

UNIVERSIDADE FEDERAL DE MINAS GERAIS – UFMG
FACULDADE DE LETRAS – FALE
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGÜÍSTICOS – POSLIN

Luiz Egito de Souza Barros

A DINÂMICA VARIÁVEL *TU/VOCÊ* EM CARTAS PIAUIENSES
NOVECENTISTAS: um percurso histórico.

Orientadora: Profa. Dra. Márcia Cristina de Brito Rumeu

Belo Horizonte

2024

Luiz Egito de Souza Barros

**A DINÂMICA VARIÁVEL *TU/VOCÊ* EM CARTAS PIAUIENSES
NOVECENTISTAS: um percurso histórico.**

Tese de Doutorado apresentada ao Programa
de Pós-Graduação em Estudos Linguísticos
da Faculdade de Letras da Universidade
Federal de Minas Gerais da UFMG.

Área de concentração: Linguística Teórica e Descritiva.

Linha de pesquisa: (1A) Estudo da Variação e Mudança Linguística.

Orientadora: Profa. Dra. Márcia Cristina de Brito Rumeu.

Belo Horizonte

2024

B277d

Barros, Luiz Egito de Souza.

A dinâmica variável *tu/você* em cartas piauienses novecentistas [manuscrito]: um percurso histórico / Luiz Egito de Souza Barros. – 2024. 1 recurso online (174 f.: il., tabs., grafs., maps., p&b., color.): pdf.

Orientadora: Márcia Cristina de Brito Rumeu.

Área de concentração: Linguística Teórica e Descritiva.

Linha de pesquisa: Estudo da Variação e Mudança Linguística.

Tese (doutorado) – Universidade Federal de Minas

Gerais, Faculdade de Letras.

Bibliografia: f. 133-139.

Apêndices: f. 140-174.

Exigências do sistema: Adobe Acrobat Reader.

1. Língua portuguesa – Pronomes – Teses. 2. Sociolinguística – Teses. 3. Mudanças linguísticas – Teses. 4. Língua portuguesa – Variação – Piauí – Teses. 5. Linguística histórica – Teses. I. Rumeu, Márcia Cristina de Brito. II. Universidade Federal de Minas Gerais. Faculdade de Letras. III. Título.

CDD: 469.798



UNIVERSIDADE FEDERAL DE MINAS GERAIS

FACULDADE DE LETRAS

PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGUÍSTICOS

FOLHA DE APROVAÇÃO

A DINÂMICA VARIÁVEL TU/VOCÊ EM CARTAS PIAUIENSES

NOVECENTISTAS: um percurso histórico.

LUIZ EGITO DE SOUZA BARROS

Tese submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ESTUDOS LINGUÍSTICOS, como requisito para obtenção do grau de Doutor em ESTUDOS LINGUÍSTICOS, área de concentração LINGUÍSTICA TEÓRICA E DESCRITIVA, linha de pesquisa Estudo da Variação e Mudança Linguística.

Aprovada em 28 de fevereiro de 2024, pela banca constituída pelos membros:

Profa. Marcia Cristina de Brito Rumeu - Orientadora

UFMG

Profa. Célia Regina dos Santos Lopes

UFRJ

Profa. Maria do Socorro Vieira Coelho

Unimontes

Profa. Juliana Costa Moreira

CEFET-MG

Prof. Lorenzo Teixeira Vitral

UFMG

Belo Horizonte, 28 de fevereiro de 2024.



Documento assinado eletronicamente por **Marcia Cristina de Brito Rumeu, Professora do Magistério Superior**, em 29/02/2024, às 15:30, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Maria do Socorro Vieira Coelho, Usuário Externo**, em 29/02/2024, às 16:17, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Lorenzo Teixeira Vitral, Professor do Magistério Superior**, em 29/02/2024, às 17:48, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Juliana Costa Moreira, Usuária Externa**, em 01/03/2024, às 09:42, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Célia Regina dos Santos Lopes, Usuário Externo**, em 01/03/2024, às 16:25, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2912105** e o código CRC **EC2953F8**.

Referência: Processo nº 23072.278071/2023-72

SEI nº 2912105

AGRADECIMENTOS

Ajudar o outro é um gesto de bondade e de generosidade, aceitar ajuda é um gesto de humildade, agradecer a ajuda recebida é reconhecer a grandeza e a generosidade do outro. Por esse caminho, agradeço primeiramente a Deus, por tudo que me tem proporcionado durante toda a minha existência. Um agradecimento todo especial à minha orientadora, Professora Márcia Cristina de Brito Rumeu, que fez tudo para que esta Tese chegasse à sua conclusão, restando a mim agradecer pelo acolhimento e por tamanha generosidade e dedicação. Gratidão também à professora Thaís Cristófaró Silva, que soube entender o momento difícil e sugerir minha mudança de área no momento certo, dadas as adversidades da pandemia. Agradeço também aos demais professores e colegas do POSLIN, pelos conhecimentos compartilhados e pelo companheirismo posto em prática.

Sem as fontes históricas nada teria sido feito. Por isso, agradeço aos proprietários dos acervos, que, agindo como historiadores, tiveram o cuidado de preservar os documentos escritos e a bondade de gentilmente nos cederem tão preciosas fontes documentais. Os colegas da UFPI/ curso de História, Johny Santana, Francisco Gleyson, Olivia Candeia, Audrey Tapety, Mairton Celestino, e do curso de Pedagogia, Maria Alveni, que agiram como verdadeiros guias na busca por fontes históricas, também são dignos de agradecimentos. Devo gratidão também ao professor Paulo César L. Sales, que, com paciência e companheirismo, compartilhou comigo os seus conhecimentos a respeito do RBrul. Agradeço também ao professor Manoel de Jesus Barros Martins (UFMA), que permitiu acesso às suas aulas na disciplina de Paleografia, que muito nos ajudaram no processo de transcrição e edição das cartas piauienses.

Meus sinceros agradecimentos também à UFMG/POSLIN, pelo acolhimento e pelos saberes a mim disponibilizados, à UFPI e à Direção do Campus Senador Helvídio Nunes de Barros, na pessoa do professor Juscelino Francisco do Nascimento, à Coordenação e aos colegas do Curso de Letras, que me apoiaram, sempre que se fez necessário.

Meu muito obrigado também a minha família, a começar pela memória dos meus pais, Domingos de Souza Araújo e Raimunda Borges Leal, e de meus irmãos já idos, José Barros de Araújo e Sebastião Barros Araújo, pelo exemplo, a José Alberto e a Maria Neusa, pelo apoio nas idas e vindas de Belo Horizonte; a Manoel, a Maria dos Anjos, a Maria da Penha e aos sobrinhos e sobrinhas, pelo cuidado pessoal. Agradeço imensamente a minha esposa Cristina, aos meus filhos Esaú, Efigênia e Isaac, pela

compreensão e apoio. Por tudo isso, meus sinceros agradecimentos a todos que deram sua parcela de contribuição.

BARROS, Luiz Egito de Souza. *A DINÂMICA VARIÁVEL TU/VOCÊ EM CARTAS PIAUIENSES NOVECENTISTAS: um percurso histórico*. 2024. 174 fls. Tese (Doutorado em Estudos Linguísticos) – Universidade Federal de Minas Gerais, Belo Horizonte, 2024.

RESUMO

Nesta Tese de Doutorado, a proposta principal é o estudo, à luz dos princípios da sociolinguística histórica (HERNÁNDEZ-CAMPOY & SCHILLING, 2012), da variação entre as formas de referência ao sujeito de 2ª pessoa do singular (*tu/você*) com base na produção escrita de piauienses ilustres do século XX, mais especificamente entre as décadas de 20 e 90 (1920 e 1990). Considerando que o foco esteja voltado a apresentar uma descrição-analítica da alternância *tu/você*, à luz da Teoria do Poder e da Solidariedade (BROWN & GILMAN, 1960), em cartas produzidas, no decorrer do século XX (1920-1990), por piauienses, tendo em vista as relações sociais que as embasam, entendemos que o inovadorismo do tema se deixa evidenciar também em função de não ter sido tratado em relação aos redatores do Piauí em sincronias passadas. Acrescentemos ainda o fato de que detectamos evidências, fundamentadas no Atlas Linguístico do Brasil (ALiB), que nos revelam a baixa produtividade do *tu* (em tão somente 13% dos dados), cf. discutido por Lopes *et alii* (2018, p. 118), orientando o ponto de partida desta análise que é o de voltar-se ao “presente para explicar o passado” (TRUDGILL & HERNÁNDEZ CAMPOY 2007, p. 299). A hipótese principal é a de que o *você* já se deixaria entrever como mais produtivo do que o *tu* na era novecentista do PB como já observado por Scherre *et alii* (2021) à luz de Deus (2009) e Cardoso (2008) para os dados do ALiB. De um modo geral, constatamos a prevalência do *você* (60,03%), ainda que em intensa alternância entre com o *tu* (39,97%) nas cartas piauienses. No que diz respeito aos contextos das relações sociais que subsidiam as formas *tu* e *você*, constatamos a produtividade do *você* não só nas relações simétricas, mas também nas assimétricas descendentes, evidencia a semântica da Solidariedade na realidade do português brasileiro novecentista. Ainda em relação ao *você*, constatamos que a sua produtividade é favorecida nos contextos da *terceira pessoa verbal*, do *você-sujeito* e da produção escrita de *redatores idosos e adultos*, acompanhando os resultados de Souza (2021) e de Souza (2012) conduzidos também pela análise de amostras históricas. Assim sendo, acreditamos que esta Tese venha a preencher uma lacuna em relação aos estudos sociolinguísticos acerca da dinâmica *tu/você* em terras piauienses de sincronias passadas.

Palavras-chave: Variação *tu/você*, Relações sociais, Sociolinguística Histórica.

BARROS, Luiz Egito de Souza. *THE DYNAMICS OF TU/VOCÊ IN LETTERS OF 19TH CENTURY FROM PIAUÍ: a historical trajectory*. 2024. 174 fls. Thesis (Ph.D. in Linguistics studies) – Universidade Federal de Minas Gerais, Belo Horizonte, 2024.

ABSTRACT

In this Doctoral Thesis, the main proposal is the study, in light of the principles of historical sociolinguistics (HERNÁNDEZ-CAMPOY & SCHILLING, 2012), of the variation between forms of reference to the 2nd person singular subject (*tu (you)/você (you)*) based in the written production of illustrious people from Piauí in the 20th century, more specifically between the 1920s and 1990s (1920s and 1990s). Considering that the focus is on presenting an analytical description of the *tu/você (you)* alternation, in the light of the Theory of Power and Solidarity (BROWN & GILMAN, 1960), in letters produced, during the 20th century (1920-1990), by Piauí, considering the social relations that underpin them, we understand that the innovativeness of the theme is also evident due to the fact that it has not been addressed in relation to Piauí editors in past synchronies. Let us also add the fact that we detected evidence, based on the Linguistic Atlas of Brazil (ALiB), which reveals the low productivity of *tu* (in only 13% of the data), cf. discussed by Lopes *et alii* (2018, p. 118), guiding the starting point of this analysis, which is to turn to the “present to explain the past” (TRUDGILL & HERNÁNDEZ CAMPOY 2007, p. 299). The main hypothesis is that *você (you)* would already be seen as more productive than *tu (you)* in the 19th century era of PB, as already observed by Scherre *et alii* (2021) in the light of Deus (2009) and Cardoso (2008) for the data from ALiB. In general, we found the prevalence of *você (you)* (60.03%), although in intense alternation between *tu (you)* (39.97%) in letters from Piauí. With regard to the contexts of social relations that subsidize the forms *tu* and *você (you)*, we see the productivity of *você (you)* not only in symmetrical relationships, but also in descending asymmetrical ones, highlighting the semantics of Solidarity in the reality of nineteenth-century Brazilian Portuguese. Still in relation to *você (you)*, we found that your productivity is favored in the contexts of the third person verbal, *você*-subject and the written production of elderly and adult writers, following the results of Souza (2021) and Souza (2012) also conducted by analyzing historical samples. Therefore, we believe that this Thesis will fill a gap in relation to sociolinguistic studies on the *tu/você (you)* dynamics in Piauí lands of past synchronies.

Keywords: variation *tu/você*, social relations, historical sociolinguistics.

LISTA DE ABREVIATURAS E SIGLAS

FNT = Forma Nominal de Tratamento

PB = Português Brasileiro

2PL = 2ª pessoa do plural

2SG = 2ª pessoa do singular

SH = Sociolinguística Histórica

SV = Sociolinguística Variacionista

LISTAS DE GRÁFICOS, IMAGENS, QUADROS E TABELAS

Lista de Gráficos

Gráfico 1	As formas <i>vossa mercê, você, tu, senhor(a)</i> distribuídas pela produção escrita de brasileiros do nordeste e do sudeste entre os séculos XIX e XX	28
Gráfico 2	As formas <i>tu</i> e <i>você</i> correlacionadas às dinâmicas sociais de simetria e assimetria social descendente nas cartas piauienses.	109
Gráfico 3	Distribuição das formas <i>tu</i> e <i>você</i> pelas relações sociais piauienses (simétricas e assimétricas descendentes).	112

Lista de Figuras

Figura 1	Mapa da distribuição dos pronomes pessoais de 2SG <i>você, ocê, cê</i> e <i>tu</i> (com e sem concordância) por 9 capitais nordestinas, dentre elas o Piauí, Scherre <i>et alii</i> (2021, p. 175)	33
Figura 2	Mappa Geographico da Capitania Piauhy, e parte das do Maranhão, e do Gran Pará	80
Figura 3:	Mapa do estado do Piauí pelo IBGE.	81

Lista de Imagens

Imagem 1	Evidência de aspectos ortográficos e de hipercorreção. JMF. Tatu [Oeiras-PI], 19.08.1920	38
Imagem 2	Ficha de apresentação das cartas piauienses	39
Imagem 3	Carta de amizade piauiense. ABSB. RJ, 06.06.1946	39
Imagem 4	Excerto da carta de CRJM ao noivo. JMF. S.1, 14.11.1921.	41
Imagem 5	Excerto da carta do noivo à noiva. JMF. Simplicio Mendes - PI, 16.09.1920.	41
Imagem 6	Excerto da carta de AVV. Brasília, 23.09.1981.	41
Imagem 7	Excerto da carta de AVV. Brasília, 30.10.1984.	42

Imagem 8	Excerto da carta de PNQ. Oeiras, 20.04.1985.	42
Imagem 9	Excerto da carta de PNQ. Oeiras, 10.05.1985.	42
Imagem 10	Excerto da carta de ABSB. Piauí-PI, 31.01.1981.	43
Imagem 11	Fechamento da carta de ABSB. Piauí-PI, 31.01.1981.	43
Imagem 12	Excerto da carta de VVS. São Caetano do Sul, 30.07.1953.	47
Imagem 13	Excerto da carta de ABSB. Terezina, 05.03.1976.	48
Imagem 14	Excerto da carta de JMF. Simplício Mendes, 06.09.1920.	49
Imagem 15	Evidência de marca de oralidade em carta <i>familiar</i> endereçada a uma de suas netas. PNQ, Oeiras-PI, 22.05.1986.	56
Imagem 16	Excerto da carta de AVV. Brasília, 25.02.1983.	57
Imagem 17	Excerto da carta de LJL. Bocaina-PI, 22.06.1951.	57
Imagem 18	Excerto da carta de VVS. São Caetano do Sul, 25.04.1954.	63
Imagem 19	Excerto da carta de JPB. São Bernardo do Campo, 18.04.1978.	64
Imagem 20	Carta de JATF. Teresina-PI, 11.01.1972.	65
Imagem 21	Interface do RStudio – R versão 4.3.2 (2023-12-29).	83
Imagem 22	Excerto da carta de MNRSR. 10.05.1978	85
Imagem 23	Excerto da carta de JATF. Teresina-PI, 05.03.1976	105
Imagem 24	Excerto da carta de JPB. São Bernardo do Campo, SP, 05.06.1978.	106
Imagem 25	Excerto da carta de ABSB. RJ, 03.09.1984.	107
Imagem 26	Carta de PNQ. Oeiras, 20.04.1985.	110

Lista de Quadros

Quadro 1	Síntese dos pronomes pessoais na perspectiva da prescrição gramatical	26
Quadro 2	Amostras históricas piauienses: o século XX em cena.	50
Quadro 3	Quadro-síntese da correlação entre os níveis de escolarização e os perfis sociais dos redatores	62
Quadro 4	Relação das cartas do Acervo “Autores Piauienses”.	70
Quadro 5	Relação das cartas do Acervo “Batista”	71
Quadro 6	Relação das cartas do Acervo de “Cartas Pessoais Variadas”	72
Quadro 7	Relação das cartas do Acervo “Moura Fé”	73
Quadro 8	Relação das cartas do Acervo “Queiroz”	74
Quadro 9	Relação das cartas do Acervo “Santos Rego”	75
Quadro 10	Relação das cartas do acervo “Vieira do Vale”	76

Lista de Tabelas

Tabela 1	Média da distribuição de pronomes explícitos na segunda pessoa do singular no português brasileiro falado nas capital do Piauí (PI), região Nordeste: médias de amostras diversificadas, cf. Scherre <i>et alii</i> (2021, p. 176).	32
Tabela 2	As formas de referência ao sujeito de 2SG nas cartas piauienses (século XX)	88
Tabela 3	A expressão dos pronomes-sujeito de 2SG: nulo ou pleno.	89
Tabela 4	As formas de referência ao sujeito de 2SG e a pessoa verbal (concordância)	90
Tabela 5	As formas de referência ao sujeito de 2SG e o paralelismo formal e semântico	92
Tabela 6	As formas de referência ao sujeito de 2SG correlacionadas ao sexo e à faixa etária dos escreventes	99
Tabela 7	As formas <i>tu</i> e <i>você</i> em relação ao gênero, à faixa etária e ao nível de escolaridade dos redatores	101
Tabela 8	As formas <i>tu</i> e <i>você</i> em relação às práticas de escrita dos redatores	104
Tabela 9	As formas de referência ao sujeito de 2SG correlacionadas às relações sociais simétricas e assimétricas (descendentes)	108
Tabela 10	Distribuição das formas de referência ao sujeito de 2SG (<i>tu</i> e <i>você</i>) pelas cartas piauienses no eixo do tempo: de 1920 a 1990.	111
Tabela 11	As formas de referência ao sujeito de 2SG (<i>tu</i> e <i>você</i>) distribuídas pelas cartas amorosas, de amizade e familiares	115
Tabela 12	As formas de referência ao sujeito de 2SG (<i>tu</i> e <i>você</i>) distribuídas pelas temáticas das cartas pessoais piauienses	116
Tabela 13	Distribuição das formas <i>tu</i> e <i>você</i> pelos acervos piauienses e pelos subgêneros das cartas pessoais	118
Tabela 14	O efeito da variável <i>Pessoa</i> sobre o <i>você-sujeito</i>	123
Tabela 15	O efeito da variável <i>Paralelismo formal e semântico</i> sobre o <i>você-sujeito</i>	124
Tabela 16	O efeito da variável <i>Expressão do sujeito</i> sobre o <i>você-sujeito</i>	125
Tabela 17	O efeito da variável <i>faixa etária</i> sobre o <i>você-sujeito</i>	126
Tabela 18	O efeito da variável <i>Escolarização</i> sobre o <i>você-sujeito</i>	127
Tabela 19	Síntese das variáveis independentes estatisticamente relevantes à aplicação do <i>você-sujeito</i>	129

SUMÁRIO

CONSIDERAÇÕES INICIAIS	16
------------------------------	----

CAPÍTULO 1. REVISÃO DO TEMA: UMA BREVE RETOMADA DA ALTERNÂNCIA *TU/VOCÊ*.

1.1 Um pouco da história: o processo de inserção do <i>você</i> no sistema pronominal do português	20
1.2 Os pronomes de referência ao sujeito de 2ª pessoa do singular na perspectiva da tradição gramatical	25
1.3 O <i>Você</i> no Português Brasileiro de sincronias passadas e recentes	27
1.4. O <i>Você</i> no Português Brasileiro de sincronias recentes: algumas evidências na fala piauiense.	32

CAPÍTULO 2. FUNDAMENTAÇÃO TEÓRICO-METODOLÓGICA

2.1 Alguns princípios teóricos da Sociolinguística Histórica aplicados às cartas piauienses	34
2.1.1 A autenticidade das cartas piauienses	37
2.1.2 A autoria das cartas piauienses	40
2.1.3 A validade social e histórica das cartas piauienses	44
2.2 A Teoria do Poder e da Solidariedade	44
2.3 O trabalho com o gênero <i>carta pessoal</i> e os seus subtipos: amostras linguísticas de sincronias passadas	46
2.4 Uma breve exposição das amostras de cartas pessoais piauienses novecentistas: quadro geral	49
2.5 A referência ao sujeito de 2SG: os parâmetros de uma análise linguístico-social	52
2.5.1 As formas de referência ao sujeito de 2SG: <i>tu/você</i>	53
2.5.2 A expressão do pronome-sujeito de 2SG: nulo/pleno.	53

2.5.3 A pessoa verbal (concordância)	54
2.5.4 O paralelismo formal e semântico	54
2.5.5 Os subgêneros das cartas pessoais: amorosa, amizade e familiar	55
2.5.6 As temáticas prevalentes das cartas pessoais piauienses	58
2.5.7 O sexo e a faixa etária dos redatores	60
2.5.8 Os níveis de escolarização dos redatores piauienses	61
2.5.9 Os níveis de prática de escrita	63
2.5.10 As relações sociais entre os piauienses: a simetria e a assimetria social. ...	66
2.5.11 O fator <i>tempo</i> : os períodos das cartas.	67
2.5.12 Os conjuntos documentais de cartas pessoais: descrição pormenorizada das fontes históricas piauienses.	69
2.5.12.1 Acervo “Autores Piauienses”	69
2.5.12.2 Acervo “Batista”	70
2.5.12.3 Acervo “Cartas Pessoais Variadas”	71
2.5.12.4 Acervo “Moura Fé”	72
2.5.12.5 Acervo “Queiroz”	73
2.5.12.6 Acervo “Santos Rêgo”	74
2.5.12.7 Acervo “Vieira do Vale”	75
2.6 Piauí: um pouco da sua História	77
2.7 Metodologia: a submissão dos dados históricos ao RBrul.	82

CAPÍTULO 3. A VARIAÇÃO TU/VOCÊ: DESCRIÇÃO-ANALÍTICA DOS RESULTADOS ESTATÍSTICOS	85
3.1 Os fatores linguísticos: a expressão do sujeito, a pessoa verbal e o paralelismo formal e semântico.	88
3.2 Os fatores extralinguísticos: o sexo, a faixa etária, a escolaridade, os níveis de prática de escrita, as relações sociais dos informantes, os subgêneros das “cartas pessoais”, os assuntos prevalentes e o tempo.	98
CAPÍTULO 4. A VARIAÇÃO TU/VOCÊ: DESCRIÇÃO-ANALÍTICA DOS RESULTADOS PROBABILÍSTICOS	121
CONSIDERAÇÕES FINAIS	129
REFERÊNCIAS BIBLIOGRÁFICAS	133 - 139
APÊNDICE (Rodada nos níveis <i>stepping up</i> e <i>stepping down</i>)	140 - 174

CONSIDERAÇÕES INICIAIS

Ao lermos textos de sincronias passadas, já é possível nos depararmos com formas de referência ao sujeito pronominal de 2ª pessoa que atualmente também estão em variação no português brasileiro (doravante PB). Podemos destacar a forma *tu* (para a referência à 2ª pessoa do singular (doravante 2SG)) que, por sua vez, vem, no decorrer do tempo, cedendo espaço funcional à forma pronominal *você* (LOPES *et alii*, 2018). A reorganização do sistema pronominal do PB especificamente em relação à 2SG vem despertando o interesse de pesquisadores no âmbito da sociolinguística e da sociolinguística histórica (doravante SH). Nesse sentido, temos as análises de Rumeu (2011), Souza (2012), Lopes & Rumeu (2015), Batista (2017), Gomes & Lopes (2018), Moura & Martins (2018), Souza (2021) como alguns dos trabalhos que se voltam à alternância *tu/você* na posição de sujeito em distintas e diversificadas amostras históricas (cartas pessoais) de sincronias passadas (séculos XIX e XX).

Assim sendo, apresentamos como o intuito principal desta Tese o estudo da variação entre as formas pronominais *tu* e *você* com base na produção escrita (cartas pessoais) de redatores nascidos no espaço geográfico do Piauí, no decorrer das décadas de 20 e 90 do século XX, à luz dos parâmetros de análise e de composição das amostras no âmbito da sociolinguística histórica (HERNÁNDEZ CAMPOY & CONDE-SILVESTRE, 2012) e da Teoria do Poder e da Solidariedade (BROWN & GILMAN, 1960). De (a) a (e), apresentamos algumas evidências do fenômeno variável em análise nas amostras de cartas piauienses. Convém esclarecermos que, em itálico, estão, no decorrer de todo este texto, as formas pronominais *tu* e *você* (nulas ou plenas). Acrescentemos ainda o fato de que os nomes de todos os redatores (autores) das cartas pessoais piauienses, que são, na verdade, os sujeitos desta pesquisa, estão em linguagem siglada no intuito de preservar as suas identidades.

(a) “[...] Só *tu*, minha vida, és capaz de me alegrar [...]” JMF. São Raimundo, 25.11.1921.

(b) “[...] Eu não esperava nem encontrava outro rapaz melhor que *tu*, portanto filhinho te pesso que não diga mais que *tu* es indigno de minha pessoa [...]” CRJM. Curral de Pedras, 11.03.1921.

(c) “[...] Aí, *você* pode deixar de andar em festas. Muitas vezes *você* saía com Vitória para festa [...] Mas pela manhã [...] já *você* havia chegado [...]” PNQ. Oeiras-PI, 12.01.1988.

(d) “[...] Como *você* sabe: houve na sessão estes números ou ocorrências: um recitativo (um soneto antigo de minha lavra cuja declamadora não gravei o nome), a execução do hino de Oeiras, a execução da valsa Yayazinha Bugyja [...]” ABSB. RJ, 16.06.1987.

(e) “[...] Olá! Beijos a todos. Acho que demorei um pouco a te escrever, mas tenho certeza que *entendes*. Logo que cheguei, reiniciei as atividades normais, trabalho, faculdade, inglês, e no serviço público *você* sabe que nessa época é fogo. [...]” MNRSR. Sobradinho, 17.03.79.

Assumimos como ponto de partida o fato de as frequências de uso das formas *tu* e *você* no PB falado terem evidenciado à *Scherre et alii* (2015 [2009]) seis subsistemas tratamentais no ¹território brasileiro que foram, por sua vez, sintetizados por Lopes & Cavalcante (2011) em três subsistemas. São eles: (I) o subsistema de *você* (regiões centro-oeste, sudeste), (II) o subsistema de *tu* (nordeste, norte, sul) e (III) o subsistema da alternância *você/tu* (regiões centro-oeste, sudeste, sul, nordeste, norte). Acrescentemos ainda a análise histórica proposta por Lopes & Rumeu (2015) em que as autoras discutem, à luz da *Teoria do Poder e da Solidariedade* (BROWN & GILMAN, 1960), a produtividade das estratégias pronominais de 2SG e as interações entre missivistas cariocas e mineiros no decorrer dos séculos XIX e XX. Nessa análise histórica, temos a difusão do *você* como uma possível evidência da sedimentação, no Brasil do século XIX, da *Semântica da Solidariedade*.

Considerando que o foco desta análise seja o de apresentar, à luz da *Teoria do Poder e da Solidariedade* (BROWN & GILMAN, 1960), uma descrição-analítica da alternância entre as formas *tu* e *você* e das relações sociais que as contextualizam em sincronias passadas, entendemos que o inovadorismo desta Tese esteja justamente no fato de o foco estar direcionado à produção escrita de redatores piauienses da era novecentista do PB. Nesse sentido, esta Tese preenche uma lacuna acerca do que se sabe sobre a referência ao sujeito de 2SG no PB novecentista dos piauienses.

Uma vez expostos o panorama diversificado de subsistemas tratamentais no PB falado atual (SCHERRE *et alii*, 2015) e o caráter científico deste trabalho que o expõe como uma Tese, esclarecemos que assumimos também como ponto de partida o princípio de que as línguas estão em constante processo de mudança linguística (LABOV, 1994).

¹ O subsistema de *Você*, o subsistema de *Tu* com baixo nível de concordância, o subsistema de *Tu* com alto nível de concordância, o subsistema *Tu/Você* com concordância baixa, o subsistema *Tu/Você* com concordância média, o subsistema *Você/Tu* sem concordância, distribuídos pelas regiões brasileiras.

Assim sendo, pretendemos, depreender até que ponto as variáveis linguísticas (a *expressão do sujeito*, a *pessoa verbal* e o *paralelismo formal e semântico*) e as extralinguísticas (o *sexo*, a *faixa etária*, as *relações interpessoais entre os missivistas*, o *período*, a *escolaridade*, os *níveis de prática de escrita*, o *subgênero da carta pessoal* e os *acervos*) tendem a influenciar as escolhas pronominais dos redatores piauienses em relação referência ao sujeito de 2SG em suas cartas pessoais.

Nesse sentido, envolvemo-nos com as seguintes questões de pesquisa. São elas:

(1^a) As formas pronominais *tu* e *você* também se mostrariam proficuas na escrita novecentista dos piauienses, considerando o fato de que o *você* prevalece na fala atual do Piauí, ainda que em alternância com o *tu* (com concordância média entre 10% e 39%), cf. Scherre *et alii* (2021, p. 165)?

(2^a) Quais contextos linguísticos e extralinguísticos promoveriam a produtividade da forma *você* na produção novecentista piauiense?

(3^a) Quais seriam os tipos de relações sociais que subjazem as escolhas pronominais *tu* e *você* na realidade da produção escrita novecentista dos piauienses?

Passamos às hipóteses sustentadas pelas duas questões anteriormente expostas.

(1^a) Assumimos a hipótese de que o *você* já se deixaria entrever como mais produtivo do que o *tu* na era novecentista do PB, cf. já levantado por Scherre *et alii* (2021) em relação aos dados da fala piauiense;

(2^a) Tendo em vista o fato de o *você* ter sido impulsionado por *formas verbais* de 3SG (variável *concordância*), pelo *você-sujeito*, por *clíticos* e *possessivos* de 3SG, por *formas de você-não-sujeito* em estruturas precedentes (variável *paralelismo formal e semântico*) e pela *faixa etária* do informante nas cartas mineiras dos séculos XIX e XX conforme a discussão de Souza (2021), conjecturamos que o *você* também seria conduzido por alguns desses fatores internos (linguísticos) e externos (sociais) ao sistema linguístico.

(3^a) Considerando a constatação de Souza (2021, p. 101) em relação ao *você* ter se mostrado mais produtivo do que o *tu* nas relações simétricas, cf. também observado por Lopes & Rumeu (2015, p. 23) para as cartas mineiras e por Souza (2012, p. 86) para as cartas cariocas, prevemos que também esse mesmo panorama de simetria social contextualizaria a produtividade do *você* nas cartas piauienses.

Estruturalmente, este estudo está organizado em 4 capítulos. Nas considerações iniciais, trazemos uma breve exposição do objeto de estudo, bem como das questões norteadoras e das respectivas hipóteses da pesquisa. No capítulo 1, apresentamos a revisão do tema em relação à história de formação do *você*, aos pronomes de 2SG na perspectiva da tradição e à produtividade do inovador *você* no PB de sincronias passadas e recentes. No capítulo 2, descrevemos a fundamentação *teórico-metodológica* nos âmbitos da SH (HERNÁNDEZ-CAMPOY & SCHILLING, 2012), da *Teoria do Poder e da Solidariedade* (BRONW & GILMAN, 1960), apresentamos pormenorizadamente as amostras de cartas piauienses, expomos os parâmetros linguísticos e sociais (grupos de fatores), resgatamos brevemente a história de constituição territorial do Piauí e divulgamos o processo de submissão dos dados históricos ao RBrul (RStudio). No capítulo 3, trazemos os resultados gerais (em termos estatísticos) voltados às variáveis linguísticas e sociais. No capítulo 4, voltamo-nos aos resultados probabilísticos em relação à regra variável em análise (*tu/você*). Por fim, chegamos, nas *considerações finais*, às breves generalizações, tendo em vista as questões principais e as principais hipóteses. Na sequência, trazemos as referências bibliográficas desta Tese e apresentamos a rodada probabilística gerada pelo RBrul/RStudio.

CAPÍTULO 1. REVISÃO DO TEMA: UMA BREVE RETOMADA DA ALTERNÂNCIA *TU/VOCE*.

Neste capítulo, apresentamos, em 1.1, uma brevíssima incursão pela história da origem da forma pronominal *você* até se tornar uma variante pronominal em competição com o *tu* na posição de sujeito, passamos ainda, em 1.2, pela perspectiva da tradição gramatical, em 1.3, pela descrição de algumas evidências históricas (sincronias passadas) e mais recentes (sincronias recentes), à luz de estudos descritivistas, acerca da atuação do *você* no PB, e, em 1.4, voltamo-nos aos dados da fala piauiense à luz dos dados do Atlas Linguístico do Brasil – AliB: inquéritos geolinguísticos, cf. a pesquisa bibliográfica feita por Scherre *et alii* (2021).

1.1 Um pouco de história: o processo de inserção do *você* no sistema pronominal do português.

A história do pronome pessoal *você* e a sua gênese está no português europeu (PE). Para apresentar esse percurso, voltamo-nos aos estudos de Faraco (2017 [1996]), Lopes *et alii* (2018), Rumeu (2011) e Menon (1995).

Em seu artigo intitulado “O tratamento *Você* em português: uma abordagem histórica”, Faraco (2017 [1996]) apresenta os aspectos formais, históricos e pragmáticos propulsores da forma pronominal *você* e as consequências gramaticais desencadeadas nos componentes morfológico e sintático da língua portuguesa. No que se refere à origem do pronome *você*, Faraco (2017 [1996]) entende que mudanças na sociedade portuguesa motivaram as mudanças no sistema pronominal da língua portuguesa. Retomando o percurso histórico, a partir do antigo sistema pronominal do latim tardio, o autor mostra que o antigo sistema de formas de tratamento era composto pelo paradigma da forma *tu*, para tratamento singular menos formal, e pela forma *vos*, tanto para a referência a mais de um interlocutor, quanto para um único interlocutor. A partir do século XII, com o desenvolvimento das atividades mercantis corporificadas e aliadas também à força da burguesia temos em evidência também o poder do reinado português. Segundo Faraco (2017 [1996]),

Em Portugal, já no período de Afonso III (1246-1279), dois sinais dessas mudanças econômicas, sociais e políticas eram observáveis. De um lado, a corte estava desenvolvendo uma vida mais luxuosa; e, de outro, a burguesia começou a ter, a partir de 1254, representantes nas Cortes, ao lado da nobreza e do clero. A crescente riqueza da burguesia lhe permitia competir com a nobreza em termos de poder econômico já na metade do século XIV. (FARACO 2017 [1996], p. 117).

Assim sendo, a estrutura administrativa do governo teve que ser ampliada e os cargos públicos passaram a ser distribuídos entre essa nova aristocracia, uma nobreza burocrática. Novos padrões de comportamento social, inclusive de tratamento do interlocutor, foram introduzidos entre a aristocracia portuguesa, não só como consequência do grande fluxo de riqueza, mas também como uma forma de estratificação social (FARACO 2017 [1996], p. 118).

A crescente conquista de poder político e econômico pela burguesia, que passou a ocupar mais espaços políticos e sociais, conduziu entre os séculos XIV e XV, através das leis de cortesia, a instituição de formas nominiais de tratamento específicas aos membros da Corte Portuguesa assumindo o *Vossa Mercê* inicialmente a referência exclusiva ao Rei na sociedade portuguesa. Assim sendo, no Português europeu, o sistema herdado do Latim – *tu/vos* – foi paulatinamente substituído por expressões nominiais de referência ao interlocutor, tais como *Vossa Mercê*, *Vossa Senhoria*, *Vossa Alteza*, *Vossa Majestade* e *Vossa Excelência*. Percorrendo essa trajetória histórica, trazemos à cena a análise de Menon (1995) a partir da correlação entre as transformações sociais de Portugal e sua repercussão na história do português através do uso institucionalizado de novas formas nominiais de tratamento.

[...] com as modificações havidas na sociedade portuguesa, ambas as formas, em função da alteração de valores que elas expressavam, passaram a ser empregadas como formas habituais de tratamento não-íntimo entre os nobres (isto é, entre *iguais*), os quais também exigiam essa forma de tratamento respeitosa da parte de pessoas de posição social inferior, quando se dirigiam aos nobres. As pessoas das categorias sociais mais baixas imitavam os nobres e também utilizavam as formas respeitosas na mesma escala, isto é, servos e artesãos se tratavam “respeitosamente” como os nobres o faziam entre si. Assim, tais formas acabaram perdendo seu valor honorífico e, em seguida, passaram, sobretudo *Vossa Mercê*, a ser empregadas por todo mundo. (MENON 1995, p. 94)

A forma pronominal *vós* (2PL), que já não era mais exclusiva para o rei, cedeu lugar às formas nominiais de tratamento estabelecidas *pelas leis de cortesia* durante o reinado

de Afonso V (1438-1481). O primeiro registro da forma de tratamento *Vossa Mercê* data de 1331, no texto das Cortes e, na sequência, temos *Vossa Senhoria*, em 1434, *Vossa Majestade*, em 1442, *Vossa Alteza*, em 1450, e *Vossa Excelência*, em 1455. Tais expressões nominais de tratamento, por se constituírem como sintagmas nominais, se combinavam com formas verbais de 3ª pessoa, ainda que semanticamente evocassem a 2SG. A inserção das expressões de tratamento e a instabilidade nos paradigmas pronominal e verbal conduziram a “revolução da terceira pessoa”, cf. Santos Luz (1957, p. 229 *apud* FARACO, 2017 [1996], p. 118). Santos Luz utilizou a expressão “revolução da terceira pessoa para denominar o conjunto de mudanças pelas quais passou o sistema de tratamento do português a partir do século XIV. Trata-se, sem dúvida, de um termo que exprime a efervescência pela qual passou a língua portuguesa, se considerarmos a variada cadeia de mudanças morfológicas e até sintáticas que tal processo desencadeou na língua (FARACO, 2017 [1996], p. 116). Ainda segundo Santos Luz (1957, p. 229 *apud* Faraco, 2017 [1996], o uso da forma *Vossa Mercê* atingiu os seguintes índices percentuais: 19% em 1455; 37% em 1472-3; 18% em 1477; 7% em 1481-2; e 0% em 1490. Uma vez implementadas essas mudanças na sociedade portuguesa e, conseqüentemente, no sistema pronominal da língua portuguesa, temos também mudanças em outros componentes da gramática da língua, como nos esclarece Faraco (2017, p. 115 [1996]).

No caso particular da mudança do sistema de tratamento em português e de suas repercussões gramaticais, estamos, portanto, numa posição bastante privilegiada: temos condições de observar como certas mudanças sociais podem exercer pressões sobre a língua (isto é, como mudanças sociais podem ser determinantes de mudanças linguísticas); e temos também condições de observar como essas mudanças linguísticas, uma vez implementadas, podem desencadear uma cadeia de mudanças internas. (FARACO 2017, p. 115 [1996])

A difusão de *vossa mercê* disparou a perda do seu valor honorífico, conduzindo-o por processo gradual e paulatino de perda de substância fonética e de potencial semântico de honorificência, de modo a conduzir a formação do pronome *você* que é, atualmente, prevalente no PB para a referência ao sujeito de 2SG (FARACO 2017 [1996], p. 121). Em outras palavras, com as mudanças fonéticas (*vossa mercê* > *vosmecê* > *você*) e as mudanças semântico-pragmáticas, o *vossa mercê* se gramaticalizou, à medida que passou a desempenhar a função de pronome pessoal de 2SG (LOPES *et alii* 2018, p. 34).

O novo pronome *você*, por ter origem em uma expressão nominal que estabelece relação de concordância com o verbo em 3SG, passou a ter comportamento sintático idêntico ao da expressão nominal de tratamento que o originou (*vossa mercê*), acionando também concordância com a 3ª pessoa verbal. Ainda sobre esse ponto, Lopes & Rumeu (2007) afirmam que:

Por derivar de uma forma nominal que leva o verbo para a terceira pessoa do singular, o emprego de *você* na interlocução acarretou, por exemplo, um rearranjo no sistema pronominal com a fusão do paradigma de 2ª com o de 3ª pessoa do singular e com a eliminação do paradigma de 2ª pessoa do plural. (LOPES & RUMEU 2007, p. 419)

Assim sendo, temos uma forma pronominal (*você*) que faz referência ao interlocutor em termos do traço de pessoa semântica [-EU], cf. Lopes & Rumeu (2007), Rumeu (2006), mas que se combina formalmente com o verbo na 3ª pessoa como evidência de persistência (nos termos de HOPPER, 1991) de um traço formal da original no item gramaticalizado, cf. Lopes *et alii* (2018, p. 34). Assim sendo, a mudança no sistema pronominal motivou uma alteração no paradigma verbal, de modo que o português passou a ter duas formas para a referência ao sujeito de 2SG: uma delas (*tu*) morfologicamente marcada através do morfe [-s] como desinência número pessoal na forma verbal e outra forma (*você*) correlacionada a uma forma verbal na 3SG, cf. Menon (1995, p. 97). Isso quer dizer que temos não só o uso do etimológico *tu* (*canta*[-s]), mas também a referência ao sujeito de 2SG com formas verbais de 3SG como observamos, por exemplo, em *você canta*[-Ø]; *tu canta*[-Ø], de modo a neutralizarmos as distinções formais entre as formas verbais de 2ª e 3ª pessoas do singular justamente pela ausência de morfemas específicos que as especializem como já amplamente discutido por Lopes *et alii* (2018), Lopes & Cavalcante (2012), Lopes & Rumeu (2015), Lopes (2015), à luz também de Hopper (1991) especificamente em relação ao princípio da especialização.

Faraco entende que o *você* guarda, em algumas regiões de Portugal (PE), alguma marca negativa, o que nos leva a crer que essa forma pronominal deve ter tido origem no estilo mais informal da pequena burguesia (2017, p. 121 [1996]).

A situação no Brasil é bastante diferente. *Você* é o pronome de uso comum para o tratamento íntimo, estando o pronome *tu*

restrito a algumas variedades regionais. A razão para esse uso tão amplo de *você* no Brasil deve ser encontrada, certamente, na história da formação do País. É fato que não temos documentos das formas linguísticas correntes nos primeiros tempos da ocupação europeia. Alguns dados, porém, são importantes para um processo de reconstrução hipotética de tais fatos. (FARACO 2017, p. 121 [1996])

A realidade discutida por Faraco (2017 [1996]) deve-se provavelmente ao fato de que, no final século XV, em Portugal, o *vossa mercê* e suas variantes passaram a ser de uso comum entre a população. Nesse momento, temos também a vinda dos primeiros colonizadores para o Brasil, a partir de segunda metade do século XVI, período em que a simplificação fonética do *vossa mercê* se achava em fase bastante avançada (FARACO 2017, p. 122 [1996]). A complexidade do processo de gramaticalização do *você* parece ter sido orientado pela dinâmica sintetizada por Hopper e Traugott (1993) através de ciclos qual seja *item lexical* > *item gramatical* > *clítico* > *afixo* que corresponderia (à exceção da fase do processo “afixo”) à seguinte cadeia evolutiva, cf. discutido por Vitral & Ramos (2008): *Vossa Mercê* > *vosmecê* > *vancê* > *você* > *ocê* > *cê*.

Na perspectiva da análise de Rumeu (2004), embasada em cartas setecentistas e oitocentistas, o *Vossa Mercê* já disputava espaço com a forma *tu*, no campo da intimidade. Nas cartas brasileiras novecentistas, por outro lado, o *você* concorre com o *tu* em relações simétricas, tanto entre os missivistas cultos, quanto entre os menos escolarizados (LOPES *et alii*, 2018).

No Brasil do século XIX, a concorrência passou a ser maior entre *tu* e *você* em relações solidárias mais íntimas e de confiança, ao passo que *Vossa Mercê* foi-se tornando gradativamente uma forma obsoleta. (LOPES *et alii*, 2018, p.34)

Ao retomarmos as análises de Rumeu (2004) e de Lopes *et alii* (2018), constatamos que estas, com seus achados de pesquisa sobre a inserção de *você* na língua portuguesa, convergem com a discussão de Faraco (2017 [1996]), que se volta a pormenorizar o contexto histórico-social de evolução da forma nominal de tratamento *vossa mercê*, que, paralelamente às modificações fonéticas e semântico-pragmáticas, ganha nova função na gramática, a de pronome pessoal, ao evoluir para a forma *você*.

Analisando cartas cariocas oitocentistas e novecentistas, à luz da Teoria do Poder e da Solidariedade (Brown & Gilman 1960), Rumeu (2011) constata que, no Brasil, nas cartas setecentistas e oitocentistas, nas relações sociais assimétricas ascendentes (de

inferior para superior), a forma nominal de tratamento *Vossa Senhoria* concorre com *Vossa Excelência*. Já nas relações sociais assimétricas descendentes, a forma *você*, no século XVIII, apresenta-se como uma estratégia de cortesia descendente (de superior para inferior) bastante produtiva. Ainda em relação ao período oitocentista, *Vossa Senhoria* passa a competir com as formas *você*, *tu* e *o senhor*, em relações simétricas.

Recorrendo aos dados de Soto (2001), Rumeu (2013) nos evidencia que, a partir do século XIX, a aristocracia brasileira emprega o *você* diferentemente da portuguesa. O fato de o *você* preservar o prestígio de *vossa mercê* revela um certo conservadorismo no PB. Por outro lado – prossegue a autora – agora questionando se o emprego de *você*, pela Condessa de Barral ao se referir ao Imperador D. Pedro II, e de Rui Barbosa ao se dirigir a José Marcelino (senador e governador da Bahia) já não seriam indícios do acelerado processo de dessemantização do *você* com o inovador uso no domínio da solidariedade. Nesse sentido, Lopes *et alii* (2018), também se fundamentando em Soto (2001), acrescentam que, diante desse panorama tratamental variável entre *tu* e *você*, a era oitocentista é tida como o ponto de partida para o início da vivacidade do *você* como pronome de referência ao sujeito de 2SG no PB.

