

PROJETO DE VARIAÇÃO LINGÜÍSTICA NO CENTRO-OESTE (VALCO)

Rachel do Valle DETTONI*

Cíntia da Silva PACHECO**

Carolina Queiroz ANDRADE***

Maria Marta Pereira SCHERRE****

- **RESUMO:** O objetivo deste artigo é apresentar o Projeto Variação Linguística no Centro-Oeste (VALCO) e sintetizar resultados de pesquisa sociolinguística em áreas geográficas distintas da região Centro-Oeste e em temáticas também distintas. Quatro pesquisas foram realizadas com dados da grande Brasília ou Distrito Federal e analisam os pronomes de segunda pessoa *tu/você/cê/ocê*. Nestas pesquisas, apresenta-se a ideia central de que a variedade brasiliense em formação está introduzindo a variante *tu* em seu repertório linguístico, com focalização de um traço marcado no português brasileiro. Duas outras pesquisas foram realizadas com dados da Baixada Cuiabana, Estado do Mato Grosso, e analisam a variação de gênero no plano discursivo e sintagmático. Nestas duas pesquisas, diferentemente, apresentam-se traços linguísticos de natureza semicategórica, em provável situação de extinção, mas evidencia-se que a variação de gênero nos dados analisados é linguisticamente encaixada no sistema da língua portuguesa. Sendo assim, o presente artigo, que se orienta pela Teoria da Variação e da Mudança Linguística, apresenta para a comunidade científica um pouco mais da riqueza da língua portuguesa em terras brasileiras, mais especificamente, no Brasil central tradicional e no Brasil central expandido.
- **PALAVRAS-CHAVE:** Variação linguística. Região Centro-Oeste. Variedade brasiliense. Variedade Cuiabana. Pronomes de segunda pessoa. Concordância de gênero. Focalização dialetal.

Breve histórico do projeto VALCO

A preocupação em se pesquisar a fala de Brasília não é recente (BORTONI-RICARDO, 1985; HANNA, 1986; MALVAR, 1992; DIAS, 1993; CORRÊA, 1998), e essa preocupação cresceu à medida que os brasilienses foram sentindo necessidade de uma identidade linguística frente aos falantes de outras regiões do país. Como essa necessidade de “identificação” é algo natural nos brasilienses, foi natural também

* UnB – Universidade de Brasília. Departamento de Linguística, Português e Línguas Clássicas. Brasília – Distrito Federal – Brasil. 71000-000 - rdettoni@gmail.com

** UnB – Universidade de Brasília. Departamento de Linguística, Português e Línguas Clássicas. Brasília – Distrito Federal – Brasil. 71000-000 - cintialetras@yahoo.com.br

*** UnB – Universidade de Brasília. Departamento de Linguística, Português e Línguas Clássicas. Brasília – Distrito Federal – Brasil. 71000-000 - carollwith@gmail.com

****UFES – Universidade Federal do Espírito Santo. Centro de Ciências Humanas e Naturais – Departamento de Línguas e Letras. Vitória – Espírito Santo – Brasil. 29075-910 - mscherre@terra.com.br

que algumas pesquisas alocadas na Universidade de Brasília se propusessem a entender esse dialeto nascituro.

Pesquisas mais recentes (ANDRADE, 2004; LUCCA, 2005; DIAS, 2007; ANDRADE, 2010) corroboram o interesse dos pesquisadores de analisar cientificamente o que pode ocorrer linguisticamente em uma cidade planejada, com um crescimento diferenciado, como é o caso de Brasília.

Portanto, pensando em um projeto que pudesse identificar, documentar e caracterizar uma variedade linguística própria do Distrito Federal, dentro do cenário linguístico nacional, no segundo semestre de 2008, reuniu-se um grupo de professores e estudantes de pós-graduação da Universidade de Brasília ligados à área de pesquisa da Sociolinguística Variacionista para discutir a implementação de um projeto de registro de fala. Dessa reunião, surgiu a ideia de agregar, aos corpora do Distrito Federal, corpora de outras cidades da região Centro-Oeste. Desse modo, a exemplo do que fizeram pesquisadores paraibanos, com o VALPB, e sulistas, com o VARSUL, esses pesquisadores, sob coordenação da professora Rachel Dettoni (UnB), resolveram organizar um material de fala já coletado por pesquisadores envolvidos em projetos de pesquisa das universidades de Brasília, Goiás, Mato Grosso e Mato Grosso do Sul para compor um banco de dados on-line. Assim, surgiu oficialmente, em julho de 2008, o Projeto Variação Linguística no Centro-Oeste (VALCO).

O primeiro objetivo do VALCO é reunir e socializar o conhecimento e a pesquisa linguística que vem sendo desenvolvida no âmbito da variação linguística, disponibilizando material de fala coletado por meio de entrevistas labovianas típicas para todos os profissionais que atuem na pesquisa e no ensino de língua portuguesa, em nível fundamental, médio ou superior, uma vez que, em geral, esse conhecimento permanece restrito ao ambiente acadêmico. Um banco de dados on-line possibilitará o desenvolvimento de novas pesquisas sobre a diversidade linguística do Centro-Oeste, propiciando a ampliação do conhecimento acerca da identidade linguística e sociocultural dessa região do Brasil.

A formulação desse projeto contou com a colaboração fundamental da professora Marta Scherre (UnB/UFES), como uma figura que semeou boa parte do escopo do Valco (deixemos claro, no entanto, que o Valco é um projeto aberto a novas pesquisas de outras áreas da linguística), atuando como orientadora de todas as pesquisas descritas neste artigo e outras que aqui não entraram, sendo ao todo oito dissertações de mestrado e uma tese de doutorado com dados coletados na região. Scherre continua orientando alguns doutorandos alocados na UnB, sendo dois deles estudos de dados de fala do Centro-Oeste. Além de Scherre, o Valco também conta com a coordenadora e idealizadora Rachel Dettoni (UnB), que entendeu a necessidade de alocar em um ambiente todo esse nicho de pesquisa linguística. Compõem ainda o grupo Valco as

pesquisadoras-colaboradoras, doutorandas da UnB, Carolina Andrade e Cíntia Pacheco. O intuito maior é ainda buscar colaboradores de outras universidades federais e/ou estaduais da região Centro-Oeste.

Os dados coletados estão disponíveis on-line em arquivos de áudio e de transcrição, que poderão ser acessados mediante autorização prévia dos administradores do site apenas para audição e visualização on-line, não sendo permitido o download de quaisquer conteúdos sem prévia autorização. Os resultados de pesquisas realizadas com dados provenientes do projeto VALCO devem dar os devidos créditos ao Projeto e podem ser divulgados no site.

Esse site é hospedado pelo servidor da Universidade de Brasília, e o Projeto VALCO é financiado com recursos da Fundação de Apoio à Pesquisa do Governo do Distrito Federal (FAPDF).

Resumo do VALCO

O Projeto VALCO pertence à área de Linguagem e Sociedade do Programa de Pós-graduação em Linguística da Universidade de Brasília, e sua linha de pesquisa é língua, interação sociocultural e letramento.

