



UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
Campus de São José do Rio Preto

Alessandra Regina Guerra

Diacronia do grau de transparência do sistema de referência
por expressão pronominal e desinencial do argumento-sujeito
de 1^a e 2^a pessoas no português brasileiro

São José do Rio Preto
2017

Alessandra Regina Guerra

Diacronia do grau de transparência do sistema de referência
por expressão pronominal e desinencial do argumento-sujeito
de 1ª e 2ª pessoas no português brasileiro

Tese apresentada como parte dos requisitos para obtenção do título de Doutor em Estudos Linguísticos, junto ao Programa de Pós-Graduação em Estudos Linguísticos do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de São José do Rio Preto.

Financiadora: Capes

Orientador: Prof. Dr. Sebastião Carlos Leite Gonçalves.

São José do Rio Preto
2017

Guerra, Alessandra Regina.

Diacronia do grau de transparência do sistema de referência por expressão pronominal e desinencial do argumento-sujeito de 1ª e 2ª pessoas no português brasileiro / Alessandra Regina Guerra. – São José do Rio Preto, 2017

170 f. : il., tabs.

Orientador: Sebastião Carlos Leite Gonçalves

Tese (doutorado) – Universidade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências, Letras e Ciências Exatas

1. Linguística. 2. Língua portuguesa – Brasil. 3. Economia. 4. Iconicidade (Linguística) 5. Língua portuguesa – Pronomes. 6. Gramática discursivo funcional. I. Universidade Estadual Paulista "Júlio de Mesquita Filho". Instituto de Biociências, Letras e Ciências Exatas. II. Título.

CDU – 469.0-5

Ficha catalográfica elaborada pela Biblioteca do IBILCE
UNESP - Campus de São José do Rio Preto

Alessandra Regina Guerra

Diacronia do grau de transparência do sistema de referência
por expressão pronominal e desinencial do argumento-sujeito
de 1ª e 2ª pessoas no português brasileiro

Tese apresentada como parte dos
requisitos para obtenção do título de
Doutor em Estudos Linguísticos, junto ao
Programa de Pós-Graduação em Estudos
Linguísticos do Instituto de Biociências,
Letras e Ciências Exatas da Universidade
Estadual Paulista “Júlio de Mesquita
Filho”, Campus de São José do Rio Preto.

Financiadora: Capes

COMISSÃO EXAMINADORA

Prof. Dr. Sebastião Carlos Leite Gonçalves – Orientador
Ibilce/Unesp – São José do Rio Preto (SP)

Prof. Dr. Edson Rosa Francisco de Souza
Ibilce/Unesp – São José do Rio Preto (SP)

Profª. Drª. Marize Mattos Dall’Aglio-Hattner
Ibilce/Unesp – São José do Rio Preto (SP)

Prof. Dr. Aquiles Tescari Neto
IEL/Unicamp – Campinas (SP)

Profª. Drª. Rosane de Andrade Berlinck
FCL/Unesp – Araraquara (SP)

São José do Rio Preto
28 de agosto de 2017

para Eduardo, presente desde o início.

AGRADECIMENTOS

a Sebastião Carlos Leite Gonçalves, pela segurança que sua orientação representa, pela receptividade diante de novas ideias, pela leveza de espírito;

a Aquiles Tescari Neto, pela leitura dedicada da tese, pela análise inspiradora, pelo vigor acadêmico;

a Edson Rosa Francisco de Souza, pela oportunidade sempre acessível de diálogo valioso, pelo respeito e pela paciência em qualquer discussão;

a Marize Mattos Dall'Aglio-Hattner, pela sua presença constante e definidora ao longo da minha trajetória acadêmica;

a Rosane Andrade Berlinck, pela disposição em oferecer sua reconhecida experiência para a leitura do trabalho em diferentes fases de seu desenvolvimento.

Watching language move through time, it becomes clear that the more it changes, the more it remains the same (SLOBIN, 1977, p. 212).

RESUMO

O objetivo desta tese de doutorado é analisar a diacronia do grau de transparência do sistema de referência por expressão pronominal e desinencial do argumento-sujeito de 1ª e 2ª pessoas no português brasileiro. Especificamente, analisa-se a variação diacrônica do grau de transparência desse sistema decorrente da interação entre três mudanças ocorridas na história da língua: aumento da frequência de expressão pronominal do argumento-sujeito, aumento da frequência de uso do pronome *você* em detrimento de *tu* e aumento da frequência de emprego da forma pronominal *a gente* em prejuízo de *nós*. O quadro teórico adotado é constituído por uma articulação entre a Gramática Discursivo-Funcional (HENGEVELD; MACKENZIE, 2008) e um conjunto de trabalhos sobre motivações comunicativas, como *iconicidade* e *economia*, determinantes da estrutura das línguas e da dinâmica de processos de mudança diacrônica. O período recortado para estudo estende-se da primeira metade do século XIX ao início do século XXI, e o *corpus* da pesquisa é composto por peças de teatro brasileiras produzidas nesse período. Os dados apresentados e discutidos mostram que o grau de transparência do referido sistema, no recorte temporal selecionado, oscila em torno de um determinado eixo, variando ora em direção a diminuição, ora em direção a aumento de transparência, não se alterando de forma unidirecional. Argumenta-se, então, que esse resultado estaria vinculado, em grande medida, à atuação igualmente relevante das motivações comunicativas de iconicidade e economia.

Palavras-chave: transparência linguística, referência cruzada, expressão pronominal de argumento-sujeito, iconicidade, economia.

ABSTRACT

This dissertation aims at analyzing the diachronic variation in the degree of transparency of the referential system in Brazilian Portuguese, taking into account the reference to first and second person attained in the expression of Subject argument through pronoun and verbal affix. Specifically, the variation in transparency caused by three diachronic changes is addressed: percentage increase in the pronominal expression of Subject argument, percentage increase in the use of the pronoun você (you) to substitute tu (you), percentage increase in the use of the pronoun a gente (we) in order to replace nós (we). The theoretical framework carried out here is drawn from an articulation between Functional Discourse Grammar (Hengeveld & Mackenzie, 2008) and works on communicative motivations such as iconicity and economy, which are considered as forces that determine language structure and diachronic change processes. The historical period analyzed covers the period between the first half of the nineteenth century and the beginning of the twenty-first century, and the corpus of investigation includes Brazilian plays produced during that period. The data show that the degree of transparency of the aforementioned system oscillates around a certain axis, since it sometimes varies towards decrease, sometimes towards increase of transparency, rather than moving in a unidirectional movement. This study argues that this result is associated to a great extent with an equally relevant influence of iconic and economic motivations.

Keywords: *linguistic transparency, cross-reference, pronominal expression of Subject argument, iconicity, economy.*

SUMÁRIO

Lista de Figuras	12
Lista de Gráficos	12
Lista de Quadros	12
Lista de Tabelas	13
INTRODUÇÃO	14
CAPÍTULO I: A Gramática Discursivo-Funcional e a propriedade da transparência linguística	25
1.1. A Gramática Discursivo-Funcional	25
1.1.1. O modelo da Gramática Discursivo-Funcional	25
1.1.2. A representação da referência por expressão do argumento-sujeito	31
1.2. A propriedade da transparência linguística	40
1.2.1. Fatores de transparência/opacidade decorrentes de relações entre níveis	41
1.2.2. Fatores de transparência/opacidade decorrentes de operações dentro dos níveis	47
1.2.3. Quadro final de fatores de transparência/opacidade	50
CAPÍTULO II: Transparência, diacronia e princípios gerais de funcionamento da linguagem	52
2.1. Iconicidade e transparência	53
2.2. Motivações em competição	56
2.3. Teto e piso de transparência	65
CAPÍTULO III: Mudanças diacrônicas no português brasileiro interferentes no grau de transparência linguística	70
3.1. Aumento da frequência de expressão pronominal do argumento-sujeito	70
3.2. Aumento da frequência de uso do pronome <i>você</i> em detrimento de <i>tu</i>	80
3.3. Aumento da frequência de uso da forma pronominal <i>a gente</i> em detrimento de <i>nós</i>	86
CAPÍTULO IV: Metodologia	91
4.1. Metodologia de levantamento e análise de dados	91
4.2. <i>Corpus</i>	108
CAPÍTULO V: Análise de dados	117
5.1. Dados sobre mudanças diacrônicas interferentes na transparência linguística	117
5.2. Identificação da variação diacrônica do grau de transparência do sistema de referência	122
5.3. Não-unidirecionalidade da variação diacrônica do grau de transparência do sistema de referência	135
5.4. Variação diacrônica de transparência e princípios gerais de funcionamento da linguagem	139

CONCLUSÃO	157
REFERÊNCIAS	161
REFERÊNCIAS UTILIZADAS NA CONSTITUIÇÃO DO <i>CORPUS</i>	166

Lista de Figuras

Figura 01: Arquitetura geral da GFD	26
Figura 02: O Nível Interpessoal	29
Figura 03: O Nível Representacional	29
Figura 04: O Nível Morfossintático	30
Figura 05: O Nível Fonológico	31
Figura 06: Subtipos de iconicidade, segundo trabalhos de Haiman (1980, 1983, 1985) e Croft (2003)	53

Lista de Gráficos

Gráfico 01: Índices de sujeito pleno no português (KAISER, 2006, p. 19)	73
Gráfico 02: Evolução no uso de sujeito nulo, segundo dados de Duarte (1995)	74
Gráfico 03: Evolução no uso de sujeito nulo por pessoa do discurso, segundo dados de Duarte (1995)	74
Gráfico 04: Uso de sujeito nulo em diferentes gerações de falantes, segundo dados de Duarte (1995)	75
Gráfico 05: Evolução do uso de sujeito pleno no PB, segundo dados de Tarallo (1993a)	76
Gráfico 06: Evolução dos usos de <i>tu</i> e <i>você</i> , segundo dados de Souza (2012)	83
Gráfico 07: Evolução dos usos de <i>tu</i> e <i>você</i> , segundo dados de Machado (2006)	84
Gráfico 08: Percentuais de <i>nós</i> e <i>a gente</i> em diferentes capitais brasileiras, segundo dados de Vianna e Lopes (2012)	90
Gráfico 09: Variação diacrônica na frequência de expressão pronominal do argumento-sujeito de 1ª e 2ª pessoas do singular e do plural	118
Gráfico 10: Variação diacrônica no uso de <i>tu</i> e <i>você</i>	120
Gráfico 11: Variação diacrônica no uso de <i>nós</i> e <i>a gente</i>	121
Gráfico 12: Variação diacrônica do grau de transparência do sistema de referência	133
Gráfico 13: Grau de transparência do sistema de referência, considerando apenas o aumento da expressão pronominal	142
Gráfico 14: Grau de transparência do sistema de referência, considerando apenas a mudança no quadro pronominal	143
Gráfico 15: Variação diacrônica do grau de transparência do sistema de referência (média e por pessoa do discurso)	146

Lista de Quadros

Quadro 01: Parâmetros para cálculo da transparência linguística	50
Quadro 02: Possíveis correspondências forma-significado e suas motivações	60
Quadro 03: Evolução nos paradigmas flexionais no português (DUARTE, 1993, p. 109)	78
Quadro 04: Tipos de referência em termos de transparência e opacidade	92

Quadro 05: Combinações constituintes do sistema de referência por expressão pronominal e desinencial do argumento-sujeito de 1ª e 2ª pessoas no PB	93
Quadro 06: Desinências número-pessoais de verbos regulares do PB	100
Quadro 07: Peças teatrais integrantes do <i>corpus</i> da pesquisa	109
Quadro 08: Peças teatrais utilizadas no levantamento de dados geral e quantidades de dados coletados por peça	114

Lista de Tabelas

Tabela 01: Evolução dos usos de <i>tu</i> e <i>você</i> , segundo dados de Souza (2012)	82
Tabela 02: Evolução dos usos de <i>tu</i> e <i>você</i> , segundo dados de Machado (2006)	84
Tabela 03: Dados a serem obtidos sobre a variação diacrônica do grau de transparência do sistema de referência	102
Tabela 04: Dados a serem obtidos para cálculo do grau de transparência do sistema de referência em cada sincronia	103
Tabela 05: Modelo para cálculo do grau de transparência do sistema de referência no domínio da 1PS	104
Tabela 06: Variação diacrônica nas frequências de expressão e não-expressão pronominal do argumento-sujeito de 1ª e 2ª pessoas do singular e do plural	118
Tabela 07: Variação diacrônica nas frequências de uso de <i>tu</i> e <i>você</i>	119
Tabela 08: Variação diacrônica nas frequências de <i>nós</i> e <i>a gente</i>	121
Tabela 09: Cálculo do grau de transparência do sistema de referência relativamente à 1PS na sincronia 19-1	125
Tabela 10: Cálculo do grau de transparência do sistema de referência relativamente a <i>você</i> na sincronia 19-1	128
Tabela 11: Cálculo do grau de transparência do sistema de referência relativamente a <i>a gente</i> na sincronia 19-1	130
Tabela 12: Variação diacrônica do grau de transparência do sistema de referência por expressão pronominal e desinencial do argumento-sujeito de 1ª e 2ª pessoas no PB..	132

INTRODUÇÃO

Esta tese de doutorado trata da propriedade linguística da transparência no português brasileiro (PB), sob perspectiva diacrônica. O quadro teórico do trabalho é constituído por uma articulação entre a Gramática Discursivo-Funcional (GDF; HENGEVELD; MACKENZIE, 2008) e um conjunto de trabalhos de orientação funcionalista que lidam com motivações comunicativas/cognitivas da estrutura das línguas e com a influência dessas motivações sobre fenômenos diacrônicos (SLOBIN, 1977; HAIMAN, 1985; GIVÓN, 1985; CROFT, 2003).

Em termos gerais, pode-se dizer que a transparência é a propriedade que uma língua apresenta de manifestar, em maior ou menor grau, uma correspondência de um-para-um entre unidades da estrutura formal e unidades de significado dos enunciados (isto é, uma unidade formal codifica um, e apenas um, significado). Quando tende a manifestar maior incidência de relações de um-para-um entre unidades formais e unidades de significado, a língua é mais transparente (ou menos opaca). Por outro lado, se manifesta relações de um-para-um com menor incidência, a língua é menos transparente (ou mais opaca). Desse modo, toda língua apresenta um determinado grau de transparência.¹

Atualmente há uma série de trabalhos sobre transparência linguística no âmbito da GDF, os quais recobrem diferentes aspectos do fenômeno (HENGEVELD, 2011a, b; LEUFKENS, 2013, 2015; STASSI-SÉ, 2014; OLBERTZ; SOUZA, 2015; SOUZA; COMPARINI; GUIRALDELLI, no prelo; HENGEVELD; LEUFKENS, no prelo, entre outros). No entanto, uma questão ainda não estudada sobre transparência – e que se mostra relevante, tanto para a teoria sobre transparência quanto para o entendimento do funcionamento das línguas do mundo – diz respeito ao comportamento diacrônico da

¹ Construções linguísticas opacas, em termos gerais, compreendem combinações de duas unidades formais para uma unidade de significado, duas unidades de significado para uma unidade formal, uma unidade de significado sem expressão formal ou uma unidade formal sem significação.

propriedade da transparência. Trata-se de pensar o que acontece com o grau de transparência das línguas ao longo de seus desenvolvimentos diacrônicos.

Um tema como esse envolve questões como as seguintes: ao longo de sua história, as línguas apresentam mudanças em seu grau de transparência ou tendem a se manter num mesmo grau? No caso de haver mudança no grau de transparência, essa mudança apresenta uma direção única (*transparência < opacidade* ou *opacidade < transparência*) ou o processo de mudança não chega a manifestar regularidade em termos de direção (caminhando ora em uma direção ora em outra)? Em síntese, essas questões abordam a ocorrência ou não de mudança histórica no grau de transparência das línguas e a dinâmica de funcionamento desse processo.

Naturalmente, um estudo mais completo dessas questões envolveria a análise de um amplo conjunto de línguas, de diferentes tipos, para que estas pudessem ser comparadas entre si e para que se pudesse, então, fazer generalizações a respeito. Neste trabalho, dados os limites de uma pesquisa de doutorado, restringimos a análise a uma língua particular – o PB –, como forma de oferecer dados específicos de uma língua para que se possa pensar as questões gerais acima esboçadas.

Para calcular o grau de transparência de uma língua, a GDF estabelece um conjunto de fenômenos linguísticos que violam a relação de um-para-um entre unidades linguísticas e, então, afere o grau de transparência da língua verificando em que medida a língua apresenta ou não esses fenômenos. Os fenômenos estabelecidos pela GDF para esse cálculo são considerados como parâmetros de transparência/opacidade. Quanto maior for o número de fenômenos desse conjunto que a língua apresenta, menos transparente a língua é. Inversamente, quanto menor for o número de fenômenos apresentados, maior o grau de transparência da língua.²

² Na seção 1.2, apresentamos a lista completa de parâmetros de transparência/opacidade de acordo com Hengeveld e Leufkens (no prelo), exemplificando os diferentes tipos de violação de transparência.

Um dos parâmetros de transparência/opacidade estabelecidos pela GDF é a presença, na língua, do fenômeno denominado *Referência Cruzada* (RC, daqui em diante). De acordo com Hengeveld e Mackenzie (2008) e Hengeveld e Leufkens (no prelo), uma língua apresenta RC quando um marcador de pessoa no verbo é suficiente em si próprio como meio de fazer referência a uma dada entidade, podendo opcionalmente ser expandido por um sintagma nominal que faz referência a essa mesma entidade. Para a GDF, a RC é um fenômeno opaco porque, nesse caso, ocorrem, numa mesma oração, dois atos de referência a uma mesma entidade: um ato realizado pela desinência verbal, outro pelo sintagma nominal.

O PB apresenta o fenômeno da RC. Nessa língua, o verbo da oração recebe marcação de pessoa (mais especificamente, número-pessoa), a qual é suficiente em si própria como forma de fazer referência a uma entidade, no caso, à entidade correspondente ao argumento-sujeito (como se depreende da gramaticalidade de uma oração como *cantamos bem*); essa marcação de pessoa pode ser, opcionalmente, expandida pela expressão nominal ou pronominal do argumento-sujeito (como, por exemplo, em *nós cantamos bem*).³

No presente trabalho, selecionamos o parâmetro da RC para análise, como forma de delimitação de nosso objeto de estudo e porque este parâmetro, diferentemente dos demais, é nitidamente afetado por mudanças diacrônicas ocorridas no PB, conforme explicamos logo adiante, o que, então, também tem implicações sobre a diacronia do grau de transparência da língua.

A GDF, por conta de interesses mais gerais de pesquisa tipológica, ao considerar

³ Na GDF, considera-se uma diferença entre RC e Concordância Verbal. Um exemplo de Concordância verifica-se no francês, em que o verbo da oração contém desinência número-pessoal correspondente ao argumento-sujeito, mas essa desinência sozinha não é capaz de fazer referência, sendo *obrigatória* a expressão desse argumento por um sintagma nominal. No caso da RC, a desinência verbal número-pessoal do verbo, por si só, é capaz de fazer referência, sendo *opcional* a expressão do argumento-sujeito por sintagma nominal. Assim, esse fenômeno da RC é, de fato, o que se verifica no PB. Como se pode notar, a noção de Concordância Verbal da GDF é mais específica do que a noção mais tradicional. O que tradicionalmente é considerado como Concordância parece ser dividido, na GDF, entre Concordância Verbal (o caso, por exemplo, do francês) e RC (caso, por exemplo, do PB). Dessa forma, a título de esclarecimento, pode-se considerar que o que estamos estudando neste trabalho como RC (nos termos da GDF) corresponde a uma parte do que é tradicionalmente tratado como Concordância Verbal.

o parâmetro da RC na identificação do grau de transparência de uma língua, avalia apenas a manifestação ou não do fenômeno na língua. A existência do fenômeno é considerada como fator de opacidade, enquanto a inexistência é vista como um fator de transparência. No entanto, no presente trabalho, em que analisamos uma língua particular, consideramos que é possível e pertinente tratar a RC como parte de um sistema mais amplo de referência e verificar a transparência/opacidade desse sistema como um todo.

Consideramos que, no PB, a RC pode e deve ser entendida como uma possível forma de se fazer referência à entidade correspondente ao argumento-sujeito dentre um conjunto maior de possibilidades de se fazer referência a essa entidade, incluindo alternativas opacas e transparentes. A RC, como mencionado, constitui uma dupla referência a essa entidade (o que ocorre, por exemplo, em *tu canta-s*). Além dessa possibilidade, no PB pode-se fazer referência uma única vez, caso que chamaremos de *Referência Simples*, ou mesmo nenhuma vez, o que chamaremos de *Referência Zero*. A Referência Simples (RS) ocorre em duas situações: quando o argumento-sujeito é expresso apenas pela desinência número-pessoal do verbo (*Ø canta-s*); quando esse argumento é expresso apenas por sintagma nominal, e a desinência número-pessoal do verbo é representada por morfema zero (*você canta-Ø*). Já a Referência Zero (RZ) verifica-se quando não há expressão do argumento-sujeito nem por meio de desinência número-pessoal explícita (desinência correspondente a morfema zero), nem por meio de sintagma nominal (*Ø canta-Ø*, no sentido de *você canta*).

Desse modo, consideramos neste trabalho que o PB (assim como outras línguas) apresenta o que se pode chamar de *sistema de referência por expressão do argumento-sujeito*, que consiste no conjunto das combinações entre expressão/não-expressão (pro)nominal e expressão/não-expressão desinencial do argumento-sujeito, combinações que se distribuem, então, entre RC, RZ e duas possibilidades de RS.

Como se pode notar, as formas de referência aqui em pauta constituem

possibilidades de se fazer referência no que diz respeito ao processo de expressão do argumento-sujeito no PB. Por essa razão, estamos nos reportando a nosso objeto de estudo em termos de referência *por expressão do argumento-sujeito*. O ato de se fazer referência no PB, assim como em outras línguas, pode ser realizado por outras unidades linguísticas; por exemplo, num enunciado como *eu entreguei o livro para os alunos na biblioteca*, não só o pronome *eu* e a desinência *-i*, que expressam o argumento-sujeito, realizam atos de referência, mas também os sintagmas *o livro*, *os alunos* e *a biblioteca*, que expressam outros elementos, que não o argumento-sujeito. Diante disso, nosso trabalho está delimitado a um tipo de referência apenas, aquela realizada pelas unidades linguísticas que expressam o argumento-sujeito.

Além disso, por razões que explicamos no decorrer do trabalho, delimitamos nossa análise à expressão pronominal (excluindo a expressão nominal) e desinencial do argumento-sujeito, restringindo nosso estudo ainda a argumentos-sujeitos de 1ª pessoa do singular (*eu*) e do plural (*nós* e *a gente*) e de 2ª pessoa do singular (*tu* e *você*) e do plural (*vós* e *vocês*).⁴ Assim, nosso estudo trata, mais precisamente, do que estamos chamando de *sistema de referência por expressão pronominal e desinencial do argumento-sujeito de 1ª e 2ª pessoas*.

Especificamente, esse sistema compreende o conjunto de combinações entre (i) expressão/não-expressão de cada uma das formas pronominais *eu*, *tu*, *você*, *nós*, *a gente*, *vós* e *vocês* e (ii) a forma verbal correspondente a cada pronome em cada um dos nove tempos/modos verbais do PB (seis do indicativo e três do subjuntivo), tendo em vista que, dentre as formas verbais de cada tempo/moço em cada pessoa, algumas têm desinência número-pessoal explícita, outras não.

Esse sistema de referência inclui, por exemplo, combinações como as seguintes:

⁴ As explicações dessas delimitações são feitas nas seções 1.1.2 e 4.1.

- a) expressão do pronome *eu* + pres. ind. (***eu canto***) = RC
 - b) não-expressão do pronome *eu* + pres. ind. (**$\emptyset$ canto**) = RS
 - c) expressão do pronome *eu* + pret. imp. ind. (***eu cantava- $\emptyset$***) = RS
 - d) não-expressão do pronome *eu* + pret. imp. ind. (**$\emptyset$ cantava- $\emptyset$**) = RZ
 - e) expressão do pronome *tu* + pres. ind. (***tu cantas***) = RC
 - f) não-expressão do pronome *tu* + pres. ind. (**$\emptyset$ cantas**) = RS
- etc.

O conjunto completo de combinações (que é apresentado em nosso capítulo de metodologia) compreende um total de 126 possibilidades: cada uma das sete formas pronominais expressas combinada com nove tempos/modos verbais, o que resulta em 63 combinações ($7 \times 9 = 63$), somadas às versões sem expressão pronominal correspondentes a cada uma dessas 63 combinações, resultando as 126 combinações possíveis. Esse conjunto completo de combinações constitui, então, o que estamos chamando neste trabalho de *sistema de referência por expressão pronominal e desinencial do argumento-sujeito de 1ª e 2ª pessoas*.

Analizamos aqui o *grau de transparência* desse sistema, o qual é determinado pela parcela de combinações transparentes *versus* a parcela de combinações opacas dentro do sistema, assim como pelos pesos que a gramática da língua atribui às construções transparentes *versus* às opacas.

No que diz respeito às parcelas de combinações transparentes e opacas, o sistema apresenta igualdade de distribuição. Em qualquer combinação entre pronome e tempo/modo, se a expressão pronominal do argumento-sujeito acarreta combinação opaca (por exemplo, ***eu canto***), a versão correspondente sem expressão pronominal será transparente (**$\emptyset$ canto**); inversamente, se a expressão pronominal acarreta transparência (***eu cantava- $\emptyset$***), a versão correspondente sem expressão pronominal resultará em opacidade (**$\emptyset$ cantava- $\emptyset$**). Dessa forma, como o sistema prevê uma metade de combinações com expressão pronominal e outra

metade com versões correspondentes sem expressão pronominal, há mesmo número de combinações transparentes e opacas, o que significaria dizer que esse sistema de referência estaria estruturado de modo equilibrado.

No entanto, a gramática do PB, em cada sincronia, atribui pesos diferentes a cada uma das combinações do sistema, isto é, em cada sincronia, a gramática está configurada de modo a propiciar com mais frequência o uso de certas combinações em detrimento de outras. Aqui, entendemos que esse favorecimento deve-se a três características da gramática em cada sincronia: (i) frequências de expressão e não-expressão pronominal do argumento-sujeito; (ii) frequências de uso dos pronomes variantes *tu* e *você*; (iii) frequências de uso das formas pronominais variantes *nós* e *a gente*.

No que tange à primeira dessas características, observe-se que, por exemplo, o sistema de referência inclui a combinação entre expressão do pronome *eu* e verbo no presente do indicativo (***eu canto***) e a combinação entre não-expressão desse pronome também com o presente do indicativo (***Ø canto***), que constituem combinações, respectivamente, opaca (RC) e transparente (RS). No entanto, dependendo das frequências de expressão e não-expressão pronominal do argumento-sujeito, o sistema gramatical favorecerá o uso de uma ou outra dessas duas combinações – ***eu canto*** (RC) ou ***Ø canto*** (RS). Se, numa dada sincronia, a frequência de não-expressão pronominal é maior do que a de expressão pronominal, isso significa que, nessa sincronia, a gramática também propiciará, com maior frequência, a combinação ***Ø canto*** (transparente), em detrimento da combinação ***eu canto*** (opaca).

No mesmo sentido, as frequências de uso dos pronomes *tu* e *você* interferem no favorecimento do uso de combinações transparentes ou opacas. Por exemplo, o sistema de referência contempla a combinação entre expressão do pronome *tu* e verbo no presente do indicativo (***tu cantas***) e a combinação entre expressão do pronome *você* também com o presente do indicativo (***você canta-Ø***), que constituem combinações, respectivamente, opaca

(RC) e transparente (RS). Porém, a depender das frequências de uso de *tu* e de *você*, essas combinações ocorrerão com frequências diferentes. Em uma sincronia em que o uso de *tu* seja mais frequente do que o de *você*, tal preferência acarretará também uma maior frequência da combinação ***tu cantas*** (opaca) em detrimento de ***você canta-Ø*** (transparente). O mesmo tipo de favorecimento de combinações transparentes ou opacas verifica-se em relação às frequências de uso das formas pronominais *nós* e *a gente*.

Portanto, neste trabalho, analisamos o *grau de transparência* do sistema de referência calculando em que medida a gramática do PB promove a escolha de combinações transparentes (RS) *versus* combinações opacas (RC e RZ), o que será determinado pelas frequências de expressão e não-expressão pronominal do argumento-sujeito, assim como pelas frequências de uso dos pronomes *tu* e *você* e de *nós* e *a gente*. Nesse sentido, como se pode ver, nosso estudo não envolve apenas constatar a existência ou inexistência do fenômeno da RC na língua, como é o procedimento mais comum em trabalhos sobre transparência na GDF, mas envolve considerar o *grau de transparência do sistema de referência* da língua, do qual a RC é parte.

Como acabamos de explicar, numa dada sincronia o grau de transparência do sistema de referência é determinado pelas frequências de expressão/não-expressão pronominal do argumento-sujeito e pelas frequências de uso de *tu/você* e de *nós/a gente*. Desse modo, pode-se perceber que mudanças diacrônicas pelas quais o PB vem passando quanto a esses três aspectos acarretam também *variação diacrônica* no grau de transparência do sistema de referência. Assim, as seguintes mudanças tornam-se relevantes: (i) aumento na frequência de expressão pronominal do argumento-sujeito em oposição à não-expressão pronominal (fenômeno normalmente tratado em outras teorias como aumento na frequência de sujeito pronominal pleno em oposição a sujeito nulo);⁵ (ii) aumento na frequência de uso do pronome

⁵ Conforme explicaremos na seção 4.1, a mudança diacrônica no PB que afeta o grau de transparência do sistema de referência por expressão do argumento-sujeito é a mudança no aumento da frequência de expressão

você em detrimento de *tu*; (iii) aumento na frequência de uso da forma pronominal *a gente* em detrimento de *nós*.⁶

Diacronicamente, o aumento na frequência de expressão pronominal do argumento-sujeito leva, por um lado, a diminuição no grau de transparência do sistema de referência (no caso da combinação da expressão pronominal com formas verbais que têm desinência número-pessoal explícita) e, por outro lado, a aumento no grau de transparência do sistema de referência (no caso da combinação da expressão pronominal com formas verbais que não possuem desinência número-pessoal explícita – morfema zero).

Similarmente, o aumento na frequência de uso de *você* em detrimento de *tu* também afeta o grau de transparência do sistema de referência. Todas as formas verbais de 2ª pessoa do singular relativas ao pronome *tu*, em todos os tempos/modos, contêm desinência número-pessoal explícita. Assim, quando o pronome *tu* é expresso, a referência é sempre opaca e, quando esse pronome não é expresso, a referência é sempre transparente. Por outro lado, as formas verbais relativas ao pronome *você*, na maioria dos tempos/modos, não contêm desinência número-pessoal explícita. Dessa maneira, quando o pronome *você* é expresso, a referência é, na maioria das vezes, transparente, e, quando tal pronome não é expresso, a referência é, na maioria dos casos, opaca.

No caso do aumento diacrônico do uso de *a gente* em detrimento de *nós*, o grau de transparência do sistema de referência sofre uma variação diacrônica similar ao que se verifica no caso do aumento do uso de *você* em detrimento de *tu*.⁷

pronominal do argumento-sujeito. Por isso, nosso trabalho, no que diz respeito à expressão desse argumento por um sintagma nominal, restringe-se à análise apenas da expressão *pronominal* desse argumento, não incluindo os casos de expressão *nominal*.

⁶ Neste trabalho, não utilizamos o termo *mudança* no sentido técnico de mudança implementada. Com o uso do termo, pretendemos nos referir às alterações de frequência de ocorrência de determinados fenômenos, ou seja, às alterações nas frequências de expressão/não-expressão pronominal do argumento-sujeito e de uso das formas pronominais *tu*, *você*, *nós* e *a gente*.

⁷ Não incluímos em nosso estudo a análise do fenômeno da simplificação da morfologia verbal flexional que leva a construções como *tu canta*, *nós canta* e *vocês canta*. Da mesma forma, também não incluímos em nossa

Diante do exposto, o objetivo deste trabalho é mostrar como a atuação conjunta dessas três mudanças interfere diacronicamente no grau de transparência do sistema de referência por expressão pronominal e desinencial do argumento-sujeito de 1ª e 2ª pessoas no PB. A esse respeito, tentaremos demonstrar que essas três mudanças, de fato, acarretam variações diacrônicas consideravelmente significativas no grau de transparência do sistema de referência, mas que essas variações ocorrem tanto em direção a maior opacidade quanto em direção a maior transparência, oscilando em torno de um eixo de equilíbrio entre opacidade e transparência. Ademais, tentaremos mostrar que essa dinâmica diacrônica pode ser entendida como decorrente da competição entre diferentes motivações cognitivas/comunicativas que atuam sobre a língua, como a iconicidade e a economia.

Para desenvolver a discussão aqui proposta, este trabalho compreende, após esta seção de introdução, cinco capítulos e a seção de conclusão. Nos capítulos 1, 2 e 3, discutimos e sistematizamos os fundamentos teóricos do trabalho. No capítulo 1, apresentamos o modelo da GDF e a noção de transparência linguística, focalizando os pontos mais importantes para nossa pesquisa. No capítulo 2, discutimos os conceitos funcionalistas mais relevantes para estudos diacrônicos, sobretudo as motivações funcionais que atuam sobre a dinâmica diacrônica das línguas, tais como as noções de iconicidade e economia. No capítulo 3, fazemos uma incursão por trabalhos dedicados a questões de mudanças diacrônicas no PB que retratam os fenômenos aqui em pauta.

No capítulo 4, explicamos os procedimentos metodológicos, o que inclui uma delimitação mais específica de nosso objeto de estudo, e apresentamos o *corpus* de nossa pesquisa. No capítulo 5, apresentamos nossa análise e discussão de dados. Nesse capítulo, primeiramente expomos os resultados de nosso levantamento de dados sobre as três mudanças diacrônicas estudadas e, em seguida, identificamos a variação diacrônica do grau de

pesquisa o uso da forma pronominal *a gente* combinada com a desinência verbal de primeira pessoa do plural (*a gente cantamos*). No capítulo de metodologia justificamos essas exclusões.

transparência do sistema de referência aqui em questão, no período analisado; ainda nesse capítulo, desenvolvemos uma análise qualitativa sobre a variação diacrônica identificada, relacionando-a a princípios gerais de funcionamento da linguagem, principalmente às motivações de iconicidade e economia. Finalmente, na última seção da tese, apresentamos as conclusões.

CAPÍTULO I

A Gramática Discursivo-Funcional e a propriedade da transparência linguística

Este capítulo encontra-se dividido em duas partes. Na seção 1.1, apresentamos uma visão geral da GDF, uma vez que, além de esse modelo constituir um dos principais fundamentos teóricos deste trabalho, a concepção de transparência aqui adotada é desenvolvida no âmbito desse modelo. Em 1.2, sintetizamos, então, essa noção de transparência linguística.

1.1 A Gramática Discursivo-Funcional

1.1.1 O modelo da Gramática Discursivo-Funcional

A GDF (HENGEVELD; MACKENZIE, 2008) é uma teoria funcionalista, tipologicamente baseada, da estrutura gramatical das línguas. É o Componente Gramatical de um modelo mais amplo de interação verbal que envolve, além do Componente Gramatical, linguístico, três componentes não-linguísticos: o Componente Conceitual, o Componente Contextual e o Componente de Expressão. Esses três componentes interagem com o Componente Gramatical para a produção de enunciados nas línguas.

O Componente Conceitual é o *leitmotiv* do Componente Gramatical como um todo; é o responsável pela elaboração de uma intenção comunicativa relevante para o discurso e pelas conceitualizações associadas aos eventos extralinguísticos relevantes. O Componente Contextual interage com o Componente Gramatical fornecendo dois tipos de informação: (i) informação imediata recebida do Componente Gramatical sobre um enunciado particular e determinante da forma que o discurso subsequente assume; (ii) informação de longo termo advinda da interação em curso, a qual, além de determinar as distinções requeridas na língua

em questão, influenciam a formulação e a codificação naquela língua. O Componente de Expressão gera expressões acústicas, ortográficas ou de sinais com base na informação fornecida pelo Componente Gramatical.

A arquitetura geral da GDF é apresentada na figura 1 abaixo.

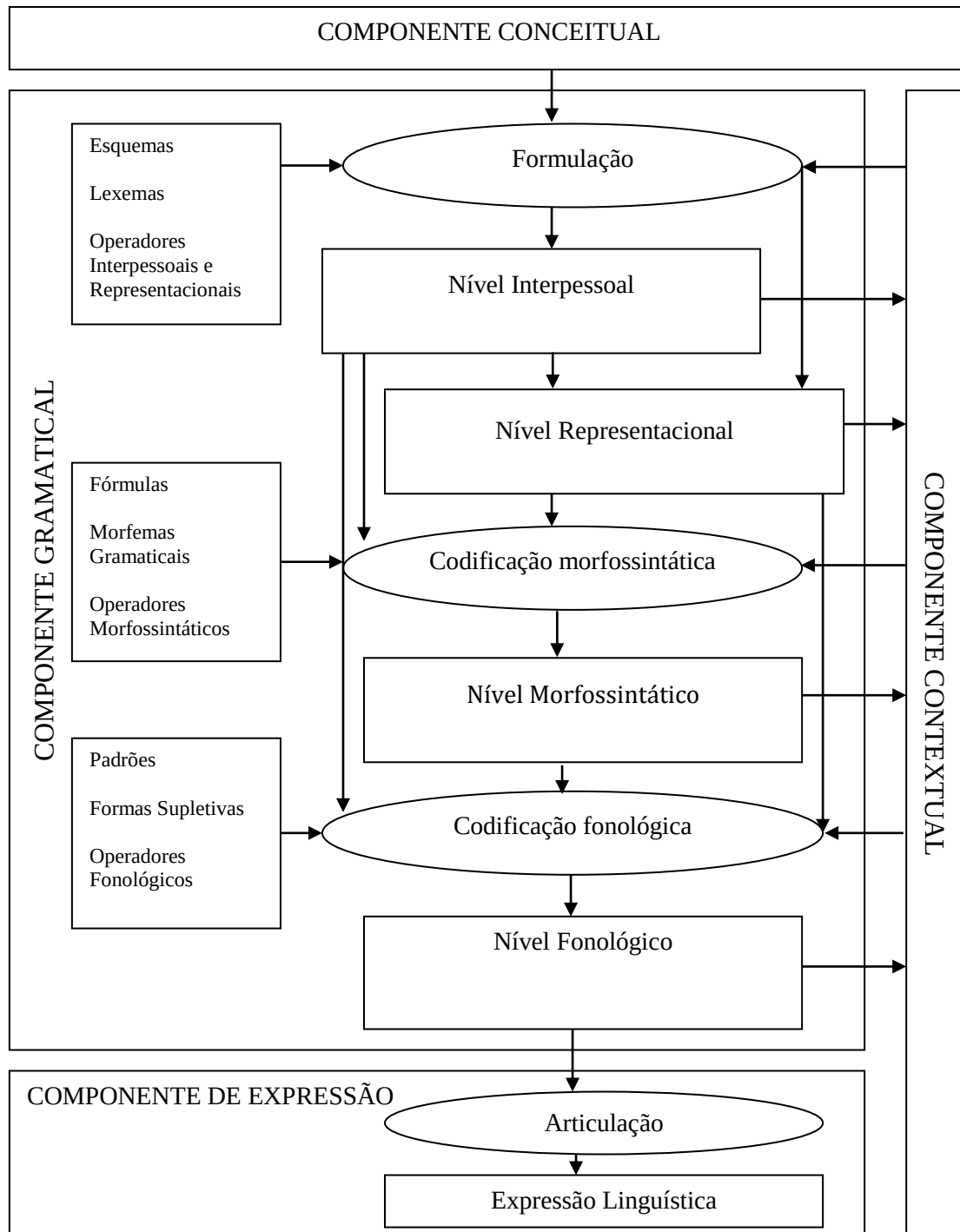


Figura 1: Arquitetura geral da GDF (HENGEVELD; MACKENZIE, 2008, p. 13).

Como se pode ver na figura, o Componente Gramatical contém: (i) três operações (representadas nas imagens ovais) – Formulação, Codificação Morfossintática e Codificação Fonológica; (ii) três conjuntos de primitivos (representados pelos quadrados) usados nas operações; (iii) quatro níveis de representação (indicados pelos retângulos), que contêm as representações construídas pelas operações, isto é, os níveis Interpessoal, Representacional, Morfossintático e Fonológico.

A GDF é um modelo que opera de cima para baixo (modelo *top-down*): seu funcionamento parte da intenção comunicativa do falante, elaborada no Componente Conceitual, e segue em direção à articulação, que ocorre dentro do Componente de Expressão. Essa proposta de organização da língua é uma tentativa de refletir o processo real de produção da linguagem apontado por estudos psicolinguísticos (LEVELT, 1989 *apud* HENGVELD; MACKENZIE, 2008), tornando essa gramática um modelo mais efetivo de representação daquilo que, de fato, ocorre na produção dos enunciados pelos falantes.

Mais detalhadamente, o primeiro passo do funcionamento do modelo ocorre quando, no Componente Conceitual, uma intenção comunicativa e sua representação mental correspondente se tornam relevantes. A operação de Formulação, utilizando elementos do primeiro conjunto de primitivos, converte essas representações conceituais em representações pragmáticas e semânticas, que são inseridas nos níveis Interpessoal e Representacional, respectivamente. As operações de Codificação, lançando mão de elementos dos outros dois conjuntos de primitivos, convertem essas configurações interpessoais e representacionais em representações morfossintáticas e fonológicas e as inserem nos níveis Morfossintático e Fonológico, respectivamente. O Nível Fonológico é o *input* para a operação de Articulação, que contém as regras fonéticas necessárias para a formação de um enunciado adequado. A articulação ocorre no Componente de Expressão, fora do Componente Gramatical.

Os vários níveis de representação contidos na gramática alimentam o Componente

Contextual, permitindo referência subsequente aos vários tipos de entidades relevantes em cada um desses níveis, uma vez introduzidas no discurso. O Componente Contextual, por sua vez, alimenta as operações de Formulação e de Codificação, e assim a disponibilidade de antecedentes pode influenciar a composição de enunciados subsequentes.

Conforme mencionado, a GDF distingue quatro níveis de análise: dois níveis de Formulação, os níveis Interpessoal e Representacional, e dois níveis de Codificação, os níveis Morfossintático e Fonológico. No Nível Interpessoal são descritos todos os aspectos pragmáticos de uma unidade linguística que são codificados na estrutura formal dessa unidade. No Nível Representacional, similarmente, são descritos todos os aspectos semânticos codificados em uma unidade linguística. O Nível Morfossintático é reservado para a descrição da estrutura morfosintática das unidades. E, no Nível Fonológico, é feita a representação da estrutura fonológica das unidades linguísticas.

