

UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGUÍSTICOS

Fillipe Dias Araújo Vieira

AS ESTRATÉGIAS DATIVAS DE SEGUNDA PESSOA DO SINGULAR
EM CARTAS BRASILEIRAS: uma análise variacionista

Belo Horizonte
Faculdade de Letras da UFMG
2022

UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGUÍSTICOS

Fillipe Dias Araújo Vieira

AS ESTRATÉGIAS DATIVAS DE SEGUNDA PESSOA DO SINGULAR
EM CARTAS BRASILEIRAS: uma análise variacionista

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Estudos Linguísticos da Faculdade de Letras da Universidade Federal de Minas Gerais, como requisito parcial para a obtenção do título de Mestre em Estudos Linguísticos. Área de Concentração: Linguística Teórica e Descritiva. Linha de Pesquisa: 1A - Estudo da Variação e Mudança Linguística.

Orientadora: Prof^ª. Dr^ª. Márcia Cristina de Brito Rumeu.

Belo Horizonte
Faculdade de Letras
2022

V658e

Vieira, Fillipe Dias Araújo.

As estratégias dativas de segunda pessoa do singular em cartas brasileiras [manuscrito]: uma análise variacionista / Fillipe Dias Araújo Vieira. – 2022.

1 recurso online (157 f.: il., tab., p&b., map., color.): pdf.

Orientadora: Márcia Cristina de Brito Rumeu.

Área de concentração: Linguística Teórica e Descritiva.

Linha de Pesquisa: Mudança e Variação Linguística.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Faculdade de Letras.

Bibliografia: f. 111-117.

Apêndices: f. 118-157.

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

. CDD: 469.798



UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGUÍSTICOS

FOLHA DE APROVAÇÃO

As estratégias dativas de segunda pessoa do singular em cartas brasileiras: uma análise variacionista

FILLIPE DIAS ARAÚJO VIEIRA

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

Aprovada em 22 de fevereiro de 2022, pela banca constituída pelos membros:

Prof(a). Márcia Cristina de Brito Rumeu - Orientadora

UFMG

Prof(a). Thiago Laurentino de Oliveira

UFRJ

Prof(a). Lorenzo Teixeira Vitral

UFMG

Belo Horizonte, 22 de fevereiro de 2022.



Documento assinado eletronicamente por **Marcia Cristina de Brito Rumeu, Professora do Magistério Superior**, em 22/02/2022, às 23:15, 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 **Thiago Laurentino de Oliveira, Usuário Externo**, em 23/02/2022, às 11:14, 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 **Lorenzo Teixeira Vitral, Professor do Magistério Superior**, em 07/03/2022, às 11:55, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, 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_org_ao_acesso_externo=0, informando o código verificador **1215972** e o código CRC **1CBFA4A6**.

UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGÜÍSTICOS

Dissertação de Mestrado intitulada “As estratégias dativas de segunda pessoa do singular em cartas brasileiras: uma análise variacionista”, defendida por Fillipe Dias Araújo Vieira, em 22 de fevereiro de 2022, à Banca Examinadora constituída pelos seguintes membros:

Prof^ª. Dra. Márcia Cristina de Brito Rumeu - Orientadora
Universidade Federal de Minas Gerais – UFMG

Prof. Dr. Thiago Laurentino de Oliveira – Examinador
Universidade Federal de Minas Gerais – UFRJ

Prof. Dr. Lorenzo Teixeira Vitral – Examinador
Universidade Federal do Rio de Janeiro – UFMG

Prof^ª. Dra. Maria Cândida Trindade Costa de Seabra – Examinadora Suplente
Universidade Federal de Minas Gerais – UFMG

Belo Horizonte
Faculdade de Letras da UFMG
2022

AGRADECIMENTOS

À Deus, toda a minha gratidão, amor e dedicação.

Agradeço aos meus familiares e amigos que tiveram a paciência de ouvir minhas angústias e a bondade de perdoar minhas abstenções necessárias ao desenvolvimento deste trabalho de Mestrado.

Agradeço, também, a todos aqueles que estiveram comigo nesse período e, de alguma forma, contribuíram para o meu trabalho.

Agradeço ao Grupo Montalvão, escritório que se tornou mais que meu trabalho, mas minha segunda casa.

Agradeço à minha orientadora, Prof^a. Márcia Rumeu.

Por fim, agradeço à Universidade Federal de Minas Gerais (UFMG) e à Faculdade de Letras (FALE) pelo suporte tão essencial a este Mestrado.

VIEIRA, Fillipe Dias Araújo. *As estratégias dativas de segunda pessoa do singular em cartas brasileiras: uma análise variacionista*. 2022. 157 fls. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Federal de Minas Gerais, Belo Horizonte, 2022.

RESUMO

Nesta Dissertação de Mestrado, buscamos descrever analiticamente, à luz dos princípios teórico-metodológicos da Teoria da Variação Linguística de orientação Laboviana (LABOV, 1972; LABOV, 1994), as formas pronominais dativas de 2ª pessoa do singular produtivas em cartas brasileiras (Cartas ao Papai Noel) redigidas por crianças, adolescentes e adultos, no espaço geográfico de Minas Gerais, no ano de 2012. A hipótese é a de que tenhamos o *te-dativo* (forma do paradigma etimológico do *tu*) em alta produtividade no contexto de *você-sujeito* das cartas brasileiras produzidas por redatores que nasceram e vivem no espaço geográfico de Minas Gerais, que é o espaço de prevalência do *você-sujeito* (SCHERRE *et alii*, 2015). Conjecturamos ainda que as estratégias dativas do paradigma de *você-sujeito* (*a/para você*) tendam a acompanhar a opção pelo *você-sujeito* nas cartas brasileiras em análise, como já discutido por Lopes & Cavalcante (2011), Lopes *et alii* (2018), Rumeu (2020) em relação às cartas brasileiras de sincronias passadas (séculos XIX e XX).

As estratégias dativas de 2ª pessoa do singular foram submetidas ao RBrul (RStudio) para a devida geração dos índices percentuais e probabilísticos voltados para a regra variável eneária em discussão: *te versus lhe, a ti, para ti, a você, para você, você e zero*. Em termos percentuais, verificamos, uma acirrada concorrência entre as formas *zero* (48,14%, 594/1232) e *te* (33,69%, 415/1232), cabendo ao *lhe* o *status* de terceira estratégia pronominal dativa mais produtiva (10,71%, 132/1232) nas cartas analisadas. Em termos probabilísticos, tão somente as variáveis linguísticas (*pronome-sujeito de 2SG, semântica do dativo e categoria semântica do predicador*) mostram-se propulsoras da regra variável em análise a partir do clítico *te* como valor de aplicação.

Em síntese, temos os contextos de *você-sujeito* (0.558), as semânticas do dativo como *meta/fonte* (0.724), como *receptor* (0.607) e os verbos de *movimento abstrato* (0.722) e de

transferência material (0.6) como dinamizadores do *te-dativo* de 2SG nas cartas brasileiras analisadas.

Palavras-chave: pronomes de 2ª pessoa do singular, complementação verbal dativa, variação linguística.

VIEIRA, Fillipe Dias Araújo. *The 2nd person dative strategies in brazilian letters: an variationist analysis*. 2022. 157 fls. Dissertation (Master in Linguistic Studies) – Universidade Federal de Minas Gerais, Belo Horizonte, 2022.

ABSTRACT

In this master's thesis, we seek to analytically describe, in the light of the theoretical-methodological principles of the Labovian Theory of Linguistic Variation (LABOV, 1972; LABOV, 1994), the productive 2nd person singular dative pronominal forms in Brazilian letters (Letters to the Santa Claus), written by children, adolescents and adults, in the geographic space of Minas Gerais, in 2012. The hypothesis is that we have the *te-dativo* (form of the etymological paradigm of *tu*) in high productivity in the context of *você-sujeito* of Brazilian letters produced by writers who were born and live in the geographic space of Minas Gerais, which is the space of prevalence of *você-sujeito* (SCHERRE et alii, 2015). We also conjecture that the dative strategies of the paradigm of *você-sujeito* (*a/para você*) tend to accompany the option for *você-sujeito* in the Brazilian letters under analysis, as already discussed by Lopes & Cavalcante (2011), Lopes et alii (2018), Rumeu (2020) in relation to Brazilian letters of past synchronies (19th and 20th centuries).

The 2nd person singular dative strategies were submitted to RBrul (RStudio) for the proper generation of percentage and probabilistic indices aimed at the ennea variable rule under discussion: *te* versus *lhe*, *a ti*, *para ti*, *a você*, *para você*, *você* and *null*. In percentage terms, we found a fierce competition between the forms *null* (48.14%, 594/1232) and *te* (33.69%, 415/1232), and *lhe* was the third most productive dative pronominal strategy (10.71%, 132/1232) in the analysed letters. In probabilistic terms, only the linguistic variables (*pronoun-subject of 2SG*, *semantics of the dative* and *semantic category of the predicator*) are shown to be propellers of the variable rule under analysis from the clitic *te* as application value.

In summary, we have the contexts of *você-sujeito* (0.558), the semantics of the dative as a *goal/source* (0.724), as a *receiver* (0.607) and the verbs of *abstract movement* (0.722) and *material transference* (0.6) as dynamizers of the *te-dativo* of 2SG in the Brazilian letters analysed.

Keywords: 2nd person singular pronouns, dative verbal complementation, linguistic variation.

LISTA DE ABREVIATURAS E SIGLAS

- 1PL = Primeira pessoa do plural
2PL = Segunda pessoa do plural
3PL = Terceira pessoa do plural
1SG = Primeira pessoa do singular
2SG = Segunda pessoa do singular
3SG = Terceira pessoa do singular
ACU = Acusativo
AEM = Acervo dos Escritores Mineiros
APM = Arquivo Público Mineiro
CR = Complemento relativo
DAT = Dativo
DE = Discurso direto para o entrevistador
DP = *Determiner phrase* - Sintagma determinante
DR3 = Discurso relatado de terceiros
DRF = Discurso relatado do falante
ECT = Empresa Brasileira de Correios e Telégrafos
FALE = Faculdade de Letras da UFMG
GT = Gramática tradicional
IHGMG = Instituto Histórico e Geográfico de Minas Gerais
NGB = Nomenclatura Gramatical Brasileira
Oco = Ocorrências
OBL = Oblíquo
PB = Português brasileiro
PN = Papai Noel
PR = Peso relativo
PREP = Preposição
OD = Objeto direto
OI = Objeto indireto

P.S. = *Post scriptum*

SU= Sujeito

UFMG = Universidade Federal de Minas Gerais

V= Verbo

LISTAS DE IMAGENS E TABELAS

Lista de Imagens e Quadros

Imagem 1: Fac-símile da Carta 930.....	42
Imagem 2: Edição proposta por Vitral & Moreira (2020, p. 44).....	45
Imagem 3: Mesorregiões de Minas Gerais propostas pelo IBGE	67
Quadro 1 Os pronomes pessoais da língua portuguesa	26
Quadro 2: Valor semântico do predicador verbal: alguns exemplos.....	81

Lista de Tabelas

Tabela 1: Estratégias pronominais dativas de 2SG produtivas nas cartas brasileiras	70
Tabela 2: Estratégias pronominais dativas de 2SG correlacionadas ao sujeito de 2SG	72
Tabela 3: Estratégias pronominais dativas de 2SG correlacionadas aos padrões sintáticos dos verbos	75
Tabela 4: Estratégias pronominais dativas de 2SG correlacionadas ao OD em estruturas ditransitivas.....	77
Tabela 5: Estratégias pronominais dativas de 2SG correlacionadas à semântica do predicador verbal	82
Tabela 6: Estratégias pronominais dativas de 2SG correlacionadas às semânticas do dativo de 2SG	88
Tabela 7: Estratégias pronominais dativas de 2SG correlacionadas às seções das cartas	92
Tabela 8: Estratégias pronominais dativas de 2SG correlacionadas à procedência das cartas	96
Tabela 9: O efeito do contexto de sujeito de 2SG em relação ao te-dativo.....	103
Tabela 10: O efeito da semântica do dativo em relação ao te-dativo	104
Tabela 11: O efeito da categoria semântica do predicador em relação ao te-dativo	105
Tabela 12: Síntese das variáveis independentes estatisticamente relevantes à aplicação do te-dativo	106

SUMÁRIO

CONSIDERAÇÕES INICIAIS	18
CAPÍTULO 1 REVISÃO HISTÓRICO-DESCRITIVA	21
1.1 Os complementos dativos de 2SG: a origem latina.....	21
1.2 Os complementos dativos de 2SG: a visão da gramática tradicional.....	25
1.3 Os complementos dativos de 2SG: a visão de alguns estudos linguísticos	27
1.3.1 Os complementos dativos de 2SG em sincronias passadas: os séculos XIX e XX em cena.....	28
1.3.2 Os complementos dativos de 2SG em sincronias recentes: algumas análises embasadas na fala brasileira.....	34
1.3.2.1 Os complementos dativos de 2SG no nordeste brasileiro	34
1.3.2.2 Os complementos dativos de 2SG no sul brasileiro	35
Síntese do capítulo	37
CAPÍTULO 2. PRESSUPOSTOS TEÓRICO-METODOLÓGICOS.....	38
2.1 A pesquisa sociolinguística: princípios e orientações metodológicas para um estudo quantitativo	38
2.2 O estudo da variação linguística em cartas: questões metodológicas	41
2.3 O <i>corpus</i>: a natureza das Cartas ao Papai Noel.....	44
2.4 O RBrul e o trabalho na interface do RStudio.....	45
2.5 O contexto de complementação verbal dativa e os critérios de análise linguística:	48
a caracterização da regra variável	48
2.5.1 O dativo de 2ª pessoa do singular: traços sintáticos e semânticos	48
2.5.2 Os complementos dativos de 2SG: a variável dependente eneária	50

2.5.3 Os complementos dativos de 2SG e os critérios de análise: as variáveis independentes	51
2.5.3.1. O contexto de sujeito de 2SG	51
2.5.3.2 O tipo de predicador em relação à estrutura argumental da sentença	53
2.5.3.3 A forma do objeto direto nas estruturas ditransitivas de complementação verbal	54
2.5.3.4 O valor semântico do predicador.....	55
2.5.3.5 Os papéis semânticos do dativo.....	62
2.5.3.6 As seções das cartas.....	65
2.5.3.7 A procedência das cartas no espaço geográfico mineiro	66
Síntese do capítulo	69
CAPÍTULO 3 ANÁLISE GERAL: RESULTADOS GERAIS EM TERMOS PERCENTUAIS	70
3.1 Os contextos linguísticos: o tipo sintático de predicador, a forma do objeto direto nas estruturas ditransitivas, o valor semântico do predicador, o papel semântico do dativo	71
3.1.1 A correlação entre as estratégias dativas e nominativas de 2SG.....	71
3.1.2 As estratégias dativas de 2SG correlacionadas aos padrões sintáticos	74
3.1.3 A forma do objeto direto em estruturas ditransitivas de complementação verbal	77
3.1.4 A semântica do predicador verbal	81
3.1.5 A semântica do dativo	87
3.2 Os contextos sociolinguísticos: a seção da carta e a procedência da carta	91
3.2.1 Os complementos dativos de 2SG e a seção da carta.....	91
3.2.2 Os complementos dativos de 2SG e a procedência das cartas	96
Síntese do capítulo	98

CAPÍTULO 4. AS ESTRATÉGIAS DATIVAS DE 2SG: DESCRIÇÃO-ANALÍTICA DOS RESULTADOS ESTATISTICAMENTE RELEVANTES	101
Síntese do capítulo	106
CONSIDERAÇÕES FINAIS.....	108
REFERÊNCIAS BIBLIOGRÁFICAS	111
APÊNDICE (Rodada nos níveis <i>stepping up</i> e <i>stepping down</i>)	118

CONSIDERAÇÕES INICIAIS

O foco desta Dissertação de Mestrado está voltado às estratégias dativas de complementação verbal de segunda pessoa do singular (doravante 2SG), à luz dos princípios teórico-metodológicos da Sociolinguística Laboviana (LABOV, 1972; LABOV, 1994), produtivas em cartas brasileiras do século XXI (2012). Tendo em vista o fato de o espaço funcional de *sujeito* já ter se mostrado como o primeiro contexto de acesso das formas gramaticalizadas *você* e *a gente* (RUMEU, 2013; SOUZA, 2012; LOPES & CAVALCANTE, 2011; LOPES & VIANNA, 2012), ao se implementarem no sistema pronominal do Português Brasileiro (doravante PB), buscamos não só descrever panoramicamente a produtividade (em termos percentuais) das estratégias pronominais dativas de 2SG, mas também identificar os condicionamentos linguísticos e/ou extralinguísticos (em termos probabilísticos) propulsores da regra variável em cena (*te versus lhe, a ti, para ti, a você, para você, você, zero*).

Tendo em vista a implementação do *você* no sistema pronominal ter se dado em distintos ritmos pelos contextos sintáticos (nominativo, complementação e adjunção), cf. Cardoso (2017), Cruz (2017), Lopes & Cavalcante (2011), Rumeu (2014), Rumeu & Oliveira (2016), a descrição panorâmica das formas pronominais dativas de 2SG se justifica também por poder contribuir para a reconstrução da história do sistema pronominal do PB (LOPES & CAVALCANTE, 2011; LOPES *et alii*, 2018; RUMEU, 2020), fornecendo mais evidências da variação no PB escrito dos anos 2000 (2012). Nesse sentido, a hipótese é a de que o *te-dativo* assumira uma alta produtividade no contexto de *você-sujeito* nas cartas produzidas por redatores que nasceram e, necessariamente, vivem no espaço geográfico de Minas Gerais (Cartas ao Papai Noel doravante PN), que é o de prevalência do *você-sujeito* (SCHERRE *et alii*, 2015). Conjecturamos ainda que as estratégias dativas do paradigma de *você-sujeito* (*a/para você*) encaminhem-se com maiores índices de produtividade nas cartas de *você-sujeito*, como já discutido por Lopes & Cavalcante (2011), Lopes *et alii* (2018), Rumeu (2020) em relação às cartas brasileiras de sincronias passadas (séculos XIX e XX).

Considerando as distintas estratégias pronominais dativas de 2SG (*te, lhe, você, a você, para você, a ti, para ti, zero*), levantam-se as seguintes questões norteadoras desta pesquisa.

- I. Quais seriam as estratégias dativas de 2SG selecionadas por escreventes brasileiros para a referência ao PN?
- II. Qual(is) seria(m) a(s) forma(s) pronominal(is) dativa(s) do paradigma de *tu* (*te, a ti, para ti*) que se mostra(m) suscetível(is) à convivência com o *você-sujeito*?
- III. Quais seriam os fatores linguísticos e/ou extralinguísticos que atuariam nesse processo de variação das formas dativas de complementação verbal de 2SG? Quais fatores estruturais tendem a refrear ou acelerar a variação em questão?

Considerando o tema (foco), as hipóteses e as questões motivadoras deste trabalho dissertativo, passamos, de (a) a (h), à ilustração das estratégias dativas de 2SG¹ monitoradas (*te, lhe, a ti, para ti, a você, para você, você e zero*) nesta análise.

- (a) “[...] Escrevo [*para você*]_{DAT} para [*te*]_{DAT} pedir uma cesta [...]” (C73)
- (b) “Meu querido papai noel venho através desta [*lhe*]_{DAT} pedir que realize meu desejo de natal. [...]” (C959)
- (c) “Gostaria de [*o*]_{DAT} pedir um trator de brinquedo para brincar com meus priminhos. [...]” (C959)
- (d) “[...] Eu gostaria de pedir [*você*]_{DAT} papai noel um tablet [...]” (C098)
- (e) “[...] A estrada da tua Felicidade não parte das pessoas e das coisas para chegar [*a ti*]_{DAT} [...]” (C834)”
- (f) “Papai Noel eu queria pedir [*o*]_{DAT} para você mim dar um Quiti de maquiagem [...]” (C85)
- (g) “Oi papai noel como vai você meu nome e [...] e estou [*lhe*]_{DAT} escrevendo para pedir [*a você*]_{DAT} que possa mim ajudar neste natal [...]” (C482)

¹ Optamos, neste trabalho, por expor o complemento dativo de 2SG entre colchetes e por sublinhar o predicador verbal responsável pela sua projeção.

(h) “Papai Noel Gosto de ti, [...] por isso eu estou dizendo [*para ti*]_{DAT} que gosto muito de presentes... te adoro” (C961)

Esta dissertação está organizada em quatro capítulos. Nas considerações iniciais, expomos, em linhas gerais, o tema de trabalho, a justificativa, as hipóteses, as questões motivadoras, a amostra de cartas brasileiras e a orientação teórico-metodológica. O capítulo 1 está voltado à revisão histórico-descritiva, assumindo como ponto de partida a posição da tradição gramatical em relação às estratégias dativas de 2SG até atingirmos os resultados de algumas pesquisas linguísticas baseadas em amostras de sincronias passadas (língua escrita) e recentes (língua falada). O capítulo 2 encaminha os principais princípios e orientações da Sociolinguística Laboviana, passando não só pela apresentação das cartas brasileiras em análise, pela exposição da linguagem RBrul na interface do RStudio em termos do tratamento metodológico concedido aos dados, mas também pela descrição-analítica das variáveis dependente eneária e das independentes (linguísticas e extralinguísticas). Destinamos os capítulos 3 e 4 à descrição-analítica dos resultados quantitativos desta dissertação. No capítulo 3, temos as estratégias dativas de 2SG percentualmente distribuídas em relação às variáveis linguísticas e extralinguísticas. No capítulo 4, expomos os resultados estatisticamente relevantes em termos probabilísticos, considerando o *te*-dativo como variável dependente. Nas considerações finais, sistematizamos respostas para as principais questões voltadas ao potencial variável das formas pronominais dativas de 2SG que impulsionam esta análise.

CAPÍTULO 1 REVISÃO HISTÓRICO-DESCRIPTIVA

Dedicamos este primeiro capítulo a uma revisão histórico-descritiva do dativo de 2SG. Partimos dos estudos históricos sobre a origem do caso dativo (subseção 1.1), passamos pela tradição gramatical (subseção 1.2) até alcançarmos as pesquisas linguísticas (subseção 1.3) baseadas em amostras de sincronias passadas (língua escrita) e recentes (língua falada).

1.1 Os complementos dativos de 2SG: a origem latina

O sistema clássico latino de flexão nominal era marcado morfologicamente não só pelos traços de *gênero* e *número*, como no português, mas também pelo traço de *caso*. O *caso*, que é a “forma tomada por uma palavra declinável para indicar precisamente qual a função que desempenha na frase” (FARIA, 1958, p. 59), deixa-se expressar através do *nominativo*, *vocativo*, *acusativo*, *genitivo*, *dativo* e *ablativo*. Considerando o fato de o objeto deste estudo ser o *dativo*, voltamos o nosso foco, neste capítulo de revisão do tema, à descrição desse caso gramatical. Em termos semânticos, na acepção de Faria (1958, p. 61) temos o dativo como o “caso da atribuição, indicando a pessoa ou coisa a quem um objeto é destinado e, daí, no interesse de quem se faz uma coisa”. O autor assume que o *dativo* é mais empregado para indicação de objeto e *complemento indireto* de verbos, substantivos e adjetivos. O autor apresenta o *dativo* latino a partir dos seus nove tipos: (1) *dativo de objeto indireto*, (2) *dativo de contato/aproximação*, (3) *dativo de interesse*, (4) *dativo de posse*, (5) *dativo de referência*, (6) *dativo ético*, (7) *dativo de agente/obrigação*, (8) *dativo de destinação* e (9) *dativo de direção*.

O *dativo de objeto indireto* é o objeto indireto ou complemento dos verbos transitivos ou dos verbos intransitivos relativos, cf. descrito por Faria (1958, p. 349) e exemplificado em (1).

(1) Hirtio cenam dedi (Cíc., Fam., 9,20,2)

“dei uma ceia [*a Hirtio*]_{DAT}”

Faria (1958, p. 350) admite que os *dativos de contato* ou *aproximação* são projetados por verbos que dão a ideia de aproximação como se observa com os verbos “misturar” (*misceo*), “ligar” (*iungo*) e “prender” (*haereo*). De (2) a (4), ilustramos evidências do *dativo de contato ou aproximação*.

(2) *Fletūmque* [cruōri]_{DAT} miscūit (Ov., Met., 4,140-1)
 “misturou o pranto [*ao sangue*]_{DAT}”

(3) *dextrae iungere* [dextram]_{DAT} (Verb., En., 1,408)
 “unir a minha destra [*à tua destra*]_{DAT}”

(4) *haerentem* [capiti]_{DAT} cum multa laude coronam (Hor., Sát., 1, 10, 49)
 “a coroa que com grande glória prende [*à sua cabeça*]_{DAT}”

Quanto ao *dativo de interesse*, este também indica a ideia de atribuição, mas se referindo a uma pessoa “interessada no fato enunciado pela oração, em cujo proveito ou prejuízo se faz uma coisa” (FARIA, 1958, p. 351). Em (5), ilustramos o *dativo de interesse*.

(5) non [tibi]_{DAT} sed *patriae natus* (Cíc., Mur., 83)
 “nascido não [*para ti*]_{DAT}, mas para a pátria”

O *dativo de posse*, que é, segundo Faria (1958, p. 352), uma extensão do dativo de interesse, serve para indicar a pessoa possuidora de uma coisa, “significando propriamente que algo existe para ou em proveito de alguém”, como observamos em (6).

(6) nihil est [mihi cum]_{DAT} eo (Cíc., Phil., 2,77)
 “nada tenho [*com ele*]_{DAT}”

O *dativo de referência* serve para indicar uma pessoa a quem a frase como um todo se refere, cf. observamos em (7), seguindo a proposta de Faria (1958, p. 352).

(7) non [tibi]_{DAT} ego exēmpli satis sum? (Ter., Heaut., 920)
 “não sou [*para ti*]_{DAT} um exemplo bastante?”

O *dativo ético* mostra-se muito produtivo nos diálogos, nas apóstrofes ou admoestações e nas interpelações, consubstanciado nos pronomes pessoais de primeira e segunda pessoas do discurso. Em (8), ilustramos, com base em Faria (1958, p. 352), o *dativo ético*.

(8) caue [mihi]_{DAT} mendaci quicquam (Plaut., As, 30)

“não [me]_{DAT} venhas com qualquer mentira!”

Faria (1958, p. 353) esclarece que o *dativo de agente* ou de *obrigação* serve para indicar “que existe uma obrigação para tal pessoa, sendo comumente usado com o gerundivo”. O autor descreve que seu uso pode se dar também com formas verbais perfectivos de voz passiva, quando indica “que para uma pessoa tal ação é considerada como um fato acabado ou consumado” (FARIA, 1958, p. 353). Em (9) e (10), ilustramos, à luz de Faria (1958), evidências do *dativo de agente* ou de *obrigação* também com formas pronominais de 1ª e 2ª pessoas do discurso.

(9) ab domo abeūdo est [mihi] (Plaut., Aul., 105)

“devo afastar-[me]_{DAT} da casa”

(10) [tibi]_{DAT} cauendum censeo (Plaut., Cas., 411)

“penso que [te]_{DAT} deves acautelar”

O *dativo de destinação* é mais comum em verbos de movimento tais como “ir” (*eo*), “enviar” (*mitto*), “conduzir” (*duco*), “trazer” (*adūco*), “vir” (*uenio*), indicando a que se destina determinada coisa (FARIA, 1958, p. 353) e acompanhando-se, comumente do *dativo de objeto indireto* ou *dativo de interesse*. Em (11) e (12), ilustramos o *dativo de destinação* projetado pelos verbos *ir* e *vir*, respectivamente.

(11) (Veientes) ... pars eunt [Sabinis subsidio]_{DAT}” (T. Lív., 2, 53, 2)

“parte dos Veientes vão [em socorro dos sabinos]_{DAT}”

(12) Cum omnibus copiis auxilio Nervii uenirent (Cés., B. Gal., 2,29,1)

“como viessem vindo [em auxílio dos nêrvios] com todas as suas tropas”

Ainda no âmbito do dativo de destinação, temos a construção do verbo *sum* com dois dativos. Trata-se do *dativo de interesse* (*dativus commodi* ou *incommodi*) e do *dativo aplicado* à indicação de efeito ou consequência, cf. Faria (1958, p. 354), cf. exemplificamos em (13) e (14).

(13) [agmen]_{DAT} hostium claudēbant et nouissimis praesidio erant (Cés., B. Gal., 1,25,6)
 “fechavam a marcha protegendo-a do inimigo e serviam de proteção [para a retaguarda]_{DAT}”

(14) [his difficultatibus duae res]_{DAT} erant subsidio (Cés., B. Gal., 2,20,3)
 “[a estas dificuldades duas coisas]_{DAT} serviam de auxílio”

O *dativo de direção* consiste, segundo informa Faria (1958, p. 354), no uso do dativo com verbos de movimento no lugar do acusativo precedido das preposições *ad* ou *in*. Expomos, em (15) e (16), os seguintes dados de *dativo de direção*.

(15) it clamor [caelo]_{DAT} (Verg., En., 5,451)
 “vai um clamor [ao céu]_{DAT}”

(16) [latēri]_{DAT} capūlo tenus abīdit ense (Verg., Em., 2,553)
 “craovou-lhe [no flanco]_{DAT} a espada até o copo”

Uma vez apresentada, à luz de Faria (1958), a taxionomia do dativo evidenciamos tratar-se de um caso latino cujas especificidades sintáticas e semânticas distintas aproximam-se em relação à sua função principal que é a de “polo para o qual a ação ou o processo referido pelo predicado se orienta”, cf. Van Hoecke (*apud* OLIVEIRA, 2014, p. 25). Em termos formais, observamos que as formas dativas latinas *mihi*, *tibi*, *sibi*, *tū*, *tē* e *tibī* resistem através das formas pronominais *mim*, *ti*, *si*, *tu*, *te* e *ti* no português como evidência da sintaxe de casos latina, conforme discutido por Oliveira (2014, p. 27-28), ainda que movidas pelo inovadorismo da combinação do complemento dativo à preposição “a”/“para” nos sintagmas preposicionados *a ti/para ti* (paradigma etimológico de *tu*) e *a você/para você* (paradigma inovador de *você*).

1.2 Os complementos dativos de 2SG: a visão da gramática tradicional

Para a descrição das formas pronominais dativas de 2SG no PB na perspectiva da tradição gramatical, passamos por Bechara (2009 [1961]), Rocha Lima (2011 [1968]) e Cunha e Cintra (2009 [1985]).

Na perspectiva da tradição gramatical, o “Pronome”, de um modo geral, é concebido como a classe gramatical responsável pela substituição dos nomes na referência às pessoas do discurso. Bechara entende que esta categoria gramatical possui um número finito de unidades que remetem “a um significado léxico pela situação ou por outras palavras do contexto” (2009, p. 138). Rocha Lima (2011 [1968]), ao apresentar a categoria Pronome, o faz a partir da retomada do conceito de Said Ali, que caracteriza como elemento destituído de conteúdo semântico próprio e que denotam ou se referem a um ente, acentuando, nesse sentido, a sua potencialidade como indicador da pessoa do discurso (SAID ALI *apud* ROCHA LIMA, 2011 [1968], p. 156). Ainda sobre os Pronomes na perspectiva da tradição gramatical, observamos que Cunha & Cintra (2009 [1985], p. 289) conceituam esta como a classe responsável por assumir as funções típicas dos elementos nominais.

Especificamente em relação aos pronomes pessoais, os gramáticos caracterizam-nos de acordo com a pessoa do discurso. Os pronomes do caso *reto* (nominativo) assumem a função de sujeito (*eu* (1SG), *nós* (1PL), *tu* (2SG), *vós* (2PL), *ele(s)*, *elas(s)* (3SG, 3PL) e os pronomes *oblíquos* assumem a função de objeto (direto (acusativo) ou indireto (dativo e oblíquos), podendo ser *átonos* ou *tônicos*. Em sua forma átona, os pronomes “são partículas inacentuadas, que se colocam antes ou depois do verbo, como se fosse uma sílaba a mais desse verbo”, cf. Rocha Lima (2011 [1968], p. 167). As formas tônicas, por outro lado, são regidas por preposições. Apresentamos, a seguir, o quadro síntese dos pronomes pessoais do português na perspectiva de Rocha Lima (2011 [1968]) que também é acompanhada por Bechara (2009 [1961]) e por Cunha & Cintra (2009 [1985]).

Quadro 1 Os pronomes pessoais da língua portuguesa

		PRONOMES PESSOAIS DO CASO RETO	PRONOMES PESSOAIS DO CASO OBLÍQUO	
			ÁTONOS	TÔNICOS
Singular	1. ^a pessoa	eu	me	mim, comigo
	2. ^a pessoa	tu	te	ti, contigo
	3. ^a pessoa	ele, ela	o, a, lhe, se	ele, ela, si
Plural	1. ^a pessoa	nós	nos	nós, conosco
	2. ^a pessoa	vós	vos	vós, convosco
	3. ^a pessoa	eles, elas	os, as, lhes, se	eles, elas, si

Fonte: Rocha Lima (2011 [1968], p. 156-158).

Nenhum dos gramáticos em discussão inclui as formas pronominais de referência à 2.^a pessoa do singular (*você*) e do plural (*vocês*) no quadro de pronomes pessoais, evidenciando o conservadorismo dos gramáticos. Por outro lado, Rocha Lima (2011[1968], p. 158) destaca que “há alguns pronomes de segunda pessoa que requerem para o verbo as terminações de terceira”, apresentando as formas *você* e *vocês*, mas expondo, por outro lado, tais formas como pronomes de tratamento familiar. Cunha e Cintra (2009 [1985], p. 308) assumem que “as formas *você* (no Brasil) e *o senhor, a senhora* (tanto em Portugal como no Brasil) estendem-se também às funções de objeto (direto ou indireto) em substituição aos clíticos átonos *o, a e lhe*”, como observamos nos seguintes dados literários expostos em (17) e (18).

(17) Há uma hora estou esperando [*você*]_{ACUS} sozinha neste escritório (C. Dos Anjos, DR, 32)

(18) Devo [*a você*]_{DAT} e ao doutor Rodrigo. (J. Amado, MM, 229)

Em síntese, no que se refere especificamente à função de objeto indireto (complemento dativo), temos em cena, na perspectiva da tradição gramatical, a forma pronominal dativa como um complemento verbal que encontra espelhamento no clítico de 3.^a pessoa *lhe* (*lhes*), bloqueando, em termos sintáticos, o licenciamento da estrutura passiva. Em (19) e (20), apresentamos evidências do dativo, expostas por Rocha Lima (2011 [1968], p. 307).

(19) a. Dar esmola [a um mendigo]_{DAT}

b. Dar-[*lhe*]_{DAT} esmola.

(20) a. Escrever [a um amigo]_{DAT}

b. Escrever-[*lhe*]_{DAT}

À luz da perspectiva da gramática tradicional, é importante preservarmos a simetria entre os paradigmas pronominais das formas *tu* e *você*. Isso quer dizer que para as formas do paradigma de *tu* temos tão somente como possibilidades para a função dativa o clítico *te* e os SPreps *a ti* e *para ti*, bem como para o paradigma de *você*, há os clíticos *o*, *a*, *lhe* (sempre na referência à 2SG) e os SPreps *a você* e *para você*. Uma vez exposta a perspectiva da tradição gramatical em relação às formas pronominais dativas de 2SG, passamos a alguns resultados das pesquisas linguísticas acerca dos usos concretos dos pronomes relacionados aos paradigmas de *tu* e de *você* em sincronias passadas e recentes do PB.

1.3 Os complementos dativos de 2SG: a visão de alguns estudos linguísticos

Após apresentarmos a origem latina do *dativo* e a visão da GT em relação ao *dativo* no PB, passamos a apresentar algumas pesquisas científicas que tomaram o *dativo de 2SG* como objeto de estudo. Nosso primeiro olhar será para as pesquisas que se valeram de *corpora* históricos, ou seja, de sincronias passadas; e depois iremos reportar resultados de pesquisas que adotaram *corpora* de sincronias mais recentes.

1.3.1 Os complementos dativos de 2SG em sincronias passadas: os séculos XIX e XX em cena

Considerando que o principal objetivo desta dissertação é o mapeamento das formas pronominais dativas de 2SG na produção escrita brasileira recente (anos 2000), assumimos também como ponto de partida alguns resultados de estudos linguísticos embasados na produção escrita de redatores nascidos e/ou ambientados no sudeste (LOPES *et alii* 2013; OLIVEIRA 2014; RUMEU 2014; CARDOSO 2017, RUMEU 2020), nordeste (LOPES *et alii* 2018) e sul (LOPES *et alii* 2018) do Brasil dos séculos XIX e XX.

Lopes *et alii* (2013) propõem uma investigação acerca do clítico *te* no PB à luz de uma perspectiva funcionalista da mudança linguística pela via da gramaticalização (BYBEE 2003; COMPANY COMPANY, 2008), assumindo que “existe uma relação direta entre frequência de uso de formas ou construções e gramaticalização, por assumir que a gramática codifica o que os falantes mais empregam” (LOPES *et alii*, 2013, p. 392). A hipótese é a de que o clítico *te* estaria em um processo de gramaticalização, o que o evidenciaria como um afixo (clítico > afixo) em função da “opacidade semântica” adquirida em virtude também da sua altíssima produtividade no PB. Com base na análise de cartas pessoais, de peças teatrais dos séculos XIX e XX e de roteiros de cinema dos séculos XX e XXI constataram os autores que a frequência evidenciou não só a afixação do *te* como marcador de *pessoa*, mas também a sua opacidade, responsável também pela condução de estruturas de redobro voltadas para a referência à 2SG (“Eu *te*; falei *pra* *você*”).