Esse projeto visa à organização de um banco de dados de língua falada da região Centro-Oeste a ser disponibilizado on-line para promover o estudo e a pesquisa das variedades linguísticas da região. A relevância desse projeto está apoiada nos seguintes aspectos:

- a) Necessidade de se identificar, documentar e caracterizar uma variedade linguística própria do Distrito Federal, dentro do cenário linguístico nacional;
- b) Necessidade de investigação científica de uma identidade linguística regional, principalmente no que diz respeito ao Distrito Federal;
- c) Urgência em estabelecer, na região Centro-Oeste, um alinhamento com uma tendência nacional da pesquisa linguística, no sentido de organizar bancos de dados representativos para documentar e estudar as manifestações linguísticas regionais, contribuindo para a construção de um perfil mais completo da diversidade do português falado no Brasil, à luz do que já vem sendo desenvolvido por outras universidades federais em projetos como, só para citar alguns, o VARSUL, na região Sul; VALPB, na Paraíba; PEUL, no Rio de Janeiro; PHPB (Projeto para a História do Português Brasileiro), que envolve várias universidades federais e estaduais (UFMG, UFRJ, UFBA, UNICAMP, USP, UFPB, UFPE, UEMS);
- d) Desenvolvimento de novas tecnologias visando à divulgação da pesquisa científica na área de sociolinguística, seus resultados e contribuições para o ensino em geral, promovendo a articulação entre a universidade e demais profissionais da área

de educação, especialmente os docentes de escola pública do Distrito Federal, considerando a inclusão dos estudos sobre variação linguística no ensino de língua portuguesa para os níveis fundamental e médio, prevista pelos Parâmetros Curriculares Nacionais (PCN) definidos pelo Ministério da Educação.

Para desenvolver o referido projeto, já dispomos de algum material coletado, em pesquisas individuais, por alunos de graduação e pós-graduação do curso de Letras (Linguística e Língua Portuguesa) da Universidade de Brasília, que forma professores de língua portuguesa para atuarem nas escolas do Distrito Federal, e também por alunos e pesquisadores de outras universidades da região Centro-Oeste – UFG, UEG, UFMT e UFMS.

A organização dos dados se dá por meio da criação de um site hospedado pelo servidor da Universidade de Brasília. As entrevistas que compõem a amostra são digitalizadas e devem estar transcritas pelo próprio pesquisador interessado, segundo as normas de transcrição do VALCO (Anexo I).

Todo acesso ao banco de dados só será efetivado mediante autorização prévia da equipe de coordenação do projeto, composta inicialmente por professores e alunos de pós-graduação do Departamento de Linguística, Português e Línguas Clássicas da UnB.

Pretende-se que a manutenção e a constante alimentação do site sejam efetuadas por monitores ligados ao Departamento de Linguística da UnB. Por razões éticas, e de acordo com o que preconiza a ética das pesquisas científicas, a identidade dos informantes das entrevistas será preservada.

Além do banco de dados de fala da região, pretende-se inserir no siteum espaço para a divulgação de artigos, teses, monografias e trabalhos advindos das pesquisas realizadas com o referido material.

Principais propostas e objetivos do VALCO

A proposta desse projeto parte do princípio de que há, no Distrito Federal, uma variedade linguística regional em processo de constituição e definição de seus traços típicos em tempo real, caso curioso e de extrema relevância para a pesquisa linguística tanto quanto para a formação dos profissionais de Letras que atuam na educação.

A documentação sobre a linguagem falada no Distrito Federal tem dado seus primeiros passos, restritos ainda ao meio acadêmico, necessitando de mais incentivo para que se possa configurar uma amostra mais representativa que nos permitirá caracterizar, com mais segurança e precisão, os traços linguísticos que estão se consolidando como formas típicas desta região e construindo uma identidade local.

Nesse sentido, partimos da hipótese de que há uma manifestação linguística em processo de definição em Brasília e em todo o Distrito Federal que ainda não foi suficientemente documentada e estudada de modo que possamos definir suas características e contornos locais.

Além disso, diferentemente do que vem sendo realizado em outros estados e regiões do Brasil, a região Centro-Oeste não dispõe ainda de nenhum grupo de pesquisa sociolinguística consolidado, apesar de muitas pesquisas isoladas já terem sido desenvolvidas ou estarem sendo feitas neste momento, nesta área, por doutorandos, mestrandos e alunos de graduação.

O material que tem sido coletado por esses pesquisadores e professores em formação representa uma amostra já bastante significativa do que é possível construir e documentar em termos de pesquisa e construção do conhecimento científico sobre a língua portuguesa falada no Brasil e suas diversas manifestações regionais.

Assim sendo, os objetivos principais do VALCO são:

- a) Ampliar a amostra de língua falada no Distrito Federal e entorno, o que permitirá mapear a diversidade linguística própria dessa região, que congrega, em seu território, um mosaico de manifestações e falares de diversas regiões do Brasil;
- b) Realizar o estudo e a sistematização dos traços linguísticos que caracterizam a(s) variedade(s) linguística(s) do Distrito Federal, trabalho ainda não realizado de modo sistemático, que terá também repercussões em nível nacional na medida em que fornecerá novos elementos para a história e o conhecimento do português do Brasil;
- c) Organizar e implementar um banco de dados linguísticos de língua falada no Distrito Federal e demais estados da região Centro-Oeste, cujo grande impacto consiste em viabilizar, para o maior número possível de profissionais (professores e pesquisadores) da área da educação, o acesso tanto aos dados linguísticos e amostras de fala de toda essa região quanto aos resultados das pesquisas e estudos (teses, dissertações e artigos) que forem sendo desenvolvidos pelos pesquisadores interessados na área;
- d) Promover e incentivar o estudo, o conhecimento e a caracterização, por meio de pesquisa científica, das variedades linguísticas do Distrito Federal e demais estados do Centro-Oeste;
- e) Capitalizar e socializar os resultados das pesquisas sociolinguísticas que vêm sendo desenvolvidas não só pelo Programa de Pós-Graduação em Linguística (PPGL) e pelo Programa de Iniciação Científica de alunos graduandos da UnB mas também pelos Programas de Pós-Graduação das outras universidades federais da região Centro-Oeste;

- f) Promover a utilização de novas tecnologias a serviço da socialização do conhecimento científico e da articulação entre a UnB e escolas de nível fundamental e médio do DF;
- g) Oferecer suporte teórico-pedagógico para professores e pesquisadores que atuam no ensino da língua portuguesa, em nível fundamental (6º a 9º ano) e médio, divulgando e incentivando o pensamento científico e o estudo da diversidade linguística nos outros níveis de ensino, além do superior.

Busca-se, então, legitimar uma identidade cultural por meio do patrimônio linguístico, caracterizando a identidade sociolinguística e cultural do Distrito Federal e demais regiões, por meio do estudo sistemático dos traços linguísticos que definem as variedades regionais locais. Assim, é importante também oportunizar a formação de crianças e jovens e a produção científica de novos pesquisadores, incentivando o estabelecimento de uma cidadania participativa por meio de reconhecimento das diversidades linguísticas e conseqüente respeito a elas.

Metodologia do projeto e das pesquisas

O projeto comporta, em princípio, duas ações básicas a serem desenvolvidas. A primeira diz respeito à criação e organização do banco de dados, o que se dará a partir da reunião seletiva das amostras de fala regional já coletadas por pesquisadores e professores em formação na UnB. Essa ação prevê que, numa segunda etapa, sejam contempladas também as amostras coletadas e organizadas por pesquisadores das outras universidades do Centro-Oeste, a saber: UFG, UEG, UFMT e UFMS.