Cada um desses níveis é dividido internamente em camadas hierarquicamente organizadas. No Nível Interpessoal (NI), a camada mais alta é a do *Move* (M), a qual, definida como “uma contribuição autônoma para uma interação em curso” (HENGEVELD; MACKENZIE, 2008, p. 50), provoca uma reação linguística ou é, ela mesma, uma reação. Um *Move* é composto de um (ou mais de um) Ato Discursivo (A). O Ato Discursivo é “a menor unidade identificável de comportamento comunicativo” (KROON, 1995, *apud* HENGEVELD; MACKENZIE, 2008, p. 60). É composto por uma *Ilocução* (F) e por um participante, o *Falante* (P1s), podendo conter ainda um segundo participante, o *Ouvinte* (P2A), e um (ou mais de um) *Conteúdo Comunicado* (C). O *Conteúdo Comunicado* contém a totalidade do que o Falante deseja evocar em sua comunicação com o Ouvinte. Essa evocação ocorre por meio de *Subatos* que, por sua vez, se dividem em *Subatos de Referência* (R), que evocam um referente, e *Subatos de Atribuição* (T), que evocam uma propriedade. A estrutura

completa do NI é representada na Figura 2.⁸

$$(M_1: [(A_1: [(F_1) (P_1)_S (P_2)_A (C_1: [(T_1)_{\{\Phi\}} \dots (T_{1+N})_{\{\Phi\}} (R_1)_{\{\Phi\}} \dots (R_{1+N})_{\{\Phi\}}](C_1)_{\{\Phi\}})] (A_1) \dots (A_{1+N})_{\{\Phi\}}] (M_1))$$

Figura 2: O Nível Interpessoal (HENGEVELD; MACKENZIE, 2008, p.15).

No Nível Representacional (NR), a camada mais alta é a do Conteúdo Proposicional (p), unidade que representa construtos mentais (conhecimentos, crenças ou desejos). Conteúdos Proposicionais contêm Episódios (ep), que são conjuntos de Estados de Coisas (e) tematicamente coerentes, no sentido de que apresentam unidade ou continuidade de tempo, de lugar e de indivíduos. Estados de Coisas incluem eventos e estados e são caracterizados pelo fato de que podem ser localizados no tempo e podem ser avaliados em termos de sua realidade. Um Estado de Coisas é constituído por uma combinação de unidades semânticas que não estão em relação hierárquica entre si. Essas unidades semânticas podem ser de vários tipos: Propriedade Lexical (f), Indivíduo (x), Lugar (l), Tempo (t), Maneira (m), Razão (r) e Quantidade (q)⁹. A estrutura completa do NR é a apresentada na figura 3.

$$(p_1: [(ep_1: [(e_1: [(f_1: [(f_2)^n (x_1)_{\Phi} \dots (x_{1+n})_{\Phi}] (f_1)) \dots (f_{1+n}) (e_1)_{\Phi}] \dots (e_{1+n})_{\{\Phi\}}] (ep_1)) \dots (ep_{1+n})_{\{\Phi\}}] (p_1))$$

Figura 3: O Nível Representacional (HENGEVELD; MACKENZIE, 2008, p.15).

O Nível Morfossintático (NM) lida com os aspectos morfossintáticos das unidades linguísticas, sem fazer distinção entre aspectos morfológicos ou sintáticos, pois a GDF considera que os princípios usados na formação de palavras são os mesmos dos usados na

⁸ No decorrer deste trabalho, as abreviações referentes a conceitos da GDF são mantidas em inglês, em adequação às indicações originais; fazemos exceção apenas às abreviações relativas aos rótulos *Nível Interpessoal*, *Nível Representacional*, *Nível Morfossintático* e *Nível Fonológico*, os quais, conforme indicado também no corpo do texto, serão referidos, respectivamente, como *NI*, *NR*, *NM* e *NF*.

⁹ O termo “Propriedades Lexicais” diz respeito a entidades que não têm existência independente e só podem ser avaliadas em termos de sua aplicabilidade a outros tipos de entidades ou à situação que descrevem. Por exemplo, a propriedade “verde” se aplica a um Indivíduo, a propriedade “bater” se aplica a dois Indivíduos, estabelecendo uma relação entre eles (HENGEVELD; MACKENZIE, 2008, p. 131).

formação de Sintagmas ou Orações. A camada mais alta desse nível é a da Expressão Linguística (Le), unidade que pode ser especificada por qualquer conjunto de, ao menos, uma unidade morfossintática, ou seja, uma Expressão Linguística pode ser composta de uma combinação de Palavras (Xw), Sintagmas (Xp) ou Orações (Cl) ou pode ser composta de apenas uma dessas unidades. Uma Oração é um agrupamento de um ou mais Sintagmas e possivelmente Palavras, podendo conter também outras Orações. Um Sintagma consiste em uma combinação sequenciada de Palavras, contendo ainda, possivelmente, outros Sintagmas e Orações. Por fim, uma Palavra pode ser constituída por um (ou mais de um) morfema. A estrutura completa do NM é representada na Figura 4. A recursividade presente neste nível permite que cada uma das unidades apresentadas na camada da Expressão Linguística possa ocorrer mais de uma vez.

$$(Le_1: [(Xw_1) (Xp_1) (Cl_1: [(Xw_2) (Xp_2: [(Xw_3) (Xp_3) (Cl_3)] (Xp_2))_{(\Phi)} (Cl_2)_{(\Phi)}] (Cl_1))] (Le_1))$$

Figura 4: O Nível Morfossintático (HENGEVELD; MACKENZIE, 2008, p.17).

No Nível Fonológico (NF), a camada mais alta é a do Enunciado (U), unidade geralmente separada de outros Enunciados por pausas bem demarcadas. O Enunciado consiste em um ou mais Sintagmas Entoacionais (IP). O Sintagma Entoacional é descrito como contendo um núcleo, um movimento tonal, localizado em uma ou mais Sílabas, essencial para a interpretação do Sintagma Entoacional como um todo. Além disso, outra propriedade identificadora dessa unidade é o fato de ser separada de outros Sintagmas Entoacionais por uma pausa (mais breve do que a que ocorre entre Enunciados). O Sintagma Entoacional é composto de um (ou mais de um) Sintagma Fonológico (PP). A caracterização de um Sintagma Fonológico depende do tipo de língua com que se está lidando. Em línguas acentuais, por exemplo, um Sintagma Fonológico é caracterizado por conter uma Sílabas que é mais fortemente acentuada (sílabas nuclear) do que as outras sílabas do Sintagma. Um

Sintagma Fonológico geralmente contém uma ou mais Palavras Fonológicas (PW) que, por sua vez, podem ser analisadas em Sílabas (S) e em Pés (F). Esse nível, assim como o NM, permite a recursividade em algumas camadas. A estrutura completa do NF é representada na Figura 5.

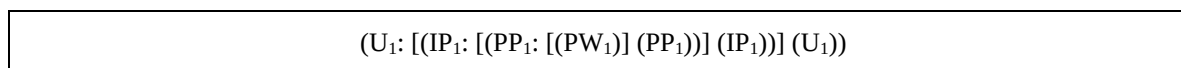


Figura 5: O Nível Fonológico (HENGEVELD; MACKENZIE, 2008, p.18).

Conforme mostraremos na seção 1.2 mais adiante, a concepção de transparência da GDF baseia-se, justamente, na existência ou não de alinhamento entre unidades desses diferentes níveis. Antes disso, na próxima seção, explicaremos como é feita a representação de um enunciado na GDF, para ilustrar melhor o funcionamento do modelo acima apresentado e para explicar alguns conceitos essenciais utilizados nesta tese.

1.1.2. A representação da referência por expressão do argumento-sujeito

Nesta seção, explicamos como seria a representação do enunciado em (1), considerando apenas os aspectos relevantes para o entendimento geral do modelo e dos conceitos usados em nosso trabalho.¹⁰

(1) Nós enviaremos as encomendas.

No NI, são representados os aspectos pragmáticos desse enunciado. Nesse caso, seria representado que o enunciado em (1) compreende um Move, que contém um Ato

¹⁰ Por essa razão, ilustraremos apenas a representação desse enunciado no NI, no NR e no NM, que são os níveis relevantes para a descrição do fenômeno aqui em estudo (não tratando da representação do NF).

Discursivo, constituído, por sua vez, pela Ilocução Declarativa, Falante, Ouvinte e um Conteúdo Comunicado. Este último compreende um Subato Atributivo (T_1) e três Subatos de Referência (R_1 , R_2 e R_3), um dos quais (R_3) contém um segundo Subato Atributivo (T_2). Esses Subatos (dentre os quais se encontram as unidades relevantes para os interesses do presente trabalho) seriam representados como em (2), em que é descrita apenas a camada do Conteúdo Comunicado¹¹.

$$(2) (C_1: [(T_1) (R_1: [+S, -A]) (R_2: [+S, -A]) (+id +s R_3: (T_2))]^C)$$

Em (2), o Subato Atributivo T_1 representa a evocação de uma propriedade realizada pelo uso do verbo *enviar*, enquanto os Subatos Referenciais R_1 , R_2 e R_3 representam as evocações de referentes realizadas, respectivamente, pelo pronome *nós*, pela desinência *-mos* e pelo sintagma nominal *as encomendas*. Os Subatos Referenciais correspondentes ao pronome (R_1) e à desinência (R_2) são preenchidos com os traços “+S” e “-A”, indicando respectivamente que o referente evocado envolve o Falante e não envolve o Ouvinte¹². Já o Subato Referencial correspondente ao sintagma nominal *as encomendas* (R_3) é preenchido com um Subato Atributivo (T_2), que representa a evocação de propriedade realizada pelo uso da palavra *encomenda*. Esse Subato Referencial (R_3) é marcado com os operadores “+id” e “+s”, indicando, respectivamente, que o Falante assume que o referente é identificável pelo Ouvinte e pelo próprio Falante (esses operadores são conjuntamente codificados pelo artigo definido *a*). Esses seriam os aspectos pragmáticos descritos no NI relativos aos Subatos.

¹¹ Em todas as representações desta seção, adotamos o modelo simplificado de representação utilizado em Keizer (2015).

¹² Na GDF, pronomes pessoais e afixos que fazem referência aos participantes do Ato Discursivo (primeira e segunda pessoas) – diferentemente dos pronomes e afixos que fazem referência à terceira pessoa – são essencialmente similares a nomes próprios e são considerados como substitutos gramaticais para o Falante nomear a si mesmo e para nomear a um destinatário. Esses pronomes e afixos são representados no NI pelos traços abstratos “+/- S” e “+/- A”, significando “envolve ou não o Falante” e “envolve ou não o Ouvinte” (cf. HENGVELD; MACKENZIE, 2008).

O fenômeno da referência estudado neste trabalho consiste na referência realizada por *Subatos de Referência*, como os exemplificados nesse enunciado. Na GDF, a noção de *referência* (assim como a de *atribuição*) é tomada como uma noção pragmática, como uma ação do Falante. A referência é a ação de evocar uma entidade (enquanto a atribuição é a evocação de uma propriedade).

Hengeveld e Mackenzie (2008) fazem menção a Dik (1978), observando que, segundo ele, a referência deve ser considerada como uma ação pragmática, cooperativa, de um Falante dentro de um modelo de interação verbal entre esse Falante e um Ouvinte. Hengeveld e Mackenzie afirmam que Dik liga o ato de referir ao que ele vê como a função primária da comunicação, qual seja, efetuar mudanças na informação pragmática do Ouvinte, considerando informação pragmática como a totalidade da informação (informação de longo termo, situacional e imediata) que os interlocutores mobilizam na interação. Hengeveld e Mackenzie definem que a GDF endossa tal posição considerando adicionalmente que a atribuição (*ascription*) é também acional – para eles, assim como a referência, a atribuição envolve uma tentativa do Falante de influenciar a informação pragmática do Ouvinte. Desse modo, os autores consideram a referência e a atribuição como acionais, como dois aspectos da ação mais global de *evocação*.

Nesse sentido, a GDF faz uma distinção crucial entre a ação de fazer referência a uma entidade (ou a uma propriedade), por um lado, e a designação dessa entidade (ou propriedade) por outro lado. No primeiro caso, tem-se um fato pragmático, descrito no NI, e no segundo, um fato semântico, descrito no NR.

Particularmente quanto à ação de fazer referência, a GDF assume como relevante a distinção entre *construção de referente* e *identificação de referente*. No primeiro caso, o Falante deseja que o Ouvinte introduza o referente em seu modelo mental, ao passo que, no segundo caso, o Ouvinte deve identificar (no sentido de reidentificar) um referente já

disponível em seu modelo mental. Essa distinção é captada na GDF pela oposição entre referência não-identificável (-id) e identificável (+id), respectivamente. Uma outra distinção tomada como relevante é aquela entre referência específica (+s) e não-específica (-s). Nesse caso, a questão é se o referente é identificável ou não para o Falante, respectivamente.

Tais aspectos são características da noção de *referência*, ou seja, a ação pragmática de fazer referência a uma entidade. A esse respeito, Hengeveld e Mackenzie (2008) explicam que, diante de uma expressão linguística referencial (como, por exemplo, *um leão*), dois fatos devem ser considerados: (i) *o Falante refere-se* a uma entidade (no caso, a uma entidade não-identificável e específica); (ii) *a expressão linguística designa* essa entidade (isto é, descreve semanticamente essa unidade, no caso, descrevendo, dentre outros traços, que se trata de um animal da classe dos leões). Desse modo, na representação da expressão *um leão* por exemplo, no NI são introduzidos os operadores “-id” e “+s”, enquanto no NR é introduzido o item lexical *leão*. Nesse sentido, explicam os autores que um termo-chave, no NI, é *evocação*, enquanto um termo-chave correspondente no NR é *designação*.¹³

Portanto, neste trabalho, usamos a noção de *referência* nesse contexto teórico definido pela GDF, tendo em vista, em particular, a noção de Subato de Referência. Mais especificamente, consideramos apenas os Subatos que fazem referência aos participantes do Ato Discursivo e, portanto, aqueles que são realizados por pronomes ou desinências verbais de 1ª e 2ª pessoa (como os Subatos R₁ e R₂, do enunciado em (1)), não analisando Subatos realizados por expressões nominais (como R₃), nem Subatos realizados por pronomes e

¹³ É importante esclarecer que, embora a noção de referência seja uma noção pragmática, isso não significa que a referência seja um fato independente dos domínios semântico e morfossintático. Na verdade, há correspondências entre os níveis Interpessoal, Representacional e Morfossintático. Um Subato de referência é realizado por alguma expressão linguística, que pertence a alguma categoria morfossintática, descrita no NM. Ao mesmo tempo, essa expressão linguística, dentre outras características semânticas, pode ocupar um papel em uma estrutura semântica (funcionando, por exemplo, como argumento de um esquema de predicação) e pode descrever semanticamente a entidade à qual faz referência, fatos descritos no NR. Nesse sentido, no decorrer desta seção, mostraremos a correspondência entre o Subato de Referência do NI e unidades do NR e do NM.

desinências de 3ª pessoa.¹⁴

Considerando agora a análise do enunciado em (1) no NR, focalizam-se, então, seus aspectos semânticos. No caso, seria descrito que o enunciado constitui um Conteúdo Proposicional, que contém um Episódio, o qual compreende um Estado de Coisas, este último sendo constituído por uma Propriedade e dois Indivíduos. As unidades constituintes do Estado de Coisas e as relações entre elas são representadas por meio de um esquema de predicação, formado, no caso, por um predicado (*enviar*) e dois argumentos (*nós/-mos* e *encomendas*). Em (3), fazemos a representação do NR desse enunciado, considerando apenas a camada do Estado de Coisas e os aspectos relevantes para a discussão.

(3) (e_1 : [(f_1 : enviar) (m x_1)_A (m x_2 : encomenda)_U] ^e)

Como se pode notar, o predicado (f_1) é preenchido com o item lexical *enviar*, que faz a descrição semântica da Propriedade constituinte do Estado de Coisas. Similarmente, o segundo argumento (x_2) é preenchido com o item lexical *encomenda*, que descreve esse argumento. Já o primeiro argumento (x_1), na medida em que corresponde a um dos participantes do Ato Discursivo, não inclui descrição semântica, permanecendo com o núcleo não preenchido; como mostrado em (2) mais acima, a indicação dos traços referentes ao Falante é feita no NI. Nessa representação, indicam-se também os traços de plural (“m”) de ambos os argumentos. São indicadas ainda as funções semânticas dos argumentos: “A” de Agente (*Actor*) e “U” de Paciente (*Undergoer*). Em nosso trabalho, a noção de *argumento* é utilizada como um elemento do esquema de predicação do NR, conforme exemplificado aqui pelos argumentos x_1 e x_2 .

No NM, é representada a organização morfossintática dos enunciados. No caso do

¹⁴ Vale esclarecer que, logo acima, ao apresentarmos as noções de Subatos Referenciais identificáveis/não-identificáveis e específicos/não-específicos, nosso intuito é apenas apresentar mais amplamente a noção de referência da GDF. Porém, conforme explicaremos no capítulo de metodologia, analisamos aqui apenas Subatos relativos a 1ª e 2ª pessoas do discurso, representados pelos traços abstratos “+/-S” e “+/-A”.

exemplo em análise, a representação indicaria que se trata de uma Expressão Linguística, composta por uma Oração (Cl), constituída por dois Sintagmas Nominais (Np) e um Sintagma Verbal (Vp), como se pode notar em (4).¹⁵

(4) (Cl₁:

(Np₁: (Nw₁: nós) ^{Np})_{Subj}

(Vp₁: (Vw₁: [(Vs₁: envi) (Aff₁: ind.fut) (Aff₂: 1.pl)] ^{Vw}) ^{Vp})

(Np₂: [(Gw₁: as) (Nw₂: encomendas)] ^{Np})

Cl)

O primeiro constituinte da Oração é um Sintagma Nominal (Np₁), formado pelo pronome *nós*, representado, de acordo com a GDF, como uma Palavra Nominal (Nw₁). O segundo constituinte é um Sintagma Verbal (Vp₁), preenchido unicamente pela Palavra Verbal *enviaremos* (Vw₁), a qual é composta pelo radical (Vs₁), pela desinência modo-temporal (Aff₁: ind.fut) e pela desinência número-pessoal (Aff₂: 1.pl).¹⁶ O terceiro constituinte é outro Sintagma Nominal (Np₂), que é formado pela Palavra Gramatical *as* (Gw₁) e pela Palavra Nominal *encomendas* (Nw₂).

No PB, um procedimento fundamental para a organização morfossintática dos enunciados, em termos da forma e da ordem dos constituintes morfossintáticos, é a atribuição da função sintática de sujeito a um dos argumentos do esquema de predicação. Em nosso trabalho, o termo *argumento-sujeito* refere-se justamente ao argumento do esquema de predicação ao qual é atribuída a função de sujeito. Como se pode ver em (4), a atribuição de sujeito é indicada na representação do NM.

¹⁵ Na representação em (4), para facilitar a visualização dos constituintes representados, dispomos cada um deles em uma linha, em vez de dispô-los linearmente como é a forma mais comum de representação.

¹⁶ A inserção da vogal temática *a*, do morfema *-re* (de futuro do presente do indicativo) e do morfema *-mos* (de primeira pessoa do plural) ocorre no NF.

O papel da atribuição de sujeito na organização morfossintática pode ser notado no exemplo aqui em pauta. Nesse caso, atribui-se a função de sujeito ao argumento x_1 , o que aciona a construção de oração na voz ativa. Considerando o mesmo esquema de predicação do enunciado em análise, a atribuição da função de sujeito ao argumento x_2 já acionaria a construção da oração na voz passiva (*as encomendas serão enviadas por nós*).

A noção de *argumento-sujeito* é crucial para o presente trabalho. No PB, a possibilidade de ocorrência do fenômeno da RC (e, por extensão, a existência do sistema de referência aqui em estudo, formado pela RC e suas formas alternativas de referência – RS e RZ) verifica-se apenas em relação ao *argumento-sujeito*, não estando relacionada a alguma outra característica de um argumento, como o estatuto de argumento 1 (x_1) ou de argumento 2 (x_2), ou à função semântica do argumento (por exemplo, Agente ou Paciente).

No exemplo em análise, que se encontra na voz ativa, a RC verifica-se em relação a x_1 . Porém, se a oração ocorre na voz passiva, a RC já envolve x_2 (*as encomendas serão enviadas por nós*). Assim, não poderíamos definir o objeto de estudo deste trabalho em termos de referência a argumento 1 ou a argumento 2. Da mesma forma, no exemplo em pauta, a RC verifica-se em relação a um argumento com função semântica de Agente. Já no caso da versão passiva desse exemplo, a RC seria realizada em relação ao argumento com função de Paciente. Assim, também não seria possível definir nosso objeto de estudo em termos de referência ao argumento com determinado tipo de função semântica. Por outro lado, tanto no exemplo aqui em análise, quanto na sua versão na voz passiva, o argumento que pode ser referido por RC é aquele a que é atribuída a função sintática de sujeito (*nós*, na voz ativa, e *as encomendas*, na passiva). Por essa razão, na constituição do nosso objeto de estudo, falamos em termos de *referência ao argumento-sujeito* (e não em referência ao argumento 1, ao argumento Agente ou a outro tipo de argumento).

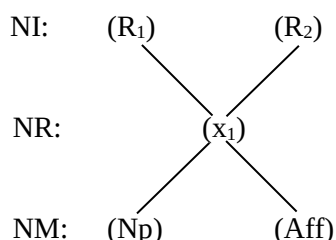
O fato de a RC estar relacionada ao argumento-sujeito deve-se naturalmente ao

sistema de expressão morfossintática do argumento-sujeito da gramática do PB, que compreende a expressão desse argumento por desinência verbal (que pode consistir em morfema explícito ou zero) e a possibilidade de expressão por um sintagma nominal inserido em uma posição na estrutura sintática da oração. Desse modo, na delimitação do nosso objeto de estudo, falamos, mais precisamente, de *referência por expressão do argumento-sujeito* – como restringimos nosso trabalho à referência aos participantes da interação verbal (o que justificaremos no capítulo de metodologia), a caracterização do nosso objeto de estudo é feita ainda em termos de *referência por expressão pronominal e desinencial do argumento-sujeito de 1ª e 2ª pessoas*.

Em síntese, o fenômeno da referência aqui em estudo, considerando sua representação na GDF, envolve, no NI, a unidade do Subato de Referência, no NR, a categoria de argumento (particularmente aquele a que é atribuída a função de sujeito) e, no NM, os usos de pronome e de desinência verbal de 1ª e 2ª pessoas.

Especificamente, tendo em vista as formas de referência aqui estudadas (RC, RS e RZ), a RC consiste na correspondência entre dois Subatos de Referência no NI, um argumento no NR e duas formas de expressão (pronome e desinência) no NM. Essa situação pode ser ilustrada no esquema em (5).

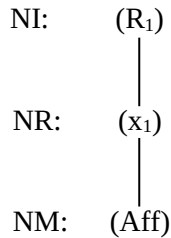
(5) Correspondência entre níveis na RC



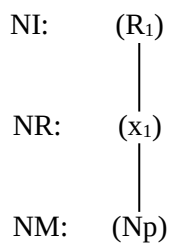
Já a RS envolve a correspondência entre um Subato de Referência no NI, um argumento no NR e uma forma de expressão no NM, que pode ser a desinência verbal explícita, conforme ilustrado em (6), ou o pronome (caso em que a desinência verbal é

realizada por morfema zero), o que pode ser ilustrado em (7).

- (6) Correspondência entre níveis na RS com expressão desinencial

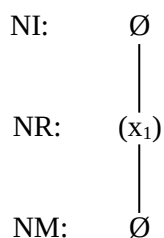


- (7) Correspondência entre níveis na RS com expressão pronominal



Já a RZ envolve um argumento no NR ao qual não corresponde a realização de um Subato Referencial, nem uma expressão morfossintática explícita, o que pode ser representado pelo esquema em (8).

- (8) Correspondência entre níveis na RZ¹⁷



¹⁷ Convém aqui fazer um esclarecimento sobre o conceito de “Referência Zero” adotado neste trabalho. Não se trata de considerar que não haveria um referente (uma entidade) envolvido na RZ. Quando, num enunciado, ocorre RZ, a significação do enunciado envolve um referente, mas esse enunciado indica esse referente justamente pela *ausência* de um ato linguístico de referência, pelo silêncio (aqui, o silêncio é significativo). Na RC, ocorrem duas ações linguísticas (dois Subatos de Referência) que evocam um mesmo referente; na RS, ocorre uma única ação linguística (um Subato) que evoca um referente; na RZ, o enunciado envolve um referente, mas o indica justamente pela não-realização de uma ação linguística (não há um Subato de Referência). Portanto, a noção de RZ representa a ausência de uma *ação linguística de referência*, e não a ausência de um *referente* na significação do enunciado.

Após termos apresentado o modelo da GDF e ilustrado seus aspectos mais relevantes para nosso trabalho, explicamos, na seção seguinte, a noção de transparência linguística.

1.2 A propriedade da transparência linguística

A transparência é frequentemente definida como uma relação de um-para-um entre unidades de significado e unidades de forma, mas, no âmbito da GDF, essa definição é um pouco mais complexa. A teoria assume que uma relação completamente transparente ocorre numa língua quando há um alinhamento de um-para-um entre unidades dos níveis Interpessoal, Representacional, Morfossintático e Fonológico. Ou seja, em vez de definir transparência simplesmente como mapeamento de um-para-um entre significado e forma, a GDF a define como mapeamento de um-para-um (na verdade, um-para-um-para-um-para-um) entre todos os quatro níveis do Componente Gramatical. De acordo com Hengeveld e Leufkens (no prelo), uma relação completamente transparente ocorre, por exemplo quando uma única unidade no NI (por exemplo um Subato de Referência) corresponde a uma única unidade no NR (por exemplo um Indivíduo), no NM (por exemplo um Sintagma Nominal) e no NF (por exemplo um Sintagma Fonológico). Portanto, todos os níveis dentro da gramática e as interações entre eles são levados em consideração.

O grau de transparência/opacidade de uma língua, nessa teoria, é calculado analisando-se em que medida a língua apresenta fenômenos linguísticos que violam a relação de um-para-um entre os quatro níveis. São distinguidos, assim, dois grandes conjuntos de fenômenos linguísticos que violam a transparência. Um primeiro conjunto inclui fenômenos que envolvem descrição, ou mapeamento, nos diferentes níveis da gramática, porém esses mapeamentos não são transparentes, isto é, não são alinhados entre si (por exemplo, um

fenômeno cuja descrição inclui duas unidades do NI, mas uma unidade do NR). Um segundo conjunto inclui fenômenos que são gerados apenas dentro dos níveis de codificação, violando, portanto, a transparência porque sua descrição envolve unidades desses níveis (NM e NF), mas não envolve os níveis Interpessoal e Representacional.

Hengeveld e Leufkens (no prelo) referem-se ao primeiro conjunto de fenômenos mencionado acima como fatores de transparência/opacidade resultantes de relações *entre níveis*. Os fenômenos do segundo conjunto são vistos como fatores de transparência/opacidade resultantes de operações que podem ocorrer *dentro dos níveis*. Nas seções 1.2.1 e 1.2.2 a seguir, explicamos e exemplificamos, respectivamente, esses dois tipos de fatores de transparência/opacidade e, em 1.2.3, apresentamos a lista completa de fatores de transparência/opacidade considerados pela GDF para o estudo de transparência.

1.2.1 Fatores de transparência/opacidade decorrentes de relações entre níveis

Em princípio, poder-se-ia considerar que a violação da transparência decorreria de uma correspondência não-biunívoca entre quaisquer dois níveis dentre os quatro níveis do modelo. Essa visão consideraria que poderia ocorrer violação de transparência entre os seguintes pares de níveis: (i) NI-NR, (ii) NI-NM, (iii) NI-NF, (iv) NR-NM, (v) NR-NF, (vi) NM-NF. No entanto, a GDF assume uma perspectiva descendente no entendimento da relação entre os níveis e considera que o mapeamento de um nível leva em consideração todos os níveis superiores a ele. Desse modo, a violação de transparência, no que diz respeito à relação entre níveis, pode ocorrer de três formas diferentes: (i) na relação entre o NR e o NI; (ii) na relação entre o NM, por um lado, e o NI e o NR, por outro; (iii) na relação entre o NF, por um lado, e o NI, o NR e o NM, por outro. Abaixo, explicamos alguns fenômenos que, segundo a GDF, representam essas três possibilidades de violação.

A) Violação de transparência entre os níveis Interpessoal - Representacional

Dentre os fenômenos linguísticos que representam violação de transparência por desalinhamento entre o NR e o NI, Hengeveld e Leufkens (no prelo) incluem a Aposição e a RC.

A Aposição, segundo os autores, consiste no processo de se fazer referência a uma única entidade duas ou mais vezes por meio de elementos lexicais, gerando uma correspondência de dois (ou mais) para um entre o NI e o NR, respectivamente. O exemplo abaixo é um caso de Aposição fornecido pelos autores.

- (9) *Peter, John's brother, has moved to Norway* (Peter, irmão de John, mudou-se para a Noruega).

(HENGEVELD; LEUFKENS, no prelo).

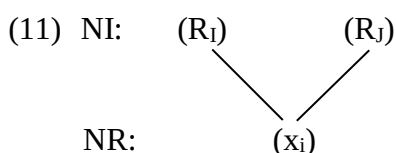
Nesse exemplo, *Peter* e *John's brother* fazem referência a uma mesma entidade e podem fazer isso de forma independente um do outro. Nos termos da GDF, isso significa que dois Subatos de Referência do NI correspondem a um único Indivíduo no NR, constituindo uma correspondência não-transparente.

A RC é um caso especial de Aposição. Por meio desse fenômeno, faz-se também referência a uma única entidade duas ou mais vezes, porém essa referência não é feita apenas por meio de elementos lexicais, como em (9), mas, ao mesmo tempo, por elementos lexicais e por elementos gramaticais. Segundo Hengeveld e Leufkens (no prelo), a noção de RC só é aplicável quando os elementos em “aposição” são capazes de fazer referência por si mesmos, isto é, quando a sentença é gramatical mesmo que seja usado apenas um desses elementos. O exemplo em (10) ilustra um caso de RC, no português europeu, segundo Leufkens (2013).

- (10) *eu* *chegue-i* (eu cheguei).
1SG arrive-1SG

(LEUFKENS, 2013, p. 340).

De acordo com a autora, em (10), faz-se referência a uma mesma entidade duas vezes, por meio do sufixo verbal e do pronome. A referência feita pelo sufixo é, segundo ela, comprovada na medida em que o uso do verbo, que recebe marcação de número-pessoa, é gramatical mesmo sem o uso do pronome, ou seja, o verbo sozinho é capaz de constituir um enunciado completo. A presença do pronome, nesse caso, é redundante. Nos termos da GDF, isso é interpretado como uma correspondência entre dois Subatos de Referência, no NI, e um Indivíduo, no NR, sendo um desses Subatos de Referência expresso gramaticalmente, por meio de sufixo verbal, e o outro expresso lexicalmente, por meio de sintagma nominal opcional. Segundo Hengeveld e Leufkens (no prelo), tanto a Aposição quanto a RC podem ser representadas pelo esquema em (11), que evidencia que esses dois fenômenos envolvem desalinhamento entre o NI e o NR.¹⁸



(HENGEVELD; LEUFKENS, no prelo).

Como definido acima, neste trabalho selecionamos o parâmetro da RC para estudo da transparência do PB. Porém, consideramos a RC como parte de um sistema de referência

¹⁸ A transparência ou opacidade da referência aqui em estudo decorre da combinação entre presença/ausência de pronome e presença/ausência de desinência número-pessoal – a coocorrência de pronome e desinência representa opacidade (RC), como estamos explicando. É preciso reconhecer ainda que o simples uso da desinência, assim como o uso do pronome, por si só, já envolve opacidade, pois a desinência, bem como pronome, acumula em si dois significados (número e pessoa). No entanto, a junção de dois significados numa mesma forma constitui um tipo de opacidade diferente da RC, considerado em alguns trabalhos e denominado de *acumulação* (cf. LEUFKENS, 2015): a RC consiste em uma dupla referência a uma mesma entidade (como a coocorrência de referência por pronome e por desinência número-pessoal aqui estudada), e a acumulação é a expressão de múltiplos significados em uma mesma forma (como é o caso da expressão de modo e tempo em desinências modo-temporais, bem como o caso da expressão de número e pessoa em desinências número-pessoais). Neste trabalho estudamos apenas o tipo de opacidade relacionado às diferentes formas de referência (o que inclui a RC). Não estudamos a opacidade decorrente da acumulação.

constituído não só pela própria RC, mas também pelo que estamos chamando de RZ e RS. Assim, nosso estudo não envolve apenas atestar ou não a presença da RC no PB, mas averiguar, em diferentes sincronias, as proporções em que a gramática do PB prevê construções de RS (transparentes, portanto) e construções de RC e RZ (opacas). Ou seja, nosso estudo focaliza o grau de transparência do sistema de referência do PB (em diferentes sincronias, considerando, como já explicado, particularmente a referência obtida no processo de expressão pronominal e desinencial do argumento-sujeito de 1ª e 2ª pessoas).

Conforme explicam Hengeveld e Leufkens (no prelo), a GDF distingue RC de Concordância. A Concordância, como mostrado abaixo, é um mecanismo puramente morfossintático que, como tal, não envolve uma dupla referência.

B) Violação de transparência entre os níveis Interpessoal/Representacional e Morfossintático

A segunda forma de violação de transparência por desalinhamento de relação entre níveis concerne à relação entre o NM, por um lado, e o NI e o NR, por outro lado. Hengeveld e Leufkens (no prelo) não chegam a explicar mais detalhadamente essa situação de desalinhamento entre níveis. No nosso entendimento, esse tipo de violação ocorre quando há alinhamento entre o NI e o NR, mas o NM está desalinhado com relação aos dois anteriores. Isso ocorreria, por exemplo, quando uma unidade do NI corresponde a uma unidade do NR, mas estas correspondem a duas unidades do NM.¹⁹

Um fenômeno representativo desse tipo de violação, segundo Leufkens (2015), é a descontinuidade de constituintes. De acordo com a autora, a descontinuidade ocorre quando uma única unidade pragmática ou semântica corresponde a uma unidade morfossintática que é dividida em mais de uma parte (ou seja, a unidade morfossintática é expressa de modo

¹⁹ Esse mesmo raciocínio se aplicaria à relação entre o NF, por um lado, e os três níveis superiores, por outro, que também não é detalhadamente explicada pelos autores.

descontínuo). Entre os subtipos de constituintes descontínuos estão fenômenos como Extraposição e Extração.

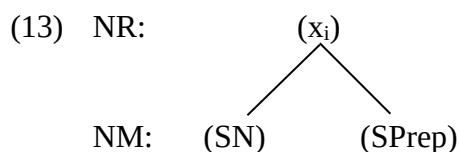
Um caso de descontinuidade seria quando um modificador, na forma de um sintagma ou de uma oração, não é expresso adjacente a seu núcleo. Se esse modificador é expresso no fim da sentença, devido a princípios morfossintáticos que posicionam elementos complexos no fim da sentença e elementos simples em seu início, tem-se a Extraposição. Se esse modificador ocupa a posição à esquerda de seu núcleo, devido a princípios pragmáticos (como topicalização, por exemplo), tem-se a Extração. Leufkens (2015) utiliza os exemplos em (12b) e (12c) para ilustrar casos de Extraposição e de Extração, respectivamente (a sentença em (12a) seria o correspondente de (12b) e (12c) sem descontinuidade).

- (12) a. We have several important *books about global warming* in stock (Nós temos muitos livros importantes sobre aquecimento global em estoque).
- b. We have several important *books* in stock [*about global warming*] (Nós temos muitos livros importantes em estoque [sobre aquecimento global]).
- c. [*About global warming*] we have several important *books* in stock ([Sobre aquecimento global] nós temos muitos livros importantes em estoque).

(VAN DE VELD, 2012 *apud* LEUFKENS, 2015, p.64).

No exemplo acima, Leufkens (2015) explica que tanto a Extraposição quanto a Extração criam um desalinhamento entre níveis. Tanto no caso da sentença (12b), quanto no caso da sentença (12c), o NR contém um Indivíduo (*books*) que tem a propriedade de ser *about global warming*. Essa propriedade funciona como um modificador do Indivíduo, como parte da entidade semântica. Já no NM, o Indivíduo e o modificador não são expressos de forma adjacente, sendo codificados como duas unidades morfossintáticas. O desalinhamento criado por esses fenômenos reside, então, no fato de uma unidade de significado ser codificada com duas unidades morfossintáticas. Hengeveld e Leufkens (no prelo) representam

essa situação pelo esquema em (13).



(HENGEVELD; LEUFKENS, no prelo).

A nosso ver, nesses casos de Extraposição e Extração representados em (13), ocorreria, no NI, assim como no NR, uma única unidade, no caso, um Subato de Referência. Assim sendo, os subtipos de descontinuidade exemplificados aqui – Extraposição e Extração – constituem fenômenos de desalinhamento envolvendo o NI e o NR, por um lado, e o NM, por outro. Nesse sentido, a presença de fenômenos de descontinuidade em uma língua constitui um fator de opacidade.

C) Violação de transparência entre os níveis Interpessoal/Representacional/Morfossintático e Fonológico

O terceiro tipo de desalinhamento entre níveis distinguido por Hengeveld e Leufkens (no prelo) é o desalinhamento entre o NF, por um lado, e os três níveis superiores, por outro. Um fenômeno que ilustra esse tipo de desalinhamento seria a ausência de paralelismo entre a estruturação sintagmática morfossintática e a estruturação sintagmática fonológica. Em certas línguas, como o achém (língua austronésia falada no território de Achém, na ilha de Sumatra, Indonésia), existe um mapeamento de um-para-um entre sintagmas morfossintáticos e sintagmas fonológicos. Em outras línguas, esse paralelismo já não se verifica, o que ocorre, por exemplo, em holandês, como mostra o exemplo em (14).

- (14) ik wou dat hij kwam
 I want.PST COMP he come.PST
 /kvau dati kvam/
 'I wish he would come' (Eu gostaria que ele viesse).

(LEUFKENS, 2015, p. 67).

Em (14), o sujeito e o verbo da oração principal formam uma palavra fonológica (kvau), assim como a conjunção e o sujeito da oração subordinada (dati), gerando um desalinhamento entre a estrutura sintagmática morfossintática e a fonológica. Os autores representam esse desalinhamento pelo esquema em (15).

- (15) NM: (Cli: [(Np_i)_{sbj} (Vp_i) (Cl_j: [(Gw_i) (Np_j)_{sbj} (Vp_j)] (Cl_j))] (Cl_i))

 NF: (PP_i: (PW_i) (PP_i)) (PP_j: [(PW_j) (PW_k)] (PP_j))

(HENGEVELD; LEUFKENS, no prelo).

A ausência de paralelismo entre as estruturações sintagmáticas morfossintática e fonológica constitui um fenômeno de desalinhamento envolvendo o NF, por um lado, e NI, NR e NM, por outro. Nesse sentido, a ausência desse paralelismo em uma língua constitui um fator de opacidade.

1.2.2 Fatores de transparência/opacidade decorrentes de operações dentro dos níveis

O segundo grupo de fenômenos linguísticos identificados por Hengeveld e Leufkens (no prelo) que violam a transparência inclui fenômenos que são gerados apenas dentro dos níveis de codificação (NM e NF), sem envolver o NI e o NR. Dentro desses níveis de codificação podem ocorrer operações puramente formais que adicionam elementos ou traços a estruturas construídas a partir de material vindo dos níveis mais altos (NI e/ou NR). A existência, em uma língua, de fenômenos gerados nesses níveis formais é um fator de

opacidade na medida em que esses fenômenos se constituem em formas sem significado pragmático ou semântico. Quanto ao NM, Hengeveld e Leufkens (no prelo) consideram os seguintes fenômenos: Expletivo, Gênero Gramatical, Concordância Sintagmática e Concordância Verbal; quanto ao NF, consideram fenômenos de adaptação fonológica, como Assimilação Nasal, Ditongação, Ensurdimento Final e Harmonia Vocálica.

A) Fatores de transparência/opacidade dentro do Nível Morfossintático

Um exemplo de fenômeno linguístico gerado dentro do NM tomado como fator de opacidade é a Concordância Verbal. Segundo a GDF, nesse fenômeno, uma propriedade de um argumento é copiada para outra unidade dentro da Oração. O exemplo em (16), do francês, extraído de Hengeveld e Mackenzie (2008), pode ilustrar esse fenômeno.

- (16) Nous chant-ons.
 We sing-1.pl
 ‘We sing’ (Nós cantamos)

(HENGEVELD; MACKENZIE, 2008, p. 350).

A sentença em (16) apresenta um verbo com uma marcação de 1ª pessoa do plural e um pronome pessoal de 1ª pessoa do plural. Conforme explicam Hengeveld e Mackenzie (2008), a presença desse pronome é, no caso dessa língua, obrigatória, ou seja, a ausência do pronome resultaria em uma sentença agramatical. Nesse caso, as informações de pessoa e número contidas no verbo são informações copiadas do pronome sujeito.

Na GDF, faz-se distinção entre Concordância Verbal e RC, como já apontado anteriormente. A Concordância Verbal é uma operação morfossintática de cópia obrigatória de informações de um argumento para outros elementos dentro da Oração. Nesse caso, a flexão verbal não é uma codificação morfossintática de significados de níveis mais altos, mas

resulta de uma operação que tem origem dentro do próprio NM, uma operação puramente formal, sem contraparte pragmática ou semântica. Nesse sentido, na Concordância, a flexão verbal não tem natureza referencial, enquanto, na RC, como visto acima, a flexão verbal é referencial por si mesma, assim como o sintagma nominal que coocorre opcionalmente com essa flexão.

A Concordância Verbal é um fenômeno não-transparente, na medida em que acarreta uma forma vazia de significado pragmático ou semântico no enunciado (a flexão verbal).

B) Fatores de transparência/opacidade dentro do Nível Fonológico

De acordo com Hengeveld e Leufkens (no prelo), no NF, assim como no NM, ocorrem várias operações que adicionam elementos ou aspectos às estruturas que foram construídas com base em material disponibilizado pelos níveis mais altos. Isso ocorre, por exemplo, quando regras fonológicas adaptam fonemas subjacentes para seu ambiente de superfície. Os autores citam como exemplos dessas operações fonológicas fenômenos como a Assimilação Nasal, a Ditongação, o Ensurdecimento e a Harmonia Vocálica. O exemplo de Ditongação no espanhol, em (17), ilustra uma dessas operações.

(17) dorm-ir
sleep-INF
'to sleep' (dormir)

duerm-e
sleep-PRS.3.SG
'(he/she) sleeps' ((ele/ela) dorme)

(HENGEVELD; MACKENZIE, 2008, p. 350 - adaptado).