Oliveira (2014) voltou-se, em sua Dissertação de Mestrado, à alternância entre as estratégias pronominais de complementação de 2SG na posição dativa (*te*, *lhe*, *a ti*, *para ti*, *a você*, *para você* e *nulo*). Embasado na análise de trezentos e dezoito cartas históricas de cariocas e fluminenses (1889 a 1980), o autor voltou-se ainda para a combinação do *te* com o *você-sujeito* (“*Você*; leu o livro que eu *te*; dei?”), tendo sempre em vista a fragilidade da noção de uniformidade tratamental tão cara à tradição gramatical e tão ineficaz em termos de produtividade no PB atual (BRITO, 2001). À luz dos pressupostos teórico-metodológicos da sociolinguística histórica (CONDE SILVESTRE, 2007; HERNÁNDEZ-CAMPOY & CONDE SILVESTRE, 2012) e do programa GOLDVARB X para os controles estatístico e probabilístico dos dados históricos de dativo de 2SG, identificou o autor, em termos gerais,

o *te* como a estratégia mais produtiva (57,2%), seguida pelo *nulo* (22,3%) e pelo *lhe* (11,3%), cabendo às demais estratégias dativas (*a ti*, *para ti*, *a você* e *para você*) frequências de uso inferiores a 5%. O *te-dativo* prevalece nas cartas oitocentistas e novecentistas distribuindo-se nas cartas de *tu-sujeito* (78,5%), de *você-sujeito* (30,9%) e nas cartas mistas (70,2%), o que permite interpretá-la como a estratégia mais produtiva independente do pronome-sujeito de 2SG. O clítico *lhe* evidenciou formalidade, os sintagmas preposicionados apresentaram baixas frequências de uso, considerando que as formas *a ti* e *para ti* se mostraram em substituição às formas *a você* e *para você* no decorrer do tempo das cartas analisadas (séculos XIX e XX). Em termos probabilísticos, Oliveira testou a oposição *te versus zero* (nulo) que expôs a amostra (cartas da Família Cruz = .860; Cartas da Família Penna = .694; Cartas da Família Pedreira-Ferraz = .585), o pronome-sujeito de 2SG (as cartas mistas (tu/você) = .595; as cartas de *tu-sujeito* = .525), a estrutura argumental do predicador (verbos de 2 lugares = .948) e o subgênero da carta particular (pessoal = .679 e amorosa = .512) como contextos propulsores do *te-dativo* (valor de aplicação). A oposição *te versus lhe* evidenciou ao autor o pronome sujeito de 2SG (*tu-sujeito* = .942), o período (1931-1955 = .932 e 1906-1930 = .683) e o subgênero da carta particular (carta amorosa = .853) como contextos condicionantes do *te-dativo* (valor de aplicação). Ainda em busca da apreensão da dinâmica da variação, testou o autor a regra variável *te versus SPreps*, o que evidenciou a *forma do objeto direto* (oracional = .874), seção da carta (captação de benevolência = .692 e núcleo da carta = .563), o pronome-sujeito de 2SG (*tu* = .733; *tu/você* = .500) e a categoria semântica do predicador (T. verbal/perceptual = .656; Mov. psicológico = .703; Mov. físico = .554) como fatores propulsores do *te-dativo* (valor de aplicação) nas cartas oitocentistas e novecentistas analisadas.

Deslocando o nosso foco para análises embasadas na produção escrita de brasileiros nascidos e/ou residentes no espaço geográfico de Minas Gerais, trazemos à cena os resultados de Rumeu (2014). Voltada também para as estratégias de complementação verbal, à luz dos pressupostos teórico-metodológicos da Sociolinguística (LABOV, 1994) e embasada na análise de cento e trinta e seis cartas mineiras (1900-1990), a autora identificou o contexto de dativo de 2SG como o mais produtivo (52%) em relação às formas do paradigma de *você*. Mostrou-se significativa a projeção de estruturas dativas com verbos ditransitivos (SU V OD OI)) como um contexto que pareceu promissor em relação à propagação do “você não-

sujeito”. Em síntese, autora assume, em termos de resultados gerais voltados ao dativo de 2SG, que o “você” parece ter adentrado, cf. Rumeu (2013), o sistema pronominal do PB, entre 1920 e 1945, atingindo, entre 1930 e 1969, as relações dativas no contexto de *você-sujeito* das cartas mineiras novecentistas analisadas.

Cardoso (2017) desenvolve, em sua Dissertação de Mestrado, um estudo panorâmico embasado nos fundamentos da sociolinguística histórica de Conde Silvestre (2007) e Hernández-Campoy e Conde Silvestre (2012) acerca das estratégias dativas de 2SG (*te, a ti, lhe, a você, para você, para ti, Ø*) produtivas em cartas pessoais de redatores nascidos e/ou residentes no espaço mineiro, entre os séculos XIX e XX. Dentre as estratégias dativas levantadas (*te, a ti, lhe, a você, para você, Ø*) a autora identificou o *lhe*, representando 43% (251/582), o *te*, em 27% (155/582) e o *zero-dativo*, 23% (134/582) dos dados. Ao correlacionar o dativo ao contexto nominativo de 2SG, Cardoso (2017), Cardoso verificou que o *lhe* foi muito produtivo para as cartas de *você sujeito*, com frequência de 62% (149/239). Este pronome também se mostrou muito produtivo nas cartas de tratamento alternativo (em que o escrevente se absteve de usar tanto o *você-sujeito*, quanto o *tu-sujeito*), também com 62% de frequência (99/160). Quanto ao *te*, a autora observou a sua predominância nas cartas de *tu-sujeito*, em 77% (120/155) e nas cartas mistas (com uso alternado entre *você-sujeito* e *tu-sujeito*), em 52% dos dados (14/27). De um modo geral, os resultados apontam para uma uniformidade formal em relação à combinação entre as formas de *você-lhe* e *tu-te* no contexto de complementação dativa. Contudo, a autora também constatou a produtividade do *zero-dativo* em todos os contextos de referência ao sujeito de 2SG (cartas de *tu-sujeito*, cartas de *você-sujeito*, cartas de *tu/você* (mistas)) como uma evidência de inovadorismo também nas cartas mineiras (CARDOSO, 2017, p. 109). A autora interpreta a questão do espriamento do *te-dativo* por todos os subgêneros das cartas mineiras (amorosas, familiares e de amizade) como uma evidência de uma “forma neutra” como também constatado por Oliveira (2014) em relação à alta produtividade do *te-dativo* nas cartas cariocas.

No estudo de Rumeu (2020), temos em cena a análise das estratégias de complementação verbal de 2SG embasada em um conjunto de cartas pessoais trocadas entre missivistas cultos (em sua maioria) que nasceram e/ou conduziram suas trajetórias de vida no espaço geográfico de Minas Gerais, no período de 1860 e 1989. A autora percorreu os três

contextos de complementação verbal (*acusativo, dativo e oblíquo*) para investigar se existe e como se manifesta a correlação entre as formas de *tu* e *você* na função de sujeito e as formas de complementação verbal de 2SG em sincronias passadas (séculos XIX e XX). Optamos por trazer a esta revisão do tema tão somente os seus resultados relacionados ao contexto de dativo de 2SG. Especificamente em relação ao contexto *dativo*, constata a autora que as formas mais produtivas no *corpus* são o *te* e o *lhe*. Ao correlacionar o contexto de dativo ao de nominativo (*tu-sujeito, você-sujeito* e alternância entre *tu* e *você*), Rumeu identificou que a estratégia de complementação mais frequente foi o *te*, em 82% dos dados, seguida pelo *zero-dativo*, representando 9%, e o *lhe*, em 6% dos dados. No contexto das cartas de *você-sujeito*, a estratégia mais frequente é o *lhe*, com 75% dos dados, seguido pela forma preposicionada *a você*, com 8,3% e o *zero-dativo*, em 5,9% dos dados. Ainda quanto ao sistema do *você-sujeito*, verificou a autora que as formas *te* e *para você* igualam-se com frequências de 5,4%. Já no sistema de cartas que alternam entre *tu-sujeito* e *você-sujeito*, Rumeu evidenciou uma maior produtividade do *te* (63%) e do *zero-dativo* (30%), os demais dados foram de *para você* (7%), as demais formas, incluindo o *lhe*, não se mostraram produtivas. Nesse sentido, a autora retoma a ideia de que ²“a coexistência de formas do paradigma de *tu* (*te*) e formas do paradigma de *você* (*você, lhe, o/a*) [...] é outra indicação da fusão dos paradigmas de segunda e terceira pessoas desde a inserção do *você* no sistema pronominal do PB” (RUMEU, 2020, p. 243). Em síntese, Rumeu (2020) admite que o pronome-complemento em função dativa acompanha o paradigma da forma nominativa também de 2SG, uma vez que temos a prevalência do *te* (paradigma de *tu*) nas cartas de *tu-sujeito* e nas cartas mistas (*tu/você*) e do *lhe* (paradigma de *você*) nas cartas do *você-sujeito*, como já discutido por Lopes & Cavalcante (2011), Oliveira (2014) e Cardoso (2017).

Lopes *et alii* (2018) estudam os complementos verbais (*acusativo, dativo e oblíquo*) em cartas pessoais dos séculos XIX e XX, embasados em cartas produzidas por redatores brasileiros nascidos e/ou residentes (ainda que por algum período de suas vidas) nos espaços geográficos do Rio de Janeiro, Minas Gerais, São Paulo, Santa Catarina, Bahia, Pernambuco e Rio Grande do Norte. Tendo em vista o escopo deste trabalho, limitamo-nos a apresentar

² “The coexistence of forms of the *tu* (*te*) paradigm and forms of the *você* (*você, lhe* and *o/a*) paradigm, as in example (36), is yet another indication of the fusion of second and third person paradigms since the emergence of *você* in the BP pronominal system.” (RUMEU, 2020, p. 243)

os resultados de Lopes *et alii* (2018) do complemento *dativo* de 2SG. De um modo geral, observam os autores a prevalência do *te* nas cartas do Rio de Janeiro (58%, 464/811), São Paulo (59%, 53/90) e Santa Catarina (70%, 70/117), bem como a prevalência do *lhe* em Minas Gerais (46%, 63/136), Bahia (75%, 188/251), Pernambuco (73%, 158/214) e Rio Grande do Norte (48%, 93/192). Em termos gerais, tendo em vista os mil oitocentos e onze dados levantados em todos os *corpora* de missivas históricas brasileiras (1.811oco), temos, de um modo geral, o *te*, em 41% dos dados (748oco), o *lhe*, em 35% (638oco), o *zero-dativo*, em 13% (241oco), os Spreps *a/para você*, em 7% (119oco), os Spreps *a/para ti*, em 3% (50oco) e *prep.+você*, em 1% (15oco). É importante atentarmos ao fato de o *zero-dativo* ter se mostrado produtivo nas cartas cariocas (22%, 181/811), mineiras (14%, 19/136) e baianas (16%, 41/251).

Voltando o foco para a Região Sudeste, Lopes *et alii* (2018) correlacionam as estratégias dativas de complementação verbal aos pronomes-sujeito nas cartas cariocas, mineiras e paulistas. O *te* é a forma mais profícua nas cartas de *tu-exclusivo* em todas as localidades, enquanto, nas cartas de *você-sujeito*, à exceção de Minas Gerais, há uma acirrada competição entre *te* e *lhe*. Nas cartas cariocas de *você-sujeito*, as formas *te* (31%, 85oco), *zero-dativo* (29%, 80oco) e *lhe* (27%, 75oco) concorrem acirradamente, ao passo que, nas cartas mineiras, prevalece o *lhe*, em 58% dos dados (62oco) e o *te*, que se mostra em 18% dos dados (19 oco). Nas cartas paulistas, *te* e *lhe* empataram em 44% dos dados (7oco). Nas cartas mistas (*tu/você*), os autores evidenciam a prevalência do *te* em relação à tão somente 3% dos dados (7oco) do *lhe*, nas cartas cariocas. Os Spreps *a/para você* mostram-se em todos os subsistemas de pronomes-sujeito de 2SG (cartas de *tu-sujeito*, cartas de *você-sujeito* e cartas mistas (*tu/você*)), exceto nas cartas mineiras de *tu-sujeito*, ao passo que as formas *a/para ti* se deixam evidenciar somente nas cartas cariocas, mas em todos os subsistemas de pronomes-sujeito de 2SG. Lopes *et alii* (2018) destacam, em síntese, a prevalência do *te*, nas cartas paulistas, e a predominância do *lhe*, nas cartas mineiras de *você-sujeito*.

Ao correlacionar as estratégias dativas ao contexto de sujeito de 2SG na produção escrita no Nordeste de sincronias passadas (séculos XIX e XX), Lopes *et alii* (2018) revelam a prevalência absoluta do *te* nas cartas de *tu-sujeito* da Bahia (07/07oco), do Rio Grande do Norte (17/17oco) e a sua prevalência nas cartas de Pernambuco, com 22 ocorrências de *te* contra 07 ocorrências dos Spreps *a/para ti*, em 76% e 24% dos dados, respectivamente. Nas

cartas de *você-sujeito*, os resultados apontam a prevalência do *lhe*, em 80% dos dados da Bahia (164/206), em 90% dos dados de Pernambuco (134/149) e em 86% dos dados do Rio Grande do Norte (42/29). Nas cartas baianas de *você-sujeito*, temos o *te* em tão somente 2% dos dados (4/206). Nestas, em todas as localidades a forma preposicionada *a/para você* se fez evidenciar, mas em níveis variados: 3% (07/206) na Bahia, 10% (13/149) em Pernambuco e 14% (07/49) no Rio Grande do Norte. O *zero-dativo* só foi identificado nas cartas de *você-sujeito* baianas (31/206, 15%). As cartas de alternância entre *tu* e *você* (mistas) evidenciam uma variação mais diversa e acirrada como observamos na Bahia em relação aos resultados que apontam para uma prevalência do *te*, em 57% dos dados (04/07), seguido pelo *lhe*, em 28% (14/20) e o *zero-dativo*, em apenas um ocorrência (14%). Para as cartas mistas de Pernambuco, temos a prevalência do *lhe*, em 70% (14/20) dos dados contra o *te*, em 25% (05/20) e de *a/para você*, em apenas uma ocorrência (5%). Já no Rio Grande do Norte, tem-se uma acirrada competição entre o *te*, em 33% (43/123), e o *lhe*, em 39% (51/123), seguidos pelo *a/para você*, com 22% (18/123). Os autores destacam o predomínio do *te*, nas cartas de *tu-sujeito*, e do *lhe*, nas cartas de *você-sujeito* (LOPES *et alii*, 2018, p. 167).

Para a Região Sul, os autores, ao correlacionarem as formas *dativas* às formas *nominativas* de 2SG nas cartas de Santa Catarina, evidenciam a predominância de formas do paradigma de *tu* em cartas de *tu-sujeito*, demonstrando a prevalência do *te* em 79% (20/63) dos dados, seguido pela forma *a/para ti* (21%, 13/63). Nas cartas de *você-exclusivo*, temos o predomínio do *lhe*, em 54% dos dados (08/15), de *a/para você*, em 26% (04/15) e *prep+você*, em 20% (03/15). Já nas cartas mistas (*tu/você*), o *te* predomina, em 34% (33/39) dos dados, seguido pelo *lhe*, em 10% (03/39), e por uma única ocorrência, em 2% dos dados, das formas preposicionadas *a/para ti* (01/39), *a/para você* (01/39) e *prep.+você* (1/39), além de não ter havido evidências do *zero-dativo* na amostra de cartas catarinenses.

Uma vez expostos alguns resultados de análises voltadas ao mapeamento das estratégias pronominais dativas de 2SG em cartas de sincronias passadas (LOPES *et alii*, 2013; RUMEU, 2014; OLIVEIRA, 2014; CARDOSO, 2017; RUMEU, 2020 e LOPES *et alii*, 2018), passamos a uma breve síntese de alguns resultados de análises voltadas ao dativo de 2SG em sincronias mais recentes a partir de amostras de língua falada.

1.3.2 Os complementos dativos de 2SG em sincronias recentes: algumas análises embasadas na fala brasileira

Esta subseção de um capítulo de Revisão do Tema está dedicada à descrição dos resultados das pesquisas linguísticas embasadas em amostras de língua falada distribuídas em relação às regiões brasileiras, quais sejam: nordeste (ALMEIDA, 2014; SILVA, 2016; PESSOA, 2017) e sul (DALTO, 2002) do Brasil.

1.3.2.1 Os complementos dativos de 2SG no nordeste brasileiro

Para o nordeste brasileiro, trazemos à cena os estudos de Almeida (2014) e Silva (2016), que embasados estão em amostras da fala baiana, e de Pessoa (2017), que tem como foco a fala piauiense.

Almeida (2014), em sua Tese de Doutorado, voltou-se à expressão variável do acusativo e dativo na fala de Santo Antônio de Jesus (Recôncavo Baiano). A pesquisadora também levou em consideração os pressupostos teórico-metodológicos da Sociolinguística Laboviana para estudar o uso variável das formas pronominais dativas de referência à 2SG. A submissão dos dados ao Pacote de Programas GoldVarb revelou à autora a seguinte distribuição: o *lhe*, em 36,2% (105/290), o *te*, em 34,5% (100/290), o *zero-dativo*, em 23,4% (68/290), o sintagma preposicionado com a forma *você*, em 4,5%, e sintagmas preposicionados com as formas *senhor/senhora*, em 1,4% dos dados. Considerando que os maiores índices de variação mais acirrada incidiram na dinâmica *lhe* versus *te*, a autora optou por submeter tal regra variável à análise probabilística a partir do *lhe* (variável dependente) independentemente da função acusativa ou dativa, incluindo, pois, todos os dados de complemento de 2SG (OD e OI). Em termos gerais, observou a autora a prevalência do *lhe* correlacionado às formas *senhor/senhora*, *você*, revelando um quadro simétrico com o *lhe*, vinculado às formas de 3ª pessoa na fala do Recôncavo Baiano. Em termos probabilísticos, conclui Almeida (2014, p. 225) que “é possível sugerir que, em Santo Antônio de Jesus, vigora um subsistema pronominal supletivo (LOPES, 2011) em que *você* pode combinar-se com *te* ou *lhe*.”

Silva (2016), em sua Dissertação de Mestrado, também se volta para a fala baiana, assumindo como foco os pronomes-objeto em funções acusativa e dativa produtivos em Feira

de Santana/BA. A autora se vale dos dados extraídos de 60 entrevistas, que se dividem em entrevistas de fala espontânea (36) e entrevistas direcionadas (24). Seguindo o aporte teórico da Sociolinguística Laboviana, Silva (2016) submeteu os dados Pacote de Programas GoldVarb para geração dos índices percentuais e probabilísticos. Em termos gerais, Silva (2016) levantou a seguinte distribuição das formas dativas de 2SG: o *te*, em 45% dos dados (92/205), o *lhe*, em 44% (91/205), seguido pela forma preposicionada *para você*, em 10% (20/205) e, por fim, o *zero-dativo*, em apenas 1% (2/205). Tendo em vista a dinâmica intensa de alternância entre as formas *te/lhe*, a autora opta também por observar a alternância de *lhe* entre as funções de OD e de OI de 2SG, o que a permitiu constatar que o dativo de 2SG, na expressão falada de Feira de Santos, mostrou-se com o PR de 0.90 em relação às suas evidências no contexto do futuro perifrástico em 66,7% dos dados (4/46), cf. Silva (2016, p. 110).

Redirecionando o foco para a fala piauiense, passamos aos resultados de Pessoa (2017), voltados ao *dativo* em suas *formas clíticas e não clíticas* de 1ª, 2ª e 3ª pessoas à luz das orientações da Teoria da Variação (LABOV [1972], 2008; 1994; 2001; 2010). Pessoa (2017), em sua Tese de Doutorado, utiliza uma amostra de 96 entrevistas constantes do Banco de Dados do Projeto *Aspectos Gramaticais do Português Falado por Estudantes Teresinenses* da Universidade Federal do Piauí (UFPI), submetendo-os ao Pacote de Programa GoldVarb para a geração dos índices percentuais e probabilísticos. No que se refere especificamente aos resultados das estratégias dativas de 2SG, temos em análise do levantamento de 307 ocorrências, sendo 169/309 da variante clítica e 138/309 da variante preposicionada. Para a 2ª pessoa, constatou a autora a sua produtividade em 80% dos dados (54/67) com o respectivo PR de .88 na fala Piauiense. Em suma, a autora conclui o fato de os clíticos *me*, *te* e *lhe* evidenciarem traços de concordância expressos morfologicamente (MACHADO ROCHA, 2011), uma vez que o pronome na função de OI (dativo) atue no ato discursivo.

1.3.2.2 Os complementos dativos de 2SG no sul brasileiro

Nesta subseção deste capítulo de Revisão do Tema, voltamos o foco aos pronomes-complemento em função dativa abordados por Dalto (2002) voltados para dados do Sul do Brasil.

Dalto (2002), em sua Dissertação de Mestrado, dedica-se aos *pronomes-objeto* de 1ª e 2ª pessoas em amostras de fala de Curitiba, Porto Alegre e Florianópolis extraídas do Banco de Dados do Projeto VARSUL (Variação Linguística Urbana da Região Sul do Brasil). Foram utilizadas, em sua pesquisa, o total de 72 entrevistas, sendo 24 para cada capital do sul do Brasil, das quais foram extraídos 915 dados de Florianópolis, 596 dados de Curitiba e 680 dados de Porto Alegre. O estudo em questão se deixa conduzir pelos princípios teórico-metodológicos da Sociolinguística Variacionista Laboviana para a análise da variação pronominal entre as formas realizadas do *objeto* e do *objeto nulo*, lidando tanto com os dados de 1ª e 2ª pessoas através do VarBrul. A principal hipótese é a variação entre as formas de pronome-objeto – os clíticos pronominais (*me, te*) e as formas tônicas (*pra mim, pra ti*), pronome-sujeito (*eu, tu*) e o objeto nulo (*zero*) no exercício das funções de OD e de OI.

Os resultados de Dalto (2002) demonstram, de uma forma geral, que a forma *te* mostra-se como a estratégia pronominal de 2SG mais produtiva nas três capitais sulistas, quais sejam 77/157 dados, em Florianópolis, 121/196, em Curitiba e 116/164, em Porto Alegre. A segunda forma pronominal mais produtiva em Florianópolis é o SPrep, *para ti*, com 29/157 ocorrências, em Curitiba é o SPrep, *para você*, com 38/196 e em Porto Alegre é o *zero dativo*, com 23/157. Os condicionamentos para a aplicação da regra do *objeto nulo* nas três capitais são as seguintes: (a) na amostra de Florianópolis - (i) tipo de interlocução; (ii) tipo de objeto; (iii) pessoa; (iv) tipo de verbo; (v) tempo e modo verbais e (vi) escolaridade, nessa ordem; (b) na amostra de Curitiba - (i) tipo de interlocução; (ii) tipo de objeto; (iii) pessoa; (iv) tipo de verbo e (v) gênero de discurso; (c) na amostra de Porto Alegre - (i) tipo de objeto; (ii) pessoa; (iii) gênero de discurso; (iv) tipo de verbo; (v) tempo e modo verbais e (vi) faixa etária. Em relação ao traço de *pessoa do discurso*, os resultados apontam para uma tendência de que objeto projetado se mantenha produtivo para desempenhar a função de objeto de segunda pessoa (DALTO, 2002, p. 104).

Em síntese, os PRs para o preenchimento do objeto são de .68 (81%, 121/153) em Florianópolis, .77 (88%, 171/195) em Curitiba e .80 (84%, 112/134) em Porto Alegre, evidenciando que a segunda pessoa condiciona fortemente o preenchimento do pronome no sul do Brasil (DALTO, 2002, p. 103), uma vez o traço de *pessoa* mostra-se como 3º grupo selecionado para Florianópolis e para Curitiba, e o 2º, para Porto Alegre.

Síntese do capítulo

Neste capítulo, partimos da expressão do caso *dativo* em sua origem latina, passamos pela proposta da tradição gramatical até alcançarmos os resultados de análises descritivas embasadas em dados de sincronias passadas e de sincronias recentes especificamente em relação à 2SG.

Em relação ao caso dativo em latim, descrevemos brevemente os seus nove tipos, tendo em vista os seus traços semânticos e sintáticos, ainda que todos estejam orientados à sua principal função como “polo para o qual a ação ou o processo referido pelo predicado se orienta”, cf. Van Hoecke (*apud* OLIVEIRA, 2014, p. 25). À luz da tradição gramatical, constatamos o seu conservadorismo, ao perpetuar o quadro pronominal inclusive na função de OI, desconsiderando, pois, por exemplo, as formas preposicionadas nucleadas a partir do *você* (*a você/para você*) como formas que integram as estratégias dativas produtivas na língua portuguesa.

Quanto aos estudos linguísticos voltados para sincronias passadas, observamos, com base nas análises de Oliveira (2014) e de Cardoso (2017), ambos voltados exclusivamente ao *dativo de 2SG* na perspectiva histórica, evidenciam as estratégias dativas de 2SG nos séculos XIX e XX com formas do paradigma de *você*, além das etimológicas formas do paradigma de *tu*. Ambos os autores correlacionaram o contexto de nominativo ao de dativo de 2SG, o que evidencia a preferência, de um modo geral, dos redatores por acompanharem as opções assumidas por formas dativas dos paradigmas de *tu* e de *você* em relação às pronomes-sujeito *tu* e *você*. À luz de Rumeu (2014), identificamos o dativo como um contexto fértil ao *você não-sujeito*.

As análises linguísticas embasadas em amostras de língua falada evidenciaram, de um modo geral, a produtividade do *te*, ao passo que o *lhe* parece se mostrar mais proeminente nos dados de sincronias recentes (século XX) e nos contextos de *você-sujeito* (ALMEIDA, 2014; SILVA, 2016).

Uma vez exposto, ainda que brevemente, o estado da questão em discussão relacionada ao dativo de 2SG, passamos aos pressupostos teórico-metodológicos adotados nesta pesquisa de Mestrado.

CAPÍTULO 2. PRESSUPOSTOS TEÓRICO-METODOLÓGICOS

Este capítulo é dedicado à apresentação dos fundamentos teóricos adotados nesta pesquisa. Inicialmente, voltamos o foco a uma breve apresentação dos fundamentos teóricos da Sociolinguística de base Laboviana em termos de princípios e orientações (2.1). Passamos à apresentação dos aspectos teóricos e metodológicos que orientam o trabalho sociolinguístico com cartas (2.2). Em seguida, vamos à descrição da amostra de *Cartas ao Papai Noel* (2.3) e da linguagem RBrul na interface do RStudio (2.4), ferramenta metodológica responsável pela geração dos índices probabilísticos da regra variável em análise. Por fim, voltamos o foco para uma descrição pormenorizada das variáveis linguísticas e extralinguísticas controladas nesta análise (2.5), considerando não só a apresentação da variável dependente eneária (*te, a ti, para ti, lhe, a você, para você, você zero*), mas também os contextos linguísticos (*o tipo de predicador em relação à estrutura argumental da sentença, a forma do objeto direto nas estruturas ditransitivas, o valor semântico do predicador, o papel semântico do dativo*) e extralinguísticos (*a seção das cartas e a procedência das cartas no espaço geográfico mineiro*) em análise.

2.1 A pesquisa sociolinguística: princípios e orientações metodológicas para um estudo quantitativo

A Sociolinguística é a área da Linguística voltada para o estudo da língua em uso, o chamado vernáculo, fundamentando-se no princípio pelo qual toda língua natural é suscetível à influência de fatores linguísticos e extralinguísticos que promovem a variação e mudança. Assim, esta subárea assume que as línguas humanas estão em constante variação, a qual ocorre em uma dinâmica de heterogeneidade ordenada (WEINREICH, LABOV & HERZOG, 2006 [1968]). Em outras palavras, a pesquisa sociolinguística toma como objeto de estudo a variação linguística, reconhecendo os seus condicionamentos estruturais e sociais.

A Sociolinguística ganhou destaque na década de 1960 em virtude da insatisfação dos pesquisadores com as propostas teórico-metodológicas existentes à época, sobretudo o Gerativismo. A propulsão da Sociolinguística também foi fruto da influência que a sociologia

passa a exercer na interpretação linguística, como apontam Hernández-Campoy e Schilling (2012, p. 64/65). Em uma quebra de paradigmas, a Sociolinguística se contrapõe ao Estruturalismo e ao Gerativismo, ao combater a ideia de que a língua seria estanque e homogênea. Assim sendo, a proposta da Sociolinguística é desenvolver uma nova metodologia para o estudo da língua, concebendo esta como um fato social, entendendo e assumindo que existe uma regularidade na sua mudança, o que só pode ser medido pelo estudo dos fatores estruturais em conjunto com fatores sociais e culturais.

Durante o século XX, o Estruturalismo e o Gerativismo vinham estabelecendo o estudo da língua em completa exclusão de fatores externos a esta. Saussure, responsável pela inauguração do Estruturalismo, fazia uma distinção entre linguística interna e linguística externa, propondo um estudo da língua de forma totalmente dissociada do estudo dos fatores externos desta, que ficariam a cargo da linguística externa. O autor afirma: “nossa definição da língua supõe que eliminemos dela tudo o que lhe seja estranho ao organismo, ao seu sistema, numa palavra: tudo quanto se designa pelo termo *linguística externa*” (SAUSSURE, 2006 [1916], p. 29).

Também os Gerativistas compreendiam a língua como homogênea, afastando fatores sociais, culturais e históricos do seu estudo. Esta corrente teórica baseava-se na concepção inatista da língua, assim concebida por Noam Chomsky, que se fundamenta na percepção de que uma criança, embora não exposta a todos os dados de uma língua, consegue desenvolver competência para operar os mecanismos básicos desta (FARACO, 2005, p. 164), o que conduz à conclusão de que são aspectos biológicos que fornecem as condições mínimas para a aquisição da linguagem. O Gerativismo também deixava à margem os aspectos sociais relacionados à língua, uma vez que estava mais preocupado em aspectos biológicos, inatos ao ser humano, para explicar a língua.

Foi o pioneiro trabalho de Weinreich, Labov e Herzog (2006 [1968]) que inaugurou as orientações teórico-metodológicas para descrever e analisar o vernáculo tendo em vista a dinâmica de heterogeneidade ordenada. Percebendo a língua como um sistema que está sempre em uma dinâmica de equilíbrio instável, os autores postularam que a mudança linguística não ocorre de forma aleatória ou abrupta, mas vai se moldando paulatinamente tendo em vista os contextos linguísticos e extralinguísticos que coocorrem e condicionam o

fenômeno. Os autores, ao pensarem a Teoria da Mudança Linguística, buscavam estudar a fala de comunidades urbanas, identificando processos de variação que ocorriam nessas comunidades e buscando compreender tais fenômenos com base nos fatores que poderiam estar atuando naquele grupo de pessoas. Baseavam-se na concepção que o estudo sociolinguístico somente pode se dar pelo estudo do falante real e da sociedade em que está inserido.

Antes de Weinreich, Labov e Herzog (1968), William Labov já vinha defendendo a relação entre língua e sociedade, inclusive estudando a variação natural da língua falada. Já em seu primeiro estudo de 1963, estudava o inglês falado na ilha de Martha's Vineyard, nos Estados Unidos. Neste e nos seus demais estudos, Labov assume a proposta de sistematizar a variação linguística em seus encaminhamentos linguísticos e sociais. Assim, a Sociolinguística se interessa pelo estudo da variação e mudança linguística, o que faz através do tratamento estatístico dos dados coletados de falantes reais de uma determinada comunidade de fala. Quanto à definição dessa comunidade, inicialmente Labov (1972, p. 120) entende que uma comunidade de fala é constituída por membros que compartilham um conjunto de normas, o que fica evidente através da atitude avaliativa dos próprios falantes, que têm consciência das normas compartilhadas. Posteriormente, tal definição foi revista por Guy (2000, p. 18), que apresentou as seguintes características para a definição da comunidade de fala.

(i) características linguísticas compartilhadas; isto é, palavras, sons ou construções gramaticais que são usados na comunidade, mas não o são fora dela; (ii) densidade de comunicação interna relativamente alta; isto é, as pessoas normalmente falam com mais frequência com outras que estão dentro do grupo do que com aquelas que estão fora dele; (iii) normas compartilhadas; isto é, atitudes em comum sobre o uso da língua, normas em comum sobre a direção da variação estilística, avaliações sociais em comum sobre variáveis linguísticas. (GUY, 2000, p. 18)

Estabelecida a comunidade a ser estudada, o pesquisador também deve atentar às orientações para a obtenção dos dados, que deve ser feita com um rigor metodológico que possibilite a obtenção dos “dados puros”, ou seja, não influenciados por fatores que não são naturais daquela comunidade. Como a sociolinguística está preocupada com o estudo da língua falada em situação naturais, a coleta de dados deve ser feita sem que o pesquisador interfira no contexto de produção dos dados. No caso dos estudos sociolinguísticos voltados

para a língua escrita, também o pesquisador deve sempre ter em mente a busca por resgatar o *vernáculo* tanto em amostras de sincronias recentes (como é a deste trabalho), quanto em amostras de sincronias passadas representativas (GUY, 2007, p. 22).

Em cena, temos a questão do *envelope da variação*. A interpretação de Tarallo é a de que cabe ao pesquisador o papel de “analisar a situação de conflito e, com base em sólidos e firmes argumentos, desmascarar a assistemática do *caos*” (TARALLO, 2004, p. 33). No âmbito da sociolinguística, voltamos o foco para as formas variantes em curso, fazendo uma descrição detalhada da *regra variável* e dos seus condicionamentos (linguísticos e/ou extralinguísticos).

Passamos à discussão de aspectos relacionados às questões metodológicas que envolvem o trabalho com cartas.

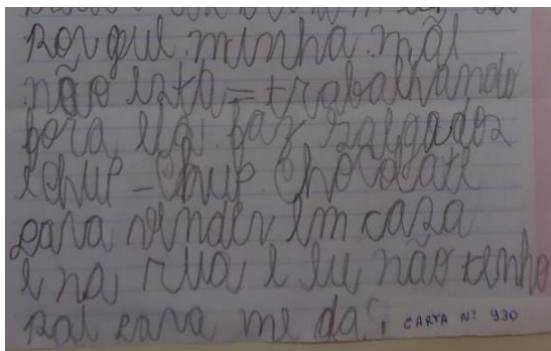
2.2 O estudo da variação linguística em cartas: questões metodológicas

Como o foco deste trabalho está voltado ao estudo sociolinguístico embasado em amostra de língua escrita do século XXI, é importante trazeremos à cena as questões da *autoria* e da *autenticidade* das cartas. Hernández-Campoy & Schilling (2012), ao tratarem especificamente dos desafios da prática da Sociolinguística Histórica, mas que também podem se aplicar à Sociolinguística não-histórica, discutem os parâmetros da *autenticidade* e da *autoria* das amostras.

A *autenticidade* da amostra está voltada para as atitudes linguísticas dos escreventes que podem não representar o seu vernáculo (“pureza dos textos”), tendo sempre em vista o caráter menos espontâneo da língua escrita. Assim sendo, é esperado do linguista-pesquisador a habilidade de detectar possíveis evidências de hipercorreção, de mistura dialetal e de “erros” da escrita (LABOV, 1994 *apud* HERNÁNDEZ-CAMPOY & SCHILLING, 2012, p. 56). Nas cartas ao PN, temos evidências da *autenticidade* das cartas relacionadas à expressão do vernáculo em suas produções escritas, tendo em vista o baixo nível de intimidade do redator com os parâmetros da escrita padrão. Na imagem 1, observamos não só a ausência de vírgulas, mas também a ausência do morfe de infinitivo

verbal [-r], em “eu não tenho pai para me dá” no lugar de “eu não tenho pai para me dar”. Trata-se do cancelamento do R em coda final que é um traço do PB já em adiantado estágio de sua manifestação, cf. Callou & Brandão (2016, p. 111), principalmente nos infinitivos verbais, independentemente de qualquer estratificação social.

Imagem 1: Fac-símile da Carta 930



“[...] minha mãe não esta = trabalhando fora ela faz salgados e chup-chup chocolate para vender em casa e na rua e eu não tenho pai para me dá.” (C930)

Em relação à *autoria* dos textos, a discussão principal é a de que os documentos podem ter sido escritos e assinados pelo próprio autor (autógrafo), por uma pessoa e assinados por outra (apógrafo) e assinados ainda por outra pessoa (testemunho idiógrafo) com a supervisão do autor (*autoria intelectual*), cf. discutido por Rumeu *et alii* (2020, p. 338). Nas cartas ao PN, também nos deparamos com a questão da *autoria* das cartas não só porque não se tem certeza sobre o fato de ter sido uma criança quem realmente redigiu o seu texto, mas também porque muitas crianças, por não saberem escrever, recorrem aos seus familiares e/ou amigos para a produção das suas cartas. Ainda que, nas cartas ao PN, não seja possível o resgate da *autoria* dos seus redatores, é possível atribuir, de um modo geral, à origem mineira, considerando o local da residência desses autores em várias localidades da Capital Belo Horizontina e do estado de Minas Gerais. Tal informação é resgatada a partir da leitura do envelope da carta endereçada ao PN.

A escolha de promover a investigação sociolinguística com base em um *corpus* constituído de *cartas* foi motivada pela natureza desse gênero discursivo. Em cartas não-oficiais, como é o caso das cartas ao PN, há uma espécie de imediatismo discursivo (KOCH & OESTERREICHER; [1985], 1994 *apud* OLIVEIRA, 2014, p. 84) que evidencia uma maior proximidade comunicativa entre remetente e destinatário, constituindo uma espécie de “conversa à distância” (OLIVEIRA, 2014, p. 84). Segundo a interpretação de Koch e

Osterreicher (2007), os textos de comunicação imediata estão mais atrelados à oralidade, informalidade e menor planejamento, ao passo que o distanciamento implica em maior formalidade e planejamento. Especificamente em relação à tradição discursiva, Koch (1997, p. 14 *apud* OLIVEIRA, 2014) assume que ³“todo discurso é exemplar de algum gênero literário, gênero textual ou forma de conversação, que revela características de um certo estilo para realizar certos atos de fala historicamente moldados” (KOCH, 1997, p. 14 *apud* OLIVEIRA, 2014). Acrescentemos, ainda à luz de Koch (1997 *apud* OLIVEIRA, 2014), o fato de todo discurso está inserido em uma dada tradição histórica, o que se ⁴“por um lado, na tradição de uma dada língua particular (ou variedade linguística), por outro lado, em uma certa tradição do discurso”.

