sintática

Estudamos o fator função sintática das variantes a fim de perceber as características gramaticais das variantes em estudo. Segundo Kato (*apud* Andrade, 2004, p. 52), somente os pronomes fortes (proeminentes na sentença) podem ocorrer em qualquer posição nas orações ou, ainda, ocupar posições fortes. Temos que a variante *cê* não pode assumir todas as posições do *você* (como vir sozinha em uma oração ou ocupar posição de objeto – se bem que em algumas variedades, como a mineira, tal variante, mesmo em pequena escala, já ocupa posição de objeto cf. Gonçalves, 2008). Sendo fraca e sintaticamente restrita, o *cê* precisa sempre de um hospedeiro para surgir na oração, sendo o seu uso mais natural e comum a posição de sujeito. Essa variante tem sido caracterizada como em processo de cliticização.

Andrade (2004) confirmou em sua análise a restrição sintática da variante *cê* (cf. tabela 11 do presente estudo, no item 4.3.5). Lucca (2005, p. 78) controlou a função sintática no sintagma nominal e esse fator não foi selecionado pelo programa. Seus resultados apontaram: 72% de *tu* em função de sujeito e 68% de *tu* em função de complemento verbal. Dias também estudou a função sintática e esse fator também foi

descartado pelo programa. É importante lembrar que somente a pesquisa de 2004 é comparável à presente pesquisa em relação à função sintática, pois as pesquisas de 2005 e 2007 analisaram as variantes *cê* e *você* como sendo uma só variante em oposição ao *tu*.

Passemos a seguir à apreciação dos resultados encontrados para esse fator no presente estudo.

Tabela 33 – Variação *você/ cê/ tu* em relação à função sintática, com todos os dados da amostra

	Você	Cê	Tu
Sujeito	285/ 779 = 36,6%	224/ 779 = 28,8%	270/ 779 = 34,7%
Sem verbo ²²	33/ 56 = 58,9%	5/ 56 = 8,9%	18/ 56 = 32,1%
Totais	318/ 835 = 38,1%	229/ 835 = 27,4%	288/ 835 = 34,5%

Cerca de 93,3% dos dados (equivalente aos resultados de Andrade e de Lucca) são dos pronomes em função de sujeito.

O controle da função sintática, originalmente, também controlou os dados de predicativo e os dados com função de objeto direto e indireto. Como não houve dado da variante *cê* nem como predicativo, nem em posição de objeto, tivemos de retirar esses casos da presente análise para poder rodar os dados no programa estatístico.

É importante mencionar, porém, que a variante *você* foi fortemente favorecida em função de objeto direto, com o percentual de 71,4% (24 pontos percentuais acima da média²³) e também pelo indireto, com 66,7%, enquanto o *tu* foi favorecido pelo predicativo, com o percentual de 52% (17 pontos percentuais acima da média). Vejamos esses resultados na tabela a seguir.

²² É importante frisar que as orações sem verbos foram assim denominadas, mas que algumas delas poderiam ter verbo e este não pôde ser ouvido quando da transcrição, em outras realmente não houve a presença do verbo.

²³ Sendo a média estabelecida pelos números em valores totais.

Tabela 34 – Variação *você/ cê/ tu* em relação à função sintática, resultados em percentuais, com todos os dados da amostra

Função sintática	Você	Cê	Tu
Sujeito	298/ 827 = 36%	236/ 827 = 28,5%	293/ 827 = 35,4%
Sem verbo	34/ 58 = 58,6%	5/ 58 = 8,6%	19/ 58 = 32,8%
Predicativo	12/ 25 = 48%	0/ 25 = 0 %	13/ 25 = 52%
Objeto indireto	24/ 36 = 66,7%	0/ 36 = 0 %	12/ 36 = 33,3%
Objeto direto	5/ 7 = 71,4%	0/ 7 = 0 %	2/ 7 = 28,6%
Totais	373/ 953 = 39,1%	241/ 953 = 25,3%	339/ 953 = 35,6%

As três variantes ocorreram livremente na função de sujeito, com seus percentuais próximos às respectivas médias, como podemos observar na tabela anteriormente colocada. Inferimos que, na posição de sujeito, as variantes estão concorrendo de forma neutra, pois as diferenças numéricas internas são muito pequenas.

Quando analisamos as sentenças sem verbo, porém, percebemos que há diferenças significativas acerca das variantes *você e cê*, de quase 20 pontos percentuais em relação às suas médias. O *você* foi favorecido, com o percentual de 58,6%, enquanto o *cê* foi sendo desfavorecido por este fator, com o percentual de 8,6%. As frases sem verbo correspondem a hesitações ou a reformulações características da oralidade. Coletamos esse tipo de dado por ser este mais um contexto de ocorrência variável dos pronomes e, portanto, esses dados são importantes do ponto de vista da quantidade real das ocorrências de cada variante em estudo²⁴. Vejamos alguns exemplos a seguir:

“D: e VOCÊ?”

“M: CÊ compra massa pra pastel, aí, aí VOCÊ...” (2º dado)

“D: CÊ... VOCÊ {init.} no computador?” (1º dado da oração)

“D: CÊ...”

“F: é bem assim, CÊ colo..

“F: não CÊ,”

“F: CÊ...”

“F: que? /eu!/ CÊ já {init.}”

²⁴ Apesar de, no presente estudo, não estudarmos esses dados do ponto de vista discursivo, algumas dessas ocorrências podem sugerir um controle do próprio informante em relação à sua fala.

“U: como TU mesmo...”

Ao realizarmos as rodadas binárias, percebemos que somente quando a variante *cê* estava em contraste é que houve a seleção do fator função sintática. Na rodada *tu/cê*, os pesos atribuídos ao *tu* foram: 0,47 para a função de sujeito e 0,92 para sentenças sem verbo. Na rodada *cê/ você*, os pesos atribuídos ao *cê* foram: 0,53 para a função de sujeito e 0,14 para sentenças sem verbo. Para visualização dos resultados atribuídos às três variantes, passemos a observar a rodada eneária a seguir.

Tabela 35 – Resultados em pesos da variação *você/ cê/ tu* quanto à função sintática, com todos os dados da amostra

Função sintática	Você	Cê	Tu
Sujeito	0,19	0,57	0,23
Sem verbo	0,46	0,15	0,37

Antes de tudo, é importante ressaltar que a variante *cê* foi a responsável pela seleção da variável função sintática (considerada como significativa pelo programa estatístico na presente pesquisa) e isso se deu, principalmente, pela restrição sintática que esta variante apresenta e isso, por si só já é um importante resultado.

A partir dos resultados apresentados na tabela anterior, podemos inferir que o *cê* é fortemente favorecido pela função de sujeito, bem como é fortemente desfavorecido pelas orações sem verbo. Aqui caberia a explicação dada por Paredes (2003) acerca da tonicidade dos pronomes. É como se o *cê* não fosse forte o suficiente para aparecer sozinho em uma oração e o *você*, ao contrário, por ser o pronome mais tônico, e de maior saliência fônica, ser fortemente favorecido nesse sentido. O *tu* fica em posição intermediária entre as demais variantes, pois, apesar de ser tônico, é monossilábico. Por fim, inferimos que as variantes *você* e *tu* ficaram abaixo do valor de referência na posição de sujeito pelo alto favorecimento da variante *cê* nesta posição.

5.5.2 Tipo de referência, presença de vocativo e entonação

a) Tipo de referência:

A hipótese para o controle desse fator foi a possibilidade de a variante *você* ser favorecida quando a referência fosse genérica. Dias (2007, p. 89-90) estudou esse fator e, em suas palavras, “pode-se se afirmar que há, de maneira geral, uma preferência do falante em usar as variantes *cê* e *você* em casos de referência genérica”. Como a pesquisadora não separou em todas as rodadas as variantes *cê* e *você*, seria interessante se se pudesse perceber semelhanças ou diferenças entre ambas as variantes em relação à referência genérica. Com relação à referência específica, esperamos que as variantes mantenham-se neutras, aproximadas de suas médias. A seguir, a tabela apresenta os resultados gerais.

