



UNIVERSIDADE ESTADUAL DE MATO GROSSO DO SUL
UNIDADE UNIVERSITÁRIA DE CAMPO GRANDE

LUIZ LEANDRO GOMES DE LIMA

GRAUS DE SALIÊNCIA DAS VARIÁVEIS E VARIANTES LINGUÍSTICAS:

A percepção da variação por falantes cultos do português

Campo Grande/MS
2016

LUIZ LEANDRO GOMES DE LIMA

GRAUS DE SALIÊNCIA DAS VARIÁVEIS E VARIANTES LINGUÍSTICAS:

A percepção da variação por falantes cultos do português

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Letras, da Universidade Estadual de Mato Grosso do Sul, Unidade Universitária de Campo Grande, como requisito parcial para obtenção do título de Mestre em Letras.

Área de concentração: Produção do Texto Escrito e Oral

Orientador: Profa. Dra. Adriana Lúcia de Escobar Chaves de Barros

Campo Grande/MS
2016

L698g Lima, Luiz Leandro Gomes de
Graus de saliência das variáveis e variantes linguísticas: a percepção da variação por falantes cultos do português/Luiz Leandro Gomes de Lima. Campo Grande: UEMS, 2016.
137p. ; 30cm.

Dissertação (Mestrado) – Letras – Universidade Estadual de Mato Grosso do Sul, 2016.

Orientadora: Profa. Dra. Adriana Lúcia de Escobar Chaves de Barros.

1.Sociolinguística 2.Variação linguística 3. Saliência linguística I.Título.

CDD 23.ed. 306.44

LUIZ LEANDRO GOMES DE LIMA

GRAUS DE SALIÊNCIA DAS VARIÁVEIS E VARIANTES LINGUÍSTICAS:

A percepção da variação por falantes cultos do português

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Letras da Universidade Estadual de Mato Grosso do Sul, Unidade Universitária de Campo Grande, como requisito parcial para obtenção do título de Mestre em Letras.

Área de concentração: Produção do Texto Escrito e Oral

COMISSÃO EXAMINADORA

Prof.^a Dr.^a Adriana Lúcia de Escobar Chaves de Barros
(Presidente)

Universidade Estadual de Mato Grosso do Sul/ UEMS

Prof. Dr. Nataniel dos Santos Gomes (Titular)

Universidade Estadual de Mato Grosso do Sul/ UEMS

Prof.^a Dr.^a Patrícia Graciela da Rocha (Titular)

Universidade Federal de Mato Grosso do Sul/ UFMS

Prof.^a Dr.^a Natalina Sierra Assencio Costa (Suplente)

Universidade Estadual de Mato Grosso do Sul/ UEMS

Prof. Dr. Ruberval Franco Maciel (Suplente)

Universidade Estadual de Mato Grosso do Sul/ UEMS

Campo Grande/MS, 08 de novembro de 2016

AGRADECIMENTOS

Terminar esta dissertação só foi possível graças à colaboração e compreensão de muitas pessoas, todas muito queridas por estarem comigo nessa jornada de quase dois anos. Gostaria de agradecer a todos que contribuíram de alguma forma para que eu pudesse levar a cabo esta pesquisa, em particular:

Aos meus pais, Maria e Raimundo, por terem me dado toda a educação que eles não tiveram, por terem me incentivado em todos os passos da minha carreira acadêmica e profissional e por abaixarem o volume da TV quando eu precisava escrever;

À profa. Adriana, minha orientadora, por ter acolhido minhas ideias, por ter me encorajado a mantê-las e por sua habilidade única de orientar uma pesquisa sem, contudo, interferir na pesquisa em si. Se o resultado desse trabalho é positivo, isso se deve à leitura e às sugestões da profa. Adriana;

Aos professores Nataniel e Ruberval, membros da banca de qualificação, por contribuírem imensamente com meu trabalho com suas sugestões e toques; à profa. Patrícia, membro da banca de defesa, pelas valiosas sugestões; aos professores das disciplinas que cursei na UEMS, por terem aberto minha mente para o mundo da pesquisa em linguagem;

Aos meus informantes, por cederem 20 minutos de seu tempo para me ajudar a entender um pouco de como funciona nossa percepção das regras variáveis do português brasileiro;

Aos professores, amigos e colegas da UFMS (do curso de Letras, da PREG), aos amigos e colegas da UEMS, aos amigos e colegas de fora do meio acadêmico;

À Camilla, por entender minha ausência em certos momentos, por me apoiar nesse longo caminho e por aturar minhas elucubrações sem fim.