Em suma, voltando o foco especificamente às *Cartas ao Papai Noel*, assumimos que essas cartas não podem ser tratadas como *cartas pessoais*, uma vez que os escreventes não dirigem a palavra a um amigo ou familiar, mas sim ao PN, o que exige, como já exposto, um certo grau de respeito e formalismo na comunicação. Por outro lado, a figura do PN não parece se manifestar como uma autoridade pública. O uso de expressões como *bom velhinho*, por outro lado, parece indicar uma tentativa aproximação entre os remetentes e a imagem que se faz do destinatário. No próprio termo *Papai Noel*, temos em evidência a referência à relação entre pai e filhos através da reduplicação de sílaba de “pai” em “papai”, denotando afetividade. Nesse aspecto, assumimos que as cartas ao PN representam espontâneos registros da produção escrita de crianças, adolescentes e adultos. Acreditamos que as cartas ao PN possam embasar interessantes análises linguísticas, porque evidenciam a variação linguística em sua expressão escrita, mais especificamente em relação às formas pronominais dativas de 2SG que representam o foco deste trabalho.

Uma vez discutidos, ainda que brevemente, as questões da *autoria* e da *autenticidade* das cartas ao PN, passamos à descrição da natureza das cartas ao PN.

³ “Letzterer Terminus besagt, dass jede Diskurs Exemplar irgendeiner Gattung, Textsorte oder Gesprächsform ist, Merkmale einer bestimmten Stillrichtung aufweist, zum Vollzug bestimmter historisch geprägter Sprechakte sprachlicher Historizität” (KOCH, 1997, p. 14).

⁴ “einerseits in der Tradition einer gegebenen Einzelsprache (oder Sprachvarietät), andererseits in einer bestimmten Diskurstadition” (KOCH, 1997, p. 14).

2.3 O corpus: a natureza das Cartas ao Papai Noel

Esta pesquisa tem como foco a descrição analítica das estratégias pronominais dativas de 2SG nas cartas ao PN, produzidas no espaço geográfico de Minas Gerais, no ano de 2012. Trata-se de cartas vinculadas ao Projeto social *Papai Noel dos Correios* promovido pela Empresa Brasileira de Correios e Telégrafos (ECT). A proposta desse projeto social é a de que crianças escrevessem cartas ao PN que seriam disponibilizadas à sociedade e, conseqüentemente, tivessem os seus pedidos natalinos realizados na época do Natal. Inicialmente, o projeto estava vinculado apenas às escolas, mas o crescimento do alcance do projeto motivou a sua abertura, de modo a entendê-las como *cartas da sociedade*, escritas não apenas por crianças, mas também por adolescentes e adultos.

As novecentos e sessenta missivas ao PN disponibilizadas ao projeto “*Cartas ao Papai Noel: a escrita em Minas Gerais*” sob a coordenação do Prof. Dr. Lorenzo Vitral têm o seu uso licenciado para fins de pesquisa linguística devidamente registrado no ⁵Comitê de Ética em Pesquisa (COEP) da UFMG. No âmbito do projeto, os seus pesquisadores atuaram na digitalização, transcrição e publicação dessas cartas, o que se concretizou a partir da disponibilização do fac-símile das cartas ao lado de suas transcrições (imagem 2), cuidando-se, obviamente, por resguardar a identidade e o endereço dos seus remetentes. A proposta de Vitral & Moreira (2020) foi a de edição conservadora dessas cartas, mantendo o seu conteúdo tal como escrito pelos seus redatores e resguardando, por exemplo, as evidências de desvios ortográficos, rasuras. As cartas são manuscritas não só por crianças, mas também por adultos cujos níveis de alfabetização são baixos, o que conduziu que a transcrição do termo ininteligível viesse identificada entre colchetes como “ilegível” ([*ilegível*]).

⁵ Registro no COEP/UFMG: CAAE 12131914-9.0000.5149.

Imagem 2: Edição proposta por Vitral & Moreira (2020, p. 44)

QUERIDO PAPAI NOEL
 GOSTARIA DE GANHAR UM ELECOPTERO
 DE CONTROLE REMOTO E
 GOSTARIA DE GANHAR UMA ARMINHA DE ÁGUA
 GOSTO MUITO DESSAS DUAS COISA
 E TAMBEM GOSTAREI MUITO DE VOCE
 SE VOCE REALIZAR O MEU
 DESEJO NÃO POR QUE VOCE
 VOCE É O PAPAI NOEL MAIS SIM POR QUE
 VOCE ESTA LENDO ESTA CARTO
 EU ESPERA QUE SEU NATAL SEJA MUITO ESPECIAL
 E QUE VOCE TENHO UM PROSPERO ANO NOVO
 ESPERA QUE VOCE REALIZE O MEU DESEJO
 E QUE ME DEIXE MUITO FELIZ NESTE NATAL!!!
 HO HO HO
 BAIRRO [...] Nº [...] NO
 RUA [...] CEP [...] LCT
 BELO HORIZONTE MINAS GERAIS MINAS GERAIS



Fonte: Vitral & Moreira (2020, p. 44).

Reconhecemos, em tais cartas, a expressão de uma escrita mais espontânea, uma vez que produzidas por pessoas ainda em processo de aquisição da língua escrita ou que já a tenham inabilmente adquirido em seu processo de aprendizagem via escola. Observamos ainda nessas cartas ao PN um acentuado nível de subjetividade dos redatores, ao exporem as suas realidades de vida como uma forma de argumentação para que alcançassem o objeto de desejo: o presente natalino.

Passamos, na sequência, ao método conduzido pelos encaminhamentos do RBrul na interface do RStudio.

2.4 O RBrul e o trabalho na interface do RStudio

O RStudio é um *software gratuito* que também leva ao trabalho com a linguagem R que, por sua vez, possui seu código aberto, convidando aos interessados a contribuir para aprimorar seus recursos como observado por Delgaard (2008, p. 7).

[...] Outros sistemas apresentam interfaces mais simples em termos de menu e modelos, mas frequentemente a facilidade de uso se torna um obstáculo a longo prazo. Ainda que estatísticas elementares sejam frequentemente apresentadas como uma coleção de procedimentos fixos, a análise de dados razoavelmente complexos requer a construção de modelos

estatísticos *ad hoc*, o que torna a flexibilidade do R muito desejada. [...]” (DALGAARD, 2008, p. 7, tradução livre)⁶

Na perspectiva de Oushiro (2014, p. 52), a linguagem computacional R é potencialmente promissora, evidenciando recursos para o tratamento de dados complexos, distribuídos em diversas variáveis de diversos níveis (JOHNSON, 2008, p. 174). Nesse sentido, à luz de Oushiro (2014, p. 134), a linguagem R permite ir além da estatística básica tendo em vista a geração de gráficos e de listas de frequência. Acrescentemos que uma das vantagens da submissão dos dados à linguagem R é o fato de termos disponível o pacote RBrul responsável também pela implementação de análise multivariável nos níveis *step-up/step-down* e *one level* através do RStudio.

Gomes (2012), ao fazer uma análise comparativa entre as ferramentas Goldvarb, Varbrul e RBrul, destaca que a última foi desenvolvida para cumprir com tudo aquilo que as demais realizavam, mas expandindo sua atuação. A autora cita como evidências a possibilidade de assumir uma variável numérica contínua como variável resposta, a inclusão de modelos mistos e, ainda, a estimativa de efeitos (GOMES, 2012, p. 262). Os resultados gerados na linguagem R em termos de peso relativo como já se tinha através do Goldvarb e do Varbrul são gerados como resultados em *log-odds*. Além de ser uma versão aprimorada de seus antecessores, o RBrul tem a vantagem de estar coligado a diversas outras funcionalidades do R, como as ferramentas de trabalho com *corpus* e seu potencial gráfico. Assim, em uma mesma plataforma, é possível tratarmos os dados, lidando, de forma simples e rápida, com os *knockouts*, a análise estatística e a geração de gráficos que representem os resultados.

Para o uso do programa, é necessário primeiro ter acesso tanto ao *software* padrão da linguagem R, que pode ser baixada pelo endereço eletrônico <<https://cran.r-project.org/>>. Na plataforma do RStudio, baixada através do endereço eletrônico <<https://www.rstudio.com/products/rstudio/download/>>, é possível fazermos a instalação do pacote adicional do RBrul, que pode ser feito dentro do próprio RStudio, pelo comando

⁶ “[...] Other systems presente simpler interfaces in terms of menus and forms, but often the apparent userfriendliness turns into a hindrance in the longer run. Althought elemental statistics is often presented as a collection of fixes procedures, analysis of moderately complex data requires ad hoc statistical model building, which makes the added flexibility of R highly desirable. [...]”

source (<http://www.danielezrajohanson.com/Rbrul.R>). Baixado o pacote, o RBrul passa a ser acessado pela função *oldbrul* ().

Avelheda Bandeira & Souza (2019) desenvolveram um *passo a passo para uso do RBrul*, no qual apresentam duas etapas para trabalhar com a ferramenta. A primeira dá-se no âmbito do *Excel*, ao passo que, a segunda etapa está voltada para a elaboração de *data frames* com os dados da pesquisa que permitirão a geração de tabelas e gráficos. Nessa segunda etapa, temos os encaminhamentos (uma espécie de guia) para o uso específico da ferramenta na qual destacam etapas como ajuste das variáveis, recodificações e as rodadas multivariáveis para a geração dos índices probabilísticos (pesos relativos) em função dos resultados dos modelos multivariados também em termos de *log-odds*. Em diálogo com Guy & Zilles (2007, p. 212), destacamos em relação aos pesos relativos que se trata de “uma representação abstrata dos efeitos de contexto independentemente dos níveis gerais de uso de um processo”. Em outras palavras, isso quer dizer que a cada fator (variável preditora) é dado um peso, também chamado de *valor do fator* ou *probabilidade*, que “caracteriza o efeito deste fator sobre a regra variável em questão” (GUY, 2007, p. 51). Ao peso que é relativo a um valor padrão, 0,50 para variáveis binárias, entendemos que a hipótese regente da variável preditora não exerce efeito sobre a resposta, não favorecendo e nem desfavorecendo o uso de uma de suas variantes. Se o valor do peso relativo for superior ao padrão, ou seja, acima de 0,50 ($> 0,5$), temos em cena a indicação de que a variável preditora favorece o uso da variante de referência (sob estudo), ao passo que caso o seu valor seja inferior a 0,50 ($< 0,5$) temos a indicação de um desfavorecimento da variável preditora em relação à variante dependente (binária ou eneária) estudada.

Em síntese, a ideia é a de utilizarmos-nos da linguagem R como método de trabalho especificamente na sua interface RStudio através do RBrul responsável por produzir, em análise multivariável, os índices probabilísticos de uma dada regra variável nos níveis *step-up & step-down*. Passamos, na sequência, à descrição da variável dependente a partir da qual submetemos a sua correlação com as variáveis independentes (linguísticas e extralinguísticas), tendo sempre em vista as hipóteses e/ou objetivos norteadores desta análise.

2.5 O contexto de complementação verbal dativa e os critérios de análise linguística: a caracterização da regra variável

À luz de Duarte & Brito (2003) e Raposo (2013) descrevemos, inicialmente, o dativo em relação aos seus traços sintáticos e semânticos (2.5.1), atentando, na sequência, à descrição da variável dependente (2.5.2) e das variáveis independentes (2.5.3), considerando ainda, à luz de Berlinck (1996), os tipos de dativo e os seus papéis semânticos com base em Company Company (2006). Em termos de variáveis extralinguísticas, atentamos, de um modo geral, às seções das cartas (2.5.3.6), considerando a sua composição estrutural, e aos locais de escritura das missivas (2.5.3.7).

2.5.1 O dativo de 2ª pessoa do singular: traços sintáticos e semânticos

A compreensão do que sejam as estratégias dativas de 2SG pressupõe o entendimento da noção de *predicador* que, à luz de Duarte & Brito (2003) e Raposo (2013), é um item lexical (verbo) capaz de projetar uma estrutura argumental ou grelha temática, cf. Raposo (2013, p. 358). No mesmo caminho de Raposo (2013), também seguem Duarte & Brito (2003, p. 183), ao conceberem os verbos como *palavras predicativas por excelência*, responsáveis pela projeção sintática das funções de sujeito e complementos.

Para a Lógica de Predicados, as proposições – e, para o linguista, as frases que as veiculam – organizam-se em redor de um elemento central, a que chamamos aqui **predicador**. A propriedade fundamental dos predicadores consiste em combinarem-se com **argumentos** [...]. Os argumentos são requeridos pelo predicador para lhe completar o sentido, formando assim uma proposição completa. (RAPOSO 2013, p. 358)

Neste trabalho, estão em discussão as estratégias dativas de 2SG, projetadas por um núcleo lexical verbal, cf. discutido por Duarte & Brito (2003) que interpreta os complementos verbais dativos como “verdadeiros argumentos”, uma vez que se trata de elementos sintática e semanticamente projetados por verbos. Em termos formais, observamos tratar-se de complementos verbais cuja implementação pode se dar a partir de verbos de *dois* ou de *três* argumentos, mostrando-se ainda cliticizável através do pronome *lhe*. Os verbos de dois

argumentos são os que projetam um sujeito e um argumento interno na forma de um OI, o que transparece no verbo *telefonar* (DUARTE 2003, p.185), cf. ilustramos em (21). Os predicadores de três argumentos (verbos ditransitivos) projetam um argumento externo e dois argumentos internos, cabendo aos internos as interpretações em relação aos complementos não-encetados a partir de preposição (OD) e encetados por preposição (OI) (DUARTE 2003, p. 185), tais como ilustramos com os verbos *mandar*, *enviar* e *pedir* de (22) a (24).

Predicadores de dois argumentos (SU V OI):

(21) ⁷“Murilo: Sabia que você andava pelo Rio, quando cheguei aqui. [eu]_{SUJ} Pretendia telefonar-[lhe]_{DAT} – quando hoje recebi sua mensagem. Espero vê-lo. [...]” (OLR. RJ, 09.01.1957)

Predicadores de três argumentos (SU V OD OI):

(22) “[...] Papai Noel meu nome e [...] e tenho 5 aninhos, todo ano [eu]_{SUJ} mando \emptyset _{DAT} [esta cartinha]_{OD} mas, ainda não ganhei o presentinho tão esperado [...]” (C424)

(23) “[...] [eu]_{SUJ} estou enviando \emptyset _{DAT} [esta carta]_{OD} para *lhe*_{DAT} pedir [um Notebook]_{OD} [...]” (C339)

(24) “[...] [a minha mãe]_{SUJ} *te*_{DAT} pede [uma cesta de natal]_{OD} [...]” (VITRAL & MOREIRA, 2020, p. 120)

Interpretamos o dativo de 2SG como argumento interno de uma relação gramatical, que, em termos sintáticos, é projetada por um predicador verbal de dois (SU V OI) ou três lugares (SU V OD OI) cliticizável através do *lhe*. Entendemos tratar-se de um complemento com o papel semântico de *alvo*, *beneficiário* da ação, marcado pelo traço [+animado], cf. Duarte (2003, p. 289-290). Nas cartas ao PN, foco de análise deste trabalho, temos em cena um único destinatário que apesar de constituir como um ser imagético, assume uma figura corporificada na sociedade principalmente na época do Natal. Trata-se de um único interlocutor marcado pelo traço [+ animado] cuja referência pronominal em contexto de complemento dativo pode se deixar evidenciar através das formas *te*, *lhe*, *a/para você*, *a/para ti*, *você* e *zero*.

⁷ Dado discutido por Cardoso (2017, p. 120).

2.5.2 Os complementos dativos de 2SG: a variável dependente eneária

Voltando o foco especificamente para as estratégias dativas de 2SG, assumimos, em termos sintáticos, o complemento dativo como um complemento verbal que é cliticizável através do clítico *lhe* sob a projeção de verbos de dois ou três lugares. Trata-se, semanticamente, de um complemento verbal com as noções semânticas de alvo, fonte, beneficiário da ação marcado pelo traço semântico [+animado]. Na perspectiva da tradição gramatical, identificamos que o complemento dativo é parcialmente contemplado sob o rótulo de *objeto indireto*, visto se tratar de um complemento cuja conexão sintática dá-se indiretamente através de uma preposição. De (25) a (36), trazemos evidências de estratégias dativas estritamente pronominais (*te*, *lhe*, *a / para você*, *você-dat.*, *a/para ti* e *zero (Ø)*) projetadas pelos verbos *desejar*, *pedir* e *agradecer*.

(25) “[...] [*te*]_{DAT} desejo um feliz natal” (C51)

(26) “[...] [*lhe*]_{DAT} desejo um feliz natal” (C60)

(27) “[...] Desejo [*a você*]_{DAT} um feliz natal [...]” (C21)

(28) “[...] [*te*]_{DAT} desejo um feliz natal *para você*” (C03)

(29) “[...] “[...] e aproveito para desejar [*você*]_{DAT} e sua família um feliz natal e um ano novo cheio de paz, amor e realizações” (C448)

(30) “[...] queria [*te*]_{DAT} pedi uma sexta de material escolar por que minha mãe não pode pagar [...]” (C87)

(31) “[...] Eu gostaria de pedir [*você*]_{DAT} papai noel um tablet [...]” (C98)

(32) “[...] Papai Noel eu queria pedir [*para você*]_{DAT} mim dar um Quiti de maquiagem [...]” (C85)

(33) “[...] Nessa véspera de natal eu gostaria de pedir [*a você*]_{DAT} que esta lendo a minha carta, um “notebook” ou um “tablet”. [...]” (C152)

(34) “[...] eu peço [*Ø*]_{DAT} uma cesta básica [...]” (C55)

(35) “[...] se não puder mim dar o presente que eu estou [*Ø*]_{DAT} pedindo, pode mim dar o que você quiser [...]” (C71)

(36) “[...] Agradeço a DEUS e [*a ti*]_{DAT} bom velinho(a) que você me Compreendeu meu sinjelo e humilde pedido e desde já Também desejo-[*lhe*]_{DAT} um Feliz Natal cheio de realizações [...]” (C94)

Considerando a apreensão de uma variável dependente (eneária) através das formas *te*, *lhe*, *a/para você*, *a/para ti*, *você* e *zero (Ø)*, passamos à descrição-analítica dos contextos (linguísticos e extralinguísticos) a partir dos quais os dados de dativos de 2SG serão controlados. São eles: *o contexto de sujeito de 2SG*; *o tipo de predicador em relação à*

estrutura argumental da sentença; o valor semântico do predicador; o papel semântico do dativo, a forma do objeto direto nas estruturas ditransitivas (S V OD OI). Além desses contextos estruturais, passamos à análise das seguintes variáveis extralinguísticas: a *seção da carta* e a *procedência da carta no espaço geográfico mineiro*. Nesta análise, não controlamos as variáveis *sexo* e *faixa etária*, uma vez que os redatores das cartas nem sempre são as crianças com as devidas referências às suas respectivas faixas etárias. Em muitas delas, os pais, avós, tios ou amigos das crianças redigiram as cartas ao PN, restringindo o controle do *sexo* e da *faixa etária* dos informantes, tendo em vista o alcance impreciso de tais informações extralinguísticas.

2.5.3 Os complementos dativos de 2SG e os critérios de análise: as variáveis independentes

Esta dissertação é conduzida pelo intuito principal de depreensão das possibilidades de manifestação dos pronomes-complemento dativo de 2SG (*te, lhe, a/para você, a/para ti, você* e *zero (Ø)*), à luz dos princípios teóricos-metodológicos da Teoria da variação e da mudança linguística (WEINREICH, LABOV, HERZOG, 2006 [1968], LABOV, 1994), passamos à descrição-analítica dos parâmetros linguísticos e extralinguísticos (GUY & ZILLES, 2007; OUSHIRO, 2017), pensados a partir dos resultados de pesquisas linguísticas acerca do dativo de 2SG (OLIVEIRA, 2014; CARDOSO, 2017; CRUZ, 2017).

2.5.3.1. O contexto de sujeito de 2SG

A primeira variável independente é o contexto de pronome-sujeito de 2SG correlacionado às estratégias pronominais dativas também de 2SG. A reorganização do sistema pronominal do PB a partir da inserção do *você*-sujeito, conduziu a reestruturação do sistema de pronomes-complemento que passou a alcançar formas clíticas como *lhe, o/a* para a referência à 2SG que, por sua vez, é o contexto etimológico reservado às formas do paradigma de *tu (te, ti, a/para ti)*. Em (37), trazemos à cena o contexto de *tu-sujeito*

combinado ao *te*-complemento, ao passo que, em (38), temos o *te*-complemento correlacionado ao *você-sujeito*.

⁸Carta de *tu-sujeito*:

(37) “[...] Quanto á escola, é como bem pódess_{SUJ} avaliar. Tenho recebido tuas cartas com muito prazer e contentamento; e, si não [*te*]_{DAT} escrevo sempre, peço-[*te*]_{DAT} dar me desconto: pouco ou nenhum tempo para escrever com tranquillidade e socêgo [...] (RAAP. Lagoa Santa, 30.10.1912) [...]”

Carta de *você-sujeito*:

(38) “[...] mas como ela me ama muito sabendo não ter condições, resolvel escrever está carta com a minha ajuda [*te*]_{DAT} pedindo amigonoel um brinquedo [...] *você*_{SUJ} pode escolher o carro e eu vou comemorar os meus 6 anos muito feliz...” (C913)

A hipótese já testada em outras análises linguísticas (LOPES & CAVALCANTE, 2012; RUMEU, 2020) é a de que as estratégias de complementação verbal de 2SG sugerem que o uso de formas do paradigma do *você-sujeito* favoreceria não só formas do paradigma de *você* (*lhe, o/a, você, prep+você*), mas também o *te-complemento* inclusive em estruturas de redobro (“Eu vou *te_i* levá *ocê_i* lá”) na fala mineira contemporânea, cf. Duarte & Diniz (2012, p. 92).

Dentre as estratégias de referência ao sujeito de 2SG adotadas pelos redatores das Cartas ao PN, temos em cena o *você-sujeito*, como exemplificado em (39), e formas de tratamento variadas como, por exemplo, o pronome de tratamento “*o senhor*” (40) que, pode inclusive assumir a função de vocativo (não-sujeito de 2SG) “*Papai Noel*” (40).

(39) Carta de *você*: “[...] E desde já gostaria de [*te*]_{DAT} agradecer muito por *você* está me compreendendo e estar me fazendo esta humilde colaboração nesse natal. Muito obrigada por *você* ler a minha carta e estar compreendendo o que eu estou falando \emptyset por meio dessa” (C152)

(40) Carta de forma nominal de tratamento: “Olá *Papai Noel*! Como *o senhor* está? Esta tudo bem! [...] Primeiro eu quero [*te*]_{DAT} desejar, que *o senhor* tenha um feliz natal, cheio de paz, alegria e tranquilidade [...]” (C090)

⁸ Como não identificamos o *tu-sujeito* nas Cartas ao PN, ilustramos, à título de exemplificação, com um dado de *te*-dativo em contexto de *tu-sujeito* levantado por Rumeu (2020, p. 242).

Ainda que a “uniformidade tratamental” não se aplique à realidade linguística do PB, Cardoso (2017, p. 148) constatou que “as cartas de *você-sujeito* tendem a motivar o *lhe* (149/239, 62%) e as cartas de *tu-sujeito* tendem a condicionar o *te* (120/155, 77%)”. Com base em tais resultados referentes à escrita mineira dos séculos XIX e XX, que evidenciam a convivência de *te-dativo* em contextos de *você-sujeito*, assume-se a hipótese que esta convivência ainda seja produtiva na escrita mineira da atual sincronia.

2.5.3.2 O tipo de predicador em relação à estrutura argumental da sentença

Considerando que o foco deste trabalho é o *dativo* de 2SG, temos dois tipos de predicadores que projetam e selecionam obrigatoriamente o *dativo*. São eles: os verbos de dois lugares (S V OI) responsáveis pela projeção de um sujeito (nominativo) e de um objeto indireto (dativo) e os verbos de três lugares (S V OD OI) impulsionadores de um de um sujeito (nominativo) e de dois argumentos internos (complementos): um objeto direto (acusativo) e um objeto indireto (dativo). De (41) a (44), trazemos evidências desses dois tipos de predicadores (2 e 3 lugares) na amostra das cartas em análise.

Verbo de dois lugares com argumento interno indireto (S V OI):

(41) “[...] Se você me dar eu [te]_{OI} agradecerei muito [...]” (Vital & Moreira, 2020, p. 62)

(42) “[...] Vou agradecer muito hà Deus é [a você]_{OI} Papai Noel. [...]” (Vital & Moreira, 2020, p. 174)

Verbo de três lugares (ditransitivos) com argumentos interno direto e indireto (S V OD OI):

(43) “[...] Quero [...] [te]_{OI} desejar [um feliz Natal e um prospero Ano Novo]_{SNs-OD} cheio de Paz [...]” (C909)

(44) “Papai Noel eu queria [te]_{OI} pedir [um presente muito Legal]_{OD} Eu queria [te]_{OI} pedir [um Kit ou um fichário]_{OD}” (C143)

Quanto à estrutura argumental do predicador verbal, este trabalho parte também da hipótese que os verbos de três lugares (S V OD OI) tendam a favorecer o dativo de 2SG, cf. também discutido por Oliveira (2014), Cardoso (2017).

2.5.3.3 A forma do objeto direto nas estruturas ditransitivas de complementação verbal

Tendo em vista o fato de as estruturas de complementação dativas de 2SG projetadas por formas verbais ditransitivas (S V OD OI) terem em sua grade argumental a projeção de um complemento verbal acusativo (OD), a forma desse complemento passa também pelos parâmetros desta análise linguística. A ideia é investigarmos se há correlação entre a forma adotada pelo escrevente para expressar o *acusativo* e a forma adotada para o *dativo*. Para o controle desta variável, temos em análise quatro (4) formas de concretização do *acusativo* tais como o *sintagma nominal*, o *sintagma oracional*, o *clítico* e o *objeto direto nulo* que estão a seguir ilustrados de (45) a (48).

(45) *Sintagma nominal*: “Papai Noel eu queria [te]_{DAT} pedir [um presente muito Legal]_{SN-ACUS} Eu queria [te]_{DAT} pedir [um Kit ou um fichário]_{SN-ACUS} [...]” (C143)

(46) *Sintagma oracional*: “Sabe papai noel estou lhe escrevendo esta carta para [lhe]_{DAT} perguntar [se você poderia me presentear com materiais escolar para o ano de 2013]_{OR-ACUS}. [...]” (C291)

(47) ⁹*Clítico*: “[...] Por estes dias, porém, enviar-t[*o*]_{CLÍTICO-ACUS-ei} pelo correio [...]”

(RAAP. Lagoa Santa, 06.04.1914)

(48) ¹⁰*Objeto direto nulo*: “[...] Gostaria de Onnar a Barbie pop-star [...] So que Não tenho diNHEiro agora, por isso estou te escrevendo \emptyset _{SN-ACUS}”. Te agradeço desde Já. (C011)

No PB, é improdutiva a contração dos clíticos principalmente em sua expressão falada ¹¹(*mo, to, lho, no-lo, vo-lo*), ainda que tenhamos evidências de tal pronome-complemento aglutinado ao verbo na produção escrita mineira novecentista tal como ilustramos em (33), mostrando-se totalmente improdutiva nas cartas ao PN em análise. Acompanhamos a hipótese de Oliveira (2014, p. 104) em relação à possibilidade de uma relação inversamente proporcional entre os complementos *dativo* e *acusativo* nas cartas cariocas (séculos XIX e

⁹ Como não levantamos objeto direto na forma de clítico nas Cartas ao PN, ilustramos, à título de exemplificação, com um dado discutido por Cardoso (2017, p. 102).

¹⁰ Leia-se “Onnar” por “ganhar”.

¹¹ À luz de Cunha & Cintra (1985, p. 300), “No Brasil, quase não se usam as combinações *mo, to, lho, no-lo, vo-lo* etc. Da língua corrente estão de todo banidas e, mesmo na linguagem literária, só aparecem em escritores um tanto artificiais”.

XX). Isso quer dizer que um sintagma oracional (forma mais “pesada” do OD) tenderia a impulsionar o *clítico* (forma mais “leve” de OI) como dativo de 2SG, hipótese também testada por Cardoso (2017) para as cartas mineiras (séculos XIX e XX).

2.5.3.4 O valor semântico do predicador

Berlinck (1996) propõe uma tipologia para o dativo que possibilite traçar os contextos de sua ocorrência no português, o que inclui, em suas palavras, “o tipo de verbo ou estrutura verbal com a qual o dativo coocorre e o tipo de relação que este mantém com o verbo ou a sentença”¹² (BERLINCK, 1996, p. 128). A autora entende que as construções transitivas são as estruturas prototípicas para a realização do dativo, que geralmente apresenta a seguinte organização sintática.

$$N_0 + V + N_1 + \{a, para, em\} + N_2$$

Nesse esquema, N_0 representa o sujeito, N_1 representa o acusativo (objeto direto) e N_2 representa o dativo (BERLINCK, 1996, p. 128). A autora faz uma ressalva quanto à ordem sintática, esclarecendo ser possível que o dativo ocorra logo após o verbo, ou, se na forma de clítico, pode até mesmo preceder o verbo. A autora assume quatro categorias semânticas básicas para as estruturas transitivas, a saber: (1) *transferência material*; (2) *transferência verbal e perceptual*; (3) *movimento físico*; e (4) *movimento abstrato*.

Quanto à categoria *transferência material*, a autora entende tratar-se da expressão prototipicamente representada pelo verbo *dar* através de um esquema em que o sujeito gramatical (N_0) e o objeto acusativo (N_1) mostram-se como propriedades do objeto dativo (N_2). Berlinck (1996) cita os seguintes verbos que compõem esse grupo: *alugar, atribuir, confiar, dar, devolver, oferecer, pagar e transferir*, dentre outros. Observa ainda a autora ser possível que o esquema da transferência material ocorra de forma reversa, em que o objeto acusativo (N_1) deixa de ser propriedade do dativo (N_2), o que ocorre em verbos como:

¹² “The present section thus intends to provide a typology of such contexts, including the type of verb or verbal construction with which the dative complement can co-occur and the kind of relationship it maintains with the verb or the sentence” (BERLINCK, 1996, p. 128)

arrebatado, comprar, confiscar, cortar, emprestar, enviar, pedir, tirar e tomar, dentre outros. Apresentam-se, de (49) a (52), evidências para cada esquema da transferência material.

(49) “não entregaram [as mercadorias]_{OD-ACUS}. [ao comprador]_{OI-DAT}.”

(50) “a gente dá [um dado]_{OD-ACUS}. [para ele]_{OI-DAT}, ele fornece [outro]_{OD-ACUS}. [para a gente]_{OI-DAT}.”

(51) “Felipe pediu [um chocolate]_{OD-ACUS}. [para o avô]_{OI-DAT}.”

(52) “Maria tomou-[lhes]_{OI-DAT} [tudo o que tinham]_{OD-oracional}.”

A autora esclarece que, nos casos de *transferência material inversa*, o complemento dativo normalmente é introduzido pela preposição *de* (como exemplificado em (53)), mas pode ser introduzido pela preposição *a* (como em (54)), mas há uma restrição ao uso da preposição *para*. Berlinck (1996) atribui essa restrição ao fato que o dativo, que normalmente tem valor de *beneficiário*, nestes casos pode ter o valor de *fonte*, que normalmente é expressa pela preposição *de*, ao passo que *para* expressa a noção de *meta*.

(53) Pedro tirou os livros_{OD} [das mãos de Joana]_{OI}.

(54) * Pedro tirou os livros_{OD} [às mãos de Joana]_{OI}.

Berlinck (1996) descreve as propriedades distributivas das formas verbais movidas pela semântica de *transferência material* conforme o seguinte esquema.

$[+/- \text{animado}] N_0 + V + [(+)/- \text{animado}] N_1 + \{a, para, de\} + [+/- \text{animado}] N_2$
--

No que diz respeito à categoria *transferência verbal e perceptual*, esta é prototipicamente representada pelo verbo *dizer* e funciona através do esquema em que se tem o sujeito gramatical (N_0) e o complemento dativo (N_2) do qual o sujeito gramatical possui um certo conhecimento, ideia ou percepção (N_1). Deparamo-nos, assim, não com a transferência de um objeto concreto, como na *transferência material*, mas de um abstrato, como efeito de um ato de comunicação. Nesse sentido, Berlinck (1996) expõe verbos como *aconselhar, anunciar, confessar, contar, dizer, ensinar, escrever, falar, perguntar, prometer, sugerir, telefonar* como verbos de *transferência verbal e perceptual*. Em (55) e (56), apresentamos

evidências de formas verbais cuja semântica é a de *transferência verbal e perceptual* em conformidade com Berlinck (1996).

(55) Pedro disse [*para seus colegas*]_{OI-DAT} [*que o diretor estava doente*]_{OD-ACUS (ORACIONAL)}

(56) Ela [*me*]_{OI-DAT} ensinou [*a técnica de leitura*]_{OD-ACUS}.

Berlinck (1996, p. 132) revela as seguintes propriedades distributivas dos verbos de *transferência verbal e perceptual*:

[+/- animado] + N_0 + V + [- animado] N_1 + {*a, para*} + [+ animado] N_2

A autora entende que N_2 sempre possui o traço [+animado] uma vez que “somente entidades animadas podem participar de um processo de transferência de conhecimento” (BERLINCK, 1996, p. 131). Quanto ao grupo dos verbos de *movimento físico*, este é representado pelo verbo prototípico *levar* e, como indica o nome, funciona em um esquema cuja noção de transferência está em efetivo movimento físico tais como evidenciam os verbos *acrescentar, conduzir, dirigir, encaminhar, levar, pôr, trazer*. O esquema do *movimento físico* implica uma *transferência* do acusativo (N_1) para o dativo (N_2) como “um movimento prototipicamente físico direcionado a um alvo” (BERLINCK, 1996, p. 132). Em (57) e (58), trazemos evidências de verbos de *movimento físico* fornecidas pela autora.

(57) No aniversário do amigo levou-[*lhe*]_{OI-DAT} [*um livro*]_{OD-ACUS}

(58) Elas [*me*]_{OI-DAT} trazem [*esse material todo*]_{OD-ACUS} para ser discutido em aula.

Berlinck (1996, p. 132) descreve as seguintes propriedades distributivas para os verbos de *movimento físico*.

[+/- animado] N_0 + V + [+/- animado] N_1 + {*a, para, em, de*} + [+/- animado] N_2

Nos casos em que o dativo (N_2) possui o traço [+ animado], a ideia que se tem é de um *locativo*, ao passo que, se for [- animado], tem-se um *beneficiário*. Para a categoria dos verbos de *movimento abstrato*, tem-se o verbo *submeter* como seu prototípico. Nesta, ainda

há a ideia de movimento, mas sem que haja transferência, ao contrário, “há uma aproximação abstrata entre uma entidade e um estado, sentimento, ideia ou outra entidade”¹³ (BERLINCK, 1996, p. 133). Em (59) e (60), trazemos evidências de verbos de *movimento abstrato* através das formas verbais “submeter” e “dedicar” como discutido por Berlinck (1996, p. 133).

[+/- animado] N_0 + V + [+/- animado] N_1 + {*a, para, em*} + [+/- animado] N_2

(59) Os trabalhos *lhe*_{OI-DAT} foram submetidos ontem.

(60) A reputação dos astros foi prejudicada [pelo amor]_{OD-ACUS} que os Nazis [*lhes*]_{OI-DAT} dedicaram.

Apresentadas as quatro categorias dadas por Berlinck (1996) para as construções dativas (*transferência material, transferência verbal e perceptual, movimento físico e movimento abstrato*), é importante ter em mente que o dativo também pode ocorrer em construções intransitivas. Berlinck (1996) expõe três tipos de verbos em estruturas intransitivas tais como os *verbos de interesse, verbos de movimento e verbos de movimento psicológicos*. Nessas estruturas, Berlinck (1996, p. 136) entende que não há *agente* ou *causa* envolvidos e o complemento dativo tem função de descrever um estado ou associação em relação aos dois argumentos do verbo.

O *dativo de interesse* está prototipicamente representado pelo verbo *obedecer* e deve ser estudado com base em dois aspectos: o envolvimento do referente do sujeito gramatical (N_1) no processo descrito pelo verbo e a referencialidade do complemento dativo (N_2). Em relação ao primeiro aspecto, a autora distingue os verbos deste grupo em dois subgrupos, uma vez que há verbos em que o sujeito gramatical (N_1) envolve-se de forma ativa e voluntária com a associação, ao passo que há verbos em que não há voluntariedade nesse envolvimento.

Quanto ao *dativo de interesse*, Berlinck (1996, p. 136) expõe as formas *obedecer, faltar, resistir, servir, aderir e assistir* como verbos que projetam tal estrutura de complementação verbal. Nesses casos, a natureza de voluntariedade que envolve o sujeito (N_1) que apesar de ser, normalmente, [+ animado], também ser [- animado], sendo o dativo

¹³ “[...] there is an abstract approximation between an entity and a state, a feeling, an idea, or another entity” (BERLINCK, 1996, p. 133).

expresso em sintagma preposicionado encetado pela preposição “a”, o que se reflete na representação esquemática a seguir exposta. Em (61) e (62), apresentamos evidências de *dativo de interesse* discutidas por Berlinck (1996, p. 137).