Tabela 36 – Variação *você/cê/tu* em relação ao tipo de referência, valores em percentuais com todos os dados da amostra

Tipo de referência	Você	Cê	Tu
Específica	266/ 754 = 35,3%	205/ 754= 27,2%	283/ 754 = 37,5%
Genérica	52/ 81 = 64,2%	24/ 81 = 29,6%	5/ 81 = 6,2%
Totais	318/ 835 = 38,1%	229/ 835 = 27,4%	288/ 835 = 34,5%

Os resultados a que chegamos foram bastante contundentes em desfavorecer o *tu* em referência genérica. Na mesma linha de observação, tais resultados também se mostraram consistentes em favorecer o *você* na mesma referência. Infere-se dos resultados, porém, que a variante *cê* favorece a referência genérica, mas de maneira mais branda se comparada ao *você*.

Na rodada binária, esse fator só foi selecionado na relação *tu/você*, ficando o *tu* com os pesos de: 0,54 para a referência específica e 0,14 para referência genérica. Esse resultado comprova o acentuado desfavorecimento do *tu* pela referência genérica. Vejamos, a seguir, os resultados da rodada eneária que refletem a seleção e a não seleção nas rodadas binárias:

Tabela 37 – Variação *você/cê/tu* em relação ao tipo de referência, valores em pesos (todos os dados)

Tipo de referência	Você	Cê	Tu
Específica	0,22	0,28	0,49
Genérica	0,45	0,35	0,20

Os resultados da tabela anterior corroboram, relativamente, as hipóteses testadas para esta variável. Por um lado, o *tu* é desfavorecido pela referência genérica enquanto o *você* é favorecido por ela. Por outro lado, porém, há o favorecimento da variante *tu* quando a referência é específica e o desfavorecimento do *você* por esta referência. A variante *cê* permanece em níveis aproximados à média, ficando levemente desfavorecida pela referência específica, ao passo que é levemente favorecida pela referência genérica, mas sem significância estatística para os resultados desta variante, como evidenciam as rodadas binárias.

b) Presença ou ausência de vocativo:

Medimos esse fator por acreditar que a presença ou não de vocativo pudesse interferir na ocorrência dos pronomes. Nossos resultados apontaram que a presença de vocativo ocorre em apenas 12,5% dos dados. Esse grupo de fatores não foi selecionado nas rodadas binárias com todos os informantes e as demais pesquisas aqui revisadas não controlaram este fator. Como exemplo dessa ocorrência, temos em nossa amostra:

“B: véi, não, só podia de ser TU, Pablo. Área cinquenta e um existe!”

Vejam, a seguir, uma tabela que apresenta resultados em percentuais, baseados em rodada geral, bem como os pesos relativos, baseados em rodada enéaria.

Tabela 38 – Variantes *você/cê/tu* em relação à presença ou ausência de vocativo (todos os dados)

Vocativo	Você	Cê	Tu
Sem vocativo	292/731 = 39,9% Peso = 0,36	209/731 = 28,6% Peso = 0,35	230/731 = Peso = 0,27
Com vocativo	26/104 = 25% Peso = 0,30	20/104 = 19,2% Peso = 0,30	58/104 = 55,8% Peso = 0,39
Totais	318/835 = 38,1%	229/835 = 27,4%	288/835 = 34,5%

Na análise dos resultados, a presença do vocativo demonstrou uma tendência a favorecer a variante *tu* e a ausência do vocativo demonstrou uma tendência a

desfavorecer a ocorrência desta variante, com um *range* de 12 (relativamente significativa para o valor de referência, que é 0,33). Porém, o fato desse fator não ter sido selecionado pelo programa nas rodadas binárias e, por outro lado, o fato de as demais variantes terem ficado muito próximas aos seus valores de referência sugerem que a presença ou a ausência de vocativo não influencia as ocorrências das variantes de forma significativa.

c) Entonação: orações interrogativas *versus* não interrogativas

A hipótese para medir esse fator surgiu a partir da percepção de que, em orações mais exaltadas, com entonação mais emotiva ou “cantada” (tom alto ou ascendente), a ocorrência de *tu* parece ser favorecida.

Andrade (2004) e Lucca (2005) estudaram o tipo de entonação. Andrade (2004, p. 58-59) mediu as interrogativas (que foram separadas em três tipos) *versus* declarativas e encontrou os seguintes resultados em função da variante *cê*: peso relativo de 0,62 para as interrogativas (que a autora denominou de ‘que que’), ao passo que o peso das declarativas foi de 0,51. Lucca (2005, p. 108-109) mediu a entonação entre as orações exclamativas, interrogativas e declarativas. A autora chegou aos seguintes resultados em função da variante *tu*: peso de 0,87 para exclamativas; 0,54 para interrogativas e 0,43 para declarativas.

Tentamos, inicialmente, separar as orações em declarativas, exclamativas e interrogativas. Acreditando-se que as exclamativas junto às interrogativas seriam as que mais favoreceriam a ocorrência de *tu*, ao passo que as declarativas desfavoreceriam a sua ocorrência. Porém, como a audição das gravações não permitiu clareza na definição entre declarativas e exclamativas, mas permitiu clareza quando se tratava de frase interrogativa, decidimos deixar esse fator separado em apenas duas formas: interrogativas por um lado e não interrogativas (ou afirmativas e exclamativas) por outro. Vejamos os resultados na tabela a seguir:

Tabela 39 – Variação *você/ cê/ tu/* em relação à entonação, resultados em percentuais para toda a amostra

Entonação	Você	Cê	Tu
Interrogativas	41/229 = 17,9%	81/229 = 35,4%	107/229 = 46,7%
Não interrogativas	277/605 = 45,8%	147/605 = 24,3%	181/605 = 29,9%
Total	318/834 = 38,1%	228/834 = 27,3%	288/834 = 34,5%

Os resultados gerais indicaram que as orações interrogativas favoreceram relativamente as variantes *cê* e *tu* de forma semelhante, ficando ambas com os percentuais acima de suas médias. Por sua vez, a variante *você* foi favorecida pelas orações não interrogativas. Ao rodar as rodadas binárias, esse fator foi selecionado somente quando a variante *você* estava em contraste com as demais, nas relações *tu/você* e *cê/você*. Na relação *tu/você* os pesos atribuídos ao *tu* foram: 0,73 em orações interrogativas e 0,42 em orações não interrogativas. Na relação *cê/você* os pesos atribuídos ao *cê* foram: 0,72 para as orações interrogativas e 0,43 para as orações não interrogativas. Os pesos foram praticamente iguais ao favorecer *cê* e *tu* e a desfavorecer o *você* nas orações interrogativas. Passemos, a seguir, à apreciação dos pesos baseados em rodada eneária.

Tabela 40 – Variação *você/ cê/ tu/* em relação à entonação, resultados em pesos

Entonação	Você	Cê	Tu
Interrogativas	0,22	0,39	0,38
Não interrogativas	0,46	0,26	0,27

A tabela acima confirma a sugestão anterior de as variantes *cê* e *tu* terem seus efeitos equivalentes, sendo favorecidas pela entonação mais ascendente presente nas orações interrogativas, ao passo que a variante *você* é desfavorecida por esse tipo de entonação.

Os resultados da presente pesquisa corroboram os resultados de Andrade e Lucca. Além de corroborados, os resultados de Lucca também nos incitam a buscar meios de analisar mais detalhadamente a entonação em pesquisas futuras, realizando a separação das exclamativas e declarativas, uma vez que a diferença em seus resultados

entre as orações declarativas, interrogativas e exclamativas se mostrou bastante significativa.

5.5.3 Paralelismo Linguístico

O paralelismo linguístico consiste na tendência de “marcas levarem a marcas e zeros levarem a zeros”, segundo Poplack e Naro (*apud* Scherre & Naro, 1993), ou, de forma mais abrangente, o paralelismo “consiste na tendência de formas gramaticais semelhantes ocorrerem juntas”, segundo Schiffrin (*apud* Scherre & Naro, 1993). Do ponto de vista metodológico, o paralelismo é a tendência de “repetição das variantes de uma mesma variável dependente” (cf. Scherre, 1998, p. 30). Em nossos dados, significa a tendência da ocorrência de um *tu* na fala, levar a outro *tu* no mesmo turno de fala, ao invés de uma continuação usando *cê* ou *você*, por exemplo. A seguir, os quatro primeiros trechos são exemplos de paralelismo em nossa amostra e, na sequência, outros quatro trechos são de não paralelismo:

P: parece, mas não é, cara, se TU for ficar lá uma semana, *véi*, TU vai querer voltar no primeiro dia, porque lá, *véi*, as brincadeiras é tudo diferente das daqui, *véi*, não tem polícia e ladrão, não tem pique esconde, não tem nada, *véi*, é muito *paia*.