Uma vez feita essa brevíssima retomada da história de formação do *você* na sociedade portuguesa, passamos, na seção 1.2, à exposição do pensamento prescritivista em relação ao quadro de pronomes pessoais e, mais especificamente, voltando-nos aos pronomes de referência ao sujeito de 2SG.

1.2 Os pronomes de referência ao sujeito de 2ª pessoa do singular na perspectiva da tradição gramatical

No que tange à perspectiva da tradição gramatical, voltamo-nos às orientações assertivas de Cunha & Cintra (2011 [1985]), de Rocha Lima (2011 [1972]) e de Bechara (2009) voltadas ao sistema pronominal no português. O sistema pronominal do Português segundo as prescrições de Cunha & Cintra (2011 [1985]), de Rocha Lima (2011 [1972]) e de Bechara (2009) assemelham-se em essência, uma vez que o intuito principal da gramática normativa é o de preservar a tradição gramatical. Nesse sentido, observamos as seguintes definições da categoria *Pronome* à luz da prescrição gramatical.

Os pronomes desempenham as funções equivalentes às exercitadas pelos elementos nominais. Servem para representar um substantivo, para acompanhar um substantivo determinando-lhe a extensão do significado. (CUNHA & CINTRA, 2011 [1985], p. 268).

“Pronome é a palavra que denota o ente ou a ele se refere, considerando-o apenas como pessoa do discurso” (ROCHA LIMA, 2011 [1991]).

Em termos semânticos, os pronomes pessoais evidenciam três pessoas discursivas que passam pela indicação de *quem* fala (“eu”, “nós”), de *com quem* se fala (“tu”, “vós”) e *de quem* se fala (“ele”/“ela”, “eles”/“elas”). Em termos funcionais, os pronomes pessoais podem assumir as formas *retas* (nominativas) para a função de sujeito da oração ou as formas *oblíquas* (acusativas e dativas) para as funções de complemento, direto ou indireto, respectivamente. Temos, ainda, as formas oblíquas que se subdividem em formas *átonas* (sem preposição) e formas *tônicas* (sempre regidas de preposição), cf. a sistematização exposta no quadro 1.

Quadro 1: Síntese dos pronomes pessoais na perspectiva da prescrição gramatical.

<p>1ª pessoa (<i>quem</i> fala): formas retas: <i>eu</i> (singular), <i>nós</i> (plural); formas oblíquas átonas: <i>me</i> (singular), <i>nos</i> (plural); formas oblíquas tônicas: <i>mim</i>, <i>comigo</i>.</p>
<p>2ª pessoa (<i>com quem</i> se fala): formas retas: <i>tu</i> (singular), <i>vós</i> (plural); formas oblíquas átonas: <i>te</i> (singular), <i>vos</i> (plural); formas oblíquas tônicas: <i>ti</i>, <i>contigo</i> (singular), <i>convosco</i> (plural).</p>
<p>3ª pessoa (<i>de quem</i> se fala): formas retas: <i>ele</i>, <i>ela</i> (singular), <i>eles</i>, <i>elas</i> (plural); formas oblíquas átonas: <i>se</i>, <i>lhe</i> o, a, (singular), <i>lhes</i>, <i>os</i>, <i>as</i> (plural); formas oblíquas tônicas: <i>si</i>, <i>consigo</i>, <i>ele</i>, <i>ela</i> (singular), <i>eles</i>, <i>elas</i> (plural).</p>

No que diz respeito à forma *você*, Cunha & Cintra (2011 [1985], p. 282) a incluem entre os pronomes de tratamento ao lado de formas tratamentais tais como *senhor* e *vossa excelência*. O gramático Rocha Lima (2011 [1972], p. 156-158, afirma que essas formas pronominais, embora sejam de 2ª pessoa, requerem o verbo na 3ª pessoa.

Para o gramático Bechara (2009, p.195), o entendimento é o de que “os pronomes estão caracterizados porque indicam dêixis [...]”. Nesse sentido, Bechara (2009, p. 196-197) nos traz a noção da *não pessoa* para a 3ª pessoa, ao passo que os pronomes pessoais, em essência, remetem-nos a 1ª e 2ª pessoas do eixo discursivo. Nesse sentido, o gramático dialoga com a interpretação de Benveniste (1995 [1976], p. 282-283).

Os chamados de ‘terceira pessoa’ são inteiramente diferentes de *eu* e *tu*, pela sua função e pela sua natureza, pois flexionam-se em gênero e número, enquanto os outros não estão sujeitos a esse processo morfossintático. As formas como *ele*, *o*, *isso* etc só servem na qualidade de substitutos abreviativos. (BENVENISTE 1995 [1976], p. 282-283)

Em síntese, os três representantes da prescrição gramatical convergem em relação aos paradigmas pronominais descritos para a referência ao sujeito do discurso. Na perspectiva da tradição, temos o *tu* como forma exclusiva para a referência ao sujeito de 2SG e o *você* (*vocês*) como uma forma de *tratamento* familiar em oposição ao cerimonioso *senhor/senhora*.

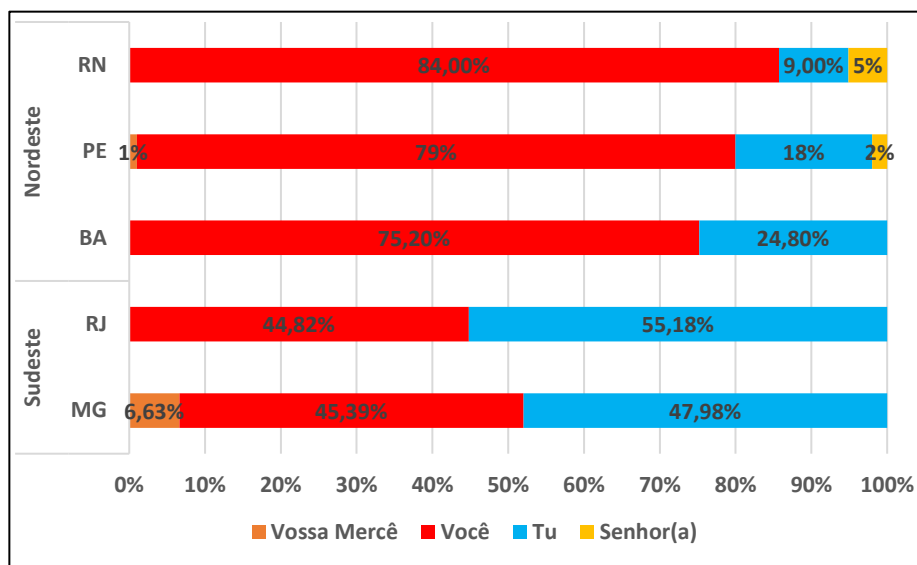
Uma vez exposta a perspectiva prescritivista em relação ao quadro de pronomes pessoais do português, passamos à descrição das formas de referência ao sujeito de 2SG não só em uso em sincronias passadas (séculos XIX e XX), mas também em uso em sincronias recentes, voltando-nos especificamente à fala piauiense à luz dos dados do Atlas Linguístico do Brasil – AliB concretamente analisados por Cardoso (2008) e por Cardoso & Deus (2009) *apud* Scherre *et alii* (2021).

1.3 O *Você* no Português Brasileiro de sincronias passadas

Ao voltarmos-nos às análises da alternância *tu/você* em sincronias passadas, passamos àquelas embasadas em amostras da produção escrita de brasileiros das regiões sudeste (Rio de Janeiro (SOUZA, 2012) e Minas Gerais (SOUZA, 2021)) e nordeste (Bahia (BATISTA, 2017), Pernambuco (GOMES & LOPES, 2018), ²Rio Grande do Norte (MOURA & MARTINS)), o que está consubstanciado no gráfico 1. A opção pela descrição dos resultados de estudos históricos acerca das formas pronominais de 2SG *tu/você* ancorada tão somente na comparação entre os resultados das regiões sudeste e nordeste está amparada no acentuado número de dados linguísticos levantados em confiáveis amostras históricas para as regiões em questão.

² Para as localidades de Pernambuco e do Rio Grande do Norte, os resultados expostos estão discutidos também em Lopes *et alii* (2018).

Gráfico 1: As formas *vossa mercê*, *ocê*, *tu*, *senhor(a)* distribuídas pela produção escrita de brasileiros do nordeste e do sudeste entre os séculos XIX e XX.



Legenda (gráfico 1):

RN = análise de Moura & Martins (2018) embasada em amostras de cartas pessoais representativas do Rio Grande do Norte;
 PE = análise de Gomes & Lopes (2018) embasada em amostras de cartas pessoais representativas de Pernambuco (séculos XIX e XX);
 BA = análise de Batista (2017) embasada em amostras de cartas pessoais representativas da Bahia (séc. XX);
 MG = análise de Souza (2021) embasada em amostras de cartas pessoais representativas de Minas Gerais (séculos XIX e XX);
 RJ = análise de Souza (2012) embasada em amostras de cartas pessoais representativas do Rio de Janeiro (séculos XIX e XX).

Com base na análise de trezentos e cinquenta e quatro cartas pessoais produzidas no intervalo de cem anos (entre 1870 e 1970) por escreventes nascidos e/ou residentes no espaço geográfico do Rio de Janeiro, Souza (2012) volta o foco da sua análise à alternância *tu/você*. Conduzida pelos princípios teórico-metodológicos da Sociolinguística Variacionista (doravante SV) e da Teoria do Poder e da Solidariedade, Souza conjecturou que a forma *tu* seria a forma pronominal mais produtiva no século XIX e no início do século XX, sendo, posteriormente, suplantada pelo inovador *você* a partir dos anos de 1930 do século XX (DUARTE 1993, 1995).

A partir da análise das mil trezentos e oitenta e uma (1381) ocorrências das formas *tu* e *você*, a autora detectou o *você*, em 44,82% (619 oco), e o uso do *tu*, em 55,18% (762 oco) entre os séculos XIX e XX, confirmando a hipótese de a década de 30 do século XX ter se mostrado como um marco da implementação crescente do *você* (DUARTE, 1993). No que se refere às relações sociais que embasam a dinâmica *tu/você*, Souza (2012, p.

103 e 111) constatou as relações sociais simétricas como propulsoras do *você* (0.571) principalmente nas relações entre parentes e amigos, ao passo que as relações sociais entre casais (0.186) se mostram inibidoras do *você* e propulsoras, por outro lado, do *tu*, nas cartas amorosas. De um modo geral, a autora admite que o *você* é menos produtivo nas dinâmicas sociais marcadas pelo maior nível de intimidade como é o caso das relações sociais entre casais (relações amorosas).

Em termos históricos, Souza (2012) detectou três períodos históricos para a dinâmica *tu/você* nas cartas pessoais cariocas analisadas. São eles: os períodos de 1870-90 (I), de 1900-30 (II), de 1930-70 (III). No período de 1870-90, o pronome *você* apresentou menor produtividade, tendo passado por uma alavancada na sua produtividade no decorrer dos períodos II (1900-30) e III (1930-70). Ao longo do século XX (entre as décadas de 30 e 70), o *você* passa a assumir a dianteira, o que permite a autora a entendê-la como a forma, de um modo geral, preferida pelo redator carioca de sincronias passadas.

Ao voltar-se às fontes históricas mineiras, Souza (2021) descreve o nível de alternância entre as formas *tu/você* correlacionando-a às relações sociais simétricas e assimétricas que as conduzem, tendo em vista os princípios da *Teoria do Poder e da Solidariedade* (BROWN & GILMAN, 1960), da Sociolinguística Histórica (HERNÁNDEZ-CAMPOY & CONDE SILVESTRE, 2012; HERNÁNDEZ-CAMPOY & SHILLING, 2012) e da Sociolinguística Laboviana (WEINREICH, LABOV & HERZOG, 2008 [1968]; LABOV, 1994). Partindo da hipótese já testada por Rumeu, Cruz & Cardoso (2018) de que o *você* seria mais produtivo do que o *tu*, Souza (2021) volta-se ao nível da alternância *tu/você* em função das relações de Poder e de Solidariedade estabelecidas por redatores nascidos e/ou residentes no espaço geográfico mineiro entre os séculos XIX e XX. Conduzido pela análise de 332 cartas pessoais a partir das quais levantou 1194 ocorrências, chegou à identificação de 573 oco de *tu* (47,98%), 542 oco de *você* (45,39%) e 79 oco de *vossa mercê* (6,63%). A distribuição dos dados no decorrer do tempo, evidenciou três fases para a manifestação da alternância pronominal na referência ao sujeito de 2SG. São elas: fase I (1840-1889), fase II (1890-1929) e fase III (1930-2009)).

Para a 1ª fase (1840-1889), o autor observa a alternância entre as formas *vossa mercê* e *tu*. Nesse período, o *vossa mercê* conserva-se com as frequências de 67%, 99%, 100% e 67% para os anos de 1840-49, 1850-59, 1860-69 e 1880-89, respectivamente. O *vossa mercê*, a partir de fins do século XIX (1890), tem o seu uso enfraquecido,

assumindo, entre os anos de 1890 e 1899 até atingir 1% de frequência de uso, entre os anos 1900 e 1909.

Para a 2ª fase (1890 e 1929), Souza (2021) detecta dados do *você*, mesmo que com baixos índices percentuais, no período que compreende os anos de 1890-99, 1900-09, 1910-19 e 1920-29, alcançando as frequências de 27%, 24%, 20%, 2%, respectivamente. Atenta o autor ao fato de que à proporção que o *você* mostra-se mais produtivo (entre 1890 e 1929), o *vossa mercê* torna-se menos frequente e o *tu*, por outro lado, mostra-se gradualmente mais frequente, ao alcançar as frequências de 69% (1890-99), 75% (1900-09), 80% (1910-19) e 98% (1920-29) em alternância com o *você*.

Para a 3ª fase (entre os anos de 1930 e 2009), as formas *tu* e *você* mostram-se em alternância. Enquanto o *tu* tem o seu uso em decréscimo com as frequências de 49% (1930-39), 9% (1940-49), 11% (1950-59), 8% (1960-69) até atingir produtividade nula para a produção escrita entre os anos 1970 e 2009, o *você*, por outro lado, passa a assumir índices mais expressivos entre os anos de 1930-39 (51%), de 1940-49 (91%), de 1950-59 (89%), de 1960-69 (92%) até alcançar produtividade categórica entre a década de 70 do século XX (1970) e o fim da 1ª década dos anos 2000 (2000-09).

Considerando que a proposta de Souza é a de submissão dos dados ao RBrul em interface com o RStudio para a geração dos pesos relativos, chegou o autor à conclusão de que as formas verbais de 3SG (concordância), o *você-sujeito*, as formas clíticas de 3SG, as formas possessivas de 3SG e as formas de *você não-sujeito* mostram-se como contextos que tendem a disparar o *você* na produção escrita mineira. No âmbito das relações sociais, Souza identificou o *vossa mercê* circunscrito às relações assimétricas ascendentes (de inferior para superior), enquanto as formas *tu* e *você* mostram-se prevalecem nas relações de simetria social, o que conduz o autor a interpretar a sociedade mineira com uma dinâmica social orientada pelos caminhos da Solidariedade (BROWN & GILMAN, 1960).

Voltando o foco especificamente a um estudo sobre a variação *tu/você* plenos embasado em cartas pessoais nordestinas, passamos à descrição dos resultados de Batista (2017) conduzida pela análise de amostras representativas da escrita baiana do século XX (1930 e 1980). Trata-se de uma análise voltada à produção escrita de escreventes baianos pouco e medianamente escolarizados no âmbito das dinâmicas sociais mais e menos

íntimas, o que está consubstanciado nas amostras de cartas familiares (entre informantes unidos por laços consanguíneos), amorosas (entre casais) e de amizade (entre amigos íntimos) no decorrer das décadas de 30 e 80 do século XX. A análise linguística das formas *tu/você* está orientada pelos princípios teóricos e metodológicos da Sociolinguística Quantitativa (LABOV, 1994) e da Teoria da Polidez de Brown & Levinson (1987), tendo em vista o intuito principal de apreensão das relações sociais conduzidas pelos remetentes e destinatários das missivas históricas baianas para a preservação da face entre os interactantes da *Família Estrela Tuy* (1930 e 1980).

Em termos de método para o levantamento dos dados linguísticos, Batista (2017) se utiliza não só da coleta manual, mas também da ferramenta de busca *E-Corp* à luz de Souza (2016). Para o tratamento dos dados linguísticos, a autora submete-os ao conjunto de programas vinculados ao GoldVarb X, tendo sempre em vista a apreensão precisa dos contextos linguísticos e extralinguísticos aos quais estão subsidiadas as formas *tu* e *você* plenas. De um modo geral, os resultados evidenciaram à autora o *você* como a estratégia preferida (75,2%) pelos redatores baianos nas relações simétricas mais solidárias e nas simétricas menos solidárias, cf. gráfico 1. Constatou a autora que as relações simétricas menos solidárias se mostram como propulsoras do uso do *você* amplo e generalizado nesse contexto de uso. Por outro lado, o *tu* mostra-se com um uso exíguo (24,8%), restringindo-se à intimidade das relações mais solidárias, cf. Batista (2017, p. 109).

Embasadas nas amostras representativas da produção de escreventes cultos do estado brasileiro de Pernambuco, Gomes & Lopes (2018) conduzem-se pelo levantamento de cento e vinte e três (123) cartas pessoais produzidas entre os anos de 1869 e 1969, configurando uma amostra composta por 47.500 palavras. Em termos gerais, as autoras chegam ao levantamento de 354 ocorrências de formas de referência ao sujeito de 2SG, sendo 279 oco de *você* (79%), 64 oco de *tu* (18%), 5 oco de *vossa mercê* (1%) e 6 oco de *senhor/senhora* (2%), cf. descrito no gráfico 1. A produtividade da forma *você* evidencia a sua alta frequência de uso nas distintas relações sociais, comprovando o seu traço polifuncional já evidenciado em dados de amostras oitocentistas do PB escrito. Isso posto, as autoras admitem que o *você* parece ter se firmado, já no início da era novecentista, como uma forma de referência ao interlocutor *neutra*, mostrando-se propícia às distintas relações sociais e às suas alterações no desenrolar do tempo.

A análise de Moura & Martins (2018) voltada à dinâmica *tu/você* em amostras de cartas pessoais representativas do estado nordestino do Rio Grande do Norte, conduziram-no pelo levantamento de dados de *tu/você* em trezentos e quatro (304) cartas pessoais novecentistas (1916-1994). Dentre as oitocentas e noventa e duas evidências pronominais de 2SG (892 oco), os autores identificam o *você*, 756 oco (84 %), acompanhada pelo *tu*, em 86 oco (9%), e pela forma *senhora*, em 50 oco (5%). Em termos de relações sociais, os autores constatarem o seguinte: a) o uso quase categórico do *você* nas relações familiares simétricas (entre irmãs) e relações simétricas entre amigos; b) o uso da forma *senhor/senhora* no âmbito das relações assimétricas ascendentes (de filho para mãe) como evidência de um tratamento mais cortês e respeitoso na produção escrita de 1940; c) o tímido uso do *tu*, que com frequência de uso de 5%, figura como uma forma pronominal de maior intimidade entre os interlocutores.

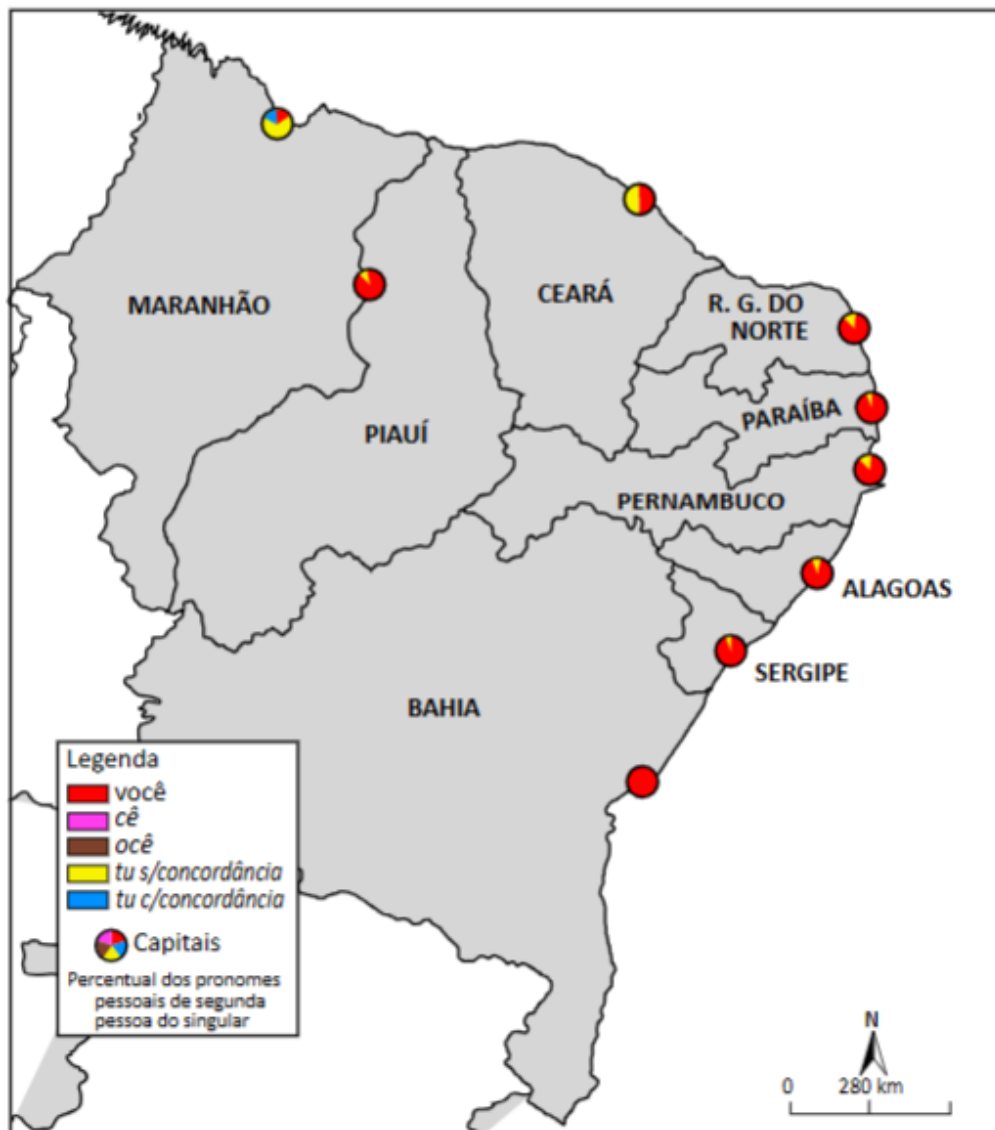
1.4. O *Você* no Português Brasileiro de sincronias recentes: algumas evidências na fala piauiense.

Ao voltarmos o foco especificamente aos usos das formas *tu/você* na fala piauiense, temos os resultados levantados da pesquisa bibliográfica feita por Scherre *et alii* (2021). Com base na análise dos dados do Atlas Linguístico do Brasil – ALiB: inquéritos geolinguísticos (CARDOSO, 2008; CARDOSO & DEUS, 2009 *apud* Scherre *et alii*, 2021), observamos, na tabela 1, a prevalência do *você*, em 92% dos dados, em relação ao *tu* sem concordância, em 8%, cf. a síntese feita por Scherre *et alii* (2021, p. 176), para a fala de Teresina (PI), cf. a figura 1.

Tabela 1: Média da distribuição de pronomes explícitos na segunda pessoa do singular no PB falado na capital do Piauí (Teresina) e região Nordeste: médias de amostras diversificadas, cf. Scherre *et alii* (2021, p. 176).

CAPITAL	AMOSTRA	VOCÊ	CÊ	OCÊ	“TU”	“TU”	100%
	/PROJETO				S/CONC.	C/CONC.	(TOTAL)
Teresina-PI	Amostra	92%	?	?	8%		(53)
	ALiB	(49)			(4)		

Figura 1: Mapa da ³distribuição dos pronomes pessoais de 2SG *você*, *ocê*, *cê* e *tu* (com e sem concordância) por 9 capitais nordestinas, dentre elas o Piauí, cf. Scherre *et alii* (2021, p.175).



Fonte: pesquisa bibliográfica, elaborado por Marta Scherre, Carolina Andrade e Rafael Catão.

À luz dos resultados expostos na tabela 1 e explicitados na figura 1, Scherre *et alii* (2021, p. 166) temos a projeção do subsistema 5 para a fala piauiense, marcada pela alternância *tu/VOCÊ*, sendo o *tu* acionado com concordância média, de 10% a 39%.

³ “Mapa dos percentuais de cinco construções com pronomes pessoais de segunda pessoa do singular (*você*, *cê*, *ocê*, *tu* sem concordância e *tu* com concordância) nas nove capitais da região Nordeste”, cf. Scherre *et alii* (2021, p. 175).

CAPÍTULO 2. FUNDAMENTAÇÃO TEÓRICO-METODOLÓGICA

Neste capítulo, apresentamos os principais encaminhamentos teórico-metodológicos que norteiam o levantamento da amostra histórica e a análise dos dados das formas pronominais de referência ao sujeito de 2SG (*tu* e *você*). Neste sentido, mantemo-nos orientados pelos princípios da SH e os apresentamos, em 2.1, à luz principalmente de Hernández-Campoy & Schilling (2012), pela proposta de Brown & Gilman (1960) em relação à *Teoria do Poder e da Solidariedade*, em 2.2, pela descrição das amostras de cartas pessoais piauienses, atentando ao gênero textual *carta pessoal* em 2.3, por uma apresentação do quadro geral de cartas piauienses, em 2.4, pela descrição dos parâmetros linguísticos e extralinguísticos (grupos de fatores) aos quais submetemos os dados linguísticos, em 2.5, por uma breve retomada da história de formação do Piauí, em 2.6, e pela apresentação do método de tratamento quantitativo dos dados históricos submetidos ao RBrul/RStudio, em 2.7.

2.1 Alguns princípios teóricos da Sociolinguística Histórica aplicados às cartas piauienses

As bases teóricas desta tese estão ancoradas nos princípios teórico-metodológicos da SH de Hernández-Campoy e Schilling (2012), Conde Silvestre (2007) que, por sua vez, se deixam orientar pelos parâmetros da SV e da Mudança à luz de Weinreich, Labov, Herzog (2006 [1968]), Labov (2008 [1972]). O ponto de encontro entre a SH e a SV parece estar não só no fato de que o princípio da variação revela formas conservadoras e inovadoras constantemente em alternância, mas também em relação à ideia de que uma dada regra variável (no âmbito de sincronias passadas ou sincronias recentes) está condicionada por variáveis linguísticas e extralinguísticas. Esses aspectos de interseção entre essas duas abordagens colocam em evidência também o fato de que cabe à SH a análise de sincronias passadas, e à SV, a análise da sincronias recentes. Com o intuito de consolidarmos esse ponto de vinculação da SH à SV (*stricto sensu*), valemo-nos de Bagno (2017, p. 430) que, em seu *Dicionário crítico de Sociolinguística*, retoma, no verbete “Sociolinguística Histórica [*HISTORICAL SOCIOLINGUISTICS*]”, um pouco da história dessa área de investigação linguística.

SOCIOLINGUÍSTICA HISTÓRICA

[*HISTORICAL SOCIOLINGUISTICS*]

Campo de investigação desenvolvido inicialmente por Suzanne Romaine (1982) e James Milroy (1992) que, combinando a **sociolinguística** com a **linguística histórica**, visa “utilizar o passado para explicar o presente, e o presente para explicar o passado das línguas” (Trudgill e Hernández Campoy, 2007: 299). Esta disciplina leva ainda mais adiante o postulado central da **sociolinguística variacionista** de que é preciso romper com a divisão rígida, herdada do estruturalismo clássico (saussuriano), entre **sincronia** e **diacronia**, uma vez que a variação linguística revela o convívio e a competição ininterrupta entre formas linguísticas conservadoras e formas linguísticas inovadoras. Seu construto epistemológico mais destacado é o princípio do **uniformitarismo**, postulado por W. Labov, segundo o qual todos os fatores sociais e linguísticos de variação e mudança ocorridos no passado continuam em ação no presente. Do ponto de vista metodológico, a função principal da sociolinguística histórica é desenvolver uma série de procedimentos que permitam reconstruir uma língua em seu **contexto social** num dado momento do passado. Para tanto, é fundamental combinar a volumosa tradição filológica de interpretação dos textos com os trabalhos mais recentes dentro da metodologia **quantitativa**. [...]

BAGNO, M. *Dicionário crítico de sociolinguística*. Parábola Editorial. Edição do Kindle. 2017, p. 430.

A pesquisa no âmbito da SH exige do pesquisador o controle de variáveis internas e externas, uma vez que tanto a variação, quanto a mudança, assim como nas análises no âmbito da SV, são condicionadas por contextos internos e externos ao sistema linguístico. Nesta análise, estão em cena as formas pronominais *tu* e *você* para a referência ao sujeito de 2SG em cartas piauienses novecentistas. Assim sendo, temos uma amostra composta por 136 (cento e trinta e seis) cartas distribuídas em relação aos subgêneros amorosas, familiares e amizade. Ao considerarmos que as cartas trocadas entre casais (amorosas), amigos (amizade) e parentes (familiares) representam um eixo mais íntimo de comunicação, entendemos a *carta pessoal* como um gênero mais transparente em relação à expressão vernacular e, conseqüentemente, mais livre da força coercitiva da norma-padrão (BARBOSA 1999 *apud* RUMEU 2013, p. 551). Nesse sentido, assumimos se tratar de uma confiável fonte para a captação de dados linguísticos de sincronias passadas.

Com o intuito de esclarecermos o que entendemos por “vernáculo” (expressão linguística vernacular), trazemos a concepção de Bagno (2017) com a qual nos identificamos para o encaminhamento dos pressupostos norteadores desta análise.

VERNÁCULO [*VERNACULAR*] O termo vernáculo tem duas acepções principais na sociolinguística. A primeira se refere a **variedades** não padronizadas, relativamente homogêneas e bem definidas, empregadas regularmente por certos **grupos sociais** (geográficos, étnicos, socioeconômicos, etários etc.) e que existem em oposição a um **padrão** dominante. [...] A segunda acepção deriva do trabalho de W. Labov, que define vernáculo como o **estilo** menos formal possível do **repertório** linguístico de um falante. O vernáculo é empregado quando a pessoa fala com amigos e familiares em contextos não monitorados. É adquirido na infância e considerado linguisticamente mais regular do que os estilos de **fala** mais monitorados, que exibem caracteristicamente graus variados de influência das variedades urbanas de prestígio ou da norma-padrão canônica. Por sofrer menos influência dessas formas de prestígio e/ou padronizadas, o vernáculo é o objeto de maior interesse para a pesquisa sociolinguística, uma vez que é nele que se encontram mais vivas as dinâmicas de **variação** que podem conduzir a **mudanças** futuras. Na metodologia de pesquisa típica da **sociolinguística variacionista**, o vernáculo é apreendido por meio das diversas técnicas de **entrevista** desenhadas para tal fim. [...] (grifos em negrito do autor)

BAGNO, M. *Dicionário crítico de sociolinguística* (Referenda) (p. 475). Parábola Editorial. Edição do Kindle. 2017.

Conscientes de que não há como resgatar a *fala* de informantes desconhecidos em sincronias passadas, entendemos a *carta pessoal* como um gênero textual propício ao levantamento de evidências de uma expressão linguística menos monitorada, uma vez que se trata de um tipo de carta marcada pelo tom de intimidade entre o remetente e o destinatário.

O levantamento de amostras históricas confiáveis pode ser obstaculizado pela própria dificuldade de encontrarmos textos preservados a ponto de permitirem a sua conservadora transcrição e análise dos dados. Acrescente-se a isso o fato de as fontes históricas poderem estar fragmentadas, o que dificulta o controle rígido das variáveis sociais, tais como *sexo, idade, escolaridade, profissão (papel social)*. Nesse sentido, Hernández-Campoy & Schilling (2012, p. 63) ao considerarem as naturais dificuldades do trabalho com fontes históricas, discutem os parâmetros de organização das fontes históricas (*representatividade, validade empírica, invariação, autenticidade, autoria, validade social e histórica e ideologia padrão*), de modo a conduzir o linguista-pesquisador ao êxito das suas análises no âmbito da SH. Dentre eles, trazemos à tona a discussão da *autoria*, da *representatividade* e da *validade social e histórica* das amostras

históricas, uma vez se mostram estreitamente vinculadas à formação de amostras históricas e ao perfil social dos redatores de sincronias passadas do PB.

Com base nos parâmetros apresentados por Hernández-Campoy & Schilling (2012), ao nos valermos do princípio do uniformitarismo⁴, pressupomos que os processos naturais de variação e mudança atuantes na língua humana, também estiveram outrora em cena, o que nos motiva a nos voltarmos ao PB escrito novecentista. Nesse resgate histórico, a confiabilidade das fontes históricas é condição *sine qua non* à consequente validação dos resultados das análises quantitativas, conduzindo-nos a descrever analiticamente os critérios da *autenticidade*, da *autoria* e da *validade social e histórica* que nos conduzem no processo de levantamento e de seleção das cartas piauienses novecentista.

2.1.1 A autenticidade das cartas piauienses

O parâmetro da *autenticidade*, segundo Hernández-Campoy & Schilling (2012, p. 68), está relacionado à “pureza dos textos”, tendo em vista os possíveis indícios de *erro*, *mistura dialetal* e *hipercorreção* (LABOV, 1994). É importante que tenhamos consciência sobre o fato de que as cartas novecentistas analisadas não refletem a fala “real” (*ipsis litteris*) dos seus autores, ou seja, não são transcrições da fala. Acrescentemos ainda o fato de que por terem sido produzidas, em sua maioria, por missivistas cultos, estão sujeitas às pressões da norma-padrão, além de estarem materializadas no registro escrito, dinamizador das suas próprias convenções articuladas ao gênero textual “carta pessoal”.

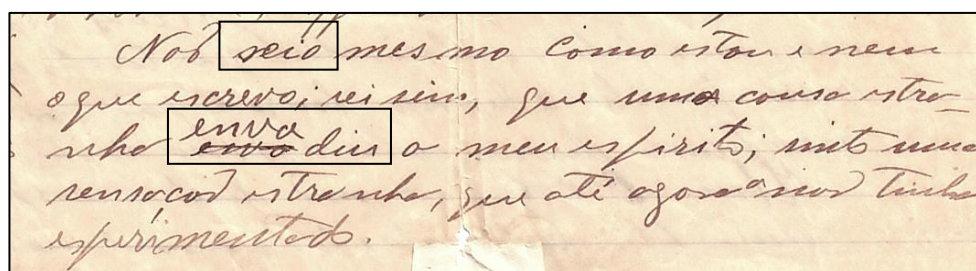
Por outro lado, o fato de serem cartas pessoais, refletindo a intimidade do escrevente, tendem, em essência, a minimizar essa forte pressão da norma-padrão. Assim sendo, as cartas pessoais constituem um contexto propício às oscilações entre o padrão e o não-padrão quer por descuido, quer por hipercorreção. No excerto exposto em (1), temos evidências que parecem apontar para a *autenticidade* da amostra. Nele, o escrevente JMF assumiu o item lexical “*seio*” no lugar da forma verbal “*sei*” e, na linha

⁴ Princípio oriundo da Geologia e aplicado à Sociolinguística, pelo qual se pode inferir que determinado fenômeno observado no presente pode refletir sua ocorrência no passado. (MILROY 1992 *apud* SOUZA 2021, p. 40.)

seguinte, segue a forma verbal “*sei*”, evidenciando a oscilação entre a grafia orientada pela norma-padrão e a sua expressão vernacular na língua falada. O mesmo escrevente, no mesmo fragmento, preferiu “*esvadiu*” no lugar da forma “*invadiu*” e, ao perceber o equívoco, corrigiu para “*envadiu*”, mas preserva a sílaba inicial “*en*” no lugar de “*in*” como um possível indício de *hipercorreção*. Além disso, temos o item lexical “*esperimentado*” grafado com “*s*” no lugar da letra “*x*” (“*experimentado*”), como também observamos no excerto (1).

Imagem 1: Evidência de aspectos ortográficos e de hipercorreção.

JMF. Tatu [Oeiras-PI], 19.08.1920.



(1) “[...] Não sei mesmo como estou e nem o que escrevo; sei sim, que uma coisa estranha *esva* <enva^{diu} o meu espirito; sinto uma sensação estanha, que até agora não tinha experimentado. [...]” JMF. Tatu [Oeiras-PI], 19.08.1920.

Convém esclarecer que o processo de transcrição das cento e trinta e seis (136) cartas que embasam esta análise deu-se à luz das normas propostas por Rumeu (2013), de modo que os “erros”, vacilações e retificações dos missivistas estão devidamente conservados tais como foram registrados pelos redatores. Em relação às amostras históricas, detivemo-nos sobre elas nas seções 2.1.1, 2.1.2, 2.1.3, 2.3, 2.4 e 2.5.12 desta Tese, mas já esclarecemos sobre a quantidade de cartas manuscritas e datiloscritas e o tratamento dado a elas no que se refere ao tipo de edição. Ocupamo-nos por conservadoramente editar as cartas piauienses em análise, através de uma edição fac-similar e diplomático-interpretativa (SPINA, 1970) a partir de uma ficha documental composta não só pelos dados biográficos dos informantes (origem e idade), mas também por um breve resumo do conteúdo da carta pessoal (RUMEU, 2013), cf. exposto na imagens 2 e 3.

Imagem 2: Ficha de apresentação das cartas piauienses.

Acervo: Pessoal
 Local e data da carta: Fazenda Tatu, Piauí, 19.08.1920.
 Autoria: JMF.
 Local e data de nascimento do(a) autor(a): Canto Alegre, Oeiras-PI, 17.12.1889.
 Nacionalidade: Brasileiro.
 Naturalidade: Piauiense.
 Idade (na ocasião da escritura desta carta): 31 anos.
 Conteúdo: Carta de J.M.F a C.R.J.M, pedindo-a em namoro.

Ilustramos, imagem 3, o formato da edição fac-similar e diplomático-interpretativa das cartas piauienses composta pelo fac-símile da carta ao lado da sua conservadora transcrição fôlio a fôlio, nos moldes de Spina (1970) e de Rumeu (2013). Nesta Tese, não trazemos à cena a publicação da edição das cartas piauienses analisadas, uma vez que não obtivemos autorização de todos os redatores (alguns deles se encontram vivos) e/ou dos seus familiares para a ampla divulgação (via publicação) das suas cartas. Reforçarmos o fato de a opção por expormos os nomes dos redatores através de suas siglas está orientada pelo intuito de preservação das suas identidades como uma prática das análises linguísticas em que buscamos sempre preservar a identidade dos sujeitos da pesquisa.

Imagem 3: Carta de amizade piauiense. ABSB. RJ, 06.06.1946.

	<p>Local e data da carta: Rio de Janeiro-RJ, 06.06.1946 Autor: Local e data de nascimento do(a) autor(a): Oeiras-PI, 21.05.1907. Nacionalidade: Brasileiro. Naturalidade: Piauiense. Idade (na ocasião da escritura desta carta): 39 anos. Conteúdo:</p> <p>[fól. 1a]</p> <p>Rio de Janeiro, 6 de Junho de 1946 Possidonio: Sem carta nenhuma a responder-lhe. O fim principal desta é pedir a você que me mande a dimensão da casa que foi do Visconde da Parnaíba. O pedido é para satisfazer a um amigo que está a publicar, aqui, um livro sobre essa figura histórica da nossa invicta Oeiras, e, daqui do Piauí. A dimensão de que falo é a quantidade de metros de frente da casa e a quantidade de metros de cada lado. O amigo nesta carta falada pediu-me também uma fotografia e descrição da casa, como eu achasse difícil, pois em nossa Oeiras talvez não haja atualmente fotografos, eu me comprometi apenas quanto a uma parte: a dimensão, e, se possível, no momento nada garantir ao mesmo amigo) alguns dados sobre a divisão interna, altura, etc, tudo da dita casa. Caso seja possível de-me resposta por telegrama. Do saldo que tenho em seu poder peço-lhe tirar a quantia de trezentos cruzeiros (Cr\$ 300,00) e entregá-la ao meu vaqueiro, José Bispo Ribeiro. Lembrança aos que lhe são caros e um abraço de P.S. Haverá algo a você me informar sobre a compra da posse de terra que estava tratada? Segundo conclusões que tirei numa carta sua de julho de 1945, o vende[do] no corrente ano completa 21 anos.</p>
--	---

Passamos por esse breve esclarecimento acerca do processo de edição das missivas pessoais piauienses em análise, mas retomamos, na sequência, os critérios de

levantamento e de seleção dessas cartas, voltando-nos especificamente à *autoria* de tais fontes históricas.

2.1.2 A autoria das cartas piauienses

Quanto à *autoria* das missivas piauienses em análise, devemos destacar que, se um dos objetivos da SH é identificar traços da língua em sincronias passadas e estabelecer uma correlação entre esses traços e a estrutura social através da qual as mudanças linguísticas também se implementam, convém que o pesquisador tenha certeza sobre quem escreveu os textos em análise, o que legitima a sua preferência pelo trabalho com textos autógrafos. Esse cuidado se deve também ao fato de que, considerando os efeitos do analfabetismo também no Piauí, era comum, em sincronias passadas, a prática de cartas ditadas (testemunho idiógrafo), o que requer uma atenção especial por parte do linguista-pesquisador.

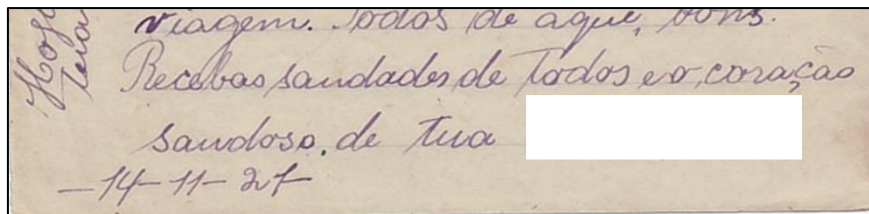
Todas as cartas piauienses novecentistas em análise são produzidas (manuscritas e datiloscritas) e assinadas pelo próprio punho do autor (autógrafas) que a redigiu. No caso das cartas manuscritas passamos pela análise comparativa entre a letra do redator no corpo da sua carta e a da sua assinatura⁵, o que permitiu também a comparação, por outro lado, tão somente com a assinatura dos autores nas cartas datiloscritas⁶. Nas imagens expostas de (4) a (9) correspondentes aos excertos (2), (3), (4) (5), (6) e (7) temos evidências das cartas produzidas por punhos autógrafos que se mostram manuscritas e datiloscritas, cabendo ao autor destas (datiloscritas) a *autoria intelectual* como observamos para as cartas de AVV (imagens 6 e 7) e de PNQ (imagens 8 e 9).

⁵ Cartas de JMF, CM, BB, ATT, JRC, JSL, LJL, JPB, VV, MNRSR, VB.

⁶ Cartas de PQ e de AVV.

Imagem 4: Excerto da carta de CRJM ao noivo.

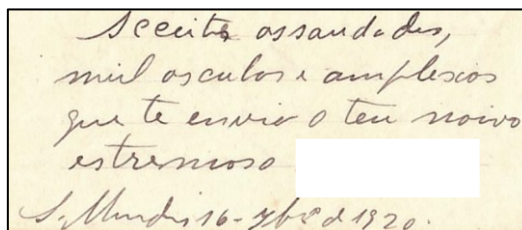
JMF. S.I, 14.11.1921.



(2) “[...] Recebas saudades de todos e o coração saudoso. de tua [...]” 14-11-21

Imagem 5: Excerto da carta do noivo à noiva.

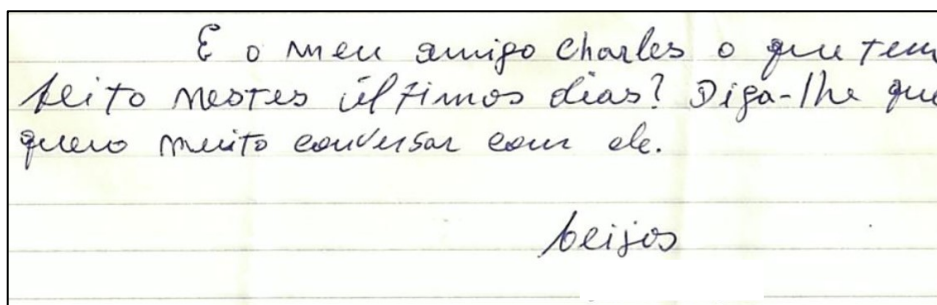
JMF. Simplicio Mendes - PI,16.09.1920.



(3) “[...] Aceite as saudades, mil osculos e amplexos que te envia o teu noivo extremoso. [...] Simplicio Mendes 16-setembro de 1920.” JMF. Simplicio Mendes-PI,16.09.1920.

Imagem 6: Excerto da carta de AVV.

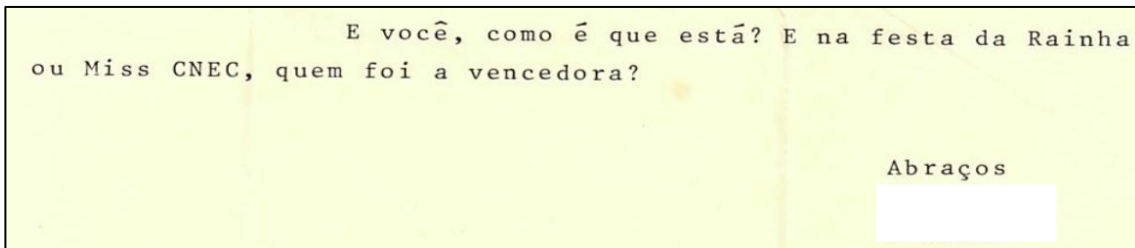
Brasília, 23.09.1981.