Em (17), a transformação da vogal “o” (do radical do verbo “dormir”) no ditongo

“ue”, chamada *irregularidade vocálica*, não ocorre em todos os radicais de verbos que contêm a vogal “o”, mas apenas naqueles radicais cuja sílaba que contém a vogal “o” é tônica. Para Hengeveld e Leufkens (no prelo), em casos como esse, a relação de um-para-um entre significado e forma é obscurecida, pois o significado expresso pelo morfema recebe uma manifestação formal diferente dependendo do contexto em que ocorre. Isso equivale a dizer que ocorre uma modificação formal que não corresponde a uma modificação de significado. Operações como essas têm origem dentro do NF e, por serem operações puramente formais, constituem fenômenos de opacidade em uma língua.

1.2.3 Quadro final de fatores de transparência/opacidade

Nas seções anteriores, exemplificamos alguns fenômenos que são levados em consideração para se calcular o grau de transparência de uma língua. Hengeveld e Leufkens (no prelo) explicam que diversos fenômenos linguísticos podem ser considerados para esse cálculo, dentre os quais os autores selecionam um conjunto de 11 fenômenos, tomados, então, como parâmetros para aferir o grau de transparência/opacidade. Esse conjunto completo de fenômenos é apresentado abaixo, no quadro 1.

Quadro 1: Parâmetros para cálculo da transparência linguística

Fenômenos não-transparentes	Interface/Nível
1. Aposição	NI-NR
2. Referência Cruzada	NI-NR
3. Relações gramaticais	NI/NR - NM
4. Alternância de raiz morfo(fono)logicamente condicionada	NI/NR - NM
5. Constituintes descontínuos	NI/NR - NM
6. Ausência de paralelismo entre as estruturas morfosintática e fonológica	NI/NR/NM - NF
7. Expletivos	NM
8. Gênero Gramatical	NM
9. Concordância Gramatical Sintagmática	NM
10. Concordância Gramatical Oracional	NM
11. Adaptações fonológicas	NF

Fonte: Hengeveld e Leufkens (no prelo).

De acordo com os autores, quanto menor o número de fenômenos que uma língua apresenta, dentre os 11 listados no quadro, mais transparente essa língua é. Ao contrário, quanto mais a língua apresentar fenômenos dentre esses 11, menos transparente é essa língua. Ou seja, o grau de transparência de uma língua é calculado analisando-se a presença ou ausência desses fenômenos. Como definido anteriormente, selecionamos neste trabalho um desses fenômenos, a RC, para, a partir dele, analisar a transparência do PB.

Cumpre ainda mencionar uma questão quanto a nossa seleção do parâmetro da RC. Hengeveld (2012), no âmbito de discussões sobre as noções de RC e de Concordância Verbal, delinea uma terceira possibilidade, chamada de *Concordância Contextual*, que ocorreria quando um marcador verbal concorda com um argumento que pode estar presente tanto na própria oração, quanto no contexto.

Diante dessa noção alternativa da Concordância Contextual, poder-se-ia considerar a possibilidade de que o PB apresentaria esse fenômeno, em vez de RC. No entanto, o próprio autor reconhece que se trata de uma noção ainda experimental. Além disso, não há trabalhos publicados que evidenciem que esse seria o tratamento adequado para o PB. Também não nos propusemos a uma discussão detalhada sobre a diferença entre essas noções, pois tal tarefa certamente extrapolaria os limites e até mesmo o foco do presente trabalho. Desse modo, optamos por interpretar o processo apresentado pelo PB como RC, seguindo a oposição já consolidada entre RC e Concordância Verbal, adotada em Hengeveld e Mackenzie (2008), Leufkens (2013), Hengeveld e Leufkens (no prelo), entre outros, sem deixar de reconhecer que a discussão sobre as noções de RC, Concordância Verbal e Concordância Contextual representa um tema relevante e promissor de pesquisa.

CAPÍTULO II

Transparência, diacronia e princípios gerais de funcionamento da linguagem

Este trabalho, ao tratar da propriedade linguística da transparência sob perspectiva diacrônica, encontra alguns pontos centrais de fundamentação teórica num conjunto de estudos funcionalistas que podem ser aqui agrupados pelo fato de teorizarem sobre as motivações cognitivas/comunicativas da estrutura formal das línguas, bem como sobre o papel dessas motivações em processos de mudança histórica. Nesse grupo, incluem-se trabalhos como os de Haiman (1980, 1983, 1985), Givón (1985), Slobin (1977, 1980) e Croft (2003). As referidas motivações compreendem, principalmente, as noções de *iconicidade* e *economia*, frequentemente tomadas como forças que competem entre si na determinação da estrutura linguística, sendo tratadas, por vezes, como *motivações em competição*. Desse modo, nesta seção, sintetizamos os pressupostos mais importantes desses trabalhos para a presente tese e apresentamos algumas conclusões teóricas que pudemos formular sobre a relação entre os conceitos desenvolvidos em tais trabalhos e nossa análise da transparência linguística, as quais contribuem para sustentar nossa tese.

Com esse propósito em mente, em 2.1 apresentamos uma sistematização da noção de *iconicidade*, para mostrar que a propriedade da transparência linguística aqui considerada equivale a uma das modalidades de iconicidade linguística, a saber, o *isomorfismo* (mais especificamente, o *isomorfismo sintagmático*). Ao identificar essa correspondência, essa seção, dentre outras contribuições, procura mostrar que a análise da transparência pode ser vinculada ao contexto do estudo da noção de iconicidade, sendo possível aproveitar a extensa bibliografia existente sobre iconicidade linguística na análise da noção de transparência.

Em 2.2, mostramos que os trabalhos em questão concebem que, em grande medida, a estrutura formal das línguas é determinada pela interação entre certas motivações,

dentre elas a iconicidade e a economia. Com base nessa concepção, consideramos, então, que o grau de transparência de uma língua, num dado momento, pode também ser entendido como sendo, em boa medida, o resultado da interação entre motivações em competição.

Na mesma direção, em 2.3, sintetizamos trabalhos de alguns autores que abordam a noção de *teto de transparência/opacidade* e argumentamos que, dada a atuação da iconicidade e da economia, o grau de transparência de uma língua e, na mesma direção, o grau de transparência do sistema de referência do PB aqui em estudo oscilariam, diacronicamente, dentro de um limite máximo (teto) e mínimo (piso) de transparência.

2.1 Iconicidade e transparência

Considerando diferentes autores que abordam a noção de iconicidade, principalmente Haiman (1980, 1983, 1985) e Croft (2003), pode-se considerar que a transparência constitui um subtipo de iconicidade.²⁰ A figura 6, que elaboramos com base nesses autores, mostra a divisão da noção de iconicidade em subtipos, sendo um deles o isomorfismo, o qual, conforme explicamos adiante, corresponde à noção de transparência adotada neste trabalho.

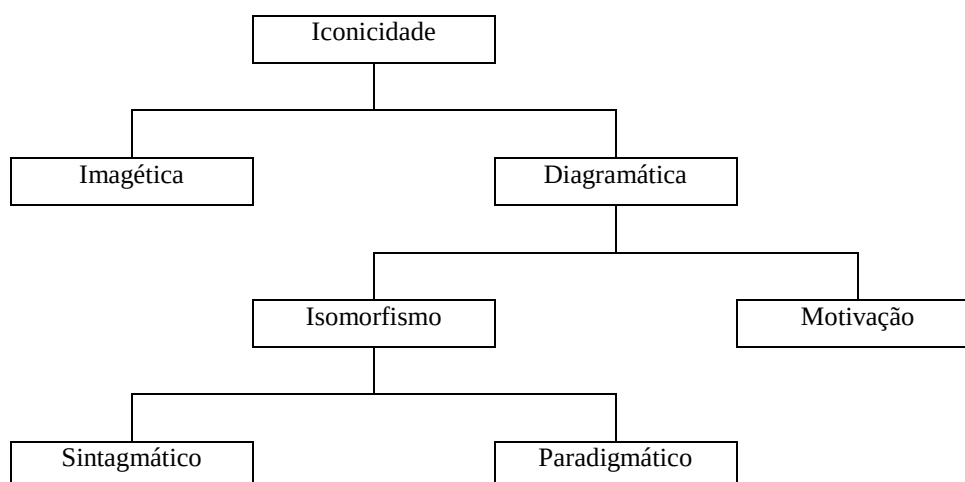


Figura 6: Subtipos de iconicidade, segundo trabalhos de Haiman (1980, 1983, 1985) e Croft (2003).

²⁰ Leufkens (2015) também discute a relação entre transparência e iconicidade.

A iconicidade pode ser definida, em termos gerais, como a existência de algum tipo de relação de similaridade entre forma e significado. Trata-se de uma propriedade linguística que consiste na existência de uma relação não-arbitrária entre forma e significado. De acordo com Croft (2003, p. 102), “a intuição por trás da iconicidade é que a estrutura da língua reflete de alguma forma a estrutura da experiência, isto é, a estrutura do mundo, incluindo a perspectiva imposta sobre o mundo pelo falante”.

Retomando Peirce (1932), Haiman (1980) considera uma distinção crucial entre dois tipos de iconicidade: *imagética* e *diagramática*. A iconicidade imagética é aquela em que um signo individual assemelha-se (como um todo) a seu referente em relação a alguma característica, como ocorre no caso das onomatopeias. A iconicidade diagramática é um arranjo sistemático de signos, nenhum dos quais necessariamente assemelham-se a seus referentes, mas cujas relações entre si espelham as relações entre seus referentes.²¹

Haiman (1980, 1985) estabelece ainda uma distinção entre *iconicidade diagramática por motivação* e *iconicidade diagramática por isomorfismo* – segundo Neves (2004), Croft (2003) estabelece uma distinção similar que denomina respectivamente de *iconicidade por correspondência de relação entre partes* e *iconicidade por correspondência entre partes*.

A *iconicidade diagramática por motivação* ocorre quando a relação entre elementos da estrutura formal assemelha-se à relação entre elementos correspondentes da estrutura conceitual. Já a *iconicidade diagramática por isomorfismo* refere-se à correspondência biunívoca entre elementos da estrutura formal e elementos da estrutura conceitual. Como se vê, a noção de transparência usada no presente trabalho equivale à noção de isomorfismo adotada em geral em trabalhos sobre o fenômeno da iconicidade.

Vários tipos de iconicidade diagramática por motivação são distinguidos: (i)

²¹ De acordo com Haiman (1985), idealmente, um diagrama icônico é homólogo àquilo que representa: (a) cada ponto no diagrama corresponde a um dado ponto da realidade descrita e (b) as relações entre esses pontos correspondem às relações entre os respectivos pontos na realidade.

iconicidade de quantidade, segundo a qual um texto maior deve conter mais informação do que um menor; (ii) *iconicidade de distância*, pela qual a distância linguística entre expressões corresponde à distância conceitual existente entre elas; (iii) *iconicidade de independência*, de acordo com a qual a separação linguística de uma expressão corresponde à independência conceitual do objeto ou evento que a expressão representa; (iv) *iconicidade de ordenação*, que estabelece que o grau de importância atribuído aos conteúdos de um texto pelo falante determina a ordenação de suas unidades formais; (v) *iconicidade de complexidade*, segundo a qual, por exemplo, a forma plural é maior que a singular; (vi) *iconicidade de categorização*, pela qual, por exemplo, sujeitos tendem a se correlacionar com agentes, objetos com pacientes etc. (NEVES, 2004).

A iconicidade diagramática por isomorfismo, a que mais nos interessa neste trabalho, é um tipo de iconicidade na medida em que envolve uma similaridade entre forma e significado. Mais especificamente, a similaridade envolvida nesse caso reside numa correspondência quantitativa: se uma determinada estrutura semântica distingue um número x de significados, a estrutura formal que a codifica também distingue esse mesmo número x de formas, havendo uma, e apenas uma, forma para cada significado.

Croft (2003) estabelece uma distinção entre isomorfismo sintagmático e isomorfismo paradigmático. O primeiro tipo, que é o que mais especificamente falando corresponde à noção de transparência em foco neste trabalho, consiste na correspondência biunívoca entre forma e significado encontrada na combinação sintagmática de formas (palavras e morfemas) em uma sentença. Já o isomorfismo paradigmático consiste na correspondência entre forma e significado no inventário de formas (palavras e formas) estocadas na mente. Em outras palavras, isomorfismo sintagmático refere-se ao fato de uma forma ter um significado (e não mais de um) em seu emprego concreto, contextualizado em uma sentença, enquanto o isomorfismo paradigmático diz respeito ao fato de uma forma ter

um significado (e não mais de um) no léxico da língua.

Em síntese, a propriedade linguística da transparência pode ser vista como uma forma de iconicidade. Desse modo, entendemos que nosso trabalho sobre a dinâmica diacrônica da transparência linguística, particularizado pela investigação da diacronia do sistema de referência aqui em análise, pode e deve levar em consideração os estudos sobre as motivações em competição na determinação da estrutura formal das línguas, dentre elas a iconicidade. Nas duas seções seguintes, bem como no decorrer de toda a tese, e principalmente no capítulo 5, pressupomos essa aproximação entre iconicidade e transparência.

2.2 Motivações em competição

Como mencionamos acima, o conjunto de trabalhos funcionalistas aqui em pauta concebe que a estrutura formal das línguas é, em grande medida, determinada por motivações em competição, como a iconicidade e a economia principalmente, além de outras. Slobin (1977) é um dos autores que consideram que as línguas estão sempre pressionadas por motivações complexas e contrárias. Segundo ele, o falante de uma língua quer se expressar de forma clara, eficiente, efetiva e razoavelmente rápida, e o ouvinte quer recuperar uma mensagem informativa e clara de forma rápida e eficiente. A partir dessas necessidades e restrições, Slobin concebe motivações fundamentais que definem a estrutura de qualquer língua humana totalmente desenvolvida: (1) ser clara, (2) ser processável, (3) ser rápida e fácil/simples, (4) ser expressiva.

O autor explica que, como essas pressões são inerentemente competitivas, as línguas estão constantemente mudando. Essas motivações, nesse sentido, ditam o direcionamento da mudança diacrônica nas línguas, mas, como estão sempre em competição,

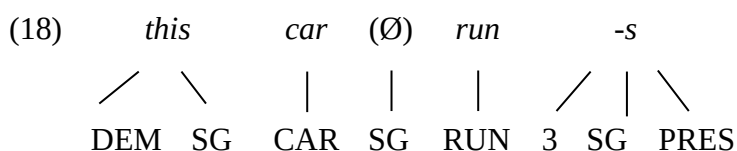
a cada momento da língua uma delas pode estar com uma maior vantagem em relação às outras.

A primeira máxima – *ser clara* – é descrita pelo autor como a tendência de se manter um mapeamento de um-para-um entre semântica e sintaxe, de se buscar a transparência semântica. A segunda máxima – *ser processável* – dita que a direção da mudança linguística é sempre no sentido de reforçar a processabilidade, de adaptar a língua às restrições de tempo real inerentes à programação e à interpretação de enunciados. A terceira máxima – *ser rápida e fácil/simples* – está relacionada com a tendência da língua de imprimir velocidade e compactar a informação, ou seja, apagar e contrair formas, combinar (juntar) formas subjacentes na estrutura superficial de maneira a aliar necessidades comunicativas de se transmitir uma grande quantidade de informação ao ouvinte a questões de restrições de memória, por exemplo. A quarta máxima – *ser expressiva* – envolve a expressividade semântica e a retórica. A expressividade semântica está relacionada com o fato de a língua ter meios para codificar ao menos o conjunto universal básico de conceitos e relações salientes. A expressividade retórica envolve fornecer ao falante meios de comandar a atenção do ouvinte, levando em consideração o conhecimento e as expectativas do ouvinte, meios para surpreender, impressionar, jogar, brincar etc.

A primeira e a terceira motivações são as que correspondem, respectivamente, às noções de *iconicidade* e *economia*, apontadas por vários autores como as principais motivações em conflito atuantes na determinação da estrutura linguística. O mapeamento de um-para-um previsto na primeira dessas motivações, conforme mostrado na seção 2.1, é o que define a iconicidade por isomorfismo. Nesse sentido, a motivação *ser clara* pode ser interpretada como uma pressão para que a estrutura da língua seja icônica, isomórfica. A necessidade de imprimir velocidade e compactar a informação descrita na terceira motivação, por sua vez, corresponde à pressão pela economia linguística.

Embora, em princípio, cada uma das quatro motivações esteja em competição com todas as outras, é possível relacioná-las em dois grupos: as motivações *ser rápida e simples* e *ser expressiva* são, segundo Slobin (1977), as que desencadeiam a mudança linguística, enquanto as duas primeiras, *ser clara* e *ser processável*, restringem as direções da mudança em uma dada língua a um conjunto limitado de possibilidades. Segundo ele, uma língua ganha em compactação ou expressividade da comunicação quando perde em facilidade de processamento ou em transparência semântica da mensagem.

Croft (2003) também considera que motivações em competição são responsáveis pela estrutura dos enunciados. O autor distingue duas forças – a economia e o isomorfismo. Para ele, essas duas forças competem para produzir a gama de correspondências entre forma e significado atestadas e não-atestadas (ou raras) nas línguas do mundo. O autor define, com base em Haiman (1980), o isomorfismo sintagmático, isto é, a correspondência de formas e significados encontrada na combinação de palavras e flexões na sentença, como o subtipo de iconicidade (conforme discutido em 2.1) que é relevante para a codificação de conceitos individuais na gramática. Croft ilustra a atuação dessas duas forças na determinação da estrutura linguística por meio da observação da correspondência forma-significado na sentença do inglês *this car runs* (esse carro corre), mostrada em (18).



(CROFT, 2003, p. 103).

Como se pode ver no esquema, há formas que codificam um significado (*car* e *run*), dois (*this*) ou três (*-s*) significados e há um significado sem uma forma correspondente (Ø para indicar singular (SG)). Para o autor, essas diferentes correlações entre formas e

significados na organização linear do enunciado são resultado da competição entre economia e isomorfismo sintagmático: (a) *car* e *run* apresentam uma relação de um-para-um entre forma e significado porque o isomorfismo é a força predominante; (b) o significado “singular” em *car* (*car-Ø* vs *car-s*) não é marcado formalmente (morfema zero) devido à predominância da motivação econômica (que, nesse caso, determina a não-expressão de um significado) sobre a icônica; (c) a forma *-s* em *run-s* acumula três significados discretos (3ª pessoa, singular e tempo presente) em uma única forma devido à predominância da motivação econômica²²; (d) a forma *this* combina dois significados (demonstrativo proximal e singular) em uma mesma forma devido à predominância da motivação econômica. Ou seja, para o autor, o modo de correlação forma-significado final de uma sentença é resultado da disputa dessas duas forças numa mesma dimensão linguística.

Para o autor, um fato que prova que a codificação formal é resultado dessas motivações em competição seria a raridade ou ausência de padrões formais, nas línguas do mundo, que não sejam explicados por nenhuma das duas motivações, como *formas sem significados* e *significados expressos por duas ou mais formas*. Esses padrões formais, segundo ele, por não serem motivados nem iconicamente, nem economicamente, são extremamente raros e, quando ocorrem, são historicamente instáveis. Os mapeamentos sintagmáticos forma-significado que são motivados por economia ou por iconicidade (isomorfismo), por outro lado, seriam amplamente encontrados nas línguas do mundo.

O quadro 2 resume a gama completa de possibilidades lógicas para expressar as correspondências entre forma e significado mediante as motivações econômica e isomórfica.

²² Embora Croft (2003) inclua o significado de presente na forma *-s*, consideramos que outra análise possível seria a de que essa marca não codificaria presente, mas apenas terceira pessoa e singular. De todo modo, essa forma codificaria mais de um significado por economia.

Quadro 2: Possíveis correspondências forma-significado e suas motivações (CROFT, 2003, p.104)

Forma(s)	Significado(s)	Iconico	Econômico	
1	1	sim	não	Estrutura icônica clássica
0	1	não	sim	Expressão zero de categoria
0	0	não	sim	Ausência de categoria
1	>1	não	sim	Acumulação/supletismo
1	0	não	não	Morfemas vazios

Na mesma direção de Slobin (1977) e Croft (2003), Haiman (1983) também considera que motivações competem na determinação da estrutura linguística; no caso, o autor considera as motivações de iconicidade e economia. O autor, inclusive, faz uma abordagem mais detalhada, especificando diferentes formas pelas quais essas motivações interagem entre si. Ele aponta duas principais possibilidades de interação: (i) a convivência harmoniosa entre as duas motivações; (ii) a competição das duas motivações num mesmo meio, tendo como resultado final a realização de apenas uma delas.²³

Haiman (1983) ilustra a convivência harmoniosa entre iconicidade e economia explicando o processo conhecido como *redução de coordenação*. Para o autor, esse processo envolve o apagamento de um ou mais constituintes linguísticos, o que seria motivado por economia, e, ao mesmo tempo, o processo envolve um reagrupamento de constituintes, o que seria motivado por iconicidade. Alguns dos exemplos discutidos pelo autor são mostrados de (19) a (21).

- (19) a. Red ribbons and white ribbons (Fitas vermelhas e fitas brancas).
b. Red and white ribbons (Fitas vermelhas e brancas).

- (20) a. We can do it quickly and we can do it well (Podemos fazer isso rapidamente e podemos fazer isso bem).
b. We can do it quickly and well (Podemos fazer isso rapidamente e bem).

²³ Há ainda a possibilidade de conflito entre diferentes motivações icônicas, sem envolvimento da motivação econômica.

- (21) a. John Smith and Mary Smith are employees of this company (John Smith e Mary Smith são empregados desta empresa).
 b. John and Mary Smith are employees of this company (John e Mary Smith são empregados desta empresa).

(HAIMAN, 1983, p. 808).

Em todos os casos acima, da construção (a) para a (b), há o apagamento de um ou mais constituintes repetidos e, assim, já conhecidos no contexto. De acordo com Haiman (1983), em termos gerais, a motivação para todos os casos de apagamento de constituintes já identificados no contexto, incluindo os casos de redução de coordenação, é a economia: não se repete o que é conhecido. Porém, destaca o autor, a redução de coordenação não envolve apenas o apagamento, mas também o reagrupamento de constituinte, o qual acompanha, iconicamente, uma alteração de significado envolvida no processo de redução de coordenação: nesse processo, há uma redução de distância entre constituintes linguísticos que acompanha iconicamente uma redução de distância conceitual entre eles.

Segundo o autor, em (19b), em que as palavras *vermelha* e *branca* estão mais próximas, a interpretação seria a de que as duas cores podem ocorrer na mesma fita – interpretação que não ocorreria ou não seria a mais provável em (19a). Similarmente, em (20a), seria possível fazer uma certa coisa rapidamente sob um conjunto de circunstâncias e fazer a mesma coisa bem (de forma bem feita) sob outro conjunto de circunstâncias; mas, em (20b), ao agrupar *rapidamente* e *bem*, a interpretação (mais provável) seria que tal coisa pode ser feita rapidamente e bem sob um mesmo conjunto de circunstâncias. Da mesma forma, é mais provável que John e Mary sejam parentes em (21b), onde os termos *John* e *Mary* estão agrupados, do que em (21a), onde não estão.

Desse modo, pode-se ver que uma mesma estrutura linguística (no caso, decorrente de um processo de redução de coordenação) é resultado da atuação, simultânea, de economia e iconicidade. É o que Haiman (1983) considera como a convivência harmoniosa entre as duas motivações. Já, em outras situações, o autor avalia que essas motivações atuam

primeira oração (*mimana*) é subordinada à segunda (*de*). De acordo com o autor, a construção em (23) é ambígua, significando tanto “Ele deu (a ela) e ela comeu” (que é a interpretação que o autor inclui na tradução do exemplo para o inglês, e que mantivemos em nossa tradução para o português), quanto “Ele deu (a ela) e ele (próprio) comeu”. Em (23), mesmo quando a interpretação é a segunda, em que os sujeitos dos dois verbos são idênticos, não ocorre apagamento, no primeiro verbo (*mimana*), do morfema que indica sujeito (no caso, *ma*). Ou seja, nessa interpretação, por economia, deveria ocorrer o apagamento desse morfema, já que a informação que ele veicula é expressa também no verbo seguinte – assim como ocorre em (22b), pelo processo de redução de coordenação. Porém, a motivação econômica não se impõe aí, não ocorrendo o processo de redução de subordinação.

Segundo Haiman (1983), a ausência de um processo de redução de subordinação (no caso, o não apagamento do morfema *ma*) é motivada por iconicidade. Conforme explica o autor, a principal característica do fenômeno da subordinação, em hua, é que orações subordinadas são conceitualmente independentes da oração seguinte (em hua, inclusive, a oração subordinada precede a oração principal). Essa independência conceitual das orações subordinadas está associada às seguintes propriedades dessas orações (em oposição às coordenadas):²⁴

- (i) elas não concordam com a oração seguinte em tempo (enquanto coordenadas concordam);
- (ii) elas não concordam com a oração seguinte em modo (ao passo que coordenadas concordam);
- (iii) elas não precisam apresentar iconicidade temporal (isto é, a ordem das orações não precisa corresponder à ordem dos eventos – já as coordenadas têm de apresentar essa iconicidade);

²⁴ Para Haiman, essas distinções morfológicas que diferenciam as coordenadas das subordinadas na língua hua, muito provavelmente, devem corresponder a definições universais de coordenação e subordinação.

- (iv) elas são sempre pressupostas e, assim, a negação da oração seguinte não afeta a validade da oração subordinada (enquanto as coordenadas não são pressupostas, de modo que a negação da oração seguinte pode também implicar a negação da oração coordenada anterior).

Conforme argumenta Haiman (1983), a oração subordinada não sofre redução para que entre ela e a oração seguinte haja uma independência formal icônica da independência conceitual entre essas orações. Como mostrado nos exemplos (19), (20) e (21) acima, a redução de coordenação promove uma aproximação formal entre certos constituintes linguísticos, e essa aproximação corresponde a uma proximidade conceitual entre os significados expressos por esses constituintes. Já no caso da subordinação, há uma independência conceitual entre a oração subordinada e a seguinte. Se ocorresse uma redução na subordinada, haveria uma integração formal entre ela e a seguinte, o que seria contrário à independência conceitual entre essas orações. Assim, o impedimento de redução na subordinada serve como uma expressão icônica da independência conceitual entre ela e a oração seguinte.

Em síntese, como Haiman (1983) procura mostrar, a redução de coordenação ilustra uma situação em que economia e iconicidade atuam de forma harmônica na determinação da estrutura linguística, enquanto o processo de subordinação exemplifica uma situação em que as duas motivações competem entre si, vencendo uma delas, no caso, a iconicidade.

Nesta seção, enfim, mostramos que, de acordo com Slobin (1977), Croft (2003) e Haiman (1983), grande parte da estrutura morfossintática das línguas é o resultado da interação entre iconicidade e economia. Segundo esses autores, essas motivações determinam a distribuição de significados entre as unidades morfossintáticas das sentenças, sendo, assim, responsáveis por definir quanto isomorfismo ocorrerá em uma sentença.

Considerando que o grau de transparência de uma língua consiste justamente na

proporção de relações isomórficas que ocorrem nas sentenças da língua, podemos concluir que o grau de transparência de uma língua e, em particular, o grau de transparência do sistema de referência aqui em estudo serão diretamente determinados pela atuação da iconicidade e da economia. Na seção seguinte, assumimos essa concepção no tratamento das noções de *teto* e de *piso de transparência linguística*.

2.3 Teto e piso de transparência

No presente trabalho, procuramos refletir sobre a questão mais geral da variação diacrônica da transparência linguística, averiguando especificamente a diacronia do grau de transparência do sistema de referência do PB (conforme delimitado anteriormente). Com base na discussão teórica exposta acima, parece ser possível concluir que a variação diacrônica da transparência, seja em termos gerais, seja no domínio específico aqui investigado, ocorreria dentro de um limite máximo e de um limite mínimo de transparência, por conta da competição entre iconicidade e economia.

Os autores que tratam de iconicidade e economia consideram que essas motivações competem entre si e atuam constantemente sobre as línguas, sendo motivações igualmente relevantes, ora vencendo uma, ora vencendo a outra, de modo a promoverem um certo equilíbrio entre estruturas icônicas (estruturas isomórficas) e econômicas na língua. Conforme defende Slobin (1977), a língua necessita desse equilíbrio para ser comunicativamente viável. Assim, com base no que diz o autor, pode-se concluir que, se uma língua, por atuação mais forte da iconicidade (ou de algum outro fator), fosse adquirindo cada vez mais relações isomórficas, em algum momento essa progressiva adoção de relações isomórficas provavelmente seria interrompida, por atuação da economia e vice-versa. Ou seja, a atuação constante da economia impediria que uma língua adquirisse, ininterruptamente,

cada vez mais relações isomórficas – haveria um nível máximo tolerável de isomorfismo, ou o que se poderia chamar de *teto de isomorfismo*. Similarmente, a atuação constante da iconicidade impediria que uma língua adquirisse ininterruptamente cada vez mais relações não-isomórficas – haveria um nível mínimo tolerável de isomorfismo, ou um *piso de isomorfismo*.

Como explicamos acima, o grau de transparência de uma língua corresponde ao grau com que essa língua manifesta relações isomórficas. Assim, consideramos que as línguas poderiam ter seu grau de transparência aumentado, mas até, no máximo, um determinado limite, que seria, então, um *teto de transparência*. Em sentido oposto, o grau de transparência das línguas poderia diminuir cada vez mais, porém até um determinado limite, que seria um *piso de transparência* (ou um *teto de opacidade*). De acordo com nossa visão, o grau de transparência das línguas variaria entre esses limites máximo e mínimo e, assim, a variação diacrônica do grau de transparência do sistema de referência aqui em estudo também ocorreria dentro desses limites.

Alguns autores, ao abordar questões relacionadas a mudança linguística, tecem considerações que apontam, mesmo que implicitamente, para as noções de teto de transparência e piso de transparência (ou teto de opacidade). Slobin (1977, 1980) argumenta na direção do reconhecimento do que se pode ver como um teto de opacidade, embora o autor não trabalhe propriamente com as noções de transparência e opacidade aqui adotadas, associando-as, na verdade, às noções de *análise* e *síntese*, respectivamente.²⁵ Quando as línguas se tornam, nos termos do autor, opacas demais, meios de expressão mais novos e mais transparentes se desenvolvem. Assim, no desenvolvimento histórico de uma língua, a transparência seria constantemente restabelecida, embora a opacidade sempre retornaria até

²⁵ Apesar de não haver equivalência exata entre as noções tratadas pelo autor e as aqui assumidas, é possível identificar correspondências importantes entre as noções por ele utilizadas e os conceitos de *transparência* e *opacidade* adotados no presente trabalho, sendo possível depreender as conclusões aqui sintetizadas.

certo ponto. Segundo o autor, esse restabelecimento da transparência pode ser visto como tendência universal, embora as leis que favoreçam esse direcionamento estejam sempre em competição com outras que indiretamente direcionariam a língua para um mapeamento mais opaco. O autor afirma que as leis em competição, que não se anulam entre si, fazem com que as línguas flutuem constantemente entre os polos da transparência e da opacidade.

Slobin (1977) considera que as línguas podem apresentar um ciclo histórico, indo do polo do que ele trata como transparência ao polo do que vê como opacidade, retornando posteriormente ao polo da transparência. Esse ciclo histórico completo entre um polo e outro seria, porém, difícil de ser depreendido, já que levaria muitas gerações para se delinear em línguas estabelecidas. O autor, com base em Hodge (1970), cita o egípcio antigo, sobre o qual haveria dados de mais ou menos 4.000 anos, como um exemplo de uma língua estabelecida que teria completado esse ciclo duas vezes, ressaltando que, na emergência de línguas crioulas (a partir de pidgins), esse processo poderia ser percebido em uma única geração.

Leufkens (2015), no contexto específico da GDF, também admite uma concepção semelhante à que assumimos aqui, falando propriamente em *teto de opacidade*. Para ela, a opacidade seria a direção esperada da mudança linguística, e a mudança em direção contrária (em direção à transparência) ocorreria em duas situações: (i) no caso de contato linguístico (em que se tem uma necessidade de inteligibilidade especialmente alta) e (ii) no caso de se atingir um certo limite máximo de opacidade. Esse teto de opacidade é explicado pela autora como um limite na aprendizagem, que faz com que determinado traço de opacidade se perca quando cruza essa linha.

Também Lightfoot (1979) considera a noção de *teto de opacidade* – o autor concebe essa noção no quadro da Gramática Gerativa, de modo que sua noção de opacidade não corresponde propriamente à noção aqui em pauta, mas a ideia geral do autor aponta na mesma direção do que consideramos aqui. Para ele, as línguas apresentam um teto de

opacidade, um grau de tolerância da língua para a opacidade, e, quando esse grau é alcançado, ocorre uma reestruturação na língua, tornando-a mais transparente. Segundo o autor, uma relação transparente seria caracterizada pela biunivocidade/proximidade entre a estrutura subjacente e sua respectiva estrutura superficial, enquanto a opacidade seria caracterizada pela distância maior ou não-biunivocidade entre essas duas estruturas. Nesse sentido, a ideia parece ser a de que, quando a relação entre essas duas estruturas se torna muito complexa/distante e ultrapassa o grau de tolerabilidade de opacidade, ocorre um processo de simplificação dessa relação, tornando-a mais clara, mais simples, menos abstrata, enfim, mais transparente nos termos do autor.

Portanto, Slobin (1977, 1980), Leufkens (2015) e Lightfoot (1979), ao mencionarem noções relacionadas a um certo limite de opacidade, apontam na direção da concepção que concluímos acima: a de que o grau de transparência das línguas e, da mesma maneira, o grau de transparência do sistema de referência aqui investigado devem oscilar diacronicamente entre um teto e um piso de transparência, concepção que está baseada no reconhecimento de que as línguas estão sempre submetidas a pressões igualmente relevantes por parte da iconicidade e por parte da economia. Essa conclusão teórica – que parece encontrar respaldo consistente nos trabalhos sintetizados neste capítulo e nas suposições que pudemos depreender a partir deles – seria uma das principais contribuições que nossa pesquisa viria a oferecer a respeito da questão da diacronia da transparência linguística.

Como definido anteriormente, o objetivo deste trabalho é analisar três mudanças diacrônicas ocorridas no PB (aumento da frequência de expressão pronominal do argumento-sujeito e das frequências de uso dos pronomes *você*, em detrimento de *tu*, e *a gente*, em oposição a *nós*), verificando como a atuação conjunta dessas três mudanças interfere diacronicamente no grau de transparência do sistema de referência aqui em estudo. Com base na concepção teórica exposta no presente capítulo – que formulamos nas etapas mais iniciais

de nossa pesquisa de doutorado – pudemos assumir a hipótese de que o grau de transparência desse sistema oscilaria entre certos limites máximo e mínimo, girando em torno de um certo eixo de transparência. Essa constitui, então, a hipótese que norteou o desenvolvimento de nossa pesquisa.

CAPÍTULO III

Mudanças diacrônicas no português brasileiro interferentes no grau de transparência linguística

Conforme estamos assumindo neste trabalho, a variação diacrônica do grau de transparência do sistema de referência por expressão pronominal e desinencial do argumento-sujeito de 1ª e 2ª pessoas é determinada por três mudanças diacrônicas verificadas no PB:

- (i) aumento da frequência de expressão pronominal do argumento-sujeito em oposição à não-expressão pronominal;
- (ii) aumento da frequência de uso do pronome *você* em detrimento do pronome *tu*;
- (iii) aumento da frequência de uso da forma pronominal *a gente* em detrimento do pronome *nós*.

Neste capítulo, sintetizamos alguns trabalhos que atestam a ocorrência dessas mudanças.

3.1 Aumento da frequência de expressão pronominal do argumento-sujeito

O fenômeno do aumento da frequência de expressão pronominal do argumento-sujeito constitui um tema muito abordado nos estudos linguísticos, principalmente em trabalhos associados à Gramática Gerativa (como no caso de trabalhos no campo da Sociolinguística Paramétrica). Na Gramática Gerativa, particularmente na versão da Teoria de Princípios e Parâmetros, considera-se que a chamada *Gramática Universal* (gramática geral, subjacente a todas as línguas naturais), conforme explica Raposo (1992), contém *princípios* rígidos, invariáveis e universais, que a gramática de qualquer língua particular terá de incorporar, assim como princípios abertos, denominados *parâmetros*, que são uma espécie

de “comutadores linguísticos” cujo valor final e definitivo apenas é atingido durante o processo de aquisição da linguagem, ao final do qual ocorre a sua fixação numa de duas posições possíveis com base na informação obtida a partir do meio linguístico.

De acordo com o autor, entre os princípios universais, encontra-se o Princípio de Projeção Estendido, segundo o qual todas as línguas humanas reservam, na estrutura subjacente da frase, uma posição sintática para a expressão do sujeito,²⁶ princípio ao qual se associa o chamado *parâmetro do sujeito nulo*. Como propriedade universal, o Princípio de Projeção Estendido não obriga que a posição de sujeito seja foneticamente preenchida, podendo ou não ficar vazia, em línguas como o italiano e o espanhol, ou ser obrigatoriamente preenchida, em línguas como o inglês e o francês. A opção entre realização fonética obrigatória *versus* realização facultativa constitui, então, uma variação paramétrica interlinguística associada ao Princípio de Projeção Estendido (RAPOSO, 1992).²⁷

A mudança diacrônica aqui em pauta é tratada, em trabalhos associados ao gerativismo (especialmente na Sociolinguística Paramétrica), em termos de diminuição do uso de sujeito nulo, ou aumento do uso de sujeito pronominal pleno. Como explica Lobo (2013), o sujeito nulo tem as mesmas propriedades que o pronome pessoal, com os traços de pessoa e número correspondentes à flexão do verbo, sendo representado pela notação *pro*. Assim, as línguas que admitem sujeito nulo são chamadas de *línguas pro-drop*, e as línguas de preenchimento obrigatório de sujeito de *línguas não pro-drop*. Dessa forma, a mudança aqui

²⁶ O Princípio de Projeção Estendido é uma expansão do antes denominado simplesmente Princípio da Projeção, segundo o qual todo núcleo lexical deve ter suas posições argumentais preservadas em qualquer dos níveis de representação da sentença (Estrutura Profunda, Estrutura Superficial e Forma Lógica), não sendo possível aumentar nem diminuir o número de argumentos ou de posições argumentais de um nível para o outro. Assim, o Princípio de Projeção Estendido diz respeito à exigência sintática de que todo predicado precisa de um sujeito acerca do qual ele predica.

²⁷ Para uma abordagem mais completa do Princípio de Projeção Estendido, bem como para reformulações desse princípio, ver Chomsky (1995), Rizzi (2006).

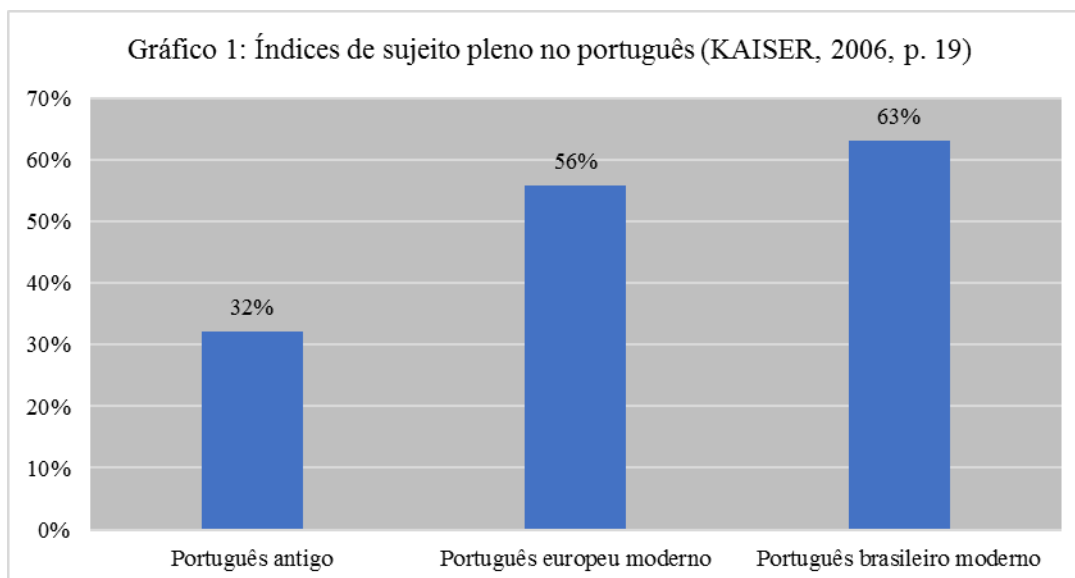
em estudo é também referida, por vezes, como envolvendo a passagem do PB de uma língua *pro-drop* para uma língua não *pro-drop*.²⁸

Embora o presente trabalho seja desenvolvido num quadro teórico-metodológico que não está vinculado ao gerativismo, entendemos que seja possível levar em conta os trabalhos relacionados à abordagem gerativista para nossa averiguação do fenômeno – em nossos termos – da mudança diacrônica do aumento da frequência de expressão pronominal do argumento-sujeito. A nosso ver, consideramos possível nos basearmos em tais trabalhos, focalizando as constatações empíricas neles descritas, sem que as particularidades teóricas desses trabalhos possam comprometer o aproveitamento de resultados que deles fazemos.

No âmbito do campo de investigação ligado ao gerativismo, então, Kaiser (2006) considera que se verifica, no PB, uma tendência cada vez mais frequente de se utilizar pronomes plenos na posição de sujeito de orações finitas. O autor afirma que essa mudança é confirmada por diversos trabalhos, tanto de cunho diacrônico, citando os estudos de Duarte (1993, 1998, 2000) – segundo ele, já clássicos – e de Tarallo (1993b), quanto de cunho sincrônico, fazendo referência a Meyer-Hermann (1998), Lira (1996), Barme (1998, 2001) e Kato e Negrão (2000).