F: já, é muito massa a guitarra lá, aí tem uma hora que (...), aí CÊ tem que fazer assim, CÊ tem que fazer assim depois ó. E aquela...a do {init.}

F: eles ensinam VOCÊ a driblar, a fazer finta, ensinam VOCÊ a dar o chapeuzinho {init.}

L: que dia que a SENHORA vai... Que dia que a SENHORA vai fazer com a gente isso aí?

B: ô, mas, ô Pablo, TU é burro demais, não sei nem como CÊ conseguiu chegar sexta-feira, doido.

L: oh TU, {obs.: repensando a fala} a SENHORA viu, eu guardei tudo lá

L: {init.} a SENHORA é alta, CÊ dá conta?

F: mas CÊ não é melhor que eu em matemática. VOCÊ não é melhor que eu em matemática.

Nos primeiros quatro trechos anteriores, tivemos exemplos do paralelismo dos itens: *tu, cê, você e a senhora*, ao passo que nos quatro últimos trechos, tivemos exemplos de não paralelismo. Andrade e Lucca, entre outros pesquisadores, também estudaram esse fator.

No estudo de Andrade (2004), em que a autora estudou a variação entre as formas *cê, ocê e você*, na rodada binária em que o *cê* está em contraste com *você*, a autora chegou aos seguintes resultados:

Tabela 41 – Variação entre *cê/você* quanto ao paralelismo, resultados de 2004

Paralelismo	Cê (versus você)	
Precedido por cê	84/ 116 = 72%	Peso = 0,88
Precedido por você	57/ 390 = 15%	Peso = 0,32
Primeiro da série	65/ 201 = 32%	Peso = 0,58
Total	206/ 707 = 29%	

(Andrade, 2004, p. 128 – com adaptações)

As análises de Andrade a respeito do paralelismo identificaram a tendência da forma *cê* levar a outra forma *cê*, desfavorecendo a ocorrência do *você* com um *range* (uma diferença de pontos em peso) de 56.

No trabalho de Lucca (2005, p. 106), a autora afirma: "(...) a posição que mais favorece a ocorrência de *tu* é o primeiro item da série (...) ao contrário do que apontam várias pesquisas variacionistas que lidam com o paralelismo lingüístico, no sentido de o

primeiro item da série tender a um efeito intermediário sobre a escolha das variantes". É fundamental explicar, no entanto, que no trabalho desenvolvido por Lucca, o foco do paralelismo foi diferente do foco da presente pesquisa. Enquanto no trabalho de 2005 as observações eram também se um item de um turno de fala influenciava um item no turno de fala vindouro, o presente estudo busca saber somente se, dentro do mesmo turno de fala, o uso de um determinado pronome influencia na escolha do próximo pronome.

A expectativa em controlar esse fator foi a de que o paralelismo influencia na ocorrência das variantes em análise. Vejamos, a seguir, os resultados alcançados por este fator na presente pesquisa:

Tabela 42 – Variação *você/cê/ tu/* em relação ao Paralelismo (percentuais)

Paralelismo	Você	Cê	Tu
Isolado na oração	164/481 = 34,1%	141/48 = 29,35	176/481 = 36,6%
1º item da série	57/138 = 41,3%	37/138 = 26,8%	44/138 = 31,9%
Precedido por <i>você</i>	79/97 = 81,4%	13/97 = 13,4%	5/97 = 5,2%
Precedido por <i>cê</i>	13/48 = 27,1%	32/48 = 66,7%	3/48 = 6,2%
Precedido por <i>tu</i>	5/71 = 7%	6/71 = 8,5%	60/71 = 84,5%
Totais	318/835 = 38,1%	229/835 = 27,4%	288/835 = 34,5%

Os resultados da tabela acima confirmam a tendência de ocorrer o paralelismo entre as variantes em estudo. Ao encontro do que se esperava, esse fator foi selecionado pelo programa em todas as rodadas binárias realizadas. Vejamos a tabela a seguir:

Tabela 43 – Resultados da variação em relação ao Paralelismo (em rodadas binárias, com todos os dados)

Paralelismo	Tu (<i>versus</i> você)	Tu (<i>vesus</i> cê)	Cê (<i>versus</i> você)
Isolado na oração	176/340 = 51,8% Peso = 0,55	176/317 = 55,5% Peso = 0,48	141/305 = 46,2% Peso = 0,55
1º item da série	44/10/ = 43,6% Peso = 0,52	44/81 = 54,3% Peso = 0,50	37/94 = 39,4% Peso = 0,49
Precedido por <i>você</i> ²⁵	5/84 = 6% Peso = 0,10	5/18 = 27,8 Peso = 0,23	13/92 = 14,1% Peso = 0,22
Precedido por <i>cê</i>	3/16 = 18,8% Peso = 0,12	3/35 = 8,6% Peso = 0,07	32/45 = 71,1% Peso = 0,74
Precedido por <i>tu</i>	60/65 = 92,3% Peso = 0,88	60/66 = 90,9% Peso = 0,88	6/11 = 54,5% Peso = 0,48
Totais	288/606 = 47,5%	288/517 = 55,7%	229/547 = 41,9%

Observamos na tabela anterior que, na relação *tu/você*, quando precedido por *tu*, o peso para a variante *tu* foi de 0,88 favorecendo sua ocorrência; quando em ocorrências isoladas e primeiro da série, o *tu* ficou com os pesos de 0,55 e 0,52 respectivamente, sua ocorrência, portanto, pode ser considerada neutra; quando precedida pelas variantes *cê* e *você*, o peso de *tu* foi de 0,12 e 0,10 respectivamente, desfavorecendo, portanto, a sua ocorrência.

Na relação *tu/cê*, quando precedido por *tu*, o peso para a variante *tu* novamente foi de 0,88, favorecendo, portanto a sua ocorrência; quando em ocorrências isoladas e primeiro da série, o *tu* ficou com pesos 0,47 e 0,52 respectivamente, sua tendência de ocorrência novamente se mostrou neutra; quando precedida pelas variantes *cê* e *você*, o peso de *tu* foi novamente desfavorecido com pesos de 0,07 e 0,23 respectivamente.

E, por fim, na relação *cê/você*, a variante *cê* quando precedida por ela, fica favorecida com o peso de 0,74; quando ocorre de forma isolada ou sendo primeira de uma série, esta variante fica neutralizada, com os pesos de 0,55 e 0,49 respectivamente, por sua vez, chama a atenção quando a forma *cê* ocorre tendo sido precedida pelo *tu*, pois a variante *cê* fica praticamente neutralizada também, com o peso de 0,48, mas quando precedida de *você* a variante *cê* fica desfavorecida com o peso de 0,22.

²⁵ Há que se perceber que existe um certo alinhamento entre as variantes *você* e *cê*.

Para melhor compreender os efeitos de uma variante em contraste com as demais, observemos, a seguir, a tabela que apresenta resultados em pesos, baseada em rodada eneária:

Tabela 44 – Variação *você/cê/tu* em relação ao Paralelismo (todos os dados, resultados em pesos)

	Você	Cê	Tu
Isolado	0,27	0,33	0,38
Primeiro da série	0,31	0,29	0,39
Precedido por <i>você</i>	0,66	0,20	0,12
Precedido por <i>cê</i>	0,23	0,68	0,07
Precedido por <i>tu</i>	0,09	0,09	0,81

Percebemos ao comparar os resultados das rodadas binárias, na tabela 43, com os resultados da rodada eneária, na tabela acima, que os valores não são diretamente correspondentes, como podemos observar na relação binária *cê/você*, em que, se precedido pelo *tu*, o *cê* demonstrou tendência a ocorrer de forma neutra ou menos restrita que se precedido pelo *você*. Os resultados eneários não apontaram tal tendência.