De posse desse material já coletado, proceder-se-á à organização do banco de dados on-line, o que permitirá o acesso de pesquisadores e professores dos vários estados.

A segunda ação refere-se à constante alimentação do site e do banco de dados com as novas coletas que forem sendo realizadas a partir dos trabalhos de novos alunos e pesquisadores interessados, além da divulgação dos trabalhos de pesquisa, como dissertações, teses, artigos e trabalhos de iniciação científica de graduandos que forem sendo desenvolvidos. A metodologia específica da pesquisa e coleta de dados será orientada pelos pressupostos teórico-metodológicos da sociolinguística variacionista de base laboviana, que prevê a coleta de dados de língua falada em contextos reais de comunicação e interação.

Em suma, os objetivos são incentivar e alimentar constantemente o espírito e a prática de pesquisa científica sobre a linguagem, ampliar o banco de dados, alimentando-o sempre com novo material coletado, divulgar resultados de

pesquisas e registros científicos sobre a linguagem do Centro-Oeste e Distrito Federal que forem sendo desenvolvidos e tantos outros. Para isso, o VALCO já tem o material sobre a Baixada Cuiabana disponível para a pesquisa. Posteriormente, serão introduzidas as entrevistas do Distrito Federal.

Para acessar o VALCO, os pesquisadores interessados deverão enviar uma mensagem para o e-mail (valco.unb@gmail.com) do projeto e solicitar permissão para o acesso às transcrições e aos áudios. Na parte das transcrições, é possível fazer uma busca pelo fenômeno linguístico desejado ou por qualquer outro aspecto que interesse ao pesquisador.

Resultados das primeiras pesquisas variacionistas que constam (constarão) no projeto VALCO

Os pronomes de segunda pessoa no Distrito Federal

Dentro do escopo do Projeto VALCO, temos as pesquisas de cunho variacionista sobre os pronomes de segunda pessoa do singular na fala de Brasília. Até o momento foram realizadas quatro dissertações sobre o tema: Andrade (2004); Lucca (2005); Dias (2007) e Andrade (2010). Alguns dos principais resultados das pesquisas serão apresentados a seguir.

Andrade (2004) realizou uma pesquisa com as variantes você, ocêecê em Brasília, com uma amostra de fala da faixa etária entre 10 e 14 anos, incluindo também os dados da pesquisadora, com, na época, 25 anos. A amostra foi constituída em 1991, feita a partir de gravações labovianas típicas, para dissertação de mestrado de Malvar (1992), a organizadora da amostra. A localidade específica de Brasília em que a amostra foi composta foi a RA Sobradinho, fundada em 1960. A amostra é composta por falantes residentes em áreas urbanas e rurais da mesma RA.

Os resultados que a autora alcançou da variação dos pronomes estão descritos na tabela a seguir:

Tabela 1 - Distribuição dos pronomes entre os falantes de 10 a 14 anos de Sobradinho em 1991

Área	TU	VOCÊ	CÊ	OCÊ	TOTAL
Área urbana	0%	63%	31%	6%	90
Área rural	0%	44%	50%	6%	32

Fonte: Andrade (2004).

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 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 foi relacionado ao sexo¹, em que as pessoas do sexo masculino favoreceram o uso do *cê*, com o peso de 0,61, enquanto as pessoas do sexo feminino desfavoreceram esse uso, com peso de 0,40 (ANDRADE, 2004).

Andrade (2004) também analisa a função sintática, e um resultado relevante é a verificação da fala da entrevistadora (com 25 anos na época da gravação). Vejamos a tabela a seguir:

Tabela 2 - Função sintática da variação *você* / *cê* / *ocê* em Brasília (dados da entrevistadora)

	VOÇÊ	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%

Fonte: 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 *ocê*, mas categoricamente não poderá ser *cê*, pois, nos dados de Andrade (2004), não há casos dessa forma como objeto. O *ocê* parece não ter restrição sintática, apesar de essa forma ser escassa nas amostras analisadas. Já o *cê* é sintaticamente restrito à função de sujeito. 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 *decê* foram amalgamados 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

¹ Nas pesquisas da sociolinguística variacionista, tem-se utilizado a codificação relacionada a sexo biológico, e Labov (2001, 2008) considera que deve ser assim senão perdemos a comparabilidade dos resultados, mas a interpretação dos dados tem sido de gênero, uma noção social e/ou cultural.

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, entre 15 a 19 anos. A seguir, apresentamos resultados gerais da pesquisa:

Tabela 3 - Distribuição dos pronomes entre os falantes de 15 a 19 anos na Grande Brasília em 2005

Região	TU	VOCÊ	CÊ	TOTAL
Ceilândia	87%	8%	5%	121
Taguatinga	66%	25%	9%	244
Brasília	68%	5%	28%	87
Total	72%	17%	11%	452

Fonte: Lucca (2005).

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 estatístico, 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 satélite 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 brasiliense em função do grande contingente de imigrantes nordestinos que hoje vive no Distrito Federal. O fato de Ceilândia ter se destacado com mais ocorrências da variante *tu* corrobora essa hipótese, 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” (LUCCA, 2005, 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). A seguir, os resultados gerais da pesquisa:

Tabela 4 - Distribuição dos pronomes por faixa etária no Plano Piloto ampliado, sem Vila Planalto, em 2007

Faixa Etária	TU	VOCÊ	CÊ	TOTAL
13-19 anos	30%	15%	55%	171
20-29 anos	13%	22%	65%	424
30-48 anos	4%	36%	60%	305

Fonte: Dias (2007).

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 peso de 0,56; (b) conversa profissional ou acadêmica, com peso de 0,17; (c) repreensões, com o peso de 0,34; e (d) observações irônicas, deboches e brincadeiras, com 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 5 - A variação tu/ você no Distrito Federal em função da idade

Faixa Etária	Frequência de TU	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%	-

Fonte: 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 essa 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 6 - A variação *tu/ você* no Distrito Federal em função do sexo

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%	-

Fonte: 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 7 - A variação *tu/ você* no Distrito Federal em função do sexo e idade

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%

Fonte: Dias (2007, p.76) – com adaptações.

Dias (2007) analisa as demais variáveis selecionadas e não selecionadas. 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 partiu exatamente dessa 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 falantes do sexo feminino.

Andrade (2010) realizou sua pesquisa baseada nas três primeiras pesquisas, sobre a variação pronominal entre *cê/tu/você*, na região administrativa nominada Vila Planalto (com apenas 30% dos dados de fora da Vila, no Plano Piloto).