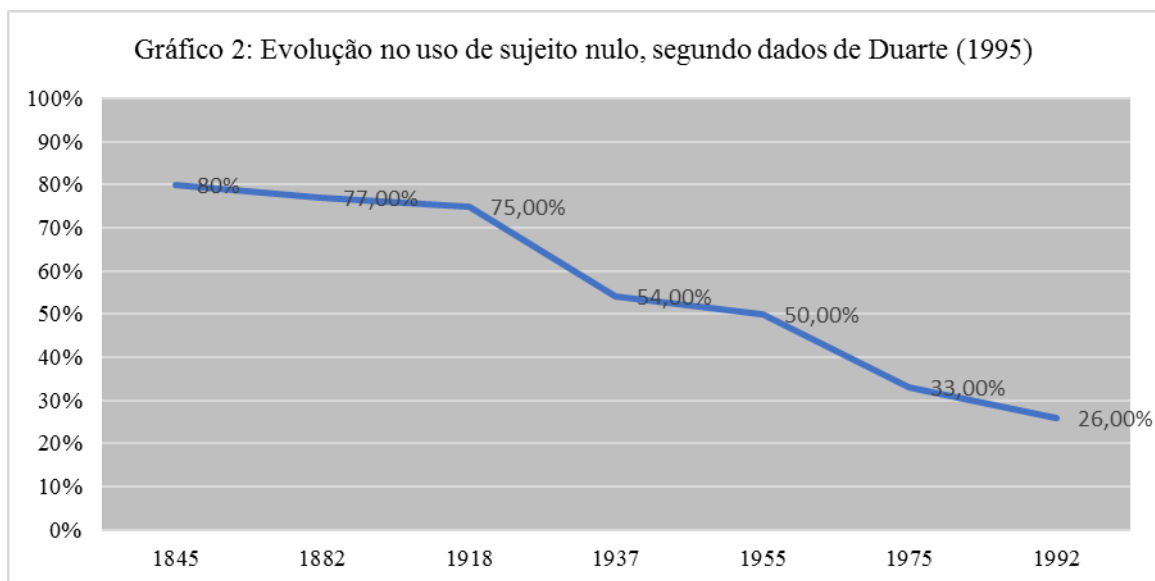
Como evidência da mudança em questão, o autor levanta dados do português antigo (século XIII), do português europeu moderno (1993) e do PB moderno (2000), extraídos de traduções portuguesas da Bíblia. No gráfico 1 a seguir, sintetizamos os dados apresentados pelo autor.

²⁸ Outras análises mais recentes referem-se ainda ao PB como língua de sujeito parcialmente *pro-drop* (ver HOLMBERG; NAYADU; SHEEHAN, 2009).



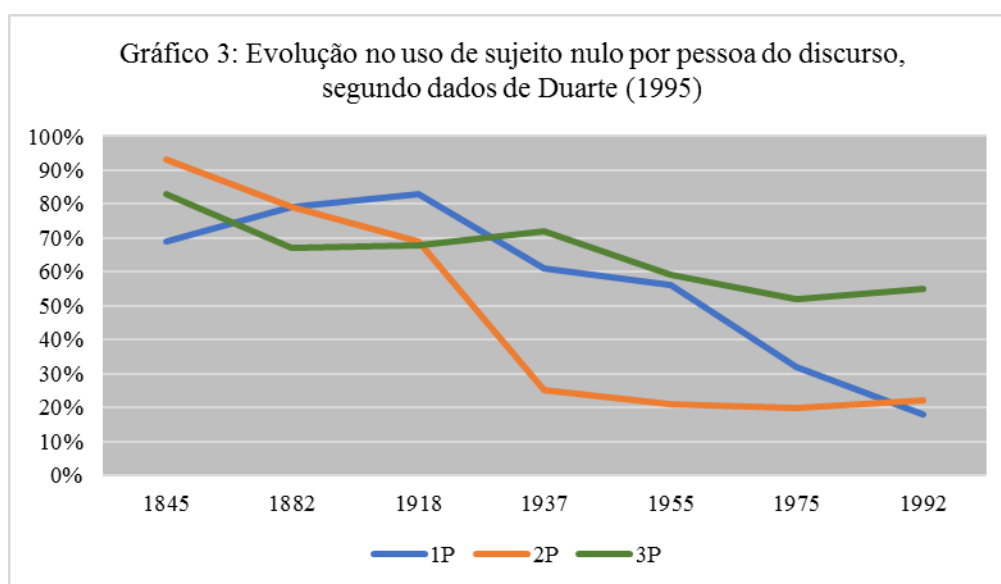
Como afirma o próprio autor, é possível observar um aumento bastante significativo no uso de sujeito pronominal pleno, tanto no português europeu, quanto no PB, em comparação com o português antigo.

Duarte (1993, 1995) também atesta o fenômeno aqui em questão. Em suas palavras, verifica-se no PB uma “crescente preferência pelo sujeito pronominal pleno” (DUARTE, 1993, p. 107). A autora (DUARTE, 1993, 1995) apresenta dados diacrônicos do PB que evidenciam essa mudança, extraídos de peças de teatro dos séculos XIX e XX. Ela levanta dados referentes aos pronomes de 1^a, 2^a e 3^a pessoas, no singular e no plural, considerando as formas *nós* e *a gente* para 1^a pessoa do plural, *tu* e *você* para 2^a pessoa do singular e as formas *vós* e *vocês* para 2^a pessoa do plural. O gráfico 2 abaixo mostra a evolução no uso do sujeito nulo no período analisado, de acordo com os dados oferecidos pela autora.



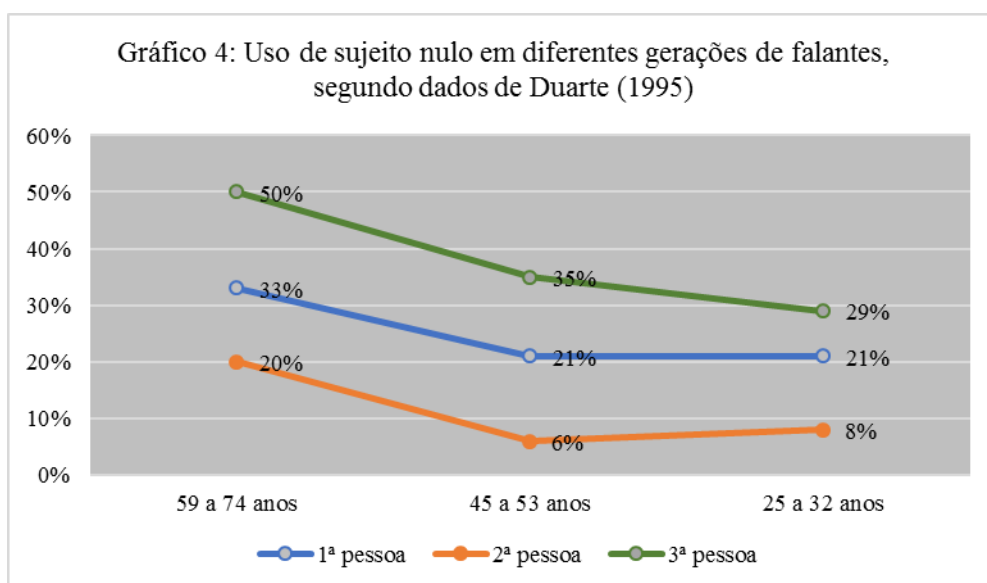
Como se pode notar, no período considerado há uma queda contínua no uso de sujeito nulo, a qual é mais significativa a partir de 1918, acentuando-se nos dois últimos períodos, de modo que praticamente se inverte a frequência de sujeito nulo e pleno em 1992.

Além dos dados expostos no gráfico 2, que reúne 1ª, 2ª e 3ª pessoas, no singular e no plural, Duarte (1993, 1995) também apresenta dados separados de cada uma das pessoas, os quais, ao menos em termos gerais, também indicam o mesmo processo de diminuição no uso de sujeito nulo. Os dados fornecidos pela autora são apresentados abaixo no gráfico 3 (os dados de cada pessoa incluem conjuntamente singular e plural).



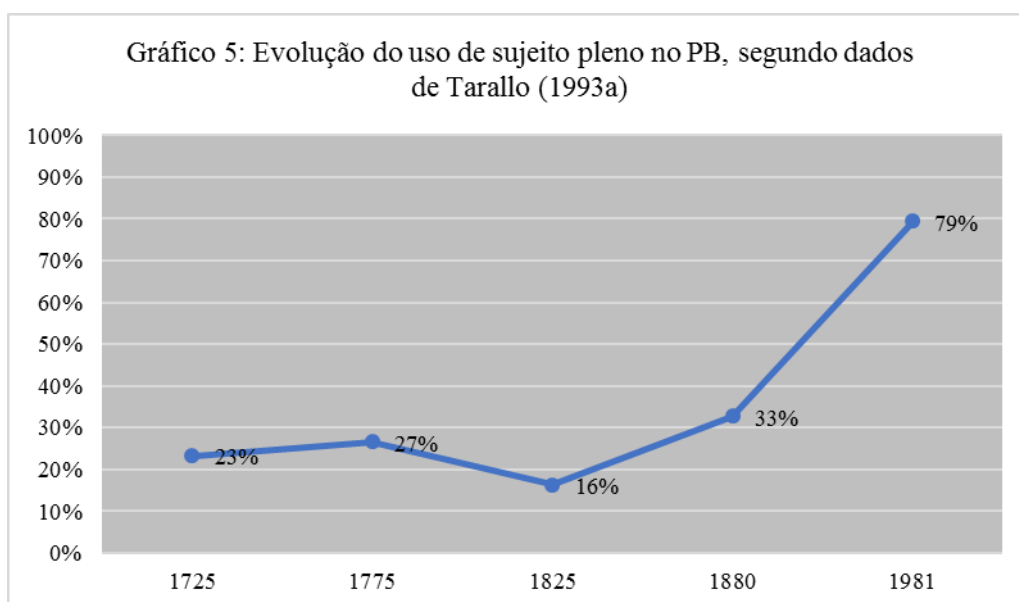
A queda mais acentuada no uso de pronome nulo ocorre na 2ª pessoa, em que se verifica a mudança de um percentual de 93% em 1845 para 22% em 1992, envolvendo um percurso quase exclusivo de queda. A segunda queda mais visível ocorre com a 1ª pessoa, com a qual se nota uma mudança de 69% de uso de sujeito nulo em 1845 para 16% em 1992. Nesse caso, há um ligeiro aumento na incidência de sujeito nulo entre 1845 e 1918, mas, em seguida, verifica-se queda constante até o último período. Por fim, a menor variação se verifica em torno da 3ª pessoa, com a qual se identifica um percentual de 83% de sujeito nulo em 1845, que passa para 55% em 1992 (percurso que inclui um ligeiro aumento no uso de sujeito nulo entre 1882 e 1937). Assim, em termos gerais, pode-se notar um claro predomínio da tendência de queda no uso de sujeito nulo com as três pessoas no período em foco.

Duarte (1995) também sustenta a ocorrência da diminuição diacrônica no uso de sujeito nulo a partir da perspectiva do tempo aparente, segundo a qual a variação no uso linguístico de falantes de diferentes gerações pode representar uma variação diacrônica em curso. A autora levanta dados de três grupos de informantes, cada um de uma faixa etária: 59 a 74 anos, 45 a 53 anos e 25 a 32 anos. Esses dados aparecem no gráfico 4. Nesse caso, ela apresenta apenas dados separados por pessoa (também aqui, em cada pessoa, estão reunidos singular e plural).



Embora se possa observar um ligeiro aumento na incidência de sujeito nulo com a 2ª pessoa da faixa etária de 45-53 anos para a de 25-32 anos, os valores mostrados no gráfico indicam uma clara predominância da tendência de queda no uso de sujeito nulo em todas as pessoas da faixa etária mais velha para a mais nova.

Na mesma direção dos trabalhos acima, Tarallo (1993a) também considera a ocorrência do fenômeno da diminuição do uso de sujeito nulo no PB – que trata em termos de aumento do sujeito pleno. O autor apresenta dados diacrônicos que evidenciam essa mudança, num contexto em que defende que o PB, na passagem do século XIX para o século XX, teria sofrido um conjunto de mudanças drásticas, caracterizando a língua como um sistema em fase de transição de língua *pro-drop* para não *pro-drop* – tais mudanças teriam feito emergir um novo sistema gramatical, diferenciando o PB do português europeu. Os dados apresentados pelo autor são sintetizados no gráfico abaixo.



Como se pode ver, a despeito de uma queda no índice de sujeito pleno entre 1775 e 1825, os dados de Tarallo (1993a) indicam o predomínio do aumento da incidência de sujeito pleno, com um aumento muito expressivo no período de 1880 a 1981.

As pesquisas relatadas acima, enfim, atestam a ocorrência, no PB, do fenômeno do aumento diacrônico do uso de sujeito pronominal pleno (ou diminuição do índice de sujeito nulo). Além dessas pesquisas, diversos outros trabalhos também assumem tal mudança, como os trabalhos mencionados por Kaiser (2006), acima listados, dentre outros, tais como Tarallo (1985), Duarte (2008), Kato (2012). Tendo em vista, então, a ocorrência dessa mudança, no presente trabalho procuraremos averiguar seu impacto, juntamente com as outras mudanças também consideradas, sobre a transparência/opacidade do PB.

Cabe aqui salientar que, embora pareça não haver maiores dúvidas sobre a ocorrência de tal mudança no PB, não é muito vasto o conjunto de trabalhos que fornecem dados diacrônicos sobre esse processo. Os principais trabalhos que identificamos com apresentação de dados diacrônicos são os que relatamos acima. Várias outras pesquisas que também sustentam tal mudança são de natureza sincrônica. Assim, parece-nos válido considerar que o próprio levantamento e a apresentação de dados diacrônicos dessa mudança feitos no presente trabalho constituiriam uma contribuição significativa do trabalho.

Convém, por fim, ainda considerar aqui uma questão muito relevante em relação a esse fenômeno do aumento nos índices de uso de sujeito pronominal pleno no PB. Trata-se da relação desse processo com o fenômeno diacrônico da simplificação dos paradigmas de flexão verbal. Alguns autores entendem que a simplificação desses paradigmas seria a causa do aumento no índice de uso de sujeito pleno, enquanto outro grupo de autores considera que não haveria essa relação de causa entre os dois fenômenos.

Duarte (1993), por exemplo, estaria incluída no primeiro grupo. Segundo ela, há no PB, diacronicamente, uma crescente simplificação nos paradigmas flexionais verbais, conforme mostrado no quadro abaixo, esboçado pela autora.

Quadro 3: Evolução nos paradigmas flexionais no português (DUARTE, 1993, p.109)

Pessoa	Núm.	Paradigma 1	Paradigma 2	Paradigma 3
1 ^a	sing.	cant-o	cant-o	cant-o
2 ^a direta	sing.	canta-s	_____	_____
2 ^a indireta	sing.	canta-0	canta-0	canta-0
3 ^a	sing.	canta-0	canta-0	canta-0
1 ^a	plur.	canta-mos	canta-mos	canta-0
2 ^a direta	plur.	canta-is	_____	_____
2 ^a indireta	plur.	canta-m	canta-m	canta-m
3 ^a	plur.	canta-m	canta-m	canta-m

Do paradigma 1 ao 3, percebe-se uma progressiva simplificação flexional. Segundo a autora, o PB evoluiu de um sistema com seis formas distintivas (paradigma 1) para um paradigma que apresenta quatro formas, em função da perda da 2^a pessoa direta *tu*, substituída pelo pronome *você* (paradigma 2) – que é uma das mudanças consideradas em nosso trabalho. Este paradigma, segundo ela, coexistiria atualmente com um terceiro, em que se veem apenas três formas, em consequência da perda do pronome de primeira pessoa do plural *nós*, substituído progressivamente pela expressão *a gente*, que se combina com formas verbais de terceira pessoa do singular (paradigma 3) – outra mudança também abordada neste trabalho.

Duarte (1993) entende que a simplificação dos paradigmas seria responsável pela diminuição no uso de sujeito nulo. Para ela, até por volta do início do século XX, estaria em funcionamento o paradigma 1, o qual seria responsável pela preferência por sujeito nulo. A partir de então, o pronome *tu* seria substituído por *você*, entrando em vigor o paradigma 2, o que levaria à diminuição no uso de sujeito nulo. Mais ao final do século XX, passaria a ocorrer a progressiva substituição do pronome *nós* pela forma *a gente*, o que levaria o paradigma 3 a coexistir com o paradigma 2, mantendo-se, assim, a tendência de diminuição no uso de sujeito nulo.

Essa relação de causa entre a simplificação dos paradigmas de flexão verbal e a diminuição do uso de sujeito nulo é também admitida em outros trabalhos, como Roberts e Kato (1993), Duarte (2000), De Oliveira (2000), Rohrbacher (1994 *apud* MODESTO, 2000),

entre outros. Porém há autores que não identificam essa relação de causa entre os dois fenômenos, como é o caso de Negrão e Viotti (2000), Modesto (2000), Negrão (1990 *apud* MODESTO, 2000), entre outros.

Na visão destes últimos autores, o reconhecimento de uma relação de causa entre simplificação flexional e redução no uso de sujeito nulo estaria assentado na ideia de que línguas de morfologia rica favoreceriam o uso de sujeito nulo, enquanto uma morfologia fraca privilegiaria o sujeito pleno. No entanto, conforme explica Kaiser (2009), a existência de uma morfologia verbal rica não necessariamente implica que uma língua exibe a propriedade de sujeito nulo. Segundo o autor, o romance suíço (*Swiss Romansch*), por exemplo, seria uma língua que exibe um paradigma verbal morfologicamente rico, mas que pertence ao grupo das línguas de sujeito não-nulo; o mesmo ocorreria com o alemão: a língua tem marcadores de concordância verbal ricos, mas só permite sujeito nulo em posição de Tópico ou em certas construções impessoais. Além disso, para ele, existem línguas que não têm morfologia verbal rica, mas que permitem sujeitos nulos, o que seria o caso do chinês e do japonês, nas quais sujeitos nulos são identificados via contexto linguístico.

Modesto (2000) assume essa visão de que não existe, necessariamente, a relação de causa entre os dois fenômenos em questão. A esse respeito, inclusive, o autor afirma que uma das premissas das teorias *pro-drop* estaria errada: ou a concordância no PB não teria se tornado fraca ou a relação de dependência entre argumentos nulos e concordância rica não seria tão forte como se costuma assumir. O autor cita também Negrão (1990 *apud* MODESTO, 2000), explicando que a autora conclui que concordância empobrecida não levaria a uma porcentagem mais alta de sujeitos plenos, o que, segundo Modesto, indicaria que o fenômeno *pro-drop* no PB não estaria relacionado a concordância.

Em nossa pesquisa, também levantamos dados diacrônicos sobre a substituição de *tu* por *você* e de *nós* por *a gente*, mudanças responsáveis pela simplificação dos paradigmas

verbais, conforme considera Duarte (1993). No capítulo 5, iremos relacionar tais dados com nossas próprias constatações sobre o uso de sujeito nulo/pleno, não propriamente para pensar a relação de causa entre esses fenômenos, mas para discutir a interação entre esses fenômenos, considerando as motivações cognitivas/comunicativas da língua.

3.2 Aumento da frequência de uso do pronome *você* em detrimento de *tu*

A emergência de *você* no quadro pronominal do PB, conforme explicam Faraco (1996), Lopes (2008), Lopes e Duarte (2003), dentre outros, decorre da evolução linguística da forma nominal de tratamento *Vossa Mercê*.

Faraco (1996) explica que o sistema de tratamento do português passou por um conjunto de mudanças a partir do século XIV, e especialmente a partir do século XV, quando foram introduzidas na língua formas de tratamento, utilizadas de início apenas para o rei, com a estrutura *Vossa + substantivo* (*Vossa Mercê*, *Vossa Senhoria*, *Vossa Alteza*, *Vossa Excelência*, *Vossa Majestade*). Segundo o autor, a mais antiga dessas formas (*Vossa Mercê*) teria evoluído, gerando, então, o pronome de segunda pessoa (*você*), com sua contraparte plural (*vocês*).

Lopes e Duarte (2003) consideram que a forma *você* teria se originado a partir de *Vossa Mercê* por meio do processo de *gramaticalização*. Em termos gerais, pode-se dizer que esse processo constitui um tipo de mudança linguística em que um item ou construção lexical torna-se gramatical, ou, se já gramatical, torna-se ainda mais gramatical. Segundo as autoras, a gramaticalização de *Vossa Mercê* a *você* teria começado timidamente no século XVIII, implementando-se de forma mais acelerada no final do século XIX. Elas procuram evidenciar a ocorrência da gramaticalização de *você* pela aplicação dos critérios propostos por Hopper (1991): *Estratificação*, *Divergência*, *Especialização*, *Persistência* e *Descategorização*.

Segundo aquelas autoras, o Princípio de Estratificação estipula a coexistência entre a forma original e a nova em um mesmo domínio funcional. No caso, como observam, não há o descarte imediato da forma mais antiga *Vossa Mercê* em detrimento da forma emergente *você*, mas um período de transição, de convivência entre as duas formas.

O Princípio de Divergência, como explicam as autoras, postula, por sua vez, que a forma original e a gramaticalizada convivem, apresentando, porém, divergências funcionais. De acordo com as autoras, o substantivo *mercê*, que existe ainda hoje na língua, conserva sua integridade fonológica e, de certo modo, sua integridade semântica, como em *Estamos à mercê de bandidos*, enquanto a forma gramaticalizada *você* sofreu perda gradual de substância fonológica (erosão), na passagem *Vossa Mercê* > *vosmecê* > *você* > *cê*, e perda semântica (dessemantização), devido à perda do caráter original de reverência e cortesia.

O Princípio de Especialização está associado à especialização de uma forma na marcação de uma determinada noção gramatical. À medida em que se implementa, esse princípio gera uma redução de formas possíveis dentro de um mesmo domínio funcional, pois a forma especializada passa a abarcar as diferentes nuances semânticas do domínio funcional de que faz parte, suprimindo as demais formas. Como resultado, essa forma, como principal representante do domínio funcional a que pertence, torna-se mais frequente.

Segundo explicam Lopes e Duarte (2003), essa especialização de uso pode ser identificada no fato de que a forma *você* passa gradativamente a ocorrer em contextos linguísticos específicos e diferentes dos contextos favorecedores de *Vossa Mercê*. Assim, as autoras mostram que o uso de *você* apresenta, a partir do século XIX, os maiores índices de uso entre membros das classes populares, enquanto *Vossa Mercê* conserva-se como forma mais produtiva nas relações socialmente assimétricas. As autoras também identificam um aumento de frequência da forma *você* na passagem do século XVIII para o século XIX.

O Princípio de Persistência está relacionado à conservação de alguns traços do

significado da forma fonte na forma gramaticalizada, quando esta não se encontra em estágios muito avançados de gramaticalização. No caso de *você*, as autoras afirmam que persiste a especificação original de 3ª pessoa, traço da sua origem nominal (*Vossa Mercê*), embora a interpretação semântico-discursiva passe a ser a de forma pronominal de 2ª pessoa.

Finalmente, o Princípio de Decategorização (ou Descategorização) determina que as formas em processo de gramaticalização tendem a neutralizar as marcas morfológicas e as propriedades sintáticas da categoria-origem e assumir os atributos da categoria-destino. Segundo Lopes e Duarte (2003), como implementação desse princípio, em vez de expressão da forma *você*, passa a ocorrer predominantemente sujeito nulo, comportando-se essa forma de modo similar a pronomes como *vós* e *tu*, num momento em que a língua seria ainda uma língua de sujeito nulo. Em contrapartida, no mesmo período, a forma *Vossa Mercê* apresenta índices relevantes como sujeito pleno. Para as autoras, esses fatos indicam que a forma *você* estaria perdendo gradativamente a autonomia sintático-discursiva de nome, passando a se comportar como pronome.

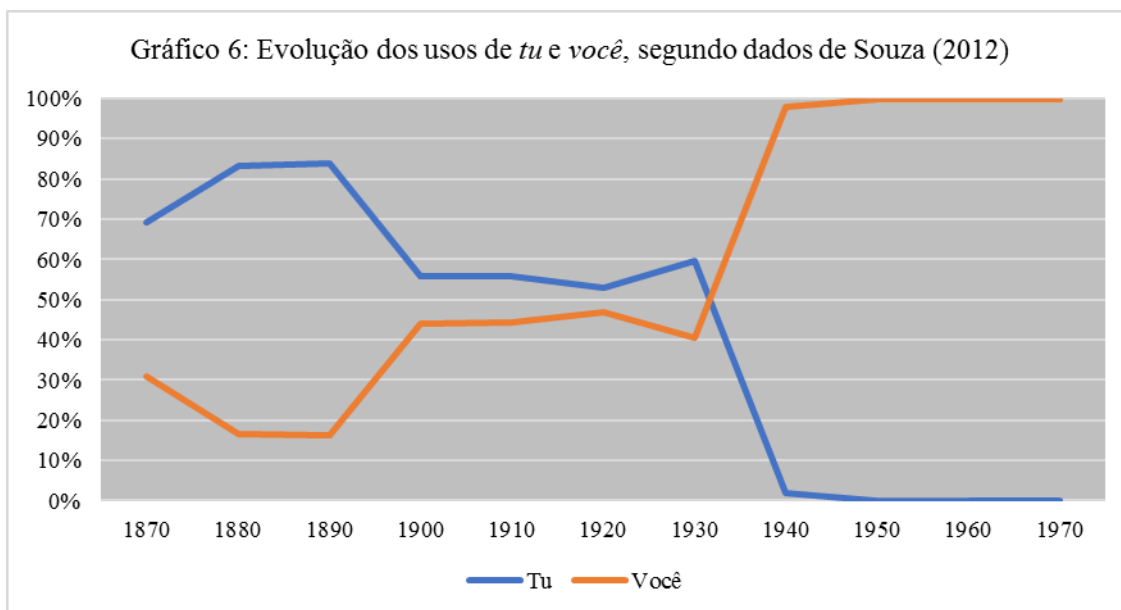
Por meio, então, desse processo de gramaticalização, de acordo com as autoras, a forma *você* torna-se pronome, “invadindo” o sistema pronominal do PB principalmente no fim do século XIX, e, a partir daí, passa a concorrer mais diretamente com o pronome *tu*, sendo cada vez mais usada, em detrimento desse outro pronome.

Alguns autores mostram dados estatísticos que atestam esse processo. Souza (2012), por exemplo, analisa as ocorrências de *tu* e *você* em posição de sujeito, em cartas pessoais escritas entre os anos de 1870 e 1970, por cariocas e moradores do estado do Rio de Janeiro. A autora levanta os dados que apresentamos abaixo na tabela 1.

Tabela 1: Evolução dos usos de *tu* e *você*, segundo dados de Souza (2012)

	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970
<i>Tu</i>	69,1%	83,4%	83,9%	55,9%	55,7%	53%	59,6%	2%	0%	0%	0%
<i>Você</i>	30,9%	16,6%	16,1%	44,1%	44,3%	47%	40,4%	98%	100%	100%	100%

Os dados da tabela 1 mostram, no período recortado, uma tendência geral de diminuição da frequência de uso do pronome *tu* e a tendência geral correspondente de aumento dos índices de uso de *você*. Verifica-se um aumento do percentual de *tu* entre 1870 e 1880 e também, de forma mais sutil, entre 1880 e 1890 e entre 1920 e 1930. Porém, predomina a orientação geral de queda no uso de *tu*, sobressaindo uma queda abrupta entre 1930 e 1940. Souza (2012), inclusive, destaca essa queda acentuada, observando que essa constatação corrobora resultados de Duarte (1993), segundo a qual entre essas décadas, de fato, o pronome *você* teria suplantado drasticamente o uso de *tu* – é interessante notar, conforme mostraremos no capítulo 5, que nossos dados indicam essa mudança entre *tu/você*, inclusive a suplantação expressiva de *você* sobre *tu* nesse mesmo período. O gráfico 6 abaixo ilustra os dados da tabela 1, evidenciando essas tendências gerais.



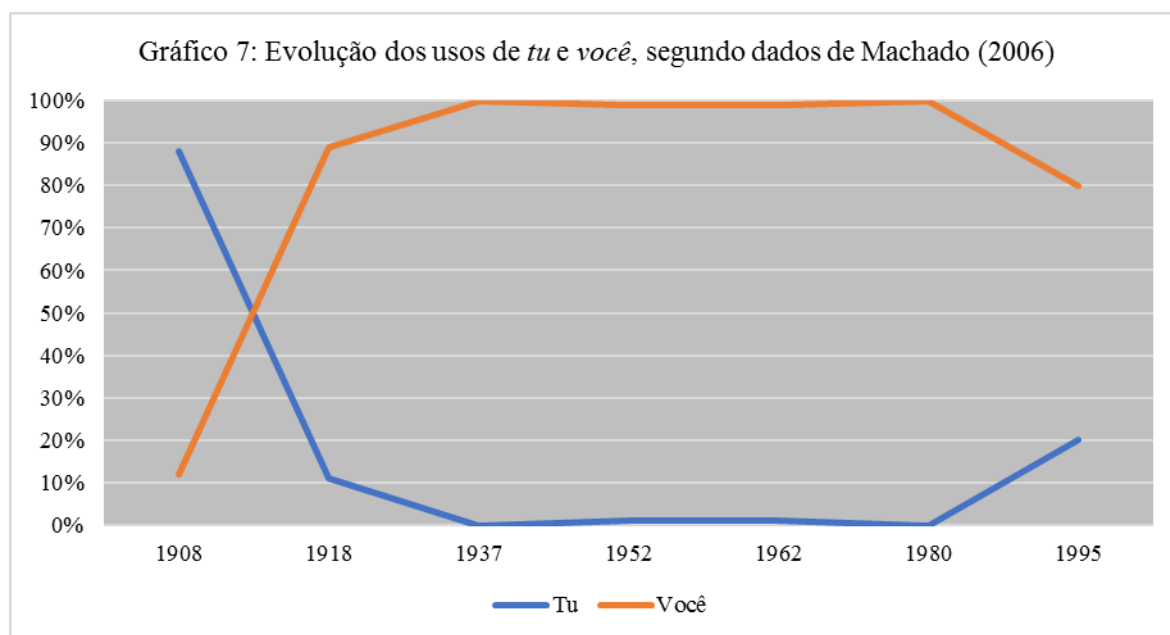
O mesmo processo de favorecimento contínuo do uso de *você* em prejuízo de *tu* é atestado por Machado (2006). A autora analisa peças teatrais, produzidas e ambientadas no Rio de Janeiro, no decorrer de todo o século XX. Parte dos dados que ela apresenta são mostrados abaixo na tabela 2 – a autora fornece dados sobre o uso desses dois pronomes em

diferentes funções sintáticas, porém transcrevemos aqui apenas os dados referentes ao emprego desses pronomes na função de sujeito.

Tabela 2: Evolução dos usos de *tu* e *você*, segundo dados de Machado (2006)

	1908	1918	1937	1952	1962	1980	1995 ²⁹
<i>Tu</i>	88%	11%	0%	1%	1%	0%	20%
<i>Você</i>	12%	89%	100%	99%	99%	100%	80%

Para facilitar a visualização da evolução desses pronomes, representamos os dados da tabela 2 também no gráfico 7.



Como se pode notar na tabela e no gráfico, Machado (2006) também identifica a tendência geral de diminuição de frequência de emprego de *tu* – e correspondente ampliação do uso de *você*. No caso, a autora detecta um grande decréscimo no percentual de *tu* mais no início do século XX, quando a frequência de uso cai de 88%, em 1908, para 0%, em 1937.

²⁹ Em seu trabalho, para cada ano analisado, Machado (2006) utiliza uma peça teatral, incluindo, apenas para o ano de 1995, duas peças. A autora apresenta separadamente os dados dessas duas peças de 1995. Porém, por se tratar de peças de um mesmo ano, optamos por calcular uma média entre os resultados, obtendo um único valor para esse ano para preenchimento da tabela 2.

Permaneça, então, praticamente estável nesse patamar até 1980, mostrando um certo aumento na passagem para 1995.

Diversos outros trabalhos, com abordagem diacrônica ou sincrônica, também registram o processo de mudança envolvendo os pronomes *tu/você*. Lopes e Cavalcante (2011) dedicam-se ao período compreendido entre 1870 e 1940, analisando os usos de *tu* e *você* na posição de sujeito, em cartas familiares. Notam um aumento gradativo e constante de *você* ao longo do período considerado, principalmente entre 1910 e 1919. Segundo as autoras, de 1870 a 1909 as taxas de *você* apresentavam índices pouco representativos (em torno de 30%), evidenciando a supremacia de *tu*, que prevalece até os anos de 1910 aproximadamente. Entre 1910 e 1919, identificam um equilíbrio no uso das formas *você* e *tu*, com valores próximos dos 50%. A partir dos anos 1920, *você* supera o uso de *tu*, com taxa de 79%.

Já como evidência sincrônica, podem ser considerados, por exemplo, dados levantados por Lopes (2008), que apura ocorrências de *tu* e *você* em roteiros cinematográficos ambientados na cidade do Rio de Janeiro no início do século XXI, chegando às taxas médias de 27,5% e 72,5% desses dois pronomes, respectivamente. Outros trabalhos que, com dados sincrônicos, também indicam a predominância de *você* sobre *tu* com a evolução do PB incluem Lucca (2005), Dias (2007), Lopes *et al.* (2009), entre outros.

Como se vê, o agrupamento de trabalhos aqui relatados, que juntos recobrem o período que vai do início do século XVIII ao início do século XXI, evidencia a queda da frequência de uso de *tu* e o aumento da de *você*, a partir da primeira metade do século XIX, e a suplantação de *tu* por *você* no decorrer do século XX. Esses trabalhos mostram ainda que essa mudança pode ser atestada por meio de diferentes tipos de *corpora* (cartas pessoais, cartas familiares, peças teatrais, roteiros cinematográficos) e, inclusive, em uma região do país em que o uso de *tu* seria bastante produtivo.

Na mesma direção dos trabalhos aqui sintetizados, em nossa pesquisa também

apuramos o aumento diacrônico da frequência de uso de *você* em detrimento de *tu*. No capítulo 5 adiante, mostraremos nossos resultados desse levantamento.

3.3 Aumento da frequência de uso da forma pronominal *a gente* em detrimento de *nós*

De modo semelhante ao que se verifica no âmbito da 2ª pessoa do singular, o quadro pronominal do PB também experiencia uma mudança diacrônica no domínio da 1ª pessoa do plural, por meio da introdução da forma *a gente*, que passa a concorrer com o pronome *nós*.

Omena e Braga (1996) explicam que o substantivo feminino *gens*, *gentis* deu origem à palavra *gente*, que nomeia, de forma coletiva, indeterminadora, mais ou menos geral, um agrupamento de seres humanos, identificados entre si por objetivos, ideias, qualidades, nacionalidade ou posição. Segundo as autoras, mais recentemente na história do PB, altera-se a semântica do termo, que passa a incorporar ao seu significado a referência à 1ª pessoa do discurso, sobretudo à 1ª do plural. Sempre determinada pelo artigo *a*, a palavra *gente* passa, então, a constituir uma expressão que propicia alternâncias com o pronome *nós*.³⁰

Novamente, o processo de gramaticalização é considerado o responsável pela alteração do quadro pronominal, agora pela introdução de *a gente*, conforme defendem alguns autores, como é o próprio caso de Omena e Braga (1996) e também de Lopes (1999; 2004). Trata-se de um processo que, de acordo com esta segunda autora, teria se implementado principalmente durante os séculos XVIII e XIX.

Omena e Braga (1996) procuram evidenciar a gramaticalização de *a gente* por meio também dos cinco princípios de Hopper (1991), abordados na seção anterior quanto ao

³⁰ Omena e Braga (1996) afirmam que o significado da expressão *a gente* incorpora também referência à 1ª pessoa do singular, sendo alternativa ao pronome *eu*, o que, porém, salientam, já constitui um uso menos frequente e bastante impreciso.

pronome *você*³¹. As autoras reportam-se ao Princípio da Estratificação pelo termo inglês *layering*, considerando, assim como mencionado acima, que tal princípio remete à coexistência de diversas camadas, isto é, alternantes, dentro de um mesmo domínio funcional. Acrescentam as autoras que essa coexistência torna-se possível porque a emergência de novas camadas/variantes não implica a eliminação das antigas. Explicam ainda que essas diversas camadas, as quais servem para codificar funções similares ou idênticas, podem estar correlacionadas a itens lexicais particulares, classes de construções ou registros sociolinguísticos e que os itens envolvidos podem ter significados ligeiramente diferentes ou ser reconhecidos como alternativas estilísticas. Nesse sentido, afirmam as autoras que o comportamento da forma *a gente* manifesta-se como um legítimo caso de *layering*, na medida em que o item *gente* teria dado origem à expressão *a gente*, a qual passaria, então, a competir com a forma *nós*, ambas exercendo funções similares.

Quanto ao Princípio da Divergência, Omena e Braga (1996) lembram que este remete à preservação, conservação da forma lexical que dá origem ao processo de gramaticalização, a qual pode permanecer como um item lexical autônomo e, enquanto tal, sofrer ou não as mudanças que atingem os itens da classe a que pertence. A esse respeito, observam que o item *gente* teria permanecido no PB como um item lexical autônomo, sem sofrer mudanças fonológicas, porém a forma inovadora é que teria sofrido alterações. Esta teria cristalizado a relação determinante-determinado, fixando-se com o artigo definido, não admitindo outras especificações e, além disso, passaria a ter uma mobilidade mais restrita, identificando-se funcionalmente com os pronomes.

No tocante ao Princípio da Especialização, que remete ao estreitamento de escolhas sofridas pelas construções, Omena e Braga (1996) focalizam a alternância entre as formas *nós* e *a gente*, identificando que a segunda, com o andamento de sua gramaticalização,

³¹ Cabe observar que é possível notar certas diferenças na interpretação que Omena e Braga (1996) fazem desses princípios, em comparação com o entendimento de Lopes e Duarte (2003) relatado acima sobre a forma *você*, embora não cheguemos a discutir tais questões aqui, por fugirem aos objetivos da presente exposição.

predomina sobre a primeira. Avaliando a distribuição dessas formas pelas funções sintáticas, com base em dados de Omena (1986), observam que a forma *a gente* é majoritária nas funções de sujeito, complemento e adjunto adverbial, prevalecendo a forma *nós* apenas na função de adjunto adnominal.

Em relação ao Princípio da Persistência, que se refere à manutenção, por parte da forma que sofre gramaticalização, de traços semânticos da forma fonte, Omena e Braga (1996) afirmam que esses traços explicam, muitas vezes, certas restrições que a forma gramaticalizada experimenta em sua distribuição. De acordo com elas, a ideia de coletividade do substantivo *gente* contribui para uma referência indeterminadora presente em certos usos da forma pronominal *a gente*, o que provoca algumas restrições em seu uso. Como apontam as autoras, enquanto o pronome *nós* admite ser modificado por quantificadores, numerais, especificadores, o mesmo não se dá com *a gente* – por exemplo, *todos*, *cada um de*, *nenhum de* podem modificar *nós*, mas não *a gente*.

Por fim, o Princípio da Descategorização, conforme tratado acima e também como sintetizam as autoras, estipula que as formas em processo de gramaticalização tendem a perder ou neutralizar os marcadores morfológicos e os privilégios sintáticos típicos das categorias plenas e a assumir os atributos característicos das categorias secundárias. Assim, a forma *gente* guarda a mobilidade de colocação dos substantivos, admite flexão de número, derivação de grau, pode ser modificada por quantificadores, determinantes, possessivos, locuções prepositivas etc., enquanto a forma *a gente*, continuam as autoras, embora tenha com o verbo a mesma relação que o substantivo *gente*, levando-o para a 3ª pessoa do singular, apresenta a variação sintática *a gente vamos*, que demonstraria maior integração ao sistema pronominal por concordar com a 1ª pessoa do plural.

Enfim, mediante a aplicação dos critérios de Hopper (1991), as autoras sustentam que a forma *a gente* gramaticaliza-se, processo, então, que seria responsável pela

introdução dessa forma no quadro pronominal do PB e por sua concorrência direta com o pronome *nós*. Assim, passaria a se instaurar na língua a mudança diacrônica de que estamos tratando no presente trabalho, referente ao aumento da frequência de uso de *a gente* em detrimento de *nós*.

Durante nossa pesquisa, no levantamento bibliográfico que realizamos, diferentemente do apurado quanto à mudança *tu/você*, não chegamos a encontrar trabalhos com dados diacrônicos que mostrem o aumento gradativo dos índices de *a gente*, em prejuízo de *nós*. De todo modo, diversos trabalhos de ordem sincrônica assumem a ocorrência desse processo, apresentando dados que, de fato, o evidenciam – a propósito, nosso estudo, na medida em que oferece dados diacrônicos sobre a mudança *nós/a gente*, representa uma contribuição significativa para o acompanhamento histórico dessa mudança.

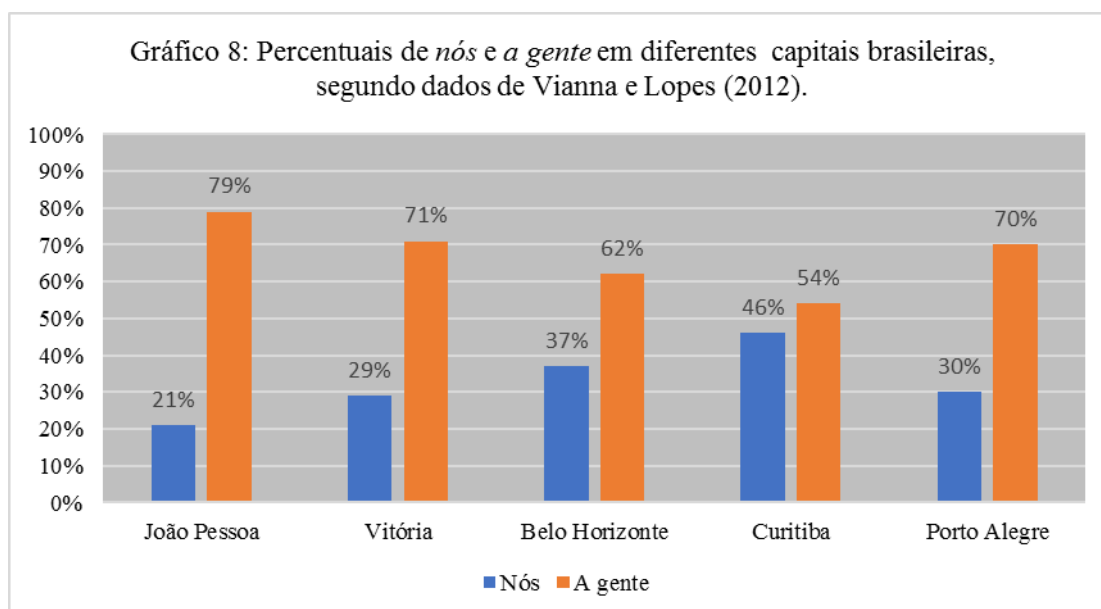
Rubio e Gonçalves (2010), por exemplo, levantam dados do Banco de Dados Iboruna (GONÇALVES, 2007), que reúne amostras de língua falada de diferentes gêneros textuais (como narrativa de experiência, relato de opinião, entre outros), recolhidas de informantes da região noroeste do Estado de São Paulo, entre os anos de 2004 e 2006. Analisando a ocorrência das formas *nós* e *a gente* na função sintática de sujeito, apuram os percentuais de 25,4% e 74,6% dessas formas, respectivamente.