Os resultados, de maneira geral, corroboraram o princípio apresentado anteriormente sobre a tendência de formas gramaticais ocorrerem juntas. A partir dos resultados da tabela acima, temos que: a) os itens isolados, assim como os primeiros da série, por ficarem em torno do número de referência, 0,33, e apresentarem as tendências das médias de suas variantes em estudo nesta amostra, nos levam a inferir que eles não exercem influência, ou seja, suas tendências são de neutralidade (mas vale dizer que não medimos se surgem por causa de um pronome usado anteriormente, em outro turno de fala); b) as variantes *cê* e *você* permutaram mais entre si, tendo um comportamento parecido entre elas mesmas, como se houvesse um subparalelismo ou uma influência interna entre as duas variáveis (uma espécie de alinhamento). Esse fato poderia ser previsível, uma vez que o *cê* deriva do *você*. Ao observar os números anteriormente colocados percebemos uma semelhança dos comportamentos numéricos das variantes *cê* e *você*, pois suas diferenças internas de peso ficaram em torno de 40 (*range*), ao passo que ambas se distanciam fortemente da variante *tu*, com diferenças de pesos em torno de 60 (*range*). Mesmo discutindo tais tendências, é necessário que haja mais

pesquisas sobre as três variantes para que se entenda melhor a relação de troca que há entre elas. Assim sendo, a variante que demonstrou obedecer mais aos critérios do paralelismo foi o *tu*, com o peso de 0,81, seguida pelas variantes *cê*, com o peso de 0,68 e *você*, com o peso de 0,66.

5.6 Considerações sobre as análises

Analizamos no presente estudo fatores externos e internos da variação entre as formas pronominais *você/cê/tu* na fala de crianças e adolescentes, de 7 a 15 anos, em Brasília. Alguns dos resultados obtidos das análises estatísticas estão compilados em uma única tabela, vejamos a seguir:

Tabela 45²⁶ – Variação *você/cê/tu* em Brasília com agrupamentos das variantes (dados da Vila Planalto – resultados em pesos relativos)

Variáveis independentes				Tipo de agrupamento entre as três variantes
Faixa etária	<i>Você</i>	<i>Cê</i>	<i>Tu</i>	<i>Você/cê/tu</i> praticamente sem distinção
Crianças: 7-11 anos	0,33	0,37	0,30	
Adolescentes: 12-15 anos	0,33	0,30	0,37	
Relação entre os interlocutores	<i>Você</i>	<i>Cê</i>	<i>Tu</i>	Você/cê vs. tu
Simétrica	0,27	0,26	0,47	
Assimétrica	0,38	0,40	0,22	
Sexo	<i>Você</i>	<i>Cê</i>	<i>Tu</i>	Cada variante com seu comportamento
Masculino	0,24	0,35	0,42	
Feminino	0,44	0,31	0,25	
Tipo de fala	<i>Você</i>	<i>Cê</i>	<i>Tu</i>	Você/cê vs. tu
Não reportada	0,28	0,24	0,48	
Reportada	0,36	0,42	0,21	
Função sintática	<i>Você</i>	<i>Cê</i>	<i>Tu</i>	Você/tu vs. cê (caso único)
Sujeito	0,19	0,55	0,26	
Estruturas sem verbo	0,48	0,17	0,35	
Tipo de referência	<i>Você</i>	<i>Cê</i>	<i>Tu</i>	Você/cê vs. tu
Específica	0,22	0,28	0,50	
Genérica	0,45	0,35	0,20	
Tipo de entonação	<i>Você</i>	<i>Cê</i>	<i>Tu</i>	Cê/tu vs. você (caso único)
Interrogativa	0,21	0,42	0,37	
Não interrogativa	0,48	0,24	0,28	
Vocativo	<i>Você</i>	<i>Cê</i>	<i>Tu</i>	
Presença de vocativo	0,26	0,34	0,40	Cada variante com seu comportamento
Ausência de vocativo	0,42	0,31	0,27	
Paralelismo	<i>Você</i>	<i>Cê</i>	<i>Tu</i>	Você/cê vs. tu
Pronome isolado	0,29	0,33	0,39	
Primeiro da série	0,31	0,26	0,43	
Precedido de <i>você</i>	0,61	0,27	0,12	
Precedido de <i>cê</i>	0,27	0,64	0,09	
Precedido de <i>tu</i>	0,10	0,10	0,81	
Origens dos pais	<i>Você</i>	<i>Cê</i>	<i>Tu</i>	Agrupamentos diversificados
Do Nordeste mesmo estado (8 fal.)	0,20	0,23	0,57	
De Brasília e do Nordeste (7 fal.)	0,25	0,15	0,60	
De Brasília (8 falantes)	0,22	0,33	0,45	
Do Nordeste estados diferentes (3 fal.)	0,33	0,44	0,23	
De Minas (1 falante)	0,15	0,70	0,15	
De Minas e Nordeste (2 falantes)	0,30	0,35	0,35	
De Brasília e Mato Grosso (1 fal.)	0,82	0,11	0,07	
Percentagem global	26%	26%	48%	

²⁶ Pode haver pequenas diferenças numéricas nas tabelas em função das rodadas utilizadas para gerá-las. Tais diferenças são normais nas diferentes rodadas e indicam as mesmas tendências.

Podemos observar, a partir da tabela anterior, que as variantes em análise formam alguns agrupamentos entre si, como indicado na quinta coluna. Assim, ora uma variante se assemelha a uma de suas opositoras, ora se assemelha a outra. É importante ressaltar, que essa influência interna dos comportamentos dos pronomes é mais consistente entre as variantes *você* e *cê* (em oposição ao *tu*), como podemos notar na tabela 45, em que há uma espécie de alinhamento (ou subparalelismo) dessas variantes nos fatores: relação entre interlocutores; tipo de fala; tipo de referência e paralelismo linguístico. O comportamento do *tu*, por sua vez, se aproxima (ou se alinha) ao comportamento do *você* no fator função sintática, ao passo que se aproxima (ou se alinha) ao comportamento do *cê* no fator tipo de entonação. Acontece também, como podemos notar nos resultados em relação ao fator sexo, o não agrupamento entre as variantes. Inferimos, assim, que as variáveis independentes, principalmente as de fatores externos, e as influências mútuas regem a variação em estudo.

Gostaríamos de comparar, de forma geral, os resultados obtidos da variante *tu* (e, quando for o caso, do *cê*) em Brasília, com resultados de algumas pesquisas aqui revisadas. Na Região Sudeste, a variação entre os pronomes se mostrou bastante heterogênea, pois em cada estado as variantes se comportaram de formas diferentes. Por exemplo, ao observar os resultados de Brasília e compará-los com os do Rio de Janeiro, reparamos que o comportamento do *tu* na capital é relativamente semelhante ao comportamento desse pronome na fala carioca. Mas, em São Paulo (cf. Modesto, 2007) e em Minas Gerais (Mota, 2008), a variante *tu* apresentou uma ocorrência menor que em Brasília. Nesses estados, os resultados também apontaram diferenças em relação ao comportamento da variante *tu*, em função de algumas variáveis sociais, como sexo e faixa etária, por exemplo, que configuram mais um indício de distância do comportamento linguístico do brasiliense quanto ao uso de *tu*.

A variante *cê*, por seu turno, apesar de ter sido pouco investigada pelos autores aqui revisados, demonstrou ser fortemente favorecida na relação *você/cê* no Rio de Janeiro, superando a ocorrência de sua forma plena (cf. Paredes, 2003). Essa forma também é fortemente favorecida em Minas Gerais, superando os números dessa variante encontrados em Brasília (cf. Golçalves, 2008, p. 188 e Andrade, 2004, p. 50). Nossos resultados corroboram esta hipótese.

Na Região Sul (cf. Loregian-Penkal, 2004), por sua vez, os estados do Rio Grande do Sul e Santa Catarina se agrupam em oposição ao Paraná em termos de

comportamento pronominal. Por um lado, há a alta ocorrência do pronome *tu*, e, por outro, a não ocorrência deste pronome. Nos estados em que há a ocorrência do pronome *tu*, esta é bastante elevada. E mesmo havendo vários fatores de natureza externa e interna incidindo sobre a variação, comprovou-se que o *tu* ali é nativo, tão típico que chega a ser característica identitária (cf. Loregian-Penkall, 2004), apesar de que, no Rio Grande do Sul, o *tu* é ainda mais recorrente que em Santa Catarina. Por outro lado, o *tu*, em Brasília, embora se assemelhe a alguns resultados do Sul (exceto Paraná) em função de alguns dos fatores investigados, como a faixa etária, por exemplo, noutros fatores as duas regiões se distanciam bastante, como no sexo, por exemplo. Em Brasília, essa variante é favorecida pelo sexo masculino enquanto no Sul é favorecida pelo sexo feminino. A ocorrência do *tu* em Brasília é significativamente menor que no Sul, sendo, dessa forma, o *tu* mais restrito na fala de Brasília. Inferimos, assim, que o *tu* em Brasília é distinto do *tu* da Região Sul. Não obtivemos resultados comparáveis da variante *cê* nesta região.