A coleta de dados de Andrade (2010) se deu por gravações de conversas tomadas como espontâneas. Porém, como os falantes tinham conhecimento da gravação, esse fato pode ter enviesado a naturalidade nos dados. Os falantes gravados foram contatados por meio de uma escola pública da Vila Planalto e compreendem as idades entre 07 a 15 anos. Andrade (2010) encontrou fatos linguísticos interessantes em função das origens de pais e mães dos falantes, pois percebeu que o uso dos pronomes pelas crianças estudadas parecia corresponder, relativamente, ao uso dos pronomes nas localidades de origens de seus pais. Filhos de pais mineiros, por exemplo, apresentavam maior tendência de uso da forma *cê*, enquanto filhos de pais nordestinos, maior tendência de uso da forma *tu*. Apesar de Andrade (2010) ter dividido as idades em duas faixas etárias – de 7 a 11 anos, as crianças; e de 12 a 15 anos, os adolescentes –, os resultados demonstraram que tanto as crianças como os adolescentes usaram os três pronomes com a mesma frequência e esse fato é relevante para o entendimento em relação à focalização dialetal, por um lado, e em relação à aquisição linguística, por outro. A seguir, apresentaremos resultados em função do sexo:

Tabela 8 - Efeito do fator sexo em percentuais e pesos (dados da Vila Planalto)

Sexo	VOCE	CE	TU
Feminino	113/330 = 34%	72/330 = 22%	145/330 = 44%
	Peso = 0,44	Peso = 0,30	Peso = 0,25
Masculino	36/244 = 15%	80/244 = 33%	128/244 = 52%
	Peso = 0,24	Peso = 0,34	Peso = 0,41

Fonte: Andrade (2010, p.100).

Os resultados obtidos em função da variável sexo em rodada eneária com três variantes (peso 0,33 como referência) indicaram que ambos os sexos tendem a usar a variante *cê* de forma neutra, com pesos de 0,30 e 0,34. 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 alcançados pela autora também indicaram que, em comparação com os resultados de 2005, vêm diminuindo as diferenças entre os sexos masculino e feminino quanto ao uso da variante *tu*. Mas esse resultado também pode ter ocorrido devido às diferenças quanto à coleta de dados. Embora os resultados de 2010 tenham apresentado os percentuais de 44% de uso de *tu* pelo sexo feminino, ficando bem acima dos percentuais alcançados por Dias em 2007, os pesos ficaram desfavorecidos por esse fator (sexo feminino), de forma semelhante aos resultados

de 2007, com pesos um pouco abaixo do peso de referência. Assim, a diminuição entre as diferenças de uso do *tu* por ambos os sexos não foi revelada na amostra de 2010 pelos resultados em pesos.

Andrade (2010), como apontado anteriormente, analisou diversos fatores que influenciam a presente variação. Entre eles, Andrade (2010) chama a atenção para a importância da origem dos pais em relação à variante mais usada pelos jovens falantes, fato apontado na análise de Lucca (2005) e posteriormente comprovado e aprofundado na análise realizada em 2010.

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

Tabela 9 - A variação *tu/você/cê* no DF, região Centro-Oeste

	VOCE	CÊ	TU
Adriana (2004)²	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%
Andrade (2010)	318/835 = 38,1%	229/835 = 27,4%	288/835 = 34,5%

Fonte: Andrade (2004, p.45); Lucca (2005, p.78); Dias (2007, p.64) e Andrade (2010, p.79) - com adaptações - Retirados os “nulos e ocês”, de Andrade (2004), e amalgamados o “você” e o “cê” nos resultados de Lucca e Dias (2005; 2007).

Com base nas quatro pesquisas, chegamos a diversas conclusões, entre as quais destacamos as seguintes:

- 1) O *tu* é uma variante inovadora no Distrito Federal, configurando, entre outros aspectos, que em Brasília, além de haver um movimento em direção à focalização dialetal, houve também uma mudança no paradigma pronominal. O *cê*, como uma variante instigante e sintaticamente restrita, necessita ser mais analisado em termos sociais e linguísticos, contudo os estudos de 2010 apontaram que este possui um papel de esquiva. O pronome *você* parece ser a variante *default*, o que significa que ela é mais neutra, menos marcada, ou que possui mais *status* de prestígio (porém essa interpretação depende de um referencial interacional).
- 2) Muitos fatores sociais são instigantes quanto ao estudo da presente variação, destacando-se entre eles: sexo, faixa etária, origem dos pais (junto à localização dentro do Distrito Federal), comunidade de prática e estilo do falante. Tais fatores necessitam de mais análises para a descrição mais acurada da presente variação.

² É importante lembrar que a pesquisa foi realizada em 2004, mas os dados foram gravados em 1991. 2

- 3) O estudo dos pronomes nos permite ter uma visão linguística e extralinguística do comportamento das variações em línguas naturais.

O site do projeto Valco, além de dar acesso às dissertações (aqui comentadas) na íntegra, em PDF, também dá acesso aos arquivos em áudio, das gravações realizadas e suas respectivas transcrições.

Vale lembrar que os arquivos em áudio e suas transcrições dão acesso à fala brasiliense em diferentes anos compreendidos entre 1991 e 2009. Os dados analisados foram em função da variação dos pronomes, mas trazem luz aos tipos de metodologias utilizadas para sua coleta. Assim, os próximos parágrafos são uma compilação do que os áudios e as transcrições trazem para o pesquisador que deseje acessá-las.

Andrade (2004) examinou o *corpus* de língua falada recolhido em parte das entrevistas constantes da dissertação de mestrado de Malvar (1992), em Sobradinho. As entrevistas foram realizadas no ano de 1991, em que foram analisados dados de quinze falantes, nas faixas etárias entre 10 e 14 anos, além da fala da própria entrevistadora, que à época das entrevistas estava com 25 anos de idade. Com relação ao sexo, oito falantes de 10 a 14 anos eram meninos e sete, meninas. A entrevistadora é natural do Rio de Janeiro, mas vivia em Brasília há cerca de 19 anos. Ao todo, foram levantados 189 dados.

Em Lucca (2005), os dados coletados compreendiam falantes da mesma faixa etária, entre as idades de 15 a 19 anos, sendo cerca de 90% dos dados de fala de meninos, residentes em três RAs distintas: Plano Piloto, Taguatinga e Ceilândia. Ao final da coleta dos dados, Lucca obteve cerca de 55 minutos de gravação com falantes de Ceilândia, 82 minutos com falantes de Taguatinga e 45 minutos com falantes da região administrativa de Brasília, somando 182 minutos de gravações e resultando em 453 dados.

Em Dias (2007), foram analisadas aproximadamente 23 horas e 30 minutos de conversa, que totalizaram 900 dados, sendo todos os falantes residentes no Plano Piloto. As faixas etárias estudadas foram: de 13 a 19 anos; de 20 a 29 anos; e mais de 30 anos. Foram escolhidos três falantes de cada sexo para cada uma das faixas etárias, resultando num total de 18 falantes.

Andrade (2010) realizou 14 transcrições diferentes (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 moradoras da Vila Planalto e 30% dos dados de moradores de outras localidades, sejam elas: 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, entre 7 e 15 anos.

As pesquisas sobre a concordância de gênero na Baixada Cuiabana

A Baixada Cuiabana é a região do Estado de Mato Grosso compreendida entre o Pantanal mato-grossense (ao Sul) e a “portentosa” Chapada dos Guimarães (ao Nordeste), constituída pelos municípios que margeiam os rios Cuiabá e Paraguai. O principal polo de desenvolvimento da baixada é a cidade de Cuiabá, atual capital de Mato Grosso.