$[+ / (-) \text{ animado}] N_1 + V + (a) + [+ / - \text{ animado}] N_2$

(61) Aos 18 anos, todos os jovens começam a servir [*ao Exército*]_{OI-DAT}.

(62) João sempre obedeceu [*às regras de trânsito*]_{OI-DAT}.

Ainda em relação ao *dativo de interesse*, observa a autora que não há envolvimento voluntário do sujeito gramatical (N_1), estando os participantes em uma associação estática ou correspondente ao *dativo* (N_2), o que pode ser verificado nos verbos *concernir*, *corresponder*, *equivaler*, *pertencer* e *sobrar*, cf. Berlinck (1996, p. 138). Há a possibilidade de o sujeito gramatical (N_1) ser [- animado], ainda que o *dativo* (N_2) possa ser tanto [+ animado], quanto [- animado], sendo o *dativo* expresso (N_2), cf. já exposto, através de um SPrep com preposição “a”, e assumindo a seguinte representação esquemática, cf. Berlinck (1996, p. 138). Apresentamos ainda, de (63) a (65), evidências do *dativo de interesse* sem traço de voluntariedade.

$[- \text{ animado}] N_1 + V + (a) + [+ / - \text{ animado}] N_2$

(63) O edifício pertence [*a um milionário do petróleo*]_{OI-DAT}.

(64) O assunto [*lhe*]_{OI-DAT} concernia muito.

(65) O preço deste anel equivale [*a dois anos do meu salário*]_{OI-DAT}.

É possível ainda que o *dativo* não tenha uma interpretação referencial, ao se constituir como lexicalmente não-expresso como se observa nos verbos como *acontecer*, *bastar*, *chegar*, *constar*, *convir*, *faltar*, *parecer*, *ser*.

Em (66) e (67), temos em evidência do *dativo de interesse* concretamente realizado especificando a entidade envolvida (N_2), ao passo que, em (68) e (69), estão não-expressos, especificidade do *dativo de interesse* que tende a focalizar o sujeito gramatical (N_1).

Apresentamos ainda, à luz de Berlinck (1996, p. 137), a representação esquemática de *verbos* que predicam o *dativo de interesse*.

[+/- animado] N₁ + V + {*a, para*} + [+ animado] N₂

- (66) a. Quem é que se importa com o que [*lhe*]_{OI-DAT} aconteceu?
 b. Quem é que se importa com o que aconteceu?
- (67) c. A questão [*me*]_{OI-DAT} pareceu algo acadêmica.
 d. A questão pareceu algo acadêmica.

Uma vez expostos os contextos os *verbos* propulsores do *dativo de interesse*, restamos apresentar os *verbos de movimento* e os *verbos de movimento psicológico*. Os *verbos de movimento* tais como os verbos *chegar, escapar, entrar, fugir, ir* e *vir* normalmente atribuem um complemento locativo ou argumento direcional esquematicamente representado. Em sendo este complemento *dativo* [+ animado], pode ser marcado também com o traço [+ afetado], cf. ilustramos em (68) e (69), à luz de Berlinck (1996, p. 140).

[+/- animado] N₁ + V + *a* + [+ animado] N₂

- (68) Veio-[*lhe*]_{OI-DAT} uma necessidade enorme de sair.
 (69) Alguns erros de ortografia [*me*]_{OI-DAT} escaparam.

Os *verbos de movimento* projetam *complementos dativos* com uma leitura referencial quando são lexicalmente realizados, cf. Berlinck (1996, p. 140). Por outro lado, os contextos de dativo nulo projetados por verbos que projetam *dativo de interesse* recebem, como observado em (66b) e (67d), um valor dêitico de 1^a ou 2^a pessoa.

Ainda no âmbito das construções intransitivas, temos o *movimento psicológico* semanticamente veiculado por predicadores verbais tais como *agradar, importar, interessar, repugnar* e *satisfazer* que atuam na projeção do complemento dativo na perspectiva de Berlinck (1996, p. 141) O complemento dativo (N₂) atua como o “portador” do movimento psicológico relativo ao verbo e o sujeito gramatical (N₁), representando a *causa* ou a *fonte* desse processo, como projetado por Berlinck (1996, p. 141) em relação à representação

esquemática de tais verbos de *movimento psicológico*. Em (70) e (71), temos as formas verbais *agradou* e *importa* em contextos de *movimento psicológico* na geração dos complementos dativos.

[+/- animado] N₁ + V + a + [+ animado] N₂

(70) A inteligência do rapaz agradou [*a todos*]_{OI-DAT}.

(71) [*A eles*]_{OI-DAT} não importa como você utiliza o seu tempo de trabalho.

Berlinck (1996) ainda discute os contextos de *dativos de posse* e as *construções se/lhe*. Entretanto, estes não serão controlados nesta variável, recorte cujo motivo se reporta a Oliveira (2014, p. 49), que alerta ao fato que tais complementos dativos se manifestam sempre da mesma forma, não havendo variação a ser estudada. Este trabalho se propõe a estudar apenas os verbos que manifestam variação na composição de seus *complementos dativos*.

Considerando que o foco deste trabalho está voltado ao complemento dativo de 2SG, temos em cena as seguintes evidências, expostas de (72) a (78), acerca das semânticas dos predicadores verbais (*transferência material, transferência verbal e perceptual, movimento físico, movimento abstrato, interesse, movimento e movimento psicológico*) responsáveis por tal estrutura de complementação verbal.

(72) *Transferência material*: “[...] Emtão quero [*ti*]_{OI-DAT} pedi um quiti escolar da Bárbi porque eu sou apaichonada com ela [...]” (C092)

(73) *Transferência verbal e perceptual*: “[...] mas sim as coisas que escrevo [*para você*]_{OI-DAT} mesmo escolher para me dar [...]” (C36)

(74) *Movimento físico*: “Assim que chegar a Paris vou dar immediato cumprimento a rua promessa e levar-[*te*]_{OI-DAT}-ei uma imagem da N. D. des Victories [...]” (24.12.1907)¹⁴

(75) *Movimento abstrato*: “Papai Noel, desde já eu [*te*]_{OI-DAT} desejo um ótimo Natal e um próspero ano novo [...]” (C190)

(76) *Interesse*: “Quero [...] que voce tome muito cuidado com tudo o que vá [*te*]_{OI-DAT} acontecer ai [...]” (02.08.1978)¹⁵

(77) *Movimento*: “A estrada da tua Felicidade não parte das pessoas e das coisas para chegar [*a ti*]_{OI-DAT}, parte sempre de ti em direção aos outros” (C834)

¹⁴ Em (56), temos uma ocorrência de dativo de 2SG levantada e analisada por Oliveira (2014, p. 103).

¹⁵ Em (58), temos uma ocorrência de dativo de 2SG levantada e analisada por Oliveira (2014, p. 103).

(78) *Movimento psicológico*: “Se você me de esse presente eu [Ø]_{OI-DAT} agradeço! Se não puder, eu [Ø]_{OI-DAT} agradeço do mesmo jeito!” (C027)

2.5.3.5 Os papéis semânticos do dativo

Company Company (2006, p. 520) descreve cinco papéis semânticos do *dativo*. São eles: *receptor*, *experienciador*, *possuidor*, *meta-fonte* e *beneficiário*. A autora entende que sua categorização está envolvida com o grau de envolvimento do *dativo* em relação ao evento que o verbo expressa, criando uma gradação em que o *receptor* seria o *dativo* com maior nível de envolvimento, seguido pelo *experienciador* e pelo *possuidor*, ao passo que os papéis semânticos de *dativo* com menor envolvimento seriam o *meta-fonte* e o *beneficiário*. A autora entende que, por vezes, a fronteira entre tais papéis semânticos pode ser tênue, sendo possível que um mesmo *dativo* assinale dois papéis semânticos ao mesmo tempo. Company Company (2006, p. 520) admite os papéis semânticos do OI ¹⁶“compartilham e estão vinculados pelo mesmo valor semântico geral de *locus* ou *meta* final da transitividade, que tem [...] esta categoria funcional como significado básico”. Passamos à descrição dos contextos de cada um desses papéis semânticos do OI (dativo) em conformidade com Company Company (2006).

O *receptor*, que é a entidade que se envolve no processo descrito pelo verbo de tal maneira que passa a ser possuidor do *acusativo*, segundo a autora, o *receptor* “recebe ou se converte no novo possuidor do objeto direto (*acusativo*)” (COMPANY COMPANY, 2006, p. 520). Este papel semântico é considerado o prototípico do *dativo* e é o mais frequente como é possível observarmos em verbos como *dar*, *outorgar*, *vender*, *comprar*. O *receptor* tem consciência e volicionalidade quanto à transferência, que pode ser física, como ocorre no exemplo (79) ou metafórica, como em (80). Há uma relação inversamente proporcional entre a agentividade do sujeito e a volicionalidade do *receptor*, quanto maior a primeira, menor será a segunda. Company Company (2006, p. 521) dá os seguintes exemplos para o papel semântico de *receptor*:

¹⁶ “Todos los papeles semánticos del OI comparten y están vinculados por el valor semántico general de *locus*, o *meta* final de la transitividad, que [...] tiene esta categoria funcional como significado básico” (COMPANY COMPANY, 2006, p. 520).

- (79) “Desi dio cartas et los presentes que traía [*a cada uno de aquellos reyes*]_{DAT}”
 “Desi deu as cartas e os presentes que trouxe [*para cada um daqueles reis*]_{DAT}”

- (80) “Pues consejo [*vos*]_{DAT} que non me lo tomedes”
 “Pois [*vos*]_{DAT} aconselho a não tirar isso de mim”

O papel semântico de *experienciador* do complemento dativo manifesta-se normalmente em situações estáticas avessa ao deslocamento e/ou transferência, expondo uma alteração de seu estado anímico, mental ou sensorial em razão do evento expresso pelo verbo. O *experienciador* prevalece nas predicções de *verbos psicológicos* tais como *pesar, gostar, assombrar, entediar e parecer* e de *experiência existencial* tais como *faltar, ficar e acontecer*, cf. Company Company (2006, p. 521). A volicionalidade do *experienciador* é menor que do *receptor*, mas maior que do *meta-fonte* e do *beneficiário*. Em (81) e (82), trazemos à cena a exemplificação fornecida por Company Company (2006, p. 521/522) para verbos *experienciadores* como predicadores do OI (dativo).

- (81) “que tiempo [*le*]_{DAT} quedará [*a v.s.*]_{DAT} para que le veamos, cuando yo torne a Toledo”
 “quanto tempo restará [*a v.s.*]_{DAT} para que possamos lhe ver, quando eu voltar a Toledo”
- (82) “y entre estos dime y dírete... [*le*]_{DAT} iba entrando sueño al arqueólogo”
 “e entre discussões... o arqueólogo ia-*[lhe]*_{DAT} alcançando o sono”.

O papel semântico de *possuidor* se manifesta quando o *dativo* se mostra como possuidor do *acusativo* independente da transitividade do verbo, o que o diferencia do *receptor*, uma vez que inexiste a ideia de transferência. Company Company (2006, p. 521) esclarece que um alto nível de afetação do *acusativo* se dá em razão da transitividade do verbo, da mesma forma que a natureza de posse inerente a essa relação também resulta em uma afetação do *dativo* como é possível observarmos em (83) e (84).

- (83) “porque le ahogaba el placer y [*le*]_{DAT} cortaba la voz em la garganta”
 “porque afogou seu prazer e cortou-*[lhe]*_{DAT}/sua voz na garganta”
- (84) “que [*me*]_{DAT} atraviesa las sienes em em clavo”
 “que [*me*] fura as têmporas com um prego”

O papel de *meta-fonte* pode ser acionado se o complemento dativo for *meta* a que algo ou alguém chega e se o dativo for *fonte* da qual a entidade em deslocamento ou em transferência procede. Company Company (2006, p. 521) entende que o *dativo meta/fonte* não exerce volicionalidade, motivo pelo qual não pode ser receptor como pode ser observado em relação aos prototípicos *chegar* (dativo-meta), *vir* (dativo-meta), *sair* (dativo-meta) e *enviar* (dativo-fonte). Em (85) e (86), ilustramos evidências de estruturas de complementação verbal projetadas pelos verbos *enviar* e *abordar* e os seus respectivos complementos dativos nas formas de dativo-meta e dativo-fonte.

(85) “Y ansi escrito a Sevilla dos o tres cartas para que las enbien [*a vuestra merçed*]_{DAT}”

“E escrevi duas ou três cartas a Sevilha para que as enviem [*a vossa mercê*]_{DAT}”.

(86) “el acólito, de puntillas, sin hacer ruido, se había acercado por detrás [*al Provisor*]_{DAT}”.

“o acólito, na ponta dos pés, sem fazer ruído, abordou por trás [*ao Provisor*]_{DAT}”.

“o acólito, na ponta dos pés, sem fazer ruído, abordou-[*lhe*]_{DAT} por trás”.

Company Company (2006, p. 522) discute ainda o papel semântico de *beneficiário* como “a entidade que recebe para seu proveito o significado da oração”, assumindo o dativo cujo papel é o de *beneficiário* com o seu alto grau de afetação em relação ao evento expresso pelo verbo. À título de ilustração, apresentamos os dados expostos em (87) e (88) como evidência do complemento dativo com o papel semântico de *beneficiário*. No que se refere ao grau de envolvimento do complemento dativo, tendo em vista o evento que o verbo expressa, Company Company (2006) entende especificidades da semântica do dativo que o evidenciam com maior envolvimento (*receptor, experienciador e possuidor*) e menor envolvimento (*meta-fonte e o beneficiário*).

(87) “...cuemo aprendi que las nuestras reynas lo adobavan [*a los nuestros reyes*]_{DAT}”.

“... como aprendi que as nossas rainhas [os] adoravam [*aos nossos reis*]_{DAT}”.

(88) “entregaba ceniza limpia y tamizada, de que se hacía consumo a causa de la gran cantidad de candelabros necesarios para las velas que ardían [*a los diversos santos que había en la casa*]_{DAT}”.

“entregava cinza limpa e peneirada, que foi consumida por causa da grande quantidade de candelabros necessários para as velas que ardiam [*aos diversos santos que havia na casa*]_{DAT}”.

Passamos às evidências de dativo de 2SG, atentando especificamente aos papéis semânticos levantados nas cartas em análise.

(89) *Receptor*: “[...] e sempre que ouiver o espirito, do Natal, lá vai está O menino Jesus [*te*]_{DAT} dando muita Saúde” (C103)

(90) *Experienciador*: “Desde já agradeço de Coração Que Deus te abençoe e não deixa Nada faltar [*para você*]_{DAT}” (C545)

(91) *Meta-fonte*: “Papai noel [*Te*]_{DAT} agradeço também por fazer muitas alegrias brigado de coração e espero que o senhor faça a minha também” (C318)

(92) *Beneficiário*: “[...] todo ano faço cartas [\emptyset]_{DAT} mas nunca recebi então espero esse ano você realize os meus pedidos [...]” (C852)

2.5.3.6 As seções das cartas

A primeira variável extralinguística em análise é a *seção da carta* em que ocorre o *complemento dativo* de 2SG. Voltar a atenção para a estrutura formal do gênero textual *carta* pode ser relevante à análise de uma dada regra variável, uma vez que é preciso discernir o que pode ser específico do gênero *carta* em função dos seus traços formulaicos (¹⁷tradição discursiva) e o que pode ser expressão vernacular. Preferimos assumir a configuração estrutural da carta marcada pela *saudação inicial com captação de benevolência*, o *núcleo da carta*, a *seção de despedida* e o possível *post-scriptum* (observações), buscando marcar a parte específica da carta em que se dá o dado linguístico de estrutura dativa de 2SG nas cartas ao PN. Conjecturamos que as seções inicial e final da carta possam ser espaços propícios às fórmulas/estruturas cristalizadas em virtude dos cumprimentos (iniciais e finais) e despedidas, atuando assim como instâncias do texto possivelmente fomentadoras de formas variantes específicas, em conformidade também com Oliveira (2014, p. 105). Em (93) e (94), temos evidências formas pronominais dativas de 2SG em contextos típicos da captação da

¹⁷ “Entendemos por Tradição Discursiva (TD) a repetição de um texto ou de uma forma textual ou de uma maneira particular de escrever ou falar que adquire valor de signo próprio (portanto é significável). Pode-se formar em relação a qualquer finalidade de expressão ou qualquer elemento de conteúdo, cuja repetição estabelece uma relação de união entre atualização e tradição; qualquer relação que se pode estabelecer semioticamente entre dois elementos de tradição (atos de enunciação ou elementos referenciais) que evocam uma determinada forma textual ou determinados elementos linguísticos empregados.” (KABATEK, 2006, p. 512 *apud* Oliveira, 2014, p. 105).

benevolência, no início da carta. Expomos de (95) a (97) evidências dos dados pronominais dativos em análise nos contextos de *núcleo da carta*, *despedida* e *post-scriptum*.

(93) *Saudação inicial com captação de benevolência*: “Querido papai Noel Saudações: Escrevo-[*lhe*]_{DAT} está para [*lhe*]_{DAT} pedir presentes que sonho com eles sempre porque a minha mãe [...] e o meu padrasto [...] não tem condição pois ganham pouco.” (C923)

(94) *Saudação inicial com captação de benevolência*: “Querido Papai Noel como vai Você meu nome e [...] tenho 12 anos e estou [*lhi*]_{DAT} escrevendo por que gostaria de ganhar umas roupas neste natal” (C922)

(95) *Núcleo da carta*: “Meu nome é [...] e estou na 4ª série estudo na Escola Municipal Rui da Costa Val. Quero [*lhe*]_{DAT} pedir um par de roupas da ADIDAS” (C197)

(96) *Despedida*: “[...] Obrigado por ler a minha cartinha que Deus te abençoe e [*te*]_{DAT} dê em dobro” (C905)

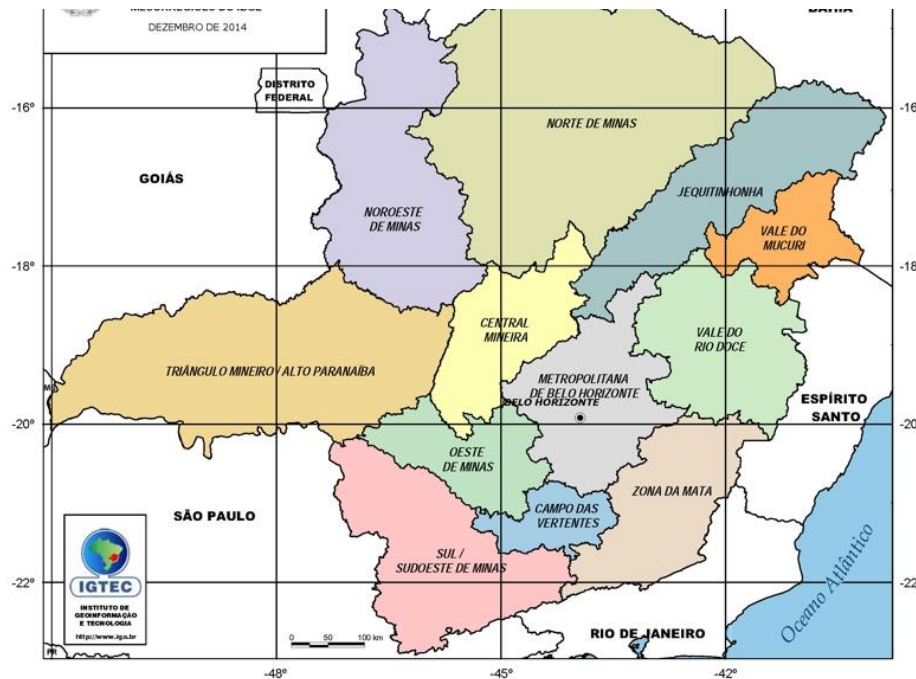
(97) *Post-scriptum*: “Prometo-[*o*]_{DAT} até lá ser Paciente obrigado. PAPAÍ NOEL” (C204)

Passamos à análise da variável *procedência da carta dentro do espaço geográfico mineiro*, tendo sempre em vista o fato de que nem sempre os redatores especificam o local e a data na forma de um cabeçalho.

2.5.3.7 A procedência das cartas no espaço geográfico mineiro

Com o intuito de mapearmos a produtividade das formas pronominais dativas de 2SG no espaço geográfico mineiro, tomamos por base a divisão proposta pelo Instituto Brasileiro de Geografia Estatística (imagem 3) em relação às mesorregiões do Estado de Minas Gerais quais sejam as seguintes: *Noroeste de Minas*, *Norte de Minas*, *Jequitinhonha*, *Vale do Mucuri*, *Triângulo Mineiro e Alto Paranaíba*, *Central Mineira*, *Metropolitana de Belo Horizonte*, *Vale do Rio Doce*, *Oeste de Minas*, *Sul e Sudeste de Minas*, *Campos das Vertentes* e *Zona da Mata*.

Imagem 3: Mesorregiões de Minas Gerais propostas pelo IBGE



Fonte: Instituto de Geoinformação e Tecnologia (Minas Gerais, 2014).

Para evitar dúvidas ou erros na categorização, levaremos em consideração, nesta pesquisa, as mesorregiões do Estado e as respectivas cidades que as compõem oficialmente expostas pelo Governo do Estado de Minas Gerais¹⁸. Tendo em vista o fato de que a maioria das cartas foram escritas na Região Metropolitana de Belo Horizonte, mais especificamente na Microrregião de Belo Horizonte, identificamos especificamente para esta microrregião, os municípios de *Contagem*, *Betim*, *Ribeirão das Neves*, *Sabará* e *Nova Lima*. O objetivo é o de fornecermos um panorama mais abrangente em relação à distribuição dos pronomes-complemento dativos de 2SG em relação ao local específico de Minas Gerais de que são provenientes.

Por fim, cumpre assinalar que os dados quanto à procedência das *cartas* serão coletados com base nas informações prestadas pelos próprios escreventes. Em algumas cartas essa informação vem já no cabeçalho como reflexo do respeito às seções específicas do gênero textual “Carta”. Contudo, para aquelas missivas cujo local de escrita foi omitido no

¹⁸ <https://www.mg.gov.br/sites/default/files/paginas/arquivos/2016/ligminas_10_2_04_listamesomicro.pdf>

texto da carta, recorreremos ao seu respectivo envelope de correspondência, em que inevitavelmente consta o endereço do remetente.

Síntese do capítulo

Neste capítulo, voltamos o foco aos encaminhamentos em relação aos aspectos teórico-metodológicos da Sociolinguística Laboviana.

Passamos à discussão acerca do método de trabalho com cartas, justificando-o e descrevendo os desafios de um estudo linguístico embasado em testemunhos da língua escrita. Na sequência, voltamo-nos à apresentação das *Cartas ao PN*, apontando algumas particularidades e justificando a sua adoção. Ainda no âmbito do método de organização e de trabalho com dados de uma regra variável, passamos à descrição das potencialidades do RBrul na interface do RStudio para a geração, no nível da rodada multivariacionista, os índices probabilísticos da regra variável em análise nos eixos *step-up & step-down*. Por fim, passamos à descrição da regra variável eneária em si (estratégias dativas de 2SG), pormenorizando não só a variável dependente (*te, lhe, você, a você, para você, zero*), mas também as variáveis independentes controladas em termos de contextos linguísticos (*o contexto de sujeito de 2SG, o tipo de predicador em relação à estrutura argumental da sentença, a forma do objeto direto nas estruturas ditransitivas, o valor semântico do predicador, o papel semântico do dativo*) e extralinguísticos (*a seção das cartas e a sua procedência no espaço geográfico mineiro*).

Uma vez expostos os pressupostos teórico-metodológicos condutores desta dissertação, passamos, no capítulo 3, à descrição dos resultados gerais das estratégias dativas de 2SG em termos percentuais.

CAPÍTULO 3 ANÁLISE GERAL: RESULTADOS GERAIS EM TERMOS PERCENTUAIS

Neste capítulo, pretendemos apresentar os resultados de uma rodada geral feita através do Excel através do qual foi possível chegar aos índices percentuais de produtividade de uma regra variável cuja variável dependente é eneária (*te, lhe, a ti/para ti, a você, você e zero*) em relação às formas pronominais dativas de 2SG. Para isso, subdividimos este capítulo em duas segmentos principais orientados para as análises dos fatores linguísticos (*o tipo sintático do predicador em relação à estrutura argumental da sentença, a forma do objeto direto nas estruturas ditransitivas, o valor semântico do predicador, o papel semântico do dativo*) e extralinguísticos (*a seção das cartas e a procedência das cartas no espaço geográfico mineiro*).

Tendo em vista o fato de que, neste capítulo, traremos à discussão, em termos percentuais, as estratégias pronominais dativas de 2SG (tabela 1), conjecturamos que as formas pronominais dativas do paradigma de *você* (principalmente *a você, para você*) seriam as mais produtivas, uma vez que se trata de cartas produzidas no espaço mineiro em que prevalece o subsistema de *você*-sujeito de 2SG, cf. Scherre *et alii* (2015).

Tabela 1: Estratégias pronominais dativas de 2SG produtivas nas cartas brasileiras

ESTRATÉGIAS PRONOMINAIS DATIVAS DE 2SG							
te	lhe	a ti	para ti	a você	para você	você	zero
415/1232	132/1232	1/1232	1/1232	23/1232	59/1232	7/1232	594/1232
33,69%	10,71%	0,08%	0,08%	1,87%	4,79%	0,73%	48,14%
1232							

De um modo geral (tabela 1), observamos que estão em concorrência mais acirrada o *zero* (48,14%, **594**/1232) e o *te* (33,69%, **415**/1232), assumindo as formas *lhe* (10,71%, 132/1232), *para você* (4,79%, 59/1232), *a você* (1,87%, 23/1232), *você* (0,73%, 7/1232), *para ti* (0,08%, 1/1232) e *a ti* (0,08%, 1/1232) baixos índices de produtividade nas cartas ao PN analisadas, cf. ilustramos de (98) a (105). Considerando que o *zero* e o *te* mostram-se

como a 1ª e a 2ª formas dativas mais produtivas, respectivamente, observamos o clítico *lhe* como a 3ª estratégia dativa mais produtiva das cartas ao PN.

(98) “[...] Papai Noel gostaria de [*te*]_{DAT} pedir uma boneca de natal [...]” (C012)

(99) “[...] Papai Noel se não for [\emptyset]_{DAT} pedir muito queria a minha boneca e um carrinho para o meu irmãozinho [...]” (C021)

(100) “[...] Eu gostaria com muito amor de [*lhe*]_{DAT} pedir um material escolar [...]” (C052)

(101) “[...] neste natal quero pedir [*para você*]_{DAT} uma cesta de natal [...]” (C918) para você

(102) “[...] Nessa véspera de natal eu gostaria de pedir [*a você*]_{DAT} que esta lendo a minha carta, um “notebook” ou um “tablet” [...]” (C152)

(103) “[...] Eu gostaria de pedir [*você*]_{DAT} papai Noel um tablet [...]” (C098)

(104) “[...] A estrada da tua Felicidade não parte das pessoas e das coisas para chegar [*a ti*]_{DAT} [...]” (C834)

(105) “Papai Noel Gosto de ti, [...] por isso eu estou dizendo [*para ti*]_{DAT} que gosto muito de presentes... te adoro” (C961)

3.1 Os contextos linguísticos: o tipo sintático de predicador, a forma do objeto direto nas estruturas ditransitivas, o valor semântico do predicador, o papel semântico do dativo

Para controlarmos a produtividade da regra variável em análise em função de uma variável dependente enérgica (*te, lhe, a você, para você, a ti, para ti, zero*), passamos à descrição da distribuição dos dados de dativo de 2SG em relação aos parâmetros linguísticos (o tipo sintático de predicador verbal, a forma do objeto direto nas estruturas ditransitivas, o valor semântico do predicador, o papel semântico do dativo), conforme as hipóteses enunciadas para cada variável independente (linguística) no capítulo 2.

3.1.1 A correlação entre as estratégias dativas e nominativas de 2SG

O controle das estratégias pronominais dativas de 2SG correlacionadas ao contexto de referência ao sujeito também de 2SG é motivado pelo intuito de averiguarmos a questão da famigerada “uniformidade tratamental”, uma utopia linguística no PB ainda legitimada pela tradição gramatical. Ao correlacionarmos os contextos de nominativo e dativo (tabela 2), temos o objetivo de testar a hipótese de produtividade da combinação entre os contextos de *você*-sujeito e de complemento dativo (*te, a ti, para ti, lhe, a você, para você*) em consonância

com os resultados das análises embasadas em dados de sincronias passadas, produzidos por redatores, em sua maioria, letrados (LOPES & CAVALCANTE, 2012; LOPES *et alii*, 2018; RUMEU, 2020) que já evidenciam a confluência entre as escolhas nominativas e dativas vinculadas aos paradigmas formais de *tu* (*te, a ti, para ti*) e *você* (*lhe, a você, para você*).

Tabela 2: Estratégias pronominais dativas de 2SG correlacionadas ao sujeito de 2SG

SUJEITO DE 2SG	ESTRATÉGIAS PRONOMINAIS DATIVAS DE 2SG								Total
	te	lhe	a ti	para ti	a você	para você	você	zero	
Cartas de Você- sujeito exclusivo	165/416 (30,64%)	41/416 (9,86%)	1/416 (0,24%)	1/416 (0,24%)	14/416 (3,37%)	27/416 (6,49%)	3/416 (0,72%)	164/416 (39,42%)	416/ 1232 (33,63%)
Cartas de FNT exclusivo	250/816 (39,66%)	91/816 (11,15%)	-	-	9/816 (1,10%)	32/816 (3,92%)	4/816 (0,49%)	430/816 (52,70%)	816/ 1232 (66,37%)

Nas cartas ao PN, mostram-se bem mais produtivas as evocações ao interlocutor com formas nominais de tratamento (“Meu querido Papai Noel”, “Querido Papai Noel”, “Querido Papai Noel dos Correios” etc.) do que com o *você*, o que constatamos através dos índices de produtividade, respectivamente, de 66,37% (816/1232) e 33,63% (416/1232), expostos na tabela 2. Voltando o foco especificamente para as cartas em que o redator optou por evocar o seu interlocutor através de uma FNT, observamos que o *zero* predomina, em 52,70% (430/816), ainda que em concorrência com o *te* que se deixa evidenciar, por outro lado, em 39,66% dos dados (250/816). Em (106) e (107), temos em análise dados de *zero* e *te* em contexto de ausência de forma pronominal (*tu* ou *você*) para fazer referência ao sujeito de 2SG, optando o redator por uma FNT.

(106) Carta de FNT: “[...] Quero deixar [Ø]_{DAT} meus votos de felicidades Um feliz natal e um próspero Ano novo [...]” (C010)

(107) Carta de FNT: “[...] Porque isso Papai noel eu [*te*]_{DAT} pesso uma material escolar [...]” (C031)

Nas cartas de *você-sujeito*, o *te* prevaleceu, em 30,46% (165/416), em concorrência mais acirrada com o *zero-dativo*, em 39,42% (164/416), o que vai ao encontro dos resultados de pesquisas linguísticas que também evidenciam o *te-dativo* com altíssima produtividade no contexto de *você-sujeito*, cf. Oliveira (2014), Lopes & Marcotulio (2019, p. 152) para as cartas cariocas, Rumeu (2015) para as cartas mineiras, Rumeu & Oliveira (2016) para as cartas amorosas, Cardoso (2017) e Rumeu (2020) para as cartas mineiras, como evidências, em distintas amostras, do *te* como um contexto de resistência do *tu* no PB. Em (108) e (109), ilustramos o *te* e o *zero* em contexto de *você-sujeito*. Não temos, nas cartas de FNT, uma concorrência entre *zero* (52,70%) versus *te* (39,66%) tão acirrada assim como a que é possível verificarmos, nas cartas de *você-sujeito*, com o *te*, em 30,64% versus o *zero*, em 39,42% dos dados. Ao evidenciarmos a produtiva correlação entre o *você-sujeito* e *te-dativo*, é possível justificarmos tal fato também em virtude de se tratar de formas que, quanto à pessoa semântica, são idênticas, uma vez que remetem ao interlocutor sempre [-EU], cf. Lopes & Rumeu (2007). Nas cartas de FNT, observamos que prevalece o *zero*, em 52,70%, ainda que em alternância com o *te*, em 39,66% dos dados, o que pode estar vinculado ao fato de o *zero dativo* não estar formalmente relacionado a um paradigma específico de formas de *tu* e de *você*, cf. também discutido por Cardoso (2017, p. 101) em relação às cartas mineiras (séculos XIX e XX) e também por Oliveira (2014) em relação às cartas cariocas. Acrescenta-se ainda o fato de o *zero-dativo* ter se mostrado a estratégia mais produtiva nas cartas ao PN, em 48,14% dos dados (594/1232), ainda que tenha, nas cartas de FNT (52,70%), prevalecido e se mantido, nas cartas de *você-sujeito*, em forte concorrência (39,42%) com o *te-dativo* (30,64%).

(108) Carta de *você-sujeito*: “[...] se *você* me dar eu [*te*]_{DAT} agradeceria muito se *você*. Ler a minha Cartinha [...]” (C008)

(109) Carta de *você-sujeito*: “[...] Se *você* me de esse presente eu [\emptyset]_{DAT} agradeço! Se não puder, eu [\emptyset]_{DAT} agradeço do mesmo jeito! (C027)

É interessante também atentarmos aos dados de *você* que, no lugar de prep+*você*, oferecem ao predicador verbal a complementação dativa de 2SG. Trata-se de 7 (sete) ocorrências distribuídas entre as cartas de *você-sujeito* (3/416, 0,72%) e as cartas de FNTSs (4/816, 0,49%).

(110) Carta de *você-sujeito*: “[...] se *você* tiver nas condisoes quero pedi [*você*]_{DAT} que meda o kit completo [...]” (C364)

(111) Carta de *você-sujeito*: “[...] papai noel eu queria pidir [*você*]_{DAT} um presente se *você* poder me dar eu gostaria de ganhar um rádio [...]” (C531)

(112) Carta de *você-sujeito*: “[...] Papai nolu Ø escrevo esta carta pra pediz [*você*]_{DAT} um vide o geme supe nitendo -2 [...]” (C207)

(113) Carta de FNT: “[...] Eu gostaria de pedir [*você*]_{DAT} *papai noel* um tablet [...]” (C098) ok

(114) Carta de FNT: “[...] *Papai nolu* escrevo esta carta pra pediz [*você*]_{DAT} um vide o geme supe nitendo -2 [...]” (C207)

(115) Carta de FNT: “[...] *Querido papai Noel*, é com muita emoção que lhe escrevo estás puças linhas para contar um pouco de minha vida [...] aproveito para desejar [*você*]_{DAT} e sua família um feliz natal e um ano novo cheio de paz, amor e realizações [...]” (C448)

(116) Carta de FNT: “*Papai Noel*, os meus pais moram na casa dos meus avos de favor, e eles não tem condições para comprar [...] por isso estou pedindo [*você*]_{DAT} uma bicicleta [...]” (C823)

3.1.2 As estratégias dativas de 2SG correlacionadas aos padrões sintáticos

Ao correlacionarmos os pronomes-complemento dativos de 2SG aos padrões sintáticos que os projetam, levantamos evidências de verbos que projetam o dativo através de grades argumentais de dois (S V OI) e de três (S V OD OI) lugares (tabela 3). A ideia é controlarmos o nível de inserção do *você* no sistema pronominal do PB em relação aos padrões de organização sintática descritos por Mateus *et alii* (2003), buscando identificar se o padrão de 2 lugares (S V OI) tenderá (ou não) a impulsionar o *te-dativo* como observado por Oliveira (2014, p. 240) em relação às cartas cariocas (.948).

Tabela 3: Estratégias pronominais dativas de 2SG correlacionadas aos padrões sintáticos dos verbos

PADRÕES SINTÁTICOS ¹⁹	ESTRATÉGIAS PRONOMINAIS DATIVAS DE 2SG								
	te	lhe	a ti	para ti	a você	para você	você	zero	total
Verbos de 2 lugares (S V OI)	58/187 (31,02%)	16/187 (8,56%)	1/187 (0,53%)	-	4/187 (2,14%)	-	-	108/187 (57,75 %)	187/ 1232 (15,50%)
Verbos de 3 lugares (S V OD OI)	357/1045 (34,16%)	116/1045 (11,10%)	-	1/1045 (0,10%)	19/1045 (1,82%)	59/1045 (5,65%)	7/1045 (0,67%)	486/1045 (46,51%)	1045/ 1232 (84,50%)

Para os verbos de 2 lugares (S V OI), observamos, com base na análise da tabela 3, a alternância mais intensa entre o *zero* e o clítico *te*, que com índices percentuais de **57,75 %** (108/187) e **31,02%** (58/187), permitem-nos constatar a prevalência do *zero*-dativo em contextos verbais que predicam 2 argumentos (S e OI), como está ilustrado em (117) e (118). Ainda no que se refere ao contexto de predicadores verbais de 2 lugares, verificamos não só a baixa produtividade do prototípico *lhe* (8,56%, 16/187), mas também as ínfimas frequências de uso das formas *a você* (2,14%, 4/187) e *a ti* (0,53%, 1/187) cujas evidências estão expostas de (119) a (121). Em relação aos verbos de dois lugares (S V OI), observamos a prevalência do verbo *agradecer* no sentido de *agradecer a...*, *agradecer-lhe*, o que nos levou a interpretá-lo como verbo *transitivo indireto* também em consonância com a de Luft (2006, p. 44), ainda que tenhamos a interpretação ditransitiva a tal verbo (“agradecer (-*lhe*) algo”) como descrito também por Luft (2006)²⁰. Observemos ainda o fato de o verbo *agradecer* ter admitido uma variedade de formas dativas de 2SG (*zero*, *te*, *lhe*, *a você* e *a ti*), o que também foi observado por Cardoso (2017, p. 130) nas cartas mineiras (séculos XIX e XX).