Seara (2000), por sua vez, investigando entrevistas de informantes florianopolitanos (SC) do Projeto VARSUL (Projeto de Variação Linguística da Região Sul do Brasil), coletadas na década de 1990, e também averiguando a função de sujeito, identifica as taxas de 28% de *nós* e 72% de *a gente*.

Analogamente, Santana (2014) dedica-se ao estudo dessas duas formas, também na função de sujeito, no português popular de Salvador (BA), utilizando como *corpus* o Programa de Estudos do Português Popular Falado em Salvador, constituído por entrevistas do tipo diálogo entre informante e documentador, em situação de fala semi-espontânea,

coletadas entre 1998 e 2000. O autor chega aos valores de 24% e 76% para as formas *nós* e *a gente*, respectivamente.

Vianna e Lopes (2012), em trabalho de orientação sociolinguística, apresentam dados de diferentes autores (os quais utilizam *corpora* diversos), para comparar os usos de *nós* e *a gente* em diferentes capitais brasileiras, com dados referentes ao período compreendido entre fins do século XX e início do século XXI. Os resultados de todas as capitais mostram o uso majoritário de *a gente* sobre *nós*. Os dados sintetizados pelas autoras são apresentados abaixo no gráfico 8.



Considerando, portanto, que, em séculos anteriores do PB, não se registrava o uso de *a gente* como pronome, ou se registrava um baixo índice desse uso, e tendo em vista, atualmente, o emprego pronominal dessa expressão, com taxas mais altas do que é exibido pelo pronome *nós*, pode-se concluir que o PB experienciou o processo de aumento da frequência de uso de *a gente* em detrimento de *nós*. No capítulo 5 adiante, mostraremos que nosso levantamento de dados comprova esse processo, especificando sua implementação gradativa do início do século XIX ao início do XXI.

CAPÍTULO IV

Metodologia

Este capítulo encontra-se dividido em duas partes. Inicialmente, na seção 4.1, apresentamos a metodologia utilizada para levantamento e análise de dados. Nessa seção, num primeiro momento, fazemos uma delimitação específica de nosso objeto de estudo; em seguida, explicamos os critérios utilizados para o levantamento de dados referentes às três mudanças diacrônicas ocorridas no PB aqui em discussão, bem como a metodologia que elaboramos para, a partir desses dados, calcular a variação diacrônica do grau de transparência do sistema de referência em estudo. Na seção 4.2, apresentamos o *corpus* de nossa pesquisa.

4.1 Metodologia de levantamento e análise de dados

Como explicado anteriormente, o objetivo central deste trabalho é analisar a variação diacrônica do grau de transparência do sistema de referência por expressão pronominal e desinencial do argumento-sujeito de 1ª e 2ª pessoas no PB. Esse sistema, em termos de transparência ou opacidade, é composto por quatro possibilidades de se fazer referência em uma oração:

- (i) *Referência Cruzada*: o argumento-sujeito é expresso por desinência verbal número-pessoal explícita e por pronome;
- (ii) *Referência Zero*: o argumento-sujeito não é expresso por desinência verbal número-pessoal explícita (correspondendo a morfema zero) nem por pronome;
- (iii) *Referência Simples 1*: o argumento-sujeito é expresso por desinência verbal número-pessoal explícita, mas não por pronome;
- (iv) *Referência Simples 2*: o argumento-sujeito não é expresso por desinência verbal número-pessoal explícita, sendo expresso apenas por pronome.

As duas primeiras combinações constituem realizações opacas da referência, enquanto as duas últimas são realizações transparentes. Os exemplos de (24) a (27) abaixo ilustram essas quatro combinações respectivamente.³²

- (24) TIBURCIO – Consentes que eu fale a teu mano?
 LUIZA (*vergonhosa*) – Não sei...
 TIBURCIO (*beijando-lhe a mão*) – Malditos tagarellas que iam-me fazendo perder este torrão d'assucar. Minha Luiza, **nós seremos** muito felizes (19-1 OIA, p. 11).
- (25) AUGUSTO – Se quer vender as suas acções, não perca a ocasião.
 ERNESTO – Vender as minhas acções?
 AUGUSTO – Sim, Sr.; acredite no que lhe digo; não valem mais do que cinco mil réis é já são bem pagas.
 ERNESTO – O Sr. quer brincar naturalmente!
 AUGUSTO – Não brinco em negocio. Para encurtar razões dou-lhe seis mil réis. Quer? Aqui estão. Quantas tem?
 ERNESTO (*a Braga*) – Deste gênero ainda não **tinha** encontrado! É pior do que os taes cambistas de loterias. (*Passeia*). (19-2 VER, p. 12-13)
- (26) ISABEL — Ludgero, não se esqueça de que vamos à casa do conselheiro, e é longe.
 LUDGERO — Tens razão, minha mulher. Vamos!
 ÂNGELO — Então não **jantam** conosco? (20-1 ODO, p. 38)
- (27) TERESA – Mas não quero que você morra nunca! Só depois de mim (com uma nova expressão, embelezada). Ou então, ao mesmo tempo, juntas. Eu e você enterradas no mesmo caixão.
 GLÓRIA – **Você gostaria?**
 TERESA (no seu transporte) – Seria tão bom, mas tão bom!
 GLÓRIA (prática) – Mas no mesmo caixão não dá – nem deixam! (20-1 ADF, p. 21)

Os tipos de referência por expressão pronominal e desinencial do argumento-sujeito no PB podem ser também visualizados no quadro abaixo.

Quadro 4: Tipos de referência em termos de transparência e opacidade

		Expressão desinencial do argumento-sujeito	
		+	-
Expressão pronominal do argumento-sujeito	+	RC	RS
	-	RS	RZ

Mais especificamente, o sistema de referência aqui delimitado para análise

³² As ocorrências exemplificativas foram extraídas do *corpus* de análise (cf. seção 4.2, deste capítulo).

compreende o conjunto de combinações entre:

- (i) a expressão/não-expressão de cada uma das formas pronominais *eu*, *tu*, *você*, *nós*, *a gente*, *vós* e *vocês* e
- (ii) a forma verbal correspondente a cada pronome em cada um dos nove tempos/modos verbais do PB (algumas das quais têm desinência número-pessoal explícita, outras não).

Esse sistema de referência completo inclui as 126 combinações mostradas no quadro 5.³³

Quadro 5: Combinações constituintes do sistema de referência por expressão pronominal e desinencial do argumento-sujeito de 1ª e 2ª pessoas no PB.

Expressão pronominal	ind. pres.		ind. pret. imp.		ind. pret. perf.		ind. pret. m.q.perf.		ind. fut. pres.		ind. fut. pret.		subj. pres.		subj. pret. imp.		subj. fut.		Totais
	des.	ref.	des.	ref.	des.	ref.	des.	ref.	des.	ref.	des.	ref.	des.	ref.	des.	ref.	des.	ref.	
<i>eu</i>	+																		9
	-																		9
<i>tu</i>	+																		9
	-																		9
<i>você</i>	+																		9
	-																		9
<i>nós</i>	+																		9
	-																		9
<i>a gente</i>	+																		9
	-																		9
<i>vós</i>	+																		9
	-																		9
<i>vocês</i>	+																		9
	-																		9
Totais																			126

Como se pode notar, a realização transparente ou opaca da referência aqui em estudo depende da combinação, na expressão do argumento-sujeito, entre presença ou ausência de desinência verbal número-pessoal explícita e presença ou ausência de pronome pessoal. Desse modo, investigamos aqui três mudanças diacrônicas já atestadas no PB que afetam justamente a combinação entre esses dois fenômenos e que, assim, interferem diretamente na variação diacrônica do grau de transparência do sistema de referência por

³³ Nesse quadro, é indicada a presença ou ausência de desinência número-pessoal explícita nas formas verbais consideradas neste trabalho. Mais adiante, no quadro 6, apresentamos o quadro específico de desinências aqui considerado.

expressão pronominal e desinencial do argumento-sujeito: (i) aumento da frequência de argumento-sujeito expresso por pronome em oposição à não-expressão por pronome; (ii) aumento da frequência de uso do pronome *você* em detrimento do pronome *tu*; (iii) aumento da frequência de uso da forma pronominal *a gente* em detrimento de *nós*. Assim sendo, em nossa pesquisa, realizamos o levantamento de dados referente a essas três mudanças para, então, analisarmos o impacto delas no comportamento diacrônico do grau de transparência do PB.

No que diz respeito ao aumento da frequência de argumento-sujeito expresso por pronome, é necessário reconhecer que tal mudança afeta a opção entre expressar o argumento por meio de pronome pessoal *versus* não o expressar, deixando a expressão apenas por conta da desinência verbal (isto é, trata-se da opcionalidade entre sujeito pronominal pleno *versus* sujeito pronominal nulo, característica das chamadas línguas *pro drop*, como mostrado no capítulo anterior). A mudança gramatical identificada no PB diz respeito a essa opcionalidade, não envolvendo a expressão *nominal* do argumento-sujeito. Por essa razão, não incluímos no presente trabalho a referência que envolve argumentos-sujeitos *nominais*.

Além disso, nosso estudo está restrito à análise de 1ª e 2ª pessoas do discurso (no singular e no plural), desconsiderando a 3ª pessoa (singular e plural). A opção por 1ª e 2ª pessoas do singular e do plural (daqui em diante, 1PS, 2PS, 1PP e 2PP) se justifica com base no fato de que é com essas pessoas (particularmente com a 2PS e com a 1PP) que se verificam, no PB, mudanças que podem interferir no grau de transparência da referência por expressão do argumento-sujeito (a inclusão da 1PS e da 2PP foi adotada para completarmos o estudo de 1ª e 2ª pessoas). Embora reconheçamos que o aumento na frequência de expressão pronominal do argumento-sujeito leve à variação diacrônica de transparência também no uso da 3ª pessoa (do singular e do plural), optamos por não incluir essa pessoa em nossa análise, como um dos recortes adotados para a delimitação de nosso objeto do estudo. Ademais, tendo

em vista que a 1ª e a 2ª pessoas constituem as chamadas *pessoas do discurso*, em oposição à 3ª, considerada a *não-pessoa*, essa delimitação circunscreve a pesquisa a um subsistema particular de referência, no caso, a referência de natureza dêitica.

Neste trabalho, estudamos as três mudanças diacrônicas elencadas acima, porque todas elas afetam a realização transparente ou opaca da referência por expressão do argumento-sujeito. Uma mudança diacrônica similar à ocorrida com as formas pronominais *tu/você* e *nós/a gente* consiste no aumento da frequência de uso de *vocês* em detrimento de *vós*. Porém, não incluímos esse caso em nossa pesquisa, uma vez que tal mudança não interfere no grau de transparência da referência. A presença ou ausência de desinência número-pessoal explícita nos diferentes tempos/modos verbais é igual com ambos esses pronomes (todos os tempos/modos verbais apresentam desinência número-pessoal explícita com esses dois pronomes), e, assim, a escolha entre um e outro pronome não interfere na realização opaca ou transparente da referência, diferentemente do que ocorre com a alternância *tu/você* e *nós/a gente*.

Quanto ao período diacrônico recortado para nosso estudo, selecionamos o período que vai do século XIX ao início do século XXI. Essa delimitação se deve a restrições de acesso a mesmo tipo de material para análise de períodos anteriores ao século XIX, já que nossa pesquisa se atém a um gênero textual particular, peças teatrais, conforme justificado na seção seguinte. Seguindo um procedimento normalmente adotado em pesquisas diacrônicas, dividimos cada século em primeira e segunda metade, de modo a possibilitar a comparação diacrônica entre cinco sincronias: primeira e segunda metades dos séculos XIX e XX e primeira metade do século XXI, referidas, ao longo deste trabalho, na maior parte das vezes, como 19-1, 19-2, 20-1, 20-2 e 21-1, respectivamente).

Cabe aqui esclarecer que a finalidade do levantamento de dados referente às três mudanças diacrônicas aqui em foco não é, primordialmente, demonstrar a existência dessas

mudanças, na medida em que diferentes autores já as atestam (ver capítulo 3). Na verdade, nosso levantamento tem como meta, em certa medida, corroborar a ocorrência dessas mudanças e, principalmente, obter nossos próprios dados para cálculo da variação diacrônica da transparência. O levantamento de dados próprios é necessário e pertinente para que tenhamos dados das três mudanças em pauta extraídos de um mesmo *corpus* e para que os dados sejam coletados a partir de um levantamento específico já metodologicamente projetado para um mesmo fim, no caso, a análise de variação diacrônica de transparência.

Após essa delimitação do tema em estudo, passamos agora à explicação de como realizamos a análise desses fenômenos. Em nossa pesquisa, procedemos a três levantamentos de dados separados, um para cada uma das três mudanças diacrônicas consideradas, uma vez que cada uma delas se pauta por diferentes critérios de levantamento e de análise.

O primeiro levantamento recolheu dados sobre o aumento diacrônico da frequência de expressão pronominal do argumento-sujeito, *versus* a sua não-expressão. Foram levantadas, em cada sincronia em estudo, 200 ocorrências entre orações com expressão ou não-expressão pronominal do argumento-sujeito, a fim de identificarmos a frequência de uso de cada uma dessas possibilidades em cada sincronia. Assim, considerando o recorte de cinco sincronias, levantamos, no total, 1000 dados sobre expressão ou não-expressão pronominal do argumento-sujeito.

Nesse levantamento, para a obtenção do total de ocorrências em cada sincronia, fomos selecionando todos os dados encontrados nos textos do *corpus*, desde o início do texto, até completarmos a quantidade estabelecida (200 dados, no caso desse levantamento). Mais detalhadamente, conforme explicaremos adiante, para cada texto do *corpus*, estipulamos o levantamento de uma quantia máxima de 25 dados. Assim, para cada texto, fomos levantando todos os dados encontrados, desde seu início até a obtenção de, no máximo, 25 dados. Por exemplo, numa sincronia em que cada texto pôde oferecer, de fato, 25 dados (tendo em vista

os critérios de levantamento abaixo definidos), foram usados, então, oito textos para se chegar ao total de 200 dados sobre expressão/não-expressão pronominal. O mesmo tipo de procedimento foi aplicado aos levantamentos sobre *tu/você* e *nós/a gente* (com a diferença explicada mais adiante de que, para cada um desses dois levantamentos, selecionamos, de cada sincronia, 100 ocorrências – e não 200).

No levantamento sobre expressão/não-expressão pronominal, os seguintes critérios foram considerados:

- (i) selecionamos ocorrências de expressão/não-expressão pronominal de argumento-sujeito apenas em orações principais, ou seja, apenas em orações absolutas, em orações matrizes de subordinadas, em orações núcleos de adverbiais e na primeira oração de estruturas coordenadas. Estabelecemos esse critério por considerarmos que as *orações principais* constituem contexto de menor restrição sobre a escolha entre expressão ou não-expressão pronominal do argumento-sujeito, isto é, um contexto em que o argumento-sujeito seria ou não expresso pronominalmente a depender mais livremente da tendência da língua para uma e outra dessas opções. Outros contextos estruturais, como o de oração subordinada, por exemplo, não apresentam esse mesmo grau de neutralidade, podendo favorecer uma dessas possibilidades, no caso, a opção pela não-expressão pronominal;
- (ii) levantamos, separadamente, dados referentes às quatro pessoas do discurso em análise, para identificar a frequência de expressão/não-expressão pronominal do argumento-sujeito com cada pessoa particularmente, tendo em vista o fato de que a frequência pode se mostrar diferente entre uma pessoa e outra. Assim, organizamos nosso levantamento das 200 ocorrências de expressão/não-expressão pronominal do argumento-sujeito de cada sincronia de modo a obter 50 ocorrências de cada uma das pessoas, garantindo, com isso, uma quantidade mínima de dados para cada uma dessas pessoas;
- (iii) selecionamos apenas ocorrências de expressão/não-expressão pronominal do argumento-sujeito em estrutura simples, desconsiderando ocorrências de expressão do

argumento-sujeito compostas por formas pronominais e outros sintagmas, como em eu e minha mãe fomos ao mercado, você e o José vão para a mesma escola etc.;

- (iv) excluímos ocorrências de orações com verbos no imperativo, por considerarmos que esse contexto estrutural favorece a não-expressão pronominal do argumento-sujeito, em detrimento da expressão;
- (v) não controlamos o tempo/modo dos verbos das orações levantadas. Reconhecemos (conforme mostram autores como Duarte (1995)) que a expressão/não-expressão pronominal do argumento-sujeito pode ser influenciada pelo tempo/modo do verbo da oração. Porém, optamos por não controlar esse fator por duas razões: (i) o *corpus* de nossa pesquisa dificilmente disponibilizaria dados suficientes de todos os tempos/modos; como explicamos mais adiante, por conta de fatores relevantes para nosso trabalho, restringimos nosso *corpus* apenas a peças de teatro, o que compromete a constituição de um *corpus* extenso para as sincronias mais antigas; (ii) o controle dessa variável também comprometeria a viabilidade operacional da pesquisa, já que levaria ao levantamento de uma quantidade muito grande de dados, pois implicaria levantar dados (em quantidades suficientes para a análise) dos nove tempos/modos verbais com cada uma das quatro pessoas do discurso. Assim, optamos por não controlar a variável tempo/modo verbal, preferindo controlar a variável pessoa do discurso, a qual, de acordo com os trabalhos sobre o tema, tem mais impacto sobre a escolha entre expressão/não-expressão pronominal do argumento-sujeito.

O segundo levantamento de dados foi dedicado a verificar o aumento diacrônico da frequência de uso do pronome *você* em detrimento de *tu*. Levantamos 100 ocorrências de orações com argumento-sujeito de 2PS (considerando conjuntamente casos de *tu* e de *você*) em cada sincronia estudada (totalizando 500 dados, sendo 100 de cada sincronia), para identificarmos as frequências de escolha de cada uma dessas formas em cada sincronia. Esse levantamento foi norteado pelos seguintes critérios:

- (i) levantamos as formas *tu* e *você* apenas quando expressam o argumento-sujeito e sem

considerar a restrição (adotada na seleção de dados sobre expressão/não-expressão pronominal do argumento-sujeito) de levantar apenas ocorrências em orações principais. No caso desse levantamento, a nosso ver, o fato de a oração analisada ser ou não uma oração principal não favoreceria o uso de *tu* ou *você*;

- (ii) selecionamos ocorrências de argumento-sujeito de 2PS apenas no caso de expressão pronominal, excluindo casos de não-expressão pronominal (isto é, excluindo os casos dos chamados sujeitos nulos). Essa restrição foi adotada para tornar esse levantamento mais uniforme em relação ao levantamento das formas pronominais *nós* e *a gente*, para as quais parece ser relevante essa restrição (conforme explicado logo adiante).

O terceiro levantamento tratou da frequência de uso da forma pronominal *a gente* em detrimento do pronome *nós*. Também levantamos 100 ocorrências de orações contendo argumento-sujeito de 1PP de cada sincronia estudada, com exceção da sincronia 19-1, da qual levantamos 58 dados³⁴ (neste caso reunindo 458 ocorrências no total)³⁵, para identificarmos a frequência de uso de uma e outra forma pronominal em cada uma dessas sincronias. Para esse levantamento, consideramos os seguintes critérios:

- (i) assim como no levantamento de *tu* e *você*, selecionamos as formas *nós* e *a gente* apenas quando expressam argumento-sujeito. Aqui também não adotamos a restrição de considerar apenas orações principais, entendendo que o fato de o argumento-sujeito ser expresso em oração principal ou em oração não-principal não favoreceria o uso de uma ou outra forma;
- (ii) seguindo Rubio e Gonçalves (2010), selecionamos as formas *nós* e *a gente* apenas nos casos em que há expressão dessas formas pronominais, excluindo casos de não-expressão desses pronomes (isto é, casos de sujeito nulo). Essa restrição foi definida

³⁴ O levantamento dessa quantia menor de dados deve-se a indisponibilidade de *corpus* suficiente para um levantamento maior (apesar de termos feito o levantamento num material bastante extenso, como explicado na seção seguinte).

³⁵ Desse modo, analisamos em nossa pesquisa um total de 1958 dados, considerando as três mudanças diacrônicas analisadas, sendo 1.000 dados referentes à expressão/não-expressão pronominal do argumento-sujeito, 500 dados relativos a *tu/você* e 458 sobre *nós/a gente*.

por considerarmos que haveria uma frequência maior de não-expressão pronominal referente ao pronome *nós* do que à forma *a gente*. Assim, o levantamento de ocorrências de não-expressão pronominal poderia resultar na identificação de uma frequência de uso de *nós* maior do que a real, por conta dessa questão sintática;

- (iii) no caso de *a gente*, consideramos apenas ocorrências em que este item se encaixa na categoria de pronome de 1PP, descartando ocorrências em que figura como forma não-gramaticalizada formada por artigo *a* e pelo substantivo *gente*;
- (iv) no caso do pronome *nós*, descartamos ocorrências em que o item é modificado por quantificadores (por exemplo, *nós dois*, *nós todos*, *cada um de nós* etc.), considerando que tais casos não constituem variantes de *a gente*.

Em cada sincronia, o grau de transparência do sistema de referência por expressão do argumento-sujeito decorre da interação entre esses três fenômenos (expressão ou não-expressão pronominal do argumento-sujeito, uso de *tu* ou *você*, uso de *nós* ou *a gente*) com a presença ou ausência de desinência verbal número-pessoal explícita. A presença/ausência dessa desinência de forma explícita depende do tempo/modo verbal e da pessoa do discurso do argumento-sujeito. A esse respeito, consideramos o quadro 6 abaixo, que, organizado com base em Câmara Júnior (1989) e Bechara (2009), mostra as desinências número-pessoais da 1ª e da 2ª pessoas do singular e do plural de todos os tempos verbais do indicativo e do subjuntivo, considerando apenas os verbos regulares.

Quadro 6: Desinências número-pessoais de verbos regulares do PB

Pessoa-número Tempo-modo	1PS	2PS		1PP		2PP	
	<i>eu</i>	<i>tu</i>	<i>você</i>	<i>nós</i>	<i>a gente</i>	<i>vós</i>	<i>vocês</i>
Pres. ind.	-o	-s	-	-mos	-	-is	-m
Pret. imp. ind.	-	-s	-	-mos	-	-is	-m
Pret. perf. ind.	-i	-ste	-u	-mos	-u	-stes	-m
Pret. m. q. perf. ind.	-	-s	-	-mos	-	-is	-m
Fut. pres. ind.	-i	-s	-	-mos	-	-is	-m (~)
Fut. pret. ind.	-	-s	-	-mos	-	-is	-m
Pres. subj.	-	-s	-	-mos	-	-is	-m
Pret. imp. subj.	-	-s	-	-mos	-	-is	-m
Fut. subj.	-	-s	-	-mos	-	-des	-m

Neste trabalho, consideramos os verbos regulares e, dos verbos irregulares, consideramos as formas verbais que apresentam o mesmo comportamento das formas regulares correspondentes com relação à presença ou ausência de desinência número-pessoal explícita. Por exemplo, no caso do verbo irregular *estar*, excluímos as formas verbais de 1PS e 2PS do pretérito perfeito do indicativo (*estive* e *esteve*, respectivamente), pois essas formas diferem das formas correspondentes de verbos regulares quanto à presença/ausência de desinência número-pessoal, já que, nos verbos regulares, as formas verbais para essas pessoas, nesse tempo/modo, têm desinência número-pessoal (por exemplo, *cante-i*, *canto-u*), enquanto, no verbo *estar*, as formas verbais em questão não têm essa desinência (*estive-Ø* e *esteve-Ø*). Já as demais formas do verbo *estar*, embora se trate de um verbo irregular, comportam-se igualmente às formas correspondentes dos verbos regulares, quanto à presença/ausência de desinência número-pessoal explícita (como se pode ver, por exemplo, nas formas do futuro do pretérito do indicativo, na comparação com o verbo regular *cantar*: eu *estaria-Ø* / eu *cantaria-Ø*; tu *estaria-s* / tu *cantaria-s*; você/a gente *estaria-Ø* / você/a gente *cantaria-Ø*; nós *estariamos* / nós *cantariamos*; vós *estarieis* / vós *cantarieis*; vocês *estariam* / vocês *cantariam*).

Ainda em relação aos paradigmas de desinência verbal número-pessoal, não consideramos o processo de simplificação da morfologia verbal flexional do PB, que leva à eliminação da desinência número-pessoal explícita nas formas verbais referentes aos pronomes *tu*, *nós* e *vocês*, resultando em construções como *tu canta*, *nós canta*, *vocês canta*. Na mesma direção, não incluimos em nossa pesquisa outro fenômeno do PB, que é o uso da flexão *-mos* com a forma pronominal *a gente* (*a gente cantamos*).

A exclusão desses dois fenômenos foi estabelecida primeiramente porque, após uma análise prévia, constatamos que esses fenômenos se verificam em nosso *corpus* muito raramente (não oferecendo quantidades de dados minimamente suficientes para análise) e

apenas em casos muito particulares.³⁶ Além disso, optamos por desconsiderar esses fenômenos em nossa pesquisa, porque a inclusão deles aumentaria substancialmente a quantidade de variáveis a serem controladas. Assim, sem deixar de reconhecer a relevância desses fenômenos para o objeto aqui em estudo, optamos por não os considerar, a fim de viabilizar metodologicamente a pesquisa (como explicamos mais adiante ainda neste capítulo, nossa pesquisa já inclui a análise integrada de uma série de fatores relevantes interferentes na variação diacrônica do grau de transparência da referência).

Uma vez feitos os levantamentos de dados referentes aos três fenômenos diacrônicos acima considerados, nossa análise se concentrou no cálculo, a partir desses dados, da variação diacrônica do grau de transparência do sistema de referência aqui delimitado.

Para isso, nosso procedimento consistiu em calcular o grau de transparência desse sistema (em termos percentuais) em cada sincronia e, em seguida, comparar uma sincronia com a outra, para demonstrar o que ocorreria diacronicamente com o grau de transparência desse sistema. Ou seja, o propósito da análise foi chegar a resultados para o preenchimento da tabela genérica (tabela 3) seguinte:

Tabela 3: Dados a serem obtidos sobre a variação diacrônica do grau de transparência do sistema de referência

	Sincronias				
	19-1	19-2	20-1	20-2	21-1
Grau de transparência	X ₁ %	X ₂ %	X ₃ %	X ₄ %	X ₅ %

Para chegar ao grau de transparência do sistema de referência em cada sincronia (por exemplo, o valor X₁% da tabela acima), em primeiro lugar consideramos, separadamente, cada pessoa do discurso e, em seguida, fizemos uma média entre os graus de transparência no uso dessas diferentes pessoas, conforme ilustra a tabela genérica seguinte (tabela 4). É

³⁶ Por exemplo, em uma peça teatral integrante do *corpus*, num determinado momento, um personagem cita brevemente, em discurso direto, a fala de determinado sujeito de dialeto claramente diferente do utilizado por todos os outros personagens da peça, e esse sujeito citado usa o pronome *nós* com simplificação da morfologia verbal (*nós canta*).

necessário fazer essa média porque o grau de transparência do sistema de referência no uso de cada pessoa do discurso é diferente do grau de transparência no uso das demais pessoas. Essa diferença ocorre porque os paradigmas de desinência verbal número-pessoal (relativamente a presença ou ausência de desinência explícita) são diferentes entre uma pessoa e outra, como mostrado no quadro 6 acima. Por exemplo, com o pronome *tu*, todos os tempos/modos têm essa desinência explícita, enquanto, com o pronome *eu*, três tempos/modos têm desinência explícita e os demais tempos/modos não a têm.

Tabela 4: Dados a serem obtidos para cálculo do grau de transparência do sistema de referência em cada sincronia

Pessoas do discurso	Sincronias				
	19-1	19-2	20-1	20-2	21-1
1PS	a %	e %	i %	m %	q %
2PS	b %	f %	j %	n %	r %
1PP	c %	g %	k %	o %	s %
2PP	d %	h %	l %	p %	t %
Graus médios de transparência	$X_1\%$ (a+b+c+d/4)	$X_2\%$ (e+f+g+h/4)	$X_3\%$ (i+j+k+l/4)	$X_4\%$ (m+n+o+p/4)	$X_5\%$ (q+r+s+t/4)

Em cada sincronia, para se chegar ao grau de transparência no uso de cada pessoa do discurso (por exemplo, o valor *a%* da tabela 4 acima), é necessário um cálculo particular.

O grau de transparência do sistema de referência no domínio específico da **1PS** depende de dois fatores:

- (i) frequência de expressão e frequência de não-expressão pronominal do argumento-sujeito;
- (ii) expressão e não-expressão do argumento-sujeito por desinência número-pessoal explícita nos diferentes tempos/modos verbais.

No âmbito da 1PS, é necessário considerar que o grau de transparência do sistema de referência será diferente entre diferentes tempos/modos verbais. Essa diferença ocorre porque, embora consideremos uma mesma frequência de expressão pronominal do

argumento-sujeito para todos os tempos/modos da 1PS, certos tempos/modos têm desinência número-pessoal explícita, enquanto outros não. Assim, para calcular esse grau, seguimos o procedimento ilustrado na tabela 5 genérica dada a seguir.

Tabela 5: Modelo para cálculo do grau de transparência do sistema de referência no domínio da 1PS.

Tempo/modo	(+) ou (-) desinência	Frequência de não-expressão pronominal	Frequência de expressão pronominal ³⁷	Grau de opacidade	Grau de Transparência
Pres. Ind.	+	<i>n</i> %	<i>p</i> %	<i>p</i> %	<i>n</i> %
Pret. imp. Ind.	-	<i>n</i> %	<i>p</i> %	<i>n</i> %	<i>p</i> %
Pret. perf. Ind.	+	<i>n</i> %	<i>p</i> %	<i>p</i> %	<i>n</i> %
Pret. m. q. perf. Ind.	-	<i>n</i> %	<i>p</i> %	<i>n</i> %	<i>p</i> %
Fut. pres. Ind.	+	<i>n</i> %	<i>p</i> %	<i>p</i> %	<i>n</i> %
Fut. pret. Ind.	-	<i>n</i> %	<i>p</i> %	<i>n</i> %	<i>p</i> %
Pres. Subj.	-	<i>n</i> %	<i>p</i> %	<i>n</i> %	<i>p</i> %
Pret. imp. Subj.	-	<i>n</i> %	<i>p</i> %	<i>n</i> %	<i>p</i> %
Fut. Subj.	-	<i>n</i> %	<i>p</i> %	<i>n</i> %	<i>p</i> %
MÉDIA				$[(p \cdot 3) + (n \cdot 6)] / 9$	$[(n \cdot 3) + (p \cdot 6)] / 9$

Como mostra a tabela, fazemos uma média entre os graus de transparência no uso de 1PS nos diferentes tempos/modos (como mostra a última linha da tabela). Num tempo/modo em que não há desinência, o que é indicado na tabela por “-” (por exemplo, no pretérito imperfeito do indicativo), a referência será transparente sempre que houver expressão pronominal do argumento-sujeito. Assim, nos tempos/modos sem desinência explícita, o grau de transparência será igual ao valor da frequência de expressão pronominal do argumento-sujeito (valor *p*% na tabela acima). Inversamente, num tempo/modo em que há desinência explícita, o que é indicado na tabela por “+” (por exemplo, no presente do indicativo), a referência será transparente sempre que não houver expressão pronominal do argumento-sujeito. Assim, nos tempos/modos com desinência explícita, o grau de transparência será igual ao valor da frequência de não-expressão pronominal do argumento-sujeito (valor *n*% na tabela acima).³⁸

³⁷ Cabe esclarecer que, nessa tabela, os valores *n* e *p* somam 100%, já que se referem, respectivamente, aos valores de não-expressão pronominal e expressão pronominal do argumento-sujeito, em uma dada sincronia.

³⁸ Por razões explicadas acima, em nosso levantamento de dados sobre as frequências de expressão/não-expressão pronominal do argumento-sujeito, não controlamos a variável tempo/modo verbal. Assim, para cada

Considerando agora o grau de transparência do sistema de referência no que se circunscreve particularmente ao uso da **2PS**, três fatores devem ser considerados:

- (i) frequência de expressão e frequência de não-expressão pronominal do argumento-sujeito;
- (ii) expressão e não-expressão do argumento-sujeito por desinência número-pessoal explícita nos diferentes tempos/modos verbais referentes ao pronome *tu* e referentes ao pronome *você*;
- (iii) frequências de uso do pronome *tu* e do pronome *você* como expressão do argumento-sujeito.

Primeiramente é preciso calcular o grau de transparência do sistema de referência relativamente ao uso do pronome *tu* e o grau de transparência no uso do pronome *você*, separadamente, para depois fazer uma média entre esses dois graus, considerando as frequências de uso de *tu* e de *você*.

No caso de *você*, assim como no caso da 1PS, o grau de transparência depende das frequências de expressão e de não-expressão pronominal do argumento-sujeito na sincronia em análise em combinação com a expressão ou não-expressão do argumento-sujeito por desinência número-pessoal explícita nos diferentes tempos/modos verbais. Desse modo, para calcular o grau de transparência da referência com o pronome *você*, seguimos um procedimento similar ao explicado acima para a 1PS. Fazemos uma média entre o grau de transparência da referência com *você* nos diferentes tempos/modos (no caso da 1PS, esse cálculo foi mostrado na última linha da tabela). Nos tempos/modos em que não há desinência explícita (todos, com exceção do pretérito perfeito), a referência será transparente sempre que houver expressão pronominal do argumento-sujeito. Assim, nos tempos/modos sem

peessoa do discurso e para cada sincronia, chegamos a um índice de expressão e a um índice de não-expressão pronominal que generalizamos para todos os tempos/modos verbais, isto é, em cada sincronia e em cada pessoa, estamos considerando um mesmo índice de expressão ($p\%$) e um mesmo índice de não-expressão pronominal ($n\%$) para todos os tempos/modos verbais.

desinência explícita, o grau de transparência da referência será igual ao valor da frequência de expressão pronominal do argumento-sujeito. Inversamente, no tempo/modo em que há desinência explícita (no caso, o pretérito perfeito), a referência será transparente sempre que não houver expressão pronominal do argumento-sujeito. Assim, nesse tempo/modo, o grau de transparência será igual ao valor da frequência de não-expressão pronominal do argumento-sujeito.

No caso de *tu*, não há necessidade de identificar, separadamente, os graus de transparência da referência em cada tempo/modo e depois fazer uma média entre eles, já que, em todos os tempos/modos, o grau de transparência da referência será o mesmo, porque, com o pronome *tu*, todos os tempos/modos apresentam desinência número-pessoal (como mostra o quadro 6) e porque, conforme explicamos acima, para cada pessoa, em cada sincronia, iremos considerar uma mesma frequência de expressão pronominal do argumento-sujeito para todos os tempos/modos verbais. Dessa forma, com o pronome *tu*, o grau de transparência da referência será igual ao valor da frequência de não-expressão pronominal do argumento-sujeito.

Uma vez obtidos os graus de transparência com *tu* e com *você*, não basta, para se chegar ao grau de transparência com a 2PS, fazer apenas uma média simples entre o grau de transparência com o pronome *tu* e o grau com o pronome *você*, já que há diferença entre a frequência de uso de *tu* e a frequência de uso de *você*. Essa diferença na frequência de uso de um e outro pronome irá influenciar o grau de transparência do sistema de referência no caso da 2PS, fazendo com que o pronome mais usado tenha mais peso sobre o grau geral de transparência do sistema de referência no que tange ao uso da 2PS. Por essa razão, para chegar a esse grau geral, fazemos uma média ponderada entre o grau de transparência da referência com *tu* e o grau com *você*, considerando, como pesos, as frequências de uso do pronome *tu* e do pronome *você* na sincronia em análise.

No domínio da **1PP**, o grau de transparência do sistema de referência depende também de três fatores:

- (i) frequência de expressão e frequência de não-expressão pronominal do argumento-sujeito;
- (ii) expressão e não-expressão do argumento-sujeito por desinência número-pessoal explícita nos diferentes tempos/modos verbais referentes às formas pronominais *nós* e *a gente*;
- (iii) frequências de uso das formas *nós* e *a gente* como expressão do argumento-sujeito.

Na análise da 1PP, o procedimento é semelhante ao adotado para cálculo da 2PS. Inicialmente, calculamos, separadamente, o grau de transparência da referência com *nós* e o grau com *a gente*. Em seguida, obtemos o grau geral de transparência da referência com a 1PP por meio da média ponderada entre os graus de transparência da referência com *nós* e com *a gente*, usando, como pesos, as frequências de uso de cada um desses pronomes, na sincronia em análise.

Finalmente, o grau de transparência do sistema de referência relativamente ao uso da **2PP** depende de dois fatores:

- (i) frequência de expressão e frequência de não-expressão pronominal do argumento-sujeito;
- (ii) expressão e não-expressão do argumento-sujeito por desinência número-pessoal explícita nos diferentes tempos/modos verbais.

No cálculo da 2PP, é preciso considerar que a variação entre *vós* e *vocês* não influencia na realização transparente ou opaca da referência, já que a questão da presença ou ausência de desinência número-pessoal explícita nos diferentes tempos/modos é igual no caso

de ambos esses pronomes – em todos os tempos/modos há desinência explícita (cf. quadro 6 acima). A exemplo do caso de *tu* e de *nós*, no caso de *vós/vocês*, como todos os tempos/modos têm desinência, não é necessário fazer uma média entre os graus de transparência nos diferentes tempos/modos. Com *vós* e *vocês*, em todos os tempos/modos verbais, a referência será opaca se houver expressão pronominal do argumento-sujeito e será transparente se não houver. Dessa forma, com os pronomes da 2PP, o grau de transparência do sistema de referência será igual à frequência de não-expressão pronominal do argumento-sujeito.

Na realização de todos esses cálculos, as frequências de expressão e não-expressão pronominal do argumento-sujeito, as frequências de uso de *tu* e de *você* e as frequências de uso de *nós* e *a gente* foram obtidas, para cada sincronia, de acordo com os critérios explicados acima (sendo esses dados apresentados no capítulo de análise, na seção 5.1), e a expressão ou não-expressão do argumento-sujeito por desinência número-pessoal explícita nos diferentes tempos/modos verbais foi considerada de acordo com o quadro 6 acima.

Após a realização desses cálculos (que são detalhadamente desenvolvidos no capítulo 5, na seção 5.2), pudemos chegar aos graus percentuais necessários para preencher a tabela 3 acima, que trata da variação diacrônica do grau de transparência do sistema de referência por expressão do argumento-sujeito de 1ª e 2ª pessoas no PB, o que, enfim, constitui o objeto central de análise deste trabalho. A partir de então, nossa pesquisa se voltou para uma análise qualitativa desse resultado, que é apresentada nas seções 5.3 e 5.4.

4.2 *Corpus*

O *corpus* deste trabalho compõe-se de peças teatrais produzidas do século XIX ao século XXI. De acordo com nosso entendimento, os fenômenos aqui em estudo (aumento na

frequência de expressão pronominal do argumento-sujeito, em oposição à não-expressão pronominal, aumento na frequência de uso de *você* em detrimento de *tu* e aumento na frequência de uso de *a gente* em detrimento de *nós*) podem ser mais bem detectados em gêneros textuais orais. Por essa razão, e tendo em vista a dificuldade (ou até mesmo impossibilidade, em alguns casos) de encontrar amostras de língua falada para estudo diacrônico, optamos pela seleção de peças de teatro para compor o *corpus*, assim como Duarte (1993), como uma forma de analisar textos mais próximos da modalidade falada (assumindo que peças de teatro, em geral, tendem a reproduzir a língua falada de sua época).

O quadro 7 a seguir apresenta a totalidade das peças utilizadas (seus respectivos autores e anos de produção) no levantamento de dados de cada uma das sincronias analisadas.

Quadro 7: Peças teatrais integrantes do *corpus* da pesquisa

Ano	Peça	Autor
Sincronia 19-1		
1842	<i>Os dois, ou o inglês maquinista</i> (OIM)	Martins Pena
1844	<i>Judas em sábado de aleluia</i> (JSA)	Martins Pena
1844	<i>O dileitante</i> (ODI)	Martins Pena
1844	<i>O namorador</i> (ONA)	Martins Pena
1844	<i>Os irmãos das Almas</i> (OIA)	Martins Pena
1844	<i>Os três médicos</i> (OTM)	Martins Pena
1845	<i>Angelica e Firmino</i> (AEF)	Araújo Porto Alegre
1845	<i>As casadas solteiras</i> (ACS)	Martins Pena
1845	<i>As desgraças de uma criança</i> (ADC)	Martins Pena
1845	<i>O cigano</i> (OCI)	Martins Pena
1845	<i>O noviço</i> (ONO)	Martins Pena
1845	<i>Os meirinhos</i> (OME)	Martins Pena
1845	<i>Quem casa quer casa</i> (CQC)	Martins Pena
1846	<i>O usurário</i> (OUS)	Martins Pena
s.d. ³⁹	<i>O jogo das prendas</i> (JDP)	Martins Pena
1849	<i>Januário Garcia ou o sete orelhas</i> (JAG)	Martim Francisco Ribeiro de Andrada
Sincronia 19-2		
1858	<i>Azas de um Anjo</i> (ADA)	José de Alencar
1858	<i>O demônio familiar</i> (ODF)	José de Alencar
1860	<i>Luxo e vaidade</i> (LEV)	Joaquim M. de Macedo
1861	<i>Desencantos</i> (DES)	Machado de Assis
1861	<i>Meia hora de cinismo</i> (MHC)	França Júnior
1863	<i>As forças caudinas</i> (AFC)	Machado de Assis

³⁹ A peça em questão pertence à sincronia 19-1 (cf. DAMASCENO, 1956), embora não seja possível precisar a data, dentro dessa sincronia, de modo mais exato.