As 30 entrevistas do Mato Grosso, mais especificamente da Baixada Cuiabana, foram realizadas por Rachel do Valle Dettoni, entre julho de 2001 e agosto de 2002, para sua tese de doutorado: “A concordância de gênero na anáfora pronominal: variação e mudança linguística no dialeto da Baixada Cuiabana – Mato Grosso” (2003). Anos depois, Cíntia da Silva Pacheco utiliza parte dessas entrevistas para seu trabalho de mestrado: “Padrões sociolinguísticos da concordância de gênero na Baixada Cuiabana” (2010).

O fenômeno linguístico analisado por Dettoni (2003) é a concordância de gênero em configurações discursivas que se destacam por apresentarem nomes explicitamente femininos retomados por pronomes na forma masculina: **Uma dor segura todinho meu corpo, beleza. Quando ele vai aliviano eu tô podre, criatura.**

Pacheco (2010) analisa a concordância de gênero nas relações internas dos sintagmas nominais (**noite inteiro, aquele argola bonito, cara feio**) e na relação entre sujeito e predicado ou nos sintagmas predicativos (**minha mãeeravivo, esta mãotavabobo**).

Ambas as pesquisas são de linha variacionista, baseadas na Teoria da Variação e da Mudança Linguística (LABOV, 1972, 2008; WEINREICH; LABOV;HERZOG, 1968). A análise quantitativa dos dados foi feita pelo pacote de programas VARBRUL (PINTZUK, 1988) na pesquisa de Dettoni (2003) e pelo GOLDFARB-X (SANKOFF; TAGLIAMONTE; SMITH, 2005) na pesquisa de Pacheco (2010).

Sobre o trabalho de Dettoni (2003), a autora afirma que não há possibilidade de gênero neutro, porque o falar cuiabano não apresenta itens lexicais e funções sintáticas específicas para um terceiro gênero (o neutro). Outra possibilidade descartada é a de que a marcação de gênero seja um fato gramatical ausente no dialeto da baixada cuiabana, porque se trata, na verdade, de uma forma variável de marcar o gênero.

A anáfora, por ser uma relação de referência que o falante faz por meio das expressões linguísticas, deve ser examinada como um fenômeno que se situa na interface entre a sintaxe, a semântica e a pragmática.

As intenções do falante, que estão associadas à representação mental que ele faz do antecedente, determinam muito mais a escolha da forma pronominal presente em uma relação anafórica do que, estritamente, as características gramaticais do antecedente. E isto tem a ver com questões de ordem pragmática.

Das 18 entrevistas codificadas por Dettoni (2003), obteve-se um total de 829 dados, sendo 233 do pronome *ele* retomando antecedentes femininos, ou seja, 28% dos dados. Exemplificando o fenômeno analisado, temos, respectivamente, a presença da concordância e a ausência da concordância na retomada anafórica:

Presença da concordância de gênero na retomada anafórica

Essa banana mariquita é só crua, **ela** num frita.

Ausência da concordância de gênero na retomada anafórica

Um dia tava entrando **uma cobra** lá dentro, a gente tava assistindo, quando vê **ele** vinha entrando, sabe.

Entre as variáveis linguísticas selecionadas estatisticamente pelo programa VARBRUL, a anáfora com a retomada do antecedente feminino pelo pronome *ele* aparece mais fortemente condicionada pela referência genérica (0,84) do que pelo traço [inanimado]: é o aspecto da referência genérica e não o da animacidade que entra em jogo no entendimento desse fenômeno variável. Assim, o *ele* retoma qualquer antecedente, seja feminino ou masculino, prioritariamente quando a referência é genérica (Se **a pessoa** já vem com aquela intenção de não respeitar você, você mostra que respeita **ele**).

A natureza morfológica do antecedente também interfere nesse fenômeno, pois nomes invariáveis ou inanimados favorecem a retomada por *ele* (0,54). Segundo Dettoni (2003), o traço [-humano, -animado] (Tira, putcha **brasa**, apaga, djoga na água, né, aí, num demora **eleta** fria.) funciona como um divisor de águas no sistema de atribuição de gênero desse dialeto.

A análise da variável presença x ausência de determinante junto ao antecedente revelou que a ausência de determinante (0,61) é um fator favorecedor da retomada por *ele*, corroborando o paralelismo formal, no sentido de que formas gramaticais semelhantes tendem a aparecer juntas (SCHERRE, 1988).

Com relação às variáveis sociais, o grau de escolaridade confirma a hipótese de que a escolaridade (0,34) mais alta tende a inibir a retomada de itens lexicais femininos por *ele*. Nas faixas etárias, a retomada anafórica com o pronome *ele* é mais generalizada na faixa etária acima de 60 anos e mais restrita na fala dos informantes de 15 a 25 anos.

Tendo em vista que a variedade linguística da Baixada Cuiabana encontra-se em processo de mudança linguística no sentido da perda de traços típicos dialetais, que são estigmatizados, em prol da aquisição e fixação de formas mais prestigiadas do português, Dettoni (2003) aprofunda a discussão sobre os cinco problemas da mudança linguística na Baixada Cuiabana, a saber, os problemas da restrição, da transição, do encaixamento, da implementação e da avaliação.

O problema da restrição refere-se a essa variação que ocorre em contextos cada vez mais restritos e tem baixa frequência na fala dos jovens, sendo mais comum na fala de pessoas menos escolarizadas. A mudança linguística em curso no dialeto da Baixada Cuiabana pode ser descrita, atualmente, como a perda da codificação morfossintática do traço semântico [\pm referência genérica].

O problema da transição envolve a descoberta do caminho pelo qual um estágio da mudança evoluiu a partir de um estágio anterior (LABOV, 1972). Segundo Weinreich, Labov e Herzog (1968, p.184): “[...] a mudança ocorre (1) quando um falante aprende uma forma alternativa; (2) enquanto duas formas coexistirem em sua competência; (3) quando uma das formas se torna obsoleta”. Logo o processo de mudança linguística na Baixada encontra-se, atualmente, na transição do segundo para o terceiro momento apontado pelos autores acima.

O problema do encaixamento diz respeito à mudança linguística na matriz social da comunidade investigada, ou seja, é possível identificar, hoje, na Baixada Cuiabana, os seguintes processos em andamento: (i) a transformação de um dialeto regional em um dialeto urbano (LABOV, 1972, p.286) comum à comunidade brasileira como um todo; (ii) o espraiamento de traços linguísticos de prestígio na fala dos usuários do dialeto, em função das pressões sociais recebidas de cima e de fora.

O problema da implementação da mudança deve-se ao pouco acesso ao nível superior em Mato Grosso até o início dos anos setenta; à assimilação de novas formas linguísticas pelos alunos que retornavam a Cuiabá; à melhoria e ao incremento das oportunidades de escolarização na região da Grande Cuiabá; ao início da re-colonização (anos 70 do século XX); ao contato com outras variedades do português.

O problema da avaliação, por meio das falas dos colaboradores, ajuda a entender o verdadeiro conteúdo subjacente às avaliações negativas de pessoas de fora que são vistas como melhores e superiores. Por isso, há três diferentes atitudes sociais em relação ao falar cuiabano: (i) uma que se identifica com o conteúdo das

pressões externas e ratifica o estigma; (ii) outra que tem consciência do estigma e se sente pressionada a anular sua identidade local; (iii) uma terceira que procura preservar a identidade cultural, apesar das atitudes contrárias.