(117) “[...] desde já [∅]_{DAT} agradeço! [...]” (C015)

¹⁹ Padrões sintáticos propostos por Duarte (2006, p. 277-321) que permitem esquemas relacionais corporificados através das funções sintáticas de SU/S = sujeito, V = verbo, OD = objeto direto, OI = objeto indireto, OBL = oblíquo.

²⁰ “AGRADECER **TD(I)**: agradecer(-lo (a...)) (OBS.); agradecer (-*lhe*) algo. **Int** ou **TI**: agradecer (a...), agradecer(-*lhe*). Dar graças ou agradecimentos; mostrar-se ou confessar-se grato; expressar ou manifestar gratidão: “Agradeça a Deus a sua salvação” (Nascentes)” (LUFT, 2006, p. 44)

- (118) “[...] se você me dar eu [te]_{DAT} agradeceria muito se você. Ler a minha Cartinha [...]” (C008)
- (119) “[...] Eu [lhe]_{DAT} agradeço de todo coração [...]” (C263)
- (120) “[...] agradeço [a voce]_{DAT} porque neus pais ganham pouco [...]” (C817)
- (121) [...] A estrada da tua Felicidade não parte das pessoas e das coisas para ²¹chegar [a ti]_{DAT} [...] (C834)

Nos contextos dos verbos ditransitivos, observamos a alternância mais acirrada entre o *zero* versus *te*, considerando que, mais uma vez, o *zero* (486/1045, **46,51%**) prevalece em relação ao *te* (357/1045, **34,16%**), conforme exemplificamos em (122) e (123). Esses resultados para os verbos ditransitivos opõem-se aos de Oliveira (2014), que observou para as cartas cariocas a prevalência do *te* (57,4%) em alternância com o *zero* (23,4%), e os de Cardoso (2017, p. 116), que com base nas amostras de cartas mineiras, constatou a alternância *lhe* (42%) versus *te* (27%) versus *zero* (24%). Nas cartas em análise, observamos, por outro lado, a prevalência do *zero*, o que conforme Oliveira (2014, p. 113), considerando que, em sua amostra, o *zero-dativo* mostra-se como a segunda forma mais produtiva nos ditransitivos, poderia referendar “uma possível hipótese de simplificação da estrutura verbal que se automatiza com um argumento externo e um interno apenas, daí o apagamento do OI.” Mostraram-se nas cartas em análise, ainda que em distintos níveis, as formas *lhe* (116/1045, 11,10%), *para você* (59/1045, 5,65%), *a você* (19/1045, 1,82%), *você* (7/1045, 0,67%) e *para ti* (1/1045, 0,10%), cf. ilustramos de (124) a (128).

- (122) [...] Eu [Ø]_{DAT} escrevi está carta para te pedir um micro-ondas [...] (C115)
- (123) “[...] eu [te]_{DAT} envei²² está carta pois Eu quero muito a casa da Barbie. [...]” (C016)
- (124) “[...] Eu gostaria com muito amor de [lhe]_{DAT} pedir um material escolar [...]” (C052)
- (125) “[...] Olá Papai Noel tenho 12 anos e escrevo [para você]_{DAT} para te pedi uma cesta de chocolate e material escolar [...]” (C073)
- (126) “Desejo [a você]_{DAT} um feliz natal.” (C021)
- (127) “[...] Eu gostaria de pedir [você]_{DAT} papai noel um tablet [...]” (C098)

²¹ “CHEGAR 2. TI: chegar a... Alcançar; atingir: A escada não chega ao teto. Sua mente não chega a esse nível de abstração.” (LUFT, 2006, p. 216)

²² Leia-se “envei” por “enviei”.

(128) “Papai Noel Gosto de ti, [...] por isso eu estou dizendo [para ti]_{DAT} que gosto muito de presentes... te adoro” (C961)

3.1.3 A forma do objeto direto em estruturas ditransitivas de complementação verbal

Controlamos a forma do objeto direto em estruturas verbais ditransitivas com o intuito de perceber se o objeto direto (acusativo) interferiria, em algum nível, em relação às estruturas dativas de 2SG. A ideia é testarmos a hipótese pensada por Oliveira (2014) e testada também por Cardoso (2017, p. 121) em relação ao fato de que a “existência do objeto direto nesse contexto morfossintático possa condicionar qual variante do dativo irá ocorrer”.

Tabela 4: Estratégias pronominais dativas de 2SG correlacionadas ao OD em estruturas ditransitivas

FORMAS DO OD (S V OD OI)	ESTRATÉGIAS PRONOMINAIS DATIVAS DE 2SG							
	te	lhe	para ti	a você	para você	você	zero	total
Sintagma nominal (SN)	276/746 (37%)	73/746 (9,79%)	-	15/746 (2,01%)	39/746 (5,23%)	6/746 (0,8%)	337/746 (45,17%)	746/1042 (71,59%)
Sintagma oracional	33/100 (33%)	16/100 (16%)	1/100 (1%)	1/100 (1%)	5/100 (5%)	1/100 (1%)	34/100 (43%)	100/1042 (9,6%)
OD nulo	46/196 (23,47%)	27/196 (13,78%)	-	3/196 (1,53%)	14/196 (7,14%)	-	106/196 (54,08%)	196/1042 (18,81%)

O padrão verbal ditransitivo projeta os argumentos acusativo e dativo que podem vir comumente, na variedade portuguesa do português, combinados (*mo, to, lho, no-lo, vo-lo*). No PB, por outro lado, a combinação entre os complementos acusativo e dativo não é produtiva, o que motiva que controlemos, neste trabalho, a possível interferência do OD na expressão da estratégia dativa de 2SG. Perseguimos a hipótese pensada por Oliveira (2014, p. 104) em relação ao fato de que “se a estratégia de realização do OD é uma forma “pesada” – um sintagma oracional, por exemplo – a concretização do dativo se daria por outra mais

“leve” (um clítico, por exemplo).” Nesta análise, levantamos o OD que se mostra, nas estruturas ditransitivas das cartas ao PN, como *sintagma nominal*, *sintagma oracional* e *objeto nulo*. Nas cartas ao PN, observamos, de um modo geral (tabela 4), a preferência pelo OD que se mostrou como SN, em 71,59% dos dados (746/1042), seguido pelo OD nulo e pelo sintagma oracional em, respectivamente, 18,81% (196/1042) e 9,6% (100/1042) dos dados.

O complemento acusativo na forma de um SN prevaleceu nos contextos de *zero* e *te* em, respectivamente, 45,17% (337/746) e 37% (276/746) dos dados, cf. ilustramos em (129) e (130). As demais estratégias dativas *lhe*, *para você*, *a você* e *você* mostraram-se em, respectivamente, 9,79% (73/746), 5,23% (39/746), 2,01% (15/746) e 0,8% (6/746) dos dados, conforme exemplificamos de (131) a (134).

Zero dativo em contexto de OD na forma de SN:

(129) “[...] Todo ano [Ø]_{DAT} pesso [uma cesta de chocolates]_{SN-OD}, mas nunca ganho [...] (C035)

O clítico *te* dativo em contexto de OD na forma de SN:

(130) “[...] sou uma menina muito feliz e eu queria [*te*]_{DAT} pedir [umas coisas]_{SN-OD} um material escolar do rebeldes [...]” (C002)

O clítico *lhe* dativo em contexto de OD na forma de SN:

(131) “[...] que Deus [*lhe*]_{DAT} der [bastante Saúde, força e coragem]_{SN-OD} para pode a judar muitas crianças que necessita [...]” (C010)

O SPrep *para você* dativo em contexto de OD na forma de SN:

(132) “[...] olha papai noel já escrevir [vars cartas]_{SN-OD} [*para voce*]_{DAT} mais ainea não fui correspondida (C087)

O SPrep *a você* dativo em contexto de OD na forma de SN:

(133) “[...] Desejo [*a você*]_{DAT} [um feliz natal]_{SN-OD} (C021)

O *você* dativo em contexto de OD na forma de SN:

(134) “[...] Eu gostaria de pedir [*você*]_{DAT} papai noel [um tablet]_{SN-OD} [...] (C098)

Com o objeto direto nulo, em estruturas ditransitivas, observamos a prevalência do *zero-dativo*, em 54,08% dos dados (106/196), e dos clíticos *te* e *lhe*, em 23,47% (46/196) e 13,78% (27/196) dos dados, respectivamente, cf. ilustramos de (135) a (139). Com índices de percentuais baixíssimos, temos os SPreps *para você* e *a você* que se mostram produtivos em 7,14% (14/196) e 1,53% (3/196) dos dados nas cartas ao PN.

Zero dativo em contexto de OD nulo:

(135) “[...] A vovo disse também que era para mim caprichar, e [Ø]_{DAT} escrever porque eu mereço ganhar” (C090)

O clítico *te* em contexto de OD nulo:

(136) “[...] é com grande alegria que [te]_{DAT} escrevo para fazer um pedido [...]” (C210)

O clítico *lhe* em contexto de OD nulo:

(137) “[...] aqui quem [lhe]_{DAT} escreve querido papai noel e uma pessoa de 50 anos [...]” (C417)

O SPrep *para você* em contexto de OD nulo:

(138) “[...] Olá Papai Noel tenho 12 anos e escrevo [para você]_{DAT} para te pedi uma cesta de chocolate e material escolar [...]” (C073)

O SPrep *a você* em contexto de OD nulo:

(139) “[...] eu só posso pedir [a você]_{DAT} porque meus pais não tem condissões [...]” (C788)

Com o objeto direto na forma sintagma oracional, em estruturas ditransitivas, observamos a prevalência do *zero* e do *te*, em 43% (34/100) e 33% (33/100) dos dados, cf. verificamos de (140) a (142). Por outro lado, no contexto de OD como sintagma oracional, as formas *lhe*, *para você*, *a você*, *para ti* e *você* mostram-se em 16% (16/100), 5% (5/100), 1% (1/100), 1% (1/100) e 1% (1/100), respectivamente, cf. exemplificamos de (143) a (147).

Zero dativo em contexto de OD na forma de sintagma oracional:

(140) “[...] Estou escrevendo a carta pra [Ø]_{DAT} dizer [que tirei notas otimas na aula e fui obediente em casa]_{OD}. or. [...]” (C021)

O clítico *te* em contexto de OD na forma de sintagma oracional:

(141) “[...] eu [*te*]_{DAT} peço. [que o senhor realize o meu sonhos]_{OD-or.} [...]” (C303)

(142) “[...] Então eu [*te*]_{DAT} peço [se tem como o senhor me dá uma cesta de chocolate]_{OD-or.} [...]” (C957)

O clítico *lhe* em contexto de OD na forma de sintagma oracional:

(143) “[...] estou *lhe* Escrevendo, esta pequena carta para [*lhe*]_{DAT} dizer [que tenho muita vontade de ganhar uma mochila de rodinha]_{OD-or.} [...]” (C095)

O SPrep *para você* em contexto de OD na forma de sintagma oracional:

(144) “[...] olha papai noel se eu falo [*para voce*]_{DAT} [que eu sou um garoto feliz]_{OD-or.}, eu estou mentindo [...]” (C205)

O SPrep *a você* em contexto de OD na forma de sintagma oracional:

(145) “[...] porção peço [*a você*]_{DAT} [que mim atenda neste natal]_{OD-or} [...]” (C482)

O SPrep *para ti* em contexto de OD na forma de sintagma oracional:

(146) “Papai Noel Gosto de ti, [...] por isso eu estou dizendo [*para ti*] [que gosto muito de presentes]_{OD-or...} te adoro” (C961)

O *você* em contexto de OD na forma de sintagma oracional:

(147) “[...] quero pedi [*você*]_{DAT} [que meda o kit completo] [...]” (C364)

De um modo geral, observamos que o *zero* (ausência de estratégia dativa) e o clítico *te* (forma mais leve) mostram-se como as mais produtivas, contextualizadas pelos contextos de OD como sintagma nominal (SN), OD nulo (ausência) e o sintagma oracional. Nesse sentido, constatamos que as formas mais pesadas (formas sintagmáticas (sintagmas nominal e oracional)) mostram-se frequentes nos contextos de dativo de 2SG mais leves (*zero-dativo* e *te-dativo*), cf. discutido por Oliveira (2014).

Uma outra constatação instigante é o fato de o dativo nulo (zero dativo), nas construções ditransitivas, ter se mostrado produtivo não só nos contextos sintagmáticos de OD (sintagmas nominal e oracional), mas também em contexto de OD nulo, o que pode configurar “como uma pista para um possível espraiamento do nulo também em contexto de

complementação verbal dativa (OI), cf. previu Berlinck (2005).” em consonância com a constatação de Cardoso (2017, p. 127).

3.1.4 A semântica do predicador verbal

À luz de Berlinck (1996, 2005) *apud* Oliveira (2014, p. 103), pretendemos descrever a semântica do predicador verbal que é responsável por condicionar as estruturas dativas nas cartas ao PN. Assumimos, com base em Oliveira (2014, p. 103), como hipótese o fato de “que os verbos com a noção de *transferência* – prototípica para o dativo – vinculem-se mais fortemente às formas mais produtivas, ao passo que as formas de ocorrência mais restrita se encaixem, na representação do dativo, a outras noções semânticas”. No quadro 2, apresentamos os tipos de especificidades semânticas do predicador verbal, ilustrando-os.

Quadro 2: Valor semântico do predicador verbal: alguns exemplos

Valor semântico do predicador verbal	Exemplos
Transferência material	dar, entregar, enviar, fazer, implorar, mandar, pedir, solicitar.
Transferência verbal e perceptual	dizer, contar, escrever, explicar, falar, perguntar, obedecer.
Movimento psicológico	agradecer.
Movimento abstrato	desejar, dedicar, deixar.
Movimento	chegar, vir.
Interesse	deixar, faltar, parecer.

Tabela 5: Estratégias pronominais dativas de 2SG correlacionadas à semântica do predicador verbal

SEMÂNTICA DO PREDICADOR	ESTRATÉGIAS PRONOMINAIS DATIVAS DE 2SG								
	te	lhe	a ti	para ti	a você	para você	você	zero	total
Interesse	-	-	-	-	-	1/3 (33,33%)	-	2/3 (66,67%)	3/1232 (0,24%)
Transferência material	271/597 (45,39%)	64/597 (10,72%)	-	-	15/597 (2,51%)	22/597 (3,69%)	6/597 (1,01%)	219/597 (36,68%)	597/1232 (48,46%)
Transferência verbal e perceptual	37/360 (10,28%)	45/360 (12,50%)	-	1/360 (0,28%)	-	26/360 (7,22%)	-	251/360 (69,72%)	360/1232 (29,22%)
Movimento abstrato	47/84 (55,95%)	7/84 (8,33%)	-	-	4/84 (4,76%)	10/84 (11,90%)	1/84 (1,19%)	15/84 (17,86%)	84/1232 (6,82%)
Movimento físico	-	-	-	-	-	-	-	-	-
Movimento	-	-	1/7 (14,29%)	-	-	-	-	6/7 (85,71%)	7/1232 (0,57%)
Movimento psicológico	60/181 (33,15%)	16/181 (8,84%)	-	-	4/181 (2,21%)	-	-	101/181 (55,80%)	181/1232 (14,69%)

De um modo geral (tabela 5), observamos, nas cartas ao PN, a prevalência dos predicadores verbais cujas semânticas são de transferência material (597/1232, **48,46%**), de transferência verbal e perceptual (360/1232, **29,22%**) e movimento psicológico (**14,69%**, 181/1232). Os demais tipos de semânticas dos predicadores verbais que se mostram funcionais, ainda que com baixos índices de produtividade, são os de movimento abstrato (6,82%, 84/1232), movimento (0,57%, 7/1232) e interesse (0,24%, 3/1232).

Nos verbos cujo valor semântico é de *transferência material* (primeira categoria semântica mais recorrente, em 48,46% dos dados (597/1232)), observamos a intensa alternância entre o *te* e o *zero* com índices percentuais de 45,39% (271/597) e 36,68% (219/597), respectivamente, nas cartas ao PN, cf. exemplificamos em (148) e (149). A semântica de *transferência material* mostrou-se produtiva com formas verbais como *dar*, *enviar*, *entregar*, *fazer*, *implorar*, *mandar*, *pedir*, *solicitar*. Esses índices de produtividades estão em consonância com os resultados de Cardoso (2017, p. 129), nas cartas mineiras

(séculos XIX e XIX), de Oliveira (2014), nas cartas cariocas (séculos XIX e XX), e, principalmente, em relação ao predomínio da noção semântica de transferência, cf. teorizado por Berlinck (1996, 2005).

As formas *te* e *zero-dativo* projetadas por verbos de *transferência material*:

(148) “[...] eu gostaria de [Ø]_{DAT} pedir uma boneca de natal [...]” (C776)

(149) “[...] Eu queria [*te*]_{DAT} pedir um material escolar [...]” (C781)

Ainda em relação aos verbos de *transferência material*, evidenciamos, mesmo que com índices baixos de produtividade, as formas *lhe* (64/597, 10,72%), *para você* (22/597, 3,69%), *a você* (15/597, 2,51%) e *você* (6/597, 1,01%), conforme verificamos de (150) a (153).

As formas *lhe*, *para você*, *a você* e *você* projetadas por verbos de transferência material:

(150) “[...] Quero [*lhe*]_{DAT} pedir um par de roupas da ADIDAS [...]” (C197)

(151) “[...] venho através desta carta pedi [*para você*]_{DAT} papai noel se puder. que me presente com uma cesta básica [...]” (C550)

(152) “[...] Porque tudo o que agente pede [*a você*]_{DAT}, você atende ao nosso pedido [...]” (C673)

(153) “[...] papai noel eu queria pidir [*você*]_{DAT} um presente [...]” (C531)

O segundo tipo de papel semântico do predicador mais produtivo nas cartas ao PN é o de *transferência verbal e perceptual* (29,22%, 360/1232) cujos verbos que se mostram na amostra em análise são, por exemplo, *dizer*, *contar*, *escrever*, *explicar*, *falar*, *perguntar*, *obedecer*. Dentre as formas dativas de 2SG projetadas por predicadores movidos pela semântica de *transferência verbal e perceptual* (transferência abstrata como efeito da comunicação através do ato da escrita, por exemplo), observamos uma alternância mais intensa entre o *zero dativo* (69,72%, 251/360) e o *lhe* (45/360, 12,50%), cf. ilustramos em (154) e (155). Por outro lado, as formas *te* (37/360, 10,28%), *para você* (26/360, 7,22%) e *para ti* (1/360, 0,28%) apresentam-se com baixos índices percentuais, cf. ilustramos de (156) a (158).

As formas *zero-dativo* e *lhe* projetadas por verbos de *transferência verbal e perceptual*:

(154) “[...] papai noel eu tenho uma irma que já [Ø]_{DAT} escrevel uma cartinha mas não foi realizada [...]” (C650)

(155) “[...] minha mãe está [*lhe*]_{DAT} escrevendo para [*lhe*]_{DAT} contar um pouquinho da minha história [...]” (C495)

As formas *te*, *para você* e *para ti* projetadas por verbos de *transferência verbal e perceptual*:

(156) “[...] estou [*te*]_{DAT} escrevendo está humilde carta para *lhe* pedir uma boneca [...]” (C741)

(157) “[...] pois sei que são muitas crianças que escreve [*para você*]_{DAT} [...]” (C585)

(158) “Papai Noel Gosto de ti, [...] por isso eu estou dizendo [*para ti*] que gosto muito de presentes... *te* adoro” (C961)

O predicador cuja semântica é a de *movimento psicológico* foi o terceiro mais produtivo na amostra de cartas em análise (181/1232, 14,69%). Nesse contexto, o dativo atua como o “portador” do movimento psicológico relativo ao verbo e ao sujeito gramatical, representando a *causa* ou a *fonte* desse processo como observamos, por exemplo, no verbo como *agradecer*. A alternância entre o *zero dativo* e o *te* mostrou-se, mais uma vez como ocorreu tão somente com o verbo *agradecer*, como formas produtivas em 55,80 % (101/181) e 33,15% (60/181) dos dados, respectivamente, cf. ilustramos em (159) e (160). Observamos, ainda que baixas frequências de uso, as formas *lhe* (161) e *a você* (162), em 8,84% (16/181) e 2,21% (4/181) dos dados.

As formas *zero-dativo* e *te* projetadas por verbos de *movimento psicológico*:

(159) “[...] ser você pode medar vou [Ø]_{DAT} agradecer muito [...]” (C914)

(160) “[...] Se voçe puder realizar esse sonho [*te*]_{DAT} agradeço [...]” (C948)

As formas *lhe* e *a você* projetadas por verbos de *movimento psicológico*:

(161) “[...] Quero *lhe* pedir se você pode me dar o kit de material escolar, se puder [*lhe*]_{DAT} agradeço muito (C066)

(162) “[...] agradeço [*a voce*]_{DAT} porque neus pais ganham pouco [...]” (C817)

O contexto semântico de *movimento abstrato* foi o quarto mais produtivo nos predicadores verbais responsáveis pelos dativos de 2SG (6,82%, 84/1232). Nesse contexto, a semântica é a de movimento, mas sem que haja transferência, ao contrário, “há uma aproximação abstrata entre uma entidade e um estado, sentimento, ideia ou outra entidade”²³ (BERLINCK, 1996, p. 133), o que se deixou evidenciar, na amostra de cartas ao PN, principalmente com verbos *desejar*, *dedicar* e *deixar*. Dentre as estratégias pronominais dativas de 2SG projetadas por verbos de *movimento abstrato*, verificamos que os maiores índices de alternância se mostram entre o *te*, em 55,95% dos dados (47/84), e o *zero*, em 17,86% dos dados (15/84), como observamos em (163) e (166). Com índices de produtividade menores, observamos as formas *para você* (11,90%, 10/84), *lhe* (8,33%, 7/84), *a você* (4/84, 4,76%) e *você* (1,19%, 1/84) para os quais apresentamos as seguintes evidências expostas de (167) a (170).

As formas *te* e *zero-dativo* projetadas por verbos de *movimento abstrato*:

(163) “[...] Primeiro eu quero [*te*]_{DAT} desejar, que o Senhor tenha um feliz natal, cheio de paz, alegria e tranquilidade [...]” (C090)

(164) “[...] [*te*]_{DAT} dedico um bom natal [...]” (C046)

(165) “[...] Quero [\emptyset]_{DAT} deixar meus votos de felicidades Um feliz natal e um próspero Ano novo [...]” (C010)

(166) “[...] desde já [\emptyset]_{DAT} deixo meus agradecimento [...]” (C245)

As formas *para você*, *lhe*, *a você* e *você* projetadas por verbos de *movimento abstrato*:

(167) “[...] e eu [*te*]_{DAT} desejo um feliz natal [*para você*]_{DAT} é sua família [...]” (C003)

(168) “[...] papai noel [*lhe*]_{DAT} desejo um feliz natal e um prospero ano Novo! (C060)

(169) “[...] Desejo [*a você*]_{DAT} um feliz natal. [...]” (C021)

(170) “[...] aproveito para desejar [*você*]_{DAT} e sua família um feliz natal e um ano novo cheio de paz, amor e realizações [...]” (C448)

²³ “[...] there is an abstract approximation between an entity and a state, a feeling, an idea, or another entity” (BERLINCK, 1996, p. 133).

O contexto semântico de *movimento* foi o quinto mais produtivo dos predicadores verbais responsáveis pelos dativos de 2SG, em tão somente 0,57% dos dados (7/1232). A perspectiva é a de que fosse possível levantarmos a projeção de um complemento argumento dativo direcional também marcado pelo traço de [+ afetado], cf. Berlinck (1996, p. 140). Para esse tipo semântico de predicador verbal, observamos tão somente seis ocorrências do *zero-dativo* (85,71%) em oposição a uma única ocorrência do *a ti* (14,29%) com os verbos *chegar* (172 e 173) e *vir* (171). A opção por entender os verbos *chegar* e *vir* como verbos semanticamente de *movimento* foi motivada não só pelo traço de [+ afetado] atribuída ao complemento dativo de 2SG, mas também pela evidência expressa pelo dado “chegar *a ti*” que poderia alternar com “chegar *a você*”, o que também pode se aplicar ao verbo *vir* através do dado “venho [Ø]_{DAT} através esta cartinha com muita fé” em oposição às alternativas “venho *a você/a ti* através esta cartinha com muita fé”.

(171) “[...] Venho [Ø]_{DAT} através esta cartinha com muita fé que desta vez o senhor vai me ajudar” (C933)

(172) “[...] Espero que esta chegue [Ø]_{DAT} ainda em tempo [...]” (C474)

(173) “[...] A estrada da tua Felicidade não parte das pessoas e das coisas para chegar [*a ti*]_{DAT} [...]” (C834)

O contexto de predicador verbal de *interesse* foi o sexto mais produtivo mais produtivo, em tão somente 0,24% dos dados (3/1232), na projeção dos complementos dativos de 2SG das cartas ao PN. Trata-se de um verbo movido pela voluntariedade do sujeito gramatical que apesar de ser, normalmente, [+ animado], também pode ser [- animado], sendo o dativo expresso em sintagma preposicionado encetado pela preposição “a”. Na amostra das cartas ao PN, observamos três evidências da semântica de interesse nos verbos *parecer* e *deixar* através dos quais prevaleceu o *zero-dativo* (66,67%, 2/3) em oposição a uma única ocorrência do SPrep “para você” (33,33%, 1/3), conforme ilustramos de (174) a (176). A ideia é a de que os verbos *deixar*, *faltar* e *parecer* tendam a projetar um complemento dativo marcado positivamente em relação ao traço de animacidade, ou seja, é [+ animado], encetado por um SPrep encetado pela preposição “a”, podendo assumir a mesma interpretação que a preposição “para”. Assim sendo, as interpretações do papel semântico dos verbos *parecer* e *deixar* em relação ao dativo de interesse por eles projetados são as seguintes: “...apesar que

minha idade não *te* pareça...”/“[...] apesar que minha idade não pareça *a você/a ti*...”; “Deus ... não deixa *a/para você/* ou *a/para ti* [faltar nada]_{OD OF}[...]”.

(174) “[...] apesar que minha idade não [∅]_{DAT} pareça, sou uma criança sim [...]” (C343)

(175) “[...] eles são doidos por carro por incrível que [∅]_{DAT} pareça até a menina [...]” (C462)

(176) “[...] Desde já agradeço de Coração Que Deus te abençoe e não deixa nada faltar [*pra você*]_{DAT} [...]” (C545)

3.1.5 A semântica do dativo

Na perspectiva de Company Company (2006, p. 520), os papéis semânticos do OI²⁴ “compartilham e estão vinculados pelo mesmo valor semântico geral de *locus* ou *meta* final da transitividade, que tem [...] esta categoria funcional como significado básico”, assumindo papéis semânticos relacionados com o grau de envolvimento da entidade do OI integrados no significado do verbo e no evento em geral. À luz de Company Company (2006), temos um *continuum* de envolvimento ou de afetação que evidenciam maior envolvimento (*receptor, experienciador e possuidor*) e menor envolvimento (*meta-fonte e o beneficiário*). Isso posto, passamos à identificação dos papéis semânticos mais produtivos especificamente em relação às estratégias dativas de 2SG das cartas ao PN.

²⁴ “Todos los papeles semánticos del OI comparten y están vinculados por el valor semântico general de *locus*, a meta final de la transitividad, que [...] tiene esta categoria funcional como significado básico” (COMPANY, 2006, p. 520).

Tabela 6: Estratégias pronominais dativas de 2SG correlacionadas às semânticas do dativo de 2SG

SEMÂNTICA	ESTRATÉGIAS PRONOMINAIS DATIVAS DE 2SG								
	te	lhe	a ti	para ti	a você	para você	você	zero	total
DO DATIVO									
Meta/Fonte	347/778 (44,60%)	79/778 (10,15%)	1/778 (0,13%)	1/778 (0,13%)	22/778 (2,83%)	23/778 (2,96%)	7/778 (0,90%)	298/778 (38,30%)	778/1232 (63,15%)
Receptor	67/443 (15,12%)	53/443 (11,96%)	-	-	1/443 (0,23%)	35/443 (7,90%)	-	287/443 (64,79%)	443/1232 (35,96%)
Beneficiário	1/8 (12,50%)	-	-	-	-	-	-	7/8 (87,50%)	8/1232 (0,65%)
Experienciador	-	-	-	-	-	1/3 (33,33%)	-	2/3 (66,67%)	3/1232 (0,24%)

Dentre as semânticas das estratégias dativas de 2SG levantadas nas cartas ao PN (tabela 6), identificamos a prevalência do papel semântico de *meta/fonte*, em 63,15% dos dados (778/1232). Trata-se de uma entidade a que algo/alguém toca/chega (*meta*) ou se origina (*fonte*), sendo possível o deslocamento ou a transferência de uma dada entidade. Nesse sentido, verbos como *agradecer*, *desejar* e *pedir* evidenciam, nas cartas ao PN, a referência à meta, ao alvo do complemento dativo (OI) de 2SG projetado pelo predicador verbal. Nesse contexto de dativo com a semântica de *meta/fonte*, observamos a alternância entre o *zero*, em 38,30% (298/778), e o *te*, em 44,60% dos dados (347/778), cf. observamos em (177) e (179). Temos ainda as formas *lhe* (10,15%, 79/778), *para você* (2,96%, 23/778) e *a você* (2,83%, 22/778) como formas produtivas, cf. de (180) a (182), mesmo que com índices inferiores às frequências de uso das formas *zero* e *te*. Por outro lado, com índices baixíssimos, estão as formas *você* (0,90%, 7/778), *a ti* (1/778, 0,13%) e *para ti* (1/778, 0,13%), ilustrados de (183) a (185).

O papel semântico de meta nas formas de *zero* e *te*:

(177) “[...] Espero que esta chegue [Ø]_{DAT} ainda em tempo [...] (C474)

(178) “[...] esse ano gostaria de [Ø]_{DAT} pedir um presente para minha mãe [...] (C376)”

(179) “[...] eu queria [*te*]_{DAT} pedir uma Bicicleta para mim não poder mais andar a pé [...] (C475)

O papel semântico de meta nas formas *lhe*, *para você* e *a você*:

(180) “[...] Gostaria de [*Lhe*]_{DAT} pedir com minha humiLde gratidão, e caLma, uma coisa que só você poderia me, der [...]” (C281)

(181) “[...] todos os anos eu peço [*para você*]_{DAT} mas nunca sou atendida gostaria que me atendesse esse ano [...]” (C099)

(182) “[...] Nessa véspera de natal eu gostaria de pedir [*a você*]_{DAT} que esta lendo a minha carta, um “notebook” ou um “tablet” [...]” (C152)

O papel semântico de meta nas formas *a ti*, *para ti* e *você*:

(183) “[...] A estrada da tua Felicidade não parte das pessoas e das coisas para chegar [*a ti*]_{DAT} [...]” (C834)

(184) “Papai Noel Gosto de ti, [...] por isso eu estou dizendo [*para ti*]_{DAT} que gosto muito de presentes... te adoro” (C961)

(185) “[...] aproveito para desejar [*você*]_{DAT} e sua família um feliz natal e um ano novo cheio de paz, amor e realizações [...]” (C448)

O segundo papel semântico do complemento dativo de 2SG mais produtivo é o *receptor* para o qual observamos, em 35,96% dos dados (443/1232), a sua frequência de uso. O complemento dativo pode assumir o papel semântico de receptor do ato expresso pelo predicador, tendo em vista a sua volicionalidade em relação à transferência (física ou metafórica) como evidência prototípica do dativo como observamos nos verbos *contar*, *dar*, *dizer*, *escrever*, *falar*, *fazer*, *mandar*, *pagar*. Nesse contexto, observamos a prevalência do *zero*, em 64,79% dos dados (287/443), em alternância com o *te* que se manifesta em 15,12% dos dados (67/443), cf. verificamos de (186) a (189).

O papel semântico de *receptor* nas formas de *zero* e *te*:

(186) “[...] papai Noel todo ano [\emptyset]_{DAT} mando mas nunca fui correspondida [...]” (C425)

(187) “[...] Estou [\emptyset]_{DAT} mandando essa carta com sentimentos bons [...]” (C852)

(188) “[...] Papai Noel eu [*te*]_{DAT} mando erra carta para te pedir esse grande favor [...]” (C070)

(189) “[...] Papai Noel vou Terminar minha carTinha [*Te*]_{DAT} mandando um forte abraço e muiTos beijinhos [...]” (C454)

Ainda que tenhamos tido a prevalência do *zero*, observamos que, em termos absolutos e percentuais, as formas *te* e *lhe* com o papel de receptor mostram-se em um grau mais intenso de variação com os índices de 15,12% 67(1443) e 11,96% (53/443), respectivamente. Em (190) e (191), observamos algumas evidências de *te* e *lhe* com a semântica de *receptor* fomentadas pelo verbo *dar*.

(190) “[...] que Deus [*lhe*]_{DAT} der bastante Saúde, força e coragem para pode ajudar muitas crianças que necessita [...]” (C010)

(191) “[...] que te abençõe, [*te*]_{DAT} de saúde te faça feliz [...]” (C913)

Atentemos ainda para a alternância *a você* e *para você* com a semântica de *receptor* em favor do SPrep *para você* com uma diversidade maior de predicadores verbais (7,90%, 35/443) em oposição a tão somente uma ocorrência (0,23%) de *a você* projetada pelo predicador verbal *entregar*, cf. ilustramos em (192) e (193).

(192) “[...] Esta é a primeira vez que Mando uma cartinha [*para você*]_{DAT} [...]” (C681)

(193) “[...] No ano passado eu mandei uma carta para a escola entregar [*a você*]_{DAT} mas não ganhei nada [...]” (C840)

O complemento dativo com o potencial de *beneficiário* foi o terceiro tipo de papel semântico mais produtivo na amostra de cartas ao PN, assumindo uma produtividade de 0,65% dos dados (8/1232). Partimos da ideia de Company Company (2006, p. 522) que reserva tal papel semântico para as estruturas de OI nucleadas pela “entidade que recebe para seu proveito o significado da oração” em alto grau de afetação tendo em vista o evento manifesto pelo predicador verbal. Das oito ocorrências do dativo *beneficiário*, temos sete delas com o *zero-dativo* e tão somente uma evidência com o clítico *te*, cf. ilustramos de (194) a (196), projetadas pelos verbos *dedicar* e *deixar*.

(194) “[...] [*te*]_{DAT} dedico um bom natal [...]” (C046)

(195) “[...] Vou [\emptyset]_{DAT} deixa aqui papai noel o que eu sonho ganhar [...]” (C498)

(196) “[...] pode me ligar, deixo [Ø]_{DAT} aqui o meus números [...]” (C428)

O quarto papel semântico mais produtivo nas cartas em análise é o de *experienciador* que se mostrou produtivo em tão somente 0,24% dos dados (3/1232). A semântica do *experienciador* deixa-se evidenciar em situações estativas em que se tem mudança de estado anímico, mental ou sensorial motivada pelo evento resguardado pelo predicador verbal, o que comumente ocorre com verbos psicológicos tais como *faltar* e *parecer*. Nesse contexto, temos em análise tão somente três ocorrências de dativo *experienciador* distribuídas entre o zero (66,67%) e o SPrep *para você* (33,33%), cf. ilustramos de (197) a (199).

(197) “[...] apesar que minha idade não [Ø]_{DAT} pareça, sou uma criança sim [...]” (C343)

(198) “[...] eles são doidos por carro por incrível que [Ø]_{DAT} pareça até amenina [...]” (C462)

(199) “[...] Desde já agradeço de Coração Que Deus te abençoe e não deixa Nada faltar [*pra você*]_{DAT} [...]” (C545)

3.2 Os contextos sociolinguísticos: a seção da carta e a precedência da carta

A seguir, passamos a análise de produtividade das variáveis sociolinguísticas controladas neste trabalho (a seção da carta e a precedência da carta).

3.2.1 Os complementos dativos de 2SG e a seção da carta

O controle das partes constitutivas da carta está motivado pelo fato de poderem condicionar, em algum nível, a produtividade de uma dada estratégia pronominal dativa de 2SG. Prevemos o fato de as seções inicial e final das cartas poderem fomentar determinadas estruturas formulaicas também através das estratégias pronominais dativas de 2SG principalmente nos espaços discursivos reservados à captação da benevolência e da despedida, cf. também discutido e analisado por Oliveira (2014, p. 105).

Tabela 7: Estratégias pronominais dativas de 2SG correlacionadas às seções das cartas

SEÇÃO DA CARTA	ESTRATÉGIAS PRONOMINAIS DATIVAS DE 2SG								
	te	lhe	a ti	para ti	a você	para você	você	zero	total
Saudação inicial	85/418 (20,33%)	54/418 (12,92%)	-	-	7/418 (1,67%)	23/418 (5,50%)	2/418 (0,48%)	247/418 (59,09%)	418/1232 (33,92%)
Núcleo da carta	207/505 (40,99%)	51/505 (10,10%)	-	1/505 (0,20%)	10/505 (1,98%)	22/505 (4,36%)	4/505 (0,79%)	210/505[2] (41,58%)	505/1232 (40,99%)
Despedida	121/296 (40,88%)	26/296 (8,78%)	-	-	5/296 (1,69%)	14/296 (4,73%)	1/296 (0,34%)	129/296 (43,58%)	296/1232 (24,03%)
<i>Post Scriptum</i>	2/13 (15,38%)	1/13 (7,69%)	1/13 (7,69%)	-	1/13 (7,69%)	-	-	8/13 (61,54%)	13/1232 (1,06%)

Dentre as partes constitutivas das cartas (tabela 7), observamos que o *núcleo da carta* e a *saudação inicial* funcionam como os contextos mais receptivos às estratégias dativas de 2SG em 40,99% (505/1232) e 33,92% (418/1232) dos dados, respectivamente. O fato de o núcleo da carta ter se mostrado também nas cartas ao PN como espaço discursivo em **que** observamos o maior número de ocorrências de estratégias dativas de 2SG coaduna-se à constatação de Oliveira (2014) que também evidenciou o *núcleo da carta* como o de prevalência das formas dativas nas cartas cariocas (séculos XIX e XX), figurando, pois, como “o espaço de maior liberdade discursiva por parte do missivista” (2014, p.129). A *despedida* e o *post-scriptum* também contextualizam os complementos dativos de 2SG, em 24,03% (296/1232) e 1,06% (13/1232) dos dados, respectivamente, ainda que com índices percentuais menos expressivos do que os observados para os contextos de *núcleo* e *saudação inicial*. Nos contextos de captação de *benevolência* e *despedida* e de *post scriptum*, observamos também nas cartas ao PN.