1863	<i>O Caminho da porta</i> (OCP)	Machado de Assis
1863	<i>O Protocolo</i> (PRO)	Machado de Assis
1864	<i>Quase ministro</i> (QUM)	Machado de Assis
1864	<i>Verso e Reverso</i> (VER)	José de Alencar
1865	<i>A torre em concurso</i> (TEC)	Joaquim M. de Macedo
1869	<i>Romance de uma velha</i> (RDV)	Joaquim M. de Macedo
1870	<i>O Defeito de família</i> (DDF)	França Júnior
1870	<i>Remissão de pecados</i> (REP)	Joaquim M. de Macedo
1883	<i>Cahio o ministério</i> (COM)	França Júnior
1884	<i>O escravocrata</i> (OES)	Arthur Azevedo
1885	<i>A lotação dos bonds</i> (ALB)	França Júnior
1888	<i>Almanjarra</i> (ALM)	Artur Azevedo
1899	<i>Não consultes médico</i> (NCM)	Machado de Assis
1899	<i>Os milagres de S. José de Riba-Mar</i> (MSJ)	Americo Azevedo
Sincronia 20-1		
1907	<i>O Dote</i> (ODO)	Artur Azevedo
1921	<i>Onde canta o sabiá</i> (OCS)	Gastão Tojeiro
1921	<i>Ministro do supremo</i> (MDS)	Armando Gonzaga
1922	<i>A flor dos maridos</i> (FDM)	Armando Gonzaga
1928	<i>Balduino</i> (BAL)	Armando Gonzaga
1929	<i>As noivas</i> (ANO)	Paulo Gonçalves
1933	<i>Deus lhe pague</i> (DLP)	Joracy Camargo
1936	<i>Álbum de família</i> (ADF)	Nelson Rodrigues
1943	<i>Vestido de noiva</i> (VDN)	Nelson Rodrigues
Sincronia 20-2		
1960	<i>A invasão</i> (AIN)	Dias Gomes
1965	<i>O berço do herói</i>	Dias Gomes
1978	<i>Opera do malandro</i> (ODM)	Chico Buarque
1979	<i>A fábrica de chocolate</i> (FDC)	Mário Prata
1979	<i>Os órfãos de Jânio</i> (ODJ)	Millôr Fernandes
1982	<i>Duas tábuas e uma paixão</i> (DTP)	Millôr Fernandes
1983	<i>Crime e impunidade</i> (CEI)	Roberto Athayde
1989	<i>Doces fragmentos de loucura</i> (DFL)	Isis Baião
Sincronia 21-1		
2001	<i>E erámos todos Thunderbirds</i> (ETT)	Mario Bortolloto
2001	<i>Intimidade Indecente</i> (INI)	Leilah Assumpção
2001	<i>Sobre a arte de cortar bifés</i> (ACB)	Hugo Possolo
2002	<i>Aquele ano das marmitas</i> (ADM)	Naum Alves de Souza
2002	<i>Capitanias Hereditárias</i> (CAH)	Maria Carmen Barbosa e Miguel Falabella
2003	<i>Ilmo. Sr.</i> (ILM)	Naum Alves de Souza
2003	<i>Síndromes</i> (SIN)	Maria Carmen Barbosa e Miguel Falabella
2007	<i>Novas diretrizes em tempos de paz</i> (NDT)	Bosco Brasil
2008	<i>Ilustríssimo Filho da mãe</i> (IFM)	Leilah Assumpção
2012	<i>Nóis Otário(s)</i> (NOO)	Hugo Possolo

Para a seleção das peças do *corpus*, consideramos os seguintes critérios:

- (i) selecionamos apenas obras oficialmente publicadas (descartando versões não oficiais disponibilizadas na internet ou em outros meios de divulgação). Com essa medida,

pretendemos evitar o uso de textos com linguagem modificada ou adaptada para fins específicos, como fins escolares etc.;

- (ii) selecionamos, sempre que possível, a primeira edição da obra ou uma versão próxima da primeira edição na tentativa de ter acesso a uma versão mais fiel àquela originalmente produzida pelo autor em sua época. Em caso de indisponibilidade das próprias obras publicadas, utilizamos trabalhos de estudiosos que compilaram as obras originais e optamos, nesse caso, por estudiosos que prezavam por manter a linguagem original dos autores;
- (iii) não selecionamos peças que retratavam um tempo diferente do tempo em que foi escrita. Ou seja, excluímos, por exemplo, peças como *Amador Bueno*, de Joaquim Norberto de Souza e Silva, que, embora tenha sido escrita em 1843, retrata o período de 1642. Essa restrição foi adotada por considerarmos a possibilidade de que o autor pudesse ter feito algum tipo de modificação no texto para adequar a linguagem da peça ao período que ela estaria retratando. Por razões similares, também excluímos peças de autores brasileiros cuja trama não se passa no Brasil. Nesse caso, o receio foi de que pudesse haver algum tipo de adaptação da linguagem da obra ao local (país) em que a trama se desenvolve. Por esse motivo não selecionamos, por exemplo, peças como *Leonor de Mendonça*, de Gonçalves Dias, cuja trama se desenvolve em Portugal;
- (iv) não selecionamos peças escritas completamente em verso (como *Noite de S. João*, de José de Alencar) ou, no caso de peças que continham trechos em verso, não coletamos dados desses trechos. Essa medida visou a evitar a coleta de dados em contextos em que, por questões de ritmo, métrica, rima etc., houvesse algum tipo de (des)favorecimento de uso de uma ou outra forma em análise neste trabalho.

Optamos por constituir nosso próprio *corpus* de análise com textos sempre de um mesmo gênero, inclusive nos séculos mais recentes (em que seria possível encontrar outros tipos de amostra de fala), como forma de manter a uniformidade da comparação entre os séculos, impedindo a interferência de fatores relacionados à diversidade de gêneros textuais na análise de dados.

Certamente, a seleção desse gênero textual/*corpus*, assim como de qualquer outro, não está livre de impasses metodológicos. Em nosso caso, os principais impasses parecem estar relacionados com a possível interferência do estilo do autor da peça nos fenômenos linguísticos em estudo, fazendo com que, por exemplo, uma determinada peça não seja representativa do recorte de tempo considerado, ou com que essa peça seja representativa apenas da fala de uma determinada região do país (caso o autor seja carioca, por exemplo). Para tentar amenizar essas questões, procuramos diversificar os autores e as peças integrantes do *corpus* (ver quadro 7 acima), coletando, para cada um dos fenômenos estudados, no máximo 25 dados de cada peça e, sempre que possível, de autores diferentes, de forma a favorecer a diversidade de dados.

A diversidade de autores, no entanto, não foi possível na sincronia 19-1. As peças desse período a que se tem acesso nos dias atuais são, em sua quase totalidade, de um único autor: Martins Pena. Após investigação bibliográfica específica sobre textos teatrais do período, chegamos a identificar e ter acesso a algumas outras peças de outros autores, porém a grande maioria dessas peças não pôde ser utilizada por não se adequar aos critérios definidos neste trabalho (elencados acima). Assim, não incluímos aqui, por exemplo, peças como *O cinco de dezembro de 1833* ou *O cônego Ignez*, de O Xará, escrita em 1835, por estar em verso, *Olgiato*, de Gonçalves de Magalhães, escrita em 1839, por retratar a cidade de Milão no período de 1476, *Caetaninho, ou o tempo colonial*, de Paulo Antonio do Vale, escrita em 1848, por retratar o período de 1781, entre outras peças com impedimentos de uso de tipo semelhante. Como resultado dessa situação, as peças disponíveis para a coleta de dados nessa sincronia não foram suficientes para a coleta prevista de 100 ocorrências de um dos fenômenos estudados (variação *nós/a gente*), e, nesse caso, os dados procedem, majoritariamente, de peças de um mesmo autor (Martins Pena).

Além disso, em algumas sincronias, não foi possível coletar, para o estudo da

variação *tu/você* e/ou *nós/a gente*, o total de 25 ocorrências numa mesma peça. De fato, algumas peças das sincronias 19-1 e 19-2, por exemplo, forneceram um número de ocorrências bastante inferior a 25: na peça *Januario Garcia ou o sete orelhas*, da sincronia 19-1, foram encontrados apenas 7 casos de *tu/você*; na peça *Meia hora de cinismo*, da sincronia 19-2, apenas 2 casos de *nós/a gente* etc. Essa coleta de pequena quantidade de dados por peça, que ocorreu mais acentuadamente nas sincronias 19-1 e 19-2, deve-se a dois fatores principais. O primeiro deles é o fato de o índice de expressão pronominal do argumento-sujeito (em todas as pessoas de maneira geral) ser mais baixo nessas duas sincronias do que nas sincronias subsequentes.⁴⁰ Como já dissemos na explicação sobre a metodologia para coleta de dados referente às variações *tu/você* e *nós/a gente*, para o estudo desses fenômenos coletamos apenas casos de expressão pronominal do argumento-sujeito, desconsiderando ocorrências de não-expressão pronominal (sujeito nulo). Assim, naturalmente, o baixo índice de expressão pronominal do argumento-sujeito das sincronias em questão implica em um baixo índice de dados sobre *tu/você* e sobre *nós/a gente* a serem coletados por peça.

Essa dificuldade para encontrarmos quantidades mínimas suficientes de dados nas peças das sincronias 19-1 e 19-2, inclusive, é responsável por utilizarmos um número maior de peças nessas duas sincronias (16 peças para a sincronia 19-1 e 20 peças para a sincronia 19-2, como se pode ver no quadro 7 acima), na tentativa de encontrarmos a mesma quantidade total de ocorrências (100 ocorrências de *tu/você* e 100 ocorrências de *nós/a gente*) levantadas nas outras sincronias (nas sincronias 20-1, 20-2 e 21-1, números menores de peças foram suficientes para o levantamento dessa quantidade de dados).

O segundo fator que acarretou um baixo número de casos coletados por peças está diretamente relacionado ao fenômeno da variação *nós/a gente*. Para a coleta desses dados, houve uma dificuldade extra: muitas peças, por questões relacionadas à própria trama da peça,

⁴⁰ Ver tabela 6 no capítulo de análise de dados (Capítulo 5).

não apresentavam contextos propícios para o uso de 1PP, ou seja, nessas peças, as interações dialógicas envolviam apenas (ou principalmente) a 1PS e a 2PS. Dessa forma, a coleta de 25 ocorrências de *nós/a gente* em cada peça muitas vezes não era possível. Por exemplo, muitas peças das sincronias 19-1 e 19-2 continham apenas uma ocorrência do fenômeno. Na sincronia 19-1, essa dificuldade, somada à dificuldade de encontrarmos expressão pronominal do argumento-sujeito de 1PP e à dificuldade de encontrarmos peças em quantidades suficientes (do período em questão), impossibilitou o levantamento das 100 ocorrências de *nós/a gente* idealizadas inicialmente. Após reunirmos todas as peças disponíveis para coleta de dados de variação *nós/a gente* nessa sincronia, chegamos a um total máximo de 58 dados.

Em resumo, para fornecer uma melhor visualização de como foi feito nosso levantamento de dados, o quadro 8 abaixo mostra as peças utilizadas para cada levantamento e a quantidade de dados extraída de cada peça.⁴¹

Quadro 8: Peças teatrais utilizadas no levantamento de dados geral e quantidades de dados coletados por peça.

Ano	Peça	Expressão/não-expressão pronominal do argumento-sujeito	Variação tu/você	Variação nós/a gente
19-1				
1842	Os dois, ou o inglês maquinista	21	22	03
1844	Judas em sábado de aleluia	24	12	01
1844	O diletante	-	-	01
1844	O namorador	-	-	01
1844	Os irmãos das Almas	22	14	02
1844	Os três médicos	-	-	02
1845	Angelica e Firmino	23	25	11
1845	As casadas solteiras	23	-	15
1845	As desgraças de uma criança	23	-	03
1845	O cigano	21	-	06
1845	O noviço	22	20	07
1845	Os meirinhos	-	-	01
1845	Quem casa quer casa	-	-	01
1846	O usurário	-	-	02
s.d.	O jogo das prendas	-	-	01
1849	Januário Garcia ou o sete orelhas	21	07	01
19-1	TOTAL DE OCORRÊNCIAS	200	100	58

⁴¹ Para cada fenômeno em análise, o levantamento de dados, em cada peça, foi realizado seguindo a ordem de ocorrência dos dados, desde o início da peça até seu final ou até completar a quantidade de dados determinada, sem excluir nenhum dado (atendendo, naturalmente, às restrições estabelecidas para o levantamento de cada fenômeno). Esse procedimento foi adotado como uma tentativa de fazer com que o levantamento retratasse a realidade dos fatos linguísticos em estudo.

19-2				
1858	Azas de um Anjo	-	12	11
1858	O demônio familiar	02	25	15
1860	Luxo e vaidade	-	22	16
1861	Desencantos	-	-	02
1861	Meia hora de cinismo	23	08	02
1863	As forças caudinas	-	-	01
1863	O Caminho da porta	-	-	01
1863	O Protocolo	-	-	01
1864	Quase ministro	-	-	01
1864	Verso e Reverso	25	04	04
1865	A torre em concurso	25	10	11
1869	Romance de uma velha	25	-	03
1870	O Defeito de família	-	05	05
1870	Remissão de pecados	25	-	03
1883	Cahio o ministério	25	13	07
1884	O escravocrata	-	-	03
1885	A lotação dos bonds	25	01	01
1888	Almanjarra	25	-	02
1899	Não consultes médico	-	-	03
1899	Os milagres de São José de Riba-Mar	-	-	08
19-2	TOTAL DE OCORRÊNCIAS	200	100	100
20-1				
1907	O Dote	25	-	-
1921	Onde canta o sabiá	25	25	23
1921	Ministro do supremo	25	22	12
1922	A flor dos maridos	-	-	05
1928	Balduino	25	-	17
1929	As noivas	25	03	02
1933	Deus lhe pague	25	25	18
1936	Álbum de família	25	25	11
1943	Vestido de noiva	25	-	12
20-1	TOTAL DE OCORRÊNCIAS	200	100	100
20-2				
1960	A invasão	25	25	25
1965	O berço do herói	25	-	-
1978	Opera do malandro	25	25	25
1979	A fábrica de chocolate	25	-	-
1979	Os órfãos de Jânio	25	-	24
1982	Duas tábuas e uma paixão	25	-	01
1983	Crime e impunidade	25	25	12
1989	Doces fragmentos de loucura	25	25	13
20-2	TOTAL DE OCORRÊNCIAS	200	100	100
21-1				
2001	E erámos todos <i>Thunderbirds</i>	25	-	-
2001	Intimidade Indecente	25	-	25
2001	Sobre a arte de cortar bifés	04	-	-
2002	Aquele ano das marmitas	25	25	25
2002	Capitanias Hereditárias	25	25	25
2003	Ilmo. Sr.	25	-	-
2003	Síndromes	25	-	10
2007	Novas diretrizes em tempos de paz	02	-	-
2008	Ilustríssimo Filho da mãe	21	25	15
2012	Nóis Otário(s)	23	25	-
21-1	TOTAL DE OCORRÊNCIAS	200	100	100
TOTAL GERAL DE OCORRÊNCIAS				
	1.000	500	458	

Em acordo com as informações do quadro 7, ao final das ocorrências exemplificativas de nossas análises, serão dadas as seguintes informações sobre elas: século – metade a que pertence, sigla que identifica a peça e indicação da página de onde foi extraída. Por exemplo, a indicação “(19-1 ADC, p. 3)” significa que a ocorrência pertence à primeira metade do século XIX e se encontra na peça *As desgraças de uma criança*, à página 3.

CAPÍTULO V

Análise de dados

Neste capítulo, apresentamos a análise de dados de nossa pesquisa. Inicialmente, na seção 5.1, apresentamos separadamente os resultados do levantamento de dados de cada uma das três mudanças diacrônicas em questão (frequência de expressão pronominal do argumento-sujeito; frequência de uso dos pronomes *você/tu*; frequência de uso de *a gente/nós*). Em seguida, em 5.2, a partir desses dados, identificamos o grau de transparência do sistema de referência por expressão pronominal e desinencial do argumento-sujeito de 1ª e 2ª pessoas no PB, em cada sincronia aqui considerada, chegando, então, à variação diacrônica do grau de transparência desse sistema.

Nas duas seções seguintes, empreendemos uma análise qualitativa a respeito do resultado obtido sobre a diacronia da transparência desse sistema. Em 5.3, argumentamos em favor do comportamento não-unidirecional desse processo e, em 5.4, relacionamos esse comportamento à atuação de motivações comunicativas em interação, particularmente iconicidade e economia.

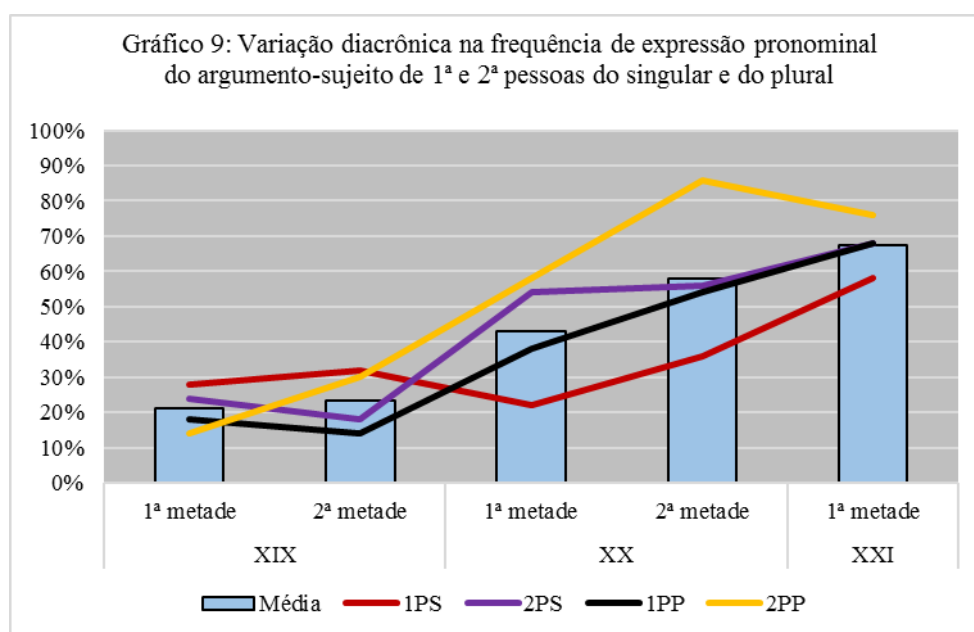
5.1 Dados sobre mudanças diacrônicas interferentes na transparência linguística

A primeira mudança diacrônica aqui considerada leva em conta o aumento na frequência de expressão pronominal do argumento-sujeito (tendo em vista 1ª e 2ª pessoas do discurso no singular e no plural). Os dados quantitativos levantados em nossa pesquisa quanto a essa mudança são sintetizados na tabela 6 a seguir.

Tabela 6: Variação diacrônica nas frequências de expressão e não-expressão pronominal do argumento-sujeito de 1ª e 2ª pessoas do singular e do plural

Pessoa do discurso	Expressão pronominal	19-1	19-2	20-1	20-2	21-1
1PS	+	28% (14/50)	32% (16/50)	22% (11/50)	36% (18/50)	58% (29/50)
	-	72% (36/50)	68% (34/50)	78% (39/50)	64% (32/50)	42% (21/50)
2PS	+	24% (12/50)	18% (09/50)	54% (27/50)	56% (28/50)	68% (34/50)
	-	76% (38/50)	82% (41/50)	46% (23/50)	44% (22/50)	32% (16/50)
1PP	+	18% (09/50)	14% (07/50)	38% (19/50)	54% (27/50)	68% (34/50)
	-	82% (41/50)	86% (43/50)	62% (31/50)	46% (23/50)	32% (16/50)
2PP	+	14% (07/50)	30% (15/50)	58% (29/50)	86% (43/50)	76% (38/50)
	-	86% (43/50)	70% (35/50)	42% (21/50)	14% (07/50)	24% (12/50)
Média	+	21% (42/200)	23,5% (47/200)	43% (86/200)	58% (116/200)	67,5% (135/200)
	-	79% (158/200)	76,5% (153/200)	57% (114/200)	42% (84/200)	32,5% (65/200)

Para efeito de ilustração, os dados da tabela 6 referentes particularmente à frequência de expressão pronominal também podem ser visualizados no gráfico 9 abaixo.



Como pode ser notado, de modo geral, da primeira metade do século XIX até o início do século XXI, há um aumento na incidência de expressão pronominal do argumento-sujeito de 1ª e 2ª pessoas do singular e do plural. Em alguns pontos, há uma certa redução nessa incidência. Isso se verifica na 1PS, entre as sincronias 19-2 e 20-1, na 2PS, entre as sincronias 19-1 e 19-2, na 1PP, entre 19-1 e 19-2, e na 2PP, entre as sincronias 20-2 e 21-1. Porém, essa redução ocorre apenas nesses pontos e com diferenças percentuais relativamente

pequenas. O que se destaca em geral é o progressivo aumento de índice de expressão pronominal do argumento-sujeito, o que, aliás, se apreende melhor nas médias entre todas as pessoas (ver última linha da tabela 6 e colunas do gráfico 9), por meio das quais se verifica aumento crescente desse índice em todo o recorte diacrônico.

Esses resultados, tanto referentes às diferentes pessoas conjuntamente quanto referentes a cada pessoa separadamente, corroboram a constatação de diversos outros autores (DUARTE, 1993, 1995; KAISER, 2006, 2009; TARALLO 1993a, b) de que se verifica no PB o aumento diacrônico do uso de sujeito pleno (ou diminuição de sujeito nulo), conforme mostramos no capítulo 3. O foco de nosso trabalho não incide propriamente na constatação da existência desse fenômeno, mas na análise de seu impacto sobre o grau de transparência do sistema de referência do PB. De todo modo, como já apontamos anteriormente, nosso levantamento de dados diacrônicos sobre esse fenômeno parece constituir uma contribuição significativa para o tema, na medida em que, embora o reconhecimento de tal fenômeno pareça ser consensual entre os autores que lidam com o assunto, não chega a haver uma diversidade mais considerável de trabalhos que propiciem dados diacrônicos desse fenômeno.

A segunda mudança aqui estudada consiste na verificação do aumento na frequência de uso do pronome de 2PS *você*, em detrimento do pronome *tu*. A tabela 7 apresenta os dados quantitativos levantados em nossa pesquisa quanto a essa mudança.

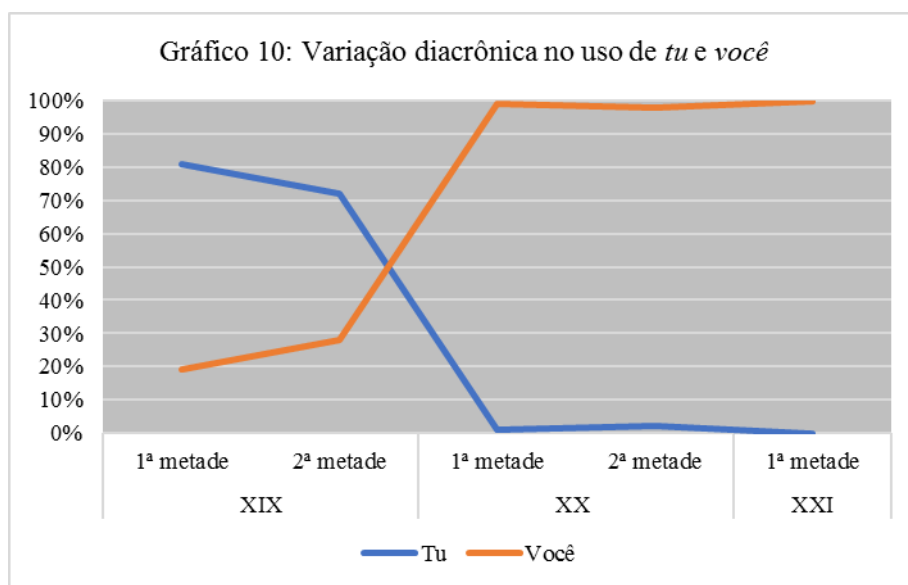
Tabela 7: Variação diacrônica nas frequências de uso de *tu* e *você*

Pronomes de 2PS	Sincronias				
	19-1	19-2	20-1	20-2	21-1
<i>tu</i>	81% (81/100)	72% (72/100)	1% (1/100)	2% (2/100)	0% (0/100) ⁴²
<i>você</i>	19% (19/100)	28% (28/100)	99% (99/100)	98% (98/100)	100% (100/100)
Total	100% (100/100)	100% (100/100)	100% (100)	100% (100)	100% (100/100)

Como é possível observar, também nesse caso há aumento, em termos gerais, na

⁴² Neste trabalho, não interpretamos o fato de não termos encontrado nenhuma ocorrência de *tu* na sincronia 21-1 como significando que esse pronome teria caído em desuso definitivo. Para nós, isso significaria, na verdade, que a frequência de uso desse pronome teria caído de forma muito expressiva.

frequência de uso do pronome *você*, da primeira metade do século XIX até a primeira metade do século XXI. Embora haja uma pequena diminuição da frequência do uso desse pronome na passagem da primeira para a segunda metade do século XX, essa diminuição ocorre apenas nessa transição entre sincronias e com uma diferença percentual insignificante (de 99% para 98%). O aumento na frequência de uso de *você* é, de fato, o que se destaca a partir das informações dispostas na tabela. Chama a atenção, inclusive, o aumento brusco do uso desse pronome na passagem da sincronia 19-2 para a 20-1 (de 28% para 99%). A partir da sincronia 20-1, o pronome se estabelece como o mais frequentemente usado, em lugar de *tu*, para referência à 2PS. Os dados da tabela acima podem ser também visualizados no gráfico 10.



Aqui também é possível notar que os dados que levantamos coincidem com os trabalhos de diferentes autores que analisam a mudança *tu/você* (SOUZA, 2012; MACHADO, 2006; LOPES; CAVALCANTE, 2011; LOPES, 2008, entre outros). Como mostrado no capítulo 3, tais autores, de modo geral, mostram a queda da frequência de *tu* e o aumento de *você*, a partir da primeira metade do século XIX, e a suplantação de *tu* por *você* no decorrer do século XX. Nossos dados também apontam na direção desse mesmo processo.

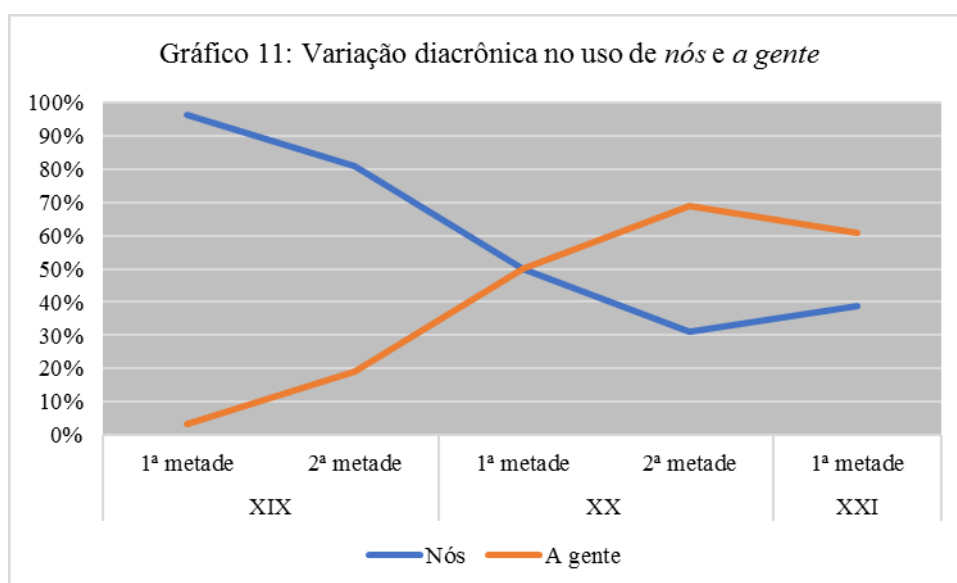
Por fim, a terceira mudança diacrônica aqui estudada consiste na verificação do

aumento da frequência de uso da forma pronominal *a gente* em detrimento do uso de *nós*. A tabela 8 apresenta os dados quantitativos levantados nesta pesquisa quanto a esse fenômeno.

Tabela 8: Variação diacrônica nas frequências de uso de *nós* e *a gente*

Formas pronominais de 1PP	Sincronias				
	19-1	19-2	20-1	20-2	21-1
<i>nós</i>	96,5% (56/58)	81% (81/100)	50% (50/100)	31% (31/100)	39% (39/100)
<i>a gente</i>	3,5% (2/58)	19% (19/100)	50% (50/100)	69% (69/100)	61% (61/100)
Total	100% (58)	100% (100/100)	100% (100/100)	100% (100/100)	100% (100/100)

A partir dos dados expostos na tabela 8, é possível observar, também aqui, de modo geral, um aumento, em termos de frequência, no uso da forma *a gente* ao longo da diacronia considerada (de 3,5% para 61%). Há uma ligeira diminuição no uso dessa forma na passagem da segunda metade do século XX para primeira metade do século XXI (de 69% para 61%), porém o que parece se destacar na diacronia considerada é o aumento expressivo no uso dessa forma. Na passagem da sincronia 19-1 para a 19-2 e na passagem da sincronia 19-2 para a 20-1, esse aumento é ainda mais significativo (respectivamente, de 3,5 para 19% e de 19% para 50%). Na sincronia 20-2, o uso da forma *a gente* já suplanta o uso de *nós*, para referência à 1PP. O gráfico 11 abaixo ilustra essas informações da tabela 8.



Desse modo, nossos dados evidenciam a mudança envolvendo *nós/a gente* assumida, conforme mostramos também no capítulo 3, por vários autores (RUBIO; GONÇALVES, 2010; SEARA, 2000; SANTANA, 2014; VIANNA; LOPES, 2012 etc.). Como explicado naquele capítulo, os trabalhos que admitem essa mudança fornecem dados sincrônicos, referentes ao período entre o final do século XX e início do século XXI. De modo geral, esses trabalhos apuram uma frequência de uso de *a gente* mais ou menos em torno de 70% nesse período. Assim, é interessante notar que esses dados coincidem com os nossos relativos à mesma época, em que encontramos taxas de 69% (20-2) e 61% (21-1) de emprego de *a gente*. Cumpre também destacar que, diante da ausência (ou escassez) de trabalhos propriamente diacrônicos sobre a mudança *nós/a gente*, o levantamento desse tipo de dado representa uma contribuição significativa de nosso estudo.

Apresentados, então, os dados diacrônicos sobre essas três mudanças, passamos agora para o cálculo do grau de transparência do sistema de referência de cada sincronia e para a identificação do desenvolvimento diacrônico do grau de transparência desse sistema.

5.2 Identificação da variação diacrônica do grau de transparência do sistema de referência

Conforme explicado anteriormente na metodologia, para calcular a variação diacrônica do grau de transparência do sistema de referência, nosso procedimento consiste em calcular o grau de transparência desse sistema (em termos percentuais) em cada sincronia e, em seguida, comparar uma sincronia com a outra, para demonstrar o que ocorreria diacronicamente com o grau de transparência.

Também conforme explicamos, dentro do sistema de referência, o grau de transparência é diferente entre uma e outra pessoa do discurso. Essa diferença ocorre porque os paradigmas de desinência verbal número-pessoal (relativamente a presença ou ausência de

desinência explícita) são diferentes de uma pessoa para outra e porque as frequências de expressão pronominal do argumento-sujeito são diferentes entre as diferentes pessoas. Desse modo, para chegar ao grau geral de transparência do sistema de referência de cada sincronia, identificamos o grau de transparência da referência com cada pessoa individualmente e, então, calculamos uma média entre os resultados dessas pessoas.

No que diz respeito, então, à **1PS**, o grau de transparência do sistema de referência, em cada sincronia, é influenciado pelas frequências de expressão e não-expressão pronominal do argumento-sujeito e pela presença ou não de desinência número-pessoal explícita nas formas verbais dos diferentes tempos/modos. Nos tempos/modos em que há desinência explícita de 1PS, a referência será opaca, se houver expressão pronominal do argumento-sujeito (como em (28)), e será transparente, se não houver expressão pronominal desse argumento (como em (29)); inversamente, nos tempos/modos em que não há desinência explícita de 1PS, a referência será transparente, se houver expressão pronominal do argumento-sujeito, e será opaca, se não houver (como em (30) e (31) respectivamente).

(28) AMBROSIO, *só de calça preta e chambre* – (...) Há oito anos, era eu pobre e miserável, e hoje sou rico, e mais ainda serei. O como não importa; no bom resultado está o mérito... Mas um dia pode tudo mudar. Oh, que **temo eu**? Se em algum tempo tiver que responder pelos meus atos, o ouro justificar-me-á e serei limpo de culpa... As leis criminaes fizeram-se para os pobres. (19-1 ONO, p. 5-6)

(29) MADALENA - A estas horas já muitos têm dormido o primeiro sono e vestem-se para a missa do galo; só eu... Ah, quem me dera poder ir também! É perder a idéia daí, que lhe acho muitos impossíveis. Cá o velho – o velho – (*ri-se*) e [a] senhora irmão, e ainda dormem; e eu nem durmo e nem vou. Mas **ficarei**, e ficando, muito bem sei o que hei-de fazer... (19-1 ADC, p. 531)

(30) LUIZA – (...) **Eu** já **podia** estar livre de tudo isto, se não fosse o maldito segredo que descobri. (19-1 OIA, p. 1)

(31) LUIZA – (...) Eu já podia estar livre de tudo isto, se não fosse o maldito segredo que descobri. Antes não **soubesse** de nada! (19-1 OIA, p. 1)

Na primeira metade do século XIX, como mostra a tabela 6 mais acima,⁴³ a frequência de expressão pronominal do argumento-sujeito de 1PS é de 28% e a frequência de não-expressão é de 72%. Sendo assim, no que diz respeito à 1PS, num tempo/modo verbal com desinência número-pessoal explícita, o sistema de referência apresentará um grau de opacidade de 28% (já que essa é a frequência de expressão pronominal do argumento-sujeito) e, naturalmente, um grau de transparência de 72% (o que corresponde justamente à frequência de não-expressão pronominal). Já num tempo/modo que não possui desinência número-pessoal explícita, o grau de transparência do sistema de referência será de 28% (justamente os 28% de frequência de expressão pronominal do argumento-sujeito), enquanto o grau de opacidade será de 72% (correspondente à frequência de não-expressão pronominal).

Desse modo, para se chegar ao grau de transparência do sistema de referência no que concerne à 1PS, pode-se calcular o grau médio, a partir dos graus de transparência que podem ser identificados nos diferentes tempos/modos verbais. A tabela 9 abaixo mostra nosso procedimento para calcular essa média. A tabela apresenta, na última coluna, os valores que estamos considerando como sendo os graus de transparência do sistema de referência no caso de cada tempo/modo verbal de 1PS, na primeira metade do século XIX, tendo em vista a ocorrência ou não de desinência número-pessoal e os percentuais de 28% e 72% de expressão e não-expressão pronominal do argumento-sujeito, respectivamente. A última linha da tabela mostra, então, o valor que obtemos como grau geral de transparência do sistema de referência relativamente à 1PS, na sincronia 19-1, calculado a partir da média simples entre os graus

⁴³ Da tabela 6, reproduzimos abaixo apenas os dados relevantes para a presente discussão. Adotaremos esse mesmo procedimento de recuperar a informação da tabela 6 nas discussões seguintes.

Cf. da tabela 6:

Pessoa do discurso	Expressão pronominal	19-1	19-2	20-1	20-2	21-1
1PS	+	28% (14/50)	32% (16/50)	22% (11/50)	36% (18/50)	58% (29/50)
	-	72% (36/50)	68% (34/50)	78% (39/50)	64% (32/50)	42% (21/50)

específicos considerados para cada tempo/modo verbal.⁴⁴

Tabela 9: Cálculo do grau de transparência do sistema de referência relativamente à 1PS na sincronia 19-1.

Tempo/modo	(+) ou (-) desinência	Frequência de não-expressão pronominal	Frequência de expressão pronominal	Grau de opacidade	Grau de Transparência
Ind. Pres.	+	72%	28%	28%	72%
Ind. Pret. imp.	-	72%	28%	72%	28%
Ind. Pret. perf.	+	72%	28%	28%	72%
Ind. Pret. m. q. perf.	-	72%	28%	72%	28%
Ind. Fut. pres.	+	72%	28%	28%	72%
Ind. Fut. pret.	-	72%	28%	72%	28%
Subj. Pres.	-	72%	28%	72%	28%
Subj. Pret. imp.	-	72%	28%	72%	28%
Subj. Fut.	-	72%	28%	72%	28%
MÉDIA				57,5%	42,5%

Conforme procuramos ilustrar na tabela acima, na sincronia 19-1, tendo em vista todos os tempos/modos verbais aqui considerados e as frequências de 28% de expressão pronominal e de 72% de não-expressão pronominal, chega-se a um valor médio de **42,5%** como sendo o grau de transparência do sistema de referência em relação à 1PS. O mesmo tipo de cálculo pode ser aplicado para identificação dos graus de transparência do sistema de referência relativamente à 1PS nas demais sincronias, o que resulta, para as sincronias 19-2, 20-1, 20-2 e 21-1, nos seguintes graus, respectivamente: **44%, 40,5%, 45,5%, 52,5%**.

Aqui convém fazer um esclarecimento sobre o que estamos entendendo como *grau de transparência do sistema de referência*. Consideramos o *potencial da língua* para a realização da referência de forma transparente, e não a ocorrência real de referências transparentes em textos específicos. Essa abordagem é válida no decorrer de todo este trabalho, para todos os momentos em que falamos em grau de transparência do sistema de referência. Ou seja, reconhecemos uma diferença entre o *potencial de transparência do sistema de referência* e a ocorrência real de referências transparentes em textos concretos, e

⁴⁴ Pelas razões explicadas no capítulo de Metodologia, estamos considerando aqui, para a 1PS, uma mesma frequência de expressão pronominal do argumento-sujeito para todos os tempos/modos verbais (e uma mesma frequência de não-expressão pronominal para todos os tempos/modos). O mesmo procedimento é adotado para cada uma das outras pessoas do discurso em análise.

focalizamos o primeiro caso.

Por exemplo, é importante levar em conta essa perspectiva no cálculo que acabamos de fazer (na tabela 9 acima) do grau de transparência do sistema de referência relativamente à 1PS. Como mostrado no cálculo, consideramos diferentes graus de transparência para os diferentes tempos/modos verbais, sendo o grau geral relativo à 1PS definido pela média simples entre os graus nos diferentes tempos/modos verbais. Como se vê, consideramos um mesmo peso para cada tempo/modo verbal. É certo que, na construção de textos, os diferentes tempos/modos são utilizados com diferentes frequências. Por exemplo, o presente do indicativo, normalmente, é mais utilizado do que o pretérito mais-que-perfeito. Assim, no cálculo do grau de transparência do sistema de referência em relação à 1PS, poder-se-ia pensar em levantar, diretamente num *corpus*, as frequências de referências opacas ou transparentes, mediante levantamento de *ocorrências reais* de orações com expressão ou não-expressão pronominal do argumento-sujeito e expressão ou não-expressão por desinência verbal explícita. Porém, a nosso ver, esse procedimento levaria a identificar a frequência de referências transparentes *em textos específicos*, e não o potencial de transparência do *sistema de referência da língua*. Se um texto qualquer apresenta mais formas verbais no presente do indicativo do que no pretérito mais-que-perfeito ou no futuro do subjuntivo, parece-nos que isso se deve a fatores contextuais, discursivos, pragmáticos da construção daquele texto, e não a uma característica decorrente da configuração gramatical da língua.

Recuperando a tabela 9 acima, consideramos que, para calcular o grau de transparência do sistema de referência em relação à 1PS, devemos comparar os graus de transparência nos diferentes tempos/modos verbais considerando, de forma abstrata, que todos eles ocorreriam num texto com a mesma frequência. Nesse sentido é que se pode fazer a média simples entre todos os tempos/modos para se obter o grau de transparência que o *sistema gramatical* da língua apresentaria.

Passamos agora a analisar o grau de transparência do sistema de referência em construções com a **2PS**. Nesse caso, há um fator extra em relação ao que ocorre com a 1PS, que é a variação entre os pronomes *tu* e *você*. Assim, para calcular o grau de transparência com a 2PS, levamos em conta as frequências de expressão e não-expressão pronominal do argumento-sujeito, a presença ou não de desinência número-pessoal explícita nos diferentes tempos/modos verbais, e a frequência de uso dos pronomes *tu* e *você*. Em primeiro lugar, calculamos, separadamente, o grau de transparência do sistema no uso do pronome *tu* e o grau de transparência no uso de *você*; em seguida, relacionamos esse resultado com as frequências de uso de *tu* e de *você*, para se chegar ao grau de transparência do sistema no âmbito da 2PS.

No caso do pronome *tu*, conforme mostra o quadro 6 (cf. capítulo 4) acima, todos os tempos/modos verbais apresentam desinência número-pessoal por meio de morfema explícito. Assim, com *tu*, em todos os tempos/modos verbais, a referência será opaca quando houver expressão pronominal do argumento-sujeito, e será transparente quando não houver. Portanto, em todos os tempos/modos verbais, o grau de transparência do sistema será o mesmo, o qual será equivalente ao valor da frequência de não-expressão pronominal do argumento-sujeito. Desse modo, com o pronome *tu*, diferentemente do que se coloca para a 1PS, não é necessário calcular uma média entre os graus de transparência do sistema nos diferentes tempos/modos verbais. Como mostra a tabela 6 acima,⁴⁵ na sincronia 19-1, a frequência de não-expressão pronominal com a 2PS é de 76%, valor, então, que irá constituir o grau de transparência do sistema no que se refere ao pronome *tu*.⁴⁶

No caso do pronome *você*, há diferença, entre os tempos/modos verbais, quanto à

⁴⁵ Cf. da tabela 6:

Pessoa do discurso	Expressão pronominal	19-1	19-2	20-1	20-2	21-1
2PS	+	24% (12/50)	18% (09/50)	54% (27/50)	56% (28/50)	68% (34/50)
	-	76% (38/50)	82% (41/50)	46% (23/50)	44% (22/50)	32% (16/50)

⁴⁶ Como explicado no capítulo de metodologia, em nosso levantamento de dados sobre sujeito pronominal pleno/nulo de 2PS, não fizemos levantamentos separados para *tu* e *você*, considerando aqui um mesmo percentual de sujeito pleno e um mesmo percentual de sujeito nulo para os dois pronomes.

presença ou ausência de desinência número-pessoal por morfema explícito: no pretérito perfeito do indicativo, há essa desinência explícita, enquanto, nos demais tempos/modos, não há. Por isso, o grau de transparência do sistema de referência no âmbito do pronome *você* será diferente entre o pretérito perfeito do indicativo e os demais tempos/modos, sendo necessário fazer uma média entre os graus nos diferentes tempos/modos para se chegar ao grau geral de transparência no que concerne ao pronome *você*.

Nosso cálculo para obtenção desse grau geral é apresentado na tabela 10 a seguir.

Tabela 10: Cálculo do grau de transparência do sistema de referência relativamente a *você* na sincronia 19-1.