Pacheco (2010) também analisa a concordância de gênero na Baixada Cuiabana, mas no interior do sintagma nominal (**a noite inteiro**) e do sintagma predicativo (**a situação** dele era muito **sério**).

Além disso, Pacheco (2010) coletou vários dados de observação participante e fez uma análise separada apenas da ausência de concordância de gênero nos sintagmas predicativos (**A situação** da Amazônia está a cada dia mais **complicado**) e no sujeito posposto (**É muito mínimo essa concordância**) do português brasileiro na língua falada formal e informal e na língua escrita formal, incluindo dados da mídia. Diante desses dados, percebe-se que há sistematicidade da ausência de concordância de gênero nos predicativos e sujeitos pospostos no português brasileiro como um todo.

Nos dados do português europeu, coletados por Lima (2007), também há esse tipo de exemplo (**A cabeça** do dedo fica muito **grosso**), o que corrobora a existência de um encaixamento linguístico muito maior e mais amplo do que apenas na Baixada Cuiabana.

Para a análise variacionista, Pacheco (2010) utilizou 21 entrevistas do *corpus* organizado por Dettoni (2003). De um total de 2.928 dados coletados, apenas 173 dados representam a ausência de concordância de gênero, totalizando 5,9%, ou seja, trata-se de um fenômeno raro e próximo à categoricidade, com 94,1% de uso pleno da concordância de gênero.

Exemplificando o fenômeno analisado, temos, respectivamente, um dado de sintagma nominal e outro de sintagma predicativo:

Uso pleno da concordância de gênero

Uma criação severa, de verdade.

A senhora olha na Bíblia, tem a Bíblia pra ler, **a Bíblia está escrita**.

Uso não pleno da concordância de gênero

Daí, **moçadinhamái novo**.

Ela é da chapada, **nascido** lá né.

Segundo Pacheco (2010), a hipótese geral é que, apesar de a concordância de gênero não ser um fenômeno tão produtivo no português brasileiro, apresenta padrões muito semelhantes aos da concordância de número, nos termos de

Scherre (1988, 1996), principalmente com relação à posição dos elementos na sentença, já que ambas as concordâncias são mais produtivas à esquerda dos sintagmas nominais. Posto isso, podemos estar diante de um padrão mais geral do português brasileiro em que a concordância de gênero e a de número seguem os mesmos padrões linguísticos.

As variáveis selecionadas como estatisticamente relevantes pelo programa VARBRUL foram, em ordem de significância, a configuração sintagmática, a natureza morfológica, a faixa etária e o grau de letramento.

Sobre as variáveis linguísticas, Pacheco (2010) fez um controle rígido de todos os tipos de sintagma (configurações sintagmáticas) que apareceram nas entrevistas, e a hipótese era de que a posição à esquerda do nome favoreceria a concordância, principalmente nos artigos e possessivos.

Através da configuração dos sintagmas, Pacheco (2010) analisou também a posição dos elementos. O resultado foi que o artigo definido *a*, adjetivo e possessivo + nome (**a roupa da minha mãe/ boa fé/ mia madrinha**) são os grandes responsáveis pela concordância de gênero, ambos à esquerda do nome, com 0,92 de peso relativo. O artigo indefinido *uma* e *numeral* + nome (**uma sala de aula/ uma hora**) vêm logo adiante, também favorecendo a concordância à esquerda, com 0,56. Os elementos que mesmo à esquerda do nome desfavoreceram relativamente a concordância foram o sintagma com dois elementos à esquerda (**o primeiro injeção**), com 0,40; o demonstrativo e indefinido (**esse menina e vários substância**), com 0,27; e o quantificador (tudo criança), com 0,05.

Todos os elementos à direita do nome desfavoreceram bruscamente a concordância de gênero, como os sintagmas com elementos à esquerda e à direita do nome (**mia casa todo desarrumado**), com 0,09; o possessivo e o adjetivo (**casa nosso, pessoa cego**), com 0,07; o predicativo (**minha família toda é cuiabano**), com 0,02; e o quantificador (**as vacina tudo**), com 0,01 (SCHERRE, 1988, 1996).

É importante registrar que, como se trata de um fenômeno morfossintático e raro, a pouca quantidade de dados da variante ausência de concordância de gênero pode interferir nos resultados finais. Assim sendo, é necessário polemizar um pouco mais sobre os elementos à esquerda que teoricamente desfavoreceram a concordância de gênero. Sintagmas de três elementos com dois elementos à esquerda (**o primeiro injeção**), os demonstrativos e indefinidos à esquerda do nome (**esse menina/ vários substância**) obtiveram um peso relativo inferior a 0,50. No entanto os dados dos sintagmas de três elementos com elementos à esquerda são muito poucos em comparação com os demais. Essa diferença também deve ser levada em consideração na análise, porque, com um *input* de 0,99, qualquer valor um pouco mais abaixo será desfavorecedor (PACHECO, 2010).

Logo, para se obter resultados mais sintéticos e gerais, os sintagmas foram reunidos em quatro fatores: (i) elementos à esquerda do nome; (ii) elementos à direita do nome; (iii) quantificadores em ambas as posições; e (iv) predicativos em ambas as posições. Assim sendo, pode-se generalizar que apenas os elementos à esquerda favorecem a concordância de gênero (0,69), em detrimento dos elementos à direita (0,16), quantificadores (0,06) e predicativos (0,04). Portanto, a hipótese da posição à esquerda para a concordância de gênero é confirmada, principalmente com o artigo + nome, adjetivo + nome e possessivo + nome.

Para a natureza morfológica, Pacheco (2010) também fez um detalhamento grande, que se polarizou entre nomes variáveis por léxico diferente (**galinha de casa só, criadoço milho**), com peso relativo de 0,88; itens lexicais morfológicamente variáveis (**esse menina**), com 0,71; e itens lexicais invariáveis (**banana frito**), com 0,43. Essa variável mostra como a flexão dos nomes e a saliência fônica, no caso dos nomes variáveis por léxico diferente, favorecem a concordância de gênero, além de os nomes variáveis também terem a presença do traço [+animado +humano], com 0,80.

Nas variáveis sociais, de acordo com Pacheco (2010), a faixa etária foi bastante consistente, em termos do favorecimento da concordância apenas para os falantes de 15 a 25 anos (0,75). Tanto a faixa etária de 30 a 45 anos (0,48) quanto a faixa acima de 60 anos (0,38) desfavorecem o uso da concordância de gênero. Esse é um grande indício de mudança linguística, em que os jovens estão à frente no uso da concordância de gênero, corroborando os resultados de Dettoni (2003).

Outra variável social significativa, em Pacheco (2010), foi o grau de letramento, constituído a partir da escolaridade e dos graus de letramento que os informantes possuíam, através de uma análise qualitativa e quantitativa. A identificação individual de cada informante também foi importante para ver a questão do desempenho linguístico diferente que muitos tinham.

O objetivo de criar uma variável social surgiu do resultado do grau de escolaridade, já que os colaboradores analfabetos (0,53), com cinco a oito anos de estudo (0,59) e mais de nove anos de estudo (0,59) favoreceriam a concordância de gênero. Assim, buscou-se interpretar, qualitativamente por meio das entrevistas, o que realmente estava em jogo diante desses resultados.