A Região Nordeste também não apresentou resultados quanto ao uso da variante *cê*. A variante *tu*, por sua vez, apareceu em todas as localidades desta região aqui revisadas, mas seus resultados são também bastante diferentes em cada um dos estados analisados. O lugar em que a variante *tu* apareceu de forma mais recorrente foi em João Pessoa e, por outro lado, o lugar em que o *tu* apareceu de forma menos recorrente foi na Bahia. Em João Pessoa (cf. Bezerra, 1994) percebemos algumas semelhanças de resultados, em alguns dos fatores sociais, com os resultados da Região Sul: além do alto índice de *tu* em João Pessoa, também há a liderança de uso do *tu* por parte das meninas. Vale ressaltar, porém, que não foi possível constatar uma nativização do pronome *tu* no Nordeste, nem mesmo onde seu uso é mais recorrente. Os resultados de Brasília também não são diretamente comparáveis a essa região como um bloco, pois, como já mencionado, cada estado é muito diferente em relação aos demais. E, observando os estados separados, não houve um resultado de *tu* no Nordeste que fosse aproximado aos resultados encontrados em Brasília na presente pesquisa. Mas vale ressaltar que os métodos das pesquisas realizadas nestas localidades foram, também, bastante distintos.

Na Região Norte, a ocorrência de *tu*, segundo os trabalhos aqui revisados, é bastante elevada (cf. Soares & Leal, 1993 e Scherre et al., 2009), sendo os resultados dessa região ainda mais próximos aos resultados da Região Sul que o Nordeste. Pelos trabalhos revisados e pela comunicação pessoal com o linguista Germano Martins, inferimos que o *tu*, na Região Norte, seja um *tu* nativo. Tal linguista tem percebido,

porém (e apresentará resultados em sua dissertação de mestrado – em breve), que há variação entre as formas *tu*, *você* e *senhor* em sua comunidade e que, talvez, essa variação caminhe na direção de uma mudança linguística nessa comunidade (Tefê, Amazonas), uma vez que fatores de ordem social interferiram na comunidade, a ponto de haver um perceptível aumento da ocorrência do pronome *você* nesta região. O fato é que os resultados da variante *tu*, na Região Norte, também se distanciam dos resultados encontrados em Brasília. Resta dizer que Martins (2010) encontrou em seus dados apenas 8 casos do pronome *cê*, que correspondeu a 0,07% de seus dados.

Podemos inferir dos resultados aqui apresentados que o *tu* em Brasília é um *tu* ainda não nativo, como o é na Região Sul e Norte. Apesar de nossos resultados se assemelharem relativamente aos resultados encontrados no Rio de Janeiro (amostra Paredes 96 – cf. Paredes, 2003), acreditamos que o *tu* tenha entrado na fala de Brasília, como já tivemos oportunidade de afirmar, pelo contingente nordestino que aqui se estabeleceu, pois não tivemos um forte contingente de fluminenses em Brasília, como tivemos de nordestinos. Nessa linha de raciocínio e, ainda, quanto aos resultados encontrados em Brasília não serem semelhantes aos resultados encontrados no Nordeste, acreditamos que: a) as pesquisas não foram realizadas da mesma forma, fato que interfere nos resultados alcançados, como já tivemos oportunidade de explicar; b) há em Brasília um estágio de focalização dialetal, que torna o jeito de falar brasiliense distinto dos demais.

6. CONSIDERAÇÕES FINAIS

Os resultados apresentados nos capítulos anteriores serviram como um retrato de como a variação entre os pronomes de segunda pessoa se comporta na oralidade de Brasília, em determinadas localidades (na Vila Planalto e fora da Vila Planalto) e em determinado tempo. Tais resultados servem como base para as considerações que faremos a seguir.

Loregian-Penkall (2004, p. 231) quando afirma “existem no PB duas formas para tratar informalmente o interlocutor: os pronomes de segunda pessoa *tu* e *você*” assume uma postura bastante consistente sobre quais pronomes ocorrem no Brasil. Assim como Loregian-Penkall, outros linguistas têm reintroduzido o *tu* como um pronome legítimo do escopo linguístico do brasileiro. Os gramáticos, por seu turno, nunca o retiraram da tradição gramatical.

Assumimos assim que o *tu* e o *você* são, de fato, os pronomes em uso mais generalizados no Brasil. Apesar de não serem os únicos, nem somente informais, são os que detêm, até o momento, legitimidade pelos estudiosos da língua, sejam gramáticos ou linguistas.

Em nossa pesquisa, o pronome *você* se destacou nas falas femininas, em referência genérica, em construções que não possuem verbos, em relações assimétricas e em enunciados não interrogativos. O pronome *cê* demonstrou não possuir exatamente as mesmas características do *você*, mesmo sendo, a rigor, uma redução deste. Este pronome se destacou em falas reportadas, em posição de sujeito, em orações interrogativas e em relações assimétricas. O *tu*, por sua vez, se destacou em falas masculinas, em falas originais, em interrogativas, em referência específica, em relações simétricas e, relativamente, também fora da função de sujeito. Em muitos momentos as variantes se agrupam em termos de tendências, pois apresentam tendências semelhantes, como pôde ser observado na tabela 45. O efeito bastante uniforme, em relação aos pronomes *você*, *cê* e *tu*, é o do paralelismo linguístico: já vimos que esta variável tem efeito semelhante em todas as pesquisas em que foi abordada, independentemente do fenômeno analisado.

O *senhor(a)*, à parte das análises estatísticas, demonstrou que ainda sobrevive nas falas casuais, mas com uma ocorrência muito baixa. Somente um único informante, dos 43 analisados, usou este pronome quando das gravações. Tal informante

demonstrou, ao longo do registro de sua fala, uma preocupação quanto ao uso dos pronomes. Houve outra informante que também demonstrou preocupação quanto ao uso dos pronomes, mas, ao invés de lançar mão do *senhor(a)*, preferiu o uso de nulos e do pronome *cê*. Assim, infere-se que o uso do *senhor(a)* está presente em situações muito específicas de uso de pronomes, ligadas, principalmente, a situações de assimetria social (em relação à idade ou a algum tipo de hierarquia, por exemplo).

Ao retomarmos as hipóteses do presente estudo, verificamos que a faixa etária analisada, que se caracterizou por ser diretamente anterior às faixas etárias estudadas por Lucca (2005) e Dias (2007), já usa de forma significativa a variante *tu* em falas casuais. Ou seja, a variante *tu* já entra, em tenra idade, na variedade brasiliense, e seu uso conta com 34,5% dos dados de pronomes. Nossa análise confirma, pois, esta hipótese.

Nos dados da Vila Planalto, foi possível verificar dados de *tu* em todas as idades estudadas, ocorrendo em 48% dos dados de pronomes. Nos dados de fora da Vila, a ocorrência de *tu* foi bem menor, contando com apenas 8,2% dos dados de pronomes. Ao testar os dados separando-os em duas faixas etárias diferentes (7 a 11 anos e 12 a 15 anos), os dados da Vila Planalto não apresentaram diferenças etárias significativas quanto ao uso do pronome *tu* em relações simétricas. Fora da Vila, porém, os dados apresentaram uma influência das faixas etárias quanto ao uso desta variante, sendo a faixa etária de 12 a 15 anos a mais propícia quanto ao uso de *tu*. Mas, como explicado anteriormente, o informante responsável por esse efeito tem família residente em Ceilândia (lugar propício ao uso da variante *tu* pelo elevado contingente de imigrantes nordestinos). Assim, a questão da idade, embora se revelando estatisticamente significativa, pode estar encobrendo variáveis de outra natureza, ainda não captadas pela análise quantitativa dos dados de fora da Vila Planalto.