LIMA, L. L. G. *Graus de saliência das variáveis e variantes linguísticas: a percepção da variação por falantes cultos do português*. 148 f. Dissertação (Mestrado em Letras) – Universidade Estadual de Mato Grosso do Sul, Campo Grande/ MS, 2016.

RESUMO

A variação linguística se apresenta nas línguas naturais por meio de regras variáveis. Cada variável tem duas ou mais variantes, e a escolha de qual variante usar depende de fatores estruturais (internos ao sistema linguístico) e sociais (extralinguísticos). Neste trabalho, investigamos em que medida esses fatores influenciam também a percepção dos usuários da língua, tentando identificar quais variáveis e variantes são mais salientes para os falantes, e quais contextos linguísticos mais contribuem para essa saliência. Utilizamos-nos do arcabouço teórico da Sociolinguística Variacionista (LABOV, 2008 [1972]; WEINREICH, LABOV e HERZOG, 2006 [1968]) e estudamos três importantes regras variáveis do português brasileiro: a concordância verbal, a retomada anafórica por objeto direto e as estratégias de relativização. A pesquisa se baseou em um *corpus* obtido a partir de testes de percepção aplicados a quarenta informantes que se encaixam na definição de "falante culto" (com escolaridade superior completa e antecedentes biográfico-culturais urbanos). Os resultados encontrados nos mostraram que, das três variáveis estudadas, a concordância verbal é a mais saliente, sendo a mais estigmatizada quando usada em sua variante não padrão. Das possibilidades de realização da variável "retomada anafórica por objeto direto", os resultados não mostraram que há uma mais saliente que outra; já para a variável "estratégias de relativização", os resultados mostraram que a variante "copiadora" é altamente estigmatizada pelos falantes cultos, ao passo que a variante "cortadora" passa despercebida na maioria dos contextos (e, mesmo quando percebida, é mais tolerada). Em geral, nosso trabalho demonstrou que de fato há variantes e variáveis mais salientes que outras, e que há fatores contextuais que podem aumentar ou diminuir essa saliência.

Palavras-chave: Sociolinguística. Variação Linguística. Saliência Linguística.

LIMA, L. L. G. *Degrees of salience in linguistic variables and variants: the perception of variation by Portuguese educated speakers*. 148 f. Dissertação (Mestrado em Letras) – Universidade Estadual de Mato Grosso do Sul, Campo Grande/ MS, 2016.

ABSTRACT

Linguistic variation manifests itself in natural languages by means of variable rules. Each variable has two or more alternative variants, and the choice between which one to use depends on structural (internal to the linguistic system) and social (extralinguistic) factors. This study investigates to what extent those factors can also influence the perception of the language users. We try to identify which variables and variants are more salient to the speakers, and which linguistic contexts contribute more to their salience. We use the theoretical framework of the Variationist Sociolinguistics (LABOV, 2008 [1972]; WEINREICH, LABOV e HERZOG, 2006 [1968]) to look into three important variable rules in Brazilian Portuguese: the subject-verb agreement, the anaphoric resumption by a direct object and the relativization strategies. We extracted the data from a *corpus* obtained from perception tests we applied to forty informants that fitted the "educated speaker" criteria (having an undergraduate degree and having an urban background). The results we found show that, among the three variable we studied, the subject-verb agreement is the most salient, being the most stigmatized when the non-standard variant is used. The results also show that, among the three possible realizations of the anaphoric resumption by a direct object, there is not a single one that is more salient. For the "relativization strategies" variable, the "resumptive pronoun" variant is highly stigmatized by the educated speakers, whereas the "pronoun-chopping" variant goes unnoticed in most contexts (and, even when it is noticed, it is often tolerated). Our study shows that there are in fact some variants and variables that are more salient than others, and that there are some context factor that can raise or decrease this salience.

Keywords: Sociolinguistics. Linguistic Variation. Linguistic Salience.