No núcleo da carta, observamos o intenso nível de alternância entre as formas *zero*, em 41,58% dos dados (210/505), e *te*, em 40,99% dos dados (207/505), evidenciando tais contextos como propícios às duas possibilidades de complementação dativa de 2SG (*zero* e *te*) mais produtivas nas cartas ao PN em análise. As formas *lhe* (51/505, 10,10%), *para você*

(22/505, 4,36%) e *a você* (10/505, 1,98%) mostram-se em alternância, ao passo que as formas *você* (4/505, 0,79%) e *a ti* (1/505 (única ocorrência), 0,20%) apresentam baixíssimas frequências de uso também no contexto do núcleo da carta. De (200) a (206), observamos evidências dessas estratégias dativas de 2SG no núcleo da carta.

As formas *zero* e *te* no núcleo da carta:

(200) “[...] Gostaria de [Ø]_{DAT} pedir um trator de brinquedo para brincar com meus priminhos [...]” (C959)

(201) “[...] Queria muito [*te*]_{DAT} pedir de presente um kit escolar [...]” (C955)

As formas *lhe*, *para você* e *a você* no núcleo da carta:

(202) “[...] Quero [*lhe*]_{DAT} pedir se você pode me dar o kit de material escolar, se puder [*lhe*]_{DAT} agradeço muito [...]” (C066)

(203) “[...] Já mandei mais de 40 cartas [*para você*]_{DAT} mas você nao min respondeu [...]” (C399)”

(204) “[...] Eu queria pedir [*avoce*]_{DAT} um Celula [...]” (C285)

As formas *você* e *para ti* no núcleo da carta:

(205) “[...] Eu gostaria de pedir [*você*]_{DAT} papai noel um tablet [...]” (C098)

(206) “Papai Noel Gosto de ti, [...] por isso eu estou dizendo [*para ti*]_{DAT} que gosto muito de presentes... te adoro” (C961)

No contexto da *saudação inicial* das cartas, observamos novamente a prevalência do *zero*, em 59,09% (247/418) em oposição mais intensa (em termos de números brutos) entre o *te* (85/418), o *lhe* (54/418) e o *para você* (23/418), em 20,33%, 12,92% e 5,50% dos dados, respectivamente. Em relação às formas *a você* e *você*, temos poucas evidências de uso nas cartas ao PN em 1,67% (7/418) e 0,48% (2/418) dos dados, respectivamente. De (207) a (212), observamos evidências dessas estratégias dativas de 2SG na saudação inicial das cartas ao PN em análise. No espaço discursivo destinado à saudação inicial, temos não só evidência da estrutura formulaica “venho através desta”, mas também de uma busca por saber como está o interlocutor (“Oi Papai Noel Como vai?”, “Oi papai noel como vai você meu nome e...”) como uma legítima expressão da captação de benevolência.

As formas *zero*, *te*, *lhe* e *para você* na saudação inicial:

(207) “Óla papai noel é com muito prazer que [Ø]_{DAT} escrevo está carta [...]” (C118)

(208) “Meu nome é [...] e venho através desta, [te]_{DAT} pedir humil dimente Se for possível atender [...]” (C306)

(209) a. “Meu querido papai noel venho através desta [lhe]_{DAT} pedir que realize meu desejo [...]” (C959)

b. “[...] tenho 31 anos steu desempregada venho através desta carta [lhe]_{DAT} pedir [...]” (C507)

(210) a. “Querido papai noel. Meu nome e: [...] vou contar um pouco da minha estoria [para você]_{DAT} [...]” (C202)

b. “Oi Papai Noel Como vai? hoje resovi Escrever uma Cartinha [Para Você] [...]” (C214)

As formas *a você* e *você* na saudação inicial:

(211) “Oi papai noel como vai você meu nome e [...] e estou lhe escrevendo para pedir [a você]_{DAT} que possa mim ajudar neste natal [...]” (C482)

(212) “Querido Papel Noel Papai noel meu nome é [...] papai noel eu tenho 36 anos papai noel eu queria pidir [você]_{DAT} um presente [...]” (C531)

Na seção de *despedida*, observamos, mais uma vez, a intensa alternância entre o *zero* e o *te* em 43,58% e 40,88% dos dados, respectivamente, cabendo ao *lhe* e ao *para você* índices mais tímidos de produtividade, em 8,78% (26/296) e 4,73% (14/296), respectivamente, mas também em intensa alternância. Ao *a você* e *você*, couberam baixíssimas frequências de uso, em tão somente 1,69% e 0,34% dos dados, reverberando o também ínfimo número de dados (5oco e 1 oco), respectivamente. De (213) a (218), ilustramos, com mais algumas evidências, as formas dativas no contexto da *despedida* da carta.

As formas *zero* e *te* na *despedida*:

(213) “[...] Caso não tiver [Ø]_{DAT} agradesso por tudo Numero: 37 e 38 tenho 10 anos” (C936)

(214) “[...] [te]_{DAT} agradeço com muito carinho espero que eu ganhe esse premio! Bjus:” (C939)

As formas *lhe* e *para você* na *despedida*:

(215) “[...] Com a sua atenção e carinho desejo-[lhe]_{DAT} um Feliz Natal e um Feliz ano Novo Papai Noel.”

(216) “[...] Desde já agradeço de Coração Que Deus te abençoe e não deixa Nada faltar [*pra você*]_{DAT} [...] (C545)

As formas *a você* e *você* na *despedida*:

(217) “[...] Eu sempre vou pedir [*a você*]_{DAT} os presentes porque eu sei que você não me decepiona! Feliz Natal ass:” (C673)

(218) “[...] aproveito para desejar [*você*]_{DAT} e sua família um feliz natal e um ano novo cheio de paz, amor e realizações obrigado (a). (C448)

Na seção *post scriptum* das cartas ao PN, levantamos treze ocorrências de estratégias dativas de 2SG, estando elas distribuídas entre o *zero* e o *te* em 61,54% (8/13) e 15,38% (2/13) respectivamente. As formas *lhe*, *a ti* e *a você* mostram-se com uma única ocorrência cada uma delas. Interessante atentar ao fator de que a única ocorrência do SPrep *a você* no contexto de *post scriptum* foi projetada pelo verbo *abençoar* cuja grade argumental não prevê complementos preposicionados no sentido em questão que é o de “conceder-lhe bençãos”, cabendo ainda a possibilidade da sua interpretação como objeto direto preposicionado. De (219) a (223), apresentamos evidências de tais formas dativas de 2SG no *post-scriptum* da carta.

As formas *zero* e *te* na seção *post scriptum*:

(219) “[...] Obs. Precisei de ajuda p/ [Ø]_{DAT} escrever esta carta.” (C439)

(220) “[...] Eu [*te*]_{DAT} peço porque eu so Pobre” (C637)

As formas *lhe*, *a ti* e *a você* na seção *post scriptum*:

(221) vi muitas flores querida mas posso [*lhe*]_{DAT} jurar: é a primeira vez na vida que eu vejo uma flor andar I love you” (C722)

(222) “[...] Agradeço e que Deus possa abençoa [*a você*]_{DAT} que está ajudando um jovem sonhador” (C479)

(223) “[...] A estrada da tua Felicidade não parte das pessoas e das coisas para chegar [*a ti*]_{DAT} parte sempre de ti em direção aos outros Papai” (C834)

3.2.2 Os complementos dativos de 2SG e a procedência das cartas

Dentre todas as ocorrências de estratégias dativas de 2SG (1232oco), temos tão somente, em 30,19% dos dados (372oco), a identificação da procedência das cartas ao PN, conforme apresentamos, na tabela 8, a partir da distribuição dos dados em análise.

Tabela 8: Estratégias pronominais dativas de 2SG correlacionadas à procedência das cartas

PROCEDÊNCIA DAS CARTAS	ESTRATÉGIAS PRONOMINAIS DATIVAS DE 2SG							
	te	lhe	a ti	a você	para você	você	zero	total
Norte de MG	-	-	-	-	-	-	1/1 (100%)	1/372 (0,27%)
Vale do Mucuri	1/1 (100%)	-	-	-	-	-	-	1/372 (0,27%)
Reg.Metropolitana de BH	129/369 (34,96%)	41/369 (11,11%)	1/369 (0,27%)	5/369 (1,36%)	21/369 (5,69%)	2/369 (0,54%)	170/369 (46,07%)	369/372 (99,19%)
Oeste de MG	-	-	-	-	-	1/1 (100%)	-	1/372 (0,27%)
								372

Dentre as estratégias dativas de 2SG levantadas nas cartas ao PN identificadas no espaço mineiro (tabela 8), verificamos, em 99,19% dos dados (369/372), a distribuição dos dados por praticamente todas os tipos de formas dativas (*te*, *a ti*, *lhe*, *a você*, *para você*, *você*, *zero*). Assim sendo, continuamos a observar a variação mais intensa, principalmente, entre o *zero* (170/369, 46,07%) e o *te* (129/369, 34,96%). O *lhe* (11,11%, 41/369) e o *para você* (5,69%, 21/369) mostram-se, ainda que com níveis de produtividade mais tímidos, também em variação. Os mais ínfimos índices de produtividade estão voltados para as formas *a você* (1,36%, 5/369), *você* (0,54%, 2/369) e *a ti* (0,27%, 1/369). De (223) a (229), ilustramos as estratégias dativas de 2SG produtivas nas cartas provenientes da região Metropolitana de BH.

As formas *zero*, *te*, *lhe*, *para você*, *a você*, *você* e *a ti* em cartas da região Metropolitana de BH:

- (223) “[...] Gostaria de [Ø]_{DAT} pedir um trator de brinquedo para brincar com meus priminhos [...] Betim, 10 de novembro de 2012” (C959)
- (224) “[...] sou uma menina muito feliz e eu queria [te]_{DAT} pedir umas coisas um material escolar do rebeldes [...] Nova Lima [...]” (C002)
- (225) “[...] Sou [...] tenho 13 anos moro em Betim Mg e estou Ø enviando esta carta para [lhe]_{DAT} pedir um Notebook [...]” (C339)
- (226) “[...] venho atraves desta carta pedi [para você]_{DAT} papai noel se puder. que me presente com uma cesta básica [...]” (C550)
- (227) “Belo Horizonte, 20 de dezembro de 2012 Querido Papai Noel, [...] Nessa véspera de natal eu gostaria de pedir [a você]_{DAT} que esta lendo a minha carta, um “notebook” ou um “tablet” [...]” (C152)
- (228) “Belo horizonte Caro papai Noel, [...] por isso estou pedindo [você]_{DAT} uma bicicleta [...]” (C823)
- (229) “Belo Horizonte dia 30 de novembro de 2013 [...] A estrada da tua Felicidade não parte das pessoas e das coisas para chegar [a ti]_{DAT} [...]” (C834)

Para as regiões do Norte de Minas Gerais, do Vale do Mucuri e do Oeste de MG, temos tão somente uma única ocorrência das formas *zero*, *te* e *você* para cada uma dessas regiões mineiras, representando tão somente 0,2% em relação a todos os dados de cartas que tiveram o seu local de produção devidamente explicitado. De (230) a (232), ilustramos tais estratégias dativas de 2SG nas regiões mineiras em foco.

As formas *zero*, no norte de Minas Gerais (Montes Claros), e *te*, no Vale do Mucuri (São Sebastião do Maranhão), e, *você*, no Oeste de Minas Gerais (São Gonçalves do Pará):

- (230) “[...] jesus te abençoa Papai noel Papai noel foi feito [Ø]_{DAT} no dia 16/10/2012 de outubro [...] montes claros estado de minas gerais” (C704)
- (231) “Papai Noel eu me chamo [...] tenho 11 anos moro em São Sebastião do Maranhão [...] Venho [te]_{DAT} pedir 2 presentes 1 notebook e um celular [...]” (C037)
- (232) “[...] papai noel eu queria pidir [você]_{DAT} um presente [...] Cidade São Gonçalves do Pará [...]” (C531)

Síntese do capítulo

Em termos de resultados gerais, podemos identificar a referência à 2SG nas cartas brasileiras em análise dá-se através das formas dativas *te*, *a ti*, *para ti*, *lhe*, *você*, *a você*, *para você* e *zero*. As estratégias mais frequentes foram o *zero* (48,14%, 594/1232) e o *te* (33,69%, 415/1232), seguidas pelo *lhe* (10,71%, 132/1232) e pelo o SPrep *para você* (4,79%, 59/1234). As demais estratégias pronominais evidenciam baixíssimos índices percentuais. São eles: o SPrep *a você* (1,87%, 23/1234), *você* (0,73%, 7/1232), o SPrep *para ti* (0,08%, 1/1232) e o SPrep *a ti* (0,08%, 1/1232).

Quando voltamos o foco para a correlação entre as formas dativas e nominativas de 2SG (subitem 3.1.1), verificamos uma alta produtividade do *te-dativo* em contextos de *você-sujeito*, confirmando a manutenção do *tu* etimológico através do *clítico te* que convivia e parece continuar a conviver com o *você-sujeito* na escrita brasileira produzida por redatores que nasceram e viveram no espaço geográfico de Minas Gerais, espaço de prevalência do *você-sujeito*, como discutido, a partir de dados da fala mineira, por Coelho (1999), Peres (2006), Herênio (2006), Gonçalves (2008), Mota (2008), Scherre *et alii* (2015), Silva (2017), Reis (2019)²⁵. Tais resultados correspondem aos previamente identificados em cartas mineiras (RUMEU, 2015; CARDOSO, 2017) e cartas cariocas de sincronias passadas (OLIVEIRA, 2014). Destacamos que os dados levantados também desmistificam a chamada “uniformidade tratamental” trazida pela tradição gramatical, dada a produtividade de formas do paradigma de *tu* em convívio com o *você-sujeito*.

No que diz respeito aos padrões sintáticos (subitem 3.1.2), buscávamos identificar se os verbos de 2 lugares (S V OI) seriam favorecedores do *te-dativo*, o que não ficou evidenciado. Verificamos que a estratégia mais produtiva nesse contexto foi o *zero* (57,75%, 108/187), enquanto o *te* ficou em segundo lugar (31,02%, 58/187). Esses dados vão de encontro aos resultados de Oliveira (2014) para as cartas cariocas, que observou em seu *corpus* o padrão de 2 lugares como impulsionador do *te-dativo*. Para os verbos *ditransitivos*

²⁵ Referências bibliográficas que fizeram parte do capítulo de “Revisão do Tema” confeccionado por Souza (2021) cujo foco de análise é a expressão variável das formas de referência ao sujeito de 2SG em cartas brasileiras (*vossa mercê/você/tu*), produzidas por punhos de redatores que nasceram e/ou viveram em Minas Gerais, entre os séculos XIX e XX.

(S V OD OI), também identificamos uma variação acirrada entre o *zero* (46,51%, 486/1045) e o *te* (34,16%, 357/1045).

Ao observarmos a relação entre as formas acusativas (OD) e as formas dativas nas estruturas ditransitivas (subitem 3.1.3), identificamos que o *zero* e o *te* foram as formas mais produtivas em convivência com o OD como sintagma nominal, OD nulo e sintagma oracional, o que confirmou nossa hipótese quanto à relação entre as funções acusativa e dativa. Os dados revelam, portanto, que as formas mais “pesadas” de OD (sintagmas nominal e oracional) favorecem o uso de formas de dativo mais “leves” (*zero-dativo* e *te-dativo*).

Em relação às categorias dos verbos propostas por Berlinck (1996, 2005) (subitem 3.1.4), confirmamos nossas hipóteses quanto à alta produtividade de verbos com valor semântico de transferência, considerada prototípica para os dativos. Especificamente para a *transferência material*, constatamos uma acirrada variação entre o *te-dativo* (45,39%, 271/597) e o *zero* (36,68%, 219/597). Para os verbos de *transferência verbal e perceptual*, a variação mais intensa se deu entre o *zero* (69,72%, 251/360) e o *lhe* (12,50%, 45/360).

A última variável linguística estudada neste trabalho é o papel semântico do *dativo* (subitem 3.1.5), cuja prevalência foi para o papel de *meta/fonte*, que assumiu 63,15% dos dados (778/1232). As formas preferidas para este papel foram o *te* (44,60%, 347/778) e o *zero* (38,30%, 298/778), também foram produtivas as formas *lhe* (10,15%, 79/778), *para você* (2,96%, 23/778) e *a você* (2,83%, 22/778), enquanto as formas *você* (0,90%, 7/778), *a ti* (1/778, 0,13%) e *para ti* (1/778, 0,13%) tiveram baixíssima produtividade. Quanto ao papel semântico de *receptor*, vê-se que houve uma predominância do *zero*, representando 64,79% dos dados (287/443), em alternância com o *te*, observado em 15,12% dos dados (67/443).

Sintetizados os principais resultados percentuais das variáveis estruturais, passamos a apresentar os percentuais relativos à seção da carta (subitem 3.1.6) e a procedência desta no espaço mineiro (subitem 3.1.7).

Quanto à seção da carta, identificamos que o contexto mais produtivo para as estratégias dativas foi o *núcleo*, com 40,99% dos dados (505/1232), seguido pela saudação (33,92%, 418/1232) e despedida (24,03%, 296/1232). O *zero* se mostrou a forma mais produtiva em todos os contextos, muito embora esteja em intensa variação com a forma *te* no *núcleo* (41,58% de *zero* em oposição a 40,99% de *te*) e na despedida (43,58% de *zero* em

oposição a 40,88% de *te*). Na saudação inicial, verificou-se uma variação acirrada em relação às formas *te* (20,33%, 85/418), *lhe* (12,92%, 54/418) e *para você* (5,50%, 23/418).

A última variável que apresentamos neste capítulo foi a procedência das cartas, das quais em apenas 372 (30,19%) delas foi possível identificar o local de escrita e a sua grande maioria (99,19%) foi firmada na Região Metropolitana de Belo Horizonte (a capital, Contagem, Betim, Nova Lima, Sete Lagoas, Santa Luzia *etc.*). Nas cartas controladas nesta etapa, observamos também a dinâmica de variação entre o *zero* (46,07%, 170/369) e o *te* (34,96%, 129/369).

Apresentamos, no próximo capítulo (capítulo 4), a distribuição dos dados em termos probabilísticos.

CAPÍTULO 4. AS ESTRATÉGIAS DATIVAS DE 2SG: DESCRIÇÃO-ANALÍTICA DOS RESULTADOS ESTATISTICAMENTE RELEVANTES

O capítulo 4 é dedicado à descrição e à análise dos resultados da alternância *te* versus *lhe, a ti, para ti, a você, para você, você, zero* (estratégias dativas de 2SG), tendo o *te* como o valor de aplicação nos níveis *step-up* & *step-down* do RBrul na interface do RStudio (OUSHIRO, 2014, p. 136). A opção pelo *te* como valor de aplicação está motivada pela conjectura da alta produtividade do *te* já comprovada nas análises voltadas em sincronias passadas (OLIVEIRA, 2014; CARDOSO, 2017) e recentes (DALTO, 2002; SILVA, 2016) como evidência inclusive da sua automação estrutural como traço da 2SG (LOPES, SOUZA & OLIVEIRA, 2013), o que fica evidente nas estruturas de redobro da fala mineira contemporânea (“Eu vou *te_i* levá *ocê_i* lá”, “Uma coisa eu vou *te_i* falá com *ocê_i*”, “Eu vou *te_i* contá pro *ocê_i* um pouquim da minha vida.”), cf. Duarte & Diniz (2012, p. 92). Isso posto, passamos aos resultados probabilísticos em relação ao *p-value*, ao *input* e aos *logodds* em relação a cada fator de cada grupo de fatores (parâmetros) selecionados e na ordem de seleção oferecida pelo programa de regra variável em questão.

Ao controlarmos a distribuição dos dados também a partir do *p-value*, assumimos a hipótese da não existência de correlação entre a variável dependente e as variáveis independentes. Se o valor do *p-value* for menor do que 0,05 (< 5%), entendemos que há o condicionamento significativo de aplicação da regra variável em análise. Por outro lado, caso o *p-value* seja igual ou maior do que 0,05 (> 5%), assumimos a hipótese da não-correlação entre a variável dependente e as independentes (modelo nulo).

O *input* representa o nível geral de produtividade do valor de aplicação assumido em relação à regra variável em análise. Índices numéricos acima de 0,40 (> 0,40) indicam a produtividade da regra variável, ao passo que índices abaixo de 0,40 (<0,40) expõem em evidência a não-produtividade da regra em questão.

Os índices de *logodds* evidenciam a aplicabilidade ou não da regra variável em análise. A identificação dos *logodds* está em conexão com os índices probabilísticos na forma dos pesos relativos. De um modo geral, identificamos os *logodds* positivos em relação à

aplicabilidade da regra variável em questão em contraposição aos valores negativos que contrariam a aplicação da regra variável em cena.

A rodada multivariada no RBrul em interface com o RStudio deu-se a partir da exclusão de dez (10) ocorrências das estratégias dativas de 2SG projetadas por verbos cujas semânticas são de movimento (sete ocorrências) e de interesse (três ocorrências), para que alcançássemos os níveis de *step up and step-down* com o devido *match*. Uma vez feito esse ajuste que nos levou ao total de 1222 ocorrências de estratégias dativas de 2SG, identificamos um *input* de 0.197 para a aplicação da regra variável movida pelo *te* como variável dependente em oposição às demais estratégias dativas de 2SG, o que nos conduz aos seguintes grupos de fatores selecionados na própria ordem de escolha oferecida pelo próprio RBrul (RStudio)²⁶: 1º) o contexto de sujeito de 2SG (0.197); 2º) a semântica do dativo (0.00219) e 3º) a categoria semântica do predicador (1.95e-34).

O fato de termos assumido o *te*, 2ª forma dativa mais produtiva na amostra em análise, em 33,69% dos dados (415/1232 oco), como valor de aplicação está em consonância com o *input* baixo de 0.197 para a aplicação da regra variável em questão (*te versus lhe, a ti, para ti, a você, para você, você, zero*), harmonizando-se ainda com o índice de *log-likelihood* negativo -691.181 (logaritmo de verossimilhança), o que é revelador em relação ao fato de termos em cena a segunda forma mais produtiva nas cartas ao PN, mas ainda de expressividade baixa em relação ao *zero-dativo*. A opção pelo *te* como variável dependente está motivada pelo fato de os estudos linguísticos terem evidenciado o *te* como uma expressiva estratégia dativa de 2SG tanto em amostras da língua falada brasileira (DALTO, 2002; ALMEIDA, 2014; SILVA, 2016), quanto em amostras da língua escrita de sincronias passadas (OLIVEIRA, 2014; RUMEU, 2014; LOPES *et alii*, 2018; CARDOSO, 2017; RUMEU, 2020). Passamos, na sequência, à análise descritivo-analítica dos resultados em termos probabilísticos voltados para as variáveis selecionadas: o contexto de sujeito de 2SG, a semântica do dativo e a semântica do predicador.

²⁶ Reprodução das informações dadas pelo RBrul (RStudio): “BEST STEP-UP MODEL OF RESPONSE pronome IS WITH PREDICTOR(S): CATEGORIA.DO.VERBO (1.95e-34) + TRATAMENTO (0.000671) + PAPEL.SEMANTICO (0.00219) [*p-values building from null model*]”

O contexto de sujeito de 2SG assume o *p-value* de 0.000671, valor muito próximo de zero, evidenciando tal contexto como propulsor do *te-dativo* (1º fator condicionante), cf. observamos na análise da tabela 9.

Tabela 9: O efeito do contexto de sujeito de 2SG em relação ao *te-dativo*

SUJEITO DE 2SG	TE (413/415 oco)	%	PR	<i>logodds</i>
Cartas de <i>você-sujeito</i> exclusivo	164 /411	40,1%	0.558	0.234
Cartas de FNTs exclusivo	249 /811	30,8%	0.442	-0.234

Valor de aplicação: *te*
p-value: **0.000671**
input: 0.197

O contexto de *sujeito* de 2SG apresenta o *p-value* de 0.00067, figurando como uma condição propícia à aplicação do *te-dativo* nas cartas ao PN, cf. tabela 8. A correlação dos valores em termos de índices percentuais, probabilísticos (PR) e de *logodds* expõe o contexto de *você-sujeito* como propulsor do *te-dativo*, corroborando os resultados de trabalhos que também se voltaram para a expressão variável do dativo de 2SG (OLIVEIRA, 2014; RUMEU, 2014; LOPES *et alii*, 2018; CARDOSO, 2017; RUMEU, 2020). No contexto de *você-sujeito*, o *te* assume o PR de 0.558 (=40,1%) aliado ao *logodd* de 0.234, encaminhando-se em uma direção diametralmente oposta ao contexto das cartas de FNTs, uma vez que o *logodd* é negativo (-0.234) e o índice probabilístico é abaixo de 0,5 (0.442 = 30,8%), o que permite interpretar as FNTs como contexto inibidor do *te-dativo* nas cartas ao PN analisadas.

Tabela 10: O efeito da semântica do dativo em relação ao te-dativo

SEMÂNTICA DO DATIVO	TE (413/415 oco)	%	PR	logodds
Meta/Fonte	346/771	45%	0.724	0.964
Receptor	66/443	15%	0.607	0.435
Beneficiário	1/8	12,5%	0.198	-1.400

Valor de aplicação: *te*
p-value: **0.00219**
input: 0.197

A semântica do dativo apresenta-se com *p-value* de **0.00219**, valor também muito próximo ao zero, o que justifica a seleção dessa variável pelo RBrul (RStudio) como o segundo contexto a impulsionar o *te-dativo* nas cartas ao PN em análise, cf. identificamos na tabela 9. No que se refere aos índices probabilísticos, estatísticos e de *logodds*, observamos que as semântica de *meta/fonte* e de *receptor* apresentam-se como propulsores do *te-dativo*. Nas cartas ao PN em análise, o *te* mostra-se fortemente condicionado pela semântica de *meta/fonte*, o que está consubstanciado nos altos índices de PR, **0.724** (= 45%), e de *logodd*, **0.964**. O dativo com a noção semântica de *receptor* mostrou-se também como contexto promissor ao *te-dativo* uma vez que obtivemos o PR de **0.607** (= 15%), ainda que o *logodd* de 0.435 tenha se mostrado abaixo de 0,5, o que, por outro lado, parece de acordo com a sua produtividade estatística baixa, em 15% dos dados, do *te-dativo* com a semântica de *receptor*. Por outro lado, observamos que a semântica de *beneficiário* com baixíssimos índices probabilístico (**0.198**), de *logodds* (-1.400) e percentual (= 12,5% = 1 ocorrência), mostra-se como um contexto improdutivo e, portanto, inibidor do *te-dativo* nas cartas ao PN em análise.

Tabela 11: O efeito da categoria semântica do predicador em relação ao te-dativo

CATEGORIA SEMÂNTICA DO PREDICADOR	TE (413/415 oco)	%	PR	logodds
Movimento abstrato	47/84	56%	0.722	0.957
Transferência material	270/597	45,4%	0.6	0.404
Movimento psicológico	59/181	33,1%	0.447	-0.214
Transferência verbal e perceptual	37/360	10,3%	0.241	-1.147

Valor de aplicação: *te*
p-value: **1.95e-34**
input: 0.197

A *semântica do predicador* mostra-se como o terceiro contexto motivador do *te-dativo*, o que está respaldado no *p-value* mais distante de zero, **1.95e-34**, como expressão do modelo nulo (tabela 11). As semânticas de *movimento abstrato* e de *transferência material* mostram-se como contextos propulsores do *te-dativo*, o que está respaldado nos índices probabilísticos de **0.722** (= 56%) e de **0.6** (= 45,4%) em correlação aos índices de *logodds* de **0.957** e **0.404**, respectivamente. Nas cartas ao PN, observamos a altíssima produtividade de verbos como *agradecer* (*movimento abstrato*), considerando que o contexto semântico das cartas ao PN é movido por pedidos para os quais já se *agradece* ao interlocutor antecipadamente. A semântica de *transferência material* também se mostra como um contexto propulsor do *te-dativo*, o que está consubstanciado principalmente no índice probabilístico de **0.6** (maior que 0,5). A noção de *transferência* caracteriza a semântica de projeção verbal do dativo, cf. observado por Berlinck (1996, 2005), por Oliveira (2014) e por Cardoso (2017).

As semânticas verbais de *movimento psicológico* e *transferência verbal e perceptual* apresentam-se como contextos inibidores do *te-dativo*. Para o significado de *movimento psicológico* observamos os baixos índices probabilísticos e de *logodds* (negativos), 0.447 (= 33,1%) e -0.214, respectivamente. Em relação à semântica de *transferência verbal e perceptual* também verificamos os baixíssimos índices probabilístico e de *logodds*, 0.241 (=

10,3%) e -1.147, respectivamente, evidenciando tais contextos semânticos como bloqueadores do *te-dativo* nas cartas ao PN em análise.

Síntese do capítulo

Organizamos, na tabela 12, os resultados probabilísticos para as variáveis independentes que se mostraram relevantes à aplicação do *te-dativo* nas cartas ao PN. São elas: *sujeito de 2SG*, *semântica de 2SG* e *categoria semântica do predicador*.

Tabela 12: Síntese das variáveis independentes estatisticamente relevantes à aplicação do *te-dativo*

VARIÁVEIS LINGUÍSTICAS	
SUJEITO DE 2SG (<i>p-value</i> : 0.000671)	Cartas de <i>você-sujeito</i> exclusivo: 0.558.
SEMÂNTICA DO DATIVO DE 2SG (<i>p-value</i> : 0.00219)	Meta/Fonte: 0.724. Receptor: 0.607.
CATEGORIA SEMÂNTICA DO PREDICADOR (<i>p-value</i> : 1.95e-34)	Movimento abstrato: 0.722. Transferência material: 0.6.
Valor de aplicação: <i>te</i> . <i>input</i> : 0.197.	

O contexto das cartas de *você-sujeito* de 2SG (exclusivo) mostra-se como uma variável dinamizadora do *te-dativo*. O alto índice probabilístico de 0.558 assumido pelo *te* nas cartas de *você-sujeito* evidencia tal contexto linguístico como propulsor desse clítico pronominal como complemento dativo de 2SG, corroborando os resultados de Oliveira (2014, p. 139) com os (3) cruzamentos de *te versus zero*, *te versus lhe*, *te versus SPreps* para as cartas cariocas, Rumeu (2015) para as cartas mineiras, Rumeu & Oliveira (2016) para as cartas amorosas, Cardoso (2017) e Rumeu (2020) para as cartas mineiras.

No que se refere à semântica do dativo de 2SG, observamos que o *te-dativo* é favorecido pelos contextos semânticos de *meta/fonte* (0.724) e de *receptor* (0.607). Considerando a ideia geral de que a semântica do OI, cf. discutido por Company Company (2006, p. 520), está orientada, de um modo geral, para a noção de *locus* ou *meta* final da transitividade, passando pelo *continuum* de envolvimento (afetação) também discutido por Company Company (2006) em relação aos contextos de *receptor* (maior envolvimento) e de *meta/fonte* (menor envolvimento).

A categoria semântica do predicador mostra-se como um contexto propulsor do *te-dativo*, tendo em vista especificamente os verbos de *movimento abstrato* (0.722) e *transferência material* (0.6). Trata-se de predicadores verbais transitivos cujas semânticas apontam não só para “uma aproximação abstrata entre uma entidade e um estado, sentimento, ideia ou outra entidade”, cf. Berlinck (1996, p. 133), mas também para estruturas sintáticas em que o sujeito gramatical e o objeto acusativo sejam propriedades do objeto dativo, cf. Berlinck (1996, p. 128).

CONSIDERAÇÕES FINAIS

Neste trabalho, investigamos a produtividade variável das formas pronominais dativas de 2SG em cartas brasileiras do século XX. Apresentamos, no capítulo 3, os resultados percentuais em relação à distribuição da variável dependente enária (*te, a ti, para ti, lhe, a você, para você, você* e *zero*) pelos contextos linguísticos e extralinguísticos. Passamos, no capítulo 4, aos resultados da rodada, níveis de *step-up* & *step-down* do RBrul na interface do RStudio, a partir do *te-dativo* como variável dependente. Isso posto, passamos às questões norteadoras desta dissertação, propondo uma resposta a cada uma delas.

- I. Quais seriam as estratégias dativas de 2SG selecionadas por escreventes brasileiros para a referência ao PN?

Nas cartas brasileiras, estão em concorrência o *zero* (48,14%, 594/1232) e o *te* (33,69%, 415/1232).

O fato de o *zero-dativo* (forma não filiada à paradigma pronominal algum) ter se mostrado como a estratégia mais produtiva se, por um lado, surpreende, uma vez que esperávamos o *te* em 1º lugar, parece acompanhar, por outro lado, os resultados das pesquisas voltadas ao PB em sua expressão escrita dos séculos XIX e XX (OLIVEIRA, 2014; CARDOSO, 2017), através das quais o *zero-dativo* também se mostrou, em níveis distintos, nas cartas de *você-sujeito* e mistas (*tu/você*), acompanhando “o avanço progressivo do dativo nulo (10% (1850-1901) → 25% (1902-1927) → 65% (1928-1950)) nas cartas mineiras de *você-sujeito*”, cf. também observado por Rumeu (2015).

A 2ª estratégia mais produtiva é o *te* (165/416, 30,64%), o que parece não só confirmar a hipótese de que a sua alta frequência de uso pode ter favorecido a automação estrutural consubstanciada através da afixação do clítico *te* como traço da 2SG (LOPES, SOUZA & OLIVEIRA, 2013), mas também corroborar a estimativa da sua alta produtividade no contexto de *você-sujeito* ter sido também alavancada por redatores que nasceram e, necessariamente, vivem, em 2012, no espaço geográfico de Minas Gerais, espaço em que prevalece o *você-sujeito* (SCHERRE *et alii*, 2015).

Para as demais estratégias dativas de 2SG, identificamos baixos índices de produtividade como observamos em relação às formas *lhe* (10,71%, 132/1232), *para você* (4,79%, 59/1234), *a você* (1,87%, 23/1234), *você* (0,73%, 7/1232), *para ti* (0,08%, 1/1232) e *a ti* (0,08%, 1/1232) nas cartas brasileiras analisadas.

II. Qual(is) seria(m) a(s) forma(s) pronominal(is) dativa(s) do paradigma de *tu* (*te*, *a ti*, *para ti*) que se mostra(m) suscetível(is) à convivência com o *você-sujeito*?

Em termos tão somente de formas dativas pronominais de 2SG, observamos o *você-sujeito* (*função nominativa*) em convivência com o clítico *te* (165/416, 30,64%) e com os sintagmas preposicionados *para você* (27/416, 6,49%) e *a você* (14/416, 3,37%) como evidências da fusão entre os paradigmas de 2ª e 3ª pessoas pronominais, conforme já discutido, à luz de amostras históricas do PB, por Lopes e Cavalcante (2011), Oliveira (2014), Rumeu (2013), Rumeu (2015), Lopes *et alii* (2018), Rumeu (2020). Ainda que o clítico *lhe* seja a forma etimológica para a expressão do dativo de 3SG, é interessante mencionar a sua produtividade na amostra brasileira de cartas do século XXI (41/416, 9,86%) para a referência ao interlocutor (2SG), o que pode, em alguma medida, apontar para traço formulaico.

III. Quais seriam os fatores linguísticos e/ou extralinguísticos que atuariam nesse processo de variação das formas dativas de complementação verbal de 2SG? Quais fatores estruturais tendem a refrear ou acelerar a variação em questão?

Nas cartas brasileiras do século XXI em análise, temos em evidência tão somente contextos linguísticos como variáveis estatisticamente relevantes à aplicação do *te-dativo*: o contexto de *sujeito de 2SG*, a *semântica do dativo de 2SG* e a *categoria semântica do predicador*. Para cada uma das variáveis linguísticas selecionadas, temos, em síntese, o *você-sujeito* (0.558), as semânticas do dativo de *meta/fonte* (0.724), de *receptor* (0.607) e os verbos de *movimento abstrato* (0.722) e de *transferência material* (0.6) como contextos propulsores do *te-dativo* de 2SG.

Em síntese, o fato de o *te-dativo* ter sido condicionado, em cartas do século XXI, pelos contextos de *você-sujeito* está em consonância com os resultados das análises embasadas em cartas produzidas por brasileiros nascidos e/ou residentes nos espaços geográficos do Rio de

Janeiro (OLIVEIRA, 2014) e de Minas Gerais (RUMEU, 2020; CARDOSO, 2017; RUMEU & OLIVEIRA, 2016; RUMEU, 2015) entre os séculos XIX e XX. Acrescentemos ainda o fato de as semânticas de *meta/fonte* e de *receptor* que como contextos prototípicos da complementação dativa também terem impulsionado o *te* em conformidade, pois, com a discussão de Company Company (2006) voltada para a semântica do objeto indireto. Além disso, temos os predicadores cujas semânticas são de *movimento abstrato* e de *transferência material* como condicionamentos do *te-dativo*, reforçando traços originais do dativo como discutido por Berlinck (1996, 2005), Oliveira (2014) e Cardoso (2017). Em suma, os contextos estruturais fomentadores do dativo de 2SG não só evidenciam a prototipicidade desse contexto de complementação verbal, mas também apontam para o caráter variável deste fenômeno linguístico no PB, como amplamente discutido por Berlinck (1996, 2005) e Gomes (2003).