(4) “[...] E o meu amigo Charles o que tem feito nestes últimos dias? Diga-lhe que quero muito conversar com ele. beijos [...]” AVV. Brasília, 23.09.1981.

Imagem 7: Excerto da carta de AVV.

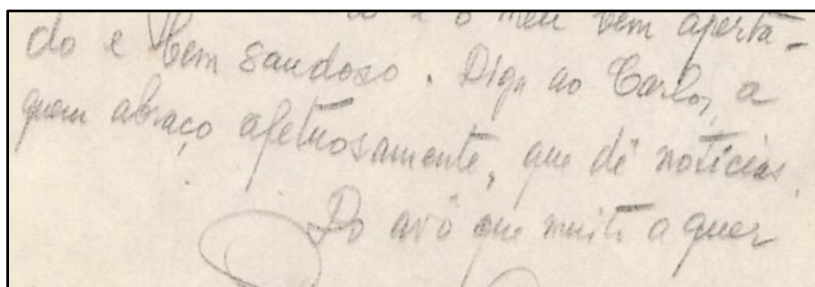
Brasília, 30.10.1984.



(5) “[...] E você, como é que está? E na festa da Rainha ou Miss CNEC, quem foi a vencedora? Abraços [...]” Brasília, 30.10.1984.

Imagem 8: Excerto da carta de PNQ.

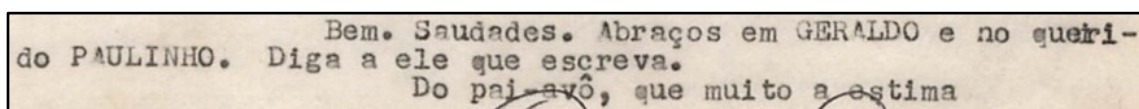
Oeiras, 20.04.1985.



(6) “[...] Diga ao Carlos, a quem abraço afetuosamente, que dê notícias. Do avô que muito a quer [...]” PNQ. Oeiras, 20.04.1985.

Imagem 9: Excerto da carta de PNQ.

Oeiras, 10.05.1985.



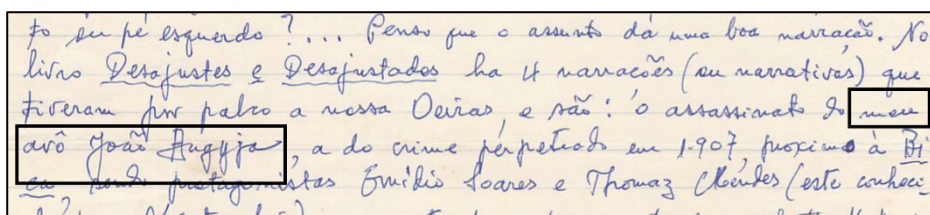
(7) “[...] Bem. Saudades. Abraços em GERALDO e no querido PAULINHO. Diga a ele que escreva. Do pai-avô, que muito a estima. Possidônio Queiroz” PNQ. Oeiras, 10.05.1985.

O rigor metodológico em relação não só à *autenticidade*, mas também à *autoria*, exige do linguista-pesquisador a busca pelo levantamento dos perfis biográficos dos redatores, a fim de comprovar que se trata de escreventes brasileiros, nascidos, no caso específico desta análise, no contexto histórico-social do Piauí. Ainda que a reconstituição

dos perfis sociais dos redatores seja uma tarefa trabalhosa, é necessária, tendo sempre em vista o fato de que “geralmente sabemos muito pouco sobre a posição social dos informantes, e não mais acerca da estrutura social da comunidade” (LABOV, 1994). Além das próprias cartas terem servido com fontes que nos permitiram confirmar informações biográficas dos redatores (imagem 8), valemo-nos não só de outras fontes documentais (fontes secundárias) tais como publicações em jornais, índices genealógicos, mas também de testemunhos orais de familiares e amigos nesse processo de reconstituição do perfil biográfico de alguns dos redatores piauienses (todos nascidos no espaço geográfico do Piauí) em análise. Na imagem 10 (excerto 8), temos o testemunho do redator ABSB acerca da referência ao assassinato do seu avô “JB”, além de, no excerto 9 (imagem 11) desta mesma carta, o redator referir-se ao seu interlocutor (PQ) como “confrade, amigo e conterrâneo”, confirmando se tratar de cartas genuinamente produzidas por escreventes nascidos no espaço geográfico do Piauí, mais precisamente, em Oeiras, ou melhor em “nossa Oeiras”, nos termos de ABSB.

Imagem 10: Excerto da carta de ABSB.

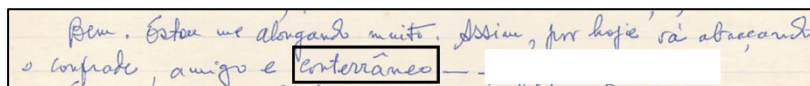
Piauí-PI, 31.01.1981.



(8) “[...] No livro Desajustes e Desajustados há 4 narrações (ou narrativas) que tiveram por palco a nossa Oeiras, e são: ‘o assassinato do meu avô João Bugyja [...]’ ABSB. Fazenda Riacho Fundo, São Pedro do Piauí-PI, 31.01.1981.

Imagem 11: Fechamento da carta de ABSB.

Piauí-PI, 31.01.1981.



(9) “[...] Bem. Estou me alongando muito. Assim, por hoje vá abraçando o confrade, amigo e conterrâneo [...]” ABSB. Fazenda Riacho Fundo, São Pedro do Piauí-PI, 31.01.1981.

Temos, nesta tese, escreventes cujos níveis de escolarização alcançam redatores tão somente alfabetizados, passando ainda por aqueles com o nível médio e com o nível universitário. Ao descrevermos as amostras das cartas correlacionando-as aos acervos que

representam, passamos, na seção 2.5.12, a uma descrição mais pormenorizada dos perfis biográficos dos redatores e das suas funções sociais, tendo em vista a descrição das amostras históricas em análise.

Em síntese, foi-nos possível contar, nesta Tese, com textos manuscritos (autógrafos) e datiloscritos datados e produzidos por punhos de redatores nascidos em terras piauienses, além de se tratar de um material escrito, muito bem preservado, de compreensão e de transcrição relativamente fáceis.

2.1.3 A validade social e histórica das cartas piauienses

Em se tratando da *validade social e histórica*, segundo Hernández-Campoy & Schilling (2012, p. 70), temos em análise um parâmetro também voltado ao conhecimento que se tem acerca da posição social dos escreventes e da estrutura social da comunidade linguística da qual faziam parte, o que legitima o trabalho de reconstrução dos perfis sociais dos informantes para que se consolidem acertadamente as interpretações relacionadas aos padrões de variação e mudança em sincronias passadas (MILROY & GORDON, 2003 *apud* HERNÁNDEZ-CAMPOY & SCHILLING, 2012). Assim sendo, a identificação da estrutura social será tão mais próxima da realidade pretérita, na proporção direta da certeza que o linguista-pesquisador tenha acerca da *autoria* e da *autenticidade* das fontes históricas.

2.2 A Teoria do Poder e da Solidariedade

A heterogeneidade social presente nas sociedades é terreno fértil para que os seres humanos, dependendo da posição social que ocupem, reivindiquem e exerçam relações “poder” ou de “solidariedade”. As relações de poder e solidariedade deixam-se evidenciar também através das escolhas pronominais de 2SG. Na perspectiva de Brown & Gilman (1960), as *relações assimétricas ascendentes* (de inferior para o superior) e *descendentes* (de superior para o inferior) estão associadas à Semântica do Poder. Por outro lado, as *relações simétricas*, marcadas por grupos sociais que estão em um mesmo nível hierárquico estão associadas à Semântica da Solidariedade.

As *semânticas do Poder* e da *Solidariedade* se deixam inferir a partir das referência ao sujeito de 2SG, já que essas escolhas por formas treatmentais específicas espelham,

em distintos níveis, uma dinâmica de hierarquização social. Brown & Gilman (1960) compreendem o conceito de semântica como sendo a “covariação entre o pronome usado e a relação social existente entre falante e destinatário”, traduzido do original: “*covariation between the pronoun used and the objective relationship existing between speaker and addressee*” (BROWN & GILMAN 1960, p. 252). O sistema pronominal latino dispunha de duas formas nominativas de 2ª pessoa: *tu* (singular) e *vos* (plural). A partir do século IV, com a reforma do imperador Diocleciano, o *vos*, incorporando a função gramatical de 2SG, passou a ser usado como forma de deferência exclusivamente ao Imperador, como um recurso de polidez e distanciamento entre classes sociais (BROWN & GILMAN 1960, p. 254). A relação de assimetria social como expressão *Semântica do Poder* ganhou força entre os séculos XI e XIV⁷. Segundo Wardhaugh (1997 *apud* RUMEU 2011, p. 117)⁸,

[...] em algumas sociedades contemporâneas, a evolução do sistema conceptual *Tu/Vous* (*tu/vós*) se direciona de um assimétrico *Tu/Vous* para o polido e simétrico *Vous/Vous* e para o mútuo e simétrico *Tu/Tu* em virtude da relevância da Solidariedade nas sociedades em geral. [...] (RUMEU 2011, p. 117)

A partir dessas formulações teóricas, é possível perceber que as correlações de força presentes nas sociedades se refletem nos usos da língua, sobretudo no sistema pronominal de referência ao sujeito de 2ª pessoa do discurso. Neste estudo, um dos objetivos é justamente a detecção das relações de Poder e Solidariedade através das escolhas pronominais de 2SG variáveis travadas entre os redatores piauienses (*tu versus você*), uma vez que sobre cada escolha pronominal repousa uma estreita relação com as dimensões do Poder e da Solidariedade. (BROWN & GILMAN 1960, p. 252).

Em síntese, entendemos que para a reconstrução de uma parte da história do português registrada nas cartas novecentistas do Piauí, é importante embasarmos este estudo tanto nos parâmetros da SH (CONDE SILVESTRE, 2012), para a organização das amostras históricas, quanto na *Teoria do Poder e da Solidariedade* (BROWN & GILMAN, 1960) para a interpretação da alternância *tu/você* conduzida pelas relações sociais que se deixam entrever na produção escrita piauiense. De acordo com Rumeu (2011, p. 118), quando se tenta interpretar as relações sociais supostamente estabelecidas entre remetente e destinatário de cartas, com base na *Teoria do poder e da solidariedade*,

⁷ BROWN & GILMAN (1960, p. 255).

⁸ WARDHAUGH (1997 *apud* RUMEU 2011, p. 117).

interpretamos indícios que nos conduzem, em algum nível, à reconstituição da história social do PB em sincronias passadas. Assim, a presente análise procura alcançar essas relações sociais que se manifestam a partir dos usos das formas pronominais utilizadas pelos missivistas piauienses.

2.3 O trabalho com o gênero *carta pessoal* e os seus subtipos: amostras linguísticas de sincronias passadas.

A *carta*, ao transitar por vários meios e suportes, assumiu no decorrer do tempo distintas funções comunicativas, o que conduziu a uma ampla variedade de subgêneros. Em termos semânticos, a *carta* se concretiza como uma enunciação sem a presença do interlocutor *in loco*, cf. Bakhtin (1997 *apud* LOPES *et alii*, 2018), tendo sempre em vista a finalidade da carta, a dinâmica emissor-receptor, o lugar, o momento de escritura, o suporte material de confecção da carta e a sua organização textual (MAINGUENEAU, 2001 *apud* LOPES *et alii* 2018, p. 39).

A *carta* está presente em distintos domínios de interação comunicativa, podendo se apresentar nos níveis oficial, comercial e pessoal. Considerando que a *carta*, por assumir diversificados propósitos, pode ser entendida como um *hipergênero*. Assim sendo, entendemos, à luz de Lopes *et alii* (2018), a *carta pessoal* como um hiperônimo que abrange os subgêneros *carta de amizade*, *carta amorosa*, e *carta familiar*.

A composição estrutural da carta pessoal é composta pelas seguintes partes constitutivas, considerando a tradição medieval dos clérigos, cf. HAROCHE-BOUZINAC (2016, p. 33): *salutatio* (saudação), *benevolentiae captatio* (captação da benevolência), *narratio* (narração), *petitio* (pedido ou objeto da mensagem) e *conclusio* (conclusão). Acrescentemos o fato que, no âmbito dessa composição estrutural está a relação estabelecida entre o emissor e o destinatário, desencadeando maior ou menor grau de cumplicidade, de afetividade, de expressividade entre o redator e o seu interlocutor, o que repercute diretamente, por exemplo, na escolha da forma nominal de tratamento do interlocutor, contextualizada por uma dada situação comunicativa da carta. Na perspectiva de Cícero, a *carta* é definida como “*familiare et jocosum* (“familiar e jocoso”), por oposição a outro tipo de carta mais solene, *severum et grave* (“severo e grave””, cf. Haroche-Bouzinac (2016, p. 40). Nessa perspectiva, Haroche-Bouzinac (2016) segue com a sua interpretação sobre a noção de *carta familiar* com a qual concordamos, sobretudo, por se aplicar às cartas piauienses em análise.

A carta é considerada familiar em virtude de seu vocativo, mas isso implica que o tom seja necessariamente familiar. Convém uma forma respeitosa no círculo familiar, particularmente nas cartas escritas aos mais velhos. [...] Ao contrário das cartas protocolares, que devem adotar um esquema respeitando as leis da troca hierárquica, ou das cartas de mercadores e da carta comercial contemporânea, que se situa na ordem de uma prática de tratativas e trocas de garantias, a carta familiar beneficia-se da liberdade permitida pela proximidade dos vínculos. O termo “familiar” sugere igualmente o surgimento do hábito de frequência. (HAROCHE-BOUZINAC 2016, p. 40)

Na imagem 12, observamos uma carta trocada entre irmãos, no início da 2ª metade do século XX. No vocativo “Perzado Irmão Antonio a Deus”, verificamos a menção ao laço familiar norteador da relação interpessoal entre o remetente e o seu destinatário, o que é corroborado também pelo desejo do redator de que o seu irmão venha a visitá-lo “na paz e feliz” (“eu espero você e todos na pais e felix”).

Imagem 12: Excerto da carta de VVS.

São Caetano do Sul, 30.07.1953.

São Caetano do Sul 30 de 7 de 1953

Saudação

Perzado irmão Antonio a Deus

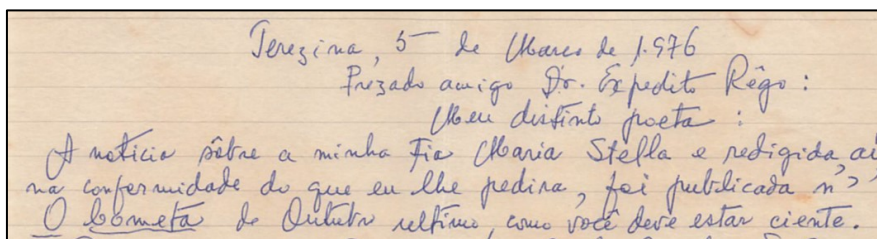
São este os meus gante votos e pego
napena e so minte para dartis as
minha notuça e como tãbem
Receber as tu notucio eu Vou gosando
gante felecedade e gal minte a Deus
eu espero você e todos na pais e felix

(10) “São Caetano do Sol 30 de ‘7’ de 1953 Saudaçõe Perzado Irmão Antonio a Deus São este os meus gante votos e pego napena e so minte para dartis as minha notisça e como tãbem Receber as tu notiscia eu vou gosando | gante felecedade e galmente a Deus eu espero você e todos na pais e felix. [...]” VVS. São Caetano do Sul, 30.07.1953.

Além das cartas familiares piauienses, temos, nas cartas de amizade, evidências de relações de amizade entre os interlocutores que se tratam como confrades de uma mesma academia de letras, colegas de infância, conterrâneos, o que nos está revelado também através do *vocativo*. Na imagem 13, ilustramos a relação de amizade travada, na 2ª metade do século XX, entre ABSB e Expedito Rêgo.

Imagem 13: Excerto da carta de ABSB.

Terezina, 05.03.1976.

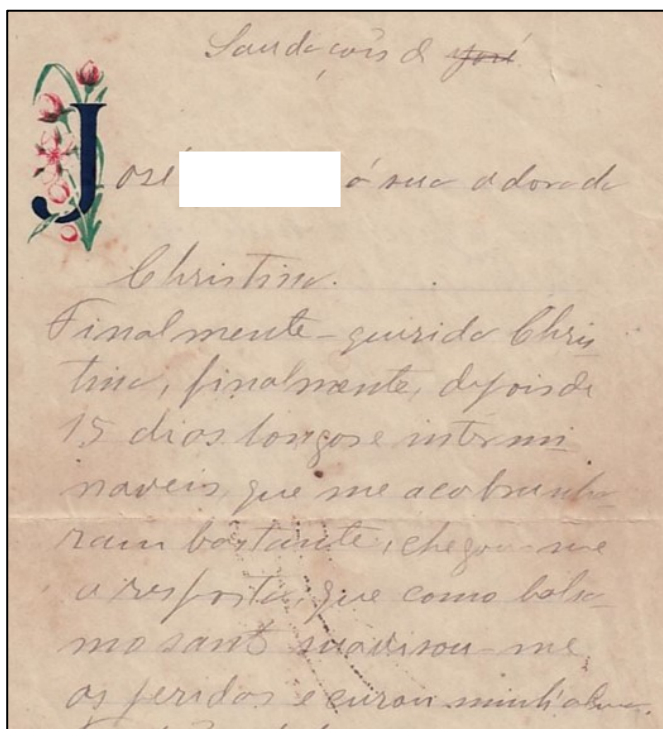


(11) “Terezina, 5 de Março de 1976 Prezados amigos Dr. Expedito Rêgo: Meu distinto poeta: A notícia sobre a minha tia Maria Stella e redigida aí, na conformidade de que eu lhe pedira, foi publicada n.º 1 O Cometa de Outubro ultimo, como você deve estar ciente. [...]” ABSB. Terezina, 05.03.1976.

As cartas *amorosas* em análise também trazem, no vocativo, a menção à relação amorosa em si, deixando-nos corroborar, no corpo da carta, a relação de intimidade travada entre os interlocutores. À título de exemplificação, apresentamos, na imagem 14, um excerto da correspondência trocada entre JMF e CRJM, na 1ª metade do século XX (1920), na qual temos a abertura da carta a partir do chamamento “Saudações de José | José Moura á sua adorada | Christina” com a expressão explícita da relação de vassalagem amorosa entre o redator e a sua amada.

Imagem 14: Excerto da carta de JMF.

Simplicio Mendes, 06.09.1920.



(12) “Saudações de José José [...] á sua adorada Christina Finalmente – querida Christina, finalmente, depois de 15 dias longos e interminaveis, que me acabrunharam bastante, chegou-me a resposta, que como balsamo santo suavizou-me as feridas e curou minh’ alma. [...]” JMF. Simplicio Mendes-PI, 06.09.1920.

Visto que apresentamos em que consiste o gênero textual *carta pessoal* e os seus subtipos (amorosas, familiares e de amizade), passamos, na sequência, a uma breve descrição panorâmica, em termos quantitativos, das cartas piauienses em análise nesta tese.

2.4 Uma breve exposição das amostras de cartas pessoais piauienses novecentistas: quadro geral.

Neste texto, temos em análise missivas pessoais produzidas por treze (13) redatores históricos com níveis de escolarização que variam do tão somente nível alfabetizado ao nível superior completo, assim distribuídos: três redatores (3) com nível superior; três (3) com ensino médio; um (1) autodidata (músico, historiador, professor, escritor, advogado

rábula) e seis (6) alfabetizados. Temos uma amostra bem diversificada no que tange à formação escolar dos redatores e às suas respectivas habilidades de leitura e escrita.

Expomos, no quadro 2, a distribuição das amostras de cartas pessoais, em seus subgêneros (*amorosas, amizade e familiares*), em relação aos sete (7) conjuntos documentais ⁹(“acervos”) e ao eixo do tempo (1920-1990).

Quadro 2: Amostras históricas piauienses: o século XX em cena.

ACERVOS	QUANTIDADE DE CARTAS EM RELAÇÃO AO SEXO DOS REDATORES		PERÍODOS	SUBGÊNERO DAS MISSIVAS				QUANTIDADE DE CARTAS
	H	M		AMOR	AMIZADE	FAMILIAR	AUTORES	
Acervo Autores piauienses	42	-	1946 - 1990	-	42	-	02	42
Acervo Batista	04	-	1078 - 1978	-	04	-	01	04
Acervo Cartas pessoais variadas	08	01	1953 -1976	-	07	02	05	09
Acervo Moura Fé	26	10	1920 - 1922	36	-	-	02	36
Acervo Queiroz	27	-	1975 - 1990	04	02	21	01	27
Acervo Santos Rego	-	07	1978 - 1981	-	07	-	01	07
Acervo Vieira do Vale	11	-	1981 - 1984	11	-	-	01	11
TOTAL	118	18	1920 - 1990	51	62	23	13	136

Nesta Tese, as amostras linguísticas são compostas por 136 cartas pessoais (*amorosa* (51), *familiar* (62) e de *amizade* (23)) produzidas por missivistas piauienses (homens e mulheres), entre os anos de 1920 e 1990. Como temos em evidência a produção escrita de redatores com distintos graus de habilidade escrita, voltamo-nos à dinâmica *tu/você* e a sua distribuição em relação aos níveis de escolarização e, conseqüentemente, de habilidade dos redatores piauienses (mais ou menos hábeis).

Tomada decisão de trabalhar com cartas pessoais (*amorosas, familiares, de amizade*), surgiu uma pergunta: onde encontrar cartas que atendam aos nossos objetivos

⁹ Acervos: Batista, Moura Fé, Queiroz, Santos Rêgo, Vieira do Vale, autores piauienses e cartas pessoais variadas.

da pesquisa? Contactamos colegas do curso de História da UFPI de quem recebemos um arquivo com inúmeras cartas oficiais digitalizadas cujas formas de tratamento do destinatário eram *Vossa Excelência*, *Vossa Reverendíssima*, *Vossa Senhoria*, dentre outros tratamentos formais. Como deparamo-nos com *cartas oficiais* conduzidas tão somente por formas nominais de tratamento, voltamo-nos aos acervos privados em busca de fontes para o levantamento de formas pronominais.

Contactamos o Museu *Ozildo Albano* (em Picos-PI) que cultivava parcerias com a Secretaria de Estado da Cultura, com Universidades da região e com escolas públicas e privadas, motivo pelo qual a Prefeitura Municipal de Picos o reconheceu oficialmente como Instituição de Utilidade Pública. O referido museu nos disponibilizou um acervo contendo mais de 600 (seiscentas) cartas pessoais todas já digitalizadas e algumas já transcritas. Ainda que se trate de interessantes amostras textuais, não as utilizamos, motivados principalmente por duas razões. São elas: a) nem todas as cartas eram comprovadamente produzidas e assinadas por punhos piauienses e b) não tínhamos como levantar rapidamente os dados biográficos dos escreventes.

Contactamos a Academia Piauiense de Letras, em Teresina, da qual recebemos a resposta de inexistência de cartas trocadas entre seus membros escritores. Entramos também em contato com o Arquivo Público do Piauí que se mantinha fechado ao público em virtude de estarmos no contexto mais crítico da Pandemia de COVID-19. Visitamos também o Centro Cultural Francelino de Sousa Araújo (Francelino Piauí), em Inhuma-Pi, mas nos deparamos principalmente com textos jornalísticos (artigos e reportagens), e com algumas poucas cartas familiares já publicadas que tinham sido submetidas ao crivo do editor do Jornal Correio Popular, em Campinas-SP, o que nos levou a não insistirmos nessa amostra.

Em conversas com uma ex-colega, já aposentada, comentei sobre a dificuldade de encontrar cartas. Seus olhos brilharam e ela exclamou: “Encontrei uma utilidade para as cartas de vovô *D* e de Vovó *C*!”, identificados, nesta análise, pelas siglas dos seus nomes (JMF e CRJM). Entregou-nos um pacote com as cartas, das quais selecionamos as que atendiam aos parâmetros de constituição de amostras confiáveis da SH. Contactamos também membros do Instituto Histórico e Geográfico de Oeiras-PI, por quem fomos informados sobre um rico acervo pertencente a um neto de um escritor local (PNQ), membro fundador do referido instituto. O referido neto de PNQ nos forneceu cartas pessoais redigidas por seu avô, endereçadas à esposa, aos netos e sobrinhos e aos amigos. Além dessas, havia outras cartas de dois outros escritores amigos de PNQ (ABSB e

JATF). Outro amigo historiador nos forneceu cartas de seus amigos que, quando jovens, moraram em São Paulo. Por fim, outras duas amigas de trabalho nos cederam algumas cartas de amor, familiares e de amizade.

Apesar das dificuldades relatadas acima, conseguimos compor um conjunto com 136 (cento e trinta e seis) cartas pessoais classificadas em relação aos subgêneros *amor*, *familiar* e *de amizade*, que foram conservadoramente transcritas tendo em vista o levantamento, a codificação e a análise estatísticas e probabilística a partir das potencialidades do RBrul e do RStudio. As referidas cartas estão distribuídas em 7 (sete) acervos pessoais e geraram 583 (quinhentos e oitenta e três) dados de *tu/você*, além de 23 (vinte e três) dados de *vós* em referência a 2SG, ou seja, *vós* em lugar de *tu*, que não foram computadas nas análises quantitativas, mas que comentamos qualitativamente no decorrer do capítulo de análise estatística dos dados.

2.5 A referência ao sujeito de 2SG: os parâmetros de uma análise linguístico-social

Com base nos parâmetros teórico-metodológicos da SH, concebemos a dinâmica *tu/você* como a regra variável em análise condicionada por variáveis linguísticas e extralinguísticas. Em um primeiro momento, fizemos o levantamento de cartas pessoais em museus e bibliotecas. Por conta do insucesso nessa busca, partimos aos acervos privados, de onde levantamos e editamos (edição fac-similar e diplomático-interpretativa) cento e trinta e seis cartas pessoais que evidenciam a produção escrita dos piauienses na era novecentista do PB. Na sequência, passamos ao levantamento das formas de referência ao sujeito de 2SG, com atenção voltada às formas pronominais *tu* e *você*. Os dados foram codificados por meio de uma planilha Excel, de modo que as formas pronominais de referência ao sujeito de 2SG constituem a *variável dependente (tu/você)*, que foram analisadas em função de variáveis independentes (internas e externas). São elas: (1) a *expressão do sujeito* de 2SG (nulo ou pleno); (2) a *pessoa verbal*; (3) o *paralelismo formal e semântico*; (4) os *subgêneros das cartas pessoais (cartas de amizade, amorosas e familiares)*; (5) *As temáticas prevalentes das cartas pessoais*; (6) o *sexo do informante*; (7) a *faixa etária* dos missivistas; (8) os *níveis escolarização dos informantes*; (9) os *níveis de prática de escrita*; (10) as *relações sociais* (simétrica e assimétrica); (11) os *períodos das cartas* e (12) os *acervos históricos*. A partir do cruzamento desse doze grupos de fatores alcançamos os resultados estatísticos (via excel) probabilísticos (via RBrul).

Apresentamos, a seguir, os critérios linguísticos e extralinguísticos, correlacionando-os às hipóteses e/ou objetivos motivadores.

2.5.1 As formas de referência ao sujeito de 2SG: *tu/você*.

Considerando o fato de, na fala piauiense, termos em evidência as formas *tu* e *você* (SCHERRE *et alii* 2021, p. 176) a partir da ¹⁰prevalência do *você* (92%), ainda que em convivência com *tu* sem concordância (8%), passamos à análise de tais formas pronominais, conjecturando que também o *você* já prevalecesse, na produção escrita novecentista piauiense. Em (13) e (14), ilustramos o *tu* e *você* nas cartas do Piauí em análise.

(13) “[...] *Achas* que eu não te amo nem a metade de que me *amas*. [...]” CRJM. Fazenda Tatu (Oeiras-PI), 21.05.1921.

(14) “[...] Dizem que os homens de inteligência são bons. *Você* confirma a sentença. [...]” JATF. Teresina-PI, 27. 07.1986.

2.5.2 A expressão do pronome-sujeito de 2SG: nulo/pleno.

A análise da *expressão do sujeito pronominal* de 2SG em suas formas *nula ou plena* está, neste estudo, orientada pela conjectura de que se o *você* leva o verbo para a 3SG o encaminha à sua expressão plena, em consonância inclusive com o histórico da sua origem nominal (*Vossa Mercê*), cf. Rumeu (2013). O *tu*, por outro lado, se mantém etimologicamente marcado pela concordância com a forma verbal munida por desinência verbal específica para a 2SG. De (15) a (17), temos evidências do *tu* e do *você* em suas expressões nula e plena. Ao voltarmos o foco à expressão (nula ou plena) dos pronomes-sujeito de 2SG (*tu* e *você*), utilizamo-nos do *zero* entre colchetes para a identificação do sujeito nulo.

(15) “[...] e *tu* o que *és*? Se não [0] *és* poeta para as outras [0] *és* para mim. [...]” CRJM. Fazenda Curral de Pedras-PI, 11.03.1921.

¹⁰ Resultados resgatados por Scherre *et alii* (2021) a partir da análise dos dados do “Atlas Linguístico do Brasil – ALiB: inquéritos geolinguísticos” para Teresina/PI (CARDOSO, 2008; CARDOSO & DEUS, 2009).

(16) “[...] Por isso [0] *fazes* bem Christina de min’halma [...] [0] *fazes* muito bem em só [0] *pensares* em mim, [...]” JMF. Simplício Mendes-PI, 11.10.1920.

(17) “[...] *Você* precisa estudar muito. [0] Precisa de aprender muito para vir a ser um homem preparado [...]” PQ. Oeiras-PI, 12. 03.1975.

2.5.3 A pessoa verbal (concordância)

Para a análise da pessoa verbal (concordância), assumimos a hipótese de que o *você*, por ter tido a sua origem na expressão nominal *vossa mercê*, tenda a ser favorecido por formas verbais de 3SG. O *tu* acionaria a concordância etimológica a partir das desinências número pessoais específicas da língua portuguesa para a identificação formal da referência ao sujeito de 2SG, o que está em conformidade com as propostas das análises de Souza (2012) e de Souza (2021), embasados em missivas cariocas e mineiras, respectivamente. Em (18) e (19), trazemos evidências das formas *tu* e *você*, respectivamente.

(18) “[...] Senti que *tu* me não eras indiferente como todas as outras virgens. [...]” JF. Simplício Mendes-PI, 24.04.1921.

(19) “[...] Eu também, como *você* sabe, estou fazendo aniversário no mez de Maio [...]” ABSB. RJ, 17.05.1988.

2.5.4 O paralelismo formal e semântico

Ao analisarmos a variável *Paralelismo formal e semântico*, assumimos a hipótese de que uma sequência linguística contextualizada por formas do paradigma de *tu* (*pronomes possessivos, clíticos, pronomes-complemento de 2SG*) favoreceria o *tu-sujeito*, cf. ilustramos de (20) a (22). Nesse mesmo sentido, entendemos que as formas do paradigma de *você* (*pronomes possessivos, clíticos, pronomes-complemento de 3SG*) contextualizariam o *você-sujeito* nas cartas piauienses, cf. verificamos de (23) a (25). Trata-se de uma variável linguística constantemente testada em análises no âmbito da morfossintaxe do português (OMENA, 2003; LOPES & VIANA, 2012) inspirados principalmente em Scherre (1998) que discute a força da variável “paralelismo linguístico” (fônico e formal e semântico) na análise de regras variáveis como, por exemplo, a *concordância* e o *imperativo*.

(20) “[...] Orgulhoso por te amar e merecer a [tua] afeição, por *seres* uma pessoa virtuosa” JMF. Simplício Mendes-PI, 16.09.1920.

(21) “[...] que com o teu convívio me corrija destes defeitos e [te] trate como *mereces*. [...]” JMF. Simplício Mendes-PI, 16.09.1920.

(22) “[...] [A ti] nada tenho que perdoar, pois nada me *fizestes*, [...] como *tiveste* coragem de dizer tal coisa, [...]” CRJM. Fazenda Curral de Pedras-PI, 11.03.1921.

(23) “[...] Eu [lhe] pedia um artigo [...] sobre o Pedro Britto - eis que *você* - como escritor brilhante e filho de Oeiras - fará (...)” BB. RJ, 16.03.1981.

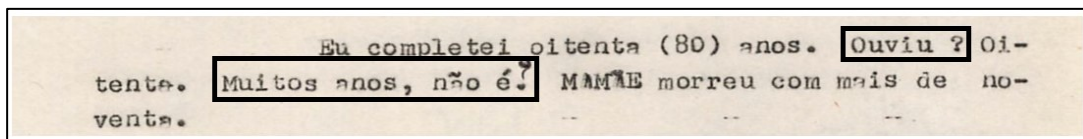
(24) “[...] O [seu] trabalho - caso *você* queira - pode ser ilustrado com um clichê da casa (...)” BB. Fazenda Riacho Fundo-São Pedro-PI, 04.02.1982.

(25) “[...] Pedia-[lhe] [...] que *você* obtivesse a transcrição [...]” BB. RJ, 25. 07.1988.

2.5.5 Os subgêneros das cartas pessoais: amorosa, amizade e familiar.

Entendemos, à luz de Barbosa (1999 *apud* RUMEU 2013, p. 56) que as cartas pessoais representam fontes históricas mais transparentes e livres da pressão da norma-padrão. Essa é uma preocupação do linguista-pesquisador que está voltado à apreensão do vernáculo em sincronias passadas. Nesse sentido, evidenciamos, através da análise de um excerto de uma das cartas de PNQ (Oeiras-PI, 22.05.1986), uma marca da oralidade do redator, ao se dirigir a sua neta (imagem 15).

Imagem 15: Evidência de marca de oralidade em carta *familiar* endereçada a uma de suas netas.
PNQ, Oeiras-PI, 22.05.1986.



(26) “[...] Eu completei oitenta (80) anos. Ouviu? Oitenta. Muito anos, não é? Mamãe morreu com mais de noventa. [...]” PNQ. Oeiras-PI, 22.05.1986.

No excerto da carta familiar exposto através da imagem 15, o missivista parece se comportar como se estivesse dialogando presencialmente com seu destinatário “*Completei oitenta (80) anos, ouviu?*”, “*muitos anos, não é?*”, o que mostra que, embora não se possa abstrair fielmente a fala a partir da escrita, esta pode nos trazer muitos dados importantes sobre os traços da oralidade em sincronias passadas. Nesse sentido, temos em análise a intimidade das relações sociais e a transparência em relação à expressão do vernáculo como traços peculiares ao gênero *carta pessoal*.

No âmbito das cartas amorosas, trazemos à cena as cartas cujos redatores mantêm uma relação amorosa com o seu interlocutor. Passamos, na sequência, a exemplificar, através de um excerto de uma carta amorosa (imagem 16), a correlação entre o pronome *você-sujeito* e o pronome possessivo *sua* (“em sua carta”), ambos em referência à 2SG. É interessante atentarmos ao fato de que uma das repercussões da inserção do *você* no sistema pronominal do PB está no nosso sistema de pronomes possessivos, ou seja, como evidência da atual norma objetiva do PB. No excerto transcrito, em (27) em relação à imagem 16, trazemos à discussão um clarividente dado de *seu* de 2SG (“receber uma carta *sua_i*, saber como *você_i* está”), como já discutido por Silva (2023) para as cartas pernambucanas (século XX), Lucena (2016) para as cartas cariocas (séculos XIX e XX), Barbosa (2018) para as cartas baianas (século XX) e Lopes & Cavalcante (2012).

Imagem 16: Excerto da carta de AVV. Brasília, 25.02.1983.

quilidade. Não deixa de ser bem gratificante receber uma carta sua, saber como você está, mas é bem melhor poder ver isto de perto, receber carta é muito bom, mas é melhor receber a pessoa que emite a carta, e, isto eu espero poder fazê-lo em breve.

(27) “[...] Não deixa de ser bem gratificante receber uma carta sua, saber como você está, mas é bem melhor poder ver isto de perto, receber carta é muito bom, mas é melhor receber a pessoa que emite a carta, e, isto eu espero poder fazê-lo em breve. [...]” AVV. Brasília, 25.02.1983.

As cartas de amizade mostram-se marcadas por relações de afetividade entre redator e interlocutor, de modo a ter total liberdade de poder buscar a veracidade ou não de um boato acerca do seu casamento, cf. observamos em missiva entre amigos a seguir exposta na imagem 17 (dado 28). Em termos de evidências linguísticas da norma objetiva do PB, temos a grafia do pronome pessoal “nós” a partir do desenvolvimento de um encontro vocálico (ditongo aberto) [oi] no lugar da sibilante [s], como está também clarividente através da sua grafia como “nóis”. Temos ainda o item lexical “boato” com o alteamento da vogal média através da sua grafia como “buato”, apontando para mais uma evidência grafofonética.

Imagem 17: Excerto da carta de LJJ.

Bocaina-PI, 22.06.1951.

Nêgo aqui corre um buato que você já se cazou nós pedimos que mande a realidade e se cazo fôr verdade que nós queremos ficar sabendo se é certo mesmo ou não. nêgo com relação ao que se falou não aguardo

(28) “[...] nêgo aqui corre um buato que você já se cazou nós pedimos que mande a realidade e se cazo fôr verdade que nós queremos ficar sabendo se é certo mesmo ou não. [...]” LJJ. Bocaina-PI, 22.06.1951.

Esta análise está embasada em cartas pessoais em seus subgêneros *amor*, *amizade* e *familiar*. Partimos da hipótese de que as cartas amorosas tendem a contextualizar o *íntimo tu* como já verificado por Souza (2012) e Silva (2012) para as cartas cariocas e por Souza (2021) para as cartas mineiras. As cartas de *amizade* e *familiares*, por outro lado, tendem a ser protagonizadas pelo *você* que é uma estratégia neutra, ou seja, não tão íntima e invasiva como o *tu*, mas também não tão distante quanto a forma nominal de tratamento originária (*vossa mercê*). Assim sendo, acreditamos que o monitoramento das cartas piauienses também em relação ao subgênero da carta pessoal possa nos oferecer um panorama da produtividade da regra variável *tu/você*, restando-nos descobrir os níveis da variação pronominal.

2.5.6 As temáticas prevalentes das cartas pessoais piauienses

Optamos ainda por testar/controlar os assuntos prevalentes nas cartas, considerando que as cartas piauienses em análise passaram por questões *sentimentais (stricto sensu)*, *políticas e/ou econômicas* (poder, dinheiro), *familiares* (saúde, compras, mensalidade escolar, conselhos), *literárias* (acervos, doações de livros, comentários sobre produção literária) e por questões de viagens e festas (convites). O intuito é não só averiguarmos se as temáticas das cartas pessoais tenderiam, em alguma medida, a impulsionar as formas *tu* e *você* para a referência ao interlocutor, mas também testarmos a hipótese de a temática das cartas amorosas mostrar-se como o contexto de resistência do *tu*, como constatado por Pereira (2012) para a *família Penna* e por Silva (2012) para o casal de namorados *Jaime e Maria* em relação ao contexto das cartas amorosas. Nesse sentido, detectamos questões sentimentais, políticas e/ou econômicas, familiares, literárias e questões de viagens e festas. Passamos à exposição das temáticas em análise, respectivamente expostas de (30) a (34).

Cartas de temática amorosa: questões amorosas (sentimentais entre casais *stricto sensu*).

(30) a. “[...] Orgulhoso por te amar e merecer a tua afeição, por *seres* uma pessoa virtuosa [...]” JMF. Simplicio Mendes-PI, 16.09.1920

b. “[...] recebi apenas duas cartas suas, apesar dos meus apelos para que *você* me escreva [...]” AVV. Brasília, 29.09.1982.

Cartas de temática familiar: questões familiares (saúde dos familiares, compras, mensalidades (escolar), conselhos).

(31) a. “[...] *Sabes* Prima, as coisas aqui em casa andaram tão ruins que pensei até que não seria possível mesmo consertá-la, mas graças ao nosso Amigo lá de cima, entramos numa boa fase. [...]” MNRSR. Brasília, 08.10.1978.

b. “[...] Quando *você* vier a Oeiras, todos ficaremos alegres, se soubermos sempre que *você* foi um menino bom [...]” PNQ. Oeiras-PI, 12.03.1975.

Cartas de temática literária: acervos, doações de livros, comentários sobre a produção literária.

(32) a. “[...] Estou de posse da Revista (número Especial do 30º aniversário), ora circulando, do Instituto Histórico e Geográfico do Rio de Janeiro, que terá ~~de~~ de ser enviada para você, por estes dias. O número destina-se à Biblioteca do nosso Instituto Histórico, caso *você* queira. [...]” ABSB. RJ, 12.12.1987.

b. “[...] Eu gostaria de combinar uma cousa com você: *você* remeter-me somente o seu discurso de saudação do dia 21 de Maio. Uma reportagem descrevendo outras peças ou números da sessão do referido dia 21 de Maio [...]” ABSB. RJ, 20.07.1987.

Cartas de temáticas política e/ou econômica

(33) a. “[...] Olha! *tu* não está sendo como nós eramos, porque está escondendo as novidades, mais *tu* pensa que não sei? [...] Sim como vai a política por aí, está quente mesmo, e o Antonio de Aristide passou mesmo para cristovão e o nezinho está mesmo no seu partido, mande pelo menos um pequeno relato. [...]” JPB. São Bernardo do Campo-SP, 18.04.1978.

b. “[...] Valdemar quero que fique bem claro que não quero receber rendas quero apenas que *você* cuide desse pequeno patrimônio que nosso pai deixou. [...]” JRC. Santo André-SP, 23.12.1976.

Cartas voltadas às questões de viagens e festas (convites)

(34) a. “[...] Ontem saí com o Hilario (ele está aqui), fomos num lugar delicioso que chama-se “Taberna”, lá, quem gosta, bebem numa boa, e quem é romântica como eu, pode curtir uma música ao vivo, com violão e uma voz muito bonita. É um lugar muito bom mesmo, qdo *pintares* por estas bandas, prometo uma noitada lá, certo! [...]” MNRS. Brasília, 20.08.1978.

b. “[...] Otacília me disse que *você* fez uma promessa: a de vir passar uns dias conosco, em julho.” PNQ. Oeiras-PI, 21.04.1987.

2.5.7 O sexo e a faixa etária dos redatores

Pretendemos observar se o gênero do escrevente se mostrou ou não relevante no processo de variação entre as formas *tu/você* nas missivas piauienses do século XX. Em relação às cartas cariocas, Rumeu (2013), através da análise de cartas cariocas oitocentistas e novecentistas, observou que as mulheres se mostraram como as propulsoras do inovador *você* que tende a figurar como uma forma mais neutra e menos invasiva que o íntimo *tu*.

Voltando o foco ao PB em sua expressão contemporânea falada, trazemos os resultados de Scherre *et alii* (2015 *apud* Souza 2021, p. 75) que identificam o inovador *você* como a forma prevalente para a referência ao sujeito de 2SG entre os jovens, dialogando com os resultados de Rumeu (2013, p. 165) nas cartas cariocas oitocentistas e novecentistas. Teríamos, nas cartas em análise, esse mesmo comportamento por parte das redatoras piauienses? Passamos a apresentar algumas evidências de dados de *tu* e *você* entre os redatores (homens e mulheres) jovens (de 14 a 30 anos), adultos (de 31 a 50 anos) e idosos (mais de 50 anos), cf. verificamos de (35) a (39). Em relação às mulheres, acrescentemos o fato de tão somente termos levantado dados de formas pronominais de 2SG em cartas de redatoras jovens e adultas, cf. dados (38) e (39).

Homem jovem

(35) a. “[...] Se *chegares* em Picos antes do dia 2 ou 3, *ligues* pra mim nos dias mencionado [...]” VB (redator com 24 anos). Fortaleza-CE, 25.06.1978.

b. “[...] Sabendo eu de que *você* não pretende mais morar aqui peço ao amigo que chêgue até aqui para *você* deslindrar suas partes de terra com Antonio seu irmão [...]”. LJL (redator com 27 anos). Bocaina-PI, 22.06.1951.

Homem adulto

(36) a. “[...] Orgulhoso por te amar e merecer a tua afeição, por *seres* uma pessoa virtuosa afavel, carinhosa, meiga e terna, e, vaidoso por *seres* um thesouro de beleza. [...]” JMF (redator com 31 anos). Simplicio Mendes-PI, 16.09.1920.

b. “[...] Casualmente, domingo fui a casa de um colega e quando voltei soube que *você* havia ligado [...]” AVV (redator com 34 anos). Brasília, 28.02.1984.

Homem idoso

(37) a. “[...] Se *queres* vencer, desperta e luta. [...]” PNQ (redator com 81 anos). Oeiras-PI, 18.07.1987.

b. “[...] Se *você* puder vir e, bem assim algum colega, será uma satisfação para mim. [...]” ABSB (redator com 69 anos). Fazenda Riacho Fundo, São Pedro, PI, 06.05.1977.

Mulher jovem

(38) a. “[...] Se cancou *tens* razão, porque eu não sou digna de tanta *finesa* que me *despenças*. [...]” CRJM (redatora com 19 anos). Fazenda Tatu-PI, 09.01.1921.

b. “[...] *Você* pensou que fosse dia 09/05, né, errou só por 4 dias, faço aniversário é hoje. (...). MNRSR (redatora com 23 anos). Brasília, 05.05.1978.

Mulher adulta

(39) “[...] Comadre, por aqui nada de importante tudo como sempre só muito trabalho e pouco lucro, pois *você* sabe como estão as coisas. [...]” JSL (redatora com 33 anos). Santo André-SP, 15.02.1964.

2.5.8 Os níveis de escolarização dos redatores piauienses

Em análises sociolinguísticas voltadas às sincronias passadas, o linguista-pesquisador volta-se não só ao nível de contato dos redatores com os modelos de escrita, mas também ao seu nível de escolarização formal. Nesse sentido, passamos a um quadro resumitivo (quadro 3) a partir do qual correlacionamos os níveis de escolarização formal dos redatores piauienses aos seus papéis sociais com o intuito de trazermos luz a questão da interferência dos aspectos sociais na expressão linguística.