Tempo/modo	(+) ou (-) desinência	Frequência de não-expressão pronominal	Frequência de expressão pronominal	Grau de opacidade	Grau de Transparência
Ind. Pres.	-	76%	24%	76%	24%
Ind. Pret. imp.	-	76%	24%	76%	24%
Ind. Pret. perf.	+	76%	24%	24%	76%
Ind. Pret. m. q. perf.	-	76%	24%	76%	24%
Ind. Fut. pres.	-	76%	24%	76%	24%
Ind. Fut. pret.	-	76%	24%	76%	24%
Subj. Pres.	-	76%	24%	76%	24%
Subj. Pret. imp.	-	76%	24%	76%	24%
Subj. Fut.	-	76%	24%	76%	24%
MÉDIA				70%	30%

No pretérito perfeito do indicativo (em que há desinência explícita, sinalizada por “+”), a referência será opaca quando houver expressão pronominal do argumento-sujeito e será transparente quando não houver. Nesse tempo, os graus de opacidade e de transparência da referência serão correspondentes, respectivamente, às frequências de expressão pronominal (24%) e de não-expressão pronominal (76%) na 2PS. Nos demais tempos/modos (em que não há desinência número-pessoal explícita, sinalizada por “-”), a referência será opaca quando não houver expressão pronominal e será transparente quando houver. Nesses tempos/modos, os graus de opacidade e transparência da referência serão equivalentes, respectivamente, às frequências de não-expressão (76%) e expressão pronominal (24%) na 2PS. A última linha da tabela 10 mostra a média do grau de transparência do sistema de referência relativamente ao

pronome *você*, calculado mediante a média simples entre os valores que consideramos como sendo os graus específicos em cada tempo/modo verbal.

Até aqui, identificamos os graus de transparência com o pronome *tu* (76%) e com o pronome *você* (30%), faltando, agora, relacionar esses dados com as frequências de uso dos pronomes *tu* e *você*, que, como mostra a tabela 7 da seção 5.1 acima,⁴⁷ são de 81% e 19%, respectivamente, na sincronia 19-1. Para isso, como explicado no capítulo de metodologia, calculamos uma média ponderada entre os graus de transparência com esses dois pronomes, considerando como pesos as frequências de uso desses pronomes, o que resulta no valor de **67,5%** como o grau médio de transparência do sistema de referência no domínio da 2PS.⁴⁸

Aplicando esse mesmo tipo de cálculo para as demais sincronias analisadas, são obtidos os seguintes graus de transparência do sistema de referência no âmbito da 2PS: 19-2: **66%**; 20-1: **53%**; 20-2: **54,5%**; 21-1: **64%**.

Passamos agora a analisar o grau de transparência do sistema de referência no que diz respeito a construções com a **1PP**. Nessa pessoa do discurso, de forma semelhante à variação que ocorre na 2PS entre *tu* e *você*, ocorre a variação entre *nós* e *a gente*. Nesse sentido, para obtenção do grau de transparência no uso da 1PP, levamos em conta as frequências de expressão e não-expressão pronominal do argumento-sujeito, a presença ou não de desinência número-pessoal explícita nos diferentes tempos/modos verbais, e a frequência de uso das formas pronominais *nós* e *a gente*. Segundo o passo a passo já estabelecido no cálculo da 2PS, em primeiro lugar calculamos, separadamente, o grau de transparência no uso do pronome *nós* e o grau no uso de *a gente*; em seguida, relacionamos

⁴⁷ Cf. da tabela 7:

Pronomes de 2PS	Sincronias				
	19-1	19-2	20-1	20-2	21-1
<i>tu</i>	81% (81/100)	72% (72/100)	1% (1/100)	2% (2/100)	0% (0/100) ⁴⁷
<i>você</i>	19% (19/100)	28% (28/100)	99% (99/100)	98% (98/100)	100% (100/100)
Total	100% (100/100)	100% (100/100)	100% (100)	100% (100)	100% (100/100)

⁴⁸ A título de esclarecimento, o cálculo realizado foi o seguinte: $[(76 \cdot 81) + (30 \cdot 19)] / 100 = 67,5$.

esse resultado com as frequências de uso de *nós* e de *a gente*, para chegarmos ao grau geral de transparência do sistema de referência no domínio da 1PP.

No caso do pronome *nós*, todos os tempos/modos verbais apresentam desinência número-pessoal por meio de morfema explícito (cf. quadro 6 do capítulo 4). Assim, com *nós*, em todos os tempos/modos verbais, quando houver expressão pronominal do argumento-sujeito, a referência será opaca, e, quando não houver, será transparente. Em todos os tempos/modos verbais, o grau de transparência será o mesmo e irá corresponder ao valor da frequência de não-expressão pronominal: 82%. Desse modo, com o pronome *nós*, não é necessário calcular uma média de grau de transparência entre os diferentes tempos/modos verbais.

No caso da forma pronominal *a gente*, os tempos/modos verbais são diferentes entre si quanto à presença ou ausência de desinência número-pessoal por morfema explícito: há desinência explícita no pretérito perfeito do indicativo, mas, nos demais tempos/modos, não há. Sendo assim, o grau de transparência será diferente entre o pretérito perfeito do indicativo e os demais tempos/modos, o que torna necessário calcular uma média de grau de transparência entre os diferentes tempos/modos, a fim de obter o grau geral de transparência no âmbito da forma *a gente*.

A tabela 11 mostra o cálculo que fizemos para obter esse grau geral.

Tabela 11: Cálculo do grau de transparência do sistema de referência relativamente a *a gente* na sincronia 19-1.

Tempo/modo	(+) ou (-) desinência	Frequência de não-expressão pronominal	Frequência de expressão pronominal	Grau de opacidade	Grau de Transparência
Ind. Pres.	-	82%	18%	82%	18%
Ind. Pret. imp.	-	82%	18%	82%	18%
Ind. Pret. perf.	+	82%	18%	18%	82%
Ind. Pret. m. q. perf.	-	82%	18%	82%	18%
Ind. Fut. pres.	-	82%	18%	82%	18%
Ind. Fut. pret.	-	82%	18%	82%	18%
Subj. Pres.	-	82%	18%	82%	18%
Subj. Pret. imp.	-	82%	18%	82%	18%
Subj. Fut.	-	82%	18%	82%	18%
MÉDIA				75%	25%

No pretérito perfeito do indicativo (em que há desinência número-pessoal explícita, sinalizada por “+”), sempre que houver expressão pronominal, a referência será opaca, e, sempre que não houver, a referência será transparente. Assim, nesse tempo, o grau de transparência será equivalente ao valor da frequência de não-expressão pronominal (82%), e o grau de opacidade será igual à frequência de expressão pronominal (18%). Nos demais tempos/modos (em que não há desinência número-pessoal explícita), a referência será opaca quando não houver expressão pronominal e será transparente quando houver. Nesses tempos/modos, os graus de opacidade e transparência do sistema de referência serão equivalentes, respectivamente, às frequências de não-expressão (82%) e de expressão pronominal (18%). A última linha da tabela 11 mostra a média do grau de transparência do sistema de referência no que tange à forma pronominal *a gente*, calculada mediante a média simples entre os valores que consideramos como graus específicos em cada tempo/modo verbal (última coluna).

Uma vez identificados os graus de transparência do sistema de referência no uso de *nós* (82%) e no uso de *a gente* (25%), resta correlacionar esses dados com as frequências de uso dos pronomes *nós* e *a gente*, que são, respectivamente, 96,5% e 3,5%, na sincronia 19-1. Para isso, de forma semelhante ao cálculo feito para a 2PS, fazemos uma média ponderada entre os graus de transparência do sistema com *nós* e com *a gente*, considerando como pesos as frequências de uso dessas formas pronominais, o que resulta no valor de **80%** como grau geral de transparência do sistema de referência no domínio da 1PP.

O mesmo tipo de cálculo aplicado para as sincronias 19-2, 20-1, 20-2 e 21-1 resulta, respectivamente, nos seguintes graus de transparência quanto ao uso da 1PP: **74%**, **51,5%**, **51%** e **51,5%**.

Por fim, no que concerne à **2PP**, em cada sincronia, o grau de transparência do sistema de referência depende das frequências de expressão e não-expressão pronominal do argumento-sujeito e da presença/ausência de desinência número-pessoal explícita nas formas

verbaux dos diferentes tempos/modos. Em todos os tempos/modos, as formas verbais relativas aos pronomes de 2PP (tanto *vós* quanto *vocês*) apresentam desinência número-pessoal explícita (ver quadro 6, no capítulo 4). Desse modo, no caso desses pronomes, em todos os tempos/modos verbais, a referência será transparente sempre que o argumento-sujeito não for expresso por pronome e será opaca quando houver expressão pronominal. Tendo em conta que, na sincronia 19-1, as frequências de não-expressão e expressão pronominal são, respectivamente, 86% e 14%,⁴⁹ nessa sincronia o grau de transparência do sistema no domínio da 2PP será de **86%** (e naturalmente o grau de opacidade será de 14%).

Seguindo o mesmo tipo de raciocínio, nas sincronias 19-2, 20-1, 20-2 e 21-1, os graus de transparência do sistema de referência no que diz respeito à 2PP são, respectivamente, **70%, 42%, 14% e 24%**.

Esse último cálculo sobre a 2PP completa o cálculo do grau de transparência do sistema de referência no domínio de cada uma das pessoas do discurso, para cada uma das sincronias. Assim, é possível calcular o grau geral de transparência do sistema de referência como um todo de cada sincronia, por uma média simples a partir dos resultados de todas as pessoas do discurso analisadas. Esse resultado geral é apresentado na tabela 12 abaixo.

Tabela 12: Variação diacrônica do grau de transparência do sistema de referência por expressão pronominal e desinencial do argumento-sujeito de 1ª e 2ª pessoas no PB

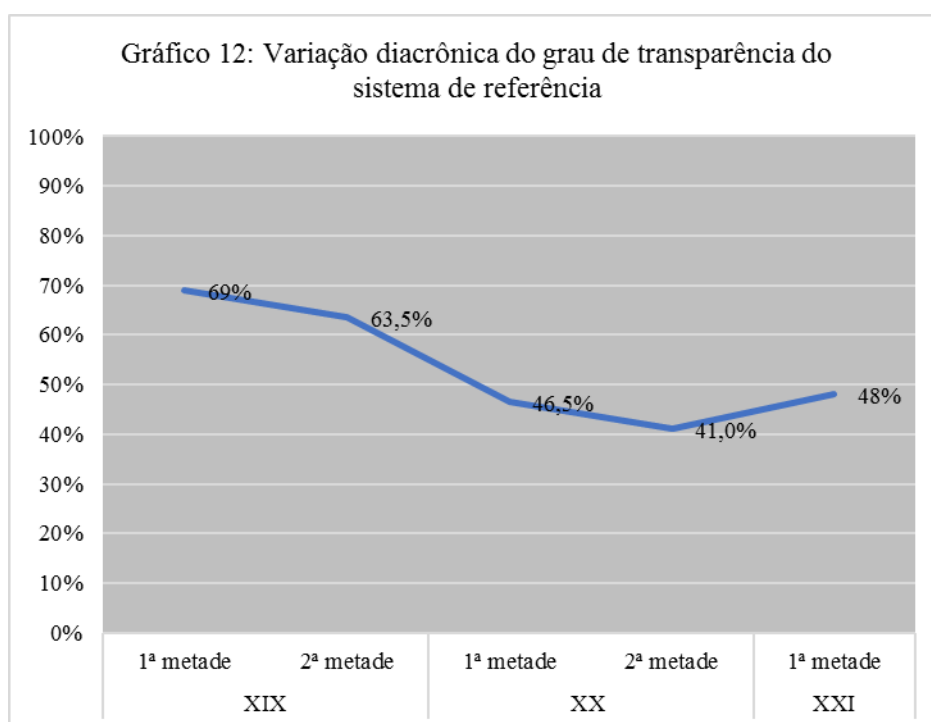
Pessoas do discurso	Sincronias				
	19-1	19-2	20-1	20-2	21-1
1PS	42,5%	44%	40,5%	45,5%	52,5%
2PS	67,5%	66%	53%	54,5%	64%
1PP	80%	74%	51,5%	51%	51,5%
2PP	86%	70%	42%	14%	24%
Média	69%	63,5%	46,5%	41%	48%

⁴⁹ Cf. da tabela 6:

Pessoa do discurso	Expressão pronominal	19-1	19-2	20-1	20-2	21-1
2PP	+	14% (07/50)	30% (15/50)	58% (29/50)	86% (43/50)	76% (38/50)
	-	86% (43/50)	70% (35/50)	42% (21/50)	14% (07/50)	24% (12/50)

Aqui convém destacar que, ao fazermos, em cada sincronia, uma média simples (e não uma média ponderada) entre os graus de transparência das diferentes pessoas do discurso para obtermos o grau geral de transparência do sistema como um todo, estamos procurando chegar ao grau de transparência da *gramática* da língua e não ao grau de transparência de textos específicos. Em textos específicos, podem ocorrer mais casos de 1PS do que de 1PP, por exemplo. Porém, isso decorreria de fatores contextuais, e não de determinações gramaticais. O sistema de referência no domínio de cada pessoa do discurso apresentaria um dado grau de transparência. Por meio de uma média simples entre esses graus, pode-se, então, chegar ao grau de transparência do sistema de referência da *gramática* da língua, que é o que procuramos identificar neste trabalho. Se cada texto particular utilizar, por exemplo, mais a 1PS do que outras pessoas, isso já seria um resultado da interação da gramática com fatores contextuais.

A média geral de variação diacrônica do grau de transparência do sistema de referência registrada na tabela 12 (resultado exibido na última linha) pode ser visualizada também no gráfico 12 abaixo.



Como definido anteriormente, o objetivo deste trabalho era mostrar como ocorre a variação diacrônica do grau de transparência do sistema de referência por expressão pronominal e desinencial do argumento-sujeito de 1ª e 2ª pessoas no PB. Assim, o resultado expresso na tabela 12 (última linha) e no gráfico 12 descreve justamente essa variação diacrônica, constituindo o resultado central deste trabalho.

Mais especificamente, nosso objetivo era mostrar que o grau de transparência desse sistema apresenta variações, mas que essas variações ocorrem tanto em direção a menor, quanto em direção a maior transparência, oscilando em torno de um certo eixo de transparência. O resultado final exposto no gráfico 12 e na tabela 12, então, demonstra justamente esse fato. Como se pode notar, o grau de transparência cai de 69%, na sincronia 19-1, para 41%, na sincronia 20-2 (o que pode ser visto como uma variação significativa), porém, na passagem seguinte para a sincronia 21-1, o grau de transparência vai se restabelecendo parcialmente, subindo para 48%. Assim, observa-se o que se pode considerar como uma variação em torno de um certo eixo de transparência.

Na seção seguinte, 5.3, discutimos mais detalhadamente esse resultado expresso na tabela 12 (e no gráfico 12), argumentando que ele indica, de fato, que a mudança no grau de transparência não apresenta uma direção única (transparência < opacidade ou opacidade < transparência), mas segue tanto em direção a diminuição, quanto em direção a aumento de transparência. Na sequência, em 5.4, completando o objetivo de nosso trabalho, relacionamos esse comportamento diacrônico do grau de transparência do sistema de referência à atuação de motivações cognitivas/comunicativas em interação, particularmente a iconicidade e a economia.

5.3 Não-unidirecionalidade da variação diacrônica do grau de transparência do sistema de referência

A ideia central que procuramos demonstrar neste trabalho é a de que o grau de transparência do sistema de referência aqui em estudo varia tanto em direção a maior quanto a menor transparência, não seguindo uma única direção (e que esse comportamento estaria associado a motivações comunicativas em interação). A nosso ver, o resultado identificado na seção anterior pode realmente ser interpretado como indicando essa tendência.

A tabela 12 e o gráfico 12 mostram, como já apontamos, uma queda de grau de transparência da primeira sincronia (19-1) à quarta (20-2), ocorrendo um aumento desta para a última sincronia (21-1). Como se verifica aumento apenas nesta última transição entre sincronias e como esse aumento, a princípio, parece pouco expressivo (7% – de 41% para 48%), poder-se-ia pensar, num primeiro momento, que os dados estariam mostrando que o sistema de referência teria, essencialmente, uma tendência de diminuição de grau de transparência e que o aumento identificado na última transição entre sincronias seria um fato pontual, isolado, dentro do que se poderia pensar como um percurso maior de queda de transparência. No entanto, alguns fatos parecem indicar que esse não seria o caso e que o movimento ascendente identificado nos dados realmente seria indicativo de que aumento de transparência é tão relevante quanto diminuição, na dinâmica da variação diacrônica do grau de transparência.

Primeiramente, é preciso observar que o aumento de transparência de 7% pode parecer um valor pouco expressivo, mas isso se ele é considerado numa escala de 0% a 100%. Porém, entendemos que esse valor deve ser situado no interior da amplitude de variação do grau de transparência na diacronia estudada. Como se pode ver nos dados, o grau de transparência varia entre um valor máximo de 69% (na sincronia 19-1) e um valor mínimo de 41% (na sincronia 20-2), o que significa que, em um intervalo de variação de 28%, o valor de

7% já se torna muito mais expressivo, correspondendo a um quarto desse intervalo. Nesse mesmo sentido, é interessante notar que o aumento de 7% identificado entre as sincronias 20-2 e 21-1 leva o grau de transparência a um nível mais alto do que o verificado na sincronia anterior, 20-1, que é de 46,5%, o que também parece indicar a relevância do aumento em foco.

Além disso, embora se verifique queda constante no grau de transparência da sincronia 19-1 à sincronia 20-2 (69% - 63,5% - 46,5% - 41%, respectivamente), essa queda é consideravelmente atenuada no seu final (de 20-1 a 20-2 = 5,5%), antes de iniciar o período de aumento de transparência (de 20-2 a 21-1 = 7%). No decorrer do período de redução de transparência, verifica-se uma diminuição bastante abrupta entre as sincronias 19-2 e 20-1, de 17%, porém, na transição seguinte (de 20-1 a 20-2), que precede o período de aumento de transparência, a redução é mais atenuada, sendo de 5,5%, o que pode ser visto como uma certa desaceleração no ritmo de queda, talvez já como parte do próprio processo seguinte de aumento. Assim, embora o aumento de transparência manifeste-se apenas na última transição entre sincronias, é possível que sinais dessa tendência já se anunciam desde a sincronia 20-1.

Levando em conta também a previsão que se pode fazer sobre a continuidade das mudanças aqui em estudo, é possível que o grau de transparência de 48% na sincronia 21-1 aumente no decorrer da sincronia. Esse prognóstico de ampliação de transparência pode ser traçado principalmente para 1PP e 2PP.

No caso da 1PP (assim como em todas as outras pessoas), vem ocorrendo um aumento no percentual de expressão pronominal do argumento-sujeito, que parte de 18% na sincronia 19-1, chegando a 68% na 21-1 (cf. tabela 6). Por um lado, esse aumento de expressão pronominal promove diminuição no grau de transparência do sistema de referência, o que ocorre no uso do pronome *nós*, já que, com esse pronome, todas as formas verbais possuem desinência número-pessoal explícita – assim, o aumento na expressão pronominal

amplia a coocorrência dessa expressão com expressão desinencial, alargando os índices de RC, contribuindo para a diminuição de transparência. Por outro lado, o aumento na expressão pronominal de 1PP leva também a ampliação do grau de transparência do sistema de referência, fato que se verifica no uso de *a gente*, visto que, com essa forma pronominal, as formas verbais de quase todos os tempos/modos (com exceção apenas do pretérito perfeito do indicativo) não têm desinência explícita – desse modo, o aumento na expressão pronominal incrementa a ocorrência de RS, promovendo aumento de transparência. Ou seja, o aumento de expressão pronominal leva a diminuição de transparência no âmbito do uso de *nós*, mas a ampliação de transparência no uso de *a gente*.

Como mostrado na tabela 8, o PB manifesta uma tendência acentuada de substituição de *nós* por *a gente* (o pronome *nós* cai de 96,5% para 39%, enquanto *a gente* sobe de 3,5% para 61%, da sincronia 19-1 para a 21-1). Portanto, a manutenção dessa tendência no decorrer da sincronia 21-1 deve levar a aumento no grau de transparência do sistema de referência, incrementando o percentual geral de 48% já identificado.

No caso da 2PP, também se constata aumento na frequência de expressão pronominal do argumento-sujeito (a tabela 6 mostra aumento de 14% para 76% nesse índice, da sincronia 19-1 para a 21-1, respectivamente). Nessa pessoa do discurso, seja com o pronome *vós*, seja com *vocês*, as formas verbais de todos os tempos/modos têm desinência número-pessoal explícita. Assim, o aumento na expressão pronominal implica em aumento no índice de RC, contribuindo para diminuição do grau de transparência do sistema de referência.

No entanto, perspectivas de mudança envolvendo a 2PP sugerem, para o decorrer da sincronia 21-1, um caminho de aumento de transparência. Embora estejamos considerando, para 2PP, os pronomes *vós* e *vocês*, nosso levantamento de dados evidencia um fato facilmente percebido no PB que é a progressiva substituição de *vós* por *vocês*, culminando com o uso quase exclusivo da forma *vocês* na sincronia 21-1. Com esse pronome, a desinência

número-pessoal, em todos os tempos/modos, consiste em uma nasalização vocálica. Trata-se, assim, de uma desinência com baixa saliência fônica (comparada, por exemplo, com desinências como *-stes* e *-des*, do pronome *vós*, ou *-mos*, do pronome *nós*) e, portanto, mais propícia a cair em desuso – possibilidade que pode ser aventada com base em autores como Scherre e Naro (1998). Com efeito, Duarte (1995, p. 41) observa que “a desinência *-m* é frequentemente omitida nos registros mais informais, sobretudo de falantes menos escolarizados”. Uma possível difusão da queda dessa desinência no PB implicaria, pois, num aumento do índice de transparência do sistema de referência, já que essa queda, associada ao aumento na frequência de expressão pronominal do argumento-sujeito, resultaria em diminuição de RC e aumento de RS. Tal fato contribuiria, então, para a ampliação do índice de 48% de transparência registrado na sincronia 21-1, fomentando, assim, a tendência de ampliação do grau de transparência do sistema de referência no decorrer dessa sincronia.

Em síntese, com base nos fatos arrolados nesta seção, entendemos que o resultado que obtivemos quanto à variação diacrônica do grau de transparência pode ser interpretado como indicativo de que essa variação não ocorre unidirecionalmente, caminhando, na verdade, em direção não só a diminuição, mas também a aumento de transparência.

Tal resultado parece, então, corroborar a hipótese que guiou nossa pesquisa, a do comportamento não-unidirecional da variação diacrônica do grau de transparência linguística, hipótese assentada em autores que teorizam sobre processos de mudança linguística histórica, conforme explicamos no capítulo 2.

É verdade que nossos dados não chegam a mostrar uma oscilação do grau de transparência em termos de repetidos movimentos de aumento e diminuição, o que certamente evidenciaria, de forma muito mais explícita, a nossa interpretação sobre a dinâmica diacrônica da transparência. A razão, no entanto, de não chegarmos a encontrar uma oscilação assim tão

explícita pode decorrer de o recorte diacrônico aqui estudado, embora significativamente representativo da história do PB, não ser tão extenso.

Por exemplo, Hodge (1970, *apud* SLOBIN, 1977) identifica no egípcio um movimento cíclico de perda e restabelecimento de construções linguística sintéticas *versus* analíticas, de modo similar ao que seria um movimento de perda e restabelecimento de transparência linguística. O autor constata dois períodos de restabelecimento de padrões de construções analíticas na história da língua, distribuídas, porém, num período de 4.000 anos de história. Nesse sentido, consideramos ser natural o resultado que encontramos, levando em conta nosso recorte diacrônico, o qual, a nosso ver, sem deixar de representar um período considerável da história do PB, é relativamente curto.

Enfim, uma vez constatado que o grau de transparência do sistema de referência varia tanto em direção a seu aumento, quanto em direção a sua diminuição, na seção seguinte procuraremos mostrar, então, que esse comportamento seria condizente com a atuação de motivações comunicativas em interação e pode ser pensado como decorrente da ação dessas forças.

5.4 Variação diacrônica de transparência e princípios gerais de funcionamento da linguagem

No capítulo 2, procuramos mostrar que, segundo teóricos que estudam fenômenos de mudança linguística histórica, a estrutura das línguas é determinada, em grande medida, pela interação entre motivações comunicativas (ou cognitivas), sendo as principais delas a *iconicidade* e a *economia*. Naquele capítulo, com base nessa premissa, concluímos que, por conta da atuação constante e igualmente relevante dessas duas forças, o grau de transparência do sistema de referência variaria ora em direção a diminuição de transparência, ora em direção a aumento (ou vice-versa), oscilando em torno de um determinado eixo, entre o que

consideramos como um piso e um teto de transparência, não caminhando em uma única direção. Realmente, esse comportamento não-unidirecional é o que identificamos nas seções anteriores deste capítulo. Na presente seção, procuramos explicar que esse resultado seria, de fato, condizente com a concepção da atuação de motivações comunicativas igualmente relevantes sobre a estrutura da língua.

Com efeito, a variação do grau de transparência entre aumento e diminuição (conforme atestado em nossa análise quantitativa) é condizente com a ideia da atuação de forças comunicativas agindo em interação, ao passo que um percurso do grau de transparência apenas em uma dessas direções já parece não ser compatível com o equilíbrio entre essas motivações.

Se o grau de transparência caminhasse apenas no sentido de aumento, o sistema de referência apresentaria um percentual cada vez mais alto de relações icônicas, no nosso caso de alguma das duas formas de RS (ou das duas), o que implicaria uma atuação muito mais predominante da motivação icônica sobre a motivação econômica. Essa seria uma situação em que a língua não estaria sob a influência igualmente relevante de iconicidade e economia.

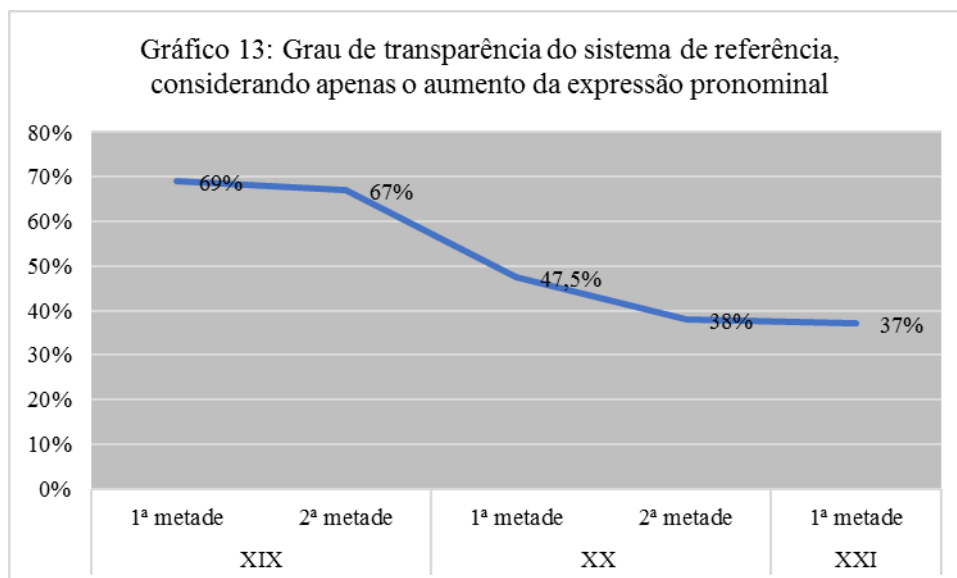
Da mesma forma, se o grau de transparência caminhasse apenas em direção a diminuição de transparência, o sistema de referência teria uma incidência muito alta de RC e/ou RZ. Um alto predomínio de RC ou RZ não seria compatível com o equilíbrio entre iconicidade e economia. Um sistema de referência com um alto índice de RC não seria icônico, nem econômico. Similarmente, um sistema com um alto índice de RZ seria muito econômico, mas pouco ou nada icônico. Portanto, de um ponto de vista teórico, o que é mais compatível com a ideia da interação igualmente relevante entre economia e iconicidade seria um sistema de referência cujo grau de transparência não caminhe unidirecionalmente para

aumento ou diminuição de transparência, justamente conforme identificamos na análise empírica deste trabalho.

As mudanças diacrônicas ocorridas no PB que interferem no grau de transparência do sistema de referência e que fazem esse grau variar de forma não-unidirecional parecem constituir um conjunto de mudanças que se relacionam de tal forma que é possível reconhecer uma interação mais equilibrada entre iconicidade e economia. O aumento na frequência de expressão pronominal do argumento-sujeito, por um lado, e as mudanças *tu/você* e *nós/a gente* – que podem ser referidas como integrando um processo de mudança no quadro pronominal do PB –, por outro lado, interagem entre si de modo a promover mudanças em que se pode reconhecer a ação conjunta e mais ou menos equilibrada de iconicidade e economia.

Ao contrário, se, hipoteticamente, ocorresse no PB apenas uma dessas mudanças, o que se passaria a notar seria um tipo de cenário no qual a variação diacrônica do grau de transparência do sistema de referência seria unidirecional e, ao mesmo tempo, no qual não seria possível reconhecer uma interação mais equilibrada entre as motivações comunicativas.

Suponha-se um cenário 1 em que se verificasse no PB, no recorte diacrônico aqui em pauta, apenas o aumento da expressão pronominal do argumento-sujeito, sem o processo de mudança no quadro pronominal. Aplicando os mesmos tipos de cálculos que utilizamos no decorrer de nossa análise de dados, é possível ver que o grau de transparência do sistema de referência apresentaria uma variação diacrônica unicamente em direção a diminuição de transparência, conforme ilustra o gráfico 13 abaixo.

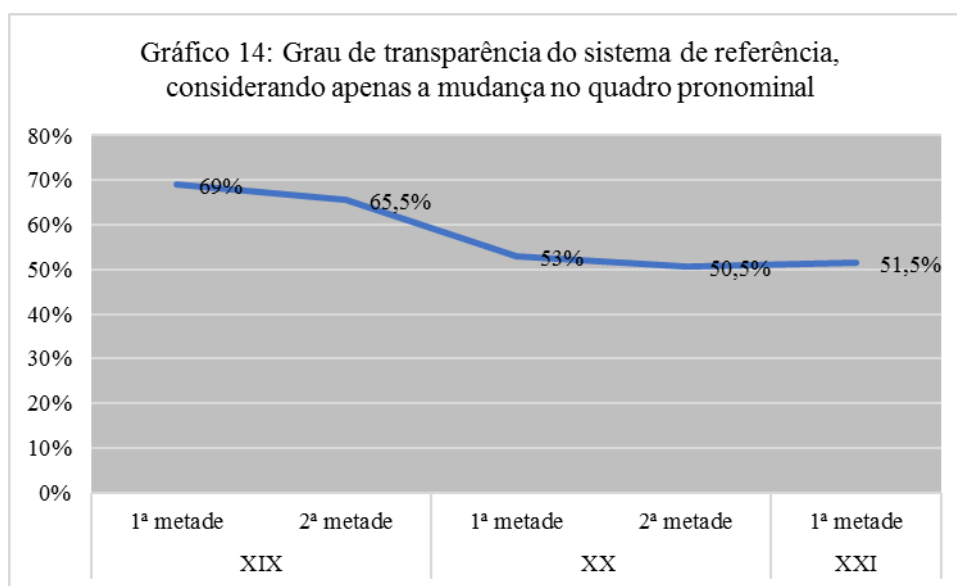


Esse percurso unidirecional de diminuição do grau de transparência representado no gráfico 13 seria o resultado de um progressivo aumento dos índices de RC no sistema de referência. Na sincronia 19-1, na 2PS, há o predomínio de uso de *tu* (81%) em detrimento de *você* (19%) e o predomínio ainda mais expressivo de *nós* (96,5%) sobre *a gente* (3,5%) – cf. tabelas 7 e 8. Desse modo, no cenário aqui suposto, em que não ocorreria mudança nas frequências de uso desses pronomes, encontraríamos, em todo o decorrer de nosso período diacrônico, um sistema de referência envolvendo o pronome *eu* de 1PS, a predominância de *tu* em 2PS e de *nós* em 1PP e os pronomes de 2PP (*vós/vocês*), todos eles combinando-se com um contínuo aumento da frequência de expressão pronominal do argumento-sujeito (cf. tabela 6). Como, com os pronomes *tu*, *nós* e *vós/vocês* (à exceção apenas do pronome *eu*), as formas verbais de todos os tempos/modos contêm desinência número-pessoal explícita, o que se verificaria, no cenário hipotético em questão, seria, pois, o predomínio do aumento de construções de RC, resultado manifestado no gráfico 13 pela diminuição ininterrupta de transparência.

Tal cenário não seria compatível com a atuação das forças de iconicidade e economia, já que a RC é um tipo de construção nem icônico, nem econômico. Com efeito,

Croft (2003), conforme explicado no capítulo 2, afirma que combinações que não sejam motivadas nem por uma nem por outra dessas duas forças tendem a ser raras ou instáveis historicamente nas línguas do mundo. Assim, a não-instauração no PB desse cenário hipotético 1 – em que a variação diacrônica do grau de transparência do sistema de referência seria unidirecional – poderia ser vista como decorrendo justamente do fato de tal cenário não ser compatível com a atuação das referidas motivações comunicativas.

Considere-se agora um cenário 2, também hipotético, em que tivesse ocorrido no PB apenas o processo de mudança do quadro pronominal, sem ocorrência do aumento da frequência de expressão pronominal do argumento-sujeito. Nesse cenário, de forma mais ou menos similar ao visto no cenário 1 acima, o grau de transparência do sistema de referência variaria diacronicamente num sentido essencialmente (embora não totalmente) unidirecional de diminuição de transparência e se trataria também de um quadro em que não se veria uma interação mais propriamente equilibrada entre iconicidade e economia. Lançando mão dos mesmos tipos de cálculos adotados nas demais partes de nosso trabalho, seria possível se chegar ao resultado esboçado no gráfico 14 abaixo.



O percurso (basicamente) unidirecional de diminuição do grau de transparência exposto no gráfico 14, diferentemente daquele previsto no gráfico 13 para o cenário 1, seria o resultado do predomínio de um progressivo aumento dos índices de RZ (não RC) no sistema de referência. Neste segundo cenário hipotético, estamos considerando a ocorrência apenas da mudança no quadro pronominal, sem o processo de aumento da expressão pronominal de argumento-sujeito. Nesse caso, consideraríamos, então, o aumento contínuo do uso de *você* em detrimento de *tu* e do uso de *a gente* em prejuízo de *nós*, culminando, ao final do período (21-1), nos índices de 100% e 61% de *você* e *a gente*, respectivamente (cf. tabelas 7 e 8). No decorrer de toda a diacronia, os índices de expressão pronominal permaneceriam baixos, já que, no cenário hipotético em questão, não se efetivaria o fenômeno do aumento desses índices, que continuariam os mesmos percentuais baixos identificados na sincronia 19-1 (21% em média – cf. tabela 6). Desse modo, ao longo de todo o período diacrônico, teríamos um quadro pronominal formado pelo pronome de 1PS *eu*, pela presença crescente de *você* e *a gente* e pelos pronomes de 2PP *vós/vocês*, todos caracterizados por um baixo índice de expressão pronominal. Como, com os pronomes *eu*, *você* e *a gente* (excetuando-se apenas os pronomes de 2PP), as formas verbais, na grande maioria dos tempos/modos, não têm desinência explícita, o que se atestaria nesse cenário 2 seria, então, o predomínio do aumento constante de construções de RZ e, assim, a diminuição, quase ininterrupta, do grau de transparência do sistema de referência, prevista no gráfico 14.

Um sistema de referência como o delineado nesse cenário hipotético 2 já seria menos incompatível com a interação entre motivações comunicativas do que aquele previsto no cenário 1 anterior. Neste cenário 2, haveria o predomínio de construções de RZ (em 1PS, 2PS e 1PP, com exceção apenas de 2PP), que constituem um tipo de construção econômica. Porém, o predomínio da economia em três das quatro pessoas do quadro pronominal parece ainda não ser a configuração mais ideal (ou ótima) em termos de equilíbrio entre as forças,

tanto que esse cenário, em que a variação do grau de transparência do sistema de referência é praticamente unidirecional, não é o que se verifica no PB.

Em contrapartida, no cenário que realmente se instala na língua – no qual, na diacronia aqui selecionada, ocorre tanto o aumento da frequência de expressão pronominal do argumento-sujeito, quanto a mudança no quadro pronominal –, o grau de transparência do sistema de referência varia em direção a diminuição, bem como em direção a aumento de transparência (conforme mostramos acima), e esse comportamento pode ser associado a uma ação mais equilibrada entre iconicidade e economia. Os dados expostos na tabela 12 parecem, de fato, indicar essa interpretação. Para facilitar a observação desses dados, repetimos abaixo a tabela 12.

Tabela 12: Variação diacrônica do grau de transparência do sistema de referência por expressão pronominal e desinencial do argumento-sujeito de 1ª e 2ª pessoas no PB

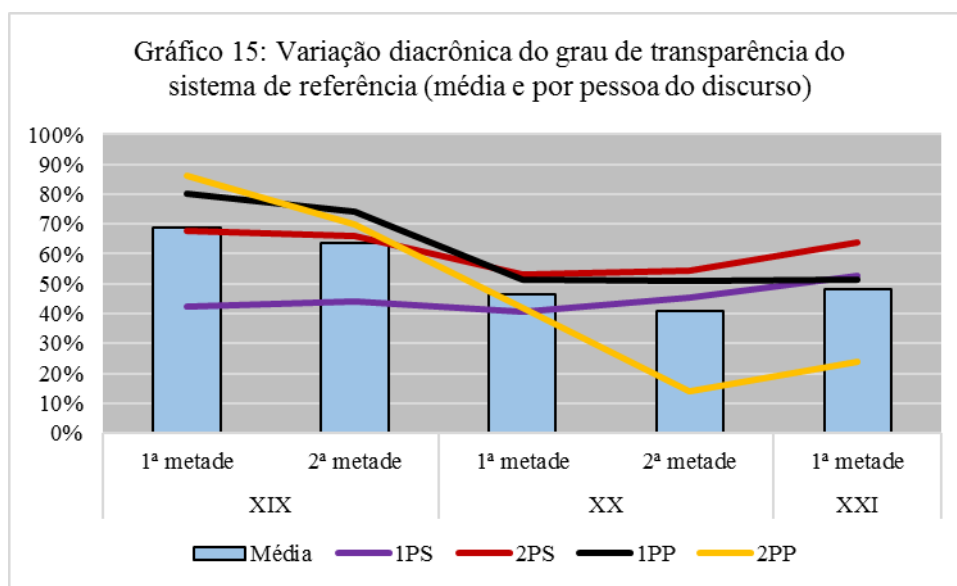
Pessoas do discurso	Sincronias				
	19-1	19-2	20-1	20-2	21-1
1PS	42,5%	44%	40,5%	45,5%	52,5%
2PS	67,5%	66%	53%	54,5%	64%
1PP	80%	74%	51,5%	51%	51,5%
2PP	86%	70%	42%	14%	24%
Média	69%	63,5%	46,5%	41%	48%

Como já comentado, o grau de transparência do sistema de referência (na média entre as quatro pessoas do discurso) decresce da sincronia 19-1 até a 20-2, de 69% para 41%, e se eleva desta para a 21-1, de 41% para 48%. Esse decréscimo decorre principalmente do comportamento de 2PS (sobretudo pela queda de transparência de 66% para 53%, da sincronia 19-2 para a 20-1), de 1PP (que mostra um decréscimo acentuado também nesse intervalo, de 74% para 51,5%) e, sobretudo, de 2PP (que exhibe acentuado decréscimo de transparência em todo o intervalo de 19-1 a 20-2). Já a 1PS pouco contribui para esse percurso de queda de transparência, visto que, no período desse percurso, o grau de transparência com essa pessoa do discurso oscila (e de forma não muito expressiva) entre subida e descida,

variando entre 42,5%, 44%, 40,5% e 45,5%.

Por sua vez, a elevação do grau geral de transparência do sistema de referência presenciada da sincronia 20-2 para a 21-1 é decorrente de 1PS (em que há ampliação de transparência de 45,5% para 52,5%), de 2PS (cujo grau de transparência passa de 54,5% para 64%) e de 2PP (em que o grau de transparência passa de 14% para 24%) – o aumento de transparência com 1PP, como se pode checar na tabela, é pouco expressivo (de apenas 0,5%).

A variação diacrônica média do grau de transparência do sistema de referência e as variações específicas de cada pessoa do discurso, que acabamos de apontar, podem ser visualizadas também no gráfico 15 abaixo.



A nosso ver, as mudanças que ocorrem com essas pessoas do discurso e que acabam por acarretar a diminuição e o aumento no grau geral de transparência do sistema de referência podem ser associadas, de modo coerente, a uma atuação mais equilibrada das forças de iconicidade e economia, comparativamente àquilo que se observa nos dois cenários hipotéticos acima.

Na tabela 12, assim como no gráfico 15, é possível ver que, a partir da sincronia 20-1, o grau de transparência com a 2PS para de diminuir, passando a subir a partir de então; da mesma forma, o grau de transparência com a 1PP se estabiliza nesse mesmo período, e se estaciona no mesmo patamar até o último período sincrônico. Esses resultados com 2PS e 1PP impactam o grau geral de transparência do sistema de referência, contribuindo para a desaceleração desse grau geral da sincronia 20-1 para a 20-2 e, finalmente, para o aumento desse grau geral a partir de 20-2 (cf. tabela 12).

Esses resultados com 2PS e 1PP são obtidos porque, a partir da sincronia 20-1, o uso de *você* suplanta (quase absolutamente) o de *tu*, e o uso de *a gente* supera o de *nós* (como mostram as tabelas 7 e 8). O predomínio de *você* e *a gente* (cujas respectivas formas verbais, na maioria dos casos, não contêm desinência explícita), combinado com o progressivo aumento na expressão pronominal do argumento-sujeito (gerando, pois, a prevalência de RS), passa, então, a contribuir para o aumento do grau de transparência com a 2PS (a partir de 20-1, cf. tabela 12) e para o estacionamento do grau de transparência com a 1PP (também a partir de 20-1, cf. tabela 12). Esses resultados com 2PS e 1PP, como dito logo acima, impactam o grau geral de transparência do sistema de referência (contribuindo para sua desaceleração e posterior aumento – cf. tabela 12).

Pode-se notar, portanto, que a combinação entre aumento de expressão pronominal e aumento do uso de *você* e *a gente* (que coocorrem no cenário real, diferentemente do que se vê nos cenários hipotéticos 1 e 2) é crucial para a alteração na direção da variação diacrônica do grau geral de transparência do sistema de referência.

É, então, interessante perceber que essa combinação de mudanças pode ser considerada, ao mesmo tempo, um movimento icônico e econômico. É essencialmente icônico, porque resulta na ampliação e predominância de construções de RS, um tipo de referência naturalmente icônico. Mas, segundo nossa visão, pode também ser visto como

econômico em certa medida. Considerando o sistema de referência como um conjunto que compreende três possibilidades de referência – RC, RS e RZ –, a RS, embora, em princípio, possa parecer apenas icônica, pode também ser reconhecida como mais econômica do que a RC, apesar de não ser tão econômica quanto a RZ. O sistema de referência pode evoluir de modo a promover RC, RS ou RZ. Quando caminha para RS, e não para RC, está se dirigindo para uma alternativa que, mesmo não sendo a mais econômica, é, de todo modo, mais econômica do que uma outra possibilidade (a RC).