REFERÊNCIAS BIBLIOGRÁFICAS

ALMEIDA, N. M. *Gramática metódica da língua portuguesa: curso único e completo*. 7. ed. São Paulo: Saraiva, 1955.

ALMEIDA, G. S. *Uso variável dos pronomes-objeto na expressão do dativo e do acusativo de segunda pessoa em Santo Antônio de Jesus – BA*. Salvador: UFBA, Tese (Doutorado em Língua e Cultura) do Programa de Pós-Graduação em Língua e Cultura, UFBA, Salvador, 2014.

AVELHEDA BANDEIRA, A. C. C.; SOUZA, S. C. G. *Passo a passo para uso do R-Brul*. 2017 (Curso de curta duração ministrado/Extensão).

BECHARA, E. *Moderna gramática portuguesa*. 37. ed. rev. e ampl. Rio de Janeiro: Editora Nova Fronteira, 2009 [1961].

BERLINCK, R. A. The Portuguese dative. In: VANBELLE, N.; VAN LANGENDONCK, N. (Ed.). *Case and grammatical relations across languages*. The dative, v. 1. Descriptive Studies. Amsterdam/Philadelphia: John Benjamins, 1996, p.119-151.

BERLINCK, R. A. O Objeto indireto no português brasileiro: um estudo diacrônico. In: MASSINI-CAGLIARI, G. *et al.* (orgs.). *Estudos de linguística histórica do português*. Araraquara: UNESP; São Paulo: Cultura Acadêmica Editora, 2005, p. 123-139.

BYBEE, J. Mechanisms of Change in Grammaticization: the role of frequency. In: JOSEPH, B. & JANDA, R. D. (eds). *The Handbook of Historical Linguistics*. Oxford: Blackwell, 2003, p. 602-623.

BRITO, O. R. M. “Faça o mundo te ouvir”. *A uniformidade de tratamento na história do Português Brasileiro*. Dissertação (Mestrado) – Universidade Estadual de Londrina, Londrina. 2001.

CALLOU, D.; BRANDÃO, S. Caracterização de áreas dialetais no Português do Brasil: análise de duas variáveis. In: SÁ JÚNIOR, L. A.; MARTINS, M. A. (Orgs.). *Rumos da*

linguística brasileira no século XXI: historiografia, gramática e ensino. 1ed. São Paulo: Blucher, 2016, v. 1, p. 97-122.

CARDOSO, N. D. *As estratégias de dativo de 2ª pessoa em cartas pessoais (séculos XIX e XX)*. 2017. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Federal de Minas Gerais, Faculdade de Letras.

COELHO, M. S. V. *Uma abordagem variacionista do uso da forma você no Norte de Minas*. 1999. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Federal de Minas Gerais, Faculdade de Letras.

COMPANY COMPANY, C. Gramaticalización, Género discursivo y otras variables en la difusión del cambio sintáctico. In: KABATEK, J. (ed.). *Sintaxis histórica del español y cambio lingüístico: Nuevas perspectivas desde las Tradiciones Discursivas*. Madrid/Frankfurt: Iberoamericana/Vervuert (= Lingüística Iberoamericana 31), 2008, p. 17-52.

COMPANY COMPANY, C. El objeto indirecto. In: COMPANY COMPANY, C. *Sintaxis histórica de la lengua española*. México: UNAM/FCE, 2006. p. 479-572.

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

CUNHA, C.; CINTRA, L. L. *Nova gramática do português contemporâneo*. 5ª ed. Rio de Janeiro: Lexikon Editora Digital, 2009 [1985].

CRUZ, I. A. A alternância tu/você em contextos sintáticos de complementação e de adjunção: estudo de cartas pessoais dos séculos XIX e XX. 2017. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Federal de Minas Gerais, Faculdade de Letras.

DALGAARD, P. *Introductory statistics with R*. 2.ed. New York, USA: Springer, 2008.

DALTO, C. D. L. *Estudo sociolingüístico dos pronomes-objeto de primeira e de segunda pessoas nas três capitais do sul do Brasil*. Curitiba: UFPR, 2002. Dissertação (Mestrado em Estudos Linguísticos) do Programa de Pós-Graduação em Letras, Faculdade de Letras, UFPR, Curitiba, 2002.

DUARTE, I.; BRITO, A. M. Predicação e classes de predicadores verbais. In: MATEUS, M. H. M. *et al.* (Orgs.). *Gramática da Língua Portuguesa*. 7. ed. Lisboa: Caminho, 2006, p. 179-203.

DUARTE, F. B.; DINIZ, C. R. Eu *te* falei *para* você: redobro de pronomes?. In: RAMOS J. M.; COELHO, S., (Org.). *Português Brasileiro Dialetal: temas gramaticais*. 1. ed. São Paulo: Mercado de Letras, 2012, p. 91-102.

DUARTE, I. Relações gramaticais, esquemas relacionais e ordem de palavras. In: MATEUS, M. H. M. *et alii.* *Gramática da língua portuguesa*, 5a ed, Lisboa, Caminho, 2003, 275-320.

ELSPASS, S. The Use of Private Letters and Diaries in Sociolinguistic Investigation. In: HERNÁNDEZ-CAMPOY, J. M.; CONDE SILVESTRE, J. C. *The Handbook of Historical Sociolinguistics*. Oxford: Wiley-Blackwell, 2012, p.156-169.

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

FARIA, E. *Gramática superior da Língua Latina*. Rio de Janeiro: Livraria Acadêmica, 1995 [1958].

GOMES, C. A. Variação e mudança na expressão do dativo no português brasileiro. In: PAIVA, M.C.; DUARTE, M.E. (org.) *Mudança linguística em tempo real*. Rio de Janeiro: Contracapa, 2003, p. 81-96.

GOMES, C. A. Para além dos Pacotes estatísticos Varbrul/GoldVarb e RBrul: qual a concepção de Gramática? In: *Revista do GELNE*, v. 14, n. 1/2, p. 257-272, 2012.

GONÇALVES, C. R. *Uma abordagem sociolinguística do uso das formas você, ocê e cê no português*. Tese (Doutorado em Letras) – Universidade de São Paulo, São Paulo, 2008.

GUY, G. Identidade Linguística da Comunidade de Fala: paralelismo interdialeto nos padrões de variação linguística. In: *Revista Organon*, v. 14, nº 28-29, fls 17-32, 2000.

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

HERNÁNDEZ-CAMPOY, J. M.; CONDE SILVESTRE, J. C. *The Handbook of Historical Sociolinguistics*. Oxford, Wiley-Blackwell, 2012.

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.

HERÊNIO, K. K. P. *Tu e você em uma perspectiva intra-linguística*. Dissertação (Mestrado em Letras) – Universidade Federal de Uberlândia, Uberlândia, 2006.

JOHNSON, K. *Quantitative methods in linguistics*. Malden. MA: Blackwell Pub., 2008.

KOCH, P. Diskurstraditionen: zu ihrem sprachtheoretischen Status und ihrer Dynamik. In: FRANK, Barbara/HAYE, Thomas/TOPHINKE, Doris (Hrsg.), *Gattungen mittelalterlicher Schriftlichkeit*, Tübingen: Narr 1997 (ScriptOralia, 99), 1997, 43-79.

KOCH. P.; OESTERREICHER, W. Oralidad y escrituralidad a la luz de la Teoría del Lenguaje. In: KOCH. P.; OESTERREICHER, W. *Lengua Hablada en La Rumania: español, francés, italiano*. Madrid: Editorial Gredos, 2007, p.20-42.

LABOV, W. *Sociolinguistics Patterns*. Philadelphia: University of Pennsylvania Press, 1991 [1972].

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

LOPES, C. R. S.; SOUZA, C. D.; MARCOTULIO, L. L.; OLIVEIRA, T. L.; LUCENA, R. O. P. A reorganização no sistema pronominal de 2ª pessoa na história do português brasileiro: outras relações gramaticais. In: LOPES, C. R. S. (Org.). *História do português brasileiro. Mudança sintática das classes de palavra: perspectiva funcionalista*. 1ed. São Paulo: Editora Contexto, 2018, v. 4, p. 106-189.

LOPES, C. R. S.; CAVALCANTE, S. R. O. A cronologia do voçamento no português brasileiro: expansão de *você-sujeito* e retenção do clítico-*te*. In: *Linguística*. Vol. 25: 30-65, 2011.

LOPES, C. R. S.; SOUZA, C. D.; OLIVEIRA, T. L. A frequência e o delineamento da gramática: a afixação do clítico *te* no português brasileiro. *Veredas* (UFJF. Online), v. 17, n. 2, p. 376-397, 2013.

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*. 17 (2): 137-161, 2012.

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. (Org.). *Descrição, história e aquisição do português brasileiro*. 1ª ed. Campinas: Pontes/FAPESP, 2007, p. 419-436.

LUFT, C. P. *Dicionário Prático de Regência Verbal*. São Paulo: Ática, 2006.

LUFT, C. P. *Moderna gramática brasileira*. 6. ed. Porto Alegre: Globo, 1985.

MACHADO ROCHA, R. Morfossintaxe de Caso nos Pronomes Pessoais do PB/MG atual. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Federal de Minas Gerais, Belo Horizonte. 2011.

MATEUS, M. H. M. *et alii*. (Orgs.). *Gramática da Língua Portuguesa*. Lisboa: Caminho, 2003, p.181-203.

MOTA, M. A. *A variação dos pronomes 'tu' e 'você' no português oral de São João da Ponte (MG)*. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Federal de Minas Gerais, Belo Horizonte, 2008.

OLIVEIRA, T. L. *Entre o Linguístico e o Social: Complementos Dativos de 2ª pessoa em Cartas Cariocas (1880-1980)*. Dissertação (Mestrado em Língua Portuguesa) - Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2014.

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.134-177.

OUSHIRO, L. *Introdução à Estatística para Linguistas (Version 1.0.2)*. Zenodo, 2017. Disponível em: < <https://zenodo.org/record/1202201>>. Acesso em: 18 nov. 2019.

PERES, E. P. *O uso do você, ocê, cê em Belo Horizonte: um estudo em tempo aparente e em tempo real*. 2006. 247 f. Tese (Doutorado em Letras: Linguística) Universidade Federal de Minas Gerais, Belo Horizonte, 2006.

PESSOA, F. C. R. *As realizações de dativo no português falado em Teresina (PI): uma análise variacionista*. Belo Horizonte: UFMG, 2017. Tese (Doutorado em Estudos Linguísticos) - Programa de Pós-Graduação em Estudos Linguísticos, UFMG, Belo Horizonte, 2017.

RAPOSO, E. B. P. Estrutura da frase. In.: RAPOSO, E. B. P.; NASCIMENTO, M. A. B.; MOTA, M. A. C.; SEGURA, L.; MENDES, A. (Orgs.). *Gramática do Português*. Volume I, Fundação Calouste Gulbenkian, Coimbra, Portugal, 2013, p. 303-398.

REIS, Z. M. *A variação 'Tu' e 'Você' no português falado e escrito em Lontra (MG)*. Dissertação (Mestrado Profissional em Letras) – Universidade Estadual de Montes Claros, Montes Claros, 2019.

ROCHA LIMA, C. H. *Gramática normativa da língua portuguesa*. 49. ed. rev. Rio de Janeiro: José Olympio, 2011 [1968].

RUMEU, M. C. B. Variation in the paradigms of *tu* and *você* subject and complements in letters from Minas Gerais, Brazil, 1860-1989. In: HUMMEL, M.; LOPES, C. R. S. (Orgs.). *Address in Portuguese and Spanish*. 1ed. GmbH, Berlin/Munich/Boston: De Gruyter, 2020, v. 1, p. 227-249. (<https://www.degruyter.com/view/title/573880>)

RUMEU, M. C. B.; OLIVEIRA, T. L. *A expressão da 2ª pessoa do singular em contextos de complementação e de adjunção: retratos do encaixamento estrutural e social*. *Revista Linguística (Online)*, v. 32, p. 25-46, 2016.

RUMEU, M. C. B. ‘Tu’ ou ‘você’, ‘te’ ou ‘lhe’?: a correlação entre as funções de sujeito e complemento verbal de 2ª pessoa. *Linguística*, v. 31-2, p. 83-109, dez. 2015.

RUMEU, M. C. B. A difusão do você pelos contextos sintáticos de complementação e de adjunção. *Revista Portuguesa de Humanidades*, v. 18, p. 91-114, 2014.

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

SAUSSURE, F. *Curso de linguística geral*. 27 ed. São Paulo: Cultrix, 2006.

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.

SILVA, D. E. R. *A variação na representação do objeto dativo e acusativo em Feira de Santana*. Dissertação (Mestrado em Estudos Linguísticos) - Universidade Estadual de Feira de Santana. 2016.

SILVA, S. C. *A variação dos pronomes tu e você na fala mineira de Ressaquinha (MG)*. Dissertação (Mestrado em Letras) – Universidade Federal de Ouro Preto, *Ouro Preto*, 2017.

TARALLO, F. *Pesquisa sociolinguística*. São Paulo: Ática, 2004.

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.

VITRAL, L.T.; MOREIRA, J. C. (Orgs). *Cartas que o Papai Noel não leu*. Belo Horizonte: LED-CEFET-MG, 2020.

WEINREICH, U.; LABOV, W.; HERZOG, M. *Fundamentos empíricos para uma teoria da mudança linguística*. São Paulo: Parábola, 2006 [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
1222	1	-0.665		0.34

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-783.027	1568.053	1568.056	0	0

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

Trying with TRATAMENTO...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
v	0.204	411		0.401		0.551
n	-0.204	811		0.308		0.449

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	2	-0.604		0.34			0.353

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-777.806	1559.612	1559.622	0.093	0.011

Run 1 (above) with TRATAMENTO is better than Run 0 without TRATAMENTO, p = 0.00123

Trying with TIPO.DE.VERBO...

\$TIPO.DE.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
d	0.041	1043		0.342		0.51
i	-0.041	179		0.324		0.49

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	2	-0.694		0.34			0.333

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-782.912	1569.825	1569.834	0.01	0

Run 2 (above) with TIPO.DE.VERBO is better than Run 0 without TIPO.DE.VERBO, p = 0.633

Trying with CATEGORIA.DO.VERBO...

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
t	0.943	84		0.560	0.72
m	0.519	597		0.454	0.627
p	0.002	181		0.331	0.5
v	-1.463	360		0.103	0.188

\$misc.1

n	df	intercept	overall	proportion	centered	input prob
1222	4	-0.703		0.34		0.331

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-703.088	1414.176	1414.209	0.377	0.198

Run 3 (above) with CATEGORIA.DO.VERBO is better than Run 0 without CATEGORIA.DO.VERBO, $p = 1.95e-34$

Trying with PAPEL.SEMANTICO...

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
m	1.090	771		0.450	0.748
r	-0.434	443		0.151	0.393
b	-0.656	8		0.125	0.342

\$misc.1

n	df	intercept	overall	proportion	centered	input prob
1222	3	-1.29		0.34		0.216

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-721.791	1449.581	1449.601	0.311	0.142

Run 4 (above) with PAPEL.SEMANTICO is better than Run 0 without PAPEL.SEMANTICO, $p = 2.54e-27$

Trying with FORMA.DO.OD...

\$FORMA.DO.OD

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
n	0.247	746		0.370	0.561
z	0.086	180		0.333	0.521
r	0.071	100		0.330	0.518
d	-0.403	196		0.235	0.401

\$misc.1

n	df	intercept	overall	proportion	centered	input prob
1222	4	-0.779		0.34		0.315

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-776.358	1560.717	1560.75	0.099	0.016

Run 5 (above) with FORMA.DO.OD is better than Run 0 without
FORMA.DO.OD, $p = 0.00396$

Trying with SECAO...

\$SECAO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
n	0.533	504		0.411		0.63
d	0.531	295		0.410		0.63
i	-0.453	412		0.206		0.389
p	-0.610	11		0.182		0.352

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	4	-0.894			0.34		0.29

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-755.893	1519.787	1519.82	0.207	0.063

Run 6 (above) with SECAO is better than Run 0 without SECAO, $p = 9.85e-12$

Trying with LOCAL...

\$LOCAL

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
2	16.625	1		1.000		> 0.999
D	16.625	1		1.000		> 0.999
W	3.157	4		0.750		0.959
Y	2.213	13		0.538		0.901
A	2.059	8		0.500		0.887
X	2.059	2		0.500		0.887
R	1.817	25		0.440		0.86
B	1.569	229		0.380		0.828
0	1.369	853		0.334		0.797
C	0.715	29		0.207		0.671
E	0.554	22		0.182		0.635
L	0.554	11		0.182		0.635
I	0.449	6		0.167		0.61
V	0.267	14		0.143		0.566
5	-12.508	1		0.000		< 0.001
F	-12.508	1		0.000		< 0.001
H	-12.508	1		0.000		< 0.001
M	-12.508	1		0.000		< 0.001

\$misc.1

	n	df	intercept	overall	proportion	centered	input	prob
	1222	18	-2.059			0.34		0.113

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-769.646	1575.292	1575.861	0.11	0.249

Run 7 (above) with LOCAL is better than Run 0 without LOCAL, p = 0.0617

Adding CATEGORIA.DO.VERBO...

STEP 2 - adding to Run 3 model with CATEGORIA.DO.VERBO

Trying with TRATAMENTO...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
v	0.23	411		0.401		0.557
n	-0.23	811		0.308		0.443

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
t	0.965	84		0.560		0.724
m	0.524	597		0.454		0.628
p	-0.018	181		0.331		0.496
v	-1.471	360		0.103		0.187

\$misc.1

	n	df	intercept	overall	proportion	centered	input	prob
	1222	5	-0.634			0.34		0.347

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-697.304	1404.609	1404.658	0.41	0.21

Run 8 (above) with CATEGORIA.DO.VERBO + TRATAMENTO is better than Run 3 without TRATAMENTO, p = 0.000671

Trying with TIPO.DE.VERBO...

\$TIPO.DE.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
d	7.151	1043		0.342		0.999	>20
i	-7.151	179		0.324		0.001	>20

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
p	10.703	181		0.331		> 0.999	>20
t	-2.624	84		0.560		0.068	>20
m	-3.048	597		0.454		0.045	>20
v	-5.030	360		0.103		0.006	>20

\$misc.1

	n	df	intercept	overall	proportion	centered	input	prob
1222	5		-4.287			0.34		0.014

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-700.857	1411.714	1411.763	0.382	0.259

Run 9 (above) with CATEGORIA.DO.VERBO + TIPO.DE.VERBO is better than Run 3 without TIPO.DE.VERBO, p = 0.0347

Trying with PAPEL.SEMANTICO...

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
t	0.941	84		0.560		0.719
m	0.401	597		0.454		0.599
p	-0.188	181		0.331		0.453
v	-1.155	360		0.103		0.24

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
m	0.960	771		0.450		0.723
r	0.453	443		0.151		0.611
b	-1.413	8		0.125		0.196

\$misc.1

	n	df	intercept	overall	proportion	centered	input	prob
1222	6		-1.474			0.34		0.186

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-697.09	1406.179	1406.248	0.405	0.206

Run 10 (above) with CATEGORIA.DO.VERBO + PAPEL.SEMANTICO is better than Run 3 without PAPEL.SEMANTICO, p = 0.00248

Trying with FORMA.DO.OD...

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
t	3.932	84		0.560		0.981	>20
m	3.483	597		0.454		0.97	>20
v	1.455	360		0.103		0.811	>20
p	-8.871	181		0.331		< 0.001	>20

\$FORMA.DO.OD

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
z	8.840	180		0.333		> 0.999	>20
d	-2.867	196		0.235		0.054	>20
r	-2.940	100		0.330		0.05	>20
n	-3.033	746		0.370		0.046	>20

\$misc.1

	n	df	intercept	overall	proportion	centered	input	prob
1222	7		-0.662			0.34		0.34

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-702.352	1418.704	1418.796	0.383	0.22

Run 11 (above) with CATEGORIA.DO.VERBO + FORMA.DO.OD is better than Run 3 without FORMA.DO.OD, $p = 0.689$

Trying with SECAO...

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
t	0.865	84		0.560		0.704
m	0.543	597		0.454		0.632
p	-0.112	181		0.331		0.472
v	-1.296	360		0.103		0.215

\$SECAO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
d	0.353	295		0.410		0.587
n	0.231	504		0.411		0.557
i	-0.050	412		0.206		0.487
p	-0.534	11		0.182		0.37

\$misc.1

	n	df	intercept	overall	proportion	centered	input	prob
1222	7		-0.905			0.34		0.288

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-700.748	1415.496	1415.588	0.403	0.203

Run 12 (above) with CATEGORIA.DO.VERBO + SECAO is better than Run 3 without SECAO, $p = 0.197$

Trying with LOCAL...

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
t	0.917	84		0.560		0.715
m	0.545	597		0.454		0.633
p	0.007	181		0.331		0.502
v	-1.470	360		0.103		0.187

\$LOCAL

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
D	16.706	1		1.000		> 0.999
2	16.168	1		1.000		> 0.999
W	3.491	4		0.750		0.97
Y	2.557	13		0.538		0.928
R	1.842	25		0.440		0.863
A	1.736	8		0.500		0.85

B	1.643	229	0.380	0.838
X	1.416	2	0.500	0.805
O	1.391	853	0.334	0.801
I	1.350	6	0.167	0.794
C	0.807	29	0.207	0.691
L	0.679	11	0.182	0.664
V	0.563	14	0.143	0.637
E	0.431	22	0.182	0.606
5	-12.426	1	0.000	< 0.001
H	-12.426	1	0.000	< 0.001
F	-12.964	1	0.000	< 0.001
M	-12.964	1	0.000	< 0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	21	-2.147		0.34			0.105

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-690.158	1422.317	1423.087	0.427	0.369

Run 13 (above) with CATEGORIA.DO.VERBO + LOCAL is better than Run 3 without LOCAL, p = 0.0771

Adding TRATAMENTO...

STEP 3 - adding to Run 8 model with CATEGORIA.DO.VERBO + TRATAMENTO

Trying with TIPO.DE.VERBO...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
v	0.223	411		0.401		0.556
n	-0.223	811		0.308		0.444

\$TIPO.DE.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
d	7.014	1043		0.342		0.999	>20
i	-7.014	179		0.324		0.001	>20

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
p	10.480	181		0.331		> 0.999	>20
t	-2.535	84		0.560		0.073	>20
m	-2.975	597		0.454		0.049	>20
v	-4.969	360		0.103		0.007	>20

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	6	-4.151		0.34			0.016

\$misc.2

```
log.likelihood      AIC      AICc    Dxy    R2
      -695.428 1402.856 1402.925 0.414 0.268
```

Run 14 (above) with CATEGORIA.DO.VERBO + TRATAMENTO + TIPO.DE.VERBO is better than Run 8 without TIPO.DE.VERBO, $p = 0.0527$

Trying with PAPEL.SEMANTICO...

\$TRATAMENTO

```
factor logodds tokens t/t+0+a+c+i+l+p+x centered factor weight
v      0.234      411                0.401                0.558
n     -0.234      811                0.308                0.442
```

\$CATEGORIA.DO.VERBO

```
factor logodds tokens t/t+0+a+c+i+l+p+x centered factor weight
t      0.957       84                0.560                0.722
m      0.404      597                0.454                0.6
p     -0.214      181                0.331                0.447
v     -1.147      360                0.103                0.241
```

\$PAPEL.SEMANTICO

```
factor logodds tokens t/t+0+a+c+i+l+p+x centered factor weight
m      0.964      771                0.450                0.724
r      0.435      443                0.151                0.607
b     -1.400       8                0.125                0.198
```

\$misc.1

```
n df intercept overall proportion centered input prob
1222 7      -1.402                0.34                0.197
```

\$misc.2

```
log.likelihood      AIC      AICc    Dxy    R2
      -691.181 1396.361 1396.453 0.436 0.219
```

Run 15 (above) with CATEGORIA.DO.VERBO + TRATAMENTO + PAPEL.SEMANTICO is better than Run 8 without PAPEL.SEMANTICO, $p = 0.00219$

Trying with FORMA.DO.OD...

\$TRATAMENTO

```
factor logodds tokens t/t+0+a+c+i+l+p+x centered factor weight
v      0.229      411                0.401                0.557
n     -0.229      811                0.308                0.443
```

\$CATEGORIA.DO.VERBO

```
factor logodds tokens t/t+0+a+c+i+l+p+x centered factor weight vif
t      3.908       84                0.560                0.98 >20
m      3.446      597                0.454                0.969 >20
v      1.404      360                0.103                0.803 >20
p     -8.758      181                0.331                < 0.001 >20
```

\$FORMA.DO.OD

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
z	8.717	180		0.333		> 0.999 >20
d	-2.803	196		0.235		0.057 >20
r	-2.939	100		0.330		0.05 >20
n	-2.975	746		0.370		0.049 >20

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	8	-0.604		0.34		0.353	

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-696.644	1409.288	1409.407	0.414	0.231

Run 16 (above) with CATEGORIA.DO.VERBO + TRATAMENTO + FORMA.DO.OD is better than Run 8 without FORMA.DO.OD, $p = 0.724$

Trying with SECAO...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
v	0.228	411		0.401	0.557
n	-0.228	811		0.308	0.443

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
t	0.888	84		0.560	0.709
m	0.550	597		0.454	0.634
p	-0.128	181		0.331	0.468
v	-1.310	360		0.103	0.212

\$SECAO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
d	0.366	295		0.410	0.59
n	0.239	504		0.411	0.56
i	-0.023	412		0.206	0.494
p	-0.582	11		0.182	0.359

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	8	-0.851		0.34		0.299	

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-695.088	1406.175	1406.294	0.424	0.215

Run 17 (above) with CATEGORIA.DO.VERBO + TRATAMENTO + SECAO is better than Run 8 without SECAO, $p = 0.218$

Trying with LOCAL...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
v	0.248	411		0.401	0.562
n	-0.248	811		0.308	0.438

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
t	0.944	84		0.560	0.72
m	0.551	597		0.454	0.634
p	-0.016	181		0.331	0.496
v	-1.479	360		0.103	0.186

\$LOCAL

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
D	16.416	1		1.000	> 0.999
2	16.346	1		1.000	> 0.999
W	3.605	4		0.750	0.974
Y	2.632	13		0.538	0.933
A	1.920	8		0.500	0.872
R	1.839	25		0.440	0.863
B	1.668	229		0.380	0.841
0	1.397	853		0.334	0.802
I	1.284	6		0.167	0.783
X	1.087	2		0.500	0.748
C	0.772	29		0.207	0.684
L	0.650	11		0.182	0.657
E	0.466	22		0.182	0.615
V	0.426	14		0.143	0.605
5	-12.220	1		0.000	< 0.001
H	-12.220	1		0.000	< 0.001
M	-12.787	1		0.000	< 0.001
F	-13.282	1		0.000	< 0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	22	-2.082		0.34			0.111

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-683.702	1411.403	1412.247	0.459	0.378

Run 18 (above) with CATEGORIA.DO.VERBO + TRATAMENTO + LOCAL is better than Run 8 without LOCAL, $p = 0.0551$

Adding PAPEL.SEMANTICO...

STEP 4 - adding to Run 15 model with CATEGORIA.DO.VERBO + TRATAMENTO + PAPEL.SEMANTICO

Trying with TIPO.DE.VERBO...

\$TRATAMENTO

```

factor logodds tokens t/t+0+a+c+i+l+p+x centered factor weight
  v    0.227    411                0.401                0.557
  n   -0.227    811                0.308                0.443

$TIPO.DE.VERBO
factor logodds tokens t/t+0+a+c+i+l+p+x centered factor weight vif
  d    7.011   1043                0.342                0.999 >20
  i   -7.011    179                0.324                0.001 >20

$CATEGORIA.DO.VERBO
factor logodds tokens t/t+0+a+c+i+l+p+x centered factor weight vif
  p   10.281    181                0.331                > 0.999 >20
  t   -2.542     84                0.560                0.073 >20
  m   -3.094    597                0.454                0.043 >20
  v   -4.645    360                0.103                0.01 >20

$PAPEL.SEMANTICO
factor logodds tokens t/t+0+a+c+i+l+p+x centered factor weight
  m    0.964    771                0.450                0.724
  r    0.435    443                0.151                0.607
  b   -1.400     8                 0.125                0.198

$misc.1
  n df intercept overall proportion centered input prob
1222 8      -4.917                0.34                0.007

$misc.2
log.likelihood      AIC      AICc  Dxy    R2
-689.31 1394.62 1394.738 0.44 0.276

Run 19 (above) with CATEGORIA.DO.VERBO + TRATAMENTO + PAPEL.SEMANTICO
+ TIPO.DE.VERBO is better than Run 15 without TIPO.DE.VERBO, p =
0.0531

Trying with FORMA.DO.OD...

$TRATAMENTO
factor logodds tokens t/t+0+a+c+i+l+p+x centered factor weight
  v    0.235    411                0.401                0.559
  n   -0.235    811                0.308                0.441

$CATEGORIA.DO.VERBO
factor logodds tokens t/t+0+a+c+i+l+p+x centered factor weight vif
  t    3.892     84                0.560                0.98 >20
  m    3.326    597                0.454                0.965 >20
  v    1.742    360                0.103                0.851 >20
  p   -8.960    181                0.331                < 0.001 >20

$PAPEL.SEMANTICO
factor logodds tokens t/t+0+a+c+i+l+p+x centered factor weight
  m    0.973    771                0.450                0.726
  r    0.420    443                0.151                0.603
  b   -1.393     8                 0.125                0.199

```

\$FORMA.DO.OD

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
z	8.730	180		0.333		> 0.999 >20
d	-2.763	196		0.235		0.059 >20
n	-2.957	746		0.370		0.049 >20
r	-3.010	100		0.330		0.047 >20

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	10	-1.387			0.34		0.2

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-690.373	1400.745	1400.927	0.437	0.24

Run 20 (above) with CATEGORIA.DO.VERBO + TRATAMENTO + PAPEL.SEMANTICO + FORMA.DO.OD is better than Run 15 without FORMA.DO.OD, $p = 0.656$

Trying with SECAO...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
v	0.233	411		0.401		0.558
n	-0.233	811		0.308		0.442

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
t	0.792	84		0.560		0.688 >5
m	0.496	597		0.454		0.621 >5
p	-0.423	181		0.331		0.396 >5
v	-0.865	360		0.103		0.296 >5

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
m	1.008	771		0.450		0.733 >2.5
r	0.376	443		0.151		0.593 >2.5
b	-1.383	8		0.125		0.201 >2.5

\$SECAO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
d	0.485	295		0.410		0.619 >2.5
n	0.146	504		0.411		0.537 >2.5
i	-0.046	412		0.206		0.489 >2.5
p	-0.586	11		0.182		0.358 >2.5

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	10	-1.655			0.34		0.16

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-688.276	1396.552	1396.734	0.445	0.225

Run 21 (above) with CATEGORIA.DO.VERBO + TRATAMENTO + PAPEL.SEMANTICO + SECAO is better than Run 15 without SECAO, $p = 0.121$

Trying with LOCAL...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
v	0.253	411		0.401	0.563
n	-0.253	811		0.308	0.437

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
t	0.934	84		0.560	0.718
m	0.431	597		0.454	0.606
p	-0.210	181		0.331	0.448
v	-1.155	360		0.103	0.24

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
m	0.953	771		0.450	0.722
r	0.424	443		0.151	0.605
b	-1.377	8		0.125	0.201

\$LOCAL

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
D	16.405	1	1.000		> 0.999
2	16.269	1	1.000		> 0.999
W	3.583	4	0.750		0.973
Y	2.603	13	0.538		0.931
A	1.928	8	0.500		0.873
R	1.869	25	0.440		0.866
B	1.653	229	0.380		0.839
0	1.396	853	0.334		0.802
I	1.282	6	0.167		0.783
X	0.946	2	0.500		0.72
C	0.833	29	0.207		0.697
L	0.594	11	0.182		0.644
E	0.405	22	0.182		0.6
V	0.381	14	0.143		0.594
5	-12.222	1	0.000		< 0.001
H	-12.222	1	0.000		< 0.001
M	-12.334	1	0.000		< 0.001
F	-13.369	1	0.000		< 0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	24		-2.835		0.34		0.055

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-677.813	1403.625	1404.628	0.476	0.384

Run 22 (above) with CATEGORIA.DO.VERBO + TRATAMENTO + PAPEL.SEMANTICO + LOCAL is better than Run 15 without LOCAL, $p = 0.0621$

No significant improvement this step, best model from last step is Run 15

BEST STEP-UP MODEL OF RESPONSE FORMA IS WITH PREDICTOR(S):
 CATEGORIA.DO.VERBO (1.95e-34) + TRATAMENTO (0.000671) +
 PAPEL.SEMANTICO (0.00219)
 [p-values building from null model]

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
v	0.234	411		0.401		0.558
n	-0.234	811		0.308		0.442

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
t	0.957	84		0.560		0.722
m	0.404	597		0.454		0.6
p	-0.214	181		0.331		0.447
v	-1.147	360		0.103		0.241

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
m	0.964	771		0.450		0.724
r	0.435	443		0.151		0.607
b	-1.400	8		0.125		0.198

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	7	-1.402		0.34			0.197

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-691.181	1396.361	1396.453	0.436	0.219

STEPPING DOWN...

STEP 0 - Run 0 - full model with TRATAMENTO + TIPO.DE.VERBO +
 CATEGORIA.DO.VERBO + PAPEL.SEMANTICO + FORMA.DO.OD + SECAO + LOCAL

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
v	0.249	411		0.401		0.562
n	-0.249	811		0.308		0.438

\$TIPO.DE.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
d	7.779	1043		0.342			> 0.999 >20
i	-7.779	179		0.324			< 0.001 >20

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
t	0.534	84		0.560	0.63	>20
m	0.339	597		0.454	0.584	>20
p	0.211	181		0.331	0.552	>20
v	-1.083	360		0.103	0.253	>20

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
m	1.014	771		0.450	0.734	>2.5
r	0.320	443		0.151	0.579	>2.5
b	-1.334	8		0.125	0.208	>2.5

\$FORMA.DO.OD

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
z	10.973	180		0.333	> 0.999	>20
d	-3.494	196		0.235	0.029	>20
n	-3.710	746		0.370	0.024	>20
r	-3.768	100		0.330	0.023	>20

\$SECAO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
d	0.519	295		0.410	0.627	>2.5
n	0.088	504		0.411	0.522	>2.5
i	-0.042	412		0.206	0.489	>2.5
p	-0.565	11		0.182	0.362	>2.5

\$LOCAL

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
D	17.495	1	1.000	> 0.999	
2	17.419	1	1.000	> 0.999	
W	3.768	4	0.750	0.977	
Y	2.644	13	0.538	0.934	
A	2.032	8	0.500	0.884	
R	2.016	25	0.440	0.882	
B	1.816	229	0.380	0.86	
O	1.543	853	0.334	0.824	
I	1.455	6	0.167	0.811	
X	1.043	2	0.500	0.739	
C	0.979	29	0.207	0.727	
L	0.741	11	0.182	0.677	
E	0.606	22	0.182	0.647	
V	0.469	14	0.143	0.615	
5	-13.140	1	0.000	< 0.001	
H	-13.140	1	0.000	< 0.001	
M	-13.665	1	0.000	< 0.001	
F	-14.080	1	0.000	< 0.001	

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	31	-7.114		0.34			0.001

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-672.616	1407.232	1408.9	0.48	0.464

STEP 1 - dropping from Run 0 model with TRATAMENTO + TIPO.DE.VERBO + CATEGORIA.DO.VERBO + PAPEL.SEMANTICO + FORMA.DO.OD + SECAO + LOCAL

Trying without TRATAMENTO...