Quadro 3: Quadro-síntese da correlação entre os níveis de escolarização e os perfis sociais dos redatores.

ACERVOS (1920-1990)	REDADORES	NÍVEIS DE ESCOLARIZAÇÃO	PAPEIS SOCIAIS (PROFISSÃO PRINCIPAL)
1. Acervo Autores piauienses (1946-87)	1. ABSB (1946, 1976-90)	Ensino Superior	Advogado
	2. JATF (1972-87)	Ensino Superior	Jornalista
2. Acervo Batista (1978)	1. VB (1978)	Ensino Médio	Estudante
3. Acervo Cartas pessoais variadas (1951-78)	1. JRC (1976)	Alfabetizado	Comerciante
	2. JSL (1964)	Alfabetizado	Do lar
	3. LJL (1951)	Alfabetizado	Agricultor
	4. JPB (1978)	Alfabetizado	Operário de fábrica
	5. VVS (1953-54)	Alfabetizado	Funcionário público municipal
4. Acervo Moura Fé (1920-21)	1. JMF (1920-22)	Ensino Superior	Farmacêutico
	2. CRJM (1920-21)	Alfabetizada	Do lar
5. Acervo Queiroz (1975-90)	1. PNQ (1975-90)	Autodidata	Funcionário Público Municipal
6. Acervo Santos Rêgo (1978-81)	1. MNRSR (1978-81)	Ensino Médio	Funcionária Pública do Distrito Federal
7. Acervo Vieira do Vale (1981-84)	1. AVV (1981-84)	Ensino Médio	Estudante
Total: 13 redatores			

De um modo geral, está em análise, nesta tese, a produção escrita de treze (13) redatores piauienses que produziram suas missivas novecentistas entre os anos de 1920 e 1990. Trata-se de redatores alfabetizados (JRC, JSL, LJL, JPB, VVS), com o ensino médio concluído (CRJM, MNRSR), inconcluso (VB e AVV), com o ensino superior (JMF, ABSB, JATF) e um redator autodidata (PNQ).

Dentre os papéis sociais dos redatores alfabetizados, temos um agricultor (LJL), um comerciante (JRC), um operário de fábrica (JPB) e um funcionário público municipal (VVS). Para os escreventes com o ensino médio concluído, observamos se tratar de uma funcionária pública do Distrito Federal (MNRSR) e de uma informante “dona de casa”

(JSL). Nessa amostra, temos dois redatores ainda em processo de conclusão do ensino superior, ou seja, atuando como estudantes do ensino superior ainda inconcluso (VB e AVV). Quanto aos informantes com o ensino superior, temos um farmacêutico (JMF), um advogado (ABSB) e um jornalista (JATF). Além disso, temos as cartas de PNQ, um autor autodidata, tendo concluído as Primeiras letras em Oeiras-PI e um breve curso de Escrituração Mercantil, em Teresina-PI.

A proposta é a de identificarmos a distribuição das formas *tu* e *você* correlacionados aos níveis de escolarização formal dos redatores piauienses na realidade brasileira novecentista.

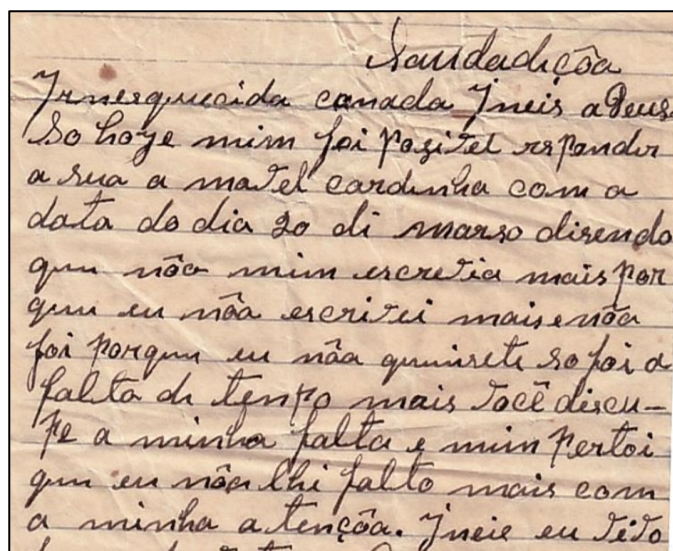
2.5.9 Os níveis de prática de escrita

O fato de termos identificado redatores piauienses com distintos níveis de prática de escrita, passamos à sua descrição a partir dos índices *baixo*, *mediano* e *alto* com o intuito de tentarmos mensurar, ainda que panoramicamente, a habilidade do redator em relação à sua expressão em língua escrita.

Assim sendo, identificamos o nível *baixo* de prática de escrita como marcado por traços da língua falada tanto na estrutura sintática das sentenças, quanto na sua representação ortográfica. A título de exemplificação, observamos traços fonéticos que reverberam também através de uma produção escrita ortográfica assumida, por exemplo, por um dado redator piauiense (VVS), em sua carta (familiar), direcionada à cunhada Inez, cf. a imagem 18.

Imagem 18: Excerto da carta de VVS.

São Caetano do Sul, 25.04.1954.

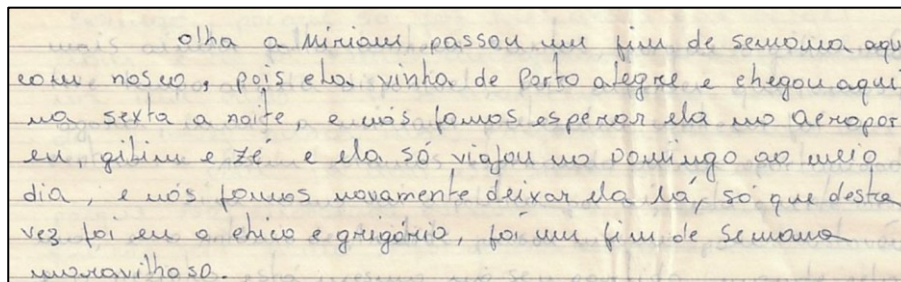


(40) “[...] Saudadiçôa Irnesquecida conada Ineis a Deus. So hoje mim foi posivel reponder a sua a mavel cardinha com a data do dia 20 di marso dizendo que nôa mim escrevia mais por que eu nôa escrevi mais nôa foi porque eu nôa quirete so foi a falta de tempo mais você discupe a minha falta e mim pertoi que eu nôa lhi falto mais com a minha atençôa. [...]” (VVS. São Caetano do Sul, 25.04.1954)

Observamos a recorrente grafia do ditongo nasal “ão” como “ôa” com uma inversão do encontro vocálico que se mostra sem a marca de nasalidade a partir do til, mas com o acento circunflexo. Identificamos também os itens “conada” por “cunhada”, “cardinha” por “cartinha”, “quirete” por “quisesse”, “discupe” por “desculpe”, “pertoi” por “perdoe” como evidências não só de inversão de fonemas (a troca do /t/ pelo /d/, em “pertoi” por “perdoe”, “card~~in~~ha” por “cartinha”, a troca do /nh/ pelo /n/, em “conada” por “cunhada”, a troca da vibrante /r/ pela sibilante /z/ e a troca da dental /t/ pela sibilante /s/, em “qu~~ir~~ete” por “qu~~is~~esse”, respectivamente), mas também de ausência de fonemas como, por exemplo, temos, em “discupe”, com a ausência da lateral /l/. Temos ainda a troca do pronome oblíquo átono “me” pelo pronome tônico “mim” nas sentenças “So hoje mim foi possível responder” por “So hoje me foi possível responder” e em “nôa mim escrevia” por “nôa me escrevia”.

Imagem 19: Excerto da carta de JPB.

São Bernardo do Campo, 18.04.1978.



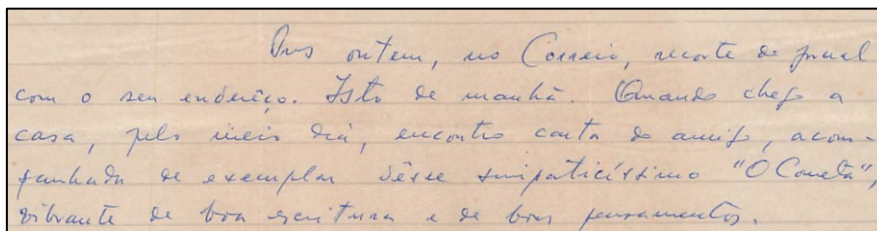
(41) “[...] Olha a Miriam passou um fim de semana aqu[i] com nosco, pois ela vinha de Porto alegre e chegou aqui na sexta feira a noite, e nós fomos esperar ela no aeropor[*to*] eu, gilim e Zé ela só viajou no Domingo ao meio dia, e nós fomos novamente deixar ela lá, só que desta vez foi eu o chico e grigório; foi um fim de semana maravilhoso. [...]” (JPB. São Bernardo do Campo, 18.04.1978)

Como evidência da produção escrita de um redator *medianamente* habilidoso (nível mediano) temos, no excerto da carta de JPB a um amigo (imagem 19), não só o pronome-sujeito *ela* na posição de objeto direto que é um traço morfossintático do português brasileiro (“esperar ela” no aeroporto, “deixar ela lá”), mas também a ausência de

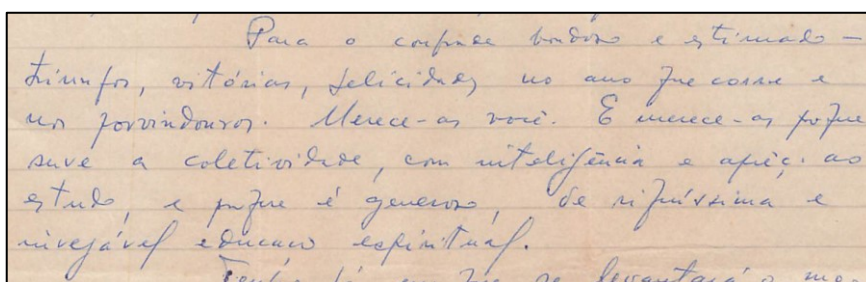
pontuação no interior do parágrafo, o que nos parece revelar que a redação da carta tenha se dado em conformidade com o fluxo do pensamento do redator. Trata-se de um redator que parece escrever sem elaboração prévia e sem monitoramento preliminares para a expressão do seu pensamento por escrito.

Imagem 20: Carta de JATF.

Teresina-PI, 11.01.1972.



[...]



(42) “[...] Pus ontem, no Correio, recorte de Jornal com o seu endereço. Isto de manhã. Quando chego a casa pelo meio dia, encontro carta do amigo, acompanhada de exemplar dêsse simpaticíssimo “O Cometa”, vibrante de boa escritura e de bons pensamentos.

[...]

“Para o confrade bondoso e estimado triunfos, estórias, felicidades, no ano que corre e nos porvindouros. Merece-as você. E merece-as porque serve a coletividade, em inteligência e apreço ao estudo, e porque é generoso, de riquíssima e invejável educação espiritual. [...]” JATF. Teresina-PI, 11.01.1972

Na imagem 20, observamos a produção escrita de um redator cujo nível de habilidade em relação à sua expressão em língua escrita mostra-se *alto*. Como evidências da habilidade de JATF temos não só as escolhas lexicais tais como o uso dos itens “porvindouros”, “riquíssima”, “simpaticíssimo”, mas também no uso da regência obsoleta do verbo *chegar a* (“chego a casa”) no lugar do uso abasileirado de *chegar* com a preposição *em* (“chego em casa”). Além disso, observemos o uso do clítico de acusativo de 3SG em sentença com o sujeito posposto ao verbo (“Merece-as você”), optando por

uma alternativa de expressão estilisticamente bem formada e, na sequência, reforçada na sentença posterior (“E merece-as porque [...]”) em que também optou o redator pelo uso do clítico de acusativo de 3SG, traço linguístico que, cf. Duarte (1989), Kato (2005), tem se mostrado como uma estratégia que chega à gramática do letrado do PB via escolarização, mostrando-se, pois, inoperante na expressão vernacular da gramática brasileira do português.

É importante atentarmos também ao fato de que a erudição demonstrada pelo escrevente JATF deve-se não só à sua formação superior (jurista, professor, jornalista e escritor) e ao desenvolvimento das habilidades relacionadas às práticas de leitura e escrita, mas também ao fato de se dirigir a um destinatário autodidata (PNQ) e extremamente ativo (professor, escritor, radialista) no exercício de atividades que tendem a favorecer o desenvolvimento das habilidades de leitura e escrita.

2.5.10 As relações sociais entre os piauienses: a simetria e a assimetria social.

O rastreamento das formas *tu* e *você* correlacionadas às dinâmicas sociais contextualizadas por redatores piauienses é o foco principal deste trabalho, buscando ampliar a discussão em relação especificamente aos resultados voltados às sincronias passadas (SOUZA, 2012; SILVA, 2012, SOUZA, 2021).

A hipótese é a de que as dinâmicas sociais movidas pela simetria (entre iguais) e assimetria descendente (de superior para inferior) impulsionariam o *você*, cf. já discutido por Lopes & Rumeu (2015, p. 23) e por Souza (2021) em relação às cartas mineiras (séculos XIX e XX) e por Souza (2012) nas cartas cariocas (séculos XIX e XX). No que diz respeito ao *tu*, a hipótese é a de que prevaleceria não só nas assimétricas ascendentes (de inferior para superior), mas também nas simétricas, tendo em vista principalmente os resultados de Souza (2012) para as cartas amorosas cariocas e os de Souza (2021) para as cartas amorosas mineiras.

Para as cartas piauienses, prevemos que as relações interpessoais travadas entre amigos (relações **simétricas solidárias** (MACHADO 2011, p. 157)), entre conhecidos, não-amigos, sem nenhum vínculo familiar, além das relações entre sogro(a) e genro/nora (relações **simétricas não-solidárias** ((MACHADO 2011, p.168), relações entre pai e filho, tio e sobrinho, patrão e empregado (relações **assimétricas descendentes** (MACHADO 2011, p. 179)), relações entre escravo e senhor, empregado e patrão, filho

e pai, neto e avô (relações **assimétricas ascendentes** (MACHADO 2011, p. 188)). Prevemos que as relações mais íntimas *entre casais* condicionariam o uso do simétrico *tu-tu*, ao passo que, nas relações assimétricas travadas entre *pai-filho*, *avô-neto*, *tio-sobrinho*, teríamos a prevalência do *você*. Na amostra de cartas piauienses, levantamos tão somente dados linguísticos de *tu/você* nas relações sociais marcadas por simetria e por assimetria descendente, conforme exemplificamos em (43) e (44).

Relações simétricas solidárias:

(43) a. Relação entre casal: “Adorada Christina [...] Orgulhoso por te amar e merecer a tua afeição, por *seres* uma pessoa virtuosa, afavel, carinhosa, meiga e terna, e, vaidoso por *seres* um thesouro de beleza [...]” J.MF. Simplício Mendes-PI, 16.09.1920.

b. Relação entre amigos: “[...] Possidônio Queiroz Ilustre Presidente do Instituto Histórico de Oeiras: Com esta rápida missiva estou convidando a você [...] Se *você* puder vir e, bem assim algum colega, será uma satisfação [...]” ABSB. Fazenda Riacho Fundo-São Pedro-PI, 06. 05.1977.

Relações assimétricas descendentes:

(44) a. “Caríssimo Neto PAULO NAZARENO Abraços [...] *Você* se lembra de que esteve aqui no fim do Ano, com a caríssima Suely? [...] A gente conversava. *Você* discorria sobre legislação trabalhista. Dizia que hoje, mesmo que o servidor contasse mais de dez (10) anos de serviços, não tinha mais estabilidade. [...]” PNQ. Oeiras-PI, 01.04.1986.

b. “Oeiras, 12 de março de 1975 Caríssimo Demontier Abraço [...] Você precisa de estudar muito. Precisa de aprender muito para vir a ser um homem preparado e que possa bem servir à nossa terra e aos seus. Não queira ser um rapaz vadio [...]” PNQ. Oeiras-PI, 12.03.1975.

2.5.11 O fator *tempo*: os períodos das cartas.

A análise da variável *tempo* está orientada pelo objetivo acompanhar a evolução na produtividade das formas pronominais *tu/você* em relação aos períodos das cartas piauienses. Nesse sentido, levantamos dados nos períodos de 20, 40, 50, 60, 70, 80 e 90 do século XX. A proposta é monitorarmos a produtividade das formas *tu* e *você* no eixo do tempo, considerando o universo de cartas piauienses novecentistas levantadas. Seguem, na sequência, evidências linguísticas de *tu* e *você* nas décadas de 20 a 90 (à exceção da década de 30) do século XX, conforme identificamos de (45) a (51).

Anos 20 (1920-29)

(45) “[...] Não me fale mais em rasgar poesias, com isso vejo que não *confias* em meu amor. [...]” CRJM. Fazenda Tatu, Piauí, 29.09.1920.

Anos 40 (1940-49)

(46) “[...] Haverá algo a *você* me informar sobre a compra da posse de terra que estava tratada? [...]” ABSB. RJ, 06.06.1946.

Anos 50 (1950-59)

(47) “[...] Sabendo eu de que *você* não pretende mais morar aqui peço ao amigo que chêgue até aqui [...]” LJL. Bocaina, Piauí, 22.06.1951.

Anos 60 (1960-69)

(48) “[...] Comadre, por aqui nada de importante tudo como sempre só muito trabalho e pouco lucro, pois *você* sabe como estão as coisas. [...]” JSL. Santo André-SP, 15.02.1964.

Anos 70 (1970-79)

(49) a. “[...] Olha! *tu* não está sendo como nós eramos, porque está escondendo as novidades, mais *tu* pensa que não sei? [...]” JPB. São Bernardo do Campo, SP, 18.04.1978.

b. “[...] Outra parte de sua carta que muito me alegrou, foi a em que *você* diz que já estava matriculado e que iria frequentar / aulas por aqueles dias. [...]” PNQ. Oeiras Piauí, 12.03.1975.

Anos 80 (1980-89)

(50) a. “[...] Faça o concurso [...] Se *queres* vencer, desperta e luta. [...]” PNQ. Oeiras-PI, 18.07.1987.

b. “[...] Não deixa de ser bem gratificante receber uma carta sua, saber como *você* está, mas é bem melhor poder ver isto de perto [...]” AVV. Brasília, 23. 09. 1981.

Anos 90 (1990-99)

(51) “[...] Penso que *você* gostará do meu trabalho, feito de amor oeirense e de amizade pela família Tapety [...]” ABSB. RJ, 22.12.1990.

2.5.12 Os conjuntos documentais de cartas pessoais: descrição pormenorizada das fontes históricas piauienses.

Nesta seção, passamos à descrição pormenorizada dos conjuntos de missivas pessoais novecentistas em análise.

2.5.12.1 Acervo “Autores Piauienses”

As cartas que compõem esse acervo foram redigidas por 02 (dois) redatores piauienses (JATF e ABSB), no período de 1946 a 1990, todas endereçadas ao amigo comum PNQ. JATF tinha escolarização de nível superior, na área de Ciências Jurídicas, mas exercia o magistério e o jornalismo, além da escrita literária. Suas cartas datam do período de 1972 e 1987, abrangendo duas faixas etárias, a de adulto e a de idoso, ou seja, entre 48 e 63 anos. ABSB tinha escolarização de nível superior na área de Ciências Jurídicas e exercia a profissão de advogado, além de suas atividades como escritor e historiador. Suas cartas datam do período de 1946 a 1990, abrangendo duas faixas etárias, a de adulto e a de idoso, ou seja, entre 43 e 83 anos. Esse acervo contém um total de 42 (quarenta e duas) *cartas de amizade*, sendo 39 (trinta e nove) de ABSB e 03 (três) de JATF. A autoria, local e data de nascimento dos missivistas, datas de escrita das cartas, tipo de relação social, subgênero e quantidade estão expostas no quadro 4.

Quadro 4: Relação das cartas do Acervo “Autores Piauienses”.

MISSIVISTA	LOCAL DE NASCIMENTO	ANO DA CARTA	REL. SOCIAL	REL. INTERPESSOAL	CARTAS PESSOAIS		
					AMOR	FAMILIAR	AMIZADE
ABSB	Oeiras-PI, 21.05.1907	1946	Simétrica	Amigo-amigo	-	-	1
		1976	Simétrica	Amigo-amigo	-	-	1
		1977	Simétrica	Amigo-amigo	-	-	3
		1981	Simétrica	Amigo-amigo	-	-	2
		1982	Simétrica	Amigo-amigo	-	-	1
		1984	Simétrica	Amigo-amigo	-	-	2
		1986	Simétrica	Amigo-amigo	-	-	1
		1987	Simétrica	Amigo-amigo	-	-	11
		1988	Simétrica	Amigo-amigo	-	-	11
		1989	Simétrica	Amigo-amigo	-	-	4
1990	Simétrica	Amigo-amigo	-	-	2		
JATF	Barras -PI 27.10.1924	1972	Simétrica	Amigo-amigo	-	-	1
		1986	Simétrica	Amigo-amigo	-	-	1
		1987	Simétrica	Amigo-amigo	-	-	1
N.º DE MISSIVISTAS 02	SUBTOTAL POR GÊNERO					-	42
TOTAL DE CARTAS						42	

2.5.12.2 Acervo “Batista”

O Acervo Batista se compõe de 4 (quatro) cartas de amizade, produzidas por VB, durante o ano de 1978, quando tinha apenas o ensino médio, era estudante universitário e estava na idade de 25 anos. Todas as cartas que compõem o presente acervo foram endereçadas à amiga ORB e trazem uma temática voltada ao cotidiano da vida de estudante. A autoria, local e data de nascimento dos missivistas, datas de escrita, tipo de relação social, subgênero e quantidade estão expostas no quadro 5.

Quadro 5: Relação das cartas do Acervo “Batista”.

MISSIVISTA	LOCAL DE NASCIMENTO	ANO DA CARTA	REL. SOCIAL	REL. INTERPESSOAL	CARTAS PESSOAIS		
					AMOR	FAMILIAR	AMIZADE
VB	Picos-PI 27.11.1954	1978	Simétrica	Amigo-amiga	-	-	4
N.º DE MISSIVISTAS	SUBTOTAL POR GÊNERO						4
01							
TOTAL DE CARTAS						4	

2.5.12.3 Acervo “Cartas Pessoais Variadas”

O Acervo Cartas Pessoais Variadas é composto de 9 (nove) cartas pessoais distribuídas nos subgêneros *de amizade e familiar*, produzidas por 5 (cinco) escreventes, a saber: a) VVS, do sexo masculino, era apenas alfabetizado, foi operário e funcionário público municipal, redigiu 4 cartas *familiares*, sendo 2 (duas) produzidas, em 1953, e 2 (duas), em 1954, período que corresponde a sua faixa etária de adulto, 31 e 32 anos respectivamente; b) L JL, do sexo masculino, era apenas alfabetizado e redigiu 1 (uma) carta *de amizade*, redigida em 1951, período que corresponde a sua faixa etária de jovem, com 27 anos; c) JPB, do sexo masculino, foi operário de fábrica, era apenas alfabetizado, redigiu 02 (duas) *cartas de amizade*, em 1978, período que corresponde à sua faixa etária de jovem, com 22 anos; d) JRC, do sexo masculino, era somente alfabetizado, foi comerciante, redigiu 1 (uma) carta *familiar*, datada de 1976, período correspondente a sua faixa etária de adulto, com 51 anos; e) JSL, do sexo feminino, era somente alfabetizada, do lar, e redigiu 1 (uma) carta *familiar* datada de 1964, período que corresponde a sua faixa etária de adulta, com 33 anos. A autoria, local e data de nascimento dos missivistas, datas de escrita, tipo de relação social, subgênero e quantidade de cartas estão expostas no quadro 6.

Quadro 6: Relação das cartas do acervo de “Cartas Pessoais Variadas”

MISSIVISTA	LOCAL DE NASCIMENTO	ANO DA CARTA	REL. SOCIAL	REL. INTERPESSOAL	CARTAS PESSOAIS		
					AMOR	FAMILIAR	AMIZADE
VVS	Bocaina-PI 14.06.1922	1953	Simétrica	Irmão-irmão	-	1	-
		1953	Simétrica	Cunhado-cunhada	-	1	-
		1954	Simétrica	Cunhado-cunhada	-	2	-
LJL	Bocaina-PI 10.11.1924	1951	Simétrica	Cunhado-cunhado	-	1	-
JPB	Bocaina-PI 10.05.1956	1978	Simétrica	Amigo-amigo	-	-	2
JRC	Bocaina-PI 05.12.1925	1976	Simétrica	Irmão-irmão	-	1	-
JSL	Bocaina-PI -. -. 1931	1964	Simétrica	Comadre-comadre	-	-	1
N.º DE MISSIVISTAS 05	SUBTOTAL POR GÊNERO						9
TOTAL DE CARTAS						9	

2.5.12.4 Acervo “Moura Fé”

Do acervo Moura Fé foram selecionadas 36 (trinta e seis) *cartas de amor*, sendo 26 (vinte e seis) redigidas por JMF e 10 (dez) por CRJM. Essas cartas foram trocadas entre o casal, durante o período de 19 de agosto de 1920 a 26 de fevereiro de 1922. Elas registram desde o pedido de namoro/noivado, feito por JMF a CRJM, até a data do casamento, 26 de fevereiro de 1922, em que JMF faz uma saudação a sua até então noiva. O Acervo Moura Fé se compõe de 2 (dois) escreventes: um do sexo masculino (JMF), com escolarização em nível universitário, adulto, na faixa etária de 31 a 32 anos; a outra (CRJM), do sexo feminino, apenas alfabetizada, mas com boa habilidade de leitura e escrita, jovem, na faixa etária entre 19 a 21 anos. A autoria, local e data de nascimento dos missivistas, datas de escrita, tipo de relação social, subgênero e quantidade de cartas estão expostas no quadro 7.

Quadro 7: Relação das cartas do Acervo “Moura Fé”.

MISSIVISTA	LOCAL E DATA DE NASCIMENTO	ANO DA CARTA	RELAÇÃO SOCIAL	RELAÇÃO INTERPESSOAL	QUANTIDADE DE CARTAS DE AMOR
JMF	Fazenda Canto Alegre – Símplicio Mendes-PI, (17/12/1889)	1920	Simétrica	Pretendente-pretendida	2
JMF		1920	Simétrica	Noivo-noiva	5
JMF		1921	Simétrica	Noivo-noiva	16
JMF		1922	Simétrica	Noivo-noiva	3
CRJM	Fazenda Tatu – Oeiras-PI, (28/08/1901)	1920	Simétrica	Noiva-noivo	2
		1921	Simétrica	Noiva-noivo	8
Nº DE MISSIVISTAS					
02					
TOTAL					36

2.5.12.5 Acervo “Queiroz”

Na busca por cartas pessoais, exploramos também o Acervo Queiroz, contendo cartas produzidas por PNQ, que é o único escrevente desse acervo. É um acervo de cunho privado onde encontramos cartas pessoais nos subgêneros *amorosa*, *familiar* e *de amizade*. Nesse acervo, foram coletadas 27 cartas endereçadas a sua esposa, seus filhos, netos, sobrinhos e amigos, durante o período de 1975 a 1990, período que corresponde à faixa etária de idoso, entre os 69 e 84 anos. PNQ não teve escolarização formal, foi um autodidata em muitas áreas/profissões e apresenta alto nível de habilidade em leitura e escrita. Foi músico (com mutas composições), ourives, professor, advogado rábula, jornalista, escritor, radialista. A autoria, local e data de nascimento do missivista, datas de escrita, tipo de relação social, subgênero e quantidade de cartas estão expostas no quadro 8.

Quadro 8: Relação das cartas do Acervo “Queiroz”

MISSIVISTA	LOCAL DE NASCIMENTO	ANO DA CARTA	REL. SOCIAL	REL. INTERPESSOAL	CARTAS PESSOAIS		
					AMOR	FAMILIAR	AMIZADE
		1975	Simétrica	Amigo-amigo	-	-	1
		1983	Descendente	Avô-neto/a	-	1	-
		1985	Descendente	Avô-neto/a	-	3	-
		1985	Descendente	Pai-filho/a	-	1	-
		1986	Descendente	Avô-neto/a	-	6	-
PNQ	Oeiras-PI, 17.05.1906	1987	Descendente	Avô-neto/a	-	6	-
		1987	Descendente	Pai-filho/a	-	1	-
		1988	Simétrica	Esposo-esposa	-	4	-
		1988	Descendente	Avô-neto/a	-	2	-
		1989	Simétrica	Amigo-amigo/a	-	-	1
		1990	Descendente	Pai-filho/a	-	1	-
N.º DE MISSIVISTAS	SUBTOTAL POR GÊNERO				-	25	2
01							
TOTAL DE CARTAS						27	

2.5.12.6 Acervo “Santos Rêgo”

O presente acervo se compõe de 7 (sete) cartas *familiares*, redigidas por MNRSR, do sexo feminino e com o ensino médio concluído. Era funcionária pública e estudante universitária em Brasília. São cartas endereçadas à sua prima ORB com temática que envolve assuntos do cotidiano da juventude, de estudante. As referidas cartas foram produzidas no período de 1978 a 1981, período que corresponde a sua faixa etária jovem. A autoria, local e data de nascimento da missivista, datas de escrita, tipo de relação social, subgênero e quantidade estão expostas no quadro 9.

Quadro 9: Relação das cartas do Acervo “Santos Rego”.

MISSIVISTA	LOCAL DE NASCIMENTO	ANO DA CARTA	REL. SOCIAL	REL. INTERPESSOAL	CARTAS PESSOAIS		
					AMOR	FAMILIAR	AMIZADE
MNRSR		1978	Simétrica	Prima-prima	-	-	5
		1979	Simétrica	Prima-prima	-	-	1
		1981	Simétrica	Prima-prima	-	-	1
N.º DE MISSIVISTAS	SUBTOTAL POR GÊNERO						
01						-	7
TOTAL DE CARTAS					7		

2.5.12.7 Acervo “Vieira do Vale”

O presente acervo é composto de 11 (onze) cartas de amor, produzidas por AVV, do sexo masculino, com escolarização de ensino médio, e estudante universitário. As referidas cartas foram redigidas no período de 1981 a 1984, período em que tinha entre 32 e 34 anos, portanto, pertencente à faixa etária adulto. Abordava temas como respostas a cartas anteriores, cobranças de atenção, planos para o futuro, cotidiano de estudante, saudades da terra natal, assuntos políticos de sua cidade. A autoria, local e data de nascimento dos missivistas, datas de escrita, tipo de relação social, subgênero e quantidade de cartas estão expostas no quadro 10.

Quadro 10: Relação das cartas do Acervo “Vieira do Vale”

MISSIVISTA	LOCAL DE NASCIMENTO	ANO DA CARTA	REL. SOCIAL	REL. INTERPESSOAL	CARTAS PESSOAIS		
					AMOR	FAMILIAR	AMIZADE
		1981	Simétrica	Namorado-namorada	-	-	1
AVV	Dom Expedito Lopes-PI	1982	Simétrica	Namorado-namorada	-	-	3
	-. -. 1949	1983	Simétrica	Namorado-namorada	-	-	4
		1984	Simétrica	Namorado-namorada	-	-	3
				Namorado-namorada			
N.º DE MISSIVISTAS	SUBTOTAL POR GÊNERO						11
01							
TOTAL DE CARTAS					11		

A partir dessa breve descrição da amostra de cartas novecentistas piauienses, constatamos que temos em análise um diversificado conjunto de cartas pessoais desarmonizado não só em relação ao número de cartas distribuídas pelos seus subgêneros textuais (cartas amorosas, familiares e de amizade), mas também no que se refere às cartas no eixo do tempo, percorrendo assim todas as décadas da era novecentista do PB escrito. Ainda que não tenhamos em análise uma amostra de cartas equanimemente equilibrada em relação aos subgêneros textuais devidamente distribuídas no eixo do tempo, passamos a uma análise probabilística tendo em vista a análise de Souza (2021) também voltada à regra variável *tu/você* em amostras históricas da escrita mineira também sob os mesmos critérios linguísticos e extralinguísticos. Assim sendo, buscamos justificar a opção pela implementação de uma análise também embasada em pesos relativos, consubstanciada em uma amostra histórica também lacunar, ainda que em distintos níveis, de modo a comparar os resultados desta Tese aos de Souza (2021).

Como descrevemos os critérios linguísticos e extralinguísticos, passamos à uma breve incursão pela história de formação do espaço geográfico do Piauí.

2.6 Piauí: um pouco da sua história de formação

Para fazer um levantamento completo da história do Piauí, seria necessário um recuo à Pré-História, em função das descobertas arqueológicas de Niède Guidon, na Serra da Capivara, em São Raimundo Nonato-PI, que revelam ser o Piauí o berço do homem americano. Segundo Dias (2006, *apud* Lima, 2020),

o povoamento colonial das Américas que representou também o despovoamento da população nativa se desenrolou sob a égide da destruição de um povo com possíveis origens pré-históricas cujos vestígios arqueológicos estão na Serra da Capivara, em São Raimundo Nonato, Piauí, Brasil. (LIMA 2020, p. 09)

Apesar de a História do Piauí ter suas raízes na pré-história, nosso interesse está circunscrito a um período mais recente, os séculos XIX e XX, uma vez que as amostras textuais desta tese datam do período compreendido entre os anos de 1920 e 1992. No entanto, para situarmos as configurações da sociedade piauiense do período em estudo, faremos um pequeno recuo ao momento da colonização (séculos XVII e XVIII), já que este deixa marcas que se prolongam para épocas posteriores.

Segundo Brandão (1995, p. 33), o Piauí passou a se constituir como unidade político-administrativa do Brasil Colônia a partir do século XVIII, período em que ocorreram a criação da capitania e a instalação do primeiro governo. Antes da criação da capitania, o território hoje piauiense fazia parte de uma área denominada Sertão de Dentro ou Sertão de Rodelas, que compreendia a área situada a oeste do rio São Francisco. Desde a sua criação, o espaço geográfico destinado à Capitania do Piauí situava-se no lado oriental (leste) da bacia do rio Parnaíba. Por ser uma área de transição climática, entre a Amazônia e o sertão (semiárido), apresentava muitos rios (Longá, Poti, Sambito, Canindé, Piauí, Gurguéia, Uruçuí, Parnaíba, além de alguns lagos) e regime de chuvas regular, que proporcionava a existência de pastagem natural abundante, um ambiente propício para a criação de gado. Assim, desde a criação da Capitania do Piauí, na segunda metade do século XVII, a região sempre foi definida como zona de produção pecuarista, cuja produção destinava-se ao abastecimento do mercado brasileiro. A estrutura (sociopolítica) e o desempenho do sistema de colonização impulsionaram a expansão dos currais no interior do Nordeste e promoveram o alargamento do processo de povoamento da região. A expansão da pecuária foi motivada ainda, segundo Brandão (1995), pela

preferência dos senhores de engenho pela mão de obra escrava, além de sua incapacidade de absorção da mão de obra livre. Assim, os trabalhadores adentravam os sertões em busca de espaços para a criação de gado, como uma alternativa à falta de trabalho na indústria açucareira. Apesar da importância do sertão para o povoamento do Piauí, apresentada por Brandão, alguns historiadores apresentam evidências de que, no mesmo período (século XVII) e até mesmo no século XVI, já havia intensa movimentação na região litorânea. Segundo Baptista (2019),

Os primeiros relatos que temos do litoral piauiense datam do século XVI, a partir do naufrágio de uma embarcação portuguesa na qual se encontrava Nicolau de Resende. Odilon José Nunes afirma que, por volta do ano de 1571, Nicolau de Resende havia se perdido na região e relata a existência de comércio clandestino entre embarcações e os índios Tremembés, senhores do litoral na época. (BAPTISTA 2019, p. 28)

Joselina L. P. Rodrigues, em seu livro *Geografia e História do Piauí* (2012), apresenta uma posição contrária à de outros historiadores piauienses, inclusive Brandão, para quem a colonização do Piauí se deu do sertão para o litoral. Rodrigues (2012) afirma que, na primeira metade do século XVII, época conhecida como pré-bandeirantismo, surgiram as primeiras expedições colonizadoras, de início no litoral, posteriormente no sertão. Tal constatação se baseia, dentre outros documentos, na obra *A Prioridade do Norte no Povoamento do Piauí*, de Pe. Cláudio Melo, segundo o qual o povoamento do norte piauiense antecedeu ao sul.

Confrontando os dados de Baptista com os dos demais historiadores, é possível fazer algumas apreensões a respeito das discordâncias sobre a direção em que se deu o processo de colonização do Piauí, se do sertão para o litoral, ou no sentido contrário. Documentos revelam incursões ao litoral ainda no século XVI (RODRIGUES 2012, p. 156), outros revelam a penetração pelo sertão (centro-sul) no século XVII (BRANDÃO, 1995, p. 32). Nossa pergunta é: qual dessas incursões motivou, de fato, o processo de colonização do estado? Considerando que as atividades econômicas sempre foram importantes para a efetivação de tal processo, questionamos: Onde se concentrava a atividade econômica mais lucrativa, no litoral ou no sertão? No litoral os invasores tinham interesses comerciais e religiosos (catequizar os Tremembés); no sertão, o interesse era a criação de gado. As discussões e divergências acerca desse tema devem nascer das incertezas sobre qual dessas atividades era mais rentável e, conseqüentemente, proporcionou maior ocupação humana no solo piauiense.

Segundo Brandão (1995, p. 32), o limite oeste do Piauí foi a bacia do rio Parnaíba, mostrando-se bem definido. Mas, pelo lado leste, a penetração dos currais, já que os criadores de gado exerciam suas atividades em um processo itinerante (estavam onde havia pastagem natural e água). Depreendemos que os limites foram ficando bastante imprecisos, de modo que o município de Príncipe Imperial, hoje Crateús-CE, durante o período colonial, pertencia à Capitania do Piauí, e A *Vila de Amarração*, hoje Luís Correia-PI, na mesma época, esteve sob a jurisdição da Capitania do Ceará. Segundo Baptista (2019, p. 23), o ano é 1880 é a data que marca a lei que garante à província do Piauí a *Vila de Amarração* em troca de *Príncipe Imperial e Independência* (Imagem &).

Figura 2: Mappa Geographico da Capitania Piauhy, e parte das do Maranhão, e do Gran Pará.

(Fonte: <http://biblioteca.interpi.pi.gov.br/Terras-war/mapa-historico.xhtml>)



Dessa discussão, inferimos que, apesar da existência de expedições no litoral, desde o século XVI, o processo de colonização do Piauí se deu a partir do século XVII, com o desenvolvimento da pecuária, no sertão, que proporcionou grande movimento de pessoas, grande produção de carne e lucro para os criadores, fato este que motivou um certo desleixo em relação ao litoral, que foi dominado pelos cearenses. Por esse motivo, em

1880, o Piauí teve que doar ao Ceará os municípios de *Príncipe Imperial* e *Independência*, em troca da *Vila de Amarração*.

Figura 3: Mapa do estado do Piauí pelo IBGE.

Fonte: https://geofp.ibge.gov.br/produtos_educacionais/mapas_tematicos/mapas_do_brasil/mapas_estaduais/politico/piaui.pdf



Essa busca por dados da História do Piauí se justifica pelo fato de que as migrações entre Piauí e os estados vizinhos, são bastante intensas desde o período colonial. A existência de áreas de litígio e de imprecisão dos limites geográficos, sobretudo com o Ceará, além do fato de que, antes de ser capitania autônoma, a área geográfica que corresponde hoje ao Piauí já esteve sob o domínio da Bahia e de Pernambuco e, posteriormente, do Maranhão (figuras 2 e 3).

2.7 Metodologia: a submissão dos dados históricos ao RBrul.

Nesta tese, os dados da alternância *tu/você* estão submetidos ao Rbrul na sua interface com o RStudio para a geração dos pesos relativos em relação aos condicionamentos intra e extralinguísticas da regra variável em cena (AVELHEDA BANDEIRA & SOUZA GUERREIRO, 2019). A linguagem computacional RBrul favorece as análises de uma dada regra variável, uma vez que oferece funções orientadas, por exemplo, à produção de gráficos e de listas de frequência, otimizando e agilizando o trabalho do linguista-pesquisador. O uso do RBrul, cf. Gomes (2012, p. 261), nas análises linguísticas tende a acentuar o alcance explanatório de algumas questões importantes como, por exemplo, o papel do indivíduo na dinâmica da variação. Está disponível, *on line*¹¹, ao acesso público para o respectivo *download* gratuito, o que facilita não só a sua instalação, mas também a sua recorrente atualização.

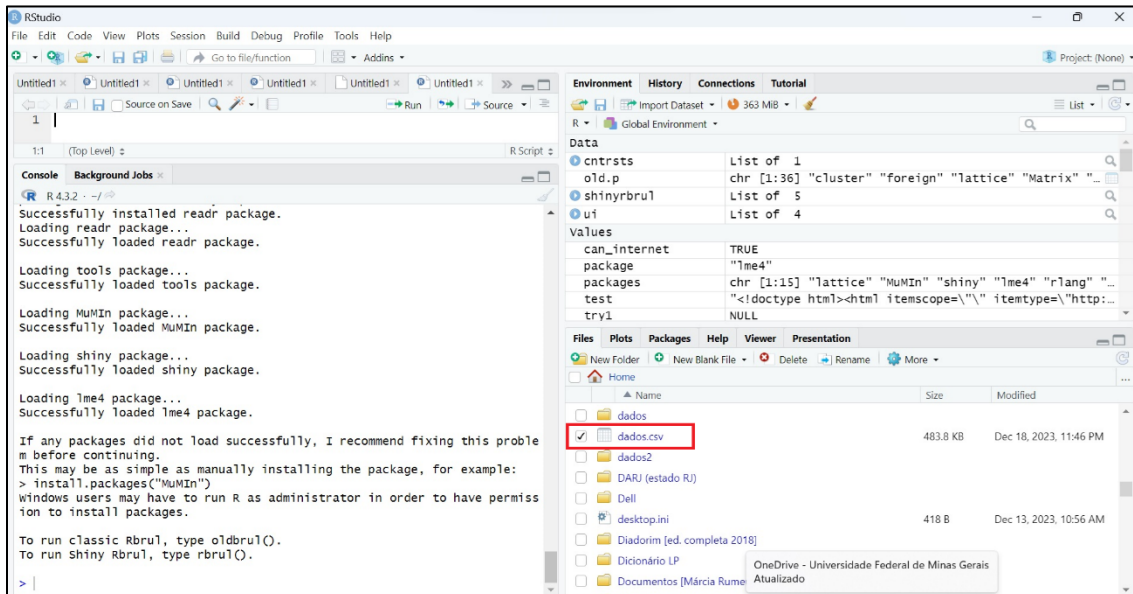
A interface do RBrul com o ¹²RStudio aumenta as suas possibilidades de uso, tendo em vista os pacotes e o histórico das linhas de comando (FREITAG 2014, p. 136). Em interface com o RStudio ampliam-se também, no RBrul, as capacidades gráficas para a geração de gráficos (GOMES, 2012). Na perspectiva de Oushiro (2014), o RBrul é uma plataforma de análise estatística completa, uma vez que reúne ferramentas fundamentais para a operacionalização dos dados, concentrando todo o processo de organização do *corpus*, das análises estatística e probabilística e da geração dos gráficos. Operacionalmente, o RBrul admite arquivos de dados nos formatos .tkn, .txt e .csv, cf. GOMES (2012, p. 262). Para a geração dos índices probabilísticos (pesos relativos),

¹¹ <http://www.danielezrajohnson.com/Rbrul.R>

¹² <http://www.rstudio.com/ide/download/>

submetemos a planilha do excel ao formato *.csv* (*comma separate values*), cf observamos, ao abriremos o RStudio que nos apresentará as seguintes janelas, na imagem 21.

Imagem 21: Interface do RStudio – RBrul
versão 4.3.2 (2023-12-29).



Na imagem 22, temos a imagem da página inicial do RStudio. Nessa página, temos a sua visualização dividida em quatro janelas. São elas: i. *Source*, espaço para a visualização e a edição dos dados através dos *scripts*; ii. *Environment* e *History*, ambiente e histórico, respectivamente, onde é possível resgatarmos os objetos carregados na memória do R e o histórico de linhas de comando executadas; iii. *Console*, espaço para as funções e a execução dos *scripts*; iv. *Files*, *Plots*, *Packages*, *Help* e *Viewer* para a implementação de arquivos, gráficos, pacotes, ajuda e visualizador, , respectivamente (FREITAG 2014, p 136). Uma vez levantados, digitados, codificados e disponibilizados em uma planilha do *excel* passamos à geração dos índices estatísticos com base na proposta de tabela dinâmica oferecida pelo próprio *excel*, passamos à submissão dessa planilha de dados em formato *.csv* (separado por vírgulas) ao RStudio, objetivando a rodada multivariada nos níveis *step-up & step down* para a geração dos pesos relativos (AVELHEDA BANDEIRA & SOUZA GUERREIRO, 2019).

Em síntese, considerando a potência dos pesos relativos para a interpretação acerca da força das variáveis independentes em relação à produtividade das formas pronominais *tu* e *você*, optamos, nesta tese, por analisar essa dinâmica variável a partir não só dos

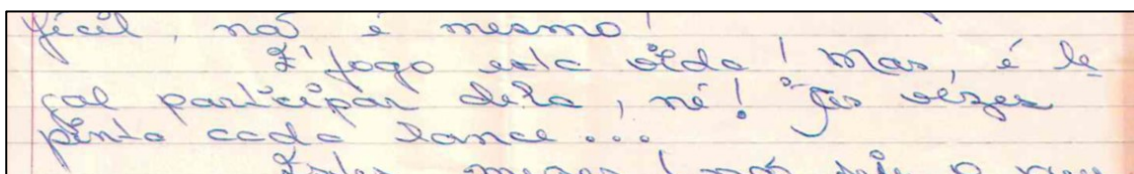
índices percentuais (capítulo 3), mas também dos índices probabilísticos (capítulo 4). Nesse sentido, assumimos a proposta de utilizarmos das potencialidades do RBrul como método de tratamento dos dados variáveis, principalmente em relação à sua interface no RStudio para a produção, em rodada multivariável, dos índices probabilísticos da regra variável (*tu/você* = variável dependente) nos níveis *step-up* & *step-down* em relação às variáveis independentes internas (a *expressão nula ou plena do pronome de 2SG*, a *pessoa verbal (concordância)*, o *paralelismo formal e semântico*) e externas/sociais (o *sexo*, a *faixa etária*, as *relações interpessoais entre os missivistas*, o *período*, a *escolaridade*, os *níveis de prática de escrita*, o *subgênero da carta pessoal* e os *acervos*).