Por exemplo, comparando o cenário real do PB com o cenário hipotético 1, em que ocorreria aumento e dominância da RC, fica mais evidente que o cenário real é mais econômico do que um outro cenário que, pelo menos teoricamente, poderia vir a ocorrer. Portanto, parece-nos pertinente considerar que a RS, quando situada no contexto das três possibilidades de referência, pode ser vista como uma alternativa que tem uma orientação, em certa medida, também econômica, além, é claro, de sua essência icônica.

Desse modo, as mudanças envolvendo 2PS e 1PP constituiriam um ponto das mudanças linguísticas aqui em análise no qual iconicidade e economia operariam em harmonia, ambas impulsionando o sistema de referência para um mesmo estado – o que é uma das possibilidades de interação entre as motivações comunicativas, conforme explicado no capítulo 2.

Portanto, no cenário real de mudanças presenciado no PB (aumento na expressão pronominal e mudança no quadro pronominal), verifica-se uma variação no grau de transparência do sistema de referência que não é unidirecional; ao mesmo tempo, levando em conta 2PS e 1PP, observa-se que tais mudanças podem associadas à iconicidade (essencialmente) e à economia. Esse panorama, a nosso ver, pode indicar justamente que a dinâmica diacrônica do grau de transparência estaria vinculada, pelo menos em certa medida, a motivações comunicativas em interação.

No caso do papel da 1PS na diacronia do grau geral de transparência do sistema de referência, particularmente no caso da sua contribuição para o aumento desse grau da sincronia 20-2 para a 21-1, também é possível perceber um movimento condizente com as motivações comunicativas.

No período da queda do grau geral de transparência do sistema de referência (19-1 a 20-2), a transparência com a 1PS, como mostram a tabela 12 e o gráfico 15, exhibe um movimento de oscilação entre aumento e diminuição, e dentro de uma margem relativamente estreita: de 42,5% (19-1), sobe para 44% (19-2), depois desce para 40,5% (20-1) e volta a subir para 45,5% (20-2). Da sincronia 20-2 para a 21-1, seguindo o padrão de oscilação percorrido até então, o grau de transparência com a 1PS deveria baixar, mas sobe para 52,5%.

Esse resultado decorre claramente do comportamento do processo de variação da frequência de expressão pronominal do argumento-sujeito com essa pessoa do discurso. Na 1PS, o grau de transparência da referência varia apenas em função de alterações na frequência de expressão pronominal (já que o pronome de 1PS não faz parte do conjunto de pronomes afetado pela mudança do quadro pronominal). Nessa pessoa, quanto maior a frequência de expressão pronominal, maior a tendência de transparência (porque, em tal pessoa, a maioria dos tempos/modos verbais não contém desinência número-pessoal explícita; assim, a combinação entre aumento na expressão pronominal e essa maioria de formas verbais sem desinência explícita vai promovendo aumento na transparência com a 1PS). A tabela 6 e o gráfico 9 mostram que a frequência de expressão pronominal na 1PS, entre as sincronias 19-1 e 20-2, oscila entre aumento e diminuição: de 28% (19-1), sobe para 32% (19-2), então desce para 22% (20-1) e sobe novamente para 36% (20-2); seguindo esse padrão, esperar-se-ia que, na passagem de 20-2 para 21-1, a expressão pronominal diminuísse, mas, na verdade, sobe, e acentuadamente, para 58%. A interrupção desse padrão, que passa a intensificar a expressão

pronominal, promove, então, o aumento da transparência com a 1PS, verificado na passagem da sincronia 20-2 para a 21-1.

O que é interessante observar aqui é, então, que a inversão desse padrão é um movimento que, neste caso, pode ser visto como iconicamente motivado. Na 1PS, o aumento da frequência de expressão pronominal promove, predominantemente, aumento de construções de RS, em detrimento de construções de RZ, isto é, a passagem de uma construção essencialmente econômica (RZ) para uma icônica e menos econômica (RS). Portanto, nota-se que a mudança ocorrida na 1PS que mais impacta o grau geral de transparência do sistema de referência, e que, inclusive, colabora para que esse grau geral varie de forma não-unidirecional, é uma mudança condizente com a atuação de forças comunicativas.

Finalmente, com a 2PP, ocorre uma situação interessante. A princípio, as mudanças envolvendo essa pessoa do discurso que afetam o grau de transparência do sistema de referência parecem ter pouca relação com as motivações icônica e econômica. Porém, sob uma análise mais detalhada, elas se mostram também plausíveis com a atuação dessas forças.

No percurso entre as sincronias 19-1 e 20-2, o grau geral de transparência do sistema de referência vai caindo, e a transparência com a 2PP certamente é um dos principais fatores dessa queda. Nesse período, particularmente com essa pessoa, o grau de transparência apresenta a seguinte sequência de queda (cf. tabela 12): 86% (19-1), 70% (19-2), 42% (20-1), 14% (20-2). Essa progressiva e acentuada diminuição de transparência deve-se ao contínuo e também intenso aumento da frequência de expressão pronominal processada nessa pessoa (cf. tabela 6): 14% (19-1), 30% (19-2), 58% (20-1), 86% (20-2). Essa correlação advém do fato de as formas verbais relativas à 2PP terem, em todos os tempos/modos, desinência explícita, o que, na combinação com o aumento da expressão pronominal, promove a ampliação de construções de RC. Esse tipo de referência, como já explicado, não é nem icônico nem

econômico. Assim, nesse período da diacronia em análise, observa-se que a mudança envolvendo a 2PP, a qual tem grande impacto sobre a variação do grau geral de transparência do sistema de referência, a princípio, não teria relação com a atuação dessas forças comunicativas.

Na passagem da sincronia 20-2 para a 21-1, quando há o aumento do grau geral de transparência do sistema de referência (41% para 48%), a 2PP também tem papel significativo, já que o grau de transparência da referência com essa pessoa sobe de 14% para 24%. Já nesse caso, ao contrário do verificado no período de 19-1 a 20-2, a mudança ocorrida com a 2PP tem clara relação com as motivações icônica e econômica. Esse aumento do grau de transparência com a 2PP de 14% para 24% é consequência da diminuição de expressão pronominal do argumento-sujeito que se verifica com essa pessoa nessa transição entre sincronias (que decresce de 86% para 76% – cf. tabela 6), já que essa diminuição de expressão pronominal promove o decréscimo de construções de RC, em favor do aumento de construções de RS (do tipo formado por não-expressão pronominal combinada com desinência explícita). A substituição de RC por RS é uma mudança que pode ser entendida como sendo motivada tanto por iconicidade, quanto por economia, operando de forma harmônica. Portanto, na passagem da sincronia 20-2 para a 21-1, encontra-se um momento de variação no grau geral de transparência do sistema de referência que evidencia o funcionamento não-unidirecional dessa variação e que tem como um de seus fatores uma mudança linguística motivada por economia e iconicidade, no caso a mudança aqui explicada sobre 2PP.

Ainda quanto à 2PP, é significativo perceber que o aumento contínuo de RC ocorrido entre 19-1 e 20-2 (e que é decisivo para a diminuição do grau geral de transparência do sistema de referência), embora, num primeiro momento, não tenha relação direta com as motivações icônica e econômica (como mencionado há pouco), pode ser também pensado

como tendo relação com essas forças. A esse respeito, é preciso considerar que, no PB, diacronicamente, há o processo de substituição do pronome *vós* por *vocês*. Em nosso trabalho, como explicado na metodologia, não controlamos essa mudança, pois ela não interfere na realização transparente ou opaca da referência. No entanto, o reconhecimento da ocorrência desse fenômeno no PB é relevante aqui.

Como já mencionado acima, alguns autores (cf. DUARTE, 1995) consideram que haveria, no PB, uma tendência para a queda da desinência número-pessoal de 3ª pessoa do plural. Na mesma direção, com base em outros trabalhos (cf. SCHERRE; NARO, 1998), pode-se considerar que desinências verbais com menor saliência fônica são mais propensas ao desuso do que desinências com maior saliência. Desse modo, não seria descabido pensar que a mudança pronominal de *vós* para *vocês*, tendo em mente uma possível perspectiva de simplificação da morfologia verbal de 3ª pessoa do plural, seria condizente com um processo de substituição de construções de RC por construções de RS com a 2PP. Tal substituição seria o resultado da passagem da combinação do alto índice de expressão pronominal com formas verbais que têm desinência explícita (***vocês cantam***) para a combinação desse alto índice de expressão pronominal com formas verbais sem desinência explícita (***vocês cantaØ***).

Sendo a passagem de RC para RS um movimento que pode ser associado à atuação de iconicidade e economia (de forma harmônica), pode-se, então, pensar que também no período entre 19-1 e 20-2, a mudança ocorrida em 2PP interferente no grau geral de transparência do sistema de referência seria uma mudança ligada a essas duas forças comunicativas.

Em resumo, com a comparação que fizemos aqui entre os cenários hipotéticos 1 e 2 e o cenário real de mudanças que se instauram no PB, procuramos mostrar que a variação diacrônica do grau geral de transparência do sistema de referência aqui atestada pode ser

entendida como estando, de alguma forma, vinculada à interação entre as motivações comunicativas da iconicidade e da economia.

No cenário hipotético 1, em que ocorreria apenas o aumento da expressão pronominal do argumento-sujeito, não se chega nem a verificar uma atuação significativa dessas forças, uma vez que esse cenário seria caracterizado pela prevalência do contínuo aumento de RC, combinação nem icônica, nem econômica, cenário em que a variação do grau de transparência do sistema de referência seria unidirecional, caminhando para a diminuição de transparência.

No cenário hipotético 2, no qual se processaria apenas a mudança no quadro pronominal, embora seja possível identificar mudanças já mais condizentes com as referidas motivações, parece haver um desequilíbrio marcante entre elas, predominando fortemente a motivação econômica. Nesse cenário, verifica-se uma variação no grau de transparência quase absolutamente unidirecional, apontada também para a diminuição de transparência, desta vez, por meio da ampliação de construções de RZ.

Já no cenário real encontrado no PB, em que coocorrem o aumento da frequência de expressão pronominal do argumento-sujeito e a mudança do quadro pronominal, é possível reconhecer uma maior interação entre as duas motivações. Ao mesmo tempo, o que se apura é uma variação diacrônica do grau de transparência do sistema de referência que não é unidirecional, mas que diminui numa primeira fase, passando a aumentar mais adiante.

Desse modo, é pertinente pensar que este último cenário é o que teria se instalado no PB por ser aquele mais condizente com uma interação mais ideal entre as motivações de iconicidade e economia, cenário no qual o grau de transparência do sistema de referência variaria de forma não unidirecional, mas ora em direção a decréscimo, ora em direção a aumento de transparência.

Quando consideramos que este último cenário é o de interação mais ideal entre as referidas motivações, não pressupomos um atendimento exatamente equilibrado entre essas forças. Como mostramos, as principais mudanças de tal cenário envolvem movimentos harmônicos entre iconicidade e economia, bem como movimento apenas icônico, o que, sem deixar de significar a interação entre as forças, indica uma certa prevalência da iconicidade. Igualmente, não assumimos que esse último cenário seria o único em que é possível reconhecer a atuação dessas forças. Como procuramos mostrar, também no cenário 2, principalmente, seria possível reconhecê-las.

Na verdade, o que entendemos é que o sistema linguístico vai se adaptando diacronicamente sempre de modo a tentar alcançar o estado mais ideal possível em termos do atendimento a essas motivações comunicativas (além de outras). Nesse sentido, Leufkens (2015, p. 28), com base em Haiman (1985) e Givón (1985), observa que motivações comunicativas competem entre si “resultando na seleção de uma forma linguística ótima”.⁵⁰ Assim, nosso entendimento não é o de que o cenário realmente atestado no PB seria de perfeito equilíbrio entre iconicidade e economia, mas o de maior equilíbrio possível entre essas forças, isto é, o cenário mais ideal, de maior *otimalidade*, em termos de interação entre essas forças.

Em síntese, na presente seção, procuramos argumentar que as mudanças linguísticas diacrônicas ocorridas no PB que interferem na diacronia do grau geral de transparência do sistema de referência são mudanças que podem ser vistas como condizentes com as forças comunicativas de iconicidade e economia. Como tais mudanças culminam com uma variação diacrônica não unidirecional de transparência, consideramos pertinente pensar que esse comportamento diacrônico da transparência seria determinado, ou delimitado, pela ação de tais forças comunicativas.

⁵⁰ Cf. o original: “[...] resulting in the selection of an optimal linguistic form” (LEUFKENS, 2015, p. 28).

Trata-se, portanto, de uma conclusão empírica que corrobora nossa conclusão teórica, formulada no capítulo 2, de que a variação diacrônica do grau de transparência do sistema de referência variaria de forma não unidirecional, por conta da existência do que chamamos de *teto* e *piso* de transparência, limites de variação de transparência existentes nas línguas justamente pela ação de iconicidade e economia.

A esse respeito, cabe destacar o que nossos dados podem apontar sobre quais poderiam vir a ser esses valores de piso e teto de transparência. Os dados permitem, mais diretamente, avaliar a noção de piso de transparência, na medida em que, de fato, mostram um percurso de diminuição de grau de transparência que, num dado momento, é desacelerado e, finalmente, interrompido (no patamar de 41%), passando o grau de transparência, a partir de então, a aumentar. O que se vê, pois, nesses dados, é que o grau de transparência decresceria a um dado nível, a partir do qual essa diminuição não continuaria, o que poderia ser pensado como um possível piso de transparência.

É certo que nossos dados não permitem determinar que o nível mais baixo de transparência aqui identificado seria o piso de transparência do sistema de referência. Para uma determinação mais fundamentada desse piso, seria necessário o estudo de um recorte diacrônico maior, que pudesse vir a mostrar mais de uma vez limites mínimos de transparência. De todo modo, embora nossos dados não permitam essa determinação, é interessante reconhecer que eles representam um resultado concordante com essa noção de piso de transparência.

Quanto à noção de teto de transparência, nossos dados não possibilitam uma avaliação mais direta. A princípio, é plausível cogitar que esse valor estaria por perto do valor máximo de transparência que encontramos, no caso, 69%. Todavia, nossos dados não mostram um aumento de transparência que atinge um pico, depois do qual ela decresce. Os dados mostram, já no início do período diacrônico recortado, um percurso já em processo de

queda (como bem ilustra o gráfico 12). Não é possível ver, obviamente, o que teria ocorrido em sincronias pretéritas à sincronia 19-1: a queda no grau de transparência que identificamos (que se inicia nessa sincronia, a partir de 69%) pode ser a continuação de um percurso anterior de queda (que estaria vindo de um grau ainda maior do que 69%) ou pode ser o início de uma queda (após o sistema de referência passar por um período de aumento de transparência que a teria levado até o patamar de 69%).

Uma avaliação mais apurada de qual patamar poderia ser pensado como um teto de transparência do sistema de referência dependeria do estudo de um recorte diacrônico maior que o delimitado neste trabalho. Por ora, acreditamos que nossa discussão teórica sobre a pertinência da noção de teto (e piso) de transparência e nossa análise empírica, que indica a variação diacrônica não unidirecional da transparência, possam já contribuir para reflexões sobre a questão do teto de transparência.

CONCLUSÃO

O objetivo central desta tese foi analisar a variação diacrônica do grau de transparência do sistema de referência por expressão pronominal e desinencial do argumento-sujeito de 1ª e 2ª pessoas no PB. Mais especificamente, procuramos analisar como o grau de transparência desse sistema é afetado por mudanças diacrônicas da língua que interferem diretamente nesse sistema, no caso, o aumento da frequência de expressão pronominal do argumento-sujeito e o aumento das frequências de uso das formas pronominais *você* e *a gente*, em detrimento de *tu* e *nós*, respectivamente.

A hipótese que guiou o desenvolvimento da nossa pesquisa foi a de que o grau de transparência do referido sistema não mudaria unidirecionalmente (para maior ou menor transparência), mas variaria dentro de certos limites máximo e mínimo de transparência, por conta da competição entre motivações comunicativas igualmente relevantes, sobretudo iconicidade e economia, competição essa defendida por diferentes teóricos como fundamental para o funcionamento da linguagem.

No decorrer desta tese, demonstramos, então, que essa hipótese se confirma. Apresentamos uma análise quantitativa de dados que integrou os resultados estatísticos que obtivemos sobre as três mudanças diacrônicas interferentes no grau de transparência do sistema de referência estudado. Essa análise, de fato, indica que o grau de transparência varia ora na direção de diminuição, ora no sentido de aumento de transparência. Conforme mostramos, identificamos, no período diacrônico analisado, um primeiro momento de queda do grau de transparência (da primeira metade do século XIX à segunda metade do XX), seguido de um período de elevação desse grau (presenciado na passagem da segunda metade do século XX para o século XXI).

Ademais, argumentamos que as mudanças diacrônicas interferentes no grau de

transparência do sistema de referência interagem de uma forma que pode ser interpretada como estando associada à atuação das forças de iconicidade e economia. Procuramos defender que a interação entre o aumento da frequência de expressão pronominal do argumento-sujeito e a mudança no quadro pronominal do PB promove uma dinâmica nos índices de RS, RC e RZ do sistema de referência condizente com a interpretação de uma atuação relevante de ambas as forças.

Dessa forma, o trabalho cumpriu o objetivo estabelecido, chegando a resultados concretos e significativos a respeito do tema focalizado. Tais resultados, além de suscitar outras generalizações, sugerem a legitimidade da concepção de que a competição entre motivações comunicativas, como iconicidade e economia, seria um dos principais fatores determinantes da estrutura das línguas, em particular, do desenvolvimento da mudança diacrônica.

Também quanto à depreensão de generalizações, recorde-se que, na introdução desta tese, explicamos que um propósito mais amplo do nosso trabalho seria oferecer resultados para que se possa pensar o funcionamento diacrônico da propriedade da transparência linguística nas línguas. Naquele momento, lançamos questionamentos gerais quanto à ocorrência ou não de variação diacrônica no grau de transparência de uma língua e quanto à direção dessa mudança. Desse modo, ao detectar variação diacrônica no grau de transparência do sistema de referência do PB e ao atestar sua oscilação não-unidirecional, nosso trabalho propicia dados específicos que contribuem para a reflexão sobre questionamentos mais gerais como os mencionados.

A constatação de que, no transcorrer de um período histórico, o grau de transparência do sistema de referência oscila entre decréscimo e elevação de transparência indica que a coexistência entre certo grau de transparência e certo grau de opacidade seja uma propriedade fundamental das línguas.

Embora nosso estudo represente um exame pontual, na medida em que averigua uma língua particular e considera apenas um parâmetro de transparência/opacidade, pode ser tomado como um ponto de partida para o entendimento de questões mais gerais sobre o tema, sobretudo quando se considera a ausência de pesquisas especificamente dedicadas ao estudo diacrônico da transparência linguística, o que confere um importante caráter de originalidade para a investigação aqui apresentada.

Nessa mesma direção, esta tese, para além do resultado sobre o grau de transparência do sistema de referência do PB, oferece contribuições relevantes justamente por formular uma abordagem consideravelmente nova dos fenômenos linguísticos tratados, praticando e chamando a atenção para uma forma diferente de olhar para esses fenômenos.

Além da perspectiva diacrônica na investigação da transparência, o trabalho manifesta uma abordagem nova na própria delimitação de seu objeto de estudo, já que a noção de *sistema de referência por expressão do argumento-sujeito* foi um conceito que não encontramos já delimitado em algum outro trabalho ou teoria, mas que precisou ser construído em nossa pesquisa, de modo a ser capaz de delimitar a parte do sistema gramatical do PB envolvida no fenômeno da RC.

Um outro aspecto do trabalho que constituiu uma decisão própria foi o tratamento do grau de transparência do PB em termos da frequência de uso dos fenômenos linguísticos envolvidos. A esse respeito, nossa pesquisa indica a possibilidade e a pertinência desse tipo de tratamento. Nesse sentido, cabe destacar também a formulação de nossa metodologia, também especificamente elaborada aqui, diante da necessidade colocada pelo objeto de estudo definido e por seu processamento quantitativo.

A realização de uma pesquisa que lida com questões pouco desenvolvidas ou ainda não abordadas em outros trabalhos ou teorias certamente impõe dificuldades e restrições, como a impossibilidade de fundamentar determinados conceitos e métodos em um

corpo de trabalhos pré-existente, mais extenso e especializado. Essa realidade, muitas vezes, exige a tomada de decisões de forma independente e demanda a elaboração de certas articulações teóricas e metodológicas que não seriam necessárias em outros tipos de pesquisa.

Por outro lado, esse problema não deixa de ser contrabalançado pelas possibilidades que investigações menos convencionais podem oferecer em termos do estímulo à formação de diferentes pontos de vista sobre os fatos linguísticos, bem como em termos do apontamento para novas perspectivas de frentes de trabalho. Com efeito, nosso estudo permite vislumbrar um amplo conjunto de alternativas de pesquisa sobre a diacronia da transparência linguística. Dentre outras opções, nosso trabalho aponta para a relevância de: (i) estender a análise do grau de transparência do sistema de referência para um recorte diacrônico maior, abarcando períodos anteriores ao início do século XIX; (ii) incluir na análise diacrônica da transparência desse sistema o fenômeno da simplificação da morfologia verbal nos usos dos pronomes *tu*, *nós* e *vocês* (que leva a construções como *tu/nós/vocês canta*), assim como o fenômeno do uso da flexão verbal de 1PP com a forma pronominal *a gente* (que origina construções como *a gente cantamos*); (iii) incluir na análise do sistema de referência os pronomes de 3ª pessoa (do singular e do plural); (iv) expandir a investigação diacrônica da transparência mediante exame de outros parâmetros de transparência/opacidade, além da RC; (v) analisar a diacronia da transparência em outras línguas, comparar os resultados e averiguar possíveis regularidades do processo etc.

Esperamos, enfim, que as conclusões teóricas e as constatações empíricas expostas nesta tese, as hipóteses que podem ser formuladas a partir desses resultados e os possíveis caminhos de pesquisa aqui listados possam servir como incentivo e base para o desenvolvimento de novos estudos sobre a diacronia da transparência linguística e sobre questões afins.

REFERÊNCIAS

BARME, S. Sujeito/ objeto explicito vs. não-realização no português brasileiro e europeu. In: GROSSE, S.; ZIMMERMANN, K. (Org.). “*Substandard*” e mudança no português do Brasil. Frankfurt: Teo Ferrer de Mesquita, 1998. p. 273-290.

_____. *Der subjektsausdruck beim verb in phonisch-nähesprachlichen varietäten des europäischen portugiesisch und brasilianischen*. Frankfurt: Lang, 2001.

BECHARA, E. *Moderna gramática portuguesa*. 37.ª ed. Rio de Janeiro: Nova Fronteira, 2009.

CÂMARA JR.; J. M. *A Estrutura da língua portuguesa*. 19.ª ed. Petrópolis: Vozes, 1989.

CHOMSKY, N. *The Minimalist Programm*. Massachussets: MIT Press, 1995.

CROFT, W. *Typology and universals*. Cambridge: Cambridge University Press, 2003.

DE OLIVEIRA, M. The pronominal subject in italian and brazilian portuguese. In: KATO, M. A.; NEGRÃO, E. V. (Ed.) *Brazilian portuguese and the null subject parameter*. Madrid: Iberoamericana; Frankfurt: Vervuert, 2000. p. 34-54.

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

DIK, S. *Functional grammar*. Amsterdam: North-Holland, 1978.

DUARTE, M. E. L. “Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil”. In: ROBERTS, I.; KATO, M. A. (Org.). *Português brasileiro: uma viagem diacrônica*. Campinas: Ed. da Unicamp, 1993. p. 107-128.

_____. *A perda do princípio “evite pronome” no português brasileiro*. 1995. 151f. Tese (Doutorado) – Universidade Estadual de Campinas, Campinas, 1995.

_____. O sujeito nulo no português do Brasil: de regra obrigatória a regra variável. In: GROSSE, S.; ZIMMERMANN, C. (Org.) “*Substandard*” e mudança no português do Brasil. Frankfurt: Teo Ferrer de Mesquita, 1998. p. 189-201.

_____. “The loss of the ‘avoid pronoun’ principle in brazilian portuguese”. In: KATO, M. A.; NEGRÃO, E. V. *Brazilian portuguese and the null subject parameter*. Madrid: Iberoamericana; Frankfurt: Vervuert, 2000. p. 17-36.

_____. Sujeito nulo/pleno e marcas de concordância. In: VOTRE, S.; RONCARATI, C. *Anthony Julius Naro e a linguística no Brasil: uma homenagem acadêmica*. Rio de Janeiro: 7 Letras, 2008. p. 265-277.

FARACO, C. A. O tratamento de você em português: uma abordagem histórica. *Fragmenta*, Curitiba, n. 13, p. 51-82, 1996.

GIVÓN, T. Iconicity, isomorphism and non-arbitrary coding in syntax. In: HAIMAN, J. *Iconicity in syntax*. Amsterdam: John Benjamins, 1985. p. 187-219.

GONÇALVES, S. C. L. *Projeto ALIP (Amostra Linguística do Interior Paulista): O português falado na região de São José do Rio Preto – constituição de um banco de dados anotado para o seu estudo*. São José do Rio Preto: UNESP (Relatório científico, FAPESP), 2007.

HAIMAN, J. The iconicity of grammar: Isomorphism and motivation. *Language*, Baltimore, v. 56, n. 3, p. 515-540, 1980.

_____. Iconic and economic motivation. *Language*, Baltimore, v. 59, n. 4, p. 781-819, 1983.

_____. *Natural syntax: iconicity and erosion*. Cambridge: Cambridge University Press, 1985.

HENGVELD, K. Transparency in Functional Discourse Grammar. *Linguistics in Amsterdam*, Amsterdam, v. 4, n. 2, p. 1-22, 2011a. Número especial (Transparent Languages).

_____. Epilogue: degrees of transparency. *Linguistics in Amsterdam*, Amsterdam, v. 4, n. 2, p.110-114, 2011b. Número especial (Transparent Languages).

_____. Referential markers and agreement markers in Functional Discourse Grammar. *Language Sciences*, Amsterdam, v. 34, n. 4, p. 468-479, 2012.

HENGVELD, K.; MACKENZIE, J. L. *Functional Discourse Grammar*. New York: Oxford University Press, 2008.

HENGVELD, K.; LEUFKENS, S. *Transparent and non transparent languages*. No prelo.

HODGE, C. T. The Linguistic cycle. *Language Sciences*. v.13, p. 1-7, 1970.

HOLMBERG, A.; NAYADU, A.; SHEEHAN, M. Three partial null-subject languages: a comparison of a Brazilian Portuguese, Finnish and Maratti. *Studia Linguistica*, v. 63, n.1, p. 59-97, 2009.

HOPPER, P. J. On some principles of grammaticalization. In: TRAUGOTT, E. C; HEINE, B. (Ed.) *Approaches to grammaticalization*. Philadelphia: John Benjamins Company, 1991. v.1, p. 17-35.

KAISER, G. A. Sobre a (alegada) perda do sujeito nulo no português brasileiro. In: LOBO, T. et al. (Ed.). *Novos dados, novas análises*. Salvador: Edufba, 2006. t. 1. p. 11-42.

_____. Losing the null subject: a contrastive study of (brazilian) portuguese and (medieval) french. In: KAISER, G. A.; REMBERGER, E. M. (Ed.). *Proceedings of the workshop null-subjects, expletives, and locatives in romance*. Konstanz: Universität Konstanz, Fachbereich Sprachwissenschaft, 2009. p. 131-156. (Arbeitspapier, 123).

KATO, M. A. O português são dois... ou três?. In: LOBO, T. et al. (Org.). *Rosae: linguística histórica, história das línguas e outras histórias* [online]. Salvador: Edufba, 2012. p. 93-108. Disponível em: <<http://books.scielo.org>>. Acesso em: 17 jun. 2015.

KATO, M. A; NEGRÃO, E. V. *Brazilian portuguese and the null subject parameter*. Frankfurt: Vervuert, 2000.

KEIZER, E. *A Functional Discourse Grammar for English*. Oxford: Oxford University Press, 2015.

LEUFENS, S. The transparency of creoles. *Journal of Pidgin and Creole Languages*, Amsterdam, v. 28, n. 2, p. 323-362, 2013.

_____. *Transparency in language: a typological study*. Utrecht: LOT, 2015.

LIGHTFOOT, D. W. A Theory of change. In: _____. *Principles of diachronic syntax*. Cambridge: Cambridge University Press, 1979. p. 120-166.

LIRA, S. A. *The subject in brazilian portuguese*. New York: Lang, 1996.

LOBO, M. Sujeito nulo: sintaxe e interpretação. In: RAPOSO, E. B. P. et al. *Gramática do português*. Lisboa: Fundação Calouste Gulbenkian, 2013. v. 2, p. 2309-2335.

LOPES, C. R. S. *A inserção de “a gente” no quadro pronominal do português: percurso histórico*. 1999. 181f. Tese (Doutorado em Língua Portuguesa) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1999.

_____. O quadro dos pronomes pessoais. In: VIEIRA, S. R.; BRANDÃO, S. F. (Org.). *Morfossintaxe e ensino de português: reflexões e propostas*. Rio de Janeiro: UFRJ, 2004. p. 151-178.

_____. Retratos da variação entre "você" e "tu" no português do Brasil: sincronia e diacronia. In: RONCARATI, C.; ABRAÇADO, J. (Org.). *Português brasileiro II: contato linguístico, heterogeneidade e história*. Niterói: EDUFF, 2008. v. 2. p. 55-71.

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

LOPES, C. R. S.; DUARTE, M. E. L. De Vossa Mercê a você: análise da pronominalização de nominais em peças brasileiras e portuguesas setecentistas e oitocentistas. In: BRANDÃO, S. F.; MOTA, M. A. (Org.). *Análise contrastiva de variedades do português: primeiros estudos*. Rio de Janeiro: In-Fólio, 2003. v. 1, p. 61-76.

LOPES, C. R. S. et al. Quem está do outro lado do túnel? Tu ou você na cena urbana carioca. *Neue Romania*, Berlin, n. 39, p.49-66, 2009.

LUCCA, N. N. G. *A variação tu/você na fala brasiliense*. 2005. Dissertação (Mestrado em Linguística). Universidade de Brasília, Brasília, 2005.

MACHADO, A. C. M. *A implementação de você no quadro pronominal: as estratégias de referência ao interlocutor em peças teatrais no século XX*. 2006. Dissertação (Mestrado em Letras Vernáculas). Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2006.

MEYER-HERMANN, R. Utilização e colocação dos pronomes pessoais no português falado do Brasil: indicações duma mudança tipológica em progresso? In: GROSSE, S.; ZIMMERMANN, K. (Org.). *"Substandard" e mudança no português do Brasil*. Frankfurt: Teo Ferrer de Mesquita, 1998. p. 203-225.

MODESTO, M. Null subjects without "rich" agreement. In: KATO, M. A.; NEGRÃO, E. V. *Brazilian portuguese and the null subject parameter*. Madrid: Iberoamericana; Frankfurt: Vervuert, 2000. p. 147-174.

NEGRÃO, E. V. A distribuição e a interpretação de pronomes na fala de crianças da escola pública. Artigo não-publicado. 1990.

NEGRÃO, E. V.; VIOTTI, E. Brazilian portuguese as a discourse-oriented language. In: KATO, M. A.; NEGRÃO, E. V. *Brazilian portuguese and the null subject parameter*. Madrid: Iberoamericana; Frankfurt: Vervuert, 2000. p. 105-126.

NEVES, M. H. M. *A Gramática Funcional*. São Paulo: Martins Fontes, 2004.

OLBERTZ, H.; SOUZA, G. C. Transparência e opacidade na variação da concordância verbal em português. In: CONGRESSO INTERNACIONAL DA ABRALIN, 9., 2015, Belém. *Caderno de Resumos do IX Congresso Internacional da ABRALIN – Simpósios temáticos*. Belém: Universidade Federal do Pará, 2015, p. 231-232.

OMENA, N. P.; BRAGA, M. L. A gente está se gramaticalizando? In: MACEDO, A. T.; RONCARATI, C.; MOLLICA, M. C. (Org.). *Variação e discurso*. Rio de Janeiro: Tempo Brasileiro, 1996. p. 75-84.

RAPOSO, E. P. *Teoria da gramática: a faculdade da linguagem*. Lisboa: Editorial Caminho, 1992.

RIZZI, L. On the form of chains: Criterial positions and ECP effects. In: CHENG, L.; CORVER, N. (Ed.). *On Wh Movement*. Cambridge: MIT Press, 2006. p. 97-134.

ROBERTS, I.; KATO, M. A. *Português brasileiro: uma viagem diacrônica*. Campinas: Ed. da UNICAMP, 1993.

RUBIO, C. F.; GONÇALVES, S. C. L. Opções metodológicas no estudo de fenômenos variáveis relacionados à primeira pessoa do discurso no plural. *Gragoatá*, Niterói, v. 15, n. 29, p. 161-182, 2010.

SANTANA, A. M. B. *Nós e a gente: um retrato do português popular de Salvador*. 2014. 114f. Dissertação (Mestrado em Estudos de Linguagens). Universidade do Estado da Bahia, Salvador, 2014.

SCHERRE, M. M. P.; NARO, A. J. Sobre a concordância de número no português falado do Brasil. In: RUFFINO, G. (Org.) *Dialettologia, geolinguística, sociolinguística. Atti del XXI Congresso Internazionale di Linguistica e Filologia Romanza*. Tübingen: Max Niemeyer Verlag, 1998. v. 5, p. 509-523.

SEARA, I. C. A variação do sujeito *nós* e *a gente* na fala florianopolitana. *Organon*, Porto Alegre, v. 14, n. 28-29, p. 179-194, 2000.

SLOBIN, D. I. Language change in childhood and in history. In: MACNAMARA, J. (Ed.). *Language learning and thought*. New York: Academic Press, 1977. p. 185-214.

_____. The repeated path between transparency and opacity in language. In: BELLUGI, U.; STUDDERT-KENNEDY, M. *Signed and spoken language: biological constraints on linguistic form*. Weinheim: Verlag Chemie, 1980. p. 229-243.

SOUZA, E. R. F.; COMPARINI, A. M.; GUIRALDELLI, L. A. Referência cruzada e concordância oracional no estudo da transparência e opacidade em línguas indígenas do Brasil. No prelo.

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

STASSI-SÉ, J. C. Transparência e opacidade na GDF: influência das línguas em contato no português europeu e brasileiro. In: CONGRESSO INTERNACIONAL ASSOCIAÇÃO DE LINGÜÍSTICA E FILOGIA DA AMÉRICA LATINA, 17., 2014, João Pessoa. *Anais da ALFAL 2014*. João Pessoa: Universidade Federal da Paraíba, 2014, p. 636.

TARALLO, F. *The filling of the gap: pro-drop rules in brazilian portuguese*. In: KING, L. D.; MALEY, C. A. *Selected papers from the XIIIth Linguistic Symposium on Romance Languages*. Amsterdam: John Benjamins, 1985. p. 355-376.

_____. Diagnosticando uma gramática brasileira: o português d'aquém e d'além-mar ao final do século XIX. In: ROBERTS, I.; KATO, M. A. (Org.). *Português brasileiro: uma viagem diacrônica*. Campinas: Ed. da Unicamp, 1993a. p. 69-105.

_____. Sobre a alegada origem crioula do português brasileiro: mudanças sintáticas aleatórias. In: ROBERTS, I.; KATO, M. A. (Org.). *Português brasileiro: uma viagem diacrônica*. Campinas: Ed. da Unicamp, 1993b. p. 35-68.

VIANA, V. *Os (des)caminhos do você: uma análise sobre variação e mudança na forma, na função e na referência do pronome você*. 2008. Tese (Doutorado em Linguística e Língua Portuguesa). Universidade Federal da Paraíba, João Pessoa, 2008.

VIANNA, J. B. S.; LOPES, C. R. S. A variação entre nós e a gente: uma comparação entre o português europeu e o português brasileiro. *Revista do GELNE*, Natal, v. 14, p. 95-116, 2012. Número especial.

REFERÊNCIAS UTILIZADAS NA CONSTITUIÇÃO DO CORPUS

ALENCAR, J. *As azas de um anjo: comedia em um prologo, quatro actos e um epilogo*. Rio de Janeiro: B. L. Garnier, Editor, 1860.

_____. *O demonio familiar: comedia em quatro actos*. Rio de Janeiro: Typographia de Soares & Irmão., 1858.

_____. *Verso e reverso: comedia em dois actos*. Rio de Janeiro: B. L. Garnier, Editor., 1864.

ANDRADA, M. F. R. *Januario Garcia: o sete orelhas: drama em 3 actos e 5 quadros*. São Paulo: Typographia do Governo., 1849.

ARAÚJO, A. M. *Teatro de Artur Azevedo*. Rio de Janeiro: INACEN; Brasília: Ministério da Cultura, 1987. (Coleção clássicos do teatro brasileiro, v. 7, t. 3).

_____. *Teatro de Artur Azevedo*. 2. ed. Rio de Janeiro: Funarte; Brasília: Ministério da Cultura, 1995. (Coleção clássicos do teatro brasileiro, v. 8, t. 6).

ASSIS, M. *Desencantos: phantasia dramatica*. Rio de Janeiro: Editor Paula Brito, 1861.

_____. *O caminho da porta; comedia em um acto*. Rio de Janeiro: Livraria C. Teixeira, 1863.

_____. *Theatro*. Rio de Janeiro: Typographia do Diario do Rio de Janeiro, 1863. v. 1.

ASSUMPÇÃO, L. *Onze peças de Leilah Assumpção*. Rio de Janeiro: Casa da Palavra, 2010.

ATHAYDE, R. *Crime e impunidade e outras peças*. Rio de Janeiro: Record, 1983.

AZEVEDO, A. *Os milagres de S. José de Riba-Mar: comedia de costumes em 3 actos e 4 quadros*. Typ. da Alfaiataria Teixeira, 1899.

BARBOSA, M. C.; FALABELLA, M. *Querido mundo e outras peças*. Rio de Janeiro: Lacerda Editores, 2003.

BORTOLLOTO, M. E éramos todos Thunderbirds. In: FRATESCHI, C. *Ágora livre dramaturgias*. São Paulo: Ágora Teatro, 2006. 146-19.

BRASIL, B. *Cheiro de chuva; Novas diretrizes em tempos de paz*. São Paulo: Aliança Francesa: Imprensa Oficial de Estado de São Paulo, 2007.

BUARQUE, C. *Ópera do malandro*. São Paulo: Cultura, 1978.

CAFEZEIRO, E.; GUERRA, R. *Teatro completo de Araújo Porto-Alegre*. Rio de Janeiro: Funarte; Brasília: Ministério da Cultura, 1997. t. 2.

CALLADO, A. *Forró no Engenho de Cananéia*. Rio de Janeiro: Editôra Civilização Brasileira S. A, 1964.

CAMARGO, J. *Deus lhe pague: comedia em 3 actos divididos em 9 quadros*. Rio de Janeiro: Livraria Educadora, Braga & Valverde, 1936.

DAMASCENO, D. *Teatro de Martins Pena: comédias*. Rio de Janeiro: Ministério da Educação e Cultura; Instituto Nacional do Livro, 1956.

FRANÇA JÚNIOR, J. J. *Meia hora cynismo: por [...] estudante do 4ª ano da Faculdade de Direito de S. Paulo*. São Paulo: J.R. de A. Marques, 1861.

_____. *O defeito de família*: comedia em um acto por França Junior. Rio de Janeiro: Typographia Americana, 1871.

_____. *Cahio o ministerio!*: comedia original de costumes em tres actos. Rio de Janeiro: Livraria Popular de A.A. da Cruz Coutinho Livreiro Editor, 1883.

_____. *A lotação dos bonds*: comedia em um acto por França Junior. Rio de Janeiro: Livraria de Cruz Coutinho, Editor, 1885.

FERNANDES, M. *Os órfãos de Jânio*. Porto Alegre: L& PM Editores, 1979. (Coleção teatro de Millôr Fernandes, v. 9).

_____. *Duas tábuas e uma paixão*. Porto Alegre: L& PM Editores, 1982. (Coleção teatro de Millôr Fernandes, v. 12).

GOMES, A. D. *A invasão; A revolução dos beatos*: teatro. Rio de Janeiro: Editora Civilização Brasileira, 1962.

_____. *O berço do herói*: um prólogo e dois atos. Rio de Janeiro: Editora Civilização Brasileira, 1965.

GONZAGA, A. *Comedias*. São Paulo: Livraria Teixeira, Vieira Pontes & Cia, 1940.

MACEDO, J. M. *Luxo e vaidade*: comedia original em 5 actos. Rio de Janeiro: Typ. de Francisco de Paula Brito, 1860.

_____. *A torre em concurso*: comedia burlesca em tres actos. Rio de Janeiro: B. L. Garnier, Editor, 1865.

_____. *Remissão de peccados*: comedia em cinco actos. Rio de Janeiro: A. A. Da Cruz Coutinho, 1870.

_____. *Romance de uma velha*: comedia em cinco actos. Rio de Janeiro: Na livraria de Cruz Coutinho – Editor, 18-?.

MARINHO, T.; GADELHA, C.; SAADI, F. *Machado de Assis*: teatro completo. Brasília: Ministério da Educação e Cultura; Rio de Janeiro: Serviço Nacional de Teatro, 1982.

PENA, L. C. M. *Os irmãos das almas*: comedia em um acto. Rio de Janeiro: [s.n.], 1847

_____. *O noviço*: comédia em 5 actos. Rio de Janeiro: Emp. Typ. Dous de Dezembro de P. Brito, 1853.

POSSOLO, H. *Sobre a arte de cortar bifes*. In: FRATESCHI, C. *Ágora livre dramaturgias*. São Paulo: Ágora Teatro, 2006. 96-114.

_____. *Nóis otário(s)*. São Paulo: Giostri Editora Ltda, 2012.

PRATA, M. *Fábrica de chocolate: um ato*. São Paulo: Editora Hucitec, 1979.

RODRIGUES, N. *Álbum de família e vestido de noiva*. Rio de Janeiro: Edições do Povo, 1946.

SOUZA, N. A. *Teatro*. Coimbra: Cena Lusófona, 2005.

TABORDA, A. N. (Org.). *Teatro (in)completo de Isis Baião*. Rio de Janeiro: Pedrazul, 2003.

TOJEIRO, G. *Onde canta o sabiá: comédia em 3 atos*. Brasília: Ministério da Educação e Cultura; Rio de Janeiro: Serviço Nacional de Teatro, 1973.

TERMO DE REPRODUÇÃO XEROGRÁFICA

Autorizo a reprodução xerográfica do presente Trabalho de Conclusão, na íntegra ou em partes, para fins de pesquisa.

São José do Rio Preto, 28/08/2017.

Assinatura do autor