LISTA DE QUADROS

| | |
|---|----|
| Quadro 1: Codificação das variáveis, variantes, fatores e categorias no teste de percepção.. | 70 |
| Quadro 2: Questões da variável "concordância verbal", bloco 1 (categoria "sujeito preposto distante do verbo")..... | 71 |
| Quadro 3: Questões da variável "concordância verbal", bloco 2 (categoria "sujeito preposto imediatamente ao verbo")..... | 72 |
| Quadro 4: Questões da variável "concordância verbal", bloco 3 (categoria "sujeito posposto ao verbo") | 72 |
| Quadro 5: Questões da variável "retomada anafórica por objeto direto" | 74 |
| Quadro 6: Questões da variável "estratégias de relativização" | 75 |
| Quadro 7: Questões extras da variável "estratégias de relativização" | 76 |

LISTA DE TABELAS

| | |
|--|-----|
| Tabela 1: Células sociais e número de informantes..... | 78 |
| Tabela 2: Marcações das sentenças não padrão para as três variáveis | 85 |
| Tabela 3: Marcação das sentenças não padrão para as três variáveis, separadas por sexo/ gênero | 87 |
| Tabela 4: Marcação das sentenças não padrão para as três variáveis, separadas por faixa etária | 88 |
| Tabela 5: Comparação das orações não padrão com sujeito preposto distante (variável "concordância verbal") | 92 |
| Tabela 6: Comparação das orações não padrão com sujeito preposto distante e com sujeito preposto imediatamente ao verbo (variável "concordância verbal") | 96 |
| Tabela 7: Comparação das orações padrão com a variável "concordância verbal" | 100 |
| Tabela 8: Comparação das orações não padrão com a variável "concordância verbal" | 102 |
| Tabela 9: Comparação das três variantes da variável "objeto direto anafórico" | 104 |
| Tabela 10: Comparação das orações não padrão com estrutura simples e estrutura de predicativo (variável "objeto direto anafórico") | 111 |
| Tabela 11: Comparação das orações padrão da variável "objeto direto anafórico" | 115 |
| Tabela 12: Comparação das orações não padrão da variável "objeto direto anafórico" | 116 |
| Tabela 13: Comparação das três variantes da variável "estratégias de relativização" | 118 |
| Tabela 14: Comparação das orações não padrão com relativa objeto direto e relativa adjunto adverbial (variável "estratégias de relativização") | 124 |
| Tabela 15: Comparação das orações padrão da variável "estratégias de relativização" | 127 |
| Tabela 16: Comparação das orações não padrão da variável "estratégias de relativização" | 127 |