CAPÍTULO 3. A VARIAÇÃO *TU/VOCÊ*: DESCRIÇÃO-ANALÍTICA DOS RESULTADOS ESTATÍSTICOS.

Neste capítulo, voltamo-nos à distribuição geral dos dados dos pronomes-sujeito de 2SG com base em uma rodada geral produzida a partir de planilhas do Excel, o que nos conduz aos índices percentuais da variável dependente binária: *tu/você*. Nesse sentido, optamos por, neste capítulo, trazer à discussão os resultados dos pronomes *tu* e *você* em relação às variáveis linguísticas (a *expressão do sujeito*, a *pessoa verbal* e o *paralelismo formal e semântico*) e às extralinguísticas (o *sexo*, a *faixa etária*, as *relações interpessoais entre os missivistas*, o *período*, a *escolaridade*, os *níveis de prática de escrita*, o *subgênero da carta pessoal* e os *acervos*), respectivamente, em 3.1 e 3.2.

Antes de passarmos à descrição dos dados pronominais de 2SG (*tu/você*), trazemos à cena as 23 ocorrências de *vós* em cartas produzidas por JMF (Acervo Moura Fé) e MNRSR (Acervo Santos Rêgo) que foram, obviamente, excluídas da análise quantitativa tanto em termos percentuais, quanto em termos probabilísticos. Passamos aos dados de *vós* expostos de 53 a (75). De (53) a (57), temos em evidência a produção escrita de MNRSR que, ao escrever à prima, assumindo um tom bastante despojado e informal marcado por gírias, por exemplo, faz uso linguístico formal do *vós* para a referência à 2SG. Para exemplificar o tom despojado e informal da produção escrita de MNRSR aos primos, transcrevemos o seguinte trecho marcado pelo uso de gírias como “legal”, “pinta” e “lance” como sinônimos de “interessante”, “aparece”, “ocasião”, respectivamente, e da redução do “não é” por “né”, traço típico da língua falada, cf. imagem (22).

Imagem 22: Excerto da carta de MNRSR. 10.05.1978.



(52) “[...] É fogo esta vida! Mas é legal participar dela, né! Às vezes pinta cada lance... [...]” MNRSR. 10.05.1978.

Por outro lado, nas cartas do casal CRJM e JMF, temos evidências (17 dados) de um *vós* cortês, ainda que no contexto de uma relação simétrica e marcada pela intimidade das confissões, das declarações no âmbito da conquista amorosa, cf. observamos de (58) a (75).

(53) “[...] Ah! deixa te contar, lembrás das compras que *fizestes* na Bi.ba.bo? [...]” MNRSR. Brasília, 05.05.1978.

(54) “[...] Chegando em casa, que grata surpresa, cartinha sua. Ontem coloquei uma para você; já *recebestes*? [...]” MNRSR. Brasília, 09.05.1978.

(55) “[...]Antes de ir para Campo Grande te escrevi uma cartinha, não sei se *recebestes*. [...]” MNRSR. Brasília, 23.07.1978.

(56) “[...] Me *pedistes* pra ver programa la na Fac. Já requeri. [...]” MNRSR. Brasília, 20.08.1978.

(57) “[...] Me pedistes pra ver programa la na Fac. [...] *Perguntastes* de Marcos, lembrás que te falei [...]” MNRSR. Brasília, 20.08.1978.

(58) “[...] Quanto me conforta e me alegraste! – *Satisfizestes* ás malícias de Zuquinha [...]” JMF. Simplício Mendes-PI, 20.09.1920.

(59) “[...] Falaste em mal educada. (...) Si não *tivestes* uma instrução vasta, recebeste uma educação [...]” JMF. Simplício Mendes-PI, 11.10.1920.

(60) “[...] Disseste que te falta um raio de intelligencia. Como *escrevestes* uma quintilha tão inspirada? [...]” JMF. Simplício Mendes-PI, 11.10.1920.

(61) “[...] porque te amo muito. *Fostes* creada para completar minh’alma. Eras a visão dos meus sonhos. [...]” JMF. Simplício Mendes-PI, 11.10.1920.

(62) “[...] Eu que só a ti amei, que só a ti amo e que só a ti amarei, jamais poderei suportar o teu sofrimento. *Dissestes* que amor não é crime. [...]” JMF. Simplício Mendes-PI, 06.01.1921.

(63) “[...] Como estavas inspirada! *Metrificastes* em prosa. (...)” JMF. Simplício Mendes-PI, 06.01.1921.

(64) “[...] Amo-te acima de tudo [...] Não me *dissestes* que havias ficado tão [...]” JMF. Simplício Mendes-PI, 06.02.1921.

(65) “[...] a felicidade de te ver agora, venho ao menos te escrever esta cartinha [...] ver se alivião-me as saudades que *deixas-tes* [...]” CRJM. Fazenda Tatu, Oeiras-PI, 18.11.1920.

(66) “[...] mas é preciso dizer-te aquilo que me cumpre. *Fostes* caipora no trabalho dos dentes da Santinha, [...]” CRJM. Fazenda Tatu, Oeiras-PI, 09.01.1921.

(67) “[...] não se esqueça de tua Christina. Eu de saude passo como me *deixas-tes*, dias melhores [...]” CRJM. Fazenda Curral de Pedras, Oeiras-PI, 11.02.1921.

(68) “[...] recibi tua carta foi a noite [...] Fiquei afflicta com tantas percas que *tivestes* [...]” CRJM. Fazenda Curral de Pedras, Oeiras-PI, 11.02.1921.

(69) “[...] agradeço-te sinseramente os pesames que a mim e a ellas *enviastes* pela morte da titia [...]” CRJM. Fazenda Curral de Pedras, Oeiras-PI, 11.03.1921.

(70) “[...] Com a pergunta da minha primeira carta que te fiz depois de tua saida *ficas-tes* supondo [...]” Acervo Família Mura Fé. CRJM. Fazenda Curral de Pedras, Oeiras-PI, 11.03.1921.

(71) “[...] A ti nada tenho que perdoar, pois nada me *fizestes* [...]” CRJM. Fazenda Curral de Pedras, Oeiras-PI, 11.03.1921.

(72) “[...] Adeus queridinho ate o fim do mez, como me *prometestes*. [...]” CRJM. Fazenda Curral de Pedras, Oeiras-PI, 11.03.1921.

(73) “[...] teus affectos e [...] nascidos de tua grande e generosa alma, do teu coração [...] *Fostes* nascido meu amor para me fazer feliz [...]” CRJM. Fazenda Tatu, Oeiras-PI, 03.05.1921.

(74) “[...] quando terei o prazer de te ver [...] porque me *disses-tes* que talvez não podesse vir. [...]” CRJM. Fazenda Tatu, Oeiras-PI, 09.09.1921.

(75) “[...] com remorço de não te ter escrito [...] Tu porque não me *decestes* antes [...]” CRJM. Fazenda Tatu, Oeiras-PI, 09.09.1921.

A proposta é acompanhar o nível de inserção do *você* no sistema pronominal do PB com base na análise de cartas piauienses de sincronias passadas (século XX), tendo em vista os resultados de Souza (2012) em relação às cartas cariocas dos séculos XIX e XX, de Souza (2021) em relação às missivas mineiras e aos resultados de Rumeu (2013) com base nas cartas cariocas também circunscritas aos séculos XIX e XX. Acrescentemos ainda o fato de que temos a constatação relatada por Scherre *et alii* (2021) acerca da alternância *tu/você*, evidenciando o *tu* com concordância média (10% a 39%), na fala piauiense (subsistema 5), à luz de dados do Atlas Linguístico do Brasil – Alib: inquéritos geolinguísticos (CARDOSO *et alii*, no prelo; CARDOSO, 2008; DEUS, 2009). Para as cartas piauienses, a hipótese é a de que o *tu* também se mostre em alternância com o *você*, o que já poderia evidenciar um traço da atual fala do Piauí que se mostra com a produtividade do *tu* abaixo de 60%, com concordância entre 10% e 39% (subsistema 5), cf. a discussão de Scherre *et alii* (2021, p. 165). A fim de darmos conta de tal hipótese,

assumimos como ponto de partida a distribuição das formas *tu* e *você* (tabela 2) como uma regra variável binária.

Tabela 2: As formas de referência ao sujeito de 2SG nas cartas piauienses (século XX).

FORMAS DE REFERÊNCIA AO SUJEITO DE 2SG	
TU	VOCÊ
233/583	350/583
(39,97%)	(60,03%)

Em termos gerais, observamos, com base na análise da tabela 2, uma instigante disputa entre as formas *tu* (233/583, 39,97%) e *você* (350/584, 60,03%) na produção escrita piauiense do século XX. Em (76) e (77), apresentamos evidências das formas de referência ao sujeito de 2SG (*tu* versus *você*) nas cartas piauienses analisadas.

(76) “[...] Adeus queridinho ate o fim do mez, como me prometestes. *Recebas* lembranças de todos e saudades e amplexos de tua noiva *amada* C. [...]” CRJM. Fazenda Curral de pedras, Piauí, 11.03.1921.

(77) “[...] A Renata, no sábado estive por lá, ela foi logo me perguntando pela “Yete”, *imaginas* o bicão dela perguntando: cadê a Yete; em !? [...] em matéria de bagunça nem preciso te falar que ela é o máximo não é mesmo, *você* já a conhece. [...]” MNRSR. Brasília, 05.05.1978.

3.1 Os fatores linguísticos: a expressão do sujeito, a pessoa verbal e o paralelismo formal e semântico.

Passamos aos resultados percentuais para o controle da variável dependente binária (*tu* e *você*) em relação às variáveis independentes linguísticas (a expressão do sujeito de 2SG, a pessoa (concordância verbal) e o paralelismo formal e semântico), tendo em vista as conjecturas previstas para uma das variáveis no capítulo 2.

Tabela 3: A expressão dos pronomes-sujeito de 2SG: nulo ou pleno.

Referência ao Sujeito de 2SG	Expressão	
	Nulo	Pleno
Tu	217/233 (93,13%)	16/233 (6,87%)
Você	93/350 (26,57%)	257/350 (73,43%)
TOTAL	310/583 (53,17%)	273/583 (46,83%)

A proposta é a de controle das formas *tu* e *você* em suas realizações nula ou plena, impulsionada pela conjectura de que à proporção que o *você* se pronominaliza, passe a adotar os mesmos espaços funcionais de um pronome do PB, o que está em harmonia com Fontanella (1977) e Lopes *et alii* (2009), cf. discutido por Rumeu (2013) para as cartas cariocas oitocentistas e novecentistas. Para as cartas cariocas, Rumeu (2013, p. 169) identificou o contexto de sujeito nulo sendo condicionado pelo *tu*, enquanto o *você* tende a se apresentar pleno mesmo que em ampla concorrência com a sua expressão nula.

Nesta análise, os resultados acerca da expressão do sujeito de 2SG evidenciam o *tu* preferencialmente nulo (217/233, 93,13%), ao passo que o *você* se mostra mais profícuo em contexto de sujeito pleno (257/350, 73,43%). O fato de o *você* mostrar-se potencialmente pleno (94,46%) reflete-nos um traço formal específico da forma nominal de tratamento *vossa mercê* que persistiu no processo de gramaticalização (*vossa mercê* > *você*), cf. já amplamente analisado também por Rumeu (2004), Rumeu (2006), Lopes & Rumeu (2007), Lopes & Cavalcante (2011, p. 46), Souza (2012), Souza (2021). Em sentido oposto, podemos interpretar o fato de o *tu* ter se mostrado preferencialmente nulo (79,58%) como expressão do *tu*-etimológico que permite ser resgatado através desinência número pessoal específica de 2SG da forma verbal. Nesse sentido, os dados deste trabalho parecem dialogar com análises embasadas em cartas históricas do PB (SOUZA, 2012; SOUZA, 2021). Em (78) e (79), apresentamos evidências do *tu* nulo e do *você* pleno, respectivamente.

(78) “[...] Orgulhoso por te amar e merecer a tua afeição, por [0] *seres* uma pessoa virtuosa [...]”. JMF. Simplicio Mendes-PI, 16.09.1920.

(79) “[...] Só os homens superiores compreendem o dever de ajudar o próximo - e *você*, Mestre, é um deles [...]” AT. Teresina-PI, 31.01.1987.

Tabela 4: As formas de referência ao sujeito de 2SG e a pessoa verbal (concordância)

Formas de referência ao Sujeito de 2SG	Pessoa verbal (Concordância)	
	2SG	3SG
Tu	223/233 (95,28%)	10/233 (4,72%)
Você	-	350/350 (100%)
Total	222/583 (38,08%)	361/583 (61,92%)

O principal objetivo do controle da pessoa verbal é determinar as evidências de uma possível concordância não-canônica do *tu* com formas verbais de 3SG (“*tu fez*”) na escrita novecentista, já que *você* alterna com o *tu*, que, por sua vez, se manifesta, em menos de 60% dos dados, com potencial médio de concordância (entre 10% e 39%), cf. Scherre *et alii* (2021, p. 165) na fala dos piauienses. Por outro lado, a previsão é a de que o *você* prevaleça em harmonia morfossintática com formas verbais de 3ª pessoa do singular (350/350, 100%), como evidência de uma espécie de uma “persistência formal (herdada do *vossa mercê*) presente na concordância *você* + 3ª pessoa, mesmo o *você* fazendo referência à segunda pessoa do discurso, como afirmam Lopes & Machado (2005) e Lopes (2003)”, cf. Souza (2012, p. 99).

De um modo geral, observamos que o *tu* prevalece em harmonia morfossintática com a morfologia flexional de 2SG (222/233, 95,28%), ainda que tenhamos levantado seis evidências, em tão somente 4,72% dos dados (11 ocorrências), do *tu* em concordância com formas verbais de 3SG, o que parece apontar para o alto nível de intimidade das cartas através das quais a norma de uso vigente atualmente já se manifesta, com distintos níveis (timidamente), nas cartas analisadas. De (80) a (89), temos em análise não só algumas evidências do *tu* em concordância com formas de 3SG (dez delas, entre (80) e

(88)), mas também dados do *tu*, etimologicamente marcado por formas de 2SG ((91) e (92)), e de *você* em harmonia com formas de 3SG nas cartas novecentistas dos piauienses ((93) e (94)).

(80) “[...] Não há razão de *teres* pensamentos desagradáveis porque *amas* e *é* verdadeiramente adorada mesmo. [...]” JMF. Simplício Mendes, 24.04.1921.

(81) “[...] Só *tu* me não *sai* da mente. [...]” JMF. Simplício Mendes, PI, 17.10.1921

(82) “[...] Mas *tu* não *acredita* que eu sofra saudade em grau alterado que me faça definhar. [...]” CRJM. Fazenda Tatu, PI, 21.05.1921.

(83) “[...] e eu calculava que *tu chegace* até 10 ou 12 [...]” CRJM. Fazenda Tatu, PI, 14.11.1921.

(84) “[...] *Tu* porque não me decestes antes de sahir que *pretendia* demorar-se um mez [...]” CRJM. Fazenda Tatu, PI, 14.11.1921.

(85) “[...] Contava que *tu viesse* com elle até aqui para melhor acertar o dia de nosso casamento [...]” CRJM. Fazenda Tatu, PI, 14.11.1921.

(86) “[...] Desejo-te saude e a todos os teus e que *tu tenha* sido feliz na viagem. [...]” CRJM. Fazenda Tatu-PI, 14.11.1921.

(87) “[...] Olha! *tu* não *está* sendo como nós eramos porque *está* escondendo as novidades [...]” JPB. São Bernardo do Campo-SP, 18.04.1978.

(88) “[...] porque *está* escondendo as novidades, mais *tu pensa* que não sei? [...]” JPB. São Bernardo do Campo-SP, 18.04.1978.

(89) “Se *queres* vencer, desperta e luta. [...]” PQ. Oeiras-PI, 18.07.1987.

(90) “[...] Mas, finalmente se aproxima a hora da felicidade suprema para mim e *tu* meu amorsinho, pois sei que me *amas* sinseramente. [...]” JMF. Simplício Mendes-PI, 12.02.1922.

(91) “[...] Vão também, mandar os retratos que *você mandou* pedir [...]” PQ. Oeiras-PI, 12.03.1975.

(92) “[...] Ante os empecilhos aparecidos, como *você relatou*, eu vinha propor uma sugestão [...]” BB. RJ, 10.05.1987.

Tabela 5: As formas de referência ao sujeito de 2SG e o paralelismo formal e semântico.

PARALELISMO FORMAL E SEMÂNTICO	FORMAS DE REFERÊNCIA	
	AO SUJEITO DE 2SG	
	TU	VOCÊ
1ª Ocorrência	3/14 (21,43%)	11/14 (78,57%)
Formas precedidas por <i>Você</i> -sujeito pleno/nulo	6/191 (3,14%)	185/191 (96,86%)
Formas precedidas por <i>Você</i> não-sujeito (complemento)	3/27 (11,11%)	24/27 (88,89%)
Formas precedidas por formas de 3SG (possessivos e clíticos)	4/107 (3,74%)	103/107 (96,26%)
Formas precedidas por <i>Tu</i> -sujeito pleno/nulo	70/73 (95,89%)	3/73 (4,11%)
Formas precedidas por <i>Tu</i> não-sujeito (complemento)	8/10 (80,00%)	2/10 (20,00%)
Formas precedidas por possessivos e clíticos de 2SG	112/117 (95,73%)	5/117 (4,27%)
Formas precedidas por formas possessivas e clíticas de 2PL	5/5 (100%)	-
Formas precedidas por Vocativo	13/25 (52,00%)	12/25 (48,00%)
Formas precedidas pelo imperativo de 2SG (Subjuntivo)	7/12 (58,33%)	5/12 (41,67%)
Formas precedidas pelo imperativo de 2SG (Indicativo)	2/2 (100%)	-
TOTAL	233/583 (39,97%)	350/583 (60,03%)

Em relação à variável paralelismo formal e semântico, a hipótese é a de que a sequência discursiva iniciada por *você* seja um contexto promissor para a sua produtividade, cf. pensado, inicialmente, por Omena em relação ao *a gente* e verificado, posteriormente, por Lopes & Vianna (2012, p. 148) à luz de Omena (2003). Resta-nos saber se as opções assumidas pelos escreventes piauienses estariam ou não motivados pela questão do paralelismo linguístico (SCHERRE, 1998) em relação às formas *você* e *tu*, por sua vez, possivelmente devem vir cercadas de formas (possessivas, clíticas, por exemplo) do paradigma de 2SG.

Em relação ao *você*, observamos a sua prevalência em contextos de precedência por formas de *você-sujeito* (185/191, 96,86%), por clíticas, por possessivos de 3SG (103/107, 96,26%), por *você* não-sujeito, complemento (24/27, 88,89%), por vocativo (12/25, 48,00%) e como 1ª ocorrência (11/14, 78,57%), cf. ilustramos de (93) a (100). Observamos, ainda que com menores índices percentuais, o *você* precedido por possessivos e clíticas de 2SG (5oco), por construções imperativas de 2SG com formas de subjuntivo (5oco), pelo *tu-sujeito* (3oco) e pelo *tu não-sujeito* (2oco), cf. verificamos de (101) a (112).

Dados de *você* precedido de *você-sujeito*:

(93) “[...] Se [você] é amigo do atual prefeito Dr. B. Sá, peça a este uma colaboração naquilo que [você] precise. Contrate logo o padre, cujo nome *você* já fez referência [...]” BB. RJ, 10. 05.1987.

Dados de *você* precedidos por clíticas, por possessivos de 3SG:

(94) “[...] Antes de mais nada, agradeço-[lhe] pelo puxão-de-orelha quando *você*, na sua carta, faz referência à miniatura de carta que lhe enviei, mas a razão da miniaturização da carta, deveu-se ao fato de que eu a fiz num dos poucos momentos de folga que tive no trabalho no dia em que lhe escrevi. [...]” AVV. Brasília, 07.04.1983.

(95) “[...] Eu queria um favor [seu]: *você* me perguntar ao Assuero Rêgo se este quer vender a propriedade Primavera, por isso que eu a desejo comprá-la. [...]” ABSB. Fazenda Riacho Fundo-São Pedro-PI, 16.03.1981.

Dados de *você* precedidos por *você* não-sujeito (complemento):

(96) “[...] À medida que o tempo foi passando [...] fui aprendendo a gostar [de você] simplesmente por gostar sem pedir nada em troca [...] o meu sentimento estava tomando conta do meu ser e, pela primeira vez na vida, pensei numa coisa mais séria e *você* foi eleita a pessoa com quem eu pretendia transformar este sonho numa coisa real. [...]” AVV. Brasília, 29.09.1982

(97) “[...] Eu pedia [à você] o obséquio de fazer-me essa distribuição [...] Os 6 restantes exemplares *você* dará a quem entender. [...]” ABSB. Rio de Janeiro-RJ, 29.07.1989.

Dados de *você* precedidos por vocativo:

(98) “[...] [Rodim] pelo que me parece, o que *você* relata nas suas cartas, eu já notei que você está pensando que eu estou usando as minhas cartas como armas de desabafo. Imagino como está magoado e também mudado. Sai Dessa cara! [...]” JPB. São Bernardo do Campo-SP, 18.04.1978.

(99) “[...] Olha [seu cara], *você* esta usando mais gíria do que mesmo eu [...]” JPB. São Bernardo do Campo-SP, 18.04.1978.

Dados de *você* em contexto de 1ª ocorrência:

(100) “[...] Carlos vai trabalhar em que? *Você* já sei que é na Assembléia Legislativa [...]” PQ. Oeiras-PI, 20. 04.1985.

Dados de *você* em contexto de possessivos e clíticos de 2SG (5 oco):

(101) “[...] nãa foi porque eu não queire-[te] so foi a falta de tempo mais *você* discupe a minha falta e mim pertoi que eu não lhe falto mais [...]” VVS. SP, 25.04.1954

(102) “[...] não tenho nada para [te] contar [...] estamos enfrentando coisas tão difíceis que *você* nem imagina [...]” JPB. São Bernardo do Campo-SP, 18.04.1978 [...]”

(103) “[...] em matéria de bagunça nem preciso [te] falar que ela é o máximo não é mesmo, *você* já conhece. [...]” MNRSR. Brasília, 05.05.1978

(104) “[...] um tempinho pra [te] escrever [...] Concurso na Caixa Econômica Federal [...] como *você* sabe, não é de se desprezar [...]” MNRSR. Sobradinho, 26.03.1979.

(105) “[...] as razões que [te] levaram a se desinteressar por mim? Só *você* mesma pode saber. [...]” AVV. Brasília, 29.09.1982.

Dados de *você* em contexto de construções imperativas com formas de subjuntivo:

(106) “[...] Tenho um pouco de pressa na aquisição. [Fale] o tabelião respectivo. Não se trata do Registro de Imóveis. Este eu tenho e de todos os pedaços de posse comprados posteriormente a Março de 1.943. Se existir outras compras lavradas em cartório *você* pode incluir tudo [...]” ABSB. Rio de Janeiro-RJ, 09.01.1989.

(107) “[...] não [esqueça] de falar ao Queiroz Neto [...] ele dirá certamente que *você* precisa de repouso [...]” PNQ. Oeiras-PI, 12.01.1988.

Dados de *você* em contexto de *tu-sujeito* (3 oco):

(108) “[...] [Sabes], naquele dia que escrevi para *você* estava tão insatisfeita interiormente, sabes quando de repente a gente se sente angustiada [...]” MNRSR. Brasília, 05.05.1978

(109) “[...] Se [chegares] em Picos antes [...] Eu tenho quase que certeza que *você* chegará em Picos pelo dia primeiro [...]” VB. Fortaleza-CE, 25.06.1978.

(110) “[...] só chegamos lá onde [tu] estais almejando caso *você* queira. [...]” JPB. São Bernardo do Campo-SP, 18.04.1978

Dados de *você* em contexto de *tu não-sujeito* (2oco):

(111) “[...] Eu gostaria muito, de estar numa praia curtindo [contigo] [...] *Você* é maravilhosa. [...]” VB. Fortaleza-CE, 15.05.1978.

(112) “[...] nos últimos dias tenho sonhado muito [contigo]. [...] numa noite em que estava havendo uma festa, sem que *você* soubesse da minha chegada. [...]”AVV. Brasília, 30.10.1984.

O *tu* mostra-se mais produtivo precedido por formas clíticas e possessivas também de 2SG (112/117, 95,73%), por *tu-sujeito* (70/73, 95,89%), por formas de *tu não-sujeito* (8/10, 80,00%) e por vocativo (13/25 oco, 52,00%), cf. exemplificamos de (113) a (120). Ainda que com baixíssimos níveis de ocorrências, temos o *tu* em contextos de *você-sujeito* (6 oco), do imperativo de 2SG com formas de subjuntivo (7oco), de formas clíticas e possessivas de 2PL (5oco), de formas possessivas e clíticas de 3SG (4oco), de *você não-sujeito* (3oco), de 1ª ocorrência (3oco) e de formas imperativas de 2SG com formas no indicativo (2oco), cf. evidenciamos de (121) a (131).

Dados de *tu* precedido por clíticos e possessivos formalmente de 2SG (112 oco):

(113) “[...] Si soubesse que [te] *affligias* tanto [...]” JMF. Simplicio Mendes-PI, 07.05.1921.

(114) “[...] respondo [tua] amorosa cartinha de 28 de Janeiro [...] Não *imaginas* o prazer que meu coração [...]” CRJM. Fazenda Tatu-PI, 09.01.1921.

Dados de *tu* precedidos por formas de *tu-sujeito* (70 oco):

(115) “[...] Faço vocetos a Deus para que [estejas], completamente curada do defluxo [...] Não *avalias* as saudades que tenho passado [...]” JMF. Simplício Mendes-PI, 01.02.1922.

(116) “[...] não posso nem dar uma corridinha que o coração fica a querer pular pela garganta a cima! Como [vês], não tem condição uma coisa dessa. Afinal o importante não é fazer um curso às pressas, mas sim fazê-lo bem feito. E isso é quase impossível de conseguir com 5 matérias, um curso de inglês e trabalho, não *concordas*? MNRSR. Brasília, 08.10.1978.

Dados de *tu* precedidos por formas de *tu* complemento (*tu não-sujeito*) (8 oco):

(117) “[...] quero saber para ficar orando [por ti], pedindo a Deus que *sejas* feliz e não se esqueça de mim. [...]”. CRJM. Fazenda Tatu-PI, 17.05.1921.

(118) “[...] Sinto-me feliz com isto, mais [por ti] do que por mim, por meu egoísmo e vaidade próprios, pois não sou tão egoísta que exija que só seja meu este amor, mas, por saber que só *puderás* ser feliz se só *amares* a um homem, ao teu esposo. [...]” JMF. Simplício Mendes-PI, 11.10.1920.

Dados de *tu* precedidos por formas vocativas (13 oco):

(119) “[...] [Querido anjo de amor], quanto bem me *fazes* com as tuas carinhosas e amáveis cartas! [...]” JMF. Simplício Mendes-PI, 06.01.1921.

(120) “[...] Olha [Rodim], acredito que a estas alturas já *estejas* aborrecido comigo [...]” JPB. São Bernardo do Campo-SP, 05.06.1978.

Dados de *tu* em contexto de *você-sujeito* (6 oco):

(121) “[...] [Você] sabe, não é? [...] Como *vês*, não tem condição uma coisa dessa. [...]” MNRSR. Brasília, 08.10.1978.

(122) “[...] Primeiro porque foi uma surpresa muito boa que [você] me fez. Pois não imaginavas que *fósses* me escrever mais. por que sempre é o que me *ameaças*. [...]” JPB. São Bernardo do Campo-SP, 05.06.1978.

Dados de *tu* em contexto forma imperativa de 2SG no subjuntivo (7 oco):

(123) (...) Não me [fale] mais em rasgar poesias, com isso vejo que não *confias* em meu amor. [...]” CRJM. Fazenda Tatu-PI, 29.09.1920.

(124) (...) não me [esqueças] por piedade; [tenhas] de mim compaixão! Si me *esqueceres* ou *deixares* de me amar deixarei de viver! [...]” JMF. Oeiras-PI, 16.11.1920.

Dados de *tu* em contextos de clíticos e possessivos formalmente de 2PL (5 oco):

(125) “[...] (...) entrego-[vos] o meu coração [...] se me não tiverdes amor e *sympathia*, si não *sentires* palpitar pormim [...]” JMF. Oeiras-PI, 20.08.1920.

(126) “[...] a certeza de uma donzela si ter unido a mim sem uma *sympathia*, sem uma *affeição* sinsera, uma paixão santa, por isso ouçais ao [vosso] coração antes de me *responderes*. [...]” JMF. Oeiras-PI, 20.08.1920.

Dados de *tu* em contextos de possessivos e clíticos formalmente de 3SG (4 oco):

(127) “[...] Não sei se esta vai [lhe] encontrar em período de aula, mas acho que sim confiando na eficiência dos Correios. [...] Se *chegares* em Picos antes do dia 2 ou 3, ligues pra mim nos dias mencionado [...]” VB. Fortaleza-CE, 25.06.1978.

(128) “[...] Nesses dois dias, seguira até Joazeiro, em propaganda do seu Collegio, o [seu] primo João Leal, porquem vou responder umas cartas do *Senhor Aarão*. [...] Como *tens* passado meu querido amorsinho? *Estas* anciosa por me veres como por te ver estou? Si podesse não me separaria um só instante de ti (...). JMF. Remanso-BA, 05.12.1921. [...]”

Dados de *tu* em contexto de *você* não-sujeito (3 oco):

(129) “[...] (...) naquele dia que escrevi [para você] estava tão insatisfeita interiormente, *sabes* quando de repente a gente se sente angustiada [...]” MNRSR. Brasília-DF, 05.05.1978.

Dados de *tu-sujeito* em contexto de 1ª ocorrência (3 oco):

(130) “[...] Querida Christina Beijo-te apaixonadamente Como *és* boa – meu amorzinho! como te sou grato! Estava já, quase em desespero de saudade, esperando o correio, uma fagulha de esperança a que minh’alma se agarrava como o naufrago se agarra aos destroços do navio para se salvar, convicto de receber noticias tuas. [...]” (JMF. Simpício Mendes, 06.02.1921.)

Dados de *tu-sujeito* em contexto de forma imperativa de 2SG no indicativo (2 oco):

(131) “[...] Me [conta] (em detalhes) como é que foi [...] Como *vês*, ando tão atarefada [...]” MNRSR. Sobradinho, 28.07.1981.

De um modo geral, esses resultados parecem anunciar que os dados de *você* estão harmonizados com formas do paradigma de 3SG. Por outro lado, o *tu* parece preferencialmente combinar-se com formas vinculadas ao paradigma de 2SG, confirmando as discussões de Omena (2003), Lopes & Vianna (2012, p. 148), Rumeu (2013), o que pode sugerir que a produção escrita piauiense, em termos gerais, tende a se mostrar mais aproximada da norma-padrão.

Uma questão interessante é o fato de termos levantado evidências do pronome *você* em convivência com formas do paradigma do *tu* tais como as formas possessivas e clíticas de *tu* (5 oco), expostas de (101) a (105), o próprio *tu* sujeito (3 oco) e o *tu* não-sujeito (2 oco), ilustradas, respectivamente, de (108) a (112). Ainda que com um baixo número de ocorrências, a expressão de um paradigma supletivo orientado pela combinação do *você* com formas de *te-teu* parece evidenciar, na produção escrita piauiense de sincronias passadas, a manifestação mais genuína da norma vernacular do PB atual.

Passamos à descrição-analítica da distribuição dos dados em relação às variáveis extralinguísticas.

3.2 Os fatores extralinguísticos: o sexo, a faixa etária, a escolaridade, os níveis de prática de escrita, as relações sociais dos informantes, os subgêneros das “cartas pessoais”, os assuntos prevalentes e o tempo.

O intuito é observarmos se o sexo (gênero) do escrevente tende ou não a impulsionar a variação *tu/você* nas cartas piauienses novecentistas. No que se refere às cartas piauienses temos um panorama não tão abrangente, uma vez que a distribuição dos dados e das cartas não estão equânimes em relação aos sexos (feminino e masculino). Aos subgêneros textuais (amizade, amor e familiar) correlacionados às faixas etárias (jovens, adultos e idosos). Isso posto, passamos a uma análise descritiva com o intuito principal de descrevermos a interferência das variáveis gênero e faixa etária dos redatores nas cartas novecentistas analisadas.

Tabela 6: As formas de referência ao sujeito de 2SG correlacionadas ao sexo e à faixa etária dos escreventes.

Faixa etária/Sexo	Formas de referência ao sujeito de 2SG			
	Tu		Você	
	Homem	Mulher	Homem	Mulher
Jovem	9/117	62/117	38/117	8/117
(de 14 a 30 anos)	(7,69%)	(52,99%)	(32,48%)	(6,84%)
Total	71/117 (60,68%)		46/117 (39,32%)	
Adulto	161/230	-	68/230	1/230
(de 31 a 50 anos)	(70,00%)		(29,57%)	(0,43%)
Total	161/230 (70%)		69/230 (30%)	
Idoso	1/236	-	235/236	-
(mais de 50 anos)	(0,42%)		(99,58%)	
Total	1/236 (0,42%)		235/236 (99,58%)	
TOTAL	171/583	62/583	341/583	9/583
	(29,33%)	(10,63%)	(58,49%)	(1,54%)
	233/583 (39,97%)		350/583 (60,03%)	

Uma vez que resguardadas as especificidades das amostras das cartas novecentistas piauienses, passamos à descrição dos pronomes-sujeito combinados às variáveis sociais *sexo* e à *faixa etária* dos missivistas. De um modo geral, observamos que o *você* se manifesta, com distintos níveis de produtividade, nas cartas dos jovens (39,32%), dos adultos (30%) e dos idosos (99,58%), o que já nos permite identificar a presença desse inovador pronome também na produção escrita piauiense, ainda que o conservador *tu* tenha sido observado principalmente entre jovens (60,68%) e adultos (70%) e, minimamente, na produção escrita dos redatores idosos em análise (0,42%).

Dentre os redatores jovens, observamos que os homens alcançaram as maiores frequências de uso para o *você* (38/117, **32,48%**), ao passo que as mulheres evidenciam o *tu* (62/117, **52,99%**) em sua produção escrita. Em (132) e (133), trazemos evidências do *você* e do *tu* nas produções de homens e mulheres piauienses, respectivamente.

Dado de *você* por homem jovem

(132) “[...] Eu gostaria muito, de estar numa praia curtindo contigo [...] *Você* é maravilhosa. [...]” VB. Fortaleza-CE, 15.05.1978.

Dado de *tu* por mulher jovem

(133) “[...] *Tu* porque não me decestes antes de sahir que pretendia demorar-se um mez [...]” CRJM. Fazenda Tatu-PI, 14.11.1921.

Na produção escrita dos homens adultos, observamos a alta produtividade do *tu* (161/230, **70,00%**). Por outro lado, verificamos o *você* ganhando espaço funcional, ainda que com um menor índice percentual (68/230, **29,57%**). Em (134) e (135), ilustramos o *tu* e o *você*, respectivamente, através dos punhos de homens adultos. Em relação ao *você*, levantamos tão somente uma única ocorrência por uma mulher adulta (136).

Dado de *tu* por um homem adulto:

(134) “(...) Soffro por ti quirida Christina, por que *teres* d’unir a uma pessoa como eu (...)” JMF. Simplício Mendes-PI, 16.09.1920.

Dado de *você* por um homem adulto:

(135) “(...) soube que *você* havia ligado. Fiquei aborrecido, pois quase não saio de casa e no momento que o faço acontece de *você* ligar. (...)” AVV. Brasília, 28.02.1984.

Dado de *você* por uma mulher adulta:

(136) “[...] Comadre, por aqui nada de importante tudo como sempre [...] pois *você* sabe como estão as coisas. [...]” JSL. Santo André-SP, 15.02.1964.

Como uma das questões do trabalho do pesquisador voltado às fontes linguísticas do passado é termos de nos restringirmos aos textos que nos restaram no interior dos acervos (públicos ou privados), não foi possível equilibrarmos as amostras equanimemente em relação ao número de cartas por homens e mulheres e por faixa etária. De qualquer modo, o que podemos constatar é a distribuição do *você* por todos os níveis etários, o que já nos pode sinalizar, a presença do *você* na produção escrita dos redatores piauienses de todas as idades (jovens, adultos e idosos).

A fim de tentarmos elucidar o que se mostra na interface *sexo, faixa etária e escolaridade*, passamos ao cruzamento dessas três variáveis sociais em relação à regra variável binária *tu/você*.

Tabela 7: As formas *tu* e *você* em relação ao sexo, à faixa etária e ao nível de escolaridade dos redatores.

Níveis de escolarização em relação ao gênero e à faixa etária	Formas de referência ao sujeito de 2SG	
	Tu	Você
Homem		
	Jovem	
Alfabetizado	-	10/10 (100%)
Ensino médio	9/37 (24,32%)	28/37 (75,68%)
Total	9/47 (19,15%)	38/47 (80,85%)
	Adulto	Tu
		Você
Alfabetizado	-	7/7 (100%)
Ensino médio	-	55/55 (100%)
Ensino universitário	161/167 (96,41%)	6/167 (3,59%)
Total	161/229 (70,31%)	68/229 (29,69%)
	Idoso	Tu
		Você
Alfabetizado	1/125 (0,80%)	124/125 (99,20%)
Ensino universitário	-	111/111 (100%)
Total	1/236 (0,42%)	235/236 (99,58%)

Mulher

	Jovem	Tu	Você
Alfabetizada		43/43 (100%)	-
Ensino médio		19/27 (70,37%)	8/27 (29,63%)
Total		62/70 (88,57%)	8/70 (11,43%)
	Adulta	Tu	Você
Alfabetizada		-	1/1 (100%)
Total			1/1 (100%)
		233/583	350/583
TOTAL		(39,97%)	(60,03%)

Ao cruzarmos as variáveis sociais *sexo*, *faixa etária* e *escolaridade* dos redatores, observamos em relação especificamente ao *você* o seu uso categórico e prevalente pelos homens jovens alfabetizados e com ensino médio a partir dos índices percentuais de 100% (10/10 oco) e 75,68% (28 oco), respectivamente. Para os redatores adultos observamos, dentre os dados de *você*, o seu uso categórico pelos redatores com ensino médio (55/55 oco, 100%) e alfabetizados (7/7, 100%), ainda que informantes com o ensino superior preferiram o *tu*, em 96,41% dos dados (161/167oco). Mais interessante são os resultados para os redatores idosos alfabetizados (124/125, 99,20%) e com o ensino superior (111/111, 100%) cujas evidências nos mostram a prevalência do *você* independentemente do nível de escolarização dos redatores, cf. (137) e (138). No âmbito da produção escrita das mulheres piauienses, observamos pouquíssimas evidências do *você*: tão somente 8 ocorrências produzidas por uma única missivista (dentre as três únicas missivistas) com o ensino médio e atuante, profissionalmente, como funcionária pública.

Considerando o baixíssimo número de dados de *você* na produção escrita feminina, o que se deu em virtude do fato de termos sido contemplados com o levantamento da produção escrita de tão somente três missivistas, foi-nos possível observar a distribuição dos dados de *você* pelos redatores jovens, adultos e idosos independentemente do seu nível de escolarização (alfabetização, médio e superior). Ainda que timidamente porque em análise temos uma amostra de cartas não equanimemente distribuída por todas as

décadas do século XX e por todas as faixas etárias e sexo dos redatores, isso parece nos apontar que, na era novecentista do PB, o *você* já estava implementado, estando presente e prevalecendo, em 60% dos dados, na expressão escrita piauiense, independentemente do nível de escolarização do seu redator.

Dado de *você* na produção escrita de redator homem com 68 anos alfabetizado:

(137) “(...) Quando *você* vier a Oeiras, todos ficaremos alegres, se soubermos sempre que *você* foi um menino bom (...)” PNQ. Oeiras-PI, 12.03.1975.

Dado de *você* na produção escrita de redator homem com 62 anos e com o nível superior completo:

(138) “[...] Dizem que os homens de inteligência são bons. *Você* confirma a sentença. [...]” JATF. Teresina-PI, 27.07.1986.

Por outro lado, o *tu* prevalece não só entre os homens adultos universitários, mas também entre as mulheres jovens alfabetizadas, cf. apresentamos em (139) e (140).

Dado de *tu* na produção escrita de redator homem adulto (31 anos) e com o nível superior completo:

(139) “(...) Si me não *tiveres* *sympathia*, não se sacrificai por mim (...)” JMF. Fazenda Tatu (Oeiras-PI), 19.08.1920.

Dado de *tu* na produção escrita de mulher jovem (19 anos) e alfabetizada (redatora):

(140) “(...) e *tu* o que *és*? Se não *és* poeta para as outras *és* para mim. (...)” CRJM. Fazenda Curral de Pedras-PI, 11.03.1921.

Passamos à análise dos níveis de prática de escrita dos redatores das missivas piauienses em análise.

Tabela 8: As formas *tu* e *você* em relação às práticas de escrita dos redatores.

Práticas de escrita (níveis)	Pronomes	
	Tu	Você
Alta	225/534 (42,13%)	309/534 (57,87%)
Média	-	18/18 (100%)
Baixa	8/31 (25,81%)	23/31 (74,19%)
Total	233/583 (39,97%)	350/583 (60,03%)

Os redatores com *alto* nível de habilidade em sua expressão escrita parecem atuar indiscriminadamente no âmbito da alternância entre as formas *tu* e *você*. Parece que os redatores mais habilidosos não só em relação ao repertório lexical, mas também em relação à composição morfossintática das sentenças na perspectiva da norma-padrão (norma subjetiva) mostram-se, nas cartas em análise, propensos ao *você* (309/534, **57,87%**), ainda que em intensa inclinação para o *tu* (225/534, 42,13%). Passamos, na imagem (23) e na sua transcrição (141), às evidências de um alto índice de prática de escrita corporificado em traços da carta ABSB. São eles: o uso conservador do pretérito mais que perfeito (“pedira”), do futuro etimológico (“demorarei”, “completarei”), do pronome relativo “onde” na retomada de lugar (“... ao Rio _ onde me demorarei...”) e a ênclise em perífrases ao verbo auxiliar (“... mande-me dizer...”).

Imagem 23: Excerto da carta de JATF.

Teresina-PI, 05.03.1976.

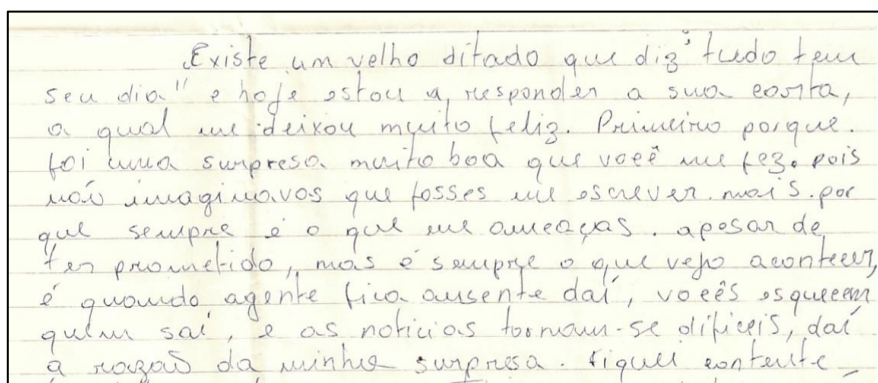
Teresina, 5 de Março de 1976
 Prezado amigo Dr. Exedito Rêgo:
 Meu distinto poeta:
 A notícia sobre a minha tia Maria Stella e redigida aí na conformidade de que eu lhe pedira, foi publicada n' O Cometa de Outubro último, como você deve estar ciente.
 De passagem por Teresina, vindo do Riacho Fundo e com destino, por estes dias, ao Rio - onde me demorarei uns 15 dias - estou lhe enviando a quantia de Cr\$ 30,00 para pagar a assinatura d' O Cometa no corrente ano. E se acontecer a quantia remetida não bastar para a assinatura - pois esta pode ter sido alterada, - mande-me dizer que eu completarei a diferença.