Além de confirmar a aquisição em tenra idade da variante *tu* por parte dos brasilienses (já aos cinco anos de idade), ao analisar o uso da variante *tu* em nossos dados, também foi possível confirmar uma tendência, já apontada em Lucca (2005), quanto ao fato de a origem dos pais influenciar a fala de seus filhos. Nossos resultados indicaram que a origem dos pais influencia de forma indubitável o uso das variantes em nossa amostra, uma vez que a fala dos filhos se assemelhou às características linguísticas típicas das origens analisadas. O maior uso do *tu* ou do *cê*, por exemplo, tem relação com o fato de ambos os pais serem, respectivamente, de um mesmo estado de cidades do Nordeste ou de Minas Gerais. Nossos resultados mostraram também que o

uso do *tu* na variedade brasiliense se descola da origem externa dos pais, tendo em vista que, na Vila Planalto, ambos os pais de Brasília também influenciam uma ocorrência maior de uso do *tu* e que, fora da Vila Planalto, em que não há nenhum pai da região Nordeste em jogo, quase não foi constatada a presença de *tu* nas falas analisadas.

Avaliar se houve algum viés na forma de levantamento dos dados em direção aos resultados aqui apresentados é o que pretendemos fazer com a continuidade da presente pesquisa²⁷. Mas já ressaltamos que descobrir que o falante de fora da Vila Planalto responsável pela maioria dos usos de *tu* nesta amostra de controle tem fortes laços familiares com Ceilândia é um fato bastante significativo para a importância da origem dos pais na variação do fenômeno analisado.

Quanto à expansão do uso de *tu* por parte do sexo feminino, nossa hipótese foi parcialmente confirmada. É fato que em nossa amostra haja mais ocorrências de *tu* por parte das mulheres quando comparada com as pesquisas de Lucca (2005) e Dias (2007). Na amostra analisada da Vila Planalto, o sexo feminino usou o *tu* em 145 das 330 ocorrências de pronome, alcançando o valor de 44% e o peso de 0,25, baseado em rodada eneária. Nas rodadas binárias, metodologicamente comparáveis às pesquisas de 2005 e 2007, nossos resultados são de 43,9% de *tu* em falas femininas com o peso de 0,41 (em oposição ao peso dos homens de 0,61). Na pesquisa de Lucca (2005, p. 83), o sexo feminino usou o *tu* em 4 de 17 ocorrências de pronomes, alcançando o valor de 23%, com o peso de 0,09 (em oposição ao peso dos homens de 0,55). Na pesquisa realizada por Dias (2007, p. 75), o sexo feminino usou o *tu* em 50 de 464 ocorrências de pronomes, alcançando o valor de 10,8%, mas com o peso de 0,40 (em oposição ao peso dos homens de 0,60). Percebemos, pelo número de ocorrências e pelos percentuais, que nossa amostra demonstra uma expansão do uso de *tu* pelo sexo feminino quanto ao uso de *tu*. O fato de nossos pesos terem ficado acima da amostra de 2005 e terem sido equivalentes ao de 2007 corrobora a hipótese de expansão. E, ainda, essa equivalência também deve ser analisada do ponto de vista das relações em jogo, uma vez que, embora haja essa expansão, as falas masculinas ainda lideram o uso de *tu* em Brasília.

Por último, sobre a hipótese de a variante *cê* poder ser usada como pronome de esQUIVA, temos os seguintes argumentos a favor dessa confirmação:

- 1) Quando pensamos sobre o tipo de relação entre os falantes, temos que, em relações assimétricas, seria natural pensar que, pela formalidade, o *você* seria

²⁷ Ao continuar a pesquisa, faremos rodadas do tipo *tu versus você e cê*; *cê versus tu e você* e *você versus cê e tu*. Tais rodadas, entre outras coisas, permitirão um aprofundamento metodológico.

o mais utilizado. Mas, em todos os nossos testes sobre o tipo de relação entre os interlocutores, o *cê* foi sempre favorecido pela assimetria, ficando com seus percentuais e pesos semelhantes aos de *você*.

- 2) Vejamos, a seguir, exemplos de fala dos dois únicos informantes que demonstraram preocupação quanto ao uso dos pronomes:

“L: oh TU, {obs: repensando a fala} a SENHORA viu, eu guardei tudo lá?”

“L: {init.} a SENHORA é alta, CÊ dá conta?”

“L: tia, a SENHORA torce pra qual time?”

“L: só se CÊ deixar meio cru”

“L: tia, amanhã a SENHORA vai buscar a gente? CÊ não ia fazer o negócio com ...”

“L: depois coloca o miojo, depois deixa ele ficar mole, quando ficar mole coloca o tempero, depois mexe, mexe, mexe /{init.}/ ontem eu comi uma panelona de miojo...”

“L: eu tenho, eu tenho. Assim, tia, a SENHORA pega a cenoura, bate no liquidificador, depois coloca, mexe, mexe, depois coloca polvilho, eu acho, aí coloca, mexe, mexe, coloca ovo, bate, bate, bate, dois ovos, bate {init.} deixa o toddy ...”

“L: a SENHORA, peraí, a SENHORA vai numa escolinha, fala “eu quero jogar futebol, eu quero fazer futebol aqui” aí eles vão {init.} consultar {init.}”

“A: CÊ não sabe?”

“A: CÊ bota óleo na panela, né?”

“A: bota óleo na panela, né? {init.} Aí deixa, deixa fritando, aí CÊ bota sal...”

“A: é, tem que...”

“A: CÊ bota o...”

Nos primeiros trechos do informante L, ele quase usa o *tu* para se referir à pesquisadora, mas o troca pelo *senhora* e só então completa a frase. Constatamos que, ao longo do registro de fala desse informante, L chama a pesquisadora por *cê* sem reformular sua fala.

A informante A usou pouquíssimos pronomes com a pesquisadora ao longo do registro de sua fala, utilizando também o artifício de nulos, mas, nas poucas ocorrências se referindo à pesquisadora, o pronome utilizado foi sempre o *cê*. Os resultados da presente pesquisa e aqui apresentados sugeriram que a variante *cê* é também utilizada na fala assimétrica e também na fala de pessoas que demonstram preocupação com o tipo de pronome a ser utilizado.

Consideramos que as hipóteses aqui estudadas foram confirmadas. No entanto, sabemos que os resultados que confirmaram a hipótese de o *cê* ser um pronome de esquiva são ainda incipientes. E, ainda, é necessário mais análises que comprovem com pesos a expansão do pronome *tu* pelo sexo feminino. Quanto ao *tu*, percebemos que este tem ganhado cada vez mais espaço no repertório linguístico do brasiliense, mas acreditamos também que mais pesquisas de acompanhamento desta variante são necessárias a fim de perceber se haverá mudança linguística, no sentido de uma variante substituir a outra. Consideramos, todavia, que o fato de o *tu* ter entrado no escopo da fala de Brasília já foi uma mudança. Não acreditamos, portanto, que estejamos diante de uma variação estável. Afinal, no período de 5 anos (período do início das investigações sobre o *tu* em Brasília), já foi possível perceber mudanças significativas quanto a esse uso por parte dos brasilienses.

Acreditamos, por fim, que a grande Brasília passa por estágio de focalização, embora ainda passe por período de difusão dialetal, pois continua recebendo imigrantes e se expandindo vertiginosamente. Ainda não temos registro de que a grande Brasília tenha características culinárias típicas, por exemplo, mas acreditamos que a variedade brasiliense já tem focalizado algumas características linguísticas, tanto lexicalmente quanto gramaticalmente, e o uso do *tu*, como um pronome marcado, figura como uma característica gramatical própria.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- Almeida, Napoleão Mendes de, 1911 – 1998. *Gramática Metódica da Língua Portuguesa* – 45ª Ed. São Paulo: Saraiva, 2005.
- Andrade, Adriana Lília V.S. *A variação você, cê e ocê no português brasileiro falado*. 2004. Dissertação (Mestrado em Linguística) – Universidade de Brasília, Brasília, 2004.
- Bechara, Evanildo. *Moderna Gramática Portuguesa*. 37. ed. rev. e ampl. 14ª reimpr. Rio de Janeiro, 2004.
- Bortoni-Ricardo, Stella Maris. O falar candango: contato de dialetos no Distrito Federal, Brasil. In: Große S. e Zimmerman K. (Eds.). *O português brasileiro: pesquisas e projetos*. Frankfurt am Main, TFM, 2000. p. 329-344.
- Blom, Jan-Peter; Gumperz, John J. *O significado social na estrutura linguística: Alternância de Códigos na Noruega*. In: Ribeiro & Garcez. Porto Alegre: AGE, 1998. p. 31-56.
- Brown, Roger; Gilman, Albert (1960). The pronouns of power and solidarity. In: Brat Paulston, C. e Tucker, G. R. (Eds.). *Sociolinguistics: the essential readings*. Oxford, 2003.
- Corrêa, Cíntia da Costa. *Focalização dialetal em Brasília: um estudo das vogais pretônicas e do /s/ pós-vocálico*. 1998. Dissertação (Mestrado) – Universidade de Brasília, Brasília, 1998.
- Cunha, Celso. *Gramática do português contemporâneo*. 9. ed. Rio de Janeiro: Padrão – Livraria, 1981.