LISTA DE GRÁFICOS

| | |
|--|-----|
| Gráfico 1: Questão CV-1-ADF "Esses políticos aqui no Brasil não entendem que quem manda é o povo" | 89 |
| Gráfico 2: Questão CV-2-ADF "Eles sem nenhum pudor leva materiais de escritório para casa" | 89 |
| Gráfico 3: Questão CV-1-ADG, "Os clientes e os vendedores neste nosso contexto precisam falar a mesma língua" | 90 |
| Gráfico 4: Questão CV-2-ADG, "Os chefes e os funcionários por incrível que pareça se comunica muito bem por aqui" | 90 |
| Gráfico 5: Questão CV-1-AEF, "Sobre os malotes, eles em pouco tempo chegam para o cliente" | 91 |
| Gráfico 6: Questão CV-2-AEF, "As impressoras na pior hora sempre para de funcionar" | 91 |
| Gráfico 7: Questão CV-1-AEG, "O caminhão e a van quando eu dirijo não levam nenhuma multa" | 92 |
| Gráfico 8: Questão CV-2-AEG, "O computador e a impressora pelo que vi ontem precisa de conserto" | 92 |
| Gráfico 9: Questão CV-1-BDF, "Meus amigos conhecem esse novo estoquista" | 93 |
| Gráfico 10: Questão CV-2-BDF, "Esses atendentes acredita mesmo nessa abordagem?" | 94 |
| Gráfico 11: Questão CV-1-BDG, "Os familiares e amigos podem usar as instalações" | 94 |
| Gráfico 12: Questão CV-2-BDG, "Os colegas e clientes considera sua saída uma pena" | 94 |
| Gráfico 13: Questão CV-1-BEF, "Gosto desses novos produtos, mas eles precisam de uma boa publicidade" | 95 |
| Gráfico 14: Questão CV-2-BEF, "As peças chega em fevereiro" | 95 |
| Gráfico 15: Questão CV-1-BEG, "Aposto que o carro e a moto pertencem ao chefe" | 95 |
| Gráfico 16: Questão CV-2-BEG, "A pista e o depoimento completa a solução do problema" | 95 |
| Gráfico 17: Questão CV-1-CDF, "Nascem muitos bebês nessa época do ano" | 97 |
| Gráfico 18: Questão CV-2-CDF, "Todos os dias chega vários compradores de São Paulo" .. | 97 |
| Gráfico 19: Questão CV-1-CDG, "Existem um gerente e um contador exatamente para isso" | 98 |
| Gráfico 20: Questão CV-2-CDG, "Se alguém fica sabendo desse erro, morre eu e você" | 98 |
| Gráfico 21: Questão CV-1-CEF, "Existem muitos problemas com essa nova abordagem" ... | 99 |
| Gráfico 22: Questão CV-2-CEF, "Antigamente saía várias vendas por dia. Hoje está mais difícil" | 99 |
| Gráfico 23: Questão CV-1-CEG, "É possível que cresçam o interesse e a procura por nossos produtos" | 100 |
| Gráfico 24: Questão CV-2-CEG, "Causou-me estranheza o acidente e a morte numa rua tão calma" | 100 |
| Gráfico 25: Questão OD-1-HL, "Quando a secretária voltou do almoço, eu a cumprimentei pelo bom trabalho" | 105 |
| Gráfico 26: Questão OD-2-HL, "Os clientes chegaram e levei eles até a sala de espera" | 105 |
| Gráfico 27: Questão OD-3-HL, " Onde está o office-boy? Faz muito tempo que não vejo" .. | 106 |
| Gráfico 28: Questão OD-1-HM, "Comprei os materiais e já os deixei no depósito" | 106 |
| Gráfico 29: Questão OD-2-HM, "Sempre que uma cadeira quebra, eu joga ela fora" | 107 |

| | |
|---|-----|
| Gráfico 30: Questão OD-3-HM, "Um dos celulares ainda funciona. Quer que eu deixe aqui com você?" | 107 |
| Gráfico 31: Questão OD-1-IL, " O estagiário andava muito distraído, mas eu o deixei em estado de alerta" | 108 |
| Gráfico 32: Questão OD-2-IL, "Acho a Maria uma pessoa difícil, mas o João acha ela simpática" | 108 |
| Gráfico 33: Questão OD-3-IL, "Conheço o Carlos, mas não considero um bom administrador" | 109 |
| Gráfico 34: Questão OD-1-IM, "Mostramos os três currículos para o chefe, mas nenhum o deixou satisfeito" | 109 |
| Gráfico 35: Questão OD-2-IM, "Esta sala costumava ser horrível, mas a Sílvia tornou ela agradável" | 110 |
| Gráfico 36: Questão OD-3-IM, "A nova campanha não me agradou muito, mas o cliente achou ótima" | 110 |
| Gráfico 37: Questão OD-1-JL, "O novo atendente já está bem treinado. Eu já o deixo fazer tudo sozinho" | 112 |
| Gráfico 38: Questão OD-2-JL, "Eu não confio nesses políticos porque eu já vi eles falando mal do povo" | 112 |
| Gráfico 39: Questão OD-3-JL, "O chefe brigou com o office-boy e mandou sair da sala" .. | 112 |
| Gráfico 40: Questão OD-1-JM, "Não sei para que serve essa máquina, eu nunca a vi funcionando" | 113 |
| Gráfico 41: Questão OD-2-JM, "A impressora estava travando, mas eu fiz ela funcionar de novo" | 114 |
| Gráfico 42: Questão OD-3-JM, "Essa sua cadeira precisa de óleo! Você não ouve rangendo?" | 115 |
| Gráfico 43: Questão ER-1-NQ, "O distribuidor em quem eu confiei está nos passando para trás" | 119 |
| Gráfico 44: Questão ER-2-NQ, "O Messias é o vendedor que eu mais gosto!" | 119 |
| Gráfico 45: Questão ER-3-NQ, "O diretor que eu discordei dele estava muito seguro de sua opinião" | 120 |
| Gráfico 46: Questão ER-1-NR, "As regras com as quais eu concordei não são essas" | 120 |
| Gráfico 47: Questão ER-2-NR, "Sinto muito, mas esse não é o modelo que eu gosto" | 120 |
| Gráfico 48: Questão ER-3-NR, "Amanhã já é a reunião que você precisa participar dela" .. | 120 |
| Gráfico 49: Questão ER-1-OQ, "O rapaz com que eu falei é loiro e meio gordinho" | 122 |
| Gráfico 50: Questão ER-2-OQ, "O nome do cliente que eu falei é Marcelo" | 122 |
| Gráfico 51: Questão ER-3-OQ, "A menina que nós estudamos com ela já está fazendo doutorado" | 122 |
| Gráfico 52: Questão ER-1-OR, "O dia em que você chegar na hora será um dia chuvoso" . | 123 |
| Gráfico 53: Questão ER-2-OR, "A cidade que eu nasci é a 100 km daqui" | 123 |
| Gráfico 54: Questão ER-3-OR, "Esse não é o carro que eu vim ontem com ele" | 123 |
| Gráfico 55: Questão ER-1-PQ, "O cliente cuja compra eu fechei gastou mais de 10 mil reais!" | 125 |
| Gráfico 56: Questão ER-2-PQ, "Os dois candidatos que eu olhei os currículos são ótimos" | 125 |
| Gráfico 57: Questão ER-3-PQ, "Os diretores que a reunião deles foi cancelada já tinham embarcado" | 125 |
| Gráfico 58: Questão ER-1-PR, "Aquele é o telefone cuja linha está com defeito" | 126 |
| Gráfico 59: Questão ER-2-PR, "O caminhão que eu consertei a roda já estragou de novo" | 126 |