As práticas sociais dos colaboradores revelaram que os traços não marcados do letramento e da escolaridade juntamente (0,32), ou seja, ausência de letramento e ausência de escolaridade, é que foram desfavorecedores da concordância de gênero, enquanto os traços [-escolaridade +-letramento], [+escolaridade +letramento], [+escolaridade +letramento] favoreceram a concordância, respectivamente com 0,59, 0,55 e 0,52.

Essa variável não é uma categoria social fixa como as que normalmente são analisadas pela Sociolinguística Variacionista, pois é uma reanálise da escolaridade com base em outros fatores que vão além da sala de aula, como o nível sociocultural, econômico, acesso à mídia, ocupação, participação em eventos sociais, viagens, moradia na zona rural ou urbana etc. Nesse ponto, trata-se de uma visão ampla e mais complexa de vários fatores sociais inseridos no grau de letramento.

Com base nas pesquisas de Dettoni (2003) e Pacheco (2010), é possível analisar o encaixamento linguístico da concordância de gênero na Baixada Cuiabana, fazendo o seguinte paralelo:

- 1) Segundo Dettoni (2003), do ponto de vista do encaixamento linguístico, a mudança na concordância de gênero na Baixada Cuiabana inicia-se pelo contexto Determinante + Nome, espraiando-se, em seguida, para outras estruturas; as construções Nome + Adjetivo e Sujeito + Predicativo apresentam-se como os contextos mais resistentes à mudança;
- 2) Pacheco (2010) discute a possibilidade de a mudança – presença categórica de concordância – não se concretizar nos sintagmas predicativos ou, pelo menos, demorar mais a chegar, uma vez que a ausência da concordância nesses casos já acontece de forma mais geral e não se limita à área geográfica de Cuiabá, ocorrendo em todo o território brasileiro, tanto na fala como na escrita. Juntamente com os predicativos, os quantitativos também são os grandes obstáculos para a concordância de gênero, uma vez que o peso relativo desses fatores fica abaixo de 0,06.
- 3) De acordo com Pacheco (2010), o peso relativo baixo dos predicativos indica que eles têm menos probabilidade de ter concordância de gênero do que os SNs. Logo os SNs estão na direcionalidade contrária da mudança, já que a probabilidade da presença da concordância de gênero é maior, principalmente nos artigos, possessivos, adjetivos e numerais à esquerda do nome. Sendo assim, percebe-se que os predicativos têm um encaixamento linguístico maior que o encaixamento social dentro do sistema linguístico da língua portuguesa.
- 4) Na pesquisa de Dettoni (2003), o contexto mais favorável à retomada pelo anafórico *ele* é a referência genérica, ausência de determinante, ausência de flexão de gênero, faixa etária mais idosa e escolaridade baixa. Na pesquisa de Pacheco (2010), o contexto mais favorável à ausência de concordância de gênero é a posição à direita do nome, predicativos e quantificadores, ausência de flexão de gênero, faixa etária mais idosa e grau de letramento menor.
- 5) Diante do exposto, a concordância de gênero nos sintagmas nominais pode estar em processo de mudança linguística, no sentido de um uso cada vez mais restrito da ausência de concordância nos SNs, já que o *input* para a realização plena de concordância é de 0,99, um resultado próximo à categoricidade, e um uso um pouco menor da concordância de gênero nos predicativos isolados, com um *input* de 0,74 (PACHECO, 2010).

- 6) Em suma, as duas análises corroboram a idéia de que a variação na concordância de gênero na Baixada Cuiabana segue os padrões maiores de variação de concordância de número no português brasileiro, o que tem sido atestado pelos trabalhos de Scherre para a concordância de número. Então, o padrão linguístico é o mesmo, ou seja, formas à esquerda concordam mais do que formas à direita do núcleo. Além disso, Pacheco (2010) mostra que a ausência de concordância de gênero nos sintagmas predicativos, independentemente de estarem na ordem direta ou indireta, é generalizada no português brasileiro.

Conclusão

O Projeto Valco atingiu, até o momento, os objetivos inicialmente apresentados, uma vez que traz em seus arquivos registros de língua em uso real, com seus respectivos áudios e transcrições, adicionados às reflexões já realizadas sobre os dados. Além disso, veicula alguns artigos e trabalhos científicos sobre o Centro-Oeste.

As pesquisas aqui apresentadas formam o escopo de estudos amplamente analisados e divulgados no site do Projeto Valco. Tal projeto, porém, ainda precisa ser plenamente difundido no meio acadêmico e alimentado com mais entrevistas, a fim de se fortalecer e se estabelecer como importante veículo difusor de conhecimentos sobre a língua em uso no Centro-Oeste brasileiro.

Como afirmado anteriormente, acreditamos que esse projeto tem muito a contribuir para o estudo do português em diversas áreas linguísticas, bem como para historiografia do português em terras brasileiras, entre tantos outros estudos na área de Letras; e, por fim, acreditamos que esse projeto seja um meio de desenvolver e difundir resultados dos estudos sociolinguísticos realizados nessa região do país.

DETTONI, R. do V.; PACHECO, C. da S.; ANDRADE, C. O.; SCHERRE, M. M. P. Linguistic variation project in the midwest region of Brazil – Valco. *Alfa*, São Paulo, v.56, n.3, p.797-823, 2012.

ABSTRACT: The aim of our article is to present Valco Project (Linguistic Variation of Brazil's Midwest Region) and summarize the results of some researches in sociolinguistics from different geographical areas (in the Midwest) and different themes. Four researches were developed with data from Brasilia and they analyze the second person pronouns: Tu/Você/Ocê/Cê. In these studies, we present a central idea that the variety from Brasilia is currently introducing the new variant TU in its linguistic repertoire, focusing on a trend in Brazilian Portuguese. Two other researches were developed with data from the Baixada Cuiabana region, in Mato Grosso, and they analyzed the gender variation present on the discourse and syntagmatic plans. Unlike

the previous researches mentioned here, those two studies present linguistic features of semi categorical nature, which are probably in an extinction situation. But it is evident that gender variance in the analyzed data is linguistically embedded in the system of Portuguese language. Thus, the present article, which is based on Language Variation and Change Theory, brings to the scientific community a wider knowledge of Portuguese language in Brazilian lands, more specifically, in traditional Central Brazil and in expanded Central Brazil.

- **KEYWORDS:** *Linguistics Variation. Midwest region. Linguistic variety from Brasília. Linguistic variety from Cuiabá. Second person pronouns. Gender agreement. Dialectfocusing.*

REFERÊNCIAS

ANDRADE, A. L. V. S. *A variação você, cé e ocê no português brasileiro falado.* 145f. 2004. Dissertação (Mestrado em Linguística) – Universidade de Brasília, Brasília, 2004.

ANDRADE, C. Q. *Tu e mais quantos? A segunda pessoa na fala brasiliense.* 132f. 2010. Dissertação (Mestrado em Linguística) – Universidade de Brasília, Brasília, 2010.

BORTONI-RICARDO, S. M. *The urbanization of rural dialects speakers: a sociolinguistic study in Brazil.* Cambridge: CUP, 1985.

CORRÊA, C. da C. *Focalização dialetal em Brasília: um estudo das vogais pré-tonicas e do /s/ pós-vocálico.* 98f. 1998. Dissertação (Mestrado em Linguística) – Universidade de Brasília, Brasília, 1998.