(141) “Teresina, 5 de Março de 1976 Prezado amigo Dr. Exedito Rêgo: Meu distinto poeta: A notícia sobre a minha tia Maria Stella e redigida aí, na conformidade de que eu lhe pedira, foi publicada n' O Cometa de Outubro último, como *você* deve estar ciente. De passagem por Teresina, vindo do Riacho Fundo e com destino, por estes dias, ao Rio, - onde me demorarei uns 15 dias, - estou lhe enviando a quantia de Cr\$ 30,00 para pagar a assinatura d' O Cometa, no corrente ano. E se acontecer a quantia remetida não bastar para a assinatura, -pois esta pode ter sido alterada, - mande-me dizer que eu completarei a diferença.” (JATF. Teresina-PI, 05.03.1976)

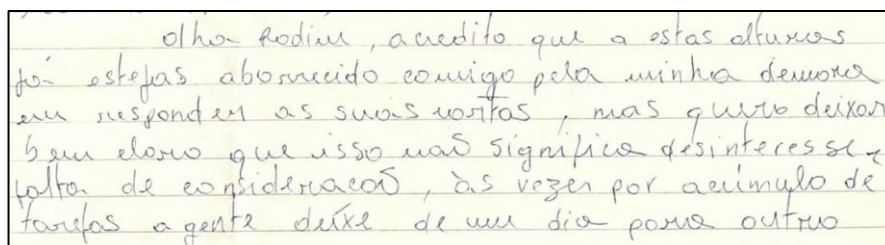
Os redatores interpretados como *medianamente* habilitados em relação a sua expressão mostram-se categoricamente propensos ao *você-sujeito*, ainda que tenhamos levantado pouquíssimas ocorrências, tão somente dezoito dados (18/18, 100%). À título de ilustração acerca da produção escrita de um redator *medianamente* habilitado (nível mediano) trazemos à cena, na imagem 24 e no excerto 142, evidências de alguns traços linguísticos que o evidenciam como um legítimo detentor da norma objetiva do PB. Trata-se das seguintes: a convivência (“mistura” nos termos tradicionais) das formas *você* e *tu*, o uso do possessivo “sua” na referência ao interlocutor (“hoje estou a responder a *sua* carta”, “pela minha demora em responder as suas cartas”). Além desses aspectos morfossintáticos do PB, observamos não só aspectos gráficos tal como a escrita do pronome “a gente” junto (agente) e separado “a gente” (“... é quando *agente* fica ausente daí...”, “às vezes por acúmulo de tarefas *a gente* deixa de um dia para o outro”), mas

também aspectos de pontuação como um ponto final de teor continuativo separando a oração principal da oração subordinada adverbial causal (“[...] foi uma surpresa muito boa que *você* me fez. Pois não *imaginavas* que *fosses* me escrever mais. por que sempre é o que me ameaças.”). Além disso, temos ainda a ausência de vírgulas duplas, ao anunciar o seu interlocutor (“Rodim”) através do vocativo (“Olha Rodim, acredito que a estas alturas já *estejas* aborrecido comigo [...]”) e o fechamento do parágrafo com a ausência do ponto final no seguinte trecho: “[...] às vezes por acúmulo de tarefas a gente deixe de um dia para o outro”.

Imagem 24: Excerto da carta de JPB.
São Bernardo do Campo, SP, 05.06.1978.



[...]



(142) “[...] Existe um velho ditado que diz “tudo tem seu dia” e hoje estou a responder a **sua** carta, a qual me deixou muito feliz. Primeiro porque foi uma surpresa muito boa que *você* me fez. Pois não *imaginavas* que *fosses* me escrever mais. por que sempre é o que me ameaças. apesar de ter prometido, mas é sempre o que vejo acontecer, é quando agente fica ausente daí, vocês esquecem quem sai, e as notícias tornam-se difíceis, daí a razão da minha surpresa.

[...]

Olha Rodim, acredito que a estas alturas já *estejas* aborrecido comigo pela minha demora em responder as *suas* cartas, mas quero deixar bem claro isso não significa desinteresse e falta de consideração, às vezes por acúmulo de tarefas a gente deixa de um dia para o outro [...]

(JPB. São Bernardo do Campo, SP, 05.06.1978)

Os redatores que se mostram altamente habilidosos em relação às práticas de escrita apresentam-se em intensa alternância entre as formas *tu/você*, ainda que tenham preferido o *você*, em 57,87% dos dados (309 oco), cf. ilustramos na imagem 25 correspondente ao excerto (143).

Imagem 25: Excerto da carta de ABSB.

RJ, 03.09.1984.

[...]

(143) “[...] O segundo pedido: receber o que *você* está me devendo; é a reportagem sobre o que houve por ocasião do meu aniversário (70º). Houve, como *você* sabe porque colaborou no caso, a 21 de Maio de 1977 uma missa solene, - solene e original se assim posso dizer – na Igreja de N. Senhora da Conceição. [...]”

[...]

[...] Mas eu posso ter me enganado quando escrevi a narrativa, há uns 3 anos. A verdade, porém, poder-se-á apurar através dos jornais da época, de Terezina. Em Novembro de 1927 eu já era redator do jornal O Piauí, para cuja redação eu entrara, como revisor, em 1926 (Agosto). [...]”

(ABSB. RJ, 03.09.1984)

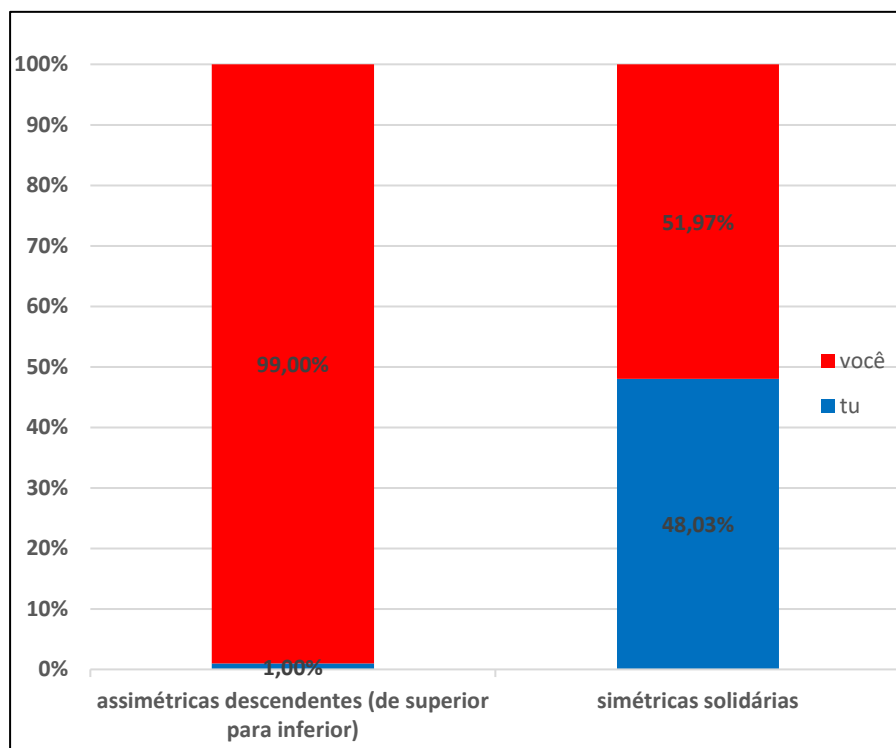
As evidências do alto grau de envolvimento do redator com a língua escrita podem ser observadas principalmente em virtude de dois aspectos morfosintáticos. São eles: o uso da mesóclise (“poder-se-á apurar através dos jornais da época”), do pretérito mais que perfeito (“eu entrara”) e do pronome relativo “cujo” em consonância com a regência do verbo *entrar* (“entrar para”) em virtude do encaixamento entre as orações: “[...] eu já era redator do jornal O Piauí, para cuja redação eu entrara, como revisor [...]”.

Tabela 9: As formas de referência ao sujeito de 2SG correlacionadas às relações sociais simétricas e assimétricas (descendentes).

Relações Sociais	Formas de referência ao sujeito de 2SG	
	Tu	Você
Assimétrica descendente (de superior para inferior)	1/100 (1,00%)	99/100 (99,00%)
Simétrica solidária (entre iguais)	232/483 (48,03%)	251/483 (51,97%)
TOTAL	233/583 (39,97%)	350/583 (60,03%)

Assumimos, em harmonia com Souza (2021, p. 95-96), a hipótese de que o *você* se mostraria proficiente nas relações assimétricas descendentes piauienses (de superior para inferior), e nas simétricas, cf. discutido por Lopes & Rumeu (2015, p. 23), conforme observado para as cartas mineiras dos séculos XIX e XX. Para o *tu*, entendemos que as relações simétricas o contextualizariam, tendo em vista a conjectura de que o *tu* e o *você* incorporariam uma dinâmica variável como também observado por Souza (2012, p. 86) em relação às missivas cariocas (séculos XIX e XX). Além disso, temos o fato de as cartas amorosas nutridas por casais (entre marido-mulher, entre noivos, entre namorados) apresentarem-se preferencialmente orientadas ao uso do *tu*, cf. discutido por Silva (2012), Lopes, Marcotulio & Oliveira (2018) para as missivas cariocas.

Gráfico 2: As formas *tu* e *você* correlacionadas às dinâmicas sociais de simetria e assimetria social descendente nas cartas piauienses.



De um modo geral, observamos que, nas relações simétricas solidárias, o *você* alterna, em 51,97% dos dados (251/483), com o *tu*, que, por sua vez, se mostra em 48,04% dos dados (232/483). Além desse contexto de simetria social, observamos que o *você*, no âmbito das assimétricas descendentes prevalece quase categoricamente, em 99,00% dos dados (99/100), ficando o *tu* restrito a 1% dos dados (1 oco). Assim sendo, esses resultados parecem evidenciar o *você* ter se irradiado pelas relações simétricas e assimétricas, equiparando-se, em certa medida, ao *tu* nas simétricas (solidárias) como a forma etimologicamente marcada para as dinâmicas sociais marcadas pela intimidade tal como o *tú/vous* do francês (BROWN & GILMAN, 1960). Em relação ao *você*, observamos a sua produtividade como uma forma neutra nas relações assimétricas descendentes e simétricas, ou seja, como uma forma menos invasiva que o íntimo e pessoal *tu* (como uma estratégia coringa) para a referência ao interlocutor, o que pode ser interpretado como persistência do tratamento polido que a originou (“*vossa mercê*”), cf. discutido por Rumeu (2011), com uma estratégia de cortesia descendente como também

discutido por Souza (2021, p. 96-97) em relação às cartas mineiras (séculos XIX e XX) e por Souza (2012, p. 139) em relação às cartas cariocas (séculos XIX e XX). Apresentamos, em (144), um único dado da forma *tu* nas relações assimétricas descendentes, em (145), o *você* em relação assimétrica descendente. De (146) a (148), temos evidências das formas *tu* e *você* nas simétricas solidárias, respectivamente.

Dado de *tu* em relação assimétrica descendente (relação social entre avô e neto):

(144) “Querido e sempre lembrado Neto CARLOS [...] Faça o concurso para o BANCO DO BRASIL, CAIXA ECONÔMICA ou qualquer atividade honesta e lucrativa. [...] Se *queres* vencer, desperta e luta. [...]” PNQ. Oeiras-PI, 18.07.1987.

Dado de *você* em relação assimétrica descendente (relação social entre avó e neta):

Imagem 26: Carta de PNQ.

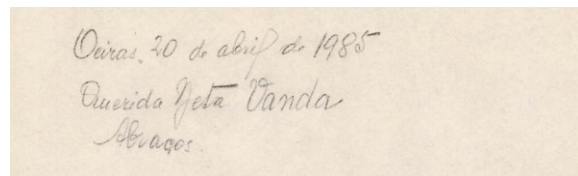
Oeiras, 20.04.1985.

(145) “[...]”

Oeiras, 20 de abril de 1985

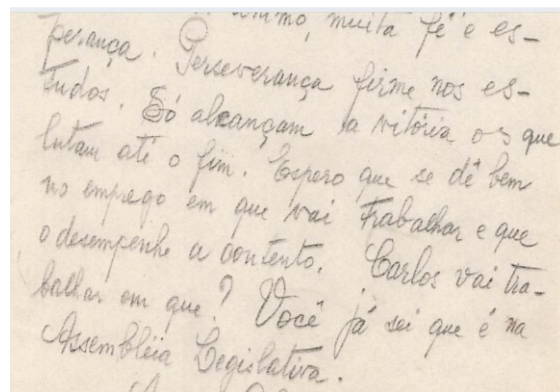
Querida Neta Vanda

Abraços



[...]

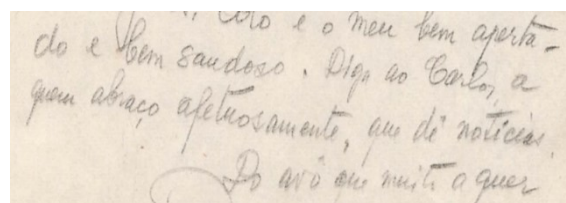
[...] Perseverança firme nos estudos. Só alcançam a vitória os que lutam até o fim. Espero que se dê bem no emprego em que vai trabalhar e que o desempenhe a contento. Carlo vai trabalhar em que? *Você* já sei que é na Assembleia Legislativa.



[...]

[...] Diga ao Carlos, a quem abraço afetuosamente, que dê notícias.

Do avô que muito a quer”



Dado de *tu* em relação simétrica solidária (relação entre noivos):

(146) “[...] Deca, amado noivo Beijo-te com amor Respondo tua estimada cartinha vinda por Dimas. Não sei o que hei de fazer quiridinho para *receberes* as minhas cartas. [...]” CRJM. Fazenda Tatu-PI, 21.05.1921.

Dado de *você* em relação simétrica solidária (relação entre amigos):

(147) “Mestre Possidônio [...] Só os homens superiores compreendem o dever de ajudar o próximo - e *você*, Mestre, é um deles. Quanta sabedoria na bondade! [...]” AT. Teresina-PI, 31.01.1987.

Dado de *você* em relação simétrica solidária (relação entre amigos):

(148) “Prezado cunhado e amigo Vicente um forte abraço. e mil saudações. [...] Vicente recebemos uma cartinha sua ano passado e desde desse tempo que esperamos por você que nos disse na carta que vinha em passeio por aqui e nunca mais veio [...] Sabendo eu de que *você* não pretende mais morar aqui peço ao amigo que chêgue até aqui para *você* deslindrar suas partes de terra [...]” LJJ. Bocaina-PI, 22.06.1951.

Tendo em vista as 583 ocorrências das formas *tu* e *você*, observamos, em termos gerais, uma estimulante concorrência entre as formas *tu* (233 oco) e *você* (250 oco) distribuídas no decorrer do tempo, cf. apresentamos na tabela 10.

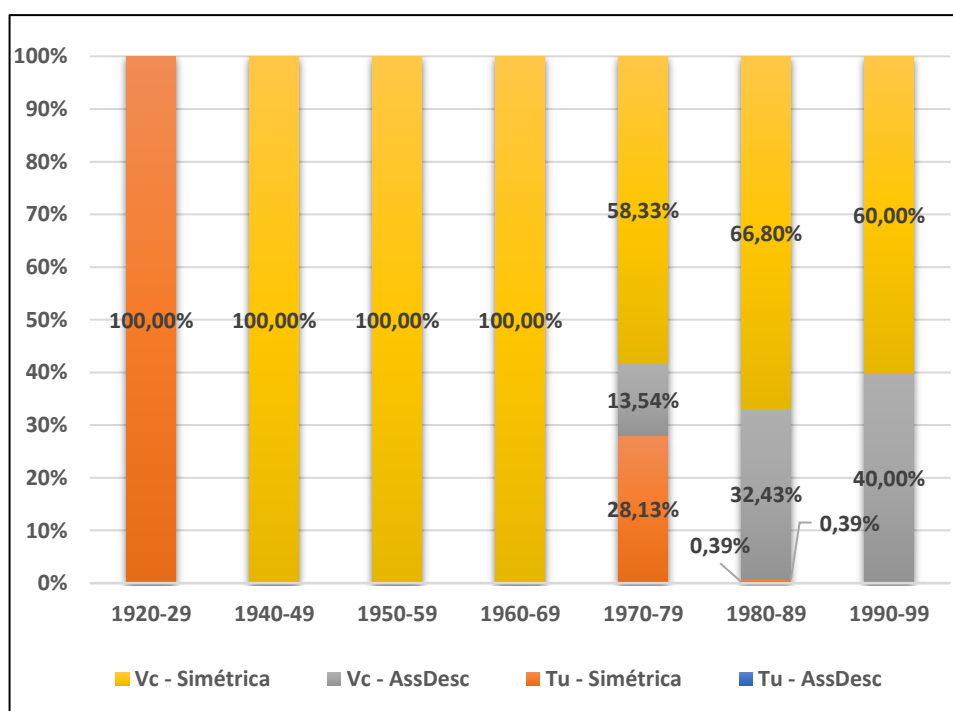
Tabela 10: Distribuição das formas de referência ao sujeito de 2SG (*tu* e *você*) pelas cartas piauienses no eixo do tempo: de 1920 a 1990.

		SÉCULO XX (1920-1990)						TOTAL (136)	
ANOS		1920-29	1940-49	1950-59	1960-69	1970-79	1980-89		1990-99
Nº CARTAS		36 cartas	1 carta	5 cartas	1 carta	19 cartas	71 cartas	3 cartas	
DADOS	TU	204/233 (87,55%)	-	-	-	27/233 (11,59%)	2/233 (0,86%)	-	233/583 (39,97%)
	VOCÊ	-	1/350 (0,29%)	17/350 (4,86%)	1/350 (0,29%)	69/350 (19,71%)	257/350 (73,43%)	5/350 (1,43%)	250/583 (60,03%)

Para os dados da década de 20 do século XX, observamos que o casal de noivos (ele, com formação superior; ela, apenas escolarizada, na década de 1920), prefere o *tu*, em 87,55% (204/233) dos dados analisados. Para os demais períodos históricos, posteriores à década de 20 século XX, temos o *você*-exclusivo para os informantes cuja produção está compreendida principalmente entre as décadas de 1970 e 1980 (19,71% de *tu* e 73,43% de *você*), o que se dá principalmente em virtude de termos conseguido levantar mais cartas justamente no período em questão e, conseqüentemente, o número ínfimo de dados para as décadas de 40, 50, 60 e 90 do século XX para os quais levantamos 1 ocorrência (0,29%), 17 ocorrências (4,86%), 1 única ocorrência (0,29%) e 5 ocorrências (1,43%), respectivamente.

Uma vez que passamos pela distribuição dos dados de *tu* e de *você* não só pelas relações sociais levantadas nessa amostra (simétrica e assimétrica) através do gráfico 2, mas também pelo eixo do tempo (1920-29, 1940-49, 1950-59, 1960-69, 1970-79, 1980-89, 1990-99), através da tabela 12, passamos ao cruzamento dessas variáveis. Como constatamos, no gráfico 2, a alternância *tu* e *você* nas relações simétricas, apesar de o *você* também ter se mostrado, nas relações assimétricas descendentes, praticamente categórico, o intuito é detectarmos em qual(is) período(s) é observada a dinâmica da variação nas cartas piauienses em análise (gráfico 3).

Gráfico 3: Distribuição das formas *tu* e *você* pelas relações sociais piauienses (simétricas e assimétricas descendentes).



Ao observarmos a distribuição dos dados de *tu* e de *você* em função dos eixos das relações sociais distribuídas no eixo do tempo, observamos que, de um modo geral, temos o uso categórico do *tu* simétrico, entre os anos de 1920-29 (204/204 oco), e do *você* simétrico, nas décadas de 40 (1/1 oco), 50 (17/17 oco) e 60 (1/1 oco) do século XX.

Nas cartas das décadas de 70, 80 e 90 do século XX, verificamos a alternância *tu/você* distribuída pelas relações sociais de simetria e assimetria descendente. Na década de 70, temos a convivência do *você* simétrico (58,33%, 56/96 oco) com o *tu* simétrico (28,13%, 27/96 oco) e com *você* assimétrico, em 13,54% (13/96 oco) dos dados (assimetria descendente), cf. ilustramos em (149), (150) e (151). Na década de 80, temos o *você*, nos contextos de simetria (66,80%, 173/259 oco) e de assimetria descendente (32,43%, 84/259 oco), ficando o *tu* minimamente evidente nos contextos de simetria (0,39%, 1/259 oco) e de assimetria descendente (0,39%, 1/259 oco), cf. apresentamos em (152) a (155). Entre os anos de 1990 e 1999, observamos a prevalência do *você* no contexto de simetria social (60%, 3/5 oco), apesar de, em contexto de assimetria descendente (40%, 2/5 oco), o *você* também ter sido acionado na escrita piauiense em análise, cf. exemplificamos em (156) e (157).

Cartas da década de 70:

Dado de *você* simétrico na década de 70:

(149) “[...] Olha talvez *você* nunca mais irá falar comigo [...] Olha seu cara, *você* esta usando mais gíria do que mesmo eu. se *você* viesse para São Paulo, talvez não sentisse uma dificuldade para se adaptar nas gírias [...]” JPB. São Bernardo do Campo-SP, 18.04.1978. (Carta entre amigos)

Dado de *você* em contexto de *assimetria descendente* na década de 70:

(150) “[...] Não queira ser um rapaz vadio (...) Quando *você* vier a Oeiras todos ficaremos alegres, se soubermos sempre que / *você* foi um menino bom, obediente, estudioso, respeitador. [...]” PNQ. Oeiras-PI, 12.03.1975. (Carta de um amigo mais velho a um amigo mais jovem em tom aconselhativo)

Dado de *tu* simétrico na década de 70:

(151) “[...] Me fale aí da nova faculdade, o que está achando. Da vida noturna aí de Teresina, afinal *deves* estar saindo muito, não? É bom umas férias, Vete, penso que depois de quatro meses *você* vai ter um tempinho para a vida noturna [...]” MNRSR. Sobradinho-DF, 17.03.1979. (Carta entre primas)

Cartas da década de 80:

Dado de *você simétrico* na década de 80:

(152) “[...] O único consolo é que nos últimos dias tenho sonhado muito contigo. [...] sonhei que eu chegava em Dom Expedito, numa noite em que estava havendo uma festa, sem que *você* soubesse da minha chegada. [...]” AVV. Brasília, 30.10.1984. (Carta entre namorados)

Dado de *você* em contexto de *assimetria descendente* na década de 80:

(153) “[...] Na carta *VOCE* explicou o motivo da demora, é que vieram receber os convites com atraso. [...]” PNQ. Oeiras-PI, 15.01.1985. (Carta entre avô e neto)

Dado de *tu simétrico* na década de 80:

(154) “[...] Como *vês*, ando tão atarefada que não tenho nem mesmo tempo pra namorar [...]” MNRSR. Sobradinho-DF, 28.07.1981. (Carta entre primas)

Dado de *tu* em contexto de *assimetria descendente* na década de 80:

(155) “[...] “Querido e sempre lembrado Neto CARLOS [...] Faça o concurso para o BANCO DO BRASIL, CAIXA ECONÔMICA ou qualquer atividade honesta e lucrativa. [...] Se *queres* vencer, desperta e luta. [...]” PNQ. Oeiras-PI, 18.07.1987. (Carta entre avô e neto)

Cartas da década de 90:

Dado de *você simétrico* na década de 90:

(156) “[...] Mas eu gostaria que *você* me mandasse o Camoneano e uma poesia notável saudando a Ilha da Madeira [...] Eu irei falar, irei fazer um discurso, e, - eu espero que *você* também fale, fazendo uma brilhante peça literária – durante os festejos e homenagens em louvor do centenário de Tapety. [...]” ABSB. RJ, 10.10.1990. (Carta entre amigos)

Dado de *você* em contexto de *assimetria descendente* na década de 90:

(157) “[...] Não sei como agradecer sua mensagem, a mim enviada pelo transcurso do “DIA DOS PAIS”. Bela, vazada numa linguagem emocionante, pôs nela *você* todas as cordas do coração bondoso. [...]” PNQ. Oeiras-PI, 17.08.1990. (Carta entre pai e filha)

O objetivo do controle dos subgêneros das cartas pessoais é o de tentar interpretar se haveria alguma conexão entre os pronomes *tu* e *você* nos contextos das cartas amorosas, de amizade e familiares piauienses, cf. já testado por Souza (2012) e Souza (2021) e em relação às cartas cariocas e mineiras, respectivamente. A hipótese, à luz de

Lopes, Marcotulio & Oliveira (2018, p. 41), é a de que as cartas amorosas favoreçam o *tu*. De (158) a (160), ilustramos o *tu*, em carta amorosa, e o *você*, em cartas de amizade e familiar, respectivamente.

Tabela 11: As formas de referência ao sujeito de 2SG (*tu* e *você*) distribuídas pelas cartas amorosas, de amizade e familiares.

As cartas pessoais e os seus subgêneros	Formas de referência ao sujeito de 2SG	
	Tu	Você
Cartas amorosas	204/280 (72,86%)	76/280 (27,14%)
Cartas de amizade	9/172 (5,23%)	163/172 (94,77%)
Cartas familiares	20/131 (15,27%)	111/131 (84,73%)
Total	233/583 (39,97%)	350/583 (60,03%)

Dado da forma *tu* em carta amorosa:

(158) a. “[...] que com o teu convívio me corrija destes defeitos e te trate como mereces. [...]” JMF. Simplício Mendes-PI, 16.09.1920.

b. “Querida Otacília Abraços [...] Soube, por Ceíça, de uma notícia que muito me alegrou. *Você* sairá daí, no dia 29, devendo chegar a Oeiras no dia 31 deste. [...] Do esposo que a estima [...]” PNQ. Oeiras-PI, 19.01.1988.

Dado da forma *você* em carta de amizade:

(159) “[...] Se *você* é amigo do atual prefeito Dr. B. Sá, peça a este uma colaboração naquilo que *você* precise. Contrate logo o padre, cujo nome *você* já fez referência [...]” ABSB. RJ, 10.05.1987.

Dado da forma *você* em carta familiar:

(160) “[...] *Você* pensou que fosse dia 09/05, né, errou só por 4 dias, faço aniversário é hoje. [...]” MNRSR. Brasília, 05.05.1978.

Nas cartas amorosas piauienses, observamos a prevalência do íntimo *tu* (72,86%), cf. também observado por Souza (2012), Silva (2012), Souza (2021) em relação às missivas cariocas e mineiras dos séculos XIX e XX. Por outro lado, nas cartas de amizade e familiares, o *você* prevalece com índices de produtividade altos e próximos (94,77% (111/131), 84,73% (**163/172**)), ainda que o *você* tenha prevalecido principalmente nas cartas familiares (94,77%), o que parece confirmar a hipótese inicial de as cartas amorosas representarem o contexto de intimidade propulsor do *tu*.

Uma vez expostos os resultados da distribuição das formas *tu* e *você* em relação aos subgêneros das cartas pessoais, passamos à análise dos assuntos (tópicos) das cartas pessoais que contextualizam os usos das formas pronominais *tu* e *você* na escrita piauiense novecentista.

Tabela 12: As formas de referência ao sujeito de 2SG (*tu* e *você*)
distribuídas pelas temáticas das cartas pessoais piauienses

Temáticas prevalentes das cartas pessoais	Formas de referência ao sujeito de 2SG	
	Tu	Você
Questões amorosas	204/257 (79,38%)	53/257 (20,62%)
Questões familiares	10/141 (7,09%)	131/141 (92,91%)
Questões literárias	-	87/87 (100%)
Questões políticas e/ou econômicas	5/49 (10,20%)	44/49 (89,80%)
Questões de viagens e festas (convites)	14/49 (28,57%)	35/49 (71,43%)
Total	233/583 (39,97%)	350/583 (60,03%)

A alternância entre as formas *tu* e *você* mostra-se mais produtiva nas cartas orientadas pela semântica das questões amorosas, cabendo ao *tu* o maior nível de produtividade (79,38%, 204/257), ainda que em variação com o *você* (20,62%, 53/257). Esse resultado vai ao encontro do fato de as cartas amorosas terem se mostrado como o contexto de resistência do *tu*, como já verificado por Pereira (2012) em relação à *família Penna*, por Silva (2012) em relação ao casal de namorados *Jaime e Maria*, corroborando

a hipótese inicial em análise. Ilustramos, em (161) e (162), evidências das formas *tu* e *você* no contexto semântico das questões amorosas.

(161) Dado da forma *tu* em carta de temática amorosa:

“Adorada Christina. Saúdo-a respeitosamente [...] Soffro por ti querida Christina, por que sendo uma flor carinhosa e boa, um anjo perfeito virtuoso, acostumada a esta vida de carícia e ternura, com pessoas de tractos finos, *teres* d’unir a uma pessoa como eu, acostumado nesta vida rustica, e portanto grosseiro e indilicado; sem carinho e affectos. [...]” JMF. Simplício Mendes-PI, 16.09.1920.

(162) Dado da forma *você* em carta de temática amorosa:

“Liinha [...] não deixa de ser bem gratificante receber uma carta sua, saber como *você* está [...] Em outras palavras, eu quero dizer que, embora seja muito bom ficarmos escrevendo cartas um para o outro, estas já não resolvem os meus problemas, pois eu quero é ter você perto de mim. [...]” AVV. Brasília, 23.09.1981.

Nos demais contextos das questões familiares (92,91%), literárias (100%), políticas e/ou econômicas (89,80%) e de viagens/festas (71,43%), observamos a prevalência do *você*. Apresentamos, de (163) a (166), dados que ilustram o *você*.

(163) Dado da forma *você* em carta de temática familiar:

“[...] Querida Neta Vanda Abraços VANDINHA, recebi o cartão que me mandou. Um postal da Assembléia Legislativa. Belo prédio. Fiquei sabendo que a querida neta trabalha nesse prédio. Muita satisfação nos trouxe as suas letras, bem como o saber que já está trabalhando. Estude, minha Filha, estude sempre, mas sem prejudicar a saúde. Aprenda e passe, e muito bem, no vestibular para a carreira escolhida. Otacília manda lhe dizer que recebeu tudo o que *VOCÊ* lhe remetera. Está muito satisfeita. Escreverá depois. Está agora muito atarefada. [...]” PQ. Oeiras-PI, 10.05.1985.

(164) Dado da forma *você* em carta de temática literária:

“[...] Meu caríssimo escritor – Possidônio Queiroz: [...] Eu irei falar, irei fazer um discurso, e, - eu espero que *você* também fale, fazendo uma brilhante peça literária – durante os festejos e homenagens em louvor do centenário de Tapety. [...] Logo que seja publicado o meu discurso de saudação ao escritor J. Romão da Silva, pronunciado em 21-5-1990, eu lhe mandarei o órgão de comunicação a publicar. [...]” BSB. RJ, 10.10.1990.

(165) Dado da forma *você* em carta de temática política e/ou econômica:

“[...] Golbery é um homem de alto valor mental. Sociólogo, pensador e escritor de vastos recursos. Acresce que é um homem de prestígio político e poderá, portanto, auxiliar o nosso Instituto Histórico, amanhã, numa oportunidade qualquer. Eu votaria, por exemplo, para ele ser nosso associado como membro correspondente ou receber uma homenagem qualquer do Instituto. Golbery é um homem de alto valor mental. Sociólogo, pensador e escritor de vastos recursos. Acresce que é um homem de prestígio político e poderá, portanto, auxiliar o nosso Instituto Histórico, amanhã, numa oportunidade qualquer. Eu votaria, por exemplo, para ele ser nosso associado como membro correspondente ou receber uma homenagem qualquer do Instituto. [...] Mudando de assunto. Há um fato que desejo lhe explicar. Você, aliás, quando eu estava aí (no dia 25-1-1981) teve ocasião, em caráter informal, de fazer referência sobre o acontecido...” ABSB. Fazenda Riacho Fundo (São Pedro do Piauí-PI), 28.02.1981.

(166) Dado da forma *você* em carta de temáticas relacionadas às questões de viagens e festas (convites):

“[...] Ciente de todas ocorrências que *você* descreve com aquela ordem de que você é mestre. [...] Estou lhe remetendo, à parte, um impresso contendo 4 exemplares, dactilografados, da minha conferencia sobre o Visconde da Parnaíba. [...]” ABSB. Oeiras-PI, 21.05.1907.

Tabela 13: Distribuição das formas *tu* e *você* pelos acervos piauienses e pelos subgêneros das cartas pessoais.

Acervos documentais	Formas de referência ao sujeito de 2SG	
	Tu	Você
Cartas de Amor		
Acervo “Moura Fé”	204/204 (100%)	-
Acervo “Queiroz”	-	21/21 (100%)
Acervo “Vieira do Vale”	-	55/55 (100%)
Cartas de amizade		
Acervo “autores piauienses”	-	118/118 (100%)
Acervo “cartas pessoais variadas”	8/29 (27,59%)	21/29 (72,41%)
Acervo “Queiroz”	-	16/16 (100%)
Acervo “Valmir Batista”	1/9 (11,11%)	8/9 (88,89%)

Cartas familiares	Tu	Você
Acervo “autores piauienses”	-	1/1 (100%)
Acervo “cartas pessoais variadas”	-	20/20 (100%)
Acervo “Queiroz”	1/83 (1,20%)	82/83 (98,80%)
Acervo “Santo Rêgo”	19/27 (70,37%)	8/27 (29,63%)
Total	233/583 (39,97%)	350/583 (60,03%)

Com base na leitura da tabela 13, observamos a prevalência categórica do *você*, nas cartas amorosas dos acervos “Queiroz” e “Vieira do Vale”, cabendo o *tu* categórico nas cartas amorosas do “Acervo Moura Fé”, o que está de acordo com as análises de Silva (2012) que também evidenciou o contexto das relações mais íntimas (amorosas) como promissor ao *tu* íntimo.

Nas cartas de amizade, o *você* prevalece categoricamente nos acervos “autores piauienses” e “Possidônio Queiroz”, mostrando-se ainda em prevalência em relação ao *tu* nos acervos “Cartas pessoais variadas” e “Valmir Batista”, em 72,41% e 88,89% dos dados, respectivamente.

Nas cartas familiares, o *você* prevalece categoricamente no acervo “autores piauienses” e “cartas pessoais variadas”, mostrando-se em alternância com o *tu*, nos acervos “Queiroz” (98,80%) e “Santos Rêgo” (29,63%), ainda que neste acervo o *tu* tenha prevalecido, em 70,30% dos dados.

Em síntese, entendemos que resguardadas as devidas proporções relacionadas ao trabalho de SH, é preciso atentarmos ao fato de que as amostras de cartas piauienses estão em desequilíbrio em relação à distribuição da quantidade por acervo privado. Ainda assim, podemos constatar que as cartas amorosas também figuram como o contexto do *tu* íntimo conforme já verificado por Souza (2012), Silva (2012) para as cartas cariocas e por Souza (2021) para as cartas mineiras. Nas cartas de amizade e nas cartas familiares, o *você* predomina, ainda que em alternância (em alguns acervos) com o *tu*, o que está em consonância com a interpretação do *você* como uma estratégia *neutra* em relação ao íntimo e direto *tu* e ao alto nível de distanciamento entre os interlocutores que está na

origem da forma *você* (*vossa mercê*), cf. já discutido por Rumeu (2013), Lopes & Cavalcante (2011).

Como finalizamos a descrição-analítica dos resultados estatísticos, passamos aos resultados probabilísticos em relação à regra variável *tu* e *você* nas cartas novecentistas piauienses.

CAPÍTULO 4. A VARIAÇÃO TU/VOCÊ: DESCRIÇÃO-ANALÍTICA DOS RESULTADOS PROBABILÍSTICOS.

Passamos, neste capítulo, aos resultados probabilísticos em função dos índices de *peso relativo*, de *p-value* e de *logodd* para cada uma das variáveis selecionadas, nos níveis *up and down* (rodada multivariada), pelo RBrul em sua interface com o RStudio (OUSHIRO, 2014, p. 136), considerando a alternância *tu/você* como regra variável e a variante *você* como valor de aplicação (*application value*). Isso posto, reforçamos, pois, o fato de que todos os resultados probabilísticos expostos nesta seção estão principalmente vinculados à produtividade da forma *você* no âmbito do conjunto de cartas pessoais novecentistas piauienses.

Entendemos o *peso relativo* como a indicação da potência dos contextos linguísticos e extralinguísticos (fatores linguísticos e extralinguísticos) para a implementação de uma dada regra variável. Nesse sentido, assumimos, em conformidade com Guy & Zilles (2007), que pesos relativos acima de 0.05 evidenciam uma tendência à aplicação da forma variante assumida como valor de aplicação (*você*), ao passo que pesos abaixo de 0.05 expõem uma tendência à inaplicabilidade da regra variável em cena, tendo sempre em vista o valor de aplicação para o qual estejamos considerando na rodada dos dados, ou seja, na interação entre todas as variáveis.

O *p-value* é o valor resultante da correspondência entre a variável dependente (variável preditora) e o conjunto das variáveis independentes (contextos linguísticos e extralinguísticos). Assim sendo, assumimos como hipótese o fato de ser possível ou não a correspondência entre a variável dependente e as variáveis independentes (grupos de fatores). Concretamente, utilizamo-nos, para mensurar, através do *p-value*, a força de um grupo de fatores em relação à aplicação da regra variável, do índice 0,05 (5%). Nesse sentido, o entendimento é o de que abaixo desse valor ($< 0,05$) temos uma variável (grupo de fatores/variável independente) de expressivo condicionamento para a produtividade da variável dependente. Por outro lado, o *p-value* igual ou acima de 0,05 ($=$ ou $> 0,05$) coloca-nos frente a um contexto refreador da regra variável em análise.

Para os *logodds* a interpretação é a de que se trata de um índice direcionado à identificação da aplicação geral de cada um dos fatores no interior de uma variável independente (grupo de fatores) selecionada como relevante na rodada multivariada.

Nesse sentido, a concretude dos resultados probabilísticos a partir dos *logodds* apontam-nos aos índices positivos e negativos como indicação dos fatores que tendem a impulsionar e a refrear, respectivamente, a dinâmica variável em análise. Uma vez expostos os critérios desta análise probabilística no nível da rodada multivariada, passamos à descrição-analítica dos resultados.

É importante também esclarecer que para que esta rodada multivariada se efetivasse coerentemente em consonância, nos níveis *up* e *down* (*step-up and step-down match*), tivemos de excluir os seguintes grupos de fatores: o *subgênero da carta pessoal*, as *temáticas* (assuntos) *prevalentes nas cartas*, os *níveis de prática de escrita dos redatores*, os *acervos* e os *períodos* (tempo). Assim sendo, permanecem em análise as seguintes variáveis (grupos de fatores): a expressão do pronome-sujeito de 2SG (nulo ou pleno), a pessoa verbal (concordância), o paralelismo formal e semântico, o sexo, a faixa etária, as relações sociais e o nível de escolarização (formal) dos redatores. Tendo em vista esse cenário de variáveis independentes, tivemos uma bem sucedida rodada implementada nos níveis *up and down* (*step-up and step-down match*) que nos expõem as seguintes variáveis (contextos linguísticos e sociais), na ordem em que foram selecionadas, como propulsoras do *você*, tendo em vista os índices de *p-value* que, por sua vez, estão diretamente vinculados à força das variáveis independentes (grupos de fatores) para a aplicação da regra variável. São elas: 1ª) Pessoa verbal (3.14e-151); 2ª) Paralelismo formal e semântico (2.76e-08); 3ª) Expressão do sujeito (0.00154); 4ª) Faixa etária (0.000491) e 5ª) Escolarização (0.00839). Passamos, na sequência, à descrição e análise dos resultados probabilísticos na ordem de seleção exposta pelo RStudio. Assim sendo, passamos à 1ª variável selecionada, a pessoa verbal, na interação entre as variáveis linguísticas e extralinguísticas. Convém comentarmos o fato de que ao submetermos os dados ao RBrul em sua interface com o Rstudio para a geração dos pesos relativos, constatamos ser possível a seleção como relevante uma *variável independente*, ainda que tenhamos, no interior desse grupos de fatores, tão somente dados que referendam um único fator para o valor de aplicação em análise *você* (*single group* nos termos do GoldVarb). Isso é o que averiguamos com relação à seleção da variável *pessoa* que, na amostra de cartas piauienses, não apresenta ocorrências do *você* em concordância formal com a 2SG (*tu* nulo ou pleno), mas tão somente com formas verbais de 3SG.

Tabela 14: O efeito da variável *Pessoa* sobre o *você-sujeito*.

Pessoa (concordância)	Tokens (v+t)	Peso relativo	Logodds
3SG	361/583 (61,92%)	0.999	103.506
2SG	222/583 (38,08%)	0.001	-103.506

Regra variável: *tu/você*.
 Valor de aplicação: *você*.
p-value: 3.14e-151

A pessoa verbal (concordância), ao apresentar o *p-value* de 3.14e-151, foi a primeira variável linguística selecionada na interação entre as variáveis. Ao correlacionarmos os índices percentuais, probabilísticos e *logodds*, é possível interpretarmos a 3ª pessoa verbal (3SG) como uma variável fomentadora do *você* nas cartas piauienses em análise. O altíssimo peso relativo no valor de 0.999 (> 0.5) aliado ao também elevado *logodd* (103.506) representam as evidências concretas do contexto de 3ª pessoa verbal como propulsor do *você* na amostra de cartas pessoais em análise. Temos, por outro lado, em oposição à 3ª pessoa, a 2ª pessoa cujos índices percentual (38,08%), probabilístico (0.001) e *logodd* (-103.506) evidenciam tal contexto como dinamizador do *tu* nas cartas piauienses. Validamos, pois, a conjectura de que, na produção escrita piauiense, evidenciaríamos um panorama de conformidade formal entre o pronome *você-sujeito* e o verbo na 3SG, conforme já discutido por Souza (2012, p.99) e por Souza (2021, p. 114) para as cartas cariocas e mineiras, respectivamente.

Tabela 15: O efeito da variável *Paralelismo formal e semântico* sobre o *você-sujeito*.

Paralelismo formal e semântico	<i>Tokens</i> (<i>v+t</i>)	Peso relativo	<i>Logodds</i>
Precedido por construções imperativas de 2SG (ind.)	2/245 (1%)	0.999	114.499
Precedido por formas pronominais de 2PL (<i>clíticos, possessivos e pronome-sujeito</i>)	5/245 (2%)	0.999	113.636
Precedido por construções imperativas de 2SG (subj.) [3°]	12/245 (5%)	0.999	107.036
Precedido por <i>Você-sujeito</i> [1°]	191/245 (78%)	0.999	80.714
Precedido por <i>Vocativos</i> [2°]	25/245 (10%)	0.993	4.948
Precedido por <i>Tu não-Sujeito</i>	10/245 (4%)	0.985	4.199
(pronomes-complemento em funções dativa, oblíquo)			
Regra variável: <i>tu/você</i>			
Valor de aplicação: <i>você</i> .			
<i>p-value</i> : 2.76e-08			

Como a segunda variável linguística selecionada temos o *paralelismo formal e semântico*, o que está também consubstanciado no *p-value* de 2.76e-08, índice menor do que 0,05. No que diz respeito ao número de ocorrências, aos índices percentuais, aos probabilísticos e aos *logodds*, observamos que o *você-sujeito* é favorecido pelos contextos de precedência pelo *você-sujeito* (191/245, 78%, 0.999, 80.714), pelo vocativo (25/245, 10%, 0.993, 4.948), por construções imperativas com formas de subjuntivo (12/245, 5%, 0.999, 107.036) e pelo *tu não-sujeito* (10/245, 4%, 0.985, 4.199), respectivamente. Em outros termos, é possível entendermos que o *você* tende, na escrita piauiense novecentista, a se implementar com mais força, em contexto de precedência pelo *você-sujeito* e pelas formas subjuntivas das construções imperativas. Com altíssimos pesos relativos (idênticos, inclusive, 0.999) e com acentuados *logodds* (80.714 e 107.036), temos o *você-*

sujeito e as construções de *imperativo surrogate* como as que se mostram mais férteis ao *você*. Nesse sentido, temos em cena estruturas, formal e semanticamente, idênticas, uma vez que o *você-sujeito* e as construções imperativas com formas de subjuntivo (*imperativo surrogate*) subsidiam o *você*. Ainda que com poucas evidências em termos de dados brutos (10 ocorrências tão somente), identificamos os pronomes-complemento de 2SG, vinculados formalmente ao paradigma do *tu*, como um contexto propulsor do *você-sujeito* (0.985), o que consubstancia a convivência entre as formas pronominais dos paradigmas de *tu* e de *você* na atual fala do PB (¹³*Você_i* disse que eu *te_i* acharia na faculdade para pegar o *teu_i* livro). Ainda no âmbito da variável *paralelismo formal e semântico*, identificamos baixíssimas evidências de uso para o *você* se precedido por formas pronominais de 2PL (5/245, 2%, 0.999, 113.636) e por construções imperativas com formas indicativas para a referência à 2SG (2/245, 1%, 0.999, 114.499).

Tabela 16: O efeito da variável *Expressão do sujeito* sobre o *você-sujeito*.

Expressão do sujeito	Tokens (v+t)	Peso relativo	Logodds
Pleno	273/583 (46,83%)	0.999	55.105
Nulo	310/583 (53,17%)	0.001	-55.105

Regra variável: *tu/você*
 Valor de aplicação: *você*.
p-value: 0.00154

Como a terceira variável selecionada na interação em análise, temos a *expressão do sujeito* cujo valor de aplicação aponta para a sua relevância (0.00154), uma vez que é menor que 0,05, atuando como uma variável dinamizadora da forma *você*. Ao voltarmos especificamente aos fatores dessa variável, identificamos que o *você* tem a sua produtividade estimulada pelo contexto da sua expressão formal (sujeito pleno), o que está amparado nos altos índices de peso relativo (0.999) e de *logodd* (55.105). O fato de o *você* ter a sua produtividade amparada pelo contexto da sua expressão fonética (sujeito pleno) dialoga com a sua origem nominal, através da FNT *vossa mercê*, que genuinamente é marcada pela sua representação plena. Por outro lado, a genuína forma pronominal *tu* é

¹³ Dado linguístico retirado de Lopes (2007).

marcada pelo traço de pessoa, podendo, em essência, apresentar-se em sua expressão nula. Para a análise dessa variável linguística, a hipótese é a de que apesar de o *você* assumir os mesmos espaços funcionais que o *tu* (etimologicamente pronome-sujeito), resguarda ainda um resquício da forma nominal de tratamento originária (*vossa mercê*) cuja expressão é essencialmente plena (típico dos nomes), o que corrobora os estudos de Lopes *et alii* (2009), Rumeu (2013) para as missivas cariocas e Souza (2021) para as cartas mineiras.

Tabela 17: O efeito da variável *faixa etária* sobre o *você-sujeito*.