| | |
|---|-----|
| Gráfico 60: Questão ER-3-PR, "Os produtos que a validade deles já venceu foram jogados fora" | 126 |
| Gráfico 61: Questão ER-QE-1, "Esse é o livro que eu te falei sobre" | 128 |
| Gráfico 62: Questão ER-QE-2, "Foi essa ideia que nós votamos contra" | 128 |

LISTA DE FIGURAS

| | |
|--|----|
| Figura 1: Questionário no Google Docs | 80 |
| Figura 2: "Randomizador" on-line | 80 |
| Figura 3: Questões já embaralhadas | 81 |
| Figura 4: Declaração de ciência | 81 |
| Figura 5: Planilha do Google Docs | 82 |

SUMÁRIO

| | |
|---|----|
| AGRADECIMENTOS | 5 |
| RESUMO | 6 |
| ABSTRACT | 7 |
| LISTA DE QUADROS..... | 8 |
| LISTA DE TABELAS..... | 9 |
| LISTA DE GRÁFICOS | 10 |
| LISTA DE FIGURAS | 13 |
| SUMÁRIO | 14 |
| INTRODUÇÃO | 16 |
| Objetivos..... | 18 |
| Justificativa..... | 19 |
| Organização da Dissertação | 20 |
| 1 FUNDAMENTAÇÃO TEÓRICA | 21 |
| 1.1 Teoria da Variação Linguística | 21 |
| 1.1.1 <i>Variação, Variante, Variável, Variedade</i> | 23 |
| 1.1.2 <i>Variação e Mudança</i> | 24 |
| 1.1.3 <i>Contínuos de Variação</i> | 25 |
| 1.2 O Conceito de Norma e os Falantes Cultos..... | 27 |
| 1.2.1 <i>Norma culta e Norma padrão</i> | 28 |
| 1.2.2 <i>Gramaticalidade e Agramaticalidade</i> | 32 |
| 1.2.3 <i>Comunidade de Fala e Atitudes Linguísticas</i> | 34 |
| 1.3 Saliência Linguística | 37 |
| 1.3.1 <i>Duarte e Tarallo (1988): Saliência e Mudança Linguística</i> | 39 |
| 1.3.2 <i>Rác (2013): Buscando um Modelo de Operacionalização da Saliência</i> | 41 |