Cunha, Celso & Cintra, Lindley. *Nova Gramática do Português Contemporâneo*. 3ª Edição. 5ª reimpressão. Ed Nova Fronteira. Rio de Janeiro, 2001.

Dias, Edilene Patrícia. *O uso do tu no português brasileiro falado*. (2007).
Dissertação de Mestrado. Brasília, UnB.

Duarte, Maria Eugênia L. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In: Roberts, I. e Kato, M.A. (Orgs.). *Português brasileiro: uma viagem diacrônica*. 2. ed. Campinas: Editora da Unicamp, 1996. p. 107-128.

Faraco, Carlos Alberto. O tratamento de você em português: uma abordagem histórica. *Fragmenta*, Curitiba, Editora da UFPR, n. 13, p. 51-82, 1996.

Goffiman, Erving. *Footing*. In: Ribeiro & Garcez (Orgs.). Porto Alegre: AGE, 1998. p. 70-97.

Gonçalves, Clézio Roberto. *Uma abordagem sociolingüística do uso das formas você, ocê e cê no português*. Tese de Doutorado - São Paulo, USP: 2008.

Guy, Gregory R.; Zilles, Ana. *Sociolingüística quantitativa*. Instrumental e análise. São Paulo: Parábola, 2007.

Hanna, Elizabeth Seixas. *Difusão e focalização dialetal: o caso de Brasília*. 1986. Dissertação (Mestrado em Linguística) – Universidade de Brasília, Brasília, 1986.

Hausen, Telma A. P. *Concordância verbal do pronome “tu” no interior do estado de Santa Catarina*. 2000. Dissertação (Mestrado em Linguística) – UFPR, Curitiba, 2000.

Hazen, Kirk. The Family. In: *The handbook of variation and Change*. Edited by Chambers, Trudgill and Schilling-Estes. Blackwell Publishing, 2005. p. 501-523.

Herênio, Kerlly Karine Pereira. “Tu” e “Você” em uma Perspectiva Intra-Linguística. Dissertação de Mestrado. Uberlândia, Universidade Federal de Uberlândia: 2006.

Ilari, R. et al. Os pronomes pessoais do português falado: roteiro para a análise. In: Castilho, A.T. e Basílio, M. (Orgs.). *Gramática do português falado*. v. 4. Estudos Descritivos. Campinas: Editora da Unicamp, 1996. p. 79-166.

Illari, Rodolfo & Basso, Renato. O português da gente. A língua que estudamos a língua que falamos. 1ª Edição. 1ª impressão - São Paulo: Contexto, 2007.

Labov, William. *Padrões sociolingüísticos*. Tradução de Marcos Bagno, Maria Marta Pereira Scherre, Caroline Rodrigues Cardoso. São Paulo: Parábola, 2008.

Labov, William. Estágios de aquisição do inglês Standard. In: *Sociolingüística*. Tradução de Luíza Leite Bruno Lobo. Maria Stella Vieira da Fonseca e Moema Facure Neves (Orgs.). Rio de Janeiro: Eldorado, 1974.

Lopes, Célia Regina dos Santos; Duarte, Maria Eugênia Lamoglia. De Vossa Mercê a você: análise da pronominalização de nominais em peças brasileiras e portuguesas setecentistas e oitocentistas”. In: Brandão, S. e Mota, M. A. (Orgs.). *Análise contrastiva de variedades do português: primeiros estudos*. Rio de Janeiro: In Fólio, 2003. p. 61-76.

Loregian-Penkal, Loremi. *Concordância verbal com o pronome tu no sul do Brasil*. Dissertação (Mestrado em Letras/Linguística) – UFSC, Florianópolis-SC, 1996.

Lorengian-Penkal, Loremi. *(Re)análise da referência de segunda pessoa na fala da região sul*. 2004. Tese (Doutorado em Letras/Linguística) – UFPR, Curitiba-PR, 2004.

Lucca, Nívia Naves Garcia. *A variação tu/você na fala brasiliense*. 2005. Dissertação (Mestrado), Universidade de Brasília, Brasília, 2005.

MATEUS, M. H. M. et al. Gramática da Língua Portuguesa. 5 ed. Lisboa: Caminho, 2003.

Menon, Odete P. da Silva. A história de você. In: Guedes, M. et al. (Orgs.). *Teoria e análise lingüísticas, novas trilhas*. São Paulo: Cultura Acadêmica, 2006. p. 99-160.

Modesto, Atarxerxes Tiago Tácito. Formas de Tratamento no Português Brasileiro: Alternância Tu/ Você na cidade de Santos, SP. Dissertação de Mestrado, São Paulo, Universidade de São Paulo, 2006.

Naro, Anthony Julius. Modelos quantitativos e tratamento estatístico. In: Mollica, M. C. e Braga, M. L. (Orgs.). *Introdução à sociolingüística: o tratamento da variação*. São Paulo: Contexto, 2003. p. 15-25.

Nascentes, Antenor. O tratamento de "você" no Brasil. *Revista Letras*, v. 5(6). Curitiba. 1956. p. 114-257.

Oliveira, Marco Antônio de. Variável Lingüística: conceituação, problemas de descrição gramatical e implicações para a construção de uma teoria gramatical. *Delta*, v. 3, n. 1, 1987, (p. 19-34).

Oliveira, Sandi Michele de. *Identidade pessoal e a relevância da análise de "frames" (molduras) para um modelo da negociação de tratamento*. Universidade de Copenhagen. Disponível em: <www.ruc.dk/isok/skriftserier/XIV-SRK-Pub/SMO/SMO03-Oliveira/>, 2006.

Pagotto, Emilio Gozze. "Variação e (‘) Identidade". Maceió: EDUFAL, 2004.

Paredes Silva et al. *Variação na 2ª pessoa: o pronome sujeito e a forma do imperativo*. Gragoatá, Niterói, n. 9, p. 115-123, 2. sem. 2000.

Paredes Silva, Vera Lúcia. O retorno do pronome tu à fala carioca. In: Roncarati & Abraçado (Orgs.). *Português Brasileiro: contato lingüístico, heterogeneidade e história*. Rio de Janeiro: 7Letras, 2003. p.160-169.

Pintzuk, Susan. *VERBRUL programs*. 1988. Inédito. Trad. Ivone Isidoro Pinto, rev. Maria Thereza G. Fioretie, coord. Maria Marta Pereira Scherre (original inglês).

Rocha Lima, Carlos Henrique da, 1915 – 1991. Gramática Normativa da Língua Portuguesa; Prefácio de Serafim da Silva Neto. 45ª Ed. – Rio de Janeiro: José Olympio, 2006.

Said Ali, Manuel. *Gramática histórica da língua portuguesa*. 6. ed. São Paulo, Melhoramentos, 1966.

Sankoff, David. Variable Rules. In: Ammon, U. Dittmar, N. & Mattheier, K. J. (Eds.), *Sociolinguistics: an international handbook of the science of language and society*. Nova York, Walter de Gruyter, 1988. p. 984-997.

Sankoff, David. Sociolinguistics and syntactic variation. In: F.J. Newmeyer (ed.), *Linguistics: The Cambridge Survey*. Cambridge: Cambridge University Press. 1988. p. 140-161.

Sankoff, David. "Statistics in sociolinguistics". In: Mesthrie, R (ed.). *Concise Encyclopedia of Sociolinguistics*. Amsterdã, Elsevier. 2001, p. 828-834. Disponível em: <<http://albuquerque.bioinformatics.uottawa.ca/Papers/stats.pdf>>.

Sankoff, Gillian. *Cross-Sectional and Longitudinal Studies in Sociolinguistics*. Disponível em: <<http://www.ling.upenn.edu/~gillian/LongitStudies.html>>, 2002.