Faixa etária	Tokens (v+t)	Peso relativo	Logodds
Idoso	236/583 (40,48%)	0.999	63.161
Adulto	230/583 (39,45%)	0.996	5.504
Jovem	117/583 (20,07%)	0.001	-68.664

Regra variável: *tu/você*
 Valor de aplicação: *você*.
p-value: 0.00154

A quarta variável selecionada é a *faixa etária* dos redatores. À luz do *p-value* 0.00154 (menor do que 0,05) entendemos que a faixa etária dos redatores parece se mostrar como um contexto favorecedor da aplicação do *você*. A *faixa etária* do redator revela os redatores idosos (236/583, 40,48%, 0.999, 63.161) e adultos (230/583, 39,45%, 0.996, 5.504) como propulsores do *você* nas cartas piauienses analisadas, o que está de acordo com as constatações de Souza (2021, p. 117) em relação às cartas mineiras oitocentistas e novecentistas. Os altos pesos relativos 0.999 e 0.996 expõem-nos as cartas produzidas por escreventes idosos e adultos como contextos estimuladores do *você*, ainda que tenhamos clareza do fato de que as amostras linguísticas de sincronias passadas estão circunscritas aos textos que estão disponíveis no interior dos arquivos públicos e privados, conduzindo-nos a construção de amostras desequilibradas em relação à distribuição do número de cartas em função do gênero e da faixa etária dos seus redatores. Ainda que as missivas históricas produzidas por punhos jovens tenham se mostrado inibidoras do *você*, o que não faz muito sentido, uma vez que se trata da variante inovadora conduzida normalmente pelos informantes mais jovens atualmente no PB, entendemos que esse

resultado tende a obviamente refletir a constituição nada homogênea da amostra de cartas novecentistas piauienses.

Tabela 18: O efeito da variável *Escolarização* sobre o *você-sujeito*.

Níveis de Escolarização	<i>Tokens</i> (<i>v+t</i>)	Peso relativo	<i>Logodds</i>
Universitário	278/583 (47,68%)	0.999	16.556
Médio	119/583 (20,41%)	0.999	19.071
Alfabetizado	186/583 (31,90%)	0.001	-35.628

Regra variável: *tu/você*
 Valor de aplicação: *você*.
p-value: 0.00839

Ainda no âmbito das variáveis extralinguísticas, temos a *escolarização* como a quinta (e última) variável selecionada como relevante à produtividade do *você*, o que justifica o *p-value* de 0.00839 desse grupo de fatores, menor do que 0,05, na interação entre as variáveis. Dentre os níveis de escolarização dos redatores das missivas novecentistas piauienses, temos os que concluíram o ensino médio (119/583, 20,41%, 0.999, 19.071) e o ensino superior (278/583, 47,68%, 0.999, 16.556) como os que mais se mostram propícios à produtividade do *você* nas cartas em análise. Nos níveis universitário e médio, verificamos um altíssimo índice probabilístico (0.999), o que mostra tais níveis de escolarização como relevantes à proeminência do *você* na interação de dados em análise. Em termos de *logodds*, observamos que os redatores com o ensino médio tendem a se destacar, ainda que tenhamos um número de ocorrências menor (119 oco) se as comparamos aos dados dos redatores com o nível universitário (278 oco). Ainda no âmbito da análise dos níveis de escolarização dos redatores piauienses, observamos que a quantidade de dados produzidos pelos alfabetizados é maior (186 ocorrências), mas, na interação entre as variáveis, mostra-se como uma variável inibidora da produtividade da variável dependente *você*, o que está consubstanciado nos baixíssimos índices probabilísticos (0.001) e de *logodd* negativo (-35.628).

É importante reforçar o fato de que apesar de termos, em análises linguísticas voltadas às sincronias passadas, de conviver com amostras de dados desarmonizadas em relação ao número de cartas (fontes históricas) que, por sua vez, também revelam a identidade dos informantes em relação ao *gênero* e à *faixa etária* dos informantes, temos de também levar em conta que os resultados probabilísticos desta Tese vão ao encontro dos resultados de Souza (2021). Nessas duas análises linguísticas embasadas em cartas de sincronias passadas (SOUZA, 2021; BARROS, 2024), temos as variáveis *pessoa verbal*, *paralelismo* e *faixa etária* como propulsoras do *você* na produção escrita dos mineiros e dos piauienses em análise.

CONSIDERAÇÕES FINAIS

O foco principal desta Tese é a análise da dinâmica *tu/você*, tendo em vista não só os encaminhamentos metodológicos da SH para a formação das amostras históricas confiáveis, mas também, os índices estatísticos e probabilísticos relacionados à regra variável em questão. Voltamo-nos ao “presente para explicar o passado” (TRUDGILL & HERNÁNDEZ CAMPOY 2007, p. 299 *apud* BAGNO 2017, p. 430), visto que, na atual realidade da fala piauiense, é o *você*, em alternância com o *tu* (subsistema 5, cf. SCHERRE *et alii*, 2021), a forma pronominal que predomina para a referência ao sujeito de 2SG. Nesse sentido, temos clareza de que os índices estatísticos e probabilísticos evidenciam uma “fotografia” do PB escrito novecentista por punhos legitimamente piauienses. Isso posto, passemos à sistematização dos principais resultados, na tabela 19.

Tabela 19: Síntese das variáveis independentes estatisticamente relevantes à aplicação do *você-sujeito*

VARIÁVEIS LINGUÍSTICAS	VARIÁVEIS EXTRALINGUÍSTICAS
<p>1º) Pessoa (concordância)</p> <p>3ª pessoa do singular (3SG): 0.999 (61,92%, 361/583oco)</p> <p><i>p-value</i>: 3.14e-151</p>	<p>4º) Faixa etária</p> <p>Idoso: 0.999 (40,48%, 236/583)</p> <p>Adulto: 0.996 (39,45%, 230/583)</p> <p><i>p-value</i>: 0.00154</p>
<p>2º) Paralelismo formal e semântico</p> <p>Precedido por <i>Você-sujeito</i>: 0.999 (78%, 191/245oco).</p> <p>Precedido por <i>Vocativos</i>: 0.993 (10%, 25/245).</p> <p>Precedido por <i>construções imperativas</i> de 2SG com <i>formas de subjuntivo</i>: 0.999 (5%, 12/245).</p> <p><i>p-value</i>: 2.76e-08</p>	<p>5º) Escolarização</p> <p>Ensino superior: 0.999 (47,68%, 278/583).</p> <p>Ensino médio: 0.999 (20,41%, 119/583).</p> <p><i>p-value</i>: 0.00839</p>

3º) Expressão do sujeito	
Precedido por formas de <i>Sujeito Pleno</i> : 0.999 (46,83%, 273/583)	-
<i>p-value</i> : 0.00154	
Regra variável: <i>tu/você</i>	
Valor de aplicação: <i>você</i> .	

Considerando que o foco desta análise é a alternância *tu/você*, buscamos descrever analiticamente os contextos linguísticos e sociais que correlacionamos à expressão da regra variável em análise. Nesse sentido, conduzimos este estudo por três questões de pesquisa, alcançando as seguintes generalizações sobre a variação *tu/você* nas cartas piauienses novecentistas.

(1ª) As formas *tu* e *você* também se mostrariam proficuas na escrita novecentista dos piauienses, considerando o fato de que o *você* prevalece na fala atual do Piauí, ainda que em alternância com o *tu* (com concordância média entre 10% e 39% = subsistema 5), cf. Scherre *et alii* (2021, p. 165)?

Confirmamos a conjectura de que o *você* prevaleceria em relação ao *tu* como já observado para a fala atual do Piauí (subsistema 5 por Scherre *et alii*, 2021). Apesar de a alternância *tu* (39,97%) e *você* (60,03%) ter se mostrado intensa, evidenciamos a preferência dos piauienses pelo *você*, anunciando, na produção escrita de sincronias passadas, aquilo que atualmente se revela na fala piauiense, cf. Scherre *et alii* (2021, p. 165) à luz das análises embasadas no ALiB (DEUS, 2009; CARDOSO, 2008).

(2ª) Quais contextos linguísticos e extralinguísticos promoveriam a produtividade da forma *você* na produção novecentista piauiense?

Nas cartas novecentistas analisadas, os contextos de *pessoa verbal*, *paralelismo formal e semântico*, *expressão do sujeito*, *faixa etária* e *nível de escolarização* mostraram-se como propulsores do *você*. Passamos às seguintes sistematizações para cada um desses parâmetros selecionados pelo RBrul/Rstudio.

(1) A articulação morfossintática do *você* às formas verbais de 3SG configura um contexto promissor da sua produtividade (0.999), evidenciando um resultado compatível com os observados por Souza (2012) e por Souza (2021) em relação às amostras carioca e mineira, respectivamente, do PB escrito em sincronias passadas (séculos XIX e XX);

(2) Em relação ao *paralelismo formal e semântico* constatamos que o fato de o *você* estar precedido por também por formas de *você-sujeito* (0.999) é altamente favorável à sua produtividade, acompanhando os resultados de Souza (2021) para as cartas mineiras oitocentistas e novecentistas. Apesar de estarmos em contexto de variação para a análise de uma dada regra variável, temos em evidência os contextos formalmente idênticos como os que tendem a conduzir a produtividade da variável dependente, o que também foi observado por Rumeu (2013) para o *tu-sujeito* das cartas cariocas, por Souza (2012) e Souza (2021) para o *você* das cartas cariocas e mineiras, respectivamente;

(3) Dentre as possibilidades de expressão do sujeito (nulo ou pleno), verificamos que o *você* tende a ser favorecido pelo *você* como sujeito pleno. Isso confirma tal traço formal como uma evidência que reverbera a FNT que o originou (*vossa mercê*) cuja tendência nominal é a de expressão plena, o que já está consubstanciado nos resultados de análises embasadas em distintos e diversificados *corpora* históricos (SOUZA, 2021; RUMEU, 2013; LOPES *et alii*, 2009);

(4) Passemos às sistematizações voltadas aos parâmetros extralinguísticos selecionados como relevantes à proficiência da regra variável em análise. São eles: a *faixa etária* e o *nível de escolarização* dos redatores. Ainda que tenhamos de ser cuidadosos em relação às amostras desarmonizadas, em algum nível, em relação aos parâmetros ¹⁴*sexo, faixa etária e nível de escolarização* dos redatores, observamos, especificamente em relação à *faixa etária*, a congruência dos resultados das análises embasadas em cartas mineiras (SOUZA, 2021) e piauienses. Isso significa que, em ambos os trabalhos, evidenciamos os redatores idosos (0.999) e adultos (0.999) como os propulsores do *você*, tendo sempre em vista os respectivos conjuntos de cartas pessoais analisadas. No que se refere ao *grau de escolarização* dos redatores piauienses, identificamos os níveis *universitário* (0.999) e *médio* (0.999) como mais propensos ao uso do *você*.

¹⁴ Nesta análise, temos treze (13) redatores, sendo nove (10) homens e três (3) mulheres. Todos nascidos no espaço geográfico do Piauí.

(3^a) Quais seriam os tipos de relações sociais que as subjazem na realidade novecentista dos piauienses?

Confirmamos a hipótese de que o *você* se espraiaria com mais força pelas relações simétricas (51,97%), apesar de ter-se sobressaído, com muito mais vigor, pelas assimétricas descendentes (de superior para inferior), em 99% dos dados.

De um modo geral, os resultados desta Tese acerca da dinâmica *tu/você* apontam para a prevalência do *você*, mesmo que em convivência com o *tu*. Ainda em relação ao *você*, constatamos que a sua produtividade é favorecida nos contextos da *terceira pessoa verbal*, do *você-sujeito* e da produção escrita de *redatores idosos e adultos*, tendo em vista os conjuntos de acervos documentais em análise, produzidos por punhos genuinamente piauienses. No âmbito das relações sociais, temos em cena um panorama que parece apontar para a semântica da Solidariedade, marcada pelo *você*, não só nas relações de simetria, mas também nas dinâmicas de assimetria social (descendente) que contextualizam a realidade linguística do PB novecentista.

REFERÊNCIAS BIBLIOGRÁFICAS

- AVELHEDA BANDEIRA, A. C. C.; SOUZA, S. C. G. *Passo a passo para uso do R-Brul*. 2017 (Curso de curta duração ministrado/Extensão).
- BAGNO, M. *Dicionário crítico de sociolinguística*. Parábola Editorial. Edição do Kindle. 2017.
- BAKHTIN, M. M; VOLOSHINOV, V. N. *Marxism and the philosophy of language*. New York: Academic, 1973.
- BARBOSA, G. M. O. [O uso dos pronomes possessivos teu e seu em cartas pessoais de sertanejos baianos do século XX](#). Dissertação (Mestrado em Estudos Linguísticos), Universidade Estadual de Feira de Santana, Feira de Santana, Bahia, 2018.
- BARBOSA, A. G. [Para uma História do Português Colonial: Aspectos Linguísticos em Cartas de Comércio](#). Tese (Doutorado em Língua Portuguesa), Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1999.
- BRANDÃO, T. M. P. *A elite colonial piauiense: família e poder*. Teresina-PI: Fundação Monsenhor Chaves, 1995.
- BATISTA, P. E. S. T. [O uso de Tu e Você em cartas baianas pessoais no século XX em relações de simetria](#). Dissertação (Mestrado em Estudos Linguísticos), Universidade Estadual de Feira de Santana, Bahia, 2015.
- BAPTISTA, M. P. C. [Amarras e desamarras: cotidiano e modernização em amarração no litoral do Piauí \(1880 - 1930\)](#), Dissertação (Mestrado em História do Brasil), Universidade Federal do Piauí, Teresina, 2019.
- BECHARA, E. *Moderna gramática portuguesa*. São Paulo: Companhia Editora Nacional, 1994.
- BROWN, L.; LEVINSON, S. *Politeness: some universal in language usage*. Cambridge: Cambridge University, 1987.
- BROWN, R.; GILMAN, A. The pronouns of Power and Solidarity. In: SEBEOK, T.A. (ed.) 1960. *Style in language*. Cambridge- Mass: MIT Press, 2003, p.253-276.
- CARDOSO, S. A. M. *et alii*. *O tratamento do interlocutor nas capitais: tu ou você*. No prelo.
- CARDOSO, S. A. M. *Caminhos dos pronomes pessoais no português brasileiro: considerações a partir dos dados do projeto ALiB*. In: SIMPÓSIO MUNDIAL DE ESTUDOS DE LÍNGUA PORTUGUESA, 2008, São Paulo.

CUNHA, C.; CINTRA, L. F. L. *Nova gramática do Português contemporâneo*. Rio de Janeiro: Nova Fronteira, 2011 [1985].

DEUS, V. G. de. [*Você ou tu? Nordeste versus Sul: o tratamento do interlocutor no português do Brasil a partir dos dados do Projeto ALiB*](#). Dissertação (Mestrado em Linguística) – Universidade Federal da Bahia, Salvador, 2009.

CONDE SILVESTRE, J. C. *Sociolinguística Histórica*. Madrid: Gredos, 2007.

COUTINHO, I. L. *Pontos de Gramática Histórica*. Rio de Janeiro: Ao Livro Técnico, 1976.

CUNHA, C.; CINTRA, L. F. L. *Nova gramática do Português contemporâneo*. Rio de Janeiro: Nova Fronteira, 1985.

DUARTE, M. E. L. [*A perda do princípio 'Evite pronome' no português brasileiro*](#). Tese (Doutorado em Linguística), Universidade Estadual de Campinas, Campinas, 1995.

DUARTE, M. E. L. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In: I. Roberts; M. Kato. (orgs.) *Português brasileiro: uma viagem diacrônica*. São Paulo: Campinas, Editora da Unicamp: 1993, p. 107-128.

DUARTE, M. E. L. Clítico acusativo, pronome lexical e categoria vazia no português do Brasil. In: Tarallo, F. (org.) *Fotografias Sociolinguísticas*. Campinas: Pontes/Ed. da Unicamp. 19-34. 1989.

FARACO, C. A. [O tratamento “você” em português: uma abordagem histórica](#). *LaborHistórico*, 2017 [1996]: 3(2), 114-132.

FARACO, C. A. *Linguística Histórica: uma introdução ao estudo da história das línguas*. São Paulo: Parábola, 2005.

FONTANELLA, M. B. La constitución del paradigma pronominal del voseo. *Thesaurus* XXXII, nº 2. 1977.

FREITAG, R. M. Ko. (Org.). [*Metodologia de Coleta e Manipulação de Dados em Sociolinguística*](#). São Paulo: Editora Edgard Blücher, 2014.

GONÇALVES, S. C; CARVALHO, C. S. Critérios de gramaticalização. In: GONÇALVES, S. C. L.; LIMA-HERNANDEZ, M. C.; CASSEB-GALVÃO, V. C. (Org.). *Introdução à gramaticalização: princípios teóricos e aplicação*. São Paulo: Parábola, 2007.

GOMES, V. S.; LOPES, C.R.S. [Formas tratamentais em cartas escritas em Pernambuco \(1869-1969\): tradição discursiva e sociopragmática](#). *Revista de Estudos da Linguagem*, v. 24, p. 137-165, 2016.

GOMES, V. S. [Para além dos pacotes estatísticos Varbrul, Goldvarb e RBrul: qual a concepção de gramática](#). *Revista Gelne*, v. 14 Especial, 2012.

GUY, G; ZILLES, A. *Sociolinguística Quantitativa – instrumental de análise*. São Paulo: Parábola Editorial, 2007.

HAROCHE-BOUZINAC, G. *Escritas Epistolares*. São Paulo: Editora da Universidade de São Paulo, 2016.

HERNÁNDEZ-CAMPOY, J. M.; SCHILLING, N. The Application of the Quantitative Paradigm to Historical Sociolinguistics: Problems with the Generalizability Principle. In: HERNÁNDEZ CAMPOY, J. M.; CONDE-SILVESTRE, J. C. *The Handbook of Historical Sociolinguistics*. Wiley-Blackwell, 2012, p. 63-79.

HOPPER, P. J. On some principles of grammaticalization. TRAUGOTT, E. C. e HEINE, B. (eds.). *Approaches to grammaticalization*. Volume I, Amsterdam/Philadelphia: John Benjamins Company, p. 17- 35. 1991.

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

KATO, M. A. A gramática do letrado: questões para a teoria gramatical. In: M. A. Marques, E. Koller; J. Teixeira & A. S. Lemos (orgs.). *Ciências da Linguagem: 30 anos de investigação e ensino*. Braga: CEHUM (Universidade do Minho), p. 131-145, 2005.

LABOV, W. *Padrões sociolinguísticos*. Tradução de BAGNO, M.; SCHERRE, M. M. P.; CARDOSO, C. R. São Paulo: Parábola Editorial, 2008 [1972].

LABOV, W. *Principles of Linguistic Change: Internal Factors*. Cambridge: Blackwell Publishers, 1994, v. I.

LOPES, C. R. S.; MARCOTULIO, L. L.; OLIVEIRA, T. L. [A atuação dos papéis sociais na mudança no sistema de tratamento no português brasileiro: análise de cartas pessoais \(1870-1979\)](#). *ESTUDOS DE LINGUISTICA GALEGA*, v. 1, p. 29-44, 2018.

LOPES, C. R. S.; MARCOTULIO, L. L.; RUMEU, M. C. B.; COELHO, I. L.; ANDRADE, A.; MARTINS, M. A.; LACERDA, M. O.; GOMES, V. S.; MONTE, V. M.; CARNEIRO, Z. O. N.; SOUZA, C. M. N.; BALSALOBRE, S.; SOUZA, J. P.; OLIVEIRA, T. L.; MOURA, K. K.; CRUZ, I. A.; CARDOSO, N. D. A reorganização do sistema pronominal de 2ª pessoa na história do português brasileiro: a posição de sujeito. In: CASTILHO, A. T. (Org.). *História do português brasileiro: mudança sintática das classes de palavra – perspectiva funcionalista*. São Paulo, Contexto, 2018, v. 4, p. 24-140.

LOPES, C. R. S.; RUMEU, M. C. B. [A identificação dos perfis socioculturais dos redatores de corpora históricos: encaminhamentos metodológicos](#). *Diadorim* (Rio de Janeiro), v. 20, p. 147-168, 2018.

LOPES, C. R. S.; RUMEU, M. C. B. [A difusão do você pelas estruturas sociais carioca e mineira dos séculos XIX e XX](#). *LaborHistórico*, v. 1, 2015, p. 12-25.

LOPES, C. R. S.; VIANNA, J. B. S. [A competição entre nós e a gente nas funções de complemento e adjunto: desvendando outras portas de entrada para o pronome inovador.](#) *Caligrama: Revista de Estudos Românicos*, v. 17, n. 2, 2012, p. 137-161.

LOPES, C. R. S.; CAVALCANTE, S. R. [A cronologia do voceamento no português brasileiro: expansão de você-sujeito e retenção do clítico-te.](#) In: *Linguística*, 2011, v.25, p. 30-65.

LOPES, C. R. S. *et alii*. Sobre Norma e Tratamento em cartas a Rui Barbosa. In: AGUILERA, V. A. (Org.). *Para a História do Português Brasileiro*. Londrina: Eduel, 2009. Volume VII, p. 45-92, 2009.

LOPES, C. R. S.; RUMEU, M. C. B. [O quadro de pronomes pessoais do português: as mudanças na especificação dos traços intrínsecos.](#) In: CASTILHO; A. T.; TORRES MORAIS, M. A.; LOPES, R. E. V.; CYRINO, S. M. L. (Orgs.) *Descrição, história e aquisição do português brasileiro*. 1ª ed. São Paulo/Campinas: FAPESP/Pontes Editores, 2007, v.1, p. 419-436.

LOPES, C. R. S.; MACHADO, A. C. M. Tradição e inovação: indícios do sincretismo entre a segunda e a terceira pessoas nas cartas dos avós. In: LOPES, C. R. S. (Org.). *A norma brasileira em construção: fatos linguísticos do século 19*. Rio de Janeiro: UFRJ; FAPERJ, 2005. p. 45-66.

LUCENA, R. O. P. [Pronomes possessivos de segunda pessoa: a variação teu/seu em uma perspectiva histórica.](#) Tese (Doutorado em Letras Vernáculas), Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2016.

MACHADO, A. C. M. *As formas de Tratamento nos teatros brasileiro e português dos séculos XIX e XX*. Tese (Doutorado em Letras Vernáculas). Rio de Janeiro, Faculdade de Letras/ UFRJ, 2011.

MACHADO, A. C. M. [As formas de tratamento no Teatro do Rio de Janeiro dos séculos XIX e XX.](#) *Cadernos de Estudos Linguísticos*, Campinas, SP, v. 60, n. 3, p. 647-668, 2018.

MAINGUENEAU, D. *Análise de textos de comunicação*. Trad. Cecília P. de Sousa e Silva, Décio Rocha. – São Paulo: Cortez, 2001.

MARQUILHAS, R. *A Faculdade das Letras: Leitura e escrita em Portugal no Séc. XVII*. Lisboa: Imprensa Nacional-Casa da Moeda, 2000.

MENON, O. [O Sistema Pronominal do Português do Brasil.](#) *Revista Letras*, Curitiba: Editora da UFPR, n. 44, 1995, p. 91-106.

MOURA, J. M. *Ney Moura Fé: Seus caminhos e sua obra*. Teresina: Halley, 2015.

MOURA, K. K. & MARTINS, M. A. (2018) Formas de tratamento em cartas do Rio Grande do Norte. In.: LOPES, *et alii*. A reorganização do sistema pronominal de 2ª pessoa

na história do português brasileiro: a posição de sujeito. In: CASTILHO, A. T. (Org.). *História do português brasileiro: mudança sintática das classes de palavra – perspectiva funcionalista*. São Paulo, Contexto, 2018, v. 4.

NASCIMENTO, F. A. *Cartografias sentimentais e narrativas históricas sobre Teresina e Oeiras*. Teresina: EdUESPI, 2021.

OMENA, N. P. A Referência à Primeira Pessoa no Plural: Variação ou Mudança? In.: PAIVA, M. C.; DUARTE, M. E. L. (2003). (Orgs.) *Mudança linguística em tempo real*. Rio de Janeiro: Contra Capa Livraria, 2003. p. 63-80.

OUSHIRO, L. [Tratamento de dados com o R para análises sociolinguísticas](#). In: FREITAG, R. M. K. (Org.). *Metodologia de Coleta e Manipulação de Dados em Sociolinguística*. São Paulo: Editora Edgard Blücher, 2014. p.133-176.

OUSHIRO, L. *Introdução à estatística para linguistas*. v1.01, dez. UNICAMP, 2017.

ROCHA LIMA, C. H. *Gramática normativa da língua portuguesa*. Rio de Janeiro: José Olympio, 2001 [1972].

R CRAN. R: [The Comprehensive R Archive Network](#).

RStudio. [Take control of your R code](#).

RUMEU, M. C. B.; CRUZ, I. A.; CARDOSO, N. D. Formas de tratamento em cartas de Minas Gerais. In.: LOPES, C. R. S. *et alii*. A reorganização do sistema pronominal de 2ª pessoa na história do português brasileiro: a posição de sujeito. In: CASTILHO, A. T. (Org.). *História do português brasileiro: mudança sintática das classes de palavra – perspectiva funcionalista*. São Paulo, Contexto, 2018, v. 4, p. 42-53.

RUMEU, M. C. B. [As relações de poder e de solidariedade na sociedade carioca dos séculos XVIII e XIX](#). São Paulo: *Todas as Letras R*, 2011, v. 13, n. 2, p. 115-126.

RUMEU, M. C. B. *Língua e sociedade: a história do pronome “Você” no português brasileiro*. Rio de Janeiro: Ítaca (FAPERJ), 2013.

RUMEU, M. C. B. [Traços formais e semântico-discursivos no processo de gramaticalização de ‘Vossa Mercê’ > ‘Você’](#). *Revista do GEL* (Araraquara), v.3, 2006, p. 67-82.

RUMEU, M. C. B. [Para uma história do português no Brasil: formas pronominais e nominais de tratamento em cartas setecentistas e oitocentistas](#). Dissertação (Mestrado em Letras Vernáculas), Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2004.

SANTOS LUZ, M. Fórmulas de tratamento no português arcaico. *Revista portuguesa de filologia*. v. 7, 1957, p. 251-363.

SCHERRE, M. M. P.; DIAS, E. P.; ANDRADE, C. Q.; MARTINS, G. F. Variação dos pronomes “tu” e “você”. In: MARTINS, M. A.; ABRAÇADO, J (Org). *Mapeamento sociolinguístico do português brasileiro*. São Paulo, Contexto, 2015, p. 133-172.

SCHERRE, M. M. P. [Paralelismo linguístico](#). *Revista de Estudos da Linguagem*. Belo Horizonte, V. 7, n.2, p. 29-59, 1998.

SILVA, E. N. [Cartas amorosas de 1930: o tratamento e o perfil sociolinguístico de um casal não-ilustre](#). Dissertação (Mestrado em Letras Vernáculas) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2012.

SOTO, E. U. M. S. *Cartas através do tempo: o lugar do outro na correspondência brasileira*. Niterói: Ed. da UFF, 2007.

SOUZA, E. Q. [As formas de referência ao sujeito de 2ª pessoa do singular em missivas mineiras dos séculos XIX e XX: uma análise linguístico-social](#). Dissertação (Mestrado em Estudos Linguísticos) – Universidade Federal de Minas Gerais, Belo Horizonte, 2021.

SOUZA, J. P. F. [Mapeando a entrada do você no quadro pronominal: análise de cartas familiares dos séculos XIX-XX](#). Dissertação (Mestrado em Letras Vernáculas) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2012.

SCHERRE, M. M. P.; ANDRADE, C. Q.; CATÃO, R. C. [Por onde transitam o tu e o você no Nordeste?](#) *Revista de Letras*, v. 40, p. 164-197, 2021.

SCHERRE, M. M. P.; DIAS, E. P.; ANDRADE, C.; MARTINS, G. F. Variação dos pronomes “tu” e “você”. In: MARTINS, M. A.; ABRAÇADO, J. *Mapeamento sociolinguístico do português brasileiro*. São Paulo: Contexto, 2015 [2009]. p.133-172.

SOUZA, I. L. *E-CORP – uma ferramenta de busca de dados para fins linguísticos: aplicação em base de dados de corpus eletrônico*. 2016. Monografia. Departamento de Exatas. UEFS, Feira de Santana.

SOUZA, I. L.; PINTO, G. R. P. R., CARNEIRO, Z. O. N., FARIA, P. P. F.; LACERDA, M. F. O. (2019). [A ferramenta de busca E-CORP aplicada ao Corpus Eletrônico de Documentos Históricos do Sertão](#). *A Cor Das Letras*, m19(2), 8-21.

SPINA, S. *Introdução à edótica: crítica textual*. 2ª ed. São Paulo: Cultrix/Edusp. 1977.

TRUDGILL, P.; HERNÁNDEZ CAMPOY, J. M. *Diccionario de sociolingüística*. Madri: Gredos. 2007.

BAGNO, M. *Dicionário crítico de sociolingüística*. Parábola Editorial. Edição do Kindle, 2017. p. 509.

VITRAL, L. T.; RAMOS, J. M. [Réplica a Peterson \(2008\): a tripartição pronominal das proformas cé, ocê e você](#). *DELTA*. 24.2: 283-308, 2008. São Paulo: PUC-SP.

WARDHAUGH, R. *An introduction to sociolinguistics*. Oxford: Blackwell Publishers Ltd, 2006.

WEINREICH, U.; LABOV, W.; HERZOG, M. I. *Fundamentos empíricos para uma teoria da mudança linguística*. Tradução Marco Bagno. São Paulo: Parábola, 2006 (1975).

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

APÊNDICE(Rodada nos níveis *stepping up* e *stepping down*)

STEPPING UP...

STEP 0 - Run 0 - model with no fixed predictors

```
$misc.1
  n df intercept overall proportion
583 1      0.407                0.6
```

```
$misc.2
log.likelihood      AIC      AICc Dxy R2
      -392.285 786.569 786.576  0  0
```

STEP 1 - adding to Run 0 model with no fixed predictors

Trying with ExprSuj...

```
$ExprSuj
factor logodds tokens v/v+t centered factor weight
      p  1.812     273 0.941                0.86
      n -1.812     310 0.300                0.14
```

```
$misc.1
  n df intercept overall proportion centered input prob
583 2      0.965                0.6                0.724
```

```
$misc.2
log.likelihood      AIC      AICc  Dxy  R2
      -250.28 504.56 504.58 0.666 0.499
```

Run 1 (above) with ExprSuj is better than Run 0 without ExprSuj, $p = 1e-63$

Trying with PVerb...

```
$PVerb
factor logodds tokens v/v+t centered factor weight
      3 12.013     361 0.97                > 0.999
      2 -12.013     222 0.00                < 0.001
```

```
$misc.1
  n df intercept overall proportion centered input prob
583 2     -8.553                0.6                0
```

```
$misc.2
log.likelihood      AIC      AICc  Dxy  R2
      -49.231 102.463 102.484 0.953 0.976
```

Run 2 (above) with PVerb is better than Run 0 without PVerb, $p = 3.14e-151$

Trying with Paral...

\$Paral

factor	logodds	tokens	v/v+t	centered	factor	weight
2	6.259	191	0.969			0.998
4	6.079	107	0.963			0.998
3	4.910	27	0.889			0.993
1	4.130	14	0.786			0.984
t	2.751	25	0.480			0.94
y	2.494	12	0.417			0.924
6	1.444	10	0.200			0.809
7	-0.278	117	0.043			0.431
5	-0.319	73	0.041			0.421
9	-13.735	5	0.000			< 0.001
z	-13.735	2	0.000			< 0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	11	-2.831			0.6		0.056

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-124.062	270.124	270.587	0.917	0.788

Run 3 (above) with Paral is better than Run 0 without Paral, $p = 7.12e-109$

Trying with Gnr...

\$Gnr

factor	logodds	tokens	v/v+t	centered	factor	weight
h	1.31	512	0.666			0.788
m	-1.31	71	0.127			0.212

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	2	-0.62			0.6		0.35

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-353.119	710.237	710.258	0.24	0.183

Run 4 (above) with Gnr is better than Run 0 without Gnr, $p = 8.71e-19$

Trying with FxEt...

\$FxEt

factor	logodds	tokens	v/v+t	centered	factor	weight
v	4.067	236	0.996			0.983
j	-1.827	117	0.393			0.139
k	-2.240	230	0.300			0.096

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	3	1.393			0.6		0.801

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-225.367	456.734	456.776	0.698	0.737

Run 5 (above) with FxEt is better than Run 0 without FxEt, $p = 3.23e-73$

Trying with RelSoc...

\$RelSoc

factor	logodds	tokens	v/v+t	centered	factor	weight
d	2.258	100	0.99			0.905
s	-2.258	483	0.52			0.095

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	2	2.337			0.6		0.912

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-340.016	684.033	684.054	0.279	0.469

Run 6 (above) with RelSoc is better than Run 0 without RelSoc, $p = 1.54e-24$

Trying with Escolar...

\$Escolar

factor	logodds	tokens	v/v+t	centered	factor	weight
m	0.502	119	0.765			0.623
k	0.495	186	0.763			0.621
u	-0.996	278	0.421			0.27

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	3	0.677			0.6		0.663

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-355.881	717.762	717.804	0.357	0.145

Run 7 (above) with Escolar is better than Run 0 without Escolar, $p = 1.55e-16$

Adding PVerb...

STEP 2 - adding to Run 2 model with PVerb

Trying with ExprSuj...

\$ExprSuj

factor	logodds	tokens	v/v+t	centered	factor	weight
p	0.417	273	0.941			0.603
n	-0.417	310	0.300			0.397

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor	weight
3	11.77	361	0.97			> 0.999
2	-11.77	222	0.00			< 0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
---	----	-----------	---------	------------	----------	-------	------

583 3 -8.43 0.6 0

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-48.361	102.723	102.764	0.962	0.977

Run 8 (above) with PVerb + ExprSuj is better than Run 2 without ExprSuj, $p = 0.187$

Trying with Paral...

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor	weight
3	21.658	361	0.97			> 0.999
2	-21.658	222	0.00			< 0.001

\$Paral

factor	logodds	tokens	v/v+t	centered	factor	weight
2	8.796	191	0.969			> 0.999
1	8.000	14	0.786			> 0.999
t	7.389	25	0.480			0.999
y	7.273	12	0.417			0.999
6	6.812	10	0.200			0.999
9	5.516	5	0.000			0.996
z	5.516	2	0.000			0.996
4	-9.599	107	0.963			< 0.001
3	-11.055	27	0.889			< 0.001
5	-14.233	73	0.041			< 0.001
7	-14.416	117	0.043			< 0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	12	-7.425			0.6		0.001

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-21.576	67.152	67.7	0.997	0.996

Run 9 (above) with PVerb + Paral is better than Run 2 without Paral, $p = 2.76e-08$

Trying with Gnr...

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor	weight
3	12.68	361	0.97			> 0.999
2	-12.68	222	0.00			< 0.001

\$Gnr

factor	logodds	tokens	v/v+t	centered	factor	weight
h	1.908	512	0.666			0.871
m	-1.908	71	0.127			0.129

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	3	-10.366			0.6		0

```
$misc.2
log.likelihood    AIC    AICc    Dxy    R2
      -36.244  78.488  78.529  0.977  0.98
```

Run 10 (above) with PVerb + Gnr is better than Run 2 without Gnr, $p = 3.46e-07$

Trying with FxEt...

```
$PVerb
factor logodds tokens v/v+t centered factor weight
      3  19.201    361  0.97                > 0.999
      2 -19.201    222  0.00                < 0.001
```

```
$FxEt
factor logodds tokens v/v+t centered factor weight
      v  12.759    236  0.996                > 0.999
      k  -5.425    230  0.300                0.004
      j  -7.334    117  0.393                0.001
```

```
$misc.1
n df intercept overall proportion centered input prob
583 4  -10.236                0.6                0
```

```
$misc.2
log.likelihood    AIC    AICc    Dxy    R2
      -33.621  75.243  75.312  0.991  0.995
```

Run 11 (above) with PVerb + FxEt is better than Run 2 without FxEt, $p = 1.66e-07$

Trying with RelSoc...

```
$PVerb
factor logodds tokens v/v+t centered factor weight
      3  18.968    361  0.97                > 0.999
      2 -18.968    222  0.00                < 0.001
```

```
$RelSoc
factor logodds tokens v/v+t centered factor weight
      d  9.025    100  0.99                > 0.999
      s -9.025    483  0.52                < 0.001
```

```
$misc.1
n df intercept overall proportion centered input prob
583 3  -6.816                0.6                0.001
```

```
$misc.2
log.likelihood    AIC    AICc    Dxy    R2
      -45.641  97.281  97.323  0.966  0.993
```

Run 12 (above) with PVerb + RelSoc is better than Run 2 without RelSoc, $p = 0.00737$

Trying with Escolar...

```
$PVerb
factor logodds tokens v/v+t centered factor weight
```


3	12.226	361	0.97	> 0.999
2	-12.226	222	0.00	< 0.001

\$Escolar

factor	logodds	tokens	v/v+t	centered	factor	weight
u	0.556	278	0.421			0.635
k	-0.167	186	0.763			0.458
m	-0.389	119	0.765			0.404

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	4	-8.713			0.6		0

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-48.559	105.118	105.187	0.961	0.977

Run 13 (above) with PVerb + Escolar is better than Run 2 without Escolar, $p = 0.51$

Adding Paral...

STEP 3 - adding to Run 9 model with PVerb + Paral

Trying with ExprSuj...

\$ExprSuj

factor	logodds	tokens	v/v+t	centered	factor	weight
p	1.908	273	0.941			0.871
n	-1.908	310	0.300			0.129

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor	weight
3	21.9	361	0.97			> 0.999
2	-21.9	222	0.00			< 0.001

\$Paral

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
2	9.392	191	0.969			> 0.999	>2.5
9	7.916	5	0.000			> 0.999	>2.5
z	7.916	2	0.000			> 0.999	>2.5
1	7.803	14	0.786			> 0.999	>2.5
y	7.628	12	0.417			> 0.999	>2.5
t	7.102	25	0.480			0.999	>2.5
6	6.101	10	0.200			0.998	>2.5
4	-10.446	107	0.963			< 0.001	>2.5
3	-12.297	27	0.889			< 0.001	>2.5
5	-15.177	73	0.041			< 0.001	>2.5
7	-15.936	117	0.043			< 0.001	>2.5

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	13	-7.673			0.6		0

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-16.558	59.117	59.757	0.999	0.996

Run 14 (above) with PVerb + Paral + ExprSuj is better than Run 9 without ExprSuj, $p = 0.00154$

Trying with Gnr...

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor	weight
3	21.662	361	0.97			> 0.999
2	-21.662	222	0.00			< 0.001

\$Paral

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
2	8.561	191	0.969			> 0.999	>2.5
1	8.220	14	0.786			> 0.999	>2.5
t	7.478	25	0.480			0.999	>2.5
y	7.017	12	0.417			0.999	>2.5
6	6.633	10	0.200			0.999	>2.5
z	5.846	2	0.000			0.997	>2.5
9	5.190	5	0.000			0.994	>2.5
4	-9.774	107	0.963			< 0.001	>2.5
3	-11.390	27	0.889			< 0.001	>2.5
7	-13.491	117	0.043			< 0.001	>2.5
5	-14.289	73	0.041			< 0.001	>2.5

\$Gnr

factor	logodds	tokens	v/v+t	centered	factor	weight
h	0.991	512	0.666			0.729
m	-0.991	71	0.127			0.271

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	13	-8.085			0.6		0

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-19.98	65.96	66.6	0.999	0.996

Run 15 (above) with PVerb + Paral + Gnr is better than Run 9 without Gnr, $p = 0.074$

Trying with FxEt...

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
3	26.518	361	0.97			> 0.999	>2.5
2	-26.518	222	0.00			< 0.001	>2.5

\$Paral

factor	logodds	tokens	v/v+t	centered	factor	weight
z	13.077	2	0.000			> 0.999
9	12.585	5	0.000			> 0.999
6	11.593	10	0.200			> 0.999
t	8.649	25	0.480			> 0.999
1	6.441	14	0.786			0.998
2	5.965	191	0.969			0.997
y	0.685	12	0.417			0.665
4	-12.847	107	0.963			< 0.001
3	-14.225	27	0.889			< 0.001

5	-15.795	73	0.041	< 0.001
7	-16.128	117	0.043	< 0.001

\$FxEt

factor	logodds	tokens	v/v+t	centered	factor weight
v	12.366	236	0.996		> 0.999
k	-5.538	230	0.300		0.004
j	-6.828	117	0.393		0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	14	-4.095			0.6		0.016

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-18.4	64.8	65.54	0.999	0.998

Run 16 (above) with PVerb + Paral + FxEt is better than Run 9 without FxEt, $p = 0.0418$

Trying with RelSoc...

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor weight	vif
3	26.276	361	0.97		> 0.999	>2.5
2	-26.276	222	0.00		< 0.001	>2.5

\$Paral

factor	logodds	tokens	v/v+t	centered	factor weight
9	12.220	5	0.000		> 0.999
z	12.220	2	0.000		> 0.999
6	11.138	10	0.200		> 0.999
t	8.372	25	0.480		> 0.999
1	7.177	14	0.786		0.999
2	6.480	191	0.969		0.998
y	2.138	12	0.417		0.895
4	-12.397	107	0.963		< 0.001
3	-13.631	27	0.889		< 0.001
5	-16.767	73	0.041		< 0.001
7	-16.949	117	0.043		< 0.001

\$RelSoc

factor	logodds	tokens	v/v+t	centered	factor weight
d	8.088	100	0.99		> 0.999
s	-8.088	483	0.52		< 0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	13	-1.421			0.6		0.194

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-21.271	68.541	69.181	0.997	0.997

Run 17 (above) with PVerb + Paral + RelSoc is better than Run 9 without RelSoc, $p = 0.434$

Trying with Escolar...

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor	weight
3	22.013	361	0.97			> 0.999
2	-22.013	222	0.00			< 0.001

\$Paral

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
2	8.876	191	0.969			> 0.999	>2.5
1	8.119	14	0.786			> 0.999	>2.5
y	7.793	12	0.417			> 0.999	>2.5
t	7.675	25	0.480			> 0.999	>2.5
6	6.694	10	0.200			0.999	>2.5
9	5.830	5	0.000			0.997	>2.5
z	5.574	2	0.000			0.996	>2.5
4	-9.639	107	0.963			< 0.001	>2.5
3	-11.426	27	0.889			< 0.001	>2.5
7	-14.440	117	0.043			< 0.001	>2.5
5	-15.055	73	0.041			< 0.001	>2.5

\$Escolar

factor	logodds	tokens	v/v+t	centered	factor	weight
m	0.690	119	0.765			0.666
u	0.181	278	0.421			0.545
k	-0.871	186	0.763			0.295

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	14	-7.564			0.6		0.001

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-20.559	69.118	69.858	0.996	0.996

Run 18 (above) with PVerb + Paral + Escolar is better than Run 9 without Escolar, $p = 0.362$

Adding ExprSuj...

STEP 4 - adding to Run 14 model with PVerb + Paral + ExprSuj

Trying with Gnr...

\$ExprSuj

factor	logodds	tokens	v/v+t	centered	factor	weight
p	1.968	273	0.941			0.877
n	-1.968	310	0.300			0.123

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor	weight
3	21.979	361	0.97			> 0.999
2	-21.979	222	0.00			< 0.001

\$Paral

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
2	8.924	191	0.969			> 0.999	>2.5
z	8.328	2	0.000			> 0.999	>2.5
1	8.040	14	0.786			> 0.999	>2.5

9	7.598	5	0.000		0.999	>2.5
y	7.332	12	0.417		0.999	>2.5
t	7.073	25	0.480		0.999	>2.5
6	5.680	10	0.200		0.997	>2.5
4	-10.296	107	0.963		< 0.001	>2.5
3	-12.849	27	0.889		< 0.001	>2.5
7	-14.751	117	0.043		< 0.001	>2.5
5	-15.079	73	0.041		< 0.001	>2.5

\$Gnr

factor	logodds	tokens	v/v+t	centered	factor	weight
h	1.251	512	0.666			0.778
m	-1.251	71	0.127			0.222

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	14	-8.469			0.6		0

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-14.243	56.485	57.225	0.999	0.996

Run 19 (above) with PVerb + Paral + ExprSuj + Gnr is better than Run 14 without Gnr, $p = 0.0314$

Trying with FxEt...