| | |
|--|-----|
| 1.3.3 Níveis de Saliência de Variantes e Variáveis | 42 |
| 2 PROCEDIMENTOS DE PESQUISA | 46 |
| 2.1 A Escolha das Regras Variáveis | 46 |
| 2.1.1 O Que Dizem as Gramáticas Normativas | 48 |
| 2.1.2 O Que Dizem as Pesquisas Linguísticas | 53 |
| 2.2 A Elaboração do Teste de Percepção | 67 |
| 2.2.1 A Confeção das Questões..... | 69 |
| 2.2.2 Se a Linguística Estuda a Fala, por que Testes Escritos? | 76 |
| 2.3 Informantes e Dados | 77 |
| 2.3.1 A Escolha dos Informantes | 77 |
| 2.3.2 Aplicação dos Questionários e Tratamento dos Dados | 79 |
| 2.3.3 Alguns problemas | 82 |
| 3 ANÁLISE DOS DADOS..... | 84 |
| 3.1 Variáveis em Conjunto | 84 |
| 3.2 A Variável Concordância Verbal | 88 |
| 3.3 A Variável Objeto Direito Anafórico | 103 |
| 3.4 A Variável Estratégias de Relativização | 117 |
| 4 CONSIDERAÇÕES FINAIS | 129 |
| REFERÊNCIAS | 134 |

INTRODUÇÃO

A Teoria da Variação Linguística, também conhecida como Sociolinguística Variacionista (WEINREICH, LABOV e HERZOG, 2006 [1968]; LABOV, 2008 [1972]), tem por princípios duas características universais das línguas naturais: a heterogeneidade inerente e sistemática e o relativismo cultural (BORTONI-RICARDO, 2014). Segundo o princípio da heterogeneidade inerente e sistemática, toda língua natural é marcada pela **variação**, e essa variação não é caótica, mas *ordenada*. Já o princípio da relatividade cultural aponta para o fato de que nenhuma língua ou variedade linguística pode ser considerada *inferior* ou *subdesenvolvida*, não obstante o nível sociocultural de seus falantes.

Essa variação inerente às línguas naturais pode ser percebida pela imensa quantidade de **regras variáveis** encontradas no uso dos falantes nativos. Cada regra variável se realiza por meio de duas ou mais **variantes**. Uma dessas variantes (raramente mais de uma) é considerada **padrão** pela comunidade de falantes, o que significa que, dentre as possibilidades de aplicação da regra, a variante padrão é a única recomendada pelos comandos normativos, nomeadamente os gramáticos tradicionais e suas gramáticas normativas. Por exemplo: no português, a concordância verbal é uma regra variável com duas possibilidades: conformação do verbo ao sujeito (“Eles vieram rápido”) *versus* não conformação do verbo ao sujeito (“Eles veio rápido”). Nesse caso, a primeira variante é considerada padrão, ao passo que a segunda não.

Segundo a Teoria da Variação Linguística, há vários fatores – linguísticos e sociais – que *condicionam* a escolha de uma variante ou outra. Os fatores de origem social, em geral, são os que mais influenciam o uso das variantes. Sabemos, por exemplo, que quanto menor a escolaridade do falante, maiores são as chances de ele usar a regra da não conformação do verbo ao sujeito que vimos acima (cf., por exemplo, LEMLE e NARO, 1977). Sabemos ainda que fatores como idade, sexo/ gênero, posição social, redes de contato, a situação de interação, etc. também influenciam a escolha de uma ou outra variante (cf. MOLLICA e BRAGA, 2013).

Já os fatores de origem linguística representam uma pressão mais sutil, por serem de mais difícil percepção, mas ainda assim determinante na escolha de qual variante usar. Esses fatores linguísticos se referem aos *contextos* em que as variáveis ocorrem – certos contextos podem inibir ou favorecer o uso de uma variante ou outra. Por exemplo: pesquisas sociolinguísticas¹ apontam que, no caso da regra variável da concordância verbal, em contextos

¹ Lemle e Naro (1977), Monguilhott (2001), Oliveira (2005) e Monte (2007), para citar algumas pesquisas.

nos quais o sujeito precede o verbo (“Os pacotes chegaram”), há uma incidência maior do uso da concordância que em contextos onde o sujeito é posposto ao verbo (“Chegou os pacotes”).