\$TIPO.DE.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
d	7.954	1043		0.342		> 0.999 >20
i	-7.954	179		0.324		< 0.001 >20

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
t	0.484	84		0.560		0.619 >20
p	0.359	181		0.331		0.589 >20
m	0.295	597		0.454		0.573 >20
v	-1.138	360		0.103		0.243 >20

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
m	1.007	771		0.450		0.732 >2.5
r	0.359	443		0.151		0.589 >2.5
b	-1.365	8		0.125		0.203 >2.5

\$FORMA.DO.OD

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
z	11.104	180		0.333		> 0.999 >20
d	-3.570	196		0.235		0.027 >20
r	-3.759	100		0.330		0.023 >20
n	-3.775	746		0.370		0.022 >20

\$SECAO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
d	0.503	295		0.410		0.623 >2.5
n	0.081	504		0.411		0.52 >2.5
i	-0.072	412		0.206		0.482 >2.5
p	-0.511	11		0.182		0.375 >2.5

\$LOCAL

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
D	17.792	1		1.000	> 0.999
2	17.249	1		1.000	> 0.999
W	3.659	4		0.750	0.975
Y	2.539	13		0.538	0.927
R	2.010	25		0.440	0.882
A	1.835	8		0.500	0.862
B	1.787	229		0.380	0.857
O	1.539	853		0.334	0.823
I	1.527	6		0.167	0.822
X	1.378	2		0.500	0.799

C	1.013	29	0.207	0.734
L	0.777	11	0.182	0.685
V	0.597	14	0.143	0.645
E	0.572	22	0.182	0.639
5	-13.340	1	0.000	< 0.001
H	-13.340	1	0.000	< 0.001
F	-13.730	1	0.000	< 0.001
M	-13.863	1	0.000	< 0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	30	-7.244		0.34		0.001	

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-678.894	1417.788	1419.349	0.462	0.457

Run 1 (above) with TIPO.DE.VERBO + CATEGORIA.DO.VERBO + PAPEL.SEMANTICO + FORMA.DO.OD + SECAO + LOCAL is worse than Run 0 also including TRATAMENTO, p = 0.000395

Trying without TIPO.DE.VERBO...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
v	0.255	411		0.401		0.563
n	-0.255	811		0.308		0.437

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
t	4.184	84		0.560		0.985	>20
m	3.967	597		0.454		0.981	>20
v	2.534	360		0.103		0.926	>20
p	-10.685	181		0.331		< 0.001	>20

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
m	1.010	771		0.450		0.733	>2.5
r	0.328	443		0.151		0.581	>2.5
b	-1.338	8		0.125		0.208	>2.5

\$FORMA.DO.OD

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
z	10.229	180		0.333		> 0.999	>20
d	-3.241	196		0.235		0.038	>20
n	-3.464	746		0.370		0.03	>20
r	-3.523	100		0.330		0.029	>20

\$SECAO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
d	0.505	295		0.410		0.624	>2.5
n	0.105	504		0.411		0.526	>2.5
i	-0.034	412		0.206		0.492	>2.5
p	-0.576	11		0.182		0.36	>2.5

\$LOCAL

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
D	16.355	1	1.000		> 0.999
2	16.306	1	1.000		> 0.999
W	3.653	4	0.750		0.975
Y	2.656	13	0.538		0.934
A	1.918	8	0.500		0.872
R	1.898	25	0.440		0.87
B	1.713	229	0.380		0.847
O	1.425	853	0.334		0.806
I	1.329	6	0.167		0.791
X	0.922	2	0.500		0.715
C	0.864	29	0.207		0.703
L	0.622	11	0.182		0.651
E	0.491	22	0.182		0.62
V	0.349	14	0.143		0.586
5	-12.268	1	0.000		< 0.001
H	-12.268	1	0.000		< 0.001
M	-12.767	1	0.000		< 0.001
F	-13.197	1	0.000		< 0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	30	-3.103		0.34			0.043

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-673.969	1407.939	1409.501	0.476	0.406

Run 2 (above) with TRATAMENTO + CATEGORIA.DO.VERBO + PAPEL.SEMANTICO + FORMA.DO.OD + SECAO + LOCAL is worse than Run 0 also including TIPO.DE.VERBO, $p = 0.0999$

Trying without CATEGORIA.DO.VERBO...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
v	0.253	411	0.401		0.563
n	-0.253	811	0.308		0.437

\$TIPO.DE.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
d	7.907	1043	0.342		> 0.999	>20
i	-7.907	179	0.324		< 0.001	>20

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
m	1.345	771	0.450		0.793
r	-0.299	443	0.151		0.426
b	-1.046	8	0.125		0.26

\$FORMA.DO.OD

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
z	10.973	180	0.333	>	0.999	>20
d	-3.548	196	0.235		0.028	>20
n	-3.587	746	0.370		0.027	>20
r	-3.839	100	0.330		0.021	>20

\$SECAO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
d	0.713	295	0.410		0.671
n	0.024	504	0.411		0.506
i	-0.202	412	0.206		0.45
p	-0.536	11	0.182		0.369

\$LOCAL

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
D	17.446	1	1.000	>	0.999
2	17.387	1	1.000	>	0.999
W	3.581	4	0.750		0.973
Y	2.498	13	0.538		0.924
A	2.167	8	0.500		0.897
R	1.970	25	0.440		0.878
B	1.802	229	0.380		0.858
0	1.551	853	0.334		0.825
I	1.246	6	0.167		0.777
C	0.972	29	0.207		0.726
X	0.971	2	0.500		0.725
L	0.659	11	0.182		0.659
E	0.590	22	0.182		0.643
V	0.376	14	0.143		0.593
M	-12.830	1	0.000	<	0.001
5	-13.181	1	0.000	<	0.001
H	-13.181	1	0.000	<	0.001
F	-14.024	1	0.000	<	0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	28	-7.257		0.34			0.001

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-683.39	1422.78	1424.142	0.46	0.446

Run 3 (above) with TRATAMENTO + TIPO.DE.VERBO + PAPEL.SEMANTICO + FORMA.DO.OD + SECAO + LOCAL is worse than Run 0 also including CATEGORIA.DO.VERBO, $p = 8.1e-05$

Trying without PAPEL.SEMANTICO...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
v	0.24	411	0.401		0.56
n	-0.24	811	0.308		0.44

\$TIPO.DE.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
d	7.719	1043		0.342		> 0.999 >20
i	-7.719	179		0.324		< 0.001 >20

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
t	0.688	84		0.560		0.665 >20
p	0.465	181		0.331		0.614 >20
m	0.411	597		0.454		0.601 >20
v	-1.563	360		0.103		0.173 >20

\$FORMA.DO.OD

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
z	10.923	180		0.333		> 0.999 >20
d	-3.508	196		0.235		0.029 >20
r	-3.680	100		0.330		0.025 >20
n	-3.734	746		0.370		0.023 >20

\$SECAO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
d	0.394	295		0.410		0.597 >2.5
n	0.200	504		0.411		0.55 >2.5
i	-0.010	412		0.206		0.497 >2.5
p	-0.583	11		0.182		0.358 >2.5

\$LOCAL

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
D	17.524	1		1.000	> 0.999
2	17.470	1		1.000	> 0.999
W	3.786	4		0.750	0.978
Y	2.666	13		0.538	0.935
A	1.998	8		0.500	0.881
R	1.945	25		0.440	0.875
B	1.805	229		0.380	0.859
0	1.529	853		0.334	0.822
I	1.467	6		0.167	0.813
X	1.189	2		0.500	0.767
C	0.896	29		0.207	0.71
L	0.796	11		0.182	0.689
E	0.673	22		0.182	0.662
V	0.527	14		0.143	0.629
5	-13.128	1		0.000	< 0.001
H	-13.128	1		0.000	< 0.001
F	-13.932	1		0.000	< 0.001
M	-14.082	1		0.000	< 0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	29		-6.26		0.34		0.002

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-679.629	1417.258	1418.718	0.466	0.459

Run 4 (above) with TRATAMENTO + TIPO.DE.VERBO + CATEGORIA.DO.VERBO + FORMA.DO.OD + SECAO + LOCAL is worse than Run 0 also including PAPEL.SEMANTICO, $p = 9e-04$

Trying without FORMA.DO.OD...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
v	0.247	411		0.401		0.561
n	-0.247	811		0.308		0.439

\$TIPO.DE.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
d	7.3	1043		0.342		0.999	>20
i	-7.3	179		0.324		0.001	>20

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
p	10.473	181		0.331		> 0.999	>20
t	-2.924	84		0.560		0.051	>20
m	-3.084	597		0.454		0.044	>20
v	-4.465	360		0.103		0.011	>20

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
m	1.008	771		0.450		0.733	>2.5
r	0.326	443		0.151		0.581	>2.5
b	-1.335	8		0.125		0.208	>2.5

\$SECAO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
d	0.517	295		0.410		0.626	>2.5
n	0.061	504		0.411		0.515	>2.5
i	-0.058	412		0.206		0.486	>2.5
p	-0.520	11		0.182		0.373	>2.5

\$LOCAL

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
D	16.392	1		1.000		> 0.999
2	16.297	1		1.000		> 0.999
W	3.624	4		0.750		0.974
Y	2.507	13		0.538		0.925
A	1.920	8		0.500		0.872
R	1.899	25		0.440		0.87
B	1.694	229		0.380		0.845
O	1.438	853		0.334		0.808
I	1.316	6		0.167		0.789
X	0.930	2		0.500		0.717
C	0.851	29		0.207		0.701
L	0.621	11		0.182		0.651
E	0.482	22		0.182		0.618
V	0.342	14		0.143		0.585
5	-12.247	1		0.000		< 0.001

H	-12.247	1	0.000	< 0.001
M	-12.609	1	0.000	< 0.001
F	-13.210	1	0.000	< 0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	28		-6.77		0.34		0.001

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-673.499	1402.998	1404.359	0.481	0.429

Run 5 (above) with TRATAMENTO + TIPO.DE.VERBO + CATEGORIA.DO.VERBO + PAPEL.SEMANTICO + SECAO + LOCAL is worse than Run 0 also including FORMA.DO.OD, p = 0.622

Trying without SECAO...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
v	0.248	411		0.401		0.562
n	-0.248	811		0.308		0.438

\$TIPO.DE.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
d	7.691	1043		0.342		> 0.999	>20
i	-7.691	179		0.324		< 0.001	>20

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
t	0.765	84		0.560		0.682	>20
p	0.382	181		0.331		0.594	>20
m	0.241	597		0.454		0.56	>20
v	-1.388	360		0.103		0.2	>20

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
m	0.962	771		0.450		0.723
r	0.411	443		0.151		0.601
b	-1.373	8		0.125		0.202

\$FORMA.DO.OD

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
z	10.903	180		0.333		> 0.999	>20
d	-3.460	196		0.235		0.03	>20
n	-3.702	746		0.370		0.024	>20
r	-3.741	100		0.330		0.023	>20

\$LOCAL

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
D	17.537	1		1.000		> 0.999
2	17.399	1		1.000		> 0.999
W	3.738	4		0.750		0.977
Y	2.618	13		0.538		0.932

A	2.038	8	0.500	0.885
R	1.992	25	0.440	0.88
B	1.769	229	0.380	0.854
O	1.508	853	0.334	0.819
I	1.442	6	0.167	0.809
X	1.074	2	0.500	0.745
C	0.963	29	0.207	0.724
L	0.729	11	0.182	0.675
E	0.529	22	0.182	0.629
V	0.517	14	0.143	0.626
5	-13.098	1	0.000	< 0.001
H	-13.098	1	0.000	< 0.001
M	-13.425	1	0.000	< 0.001
F	-14.231	1	0.000	< 0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	28	-6.776		0.34			0.001

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-675.625	1407.25	1408.611	0.477	0.461

Run 6 (above) with TRATAMENTO + TIPO.DE.VERBO + CATEGORIA.DO.VERBO + PAPEL.SEMANTICO + FORMA.DO.OD + LOCAL is worse than Run 0 also including SECAO, p = 0.111

Trying without LOCAL...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
v	0.228	411		0.401		0.557
n	-0.228	811		0.308		0.443

\$TIPO.DE.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
d	7.063	1043		0.342		0.999	>20
i	-7.063	179		0.324		0.001	>20

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
p	0.614	181		0.331		0.649	>20
t	0.458	84		0.560		0.612	>20
m	0.156	597		0.454		0.539	>20
v	-1.227	360		0.103		0.227	>20

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
m	1.017	771		0.450		0.734	>2.5
r	0.358	443		0.151		0.589	>2.5
b	-1.375	8		0.125		0.202	>2.5

\$FORMA.DO.OD

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
--------	---------	--------	-------------------	----------	--------	--------	-----

z	9.519	180	0.333	> 0.999	>20
d	-3.034	196	0.235	0.046	>20
n	-3.210	746	0.370	0.039	>20
r	-3.275	100	0.330	0.036	>20

\$SECAO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
d	0.502	295	0.410		0.623	>2.5
n	0.156	504	0.411		0.539	>2.5
i	-0.036	412	0.206		0.491	>2.5
p	-0.622	11	0.182		0.349	>2.5

\$misc.1

n	df	intercept	overall	proportion	centered	input prob
1222	14	-5.198		0.34		0.005

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-685.508	1399.016	1399.364	0.448	0.303

Run 7 (above) with TRATAMENTO + TIPO.DE.VERBO + CATEGORIA.DO.VERBO + PAPEL.SEMANTICO + FORMA.DO.OD + SECAO is worse than Run 0 also including LOCAL, p = 0.0785

Dropping FORMA.DO.OD...

STEP 2 - dropping from Run 5 model with TRATAMENTO + TIPO.DE.VERBO + CATEGORIA.DO.VERBO + PAPEL.SEMANTICO + SECAO + LOCAL

Trying without TRATAMENTO...

\$TIPO.DE.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
d	7.473	1043	0.342		0.999	>20
i	-7.473	179	0.324		0.001	>20

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
p	10.756	181	0.331	>	0.999	>20
t	-3.026	84	0.560		0.046	>20
m	-3.174	597	0.454		0.04	>20
v	-4.556	360	0.103		0.01	>20

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
m	1.006	771	0.450		0.732	>2.5
r	0.359	443	0.151		0.589	>2.5
b	-1.365	8	0.125		0.203	>2.5

\$SECAO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
d	0.502	295	0.410		0.623	>2.5
n	0.054	504	0.411		0.514	>2.5
i	-0.089	412	0.206		0.478	>2.5
p	-0.467	11	0.182		0.385	>2.5

\$LOCAL

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
D	16.686	1	1.000		> 0.999
2	16.119	1	1.000		> 0.999
W	3.511	4	0.750		0.971
Y	2.402	13	0.538		0.917
R	1.901	25	0.440		0.87
A	1.737	8	0.500		0.85
B	1.668	229	0.380		0.841
0	1.431	853	0.334		0.807
I	1.386	6	0.167		0.8
X	1.255	2	0.500		0.778
C	0.896	29	0.207		0.71
L	0.652	11	0.182		0.657
V	0.483	14	0.143		0.618
E	0.449	22	0.182		0.61
5	-12.446	1	0.000		< 0.001
H	-12.446	1	0.000		< 0.001
M	-12.813	1	0.000		< 0.001
F	-12.869	1	0.000		< 0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	27	-6.912		0.34			0.001

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-679.728	1413.456	1414.723	0.462	0.421

Run 8 (above) with TIPO.DE.VERBO + CATEGORIA.DO.VERBO + PAPEL.SEMANTICO + SECAO + LOCAL is worse than Run 5 also including TRATAMENTO, $p = 0.000416$

Trying without TIPO.DE.VERBO...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
v	0.253	411	0.401		0.563
n	-0.253	811	0.308		0.437

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
t	0.734	84	0.560		0.676	>5
m	0.552	597	0.454		0.635	>5
p	-0.446	181	0.331		0.39	>5
v	-0.839	360	0.103		0.302	>5

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
m	1.004	771	0.450		0.732	>2.5
r	0.334	443	0.151		0.583	>2.5
b	-1.338	8	0.125		0.208	>2.5

\$SECAO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
d	0.503	295	0.410		0.623	>2.5
n	0.077	504	0.411		0.519	>2.5
i	-0.050	412	0.206		0.488	>2.5
p	-0.530	11	0.182		0.37	>2.5

\$LOCAL

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
D	16.362	1	1.000		> 0.999
2	16.295	1	1.000		> 0.999
W	3.619	4	0.750		0.974
Y	2.630	13	0.538		0.933
A	1.918	8	0.500		0.872
R	1.893	25	0.440		0.869
B	1.702	229	0.380		0.846
0	1.431	853	0.334		0.807
I	1.301	6	0.167		0.786
X	0.920	2	0.500		0.715
C	0.846	29	0.207		0.7
L	0.613	11	0.182		0.649
E	0.477	22	0.182		0.617
V	0.333	14	0.143		0.582
5	-12.264	1	0.000		< 0.001
H	-12.264	1	0.000		< 0.001
M	-12.594	1	0.000		< 0.001
F	-13.216	1	0.000		< 0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	27	-3.109		0.34			0.043

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-674.891	1403.782	1405.048	0.479	0.388

Run 9 (above) with TRATAMENTO + CATEGORIA.DO.VERBO + PAPEL.SEMANTICO + SECAO + LOCAL is worse than Run 5 also including TIPO.DE.VERBO, p = 0.0952

Trying without CATEGORIA.DO.VERBO...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
v	0.25	411	0.401		0.562
n	-0.25	811	0.308		0.438

\$TIPO.DE.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
d	0.625	1043		0.342		0.651
i	-0.625	179		0.324		0.349

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
m	1.333	771		0.450		0.791
r	-0.298	443		0.151		0.426
b	-1.035	8		0.125		0.262

\$SECAO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
d	0.723	295		0.410		0.673
n	0.023	504		0.411		0.506
i	-0.200	412		0.206		0.45
p	-0.545	11		0.182		0.367

\$LOCAL

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
D	16.344	1	1.000			> 0.999
2	16.295	1	1.000			> 0.999
W	3.481	4	0.750			0.97
Y	2.433	13	0.538			0.919
A	2.009	8	0.500			0.882
R	1.824	25	0.440			0.861
B	1.686	229	0.380			0.844
0	1.442	853	0.334			0.809
I	1.142	6	0.167			0.758
X	0.879	2	0.500			0.707
C	0.807	29	0.207			0.691
L	0.556	11	0.182			0.635
E	0.474	22	0.182			0.616
V	0.224	14	0.143			0.556
M	-11.905	1	0.000			< 0.001
5	-12.288	1	0.000			< 0.001
H	-12.288	1	0.000			< 0.001
F	-13.114	1	0.000			< 0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	25		-3.46		0.34		0.03

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-684.968	1419.935	1421.022	0.458	0.364

Run 10 (above) with TRATAMENTO + TIPO.DE.VERBO + PAPEL.SEMANTICO + SECAO + LOCAL is worse than Run 5 also including CATEGORIA.DO.VERBO, p = 4.16e-05

Trying without PAPEL.SEMANTICO...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
--------	---------	--------	-------------------	----------	--------	--------

v	0.241	411	0.401	0.56
n	-0.241	811	0.308	0.44

\$TIPO.DE.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
d	7.237	1043	0.342		0.999	>20
i	-7.237	179	0.324		0.001	>20

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
p	10.690	181	0.331	>	0.999	>20
t	-2.771	84	0.560		0.059	>20
m	-3.001	597	0.454		0.047	>20
v	-4.918	360	0.103		0.007	>20

\$SECAO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
d	0.391	295	0.410		0.597	>2.5
n	0.170	504	0.411		0.542	>2.5
i	-0.030	412	0.206		0.492	>2.5
p	-0.531	11	0.182		0.37	>2.5

\$LOCAL

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
D	16.416	1	1.000	>	0.999
2	16.336	1	1.000	>	0.999
W	3.631	4	0.750		0.974
Y	2.528	13	0.538		0.926
A	1.906	8	0.500		0.871
R	1.840	25	0.440		0.863
B	1.686	229	0.380		0.844
0	1.421	853	0.334		0.805
I	1.319	6	0.167		0.789
X	1.063	2	0.500		0.743
C	0.787	29	0.207		0.687
L	0.667	11	0.182		0.661
E	0.549	22	0.182		0.634
V	0.417	14	0.143		0.603
5	-12.235	1	0.000	<	0.001
H	-12.235	1	0.000	<	0.001
M	-13.018	1	0.000	<	0.001
F	-13.077	1	0.000	<	0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	26	-5.935	0.34				0.003

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-680.48	1412.961	1414.136	0.465	0.422

Run 11 (above) with TRATAMENTO + TIPO.DE.VERBO + CATEGORIA.DO.VERBO + SECAO + LOCAL is worse than Run 5 also including PAPEL.SEMANTICO, p = 0.000929

Trying without SECAO...

\$STRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
v	0.247	411		0.401		0.561
n	-0.247	811		0.308		0.439

\$TIPO.DE.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
d	7.205	1043		0.342		0.999	>20
i	-7.205	179		0.324		0.001	>20

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
p	10.582	181		0.331		> 0.999	>20
t	-2.664	84		0.560		0.065	>20
m	-3.167	597		0.454		0.04	>20
v	-4.750	360		0.103		0.009	>20

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
m	0.954	771		0.450		0.722
r	0.425	443		0.151		0.605
b	-1.379	8		0.125		0.201

\$LOCAL

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
D	16.434	1		1.000		> 0.999
2	16.267	1		1.000		> 0.999
W	3.586	4		0.750		0.973
Y	2.477	13		0.538		0.923
A	1.931	8		0.500		0.873
R	1.873	25		0.440		0.867
B	1.644	229		0.380		0.838
0	1.401	853		0.334		0.802
I	1.296	6		0.167		0.785
X	0.955	2		0.500		0.722
C	0.838	29		0.207		0.698
L	0.601	11		0.182		0.646
E	0.410	22		0.182		0.601
V	0.393	14		0.143		0.597
5	-12.204	1		0.000		< 0.001
H	-12.204	1		0.000		< 0.001
M	-12.337	1		0.000		< 0.001
F	-13.359	1		0.000		< 0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	25	-6.445		0.34			0.002

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-676.605	1403.21	1404.297	0.477	0.425

Run 12 (above) with TRATAMENTO + TIPO.DE.VERBO + CATEGORIA.DO.VERBO + PAPEL.SEMANTICO + LOCAL is worse than Run 5 also including SECAO, p = 0.102

Trying without LOCAL...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
v	0.226	411		0.401		0.556
n	-0.226	811		0.308		0.444

\$TIPO.DE.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
d	7.073	1043		0.342		0.999	>20
i	-7.073	179		0.324		0.001	>20

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
p	10.150	181		0.331	>	0.999	>20
t	-2.749	84		0.560		0.06	>20
m	-3.023	597		0.454		0.046	>20
v	-4.377	360		0.103		0.012	>20

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
m	1.011	771		0.450		0.733	>2.5
r	0.367	443		0.151		0.591	>2.5
b	-1.378	8		0.125		0.201	>2.5

\$SECAO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight	vif
d	0.501	295		0.410		0.623	>2.5
n	0.133	504		0.411		0.533	>2.5
i	-0.050	412		0.206		0.487	>2.5
p	-0.583	11		0.182		0.358	>2.5

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	11		-5.204		0.34		0.005

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-686.262	1394.524	1394.742	0.449	0.282

Run 13 (above) with TRATAMENTO + TIPO.DE.VERBO + CATEGORIA.DO.VERBO + PAPEL.SEMANTICO + SECAO is worse than Run 5 also including LOCAL, p = 0.0835

Dropping SECAO...

STEP 3 - dropping from Run 12 model with TRATAMENTO + TIPO.DE.VERBO + CATEGORIA.DO.VERBO + PAPEL.SEMANTICO + LOCAL

Trying without TRATAMENTO...

\$TIPO.DE.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
d	7.38	1043		0.342	0.999	>20
i	-7.38	179		0.324	0.001	>20

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
p	10.872	181		0.331	> 0.999	>20
t	-2.768	84		0.560	0.059	>20
m	-3.256	597		0.454	0.037	>20
v	-4.847	360		0.103	0.008	>20

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
m	0.950	771		0.450	0.721
r	0.450	443		0.151	0.611
b	-1.399	8		0.125	0.198

\$LOCAL

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
D	16.727	1		1.000	> 0.999
2	16.095	1		1.000	> 0.999
W	3.471	4		0.750	0.97
Y	2.375	13		0.538	0.915
R	1.880	25		0.440	0.868
A	1.748	8		0.500	0.852
B	1.619	229		0.380	0.835
0	1.395	853		0.334	0.801
I	1.363	6		0.167	0.796
X	1.285	2		0.500	0.783
C	0.881	29		0.207	0.707
L	0.634	11		0.182	0.653
V	0.534	14		0.143	0.631
E	0.380	22		0.182	0.594
5	-12.406	1		0.000	< 0.001
H	-12.406	1		0.000	< 0.001
M	-12.537	1		0.000	< 0.001
F	-13.037	1		0.000	< 0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	24		-6.602		0.34		0.001

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-682.898	1413.797	1414.799	0.451	0.417

Run 14 (above) with TIPO.DE.VERBO + CATEGORIA.DO.VERBO + PAPEL.SEMANTICO + LOCAL is worse than Run 12 also including TRATAMIENTO, $p = 0.000388$

Trying without TIPO.DE.VERBO...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
v	0.253	411		0.401		0.563
n	-0.253	811		0.308		0.437

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
t	0.934	84		0.560		0.718
m	0.431	597		0.454		0.606
p	-0.210	181		0.331		0.448
v	-1.155	360		0.103		0.24

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
m	0.953	771		0.450		0.722
r	0.424	443		0.151		0.605
b	-1.377	8		0.125		0.201

\$LOCAL

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
D	16.405	1		1.000		> 0.999
2	16.269	1		1.000		> 0.999
W	3.583	4		0.750		0.973
Y	2.603	13		0.538		0.931
A	1.928	8		0.500		0.873
R	1.869	25		0.440		0.866
B	1.653	229		0.380		0.839
0	1.396	853		0.334		0.802
I	1.282	6		0.167		0.783
X	0.946	2		0.500		0.72
C	0.833	29		0.207		0.697
L	0.594	11		0.182		0.644
E	0.405	22		0.182		0.6
V	0.381	14		0.143		0.594
5	-12.222	1		0.000		< 0.001
H	-12.222	1		0.000		< 0.001
M	-12.334	1		0.000		< 0.001
F	-13.369	1		0.000		< 0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	24		-2.835		0.34		0.055

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-677.813	1403.625	1404.628	0.476	0.384

Run 15 (above) with TRATAMENTO + CATEGORIA.DO.VERBO + PAPEL.SEMANTICO + LOCAL is worse than Run 12 also including TIPO.DE.VERBO, p = 0.12

Trying without CATEGORIA.DO.VERBO...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
--------	---------	--------	-------------------	----------	--------	--------

v	0.25	411	0.401	0.562
n	-0.25	811	0.308	0.438

\$TIPO.DE.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
d	0.358	1043	0.342			0.589
i	-0.358	179	0.324			0.411

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
m	1.194	771	0.450			0.767
r	-0.515	443	0.151			0.374
b	-0.679	8	0.125			0.337

\$LOCAL

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
D	16.405	1	1.000		>	0.999
2	16.189	1	1.000		>	0.999
W	3.366	4	0.750			0.967
Y	2.424	13	0.538			0.919
A	1.996	8	0.500			0.88
R	1.743	25	0.440			0.851
B	1.625	229	0.380			0.835
0	1.385	853	0.334			0.8
X	1.123	2	0.500			0.755
I	1.048	6	0.167			0.74
C	0.773	29	0.207			0.684
L	0.477	11	0.182			0.617
E	0.321	22	0.182			0.58
V	0.258	14	0.143			0.564
M	-11.235	1	0.000		<	0.001
5	-12.227	1	0.000		<	0.001
H	-12.227	1	0.000		<	0.001
F	-13.443	1	0.000		<	0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	22	-2.925		0.34			0.051

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-695.011	1434.022	1434.866	0.44	0.35

Run 16 (above) with TRATAMENTO + TIPO.DE.VERBO + PAPEL.SEMANTICO + LOCAL is worse than Run 12 also including CATEGORIA.DO.VERBO, $p = 5.04e-08$

Trying without PAPEL.SEMANTICO...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
v	0.242	411	0.401			0.56
n	-0.242	811	0.308			0.44

\$TIPO.DE.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
d	7.199	1043		0.342		0.999 >20
i	-7.199	179		0.324		0.001 >20

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
p	10.767	181		0.331		> 0.999 >20
t	-2.651	84		0.560		0.066 >20
m	-3.045	597		0.454		0.045 >20
v	-5.071	360		0.103		0.006 >20

\$LOCAL

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
D	16.445	1		1.000		> 0.999
2	16.343	1		1.000		> 0.999
W	3.608	4		0.750		0.974
Y	2.509	13		0.538		0.925
A	1.922	8		0.500		0.872
R	1.842	25		0.440		0.863
B	1.660	229		0.380		0.84
0	1.402	853		0.334		0.802
I	1.298	6		0.167		0.785
X	1.096	2		0.500		0.749
C	0.776	29		0.207		0.685
L	0.657	11		0.182		0.658
E	0.472	22		0.182		0.616
V	0.438	14		0.143		0.608
5	-12.203	1		0.000		< 0.001
H	-12.203	1		0.000		< 0.001
M	-12.789	1		0.000		< 0.001
F	-13.273	1		0.000		< 0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	23		-5.689		0.34		0.003

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-682.503	1411.005	1411.927	0.46	0.42

Run 17 (above) with TRATAMENTO + TIPO.DE.VERBO + CATEGORIA.DO.VERBO + LOCAL is worse than Run 12 also including PAPEL.SEMANTICO, p = 0.00275

Trying without LOCAL...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
v	0.227	411		0.401		0.557
n	-0.227	811		0.308		0.443

\$TIPO.DE.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
d	7.011	1043		0.342		0.999 >20

i	-7.011	179	0.324	0.001	>20
---	--------	-----	-------	-------	-----

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight	vif
p	10.281	181	0.331	>	0.999	>20
t	-2.542	84	0.560		0.073	>20
m	-3.094	597	0.454		0.043	>20
v	-4.645	360	0.103		0.01	>20

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
m	0.964	771	0.450		0.724
r	0.435	443	0.151		0.607
b	-1.400	8	0.125		0.198

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	8	-4.917	0.34				0.007

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-689.31	1394.62	1394.738	0.44	0.276

Run 18 (above) with TRATAMENTO + TIPO.DE.VERBO + CATEGORIA.DO.VERBO + PAPEL.SEMANTICO is worse than Run 12 also including LOCAL, p = 0.0859

Dropping TIPO.DE.VERBO...

STEP 4 - dropping from Run 15 model with TRATAMENTO + CATEGORIA.DO.VERBO + PAPEL.SEMANTICO + LOCAL

Trying without TRATAMENTO...

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
t	0.915	84	0.560		0.714
m	0.429	597	0.454		0.606
p	-0.179	181	0.331		0.455
v	-1.166	360	0.103		0.238

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
m	0.948	771	0.450		0.721
r	0.449	443	0.151		0.61
b	-1.398	8	0.125		0.198

\$LOCAL

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor weight
D	16.700	1	1.000	>	0.999
2	16.092	1	1.000	>	0.999
W	3.464	4	0.750		0.97
Y	2.527	13	0.538		0.926
R	1.875	25	0.440		0.867
A	1.740	8	0.500		0.851

B	1.627	229	0.380	0.836
O	1.389	853	0.334	0.8
I	1.348	6	0.167	0.794
X	1.283	2	0.500	0.783
C	0.876	29	0.207	0.706
L	0.627	11	0.182	0.652
V	0.524	14	0.143	0.628
E	0.372	22	0.182	0.592
5	-12.432	1	0.000	< 0.001
H	-12.432	1	0.000	< 0.001
M	-12.541	1	0.000	< 0.001
F	-13.040	1	0.000	< 0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	23	-2.904		0.34		0.052	

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-684.439	1414.879	1415.8	0.45	0.374

Run 19 (above) with CATEGORIA.DO.VERBO + PAPEL.SEMANTICO + LOCAL is worse than Run 15 also including TRATAMENTO, p = 0.000272

Trying without CATEGORIA.DO.VERBO...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
v	0.237	411	0.401			0.559
n	-0.237	811	0.308			0.441

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
m	1.092	771	0.450			0.749
r	-0.457	443	0.151			0.388
b	-0.635	8	0.125			0.346

\$LOCAL

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
2	16.438	1	1.000			> 0.999
D	15.964	1	1.000			> 0.999
W	3.370	4	0.750			0.967
Y	2.390	13	0.538			0.916
A	2.043	8	0.500			0.885
R	1.899	25	0.440			0.87
B	1.738	229	0.380			0.85
O	1.485	853	0.334			0.815
X	1.398	2	0.500			0.802
C	0.902	29	0.207			0.711
I	0.860	6	0.167			0.703
L	0.559	11	0.182			0.636
E	0.413	22	0.182			0.602
V	0.244	14	0.143			0.561
M	-11.145	1	0.000			< 0.001

5	-12.694	1	0.000	< 0.001
H	-12.694	1	0.000	< 0.001
F	-13.168	1	0.000	< 0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	21	-2.727		0.34		0.061	

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-702.783	1447.566	1448.336	0.405	0.343

Run 20 (above) with TRATAMENTO + PAPEL.SEMANTICO + LOCAL is worse than Run 15 also including CATEGORIA.DO.VERBO, $p = 8.22e-11$

Trying without PAPEL.SEMANTICO...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
v	0.248	411		0.401		0.562
n	-0.248	811		0.308		0.438

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
t	0.944	84		0.560		0.72
m	0.551	597		0.454		0.634
p	-0.016	181		0.331		0.496
v	-1.479	360		0.103		0.186

\$LOCAL

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
D	16.416	1		1.000		> 0.999
2	16.346	1		1.000		> 0.999
W	3.605	4		0.750		0.974
Y	2.632	13		0.538		0.933
A	1.920	8		0.500		0.872
R	1.839	25		0.440		0.863
B	1.668	229		0.380		0.841
0	1.397	853		0.334		0.802
I	1.284	6		0.167		0.783
X	1.087	2		0.500		0.748
C	0.772	29		0.207		0.684
L	0.650	11		0.182		0.657
E	0.466	22		0.182		0.615
V	0.426	14		0.143		0.605
5	-12.220	1		0.000		< 0.001
H	-12.220	1		0.000		< 0.001
M	-12.787	1		0.000		< 0.001
F	-13.282	1		0.000		< 0.001

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	22	-2.082		0.34		0.111	

```
$misc.2
log.likelihood      AIC      AICc    Dxy    R2
      -683.702 1411.403 1412.247 0.459 0.378
```

Run 21 (above) with TRATAMENTO + CATEGORIA.DO.VERBO + LOCAL is worse than Run 15 also including PAPEL.SEMANTICO, $p = 0.00277$

Trying without LOCAL...

```
$TRATAMENTO
factor logodds tokens t/t+0+a+c+i+l+p+x centered factor weight
v      0.234     411                0.401                0.558
n     -0.234     811                0.308                0.442
```

```
$CATEGORIA.DO.VERBO
factor logodds tokens t/t+0+a+c+i+l+p+x centered factor weight
t      0.957      84                0.560                0.722
m      0.404     597                0.454                0.6
p     -0.214     181                0.331                0.447
v     -1.147     360                0.103                0.241
```

```
$PAPEL.SEMANTICO
factor logodds tokens t/t+0+a+c+i+l+p+x centered factor weight
m      0.964     771                0.450                0.724
r      0.435     443                0.151                0.607
b     -1.400      8                0.125                0.198
```

```
$misc.1
n df intercept overall proportion centered input prob
1222 7      -1.402                0.34                0.197
```

```
$misc.2
log.likelihood      AIC      AICc    Dxy    R2
      -691.181 1396.361 1396.453 0.436 0.219
```

Run 22 (above) with TRATAMENTO + CATEGORIA.DO.VERBO + PAPEL.SEMANTICO is worse than Run 15 also including LOCAL, $p = 0.0621$

Dropping LOCAL...

STEP 5 - dropping from Run 22 model with TRATAMENTO + CATEGORIA.DO.VERBO + PAPEL.SEMANTICO

Trying without TRATAMENTO...

```
$CATEGORIA.DO.VERBO
factor logodds tokens t/t+0+a+c+i+l+p+x centered factor weight
t      0.941      84                0.560                0.719
m      0.401     597                0.454                0.599
p     -0.188     181                0.331                0.453
v     -1.155     360                0.103                0.24
```

```
$PAPEL.SEMANTICO
factor logodds tokens t/t+0+a+c+i+l+p+x centered factor weight
```

m	0.960	771	0.450	0.723
r	0.453	443	0.151	0.611
b	-1.413	8	0.125	0.196

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	6	-1.474		0.34			0.186

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-697.09	1406.179	1406.248	0.405	0.206

Run 23 (above) with CATEGORIA.DO.VERBO + PAPEL.SEMANTICO is worse than Run 22 also including TRATAMENTO, $p = 0.000587$

Trying without CATEGORIA.DO.VERBO...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
v	0.219	411		0.401		0.554
n	-0.219	811		0.308		0.446

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
m	1.084	771		0.450		0.747
r	-0.450	443		0.151		0.389
b	-0.634	8		0.125		0.347

\$misc.1

n	df	intercept	overall	proportion	centered	input	prob
1222	4	-1.216		0.34			0.229

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-716.393	1440.786	1440.819	0.359	0.154

Run 24 (above) with TRATAMENTO + PAPEL.SEMANTICO is worse than Run 22 also including CATEGORIA.DO.VERBO, $p = 6.49e-11$

Trying without PAPEL.SEMANTICO...

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
v	0.23	411		0.401		0.557
n	-0.23	811		0.308		0.443

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
t	0.965	84		0.560		0.724
m	0.524	597		0.454		0.628
p	-0.018	181		0.331		0.496
v	-1.471	360		0.103		0.187

\$misc.1

	n	df	intercept	overall	proportion	centered	input	prob
1222	5		-0.634			0.34		0.347

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-697.304	1404.609	1404.658	0.41	0.21

Run 25 (above) with TRATAMENTO + CATEGORIA.DO.VERBO is worse than Run 22 also including PAPEL.SEMANTICO, $p = 0.00219$

All remaining predictors are significant, best model from last step is Run 22

BEST STEP-DOWN MODEL OF RESPONSE FORMA IS WITH PREDICTOR(S):
 CATEGORIA.DO.VERBO ($6.49e-11$) + TRATAMENTO (0.000587) +
 PAPEL.SEMANTICO (0.00219)
 [p-values dropping from full model]

\$TRATAMENTO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
v	0.234	411		0.401		0.558
n	-0.234	811		0.308		0.442

\$CATEGORIA.DO.VERBO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
t	0.957	84		0.560		0.722
m	0.404	597		0.454		0.6
p	-0.214	181		0.331		0.447
v	-1.147	360		0.103		0.241

\$PAPEL.SEMANTICO

factor	logodds	tokens	t/t+0+a+c+i+l+p+x	centered	factor	weight
m	0.964	771		0.450		0.724
r	0.435	443		0.151		0.607
b	-1.400	8		0.125		0.198

\$misc.1

	n	df	intercept	overall	proportion	centered	input	prob
1222	7		-1.402			0.34		0.197

\$misc.2

log.likelihood	AIC	AICc	Dxy	R2
-691.181	1396.361	1396.453	0.436	0.219

BEST STEP-UP MODEL OF RESPONSE FORMA IS WITH PREDICTOR(S):
 CATEGORIA.DO.VERBO ($1.95e-34$) + TRATAMENTO (0.000671) + PAPEL.SEMANTICO
 (0.00219)
 p-values building from null model]

BEST STEP-DOWN MODEL OF RESPONSE FORMA IS WITH PREDICTOR(S):
 CATEGORIA.DO.VERBO ($6.49e-11$) + TRATAMENTO (0.000587) + PAPEL.SEMANTICO
 (0.00219)
 p-values dropping from full model]

STEP-UP AND STEP-DOWN MATCH!