\$ExprSuj

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
p	29.369	273	0.941			> 0.999	>10
n	-29.369	310	0.300			< 0.001	>10

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
3	51.071	361	0.97			> 0.999	>2.5
2	-51.071	222	0.00			< 0.001	>2.5

\$Paral

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
z	53.450	2	0.000			> 0.999	>7.5
9	52.611	5	0.000			> 0.999	>7.5
2	37.949	191	0.969			> 0.999	>7.5
y	12.953	12	0.417			> 0.999	>7.5
t	-2.925	25	0.480			0.051	>7.5
4	-3.571	107	0.963			0.027	>7.5
6	-4.573	10	0.200			0.01	>7.5
1	-17.908	14	0.786			< 0.001	>7.5
5	-42.065	73	0.041			< 0.001	>7.5
3	-42.471	27	0.889			< 0.001	>7.5
7	-43.451	117	0.043			< 0.001	>7.5

\$FxEt

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
v	39.163	236	0.996			> 0.999	>5
k	-0.312	230	0.300			0.423	>5
j	-38.852	117	0.393			< 0.001	>5

```
$misc.1
  n df intercept overall proportion centered input prob
583 15      1.575                0.6                0.828
```

```
$misc.2
log.likelihood    AIC    AICc Dxy R2
      -8.939 47.878 48.725  1  1
```

Run 20 (above) with PVerb + Paral + ExprSuj + FxEt is better than Run 14 without FxEt, $p = 0.000491$

Trying with RelSoc...

```
$ExprSuj
factor logodds tokens v/v+t centered factor weight
  p   1.901     273 0.941                0.87
  n  -1.901     310 0.300                0.13
```

```
$PVerb
factor logodds tokens v/v+t centered factor weight vif
  3  25.492     361 0.97                > 0.999 >2.5
  2 -25.492     222 0.00                < 0.001 >2.5
```

```
$Paral
factor logodds tokens v/v+t centered factor weight vif
  9  12.210      5 0.000                > 0.999 >2.5
  z  12.210      2 0.000                > 0.999 >2.5
  t   9.945     25 0.480                > 0.999 >2.5
  6   9.658     10 0.200                > 0.999 >2.5
  2   9.105    191 0.969                > 0.999 >2.5
  1   6.911     14 0.786                0.999 >2.5
  y   1.626     12 0.417                0.836 >2.5
  4 -12.594    107 0.963                < 0.001 >2.5
  3 -14.206     27 0.889                < 0.001 >2.5
  5 -17.054     73 0.041                < 0.001 >2.5
  7 -17.811    117 0.043                < 0.001 >2.5
```

```
$RelSoc
factor logodds tokens v/v+t centered factor weight
  d   8.453     100 0.99                > 0.999
  s  -8.453     483 0.52                < 0.001
```

```
$misc.1
  n df intercept overall proportion centered input prob
583 14      -0.93                0.6                0.283
```

```
$misc.2
log.likelihood    AIC    AICc Dxy R2
      -16.287 60.575 61.314 0.999 0.998
```

Run 21 (above) with PVerb + Paral + ExprSuj + RelSoc is better than Run 14 without RelSoc, $p = 0.462$

Trying with Escolar...

```
$ExprSuj
factor logodds tokens v/v+t centered factor weight vif
  p  10.666     273 0.941                > 0.999 >20
```

```

n -10.666      310 0.300                < 0.001 >20

$PVerb
factor logodds tokens v/v+t centered factor weight
  3  32.116     361  0.97                > 0.999
  2 -32.116     222  0.00                < 0.001

$Paral
factor logodds tokens v/v+t centered factor weight vif
  2  20.196     191  0.969                > 0.999 >20
  z  18.198         2  0.000                > 0.999 >20
  9  17.353         5  0.000                > 0.999 >20
  y  15.721        12  0.417                > 0.999 >20
  1   1.456        14  0.786                0.811 >20
  t   0.044        25  0.480                0.511 >20
  6  -1.654        10  0.200                0.161 >20
  4  -1.896       107  0.963                0.131 >20
  3 -21.703        27  0.889                < 0.001 >20
  7 -23.777       117  0.043                < 0.001 >20
  5 -23.937        73  0.041                < 0.001 >20

$Escolar
factor logodds tokens v/v+t centered factor weight vif
  u  13.072     278  0.421                > 0.999 >20
  m  -5.744     119  0.765                0.003 >20
  k  -7.328     186  0.763                0.001 >20

$misc.1
n df intercept overall proportion centered input prob
583 15    -12.21                0.6                0

$misc.2
log.likelihood  AIC  AICc  Dxy  R2
    -13.78  57.56  58.406  0.999  0.999

Run 22 (above) with PVerb + Paral + ExprSuj + Escolar is better
than Run 14 without Escolar, p = 0.0621

Adding FxEt...

STEP 5 - adding to Run 20 model with PVerb + Paral + ExprSuj + FxEt

Trying with Gnr...

$ExprSuj
factor logodds tokens v/v+t centered factor weight vif
  p  29.084     273  0.941                > 0.999 >10
  n -29.084     310  0.300                < 0.001 >10

$PVerb
factor logodds tokens v/v+t centered factor weight
  3  50.781     361  0.97                > 0.999
  2 -50.781     222  0.00                < 0.001

$Paral
factor logodds tokens v/v+t centered factor weight vif
  z  53.172         2  0.000                > 0.999 >20
  9  52.314         5  0.000                > 0.999 >20

```

2	37.194	191	0.969		> 0.999	>20
y	12.465	12	0.417		> 0.999	>20
t	-2.803	25	0.480		0.057	>20
4	-3.284	107	0.963		0.036	>20
6	-4.455	10	0.200		0.011	>20
1	-17.193	14	0.786		< 0.001	>20
5	-41.954	73	0.041		< 0.001	>20
3	-42.616	27	0.889		< 0.001	>20
7	-42.840	117	0.043		< 0.001	>20

\$Gnr

factor	logodds	tokens	v/v+t	centered	factor	weight
h	0.443	512	0.666			0.609
m	-0.443	71	0.127			0.391

\$FxEt

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
v	39.132	236	0.996			> 0.999	>5
k	-0.796	230	0.300			0.311	>5
j	-38.337	117	0.393			< 0.001	>5

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	16	1.338			0.6		0.792

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-8.796	49.592	50.553	1	1

Run 23 (above) with PVerb + Paral + ExprSuj + FxEt + Gnr is better than Run 20 without Gnr, $p = 0.593$

Trying with RelSoc...

\$ExprSuj

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
p	29.534	273	0.941			> 0.999	>10
n	-29.534	310	0.300			< 0.001	>10

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
3	51.239	361	0.97			> 0.999	>2.5
2	-51.239	222	0.00			< 0.001	>2.5

\$Paral

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
z	50.535	2	0.000			> 0.999	>10
9	49.695	5	0.000			> 0.999	>10
y	44.938	12	0.417			> 0.999	>10
2	34.888	191	0.969			> 0.999	>10
t	-6.091	25	0.480			0.002	>10
4	-6.671	107	0.963			0.001	>10
6	-7.803	10	0.200			< 0.001	>10
1	-21.272	14	0.786			< 0.001	>10
5	-45.476	73	0.041			< 0.001	>10
3	-45.881	27	0.889			< 0.001	>10
7	-46.862	117	0.043			< 0.001	>10


```
$FxEt
factor logodds tokens v/v+t centered factor weight vif
v 39.708 236 0.996 > 0.999 >7.5
k -0.499 230 0.300 0.378 >7.5
j -39.209 117 0.393 < 0.001 >7.5
```

```
$RelSoc
factor logodds tokens v/v+t centered factor weight
s 19.325 483 0.52 > 0.999
d -19.325 100 0.99 < 0.001
```

```
$misc.1
n df intercept overall proportion centered input prob
583 16 -14.314 0.6 0
```

```
$misc.2
log.likelihood AIC AICc Dxy R2
-8.939 49.878 50.84 1 1
```

Run 24 (above) with PVerb + Paral + ExprSuj + FxEt + RelSoc is better than Run 20 without RelSoc, $p = 1$

Trying with Escolar...

```
$ExprSuj
factor logodds tokens v/v+t centered factor weight vif
p 39.057 273 0.941 > 0.999 >5
n -39.057 310 0.300 < 0.001 >5
```

```
$PVerb
factor logodds tokens v/v+t centered factor weight vif
3 79.7 361 0.97 > 0.999 >2.5
2 -79.7 222 0.00 < 0.001 >2.5
```

```
$Paral
factor logodds tokens v/v+t centered factor weight vif
z 90.270 2 0.000 > 0.999 >15
9 89.365 5 0.000 > 0.999 >15
2 55.545 191 0.969 > 0.999 >15
y 48.921 12 0.417 > 0.999 >15
t 13.358 25 0.480 > 0.999 >15
6 12.059 10 0.200 > 0.999 >15
4 -24.293 107 0.963 < 0.001 >15
1 -57.485 14 0.786 < 0.001 >15
7 -63.591 117 0.043 < 0.001 >15
5 -81.872 73 0.041 < 0.001 >15
3 -82.277 27 0.889 < 0.001 >15
```

```
$FxEt
factor logodds tokens v/v+t centered factor weight vif
v 58.577 236 0.996 > 0.999 >5
k -0.002 230 0.300 0.499 >5
j -58.574 117 0.393 < 0.001 >5
```

```
$Escolar
factor logodds tokens v/v+t centered factor weight vif
m 19.717 119 0.765 > 0.999 >20
u -0.244 278 0.421 0.439 >20
```

k -19.472 186 0.763 < 0.001 >20

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	17		3.072		0.6		0.956

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-4.159	42.318	43.401	1	1

Run 25 (above) with PVerb + Paral + ExprSuj + FxEt + Escolar is better than Run 20 without Escolar, $p = 0.00839$

Adding Escolar...

STEP 6 - adding to Run 25 model with PVerb + Paral + ExprSuj + FxEt + Escolar

Trying with Gnr...

\$ExprSuj

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
p	55.105	273	0.941			> 0.999	>7.5
n	-55.105	310	0.300			< 0.001	>7.5

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
3	103.506	361	0.97			> 0.999	>2.5
2	-103.506	222	0.00			< 0.001	>2.5

\$Paral

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
z	114.499	2	0.000			> 0.999	>20
9	113.636	5	0.000			> 0.999	>20
y	107.036	12	0.417			> 0.999	>20
2	80.714	191	0.969			> 0.999	>20
t	4.948	25	0.480			0.993	>20
6	4.199	10	0.200			0.985	>20
4	-31.350	107	0.963			< 0.001	>20
7	-87.413	117	0.043			< 0.001	>20
1	-96.924	14	0.786			< 0.001	>20
3	-104.673	27	0.889			< 0.001	>20
5	-104.673	73	0.041			< 0.001	>20

\$Gnr

factor	logodds	tokens	v/v+t	centered	factor	weight
m	8.529	71	0.127			> 0.999
h	-8.529	512	0.666			< 0.001

\$FxEt

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
v	63.161	236	0.996			> 0.999	>7.5
k	5.504	230	0.300			0.996	>7.5
j	-68.664	117	0.393			< 0.001	>7.5

\$Escolar

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
m	19.071	119	0.765			> 0.999	>20

u	16.556	278	0.421		> 0.999	>20
k	-35.628	186	0.763		< 0.001	>20

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	18	4.878			0.6		0.992

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-3.819	43.638	44.851	1	1

Run 26 (above) with PVerb + Paral + ExprSuj + FxEt + Escolar + Gnr
is better than Run 25 without Gnr, p = 0.41

Trying with RelSoc...

\$ExprSuj

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
p	39.029	273	0.941			> 0.999	>5
n	-39.029	310	0.300			< 0.001	>5

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
3	79.735	361	0.97			> 0.999	>2.5
2	-79.735	222	0.00			< 0.001	>2.5

\$Paral

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
z	87.496	2	0.000			> 0.999	>20
9	86.566	5	0.000			> 0.999	>20
y	81.545	12	0.417			> 0.999	>20
2	51.979	191	0.969			> 0.999	>20
t	10.553	25	0.480			> 0.999	>20
6	9.301	10	0.200			> 0.999	>20
4	-27.775	107	0.963			< 0.001	>20
1	-61.114	14	0.786			< 0.001	>20
7	-67.114	117	0.043			< 0.001	>20
5	-85.516	73	0.041			< 0.001	>20
3	-85.922	27	0.889			< 0.001	>20

\$FxEt

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
v	59.816	236	0.996			> 0.999	>10
k	-0.099	230	0.300			0.475	>10
j	-59.717	117	0.393			< 0.001	>10

\$RelSoc

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
s	20.093	483	0.52			> 0.999	>2.5
d	-20.093	100	0.99			< 0.001	>2.5

\$Escolar

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
m	20.391	119	0.765			> 0.999	>20
u	-1.447	278	0.421			0.191	>20
k	-18.944	186	0.763			< 0.001	>20

```
$misc.1
  n df intercept overall proportion centered input prob
583 18   -12.915                0.6                0
```

```
$misc.2
log.likelihood   AIC   AICc Dxy R2
   -4.159 44.318 45.531   1   1
```

Run 27 (above) with PVerb + Paral + ExprSuj + FxEt + Escolar + RelSoc is better than Run 25 without RelSoc, $p = 1$

No significant improvement this step, best model from last step is Run 25

BEST STEP-UP MODEL OF RESPONSE VD IS WITH PREDICTOR(S): PVerb (3.14e-151) + Paral (2.76e-08) + ExprSuj (0.00154) + FxEt (0.000491) + Escolar (0.00839)
[p-values building from null model]

```
$ExprSuj
factor logodds tokens v/v+t centered factor weight vif
  p  39.057    273 0.941                > 0.999 >5
  n -39.057    310 0.300                < 0.001 >5
```

```
$PVerb
factor logodds tokens v/v+t centered factor weight vif
  3   79.7    361 0.97                > 0.999 >2.5
  2  -79.7    222 0.00                < 0.001 >2.5
```

```
$Paral
factor logodds tokens v/v+t centered factor weight vif
  z  90.270     2 0.000                > 0.999 >15
  9  89.365     5 0.000                > 0.999 >15
  2  55.545    191 0.969                > 0.999 >15
  y  48.921    12 0.417                > 0.999 >15
  t  13.358    25 0.480                > 0.999 >15
  6  12.059    10 0.200                > 0.999 >15
  4 -24.293   107 0.963                < 0.001 >15
  1 -57.485    14 0.786                < 0.001 >15
  7 -63.591   117 0.043                < 0.001 >15
  5 -81.872    73 0.041                < 0.001 >15
  3 -82.277    27 0.889                < 0.001 >15
```

```
$FxEt
factor logodds tokens v/v+t centered factor weight vif
  v  58.577    236 0.996                > 0.999 >5
  k  -0.002    230 0.300                0.499 >5
  j -58.574    117 0.393                < 0.001 >5
```

```
$Escolar
factor logodds tokens v/v+t centered factor weight vif
  m  19.717    119 0.765                > 0.999 >20
  u  -0.244    278 0.421                0.439 >20
  k -19.472    186 0.763                < 0.001 >20
```

```
$misc.1
  n df intercept overall proportion centered input prob
583 17     3.072                0.6                0.956
```

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-4.159	42.318	43.401	1	1

STEPPING DOWN...

STEP 0 - Run 0 - full model with ExprSuj + PVerb + Paral + Gnr + FxEt + RelSoc + Escolar

\$ExprSuj

factor	logodds	tokens	v/v+t	centered	factor weight	vif
p	55.05	273	0.941		> 0.999	>7.5
n	-55.05	310	0.300		< 0.001	>7.5

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor weight	vif
3	103.491	361	0.97		> 0.999	>2.5
2	-103.491	222	0.00		< 0.001	>2.5

\$Paral

factor	logodds	tokens	v/v+t	centered	factor weight	vif
z	114.704	2	0.000		> 0.999	>20
9	113.844	5	0.000		> 0.999	>20
y	110.433	12	0.417		> 0.999	>20
2	79.899	191	0.969		> 0.999	>20
t	5.293	25	0.480		0.995	>20
6	4.515	10	0.200		0.989	>20
4	-32.026	107	0.963		< 0.001	>20
7	-88.100	117	0.043		< 0.001	>20
1	-97.642	14	0.786		< 0.001	>20
3	-105.460	27	0.889		< 0.001	>20
5	-105.460	73	0.041		< 0.001	>20

\$Gnr

factor	logodds	tokens	v/v+t	centered	factor weight
m	8.512	71	0.127		> 0.999
h	-8.512	512	0.666		< 0.001

\$FxEt

factor	logodds	tokens	v/v+t	centered	factor weight	vif
v	64.543	236	0.996		> 0.999	>10
k	5.301	230	0.300		0.995	>10
j	-69.845	117	0.393		< 0.001	>10

\$RelSoc

factor	logodds	tokens	v/v+t	centered	factor weight	vif
s	11.457	483	0.52		> 0.999	>5
d	-11.457	100	0.99		< 0.001	>5

\$Escolar

factor	logodds	tokens	v/v+t	centered	factor weight	vif
m	19.807	119	0.765		> 0.999	>20
u	15.179	278	0.421		> 0.999	>20
k	-34.986	186	0.763		< 0.001	>20

```
$misc.1
  n df intercept overall proportion centered input prob
583 19   -5.294                0.6                0.005
```

```
$misc.2
log.likelihood   AIC   AICc Dxy R2
      -3.819 45.638 46.988   1   1
```

STEP 1 - dropping from Run 0 model with ExprSuj + PVerb + Paral +
Gnr + FxEt + RelSoc + Escolar

Trying without ExprSuj...

```
$PVerb
factor logodds tokens v/v+t centered factor weight
   3  21.992    361  0.97                > 0.999
   2 -21.992    222  0.00                < 0.001
```

```
$Paral
factor logodds tokens v/v+t centered factor weight vif
   2  10.305    191  0.969                > 0.999 >2.5
   t   9.876     25  0.480                > 0.999 >2.5
   1   9.621     14  0.786                > 0.999 >2.5
   6   9.023     10  0.200                > 0.999 >2.5
   9   8.010      5  0.000                > 0.999 >2.5
   z   7.624      2  0.000                > 0.999 >2.5
   4  -9.226    107  0.963                < 0.001 >2.5
   y  -9.706     12  0.417                < 0.001 >2.5
   3 -11.352     27  0.889                < 0.001 >2.5
   7 -11.619    117  0.043                < 0.001 >2.5
   5 -12.557     73  0.041                < 0.001 >2.5
```

```
$Gnr
factor logodds tokens v/v+t centered factor weight
   h   0.404    512  0.666                0.6
   m  -0.404     71  0.127                0.4
```

```
$FxEt
factor logodds tokens v/v+t centered factor weight vif
   v  22.162    236  0.996                > 0.999 >20
   k  -0.151    230  0.300                0.462 >20
   j -22.011    117  0.393                < 0.001 >20
```

```
$RelSoc
factor logodds tokens v/v+t centered factor weight vif
   s  11.103    483  0.52                > 0.999 >2.5
   d -11.103    100  0.99                < 0.001 >2.5
```

```
$Escolar
factor logodds tokens v/v+t centered factor weight vif
   m   8.678    119  0.765                > 0.999 >20
   k   6.023    186  0.763                0.998 >20
   u -14.701    278  0.421                < 0.001 >20
```

```
$misc.1
  n df intercept overall proportion centered input prob
583 18   -7.239                0.6                0.001
```

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-13.179	62.357	63.57	0.999	0.998

Run 1 (above) with PVerb + Paral + Gnr + FxEt + RelSoc + Escolar is worse than Run 0 also including ExprSuj, $p = 1.51e-05$

Trying without PVerb...

\$ExprSuj

factor	logodds	tokens	v/v+t	centered	factor weight
p	2.198	273	0.941		0.9
n	-2.198	310	0.300		0.1

\$Paral

factor	logodds	tokens	v/v+t	centered	factor weight	vif
2	8.778	191	0.969		> 0.999	>20
4	6.350	107	0.963		0.998	>20
1	6.329	14	0.786		0.998	>20
t	4.914	25	0.480		0.993	>20
3	4.477	27	0.889		0.989	>20
5	3.436	73	0.041		0.969	>20
7	3.310	117	0.043		0.965	>20
6	3.248	10	0.200		0.963	>20
9	-13.441	5	0.000		< 0.001	>20
y	-13.666	12	0.417		< 0.001	>20
z	-13.734	2	0.000		< 0.001	>20

\$Gnr

factor	logodds	tokens	v/v+t	centered	factor weight
h	0.697	512	0.666		0.667
m	-0.697	71	0.127		0.333

\$FxEt

factor	logodds	tokens	v/v+t	centered	factor weight	vif
v	34.858	236	0.996		> 0.999	>20
k	-0.331	230	0.300		0.418	>20
j	-34.526	117	0.393		< 0.001	>20

\$RelSoc

factor	logodds	tokens	v/v+t	centered	factor weight	vif
s	25.147	483	0.52		> 0.999	>20
d	-25.147	100	0.99		< 0.001	>20

\$Escolar

factor	logodds	tokens	v/v+t	centered	factor weight	vif
m	12.188	119	0.765		> 0.999	>20
k	11.754	186	0.763		> 0.999	>20
u	-23.942	278	0.421		< 0.001	>20

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	18	-8.496			0.6		0

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-38.364	112.728	113.941	0.995	0.993

Run 2 (above) with ExprSuj + Paral + Gnr + FxEt + RelSoc + Escolar
is worse than Run 0 also including PVerb, $p = 9.41e-17$

Trying without Paral...

\$ExprSuj

factor	logodds	tokens	v/v+t	centered	factor	weight
p	0.602	273	0.941			0.646
n	-0.602	310	0.300			0.354

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor	weight
3	19.621	361	0.97			> 0.999
2	-19.621	222	0.00			< 0.001

\$Gnr

factor	logodds	tokens	v/v+t	centered	factor	weight
h	1.347	512	0.666			0.794
m	-1.347	71	0.127			0.206

\$FxEt

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
v	21.211	236	0.996			> 0.999	>20
k	-0.387	230	0.300			0.404	>20
j	-20.824	117	0.393			< 0.001	>20

\$RelSoc

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
s	10.77	483	0.52			> 0.999	>10
d	-10.77	100	0.99			< 0.001	>10

\$Escolar

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
m	7.856	119	0.765			> 0.999	>20
k	6.458	186	0.763			0.998	>20
u	-14.313	278	0.421			< 0.001	>20

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	9	-15.703		0.6			0

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-22.233	62.466	62.78	0.997	0.996

Run 3 (above) with ExprSuj + PVerb + Gnr + FxEt + RelSoc + Escolar
is worse than Run 0 also including Paral, $p = 6.06e-05$

Trying without Gnr...

\$ExprSuj

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
p	39.029	273	0.941			> 0.999	>5
n	-39.029	310	0.300			< 0.001	>5

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
--------	---------	--------	-------	----------	--------	--------	-----

3	79.735	361	0.97		> 0.999	>2.5
2	-79.735	222	0.00		< 0.001	>2.5

\$Paral

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
z	87.496	2	0.000			> 0.999	>20
9	86.566	5	0.000			> 0.999	>20
y	81.545	12	0.417			> 0.999	>20
2	51.979	191	0.969			> 0.999	>20
t	10.553	25	0.480			> 0.999	>20
6	9.301	10	0.200			> 0.999	>20
4	-27.775	107	0.963			< 0.001	>20
1	-61.114	14	0.786			< 0.001	>20
7	-67.114	117	0.043			< 0.001	>20
5	-85.516	73	0.041			< 0.001	>20
3	-85.922	27	0.889			< 0.001	>20

\$FxEt

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
v	59.816	236	0.996			> 0.999	>10
k	-0.099	230	0.300			0.475	>10
j	-59.717	117	0.393			< 0.001	>10

\$RelSoc

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
s	20.093	483	0.52			> 0.999	>2.5
d	-20.093	100	0.99			< 0.001	>2.5

\$Escolar

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
m	20.391	119	0.765			> 0.999	>20
u	-1.447	278	0.421			0.191	>20
k	-18.944	186	0.763			< 0.001	>20

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	18	-12.915			0.6		0

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-4.159	44.318	45.531	1	1

Run 4 (above) with ExprSuj + PVerb + Paral + FxEt + RelSoc + Escolar is worse than Run 0 also including Gnr, p = 0.41

Trying without FxEt...

\$ExprSuj

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
p	10.736	273	0.941			> 0.999	>20
n	-10.736	310	0.300			< 0.001	>20

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor	weight
3	31.945	361	0.97			> 0.999
2	-31.945	222	0.00			< 0.001

\$Paral

factor	logodds	tokens	v/v+t	centered	factor weight	vif
2	20.310	191	0.969		> 0.999	>20
z	19.246	2	0.000		> 0.999	>20
9	18.408	5	0.000		> 0.999	>20
y	6.275	12	0.417		0.998	>20
1	2.043	14	0.786		0.885	>20
t	1.210	25	0.480		0.77	>20
6	-0.874	10	0.200		0.295	>20
4	-1.439	107	0.963		0.192	>20
3	-20.878	27	0.889		< 0.001	>20
7	-21.490	117	0.043		< 0.001	>20
5	-22.813	73	0.041		< 0.001	>20

\$Gnr

factor	logodds	tokens	v/v+t	centered	factor weight
h	0.905	512	0.666		0.712
m	-0.905	71	0.127		0.288

\$RelSoc

factor	logodds	tokens	v/v+t	centered	factor weight
d	10.264	100	0.99		> 0.999
s	-10.264	483	0.52		< 0.001

\$Escolar

factor	logodds	tokens	v/v+t	centered	factor weight	vif
u	12.897	278	0.421		> 0.999	>20
m	-5.186	119	0.765		0.006	>20
k	-7.711	186	0.763		< 0.001	>20

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	17	-3.832			0.6		0.021

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-11.257	56.515	57.598	0.999	0.999

Run 5 (above) with ExprSuj + PVerb + Paral + Gnr + RelSoc + Escolar is worse than Run 0 also including FxEt, $p = 0.000588$

Trying without RelSoc...

\$ExprSuj

factor	logodds	tokens	v/v+t	centered	factor weight	vif
p	55.105	273	0.941		> 0.999	>7.5
n	-55.105	310	0.300		< 0.001	>7.5

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor weight	vif
3	103.506	361	0.97		> 0.999	>2.5
2	-103.506	222	0.00		< 0.001	>2.5

\$Paral

factor	logodds	tokens	v/v+t	centered	factor weight	vif
z	114.499	2	0.000		> 0.999	>20
9	113.636	5	0.000		> 0.999	>20
y	107.036	12	0.417		> 0.999	>20
2	80.714	191	0.969		> 0.999	>20

t	4.948	25	0.480		0.993	>20
6	4.199	10	0.200		0.985	>20
4	-31.350	107	0.963		< 0.001	>20
7	-87.413	117	0.043		< 0.001	>20
1	-96.924	14	0.786		< 0.001	>20
3	-104.673	27	0.889		< 0.001	>20
5	-104.673	73	0.041		< 0.001	>20

\$Gnr

factor	logodds	tokens	v/v+t	centered	factor weight	
m	8.529	71	0.127		> 0.999	
h	-8.529	512	0.666		< 0.001	

\$FxEt

factor	logodds	tokens	v/v+t	centered	factor weight	vif
v	63.161	236	0.996		> 0.999	>7.5
k	5.504	230	0.300		0.996	>7.5
j	-68.664	117	0.393		< 0.001	>7.5

\$Escolar

factor	logodds	tokens	v/v+t	centered	factor weight	vif
m	19.071	119	0.765		> 0.999	>20
u	16.556	278	0.421		> 0.999	>20
k	-35.628	186	0.763		< 0.001	>20

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	18	4.878			0.6		0.992

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-3.819	43.638	44.851	1	1

Run 6 (above) with ExprSuj + PVerb + Paral + Gnr + FxEt + Escolar
is worse than Run 0 also including RelSoc, p = 1

Trying without Escolar...

\$ExprSuj

factor	logodds	tokens	v/v+t	centered	factor weight	vif
p	29.246	273	0.941		> 0.999	>10
n	-29.246	310	0.300		< 0.001	>10

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor weight	vif
3	50.95	361	0.97		> 0.999	>2.5
2	-50.95	222	0.00		< 0.001	>2.5

\$Paral

factor	logodds	tokens	v/v+t	centered	factor weight	vif
z	50.224	2	0.000		> 0.999	>20
9	49.365	5	0.000		> 0.999	>20
y	44.756	12	0.417		> 0.999	>20
2	34.107	191	0.969		> 0.999	>20
t	-5.983	25	0.480		0.003	>20
4	-6.427	107	0.963		0.002	>20
6	-7.705	10	0.200		< 0.001	>20
1	-20.599	14	0.786		< 0.001	>20

5	-45.397	73	0.041		< 0.001	>20
3	-46.058	27	0.889		< 0.001	>20
7	-46.282	117	0.043		< 0.001	>20

\$Gnr

factor	logodds	tokens	v/v+t	centered	factor	weight
h	0.443	512	0.666			0.609
m	-0.443	71	0.127			0.391

\$FxEt

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
v	39.663	236	0.996			> 0.999	>7.5
k	-0.976	230	0.300			0.274	>7.5
j	-38.687	117	0.393			< 0.001	>7.5

\$RelSoc

factor	logodds	tokens	v/v+t	centered	factor	weight
s	19.528	483	0.52			> 0.999
d	-19.528	100	0.99			< 0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	17	-14.728			0.6		0

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-8.796	51.592	52.675	1	1

Run 7 (above) with ExprSuj + PVerb + Paral + Gnr + FxEt + RelSoc is worse than Run 0 also including Escolar, $p = 0.00689$

Dropping RelSoc...

STEP 2 - dropping from Run 6 model with ExprSuj + PVerb + Paral + Gnr + FxEt + Escolar

Trying without ExprSuj...

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
3	29.038	361	0.97			> 0.999	>2.5
2	-29.038	222	0.00			< 0.001	>2.5

\$Paral

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
9	18.310	5	0.000			> 0.999	>5
z	17.940	2	0.000			> 0.999	>5
6	17.509	10	0.200			> 0.999	>5
t	8.303	25	0.480			> 0.999	>5
2	7.049	191	0.969			0.999	>5
1	6.908	14	0.786			0.999	>5
4	-13.018	107	0.963			< 0.001	>5
3	-15.144	27	0.889			< 0.001	>5
7	-15.411	117	0.043			< 0.001	>5
y	-16.097	12	0.417			< 0.001	>5
5	-16.349	73	0.041			< 0.001	>5

\$Gnr

factor	logodds	tokens	v/v+t	centered	factor	weight
h	0.404	512	0.666			0.6
m	-0.404	71	0.127			0.4

\$FxEt

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
v	20.873	236	0.996			> 0.999	>20
k	-0.082	230	0.300			0.48	>20
j	-20.792	117	0.393			< 0.001	>20

\$Escolar

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
m	8.295	119	0.765			> 0.999	>20
k	5.640	186	0.763			0.996	>20
u	-13.935	278	0.421			< 0.001	>20

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	17	-0.226			0.6		0.444

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-13.179	60.357	61.44	0.999	0.999

Run 8 (above) with PVerb + Paral + Gnr + FxEt + Escolar is worse than Run 6 also including ExprSuj, $p = 1.51e-05$

Trying without PVerb...

\$ExprSuj

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
p	2.386	273	0.941			0.916	>2.5
n	-2.386	310	0.300			0.084	>2.5

\$Paral

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
2	6.864	191	0.969			0.999	>10
4	4.365	107	0.963			0.987	>10
1	4.142	14	0.786			0.984	>10
t	2.659	25	0.480			0.935	>10
3	2.608	27	0.889			0.931	>10
5	1.162	73	0.041			0.762	>10
7	1.059	117	0.043			0.743	>10
6	0.867	10	0.200			0.704	>10
y	-2.374	12	0.417			0.085	>10
9	-10.590	5	0.000			< 0.001	>10
z	-10.762	2	0.000			< 0.001	>10

\$Gnr

factor	logodds	tokens	v/v+t	centered	factor	weight
h	0.674	512	0.666			0.662
m	-0.674	71	0.127			0.338

\$FxEt

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
v	4.889	236	0.996			0.993	>10
k	-0.577	230	0.300			0.36	>10
j	-4.311	117	0.393			0.013	>10

\$Escolar

factor	logodds	tokens	v/v+t	centered	factor weight	vif
m	1.994	119	0.765		0.88	>7.5
k	1.437	186	0.763		0.808	>7.5
u	-3.432	278	0.421		0.031	>7.5

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	17	-1.255			0.6		0.222

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-41.405	116.81	117.894	0.992	0.948

Run 9 (above) with ExprSuj + Paral + Gnr + FxEt + Escolar is worse than Run 6 also including PVerb, $p = 4.31e-18$

Trying without Paral...

\$ExprSuj

factor	logodds	tokens	v/v+t	centered	factor weight
p	0.602	273	0.941		0.646
n	-0.602	310	0.300		0.354

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor weight
3	28.527	361	0.97		> 0.999
2	-28.527	222	0.00		< 0.001

\$Gnr

factor	logodds	tokens	v/v+t	centered	factor weight
h	1.347	512	0.666		0.794
m	-1.347	71	0.127		0.206

\$FxEt

factor	logodds	tokens	v/v+t	centered	factor weight	vif
v	19.530	236	0.996		> 0.999	>20
k	-0.477	230	0.300		0.383	>20
j	-19.053	117	0.393		< 0.001	>20

\$Escolar

factor	logodds	tokens	v/v+t	centered	factor weight	vif
m	7.235	119	0.765		0.999	>20
k	5.837	186	0.763		0.997	>20
u	-13.072	278	0.421		< 0.001	>20

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	8	-14.989			0.6		0

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-22.233	60.466	60.717	0.997	0.998

Run 10 (above) with ExprSuj + PVerb + Gnr + FxEt + Escolar is worse than Run 6 also including Paral, $p = 6.06e-05$

Trying without Gnr...

\$ExprSuj

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
p	39.057	273	0.941			> 0.999	>5
n	-39.057	310	0.300			< 0.001	>5

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
3	79.7	361	0.97			> 0.999	>2.5
2	-79.7	222	0.00			< 0.001	>2.5

\$Paral

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
z	90.270	2	0.000			> 0.999	>15
9	89.365	5	0.000			> 0.999	>15
2	55.545	191	0.969			> 0.999	>15
y	48.921	12	0.417			> 0.999	>15
t	13.358	25	0.480			> 0.999	>15
6	12.059	10	0.200			> 0.999	>15
4	-24.293	107	0.963			< 0.001	>15
1	-57.485	14	0.786			< 0.001	>15
7	-63.591	117	0.043			< 0.001	>15
5	-81.872	73	0.041			< 0.001	>15
3	-82.277	27	0.889			< 0.001	>15

\$FxEt

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
v	58.577	236	0.996			> 0.999	>5
k	-0.002	230	0.300			0.499	>5
j	-58.574	117	0.393			< 0.001	>5

\$Escolar

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
m	19.717	119	0.765			> 0.999	>20
u	-0.244	278	0.421			0.439	>20
k	-19.472	186	0.763			< 0.001	>20

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	17	3.072		0.6			0.956

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-4.159	42.318	43.401	1	1

Run 11 (above) with ExprSuj + PVerb + Paral + FxEt + Escolar is worse than Run 6 also including Gnr, p = 0.41

Trying without FxEt...

\$ExprSuj

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
p	10.617	273	0.941			> 0.999	>20
n	-10.617	310	0.300			< 0.001	>20

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor	weight
--------	---------	--------	-------	----------	--------	--------

3	32.063	361	0.97	> 0.999
2	-32.063	222	0.00	< 0.001

\$Paral

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
2	19.489	191	0.969			> 0.999	>20
z	18.265	2	0.000			> 0.999	>20
9	17.419	5	0.000			> 0.999	>20
y	15.862	12	0.417			> 0.999	>20
1	1.254	14	0.786			0.778	>20
t	0.167	25	0.480			0.542	>20
6	-1.760	10	0.200			0.147	>20
4	-1.921	107	0.963			0.128	>20
3	-22.045	27	0.889			< 0.001	>20
7	-22.885	117	0.043			< 0.001	>20
5	-23.844	73	0.041			< 0.001	>20

\$Gnr

factor	logodds	tokens	v/v+t	centered	factor	weight
h	1.071	512	0.666			0.745
m	-1.071	71	0.127			0.255

\$Escolar

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
u	12.507	278	0.421			> 0.999	>20
m	-5.452	119	0.765			0.004	>20
k	-7.055	186	0.763			0.001	>20

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	16	-12.883			0.6		0

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-12.213	56.425	57.387	0.999	0.999

Run 12 (above) with ExprSuj + PVerb + Paral + Gnr + Escolar is worse than Run 6 also including FxEt, $p = 0.000226$

Trying without Escolar...

\$ExprSuj

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
p	29.084	273	0.941			> 0.999	>10
n	-29.084	310	0.300			< 0.001	>10

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor	weight
3	50.781	361	0.97			> 0.999
2	-50.781	222	0.00			< 0.001

\$Paral

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
z	53.172	2	0.000			> 0.999	>20
9	52.314	5	0.000			> 0.999	>20
2	37.194	191	0.969			> 0.999	>20
y	12.465	12	0.417			> 0.999	>20
t	-2.803	25	0.480			0.057	>20

4	-3.284	107	0.963		0.036	>20
6	-4.455	10	0.200		0.011	>20
1	-17.193	14	0.786		< 0.001	>20
5	-41.954	73	0.041		< 0.001	>20
3	-42.616	27	0.889		< 0.001	>20
7	-42.840	117	0.043		< 0.001	>20

\$Gnr

factor	logodds	tokens	v/v+t	centered	factor	weight
h	0.443	512	0.666			0.609
m	-0.443	71	0.127			0.391

\$FxEt

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
v	39.132	236	0.996			> 0.999	>5
k	-0.796	230	0.300			0.311	>5
j	-38.337	117	0.393			< 0.001	>5

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	16	1.338			0.6		0.792

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-8.796	49.592	50.553	1	1

Run 13 (above) with ExprSuj + PVerb + Paral + Gnr + FxEt is worse than Run 6 also including Escolar, $p = 0.00689$

Dropping Gnr...

STEP 3 - dropping from Run 11 model with ExprSuj + PVerb + Paral + FxEt + Escolar

Trying without ExprSuj...

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor	weight
3	28.988	361	0.97			> 0.999
2	-28.988	222	0.00			< 0.001

\$Paral

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
9	18.279	5	0.000			> 0.999	>2.5
z	17.500	2	0.000			> 0.999	>2.5
6	17.258	10	0.200			> 0.999	>2.5
t	8.806	25	0.480			> 0.999	>2.5
2	7.421	191	0.969			0.999	>2.5
1	6.370	14	0.786			0.998	>2.5
4	-12.823	107	0.963			< 0.001	>2.5
3	-14.912	27	0.889			< 0.001	>2.5
7	-15.660	117	0.043			< 0.001	>2.5
y	-16.040	12	0.417			< 0.001	>2.5
5	-16.200	73	0.041			< 0.001	>2.5

\$FxEt

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
v	20.907	236	0.996			> 0.999	>20

k	0.093	230	0.300		0.523	>20
j	-21.000	117	0.393		< 0.001	>20

\$Escolar

factor	logodds	tokens	v/v+t	centered	factor weight	vif
m	8.362	119	0.765		> 0.999	>20
k	5.634	186	0.763		0.996	>20
u	-13.996	278	0.421		< 0.001	>20

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	16	0.046			0.6		0.511

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-13.349	58.698	59.66	0.999	0.999

Run 14 (above) with PVerb + Paral + FxEt + Escolar is worse than Run 11 also including ExprSuj, $p = 1.81e-05$

Trying without PVerb...

\$ExprSuj

factor	logodds	tokens	v/v+t	centered	factor weight	vif
p	2.395	273	0.941		0.916	>2.5
n	-2.395	310	0.300		0.084	>2.5

\$Paral

factor	logodds	tokens	v/v+t	centered	factor weight	vif
2	7.174	191	0.969		0.999	>7.5
4	4.523	107	0.963		0.989	>7.5
1	3.569	14	0.786		0.973	>7.5
3	2.832	27	0.889		0.944	>7.5
t	2.809	25	0.480		0.943	>7.5
5	1.173	73	0.041		0.764	>7.5
6	1.126	10	0.200		0.755	>7.5
7	0.658	117	0.043		0.659	>7.5
y	-2.150	12	0.417		0.104	>7.5
9	-10.475	5	0.000		< 0.001	>7.5
z	-11.239	2	0.000		< 0.001	>7.5

\$FxEt

factor	logodds	tokens	v/v+t	centered	factor weight	vif
v	4.979	236	0.996		0.993	>10
k	-0.402	230	0.300		0.401	>10
j	-4.577	117	0.393		0.01	>10

\$Escolar

factor	logodds	tokens	v/v+t	centered	factor weight	vif
m	2.037	119	0.765		0.885	>7.5
k	1.367	186	0.763		0.797	>7.5
u	-3.404	278	0.421		0.032	>7.5

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	16	-0.89			0.6		0.291

\$misc.2

```
log.likelihood      AIC      AICc    Dxy    R2
      -42.703 117.405 118.366 0.991 0.948
```

Run 15 (above) with ExprSuj + Paral + FxEt + Escolar is worse than Run 11 also including PVerb, $p = 1.64e-18$

Trying without Paral...

\$ExprSuj

```
factor logodds tokens v/v+t centered factor weight
  p      0.2      273 0.941
  n     -0.2      310 0.300
                                0.55
                                0.45
```

\$PVerb

```
factor logodds tokens v/v+t centered factor weight
  3  28.843      361 0.97
  2 -28.843      222 0.00
                                > 0.999
                                < 0.001
```

\$FxEt

```
factor logodds tokens v/v+t centered factor weight vif
  v  20.068      236 0.996
  k  -0.364      230 0.300
  j -19.703      117 0.393
                                > 0.999 >20
                                0.41 >20
                                < 0.001 >20
```

\$Escolar

```
factor logodds tokens v/v+t centered factor weight vif
  m   7.247      119 0.765
  k   5.797      186 0.763
  u -13.043      278 0.421
                                0.999 >20
                                0.997 >20
                                < 0.001 >20
```

\$misc.1

```
n df intercept overall proportion centered input prob
583 7 -14.277
                                0.6
                                0
```

\$misc.2

```
log.likelihood      AIC      AICc    Dxy    R2
      -26.898 67.797 67.992 0.996 0.998
```

Run 16 (above) with ExprSuj + PVerb + FxEt + Escolar is worse than Run 11 also including Paral, $p = 1.78e-06$

Trying without FxEt...

\$ExprSuj

```
factor logodds tokens v/v+t centered factor weight vif
  p  10.666      273 0.941
  n -10.666      310 0.300
                                > 0.999 >20
                                < 0.001 >20
```

\$PVerb

```
factor logodds tokens v/v+t centered factor weight
  3  32.116      361 0.97
  2 -32.116      222 0.00
                                > 0.999
                                < 0.001
```

\$Paral

```
factor logodds tokens v/v+t centered factor weight vif
  2  20.196      191 0.969
  z  18.198       2 0.000
  9  17.353       5 0.000
                                > 0.999 >20
                                > 0.999 >20
                                > 0.999 >20
```

y	15.721	12	0.417		>	0.999	>20
1	1.456	14	0.786			0.811	>20
t	0.044	25	0.480			0.511	>20
6	-1.654	10	0.200			0.161	>20
4	-1.896	107	0.963			0.131	>20
3	-21.703	27	0.889		<	0.001	>20
7	-23.777	117	0.043		<	0.001	>20
5	-23.937	73	0.041		<	0.001	>20

\$Escolar

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
u	13.072	278	0.421			>	0.999 >20
m	-5.744	119	0.765				0.003 >20
k	-7.328	186	0.763				0.001 >20

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
583	15	-12.21			0.6		0

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-13.78	57.56	58.406	0.999	0.999

Run 17 (above) with ExprSuj + PVerb + Paral + Escolar is worse than Run 11 also including FxEt, $p = 6.63e-05$

Trying without Escolar...

\$ExprSuj

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
p	29.369	273	0.941			>	0.999 >10
n	-29.369	310	0.300			<	0.001 >10

\$PVerb

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
3	51.071	361	0.97			>	0.999 >2.5
2	-51.071	222	0.00			<	0.001 >2.5

\$Paral

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
z	53.450	2	0.000			>	0.999 >7.5
9	52.611	5	0.000			>	0.999 >7.5
2	37.949	191	0.969			>	0.999 >7.5
y	12.953	12	0.417			>	0.999 >7.5
t	-2.925	25	0.480				0.051 >7.5
4	-3.571	107	0.963				0.027 >7.5
6	-4.573	10	0.200				0.01 >7.5
1	-17.908	14	0.786			<	0.001 >7.5
5	-42.065	73	0.041			<	0.001 >7.5
3	-42.471	27	0.889			<	0.001 >7.5
7	-43.451	117	0.043			<	0.001 >7.5

\$FxEt

factor	logodds	tokens	v/v+t	centered	factor	weight	vif
v	39.163	236	0.996			>	0.999 >5
k	-0.312	230	0.300				0.423 >5
j	-38.852	117	0.393			<	0.001 >5

```
$misc.1
  n df intercept overall proportion centered input prob
583 15      1.575                0.6                0.828
```

```
$misc.2
log.likelihood      AIC      AICc Dxy R2
      -8.939 47.878 48.725   1   1
```

Run 18 (above) with ExprSuj + PVerb + Paral + FxEt is worse than Run 11 also including Escolar, $p = 0.00839$

All remaining predictors are significant, best model from last step is Run 11

BEST STEP-DOWN MODEL OF RESPONSE VD IS WITH PREDICTOR(S): PVerb (1.64e-18) + Paral (1.78e-06) + ExprSuj (1.81e-05) + FxEt (6.63e-05) + Escolar (0.00839)
[p-values dropping from full model]

```
$ExprSuj
factor logodds tokens v/v+t centered factor weight vif
  p  39.057     273 0.941                > 0.999 >5
  n -39.057     310 0.300                < 0.001 >5
```

```
$PVerb
factor logodds tokens v/v+t centered factor weight vif
  3    79.7     361 0.97                > 0.999 >2.5
  2   -79.7     222 0.00                < 0.001 >2.5
```

```
$Paral
factor logodds tokens v/v+t centered factor weight vif
  z   90.270      2 0.000                > 0.999 >15
  9   89.365      5 0.000                > 0.999 >15
  2   55.545     191 0.969                > 0.999 >15
  y   48.921     12 0.417                > 0.999 >15
  t   13.358     25 0.480                > 0.999 >15
  6   12.059     10 0.200                > 0.999 >15
  4  -24.293     107 0.963                < 0.001 >15
  1  -57.485     14 0.786                < 0.001 >15
  7  -63.591     117 0.043                < 0.001 >15
  5  -81.872     73 0.041                < 0.001 >15
  3  -82.277     27 0.889                < 0.001 >15
```

```
$FxEt
factor logodds tokens v/v+t centered factor weight vif
  v   58.577     236 0.996                > 0.999 >5
  k   -0.002     230 0.300                0.499 >5
  j  -58.574     117 0.393                < 0.001 >5
```

```
$Escolar
factor logodds tokens v/v+t centered factor weight vif
  m   19.717     119 0.765                > 0.999 >20
  u   -0.244     278 0.421                0.439 >20
  k  -19.472     186 0.763                < 0.001 >20
```

```
$misc.1
  n df intercept overall proportion centered input prob
583 17      3.072                0.6                0.956
```

```
$misc.2
log.likelihood    AIC    AICc Dxy R2
      -4.159 42.318 43.401   1   1
```

**BEST STEP-UP MODEL OF RESPONSE VD IS WITH PREDICTOR(S): PVerb
(3.14e-151) + Paral (2.76e-08) + ExprSuj (0.00154) + FxEt
(0.000491) + Escolar (0.00839)**
[p-values building from null model]

BEST STEP-DOWN MODEL OF RESPONSE VD IS WITH PREDICTOR(S): PVerb
(1.64e-18) + Paral (1.78e-06) + ExprSuj (1.81e-05) + FxEt (6.63e-
05) + Escolar (0.00839)
[p-values dropping from full model]

STEP-UP AND STEP-DOWN MATCH!