DETTONI, R. do V. *A concordância de gênero na anáfora pronominal: variação e mudança linguística no dialeto da baixada cuiabana – Mato Grosso.* 256f. 2003. Tese (Doutorado em Estudos Linguísticos) - Universidade Federal de Minas Gerais, Minas Gerais, 2003.

DIAS, M. C. A. *A variação na concordância nominal: um contraste entre o urbano e o rural na fala brasiliense.* 174f. 1993. Dissertação (Mestrado em Linguística) – Universidade de Brasília, Brasília, 1993.

DIAS, E. P. *O uso do tu no português brasiliense falado.* 114f. 2007. Dissertação (Mestrado em Linguística) – Universidade de Brasília, Brasília, 2007.

HANNA, E. S. *Difusão e focalização dialetal: o caso de Brasília.* 184f. 1986. Dissertação (Mestrado em Linguística) – Universidade de Brasília, Brasília, 1986.

LABOV, W. *Padrões sociolinguísticos.* Tradução Marcos Bagno, Maria Marta Pereira Scherre, Caroline Rodrigues Cardoso. São Paulo: Parábola Editorial, 2008.

- _____. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press, 1972.
- LIMA, J. L. *A variação na concordância do gênero gramatical no falar cuiabano*. 248f. 2007. Tese (Doutorado em Linguística) – Universidade de Campinas, São Paulo, 2007.
- LUCCA, N. N. G. *A variação tu/você na fala brasiliense*. 139f. 2005. Dissertação (Mestrado em Linguística) - Universidade de Brasília, Brasília, 2005.
- MALVAR, E. da S. *A realização do objeto direto em cadeia anafórica no português do Brasil*. 153f. 1992. Dissertação (Mestrado em Linguística) – Universidade de Brasília, Brasília, 1992.
- PACHECO, C. da S. *Padrões sociolinguísticos da concordância de gênero da baixada cuiabana* 162f. 2010. Dissertação (Mestrado em Linguística) – Universidade de Brasília, Brasília, 2010.
- PINTZUK, S. *Programas VARBRUL*. Tradução de Ivone Isidoro Pinto. Rio de Janeiro: UFRJ, 1988.
- SANKOFF, David; TAGLIAMONTE, Sali; SMITH, Eric. Goldvarb X: *A variable rule application for Macintosh and Windows*. Department of Linguistics, University of Toronto. 2005.
- SCHERRE, M. M. P. Sobre a influência de três variáveis relacionadas na concordância nominal em português. In: SILVA, G. M. de O.; SCHERRE, M. M. P. (Org.). *Padrões sociolinguísticos - análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro*. Rio de Janeiro: Tempo Brasileiro: Ed. da UFRJ, 1996. p.41-62.
- _____. *Reanálise da concordância nominal em português*. 555f. 1988. Tese (Doutorado em Linguística) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1988.
- WEINREICH, U.; LABOV, W.; HERZOG, M. Empirical foundations for a theory in language change. In: LEHMANN, W. P.; MAKIEL, Y. (Ed.). *Directions for historical linguistics*. Austin: University of Texas Press, 1968. p.95-195.

Anexo I

Normas de Transcrição do Projeto VALCO

Cada pesquisador, ao se interessar pelo banco de dados do VALCO, deverá adaptar suas entrevistas às normas de transcrição elaboradas para o projeto VALCO, no intuito de padronizar as entrevistas e facilitar a consulta por outros interessados.

Tipos propostos de transcrição:

- a) **Fonográfica** – transcrição da fala, respeitando e reproduzindo, na medida do possível, a oralidade;
- b) **Convencional** – adaptação da ortografia aos objetivos pretendidos.

Orientações gerais:

- a) A transcrição não pode ser sobrecarregada de símbolos e deve ser limpa e legível.
- b) A transcrição deve permitir a compreensão do significado do texto.
- c) A produção oral (tá, tô, tamos, pra, eles foi, os menino, memo) deve ser respeitada, sem a necessidade de uma transcrição fonética.
- d) Os falantes devem ser indicados apenas com a sigla do nome (as três primeiras letras).
- e) Deve-se procurar facilitar ao leitor a criação de uma “imagem” do texto elaborado no plano da oralidade.

Normas para transcrição de dados:

- 1) Letra Times New Roman, 12, espaçamento 1.5.
- 2) Os participantes devem ser numerados, de acordo com as fichas catalográficas de cada interação.

Ex.:

Participante: *MAR.*

Escolaridade: *analfabeta.*

Ocupação: *lavradora, dona de casa e merendeira da escola.*

Lugar de origem: *Córrego Puba (Vendinha), em Pombal.*

Idade: *50 anos.*

Data: *27/07/1997.*

Tipo de entrevista: entrevista tipicamente laboviana.

Lugar da entrevista: Cotidiano familiar.

Duração: 40 minutos

Temas abordados: Emancipação do município de Santa Rita do Novo Destino.

- 3) Ao transcrever os dados, coloque em negrito os enunciados do pesquisador e sem destaque os do pesquisado.

Ex.: **E - Como foi isso?** (E, de entrevistador)

P - Foi trapaído, foi. (P de participante);

- 4) Ênfase - usar letras maiúsculas.

Ex.: Mais eu QUERO brincar.

- 5) Pontuação – o ponto de interrogação, nas frases interrogativas, e o de exclamação, nas exclamativas, são mantidos.

- 6) Dúvida quanto à palavra – a palavra sob dúvida deve ficar entre colchetes angulares.

Ex.: Ele <andava> muito.

- 7) Fala sobreposta (Cruzamento de vozes) – os enunciados pronunciados por dois falantes ao mesmo tempo devem estar conectados por um colchete.

Ex.: Que legal!!

Que legal!!

- 8) Alongamento de vogal – após a vogal alongada são usados quatro pontos.

Ex.: Ele gostava, e co::mo gostava!

- 9) Silabação – colocar hífen no meio da palavra.

Ex.: ca-fé, ca-mi-nha-da.

10) Repetições – reduplicar palavras ou sílabas repetidas.

Ex.: Aí, e ele foi pra casa de de Carlos.

11) Comentários do pesquisador acerca de atitudes não linguísticas do informante devem ser colocados entre parênteses.

Ex.: Ele gosta de mim, (risos)

(baixou o tom de voz)

12) Comentário de outro participante – colocar o símbolo # (3ª pessoa falando).

Ex.: # (3ª pessoa falando)_Ele, ele num cunhece. {inint} é melhó. {inint} má lá embaxo daqui lá.

13) Palavra ou trecho ininteligível – o comentário deve estar entre chaves.

Ex.: Maria queria comprar {inint}, a mãe dela falou que não queria.

14) Pausa pequena – usar vírgula.

15) Pausa média – usar ponto.

16) Reticências: suspensão do pensamento e truncamento de palavras.

Ex.: Eu gostaria de ir para...sei lá, acho melhor morar em cas...apartamento.

Referências das Normas de Transcrição do VALCO

- Projeto Vozes do Amapá.
- Núcleo de estudos da história linguística de Goiás. Normas para coleta e transcrição de dados. Faculdade de Letras, UFG.
- MARCUSCHI, L. A. **Análise da conversação**. Editora Ática, 2000.

Recebido em março de 2012.

Aprovado em julho de 2012.