Scherre, Maria Marta Pereira; Naro, Anthony Julius. Duas dimensões do paralelismo formal na concordância verbal do português popular do Brasil. *Delta*, vol. 9, n.1, 1993.

Scherre, Maria Marta Pereira; Naro, Anthony Julius. Análise quantitativa e tópicos de interpretação do *Varbrul*. In: Mollica, M.C.e Braga, M. L. (Orgs.). *Introdução à sociolinguística: o tratamento da variação*. São Paulo: Contexto, 2003. p. 147- 178. 2003.

Scherre, Maria Marta Pereira. Paralelismo linguístico. *Revista de Estudos da Linguagem*, v. 7, n. 2, jul-dez, p. 29-59. Belo Horizonte, UFMG, 1998.

Scherre, Maria Marta Pereira. *Doa-se lindos filhotes de poodle*. Variação Linguística, Mídia e preconceito. Brasília: Parábola, 2006.

Scherre et alii. **USOS DOS PRONOMES VOCÊ E TU NO PORTUGUÊS BRASILEIRO**. Simpósio *O português popular do Brasil, Portugal e África: aproximações e distanciamentos*. Évora, Portugal, Outubro de 2009.

Tagliamonte, Sali A. *Analysing Sociolinguistic Variation*. Keys Topics in Sociolinguistics. Cambridge, 2007.

Weinreich, Uriel; Labov, William & Herzog, Martin. *Empirical Foundations for a Theory Language Change. Directions for Historical linguistics: A Symposium*. Austin: University of Texas Press, 1968.

8. APÊNDICE

Arquivo de especificação –

Pesquisa realizada por ocasião do mestrado de Carolina Queiroz Andrade
carollwith@gmail.com; unbcarolinaletras@gmail.com

Análise quantitativa do fenômeno de variação linguística da segunda pessoa do singular, entre as formas TU/ CÊ/ VOCÊ, no Distrito Federal.

A análise quantitativa será realizada através do pacote de programas computacionais GoldvarbX.

Variável dependente:

1. variação linguística:

T – tu

C – cê

V – você

Variáveis independentes:

2. posição em relação ao verbo.

0 – sem verbo

a – antes do verbo

d – depois do verbo

3. idade

@ – menor que 7

1 – de 7 a 10 anos

3 – de 11 a 15 anos

4 – mais de 16

recodificando a idade:

a – 5 anos

b – 7 anos

c – 8 anos

d – 9 anos

e – 10 anos

f – 11 anos

g – 12 anos

h – 13 anos

i – 14 anos

j – 15 anos

4 – 20 anos

R – Rachel

4- vocativo:

v – com vocativo

\$ – sem vocativo

5. sexo

f – feminino

m – masculino

6. função sintática:

s – posição de sujeito

o – objeto direto

j – objeto indireto

p – predicativo do sujeito

w – quando a frase é sem verbo

7- tipo de relato:

r – fala relatada (reportada)

i – original

8 – Entonação:

? – interrogativa

c – declarativa + exclamativa

9 – polaridade (sem expectativa, mais por controle)

a afirmativa

g negativa

10 – paralelismo: a série deve ser considerada como se fosse um turno de fala.

i – isolado

1 – primeiro da série

t – precedido por tu

v – precedido por você

c – precedido por ce

11 – referência do pronome:

g – referência genérica

e – referência específica

12 – Tipo de relação entre os interlocutores:

s – simétricos = mesma idade, amigos.
+ – mesma idade, mas acabaram de se conhecer (casos da escola com a Duda).
a – assimétricos = idades diferentes, maior poder sobre a situação (pesquisador, frequentemente chamado de tia, com o mesmo ou maior grau de formalidade que o professor)

13 – localidade da amostra:

v – Vila Planalto

o – outros lugares (Asa Norte, Jardim Botânico - Lago Sul e Sudoeste)

13 – origem do pai:

n - norte

s - sul

c - centro-oeste

d - sudeste

r - nordeste

14 – origem da mãe:

n – norte

s – sul

c – centro-oeste

d – sudeste

r – nordeste

15 - código por informante(controle de informantes):

Cód./ informates	idade	sexo	Origem mãe	Origem pai	localidade
(L) Luana	5	f	RJ d	AM n	Lago Sul
(T) Tayná	7	f	RJ d	AM n	Lago Sul
@ Gabriela	7	f	BSB c	CE r	Vila Planalto
(z) Beatriz	7	f	BSB c	BSB c	Vila Planalto
(a) Aline	7	f	RN r	RN r	Vila Planalto
(s) Gustavo	7	m	BSB c	MA r	Vila Planalto
(l) Mícael	7	m	BSB c	BSB c	Vila Planalto
(t) Tainara	8	f	BSB c	PB r	Vila Planalto
(K) Keli	8	f	BSB c	BSB c	Vila Planalto
(F) Francisco	8	m	BSB c	BSB c	Vila Planalto
(C) Carlos	8	m	BSB c	BSB c	Vila Planalto
(q) Jaqueline	9	f	PE r	PE r	Vila Planalto
(n) Márcia	9	f	PE r	BSB c	Vila Planalto
(D) Duda	09 e 10	f	BSB c	RJ d	Lago Sul
(U) Udson	09	m	BSB c	BSB c	Vila Planalto
(B) Ana Beatriz	10	f	RJ d	BSB c	Sudoeste
(o) João	10	m	ES d	RJ d	Asa Norte
(i)Amigo do João	10	m	BSB c	BSB c	Asa Norte
(M)Marcos	10	m	BSB c	BSB c	Vila Planalto
(m)Milena	10	f	RN r	RN r	Vila Planalto
(A)Adelaide	11	f	TO c	Maranhão r	Vila Planalto
(h) Luiz Eduardo	11	m	BSB c	Bahia r	Vila Planalto
(y)Yuri	12	m	MG –Unaí d	BSB c	Lago Sul
(N) Brenda	12	f	BSB c	RN r	Vila Planalto
(V) Viviane	12	f	PB r	PB r	Vila Planalto
(G) Glenda	12	f	PI r	CE r	Vila Planalto
(f) Ferreira	12	m	MG d	MG d	Vila Planalto
(Y) Yan	12	m	RJ d	BSB c	Sudoeste
(E) Daniela	13	f	PE r	PE r	Vila Planalto
(R) Raissa	13	f	MG d	RN r	Vila Planalto
(k) Kleandra	13	f	BSB c	MT c	Vila Planalto
(Q)Roquera	13	f	PI r	CE r	Vila Planalto
(p)Pablo	13	m	GO c	RS s	Vila Planalto
(b)Bruna	14	f	RJ d	BSB c	Sudoeste
(l) Ingrid	14	f	BSB c	BSB c	Vila Planalto
(J) Jussara	14	f	BSB c	BSB c	Vila Planalto
(u) Bruno	14	m	PB r	MG d	Vila Planalto
(X) Jaqueline	15	f	BA r	BA r	Vila Planalto
(W) Mayara	15	f	BsB c	PB r	Vila Planalto
(1) B	15	m	-	-	Vila Planalto
(2) C	15	m	-	-	Vila Planalto
(3) ? = João	15	m	PE r	PE r	Vila Planalto
(P)Pedro Henrique	15	m	PE r	PE r	Vila Planalto
43 informantes		25 fem. / 18 mas.			

16 – refinando as origens – mãe:

Amazonas – a

Rio de Janeiro – r
Brasília – b
Ceará – c
Rio Grande do Norte – n
Maranhão – m
Pernambuco – e
Paraíba – i
Espírito Santo – s
Tocantins – t
Minas Gerais – q
Bahia – h
Pará - @
Piauí – u
Mato Grosso – o
Goiás – g
Rio Grande do Sul – S

17 - Pai: (os mesmos códigos para mãe, na coluna 16)

18 – Mesma origem do pai e da mãe:

N – ambos do Rio grande do Norte
B – ambos de Brasília
E – ambos de Pernambuco
I – ambos da Paraíba
Q – ambos de Minas Gerais
H – ambos da Bahia
/ - quando a origem não for a mesma do pai e da mãe.

Última coluna de origem dos pais, feita a partir do arquivo de condições:

f um dos pais de Minas Gerais
\$ um dos pais de Minas Gerais
P ambos do nordeste
K ambos de Brasília
n um do nordeste (de Brasília também)
k um de Mato Grosso
? de fora da Vila, sem influência do nordeste