Além de influenciar a *produção linguística* dos falantes, os contextos que favorecem o uso de uma variante não padrão podem também agir na *percepção* dos falantes, *salientando* essa forma linguística, deixando facilmente perceptível a agramaticalidade da construção de acordo com a norma padrão; ou *camuflando* a forma não padrão, deixando menos perceptível a impropriedade gramatical, mesmo para falantes cultos. No exemplo da concordância verbal, o contexto “sujeito posposto ao verbo” encobre a falta de conformação do sujeito ao verbo, deixando essa impropriedade gramatical menos perceptível, enquanto que o contexto “sujeito anteposto ao verbo” salienta essa falta de concordância e a deixa evidente. Esses fatores contribuiriam para a **estigmatização** de certas variantes e variáveis em detrimento de outras.

Outro exemplo de regra variável no português brasileiro (doravante PB) é a regra de retomada anafórica por objeto direto de 3ª pessoa (ou, para simplificar, "objeto direto anafórico"). Há três possibilidades correntes²:

- (1) a. Conheço o Pedro, mas tem muito tempo que não vejo ele
- b. Conheço o Pedro, mas tem muito tempo que não o vejo.
- c. Conheço o Pedro, mas tem muito tempo que não vejo Ø.

Nas três sentenças, o sintagma nominal *o Pedro* é retomado por três estratégias de preenchimento do objeto direto distintas: em (1a), a anáfora é feita por meio do pronome reto *ele*; em (1b), pelo pronome clítico *o*; e em (1c), por uma categoria vazia, isto é, um sintagma sem realização fonética. A sentença (1b) traz a única possibilidade de retomada prevista na norma padrão e autorizada pelas gramáticas normativas. No entanto, pesquisas linguísticas indicam que, até mesmo entre falantes cultos, a variante em (1c), chamada de *anáfora zero*, é a mais utilizada em situações de interação oral³, ao passo que a variante em (1a), chamada de *retomada lexical*, é evitada por falantes de maior escolaridade.

Assim como acontece com a regra variável da concordância verbal, os contextos em que o objeto direto anafórico ocorre também podem exercer influência sobre a escolha dentre as variantes disponíveis, podendo camuflar o uso de uma variante não padrão. Sabemos, por exemplo, que um objeto direto representado por uma entidade com o traço semântico [–

² Exemplos retirados de Bagno (2004).

³ Por exemplo: Duarte (1989), Corrêa (1991).

animado] tende a favorecer o uso da anáfora zero, ao mesmo tempo em que deixa esse uso não padrão menos evidente:

(2) Eu juntei o lixo e joguei \emptyset fora

A *percepção* do uso impróprio da realização do objeto direto anafórico também é afetada por esses contextos, na medida em que o objeto direto com traço semântico [– animado], ao mesmo tempo que induz ao uso da anáfora zero, também “camufla” essa variante, mesmo aos olhos de falantes cultos. A sentença em (2) é mais “natural”, para muitos falantes, que a sentença em (1c), apesar de em ambas acontecer o mesmo fenômeno: a categoria vazia no lugar do objeto.

Neste trabalho, buscamos evidenciar em que medida os contextos em que se inserem as variáveis linguísticas influenciam a percepção dos falantes cultos brasileiros relativamente a usos não padrão, de forma a identificar quais seriam as variantes mais salientes e os contextos que mais salientam essas variantes. Para realizar esta pesquisa, escolhemos três regras variáveis historicamente importantes no português: a **concordância verbal**, a **retomada anafórica por objeto direto**, das quais já falamos, e as **estratégias de relativização**. Por meio de testes de percepção, observamos o julgamento que falantes cultos da língua fazem a respeito das diversas possibilidades de uso dessas três regras variáveis, levantando os contextos que mais foram subjetivamente apontados como intensificadores da saliência dos usos não padrão.

A hipótese adotada é a de que certas construções, mesmo desviantes das normas gramaticais padrão, não seriam percebidas pelos falantes cultos como impropriedades gramaticais, devido aos contextos em que se encontram “camuflarem” essas impropriedades. Sendo assim, essas construções não sofreriam de estigmatização por parte da comunidade linguística, ao passo que as variáveis e variantes mais salientes sofreriam dessa discriminação.

Objetivos

O objetivo geral desta pesquisa é identificar como os falantes cultos do PB percebem os usos variáveis típicos nessa língua.

Os objetivos específicos são três: buscamos identificar e discutir, considerando as regras variáveis escolhidas para análise:

- i. Quais das *variáveis* recebem mais atenção e estigmatização dos falantes cultos, isto é, quais regras são mais salientes e quais são menos salientes;

- ii. Qual(is) das *variantes* que compõem cada variável os falantes cultos mais estigmatizam e qual(is) mais identificam como correta(s) (e mais neutra(s), se for o caso) – em outras palavras, qual variante é mais saliente e qual é menos saliente; e
- iii. Quais *contextos* (fatores condicionadores) mais afetam a percepção dos usos impróprios segundo a gramática normativa; em outras palavras, quais fatores linguísticos mais contribuem para a saliência de uma variante e quais contribuem para sua "camuflagem".

Justificativa

Esta pesquisa foi motivada por um princípio elaborado por Bortoni-Ricardo (2014), ao abordar o tema “impacto da Sociolinguística na educação”. Segundo a autora, para a aplicação dos subsídios da Sociolinguística a questões educacionais, é preciso ter em mente que “a escola deve ocupar-se principalmente das regras variáveis que recebem avaliação negativa na sociedade, enfatizando as que são *mais salientes*” (p. 160, grifo nosso). Justificamos o estudo aqui empreendido pela ausência de pesquisas que busquem evidenciar quais seriam essas regras mais salientes. Como não seria possível fazer uma pesquisa com todas as regras variáveis do português (até porque muitas delas ainda sequer foram estudadas a fundo), escolhemos três das regras mais abordadas desde que a pesquisa em Sociolinguística se iniciou no Brasil, e buscamos descobrir os graus de saliência por elas apresentados.

Além disso, há poucos trabalhos em Sociolinguística que tratam da percepção linguística (cf. CLOPPER e PISONI, 2002): a variabilidade é geralmente pensada apenas do ponto de vista do falante, e raramente da perspectiva do ouvinte, ao passo que as pesquisas sobre a percepção linguística em geral enxergam a variabilidade como ruído, como um problema a ser vencido para que a pesquisa possa continuar (cf. LOPES, 2012). Sendo assim, nosso trabalho também se justifica pela necessidade de mais pesquisas feitas sob esse enfoque, que alia os problemas sobre a percepção linguística aos pressupostos teóricos da Sociolinguística Variacionista.

Citamos ainda os possíveis frutos de que poderemos dispor ao fim do nosso trabalho: os resultados desta pesquisa podem, por um lado, contribuir para traçar um quadro geral das percepções e atitudes linguísticas dos falantes do PB frente ao fenômeno da variação linguística; e, por outro, fornecer subsídios ao ensino de língua materna, apontando as regras variáveis que merecem mais atenção por parte dos educadores, seja porque apresentam maiores graus de estigmatização por parte da sociedade, seja porque podem apresentar variantes neutras, que passam despercebidas pelo público geral, apesar de representarem impropriedades segundo a

gramática normativa. Por último, um quadro das atitudes dos falantes cultos pode ajudar a identificar tendências de mudanças linguísticas, na medida em que a avaliação social das variáveis linguísticas é um fator determinante na difusão ou desaparecimento de determinada variante (cf. 1.1.2 e 1.2.3).

Organização da Dissertação

Este trabalho está dividido em três capítulos. No primeiro, discutiremos os conceitos teóricos que fundamentaram esta pesquisa, circunscrevendo-a aos pressupostos da Sociolinguística Variacionista. Essa revisão de literatura inclui os conceitos mais básicos dessa disciplina (como "variável" e "variante"), traz uma discussão a respeito dos conceitos de "norma culta" e "norma padrão" e apresenta as ideias que subjazem a teoria da saliência linguística que utilizamos para analisar a percepção dos informantes.

No segundo capítulo, descreveremos as escolhas e os passos metodológicos seguidos no desenvolvimento da pesquisa: a escolha das regras variáveis, uma revisão do que dizem trabalhos prescritivos e descritivos sobre essas regras e os passos de elaboração e aplicação dos testes de percepção, instrumento que utilizamos para coletar os dados.

No terceiro capítulo, exporemos os dados obtidos e discutiremos os resultados que podem ser extraídos desses dados, a partir da base teórica adotada.

Por fim, nas considerações finais, faremos um resumo dos resultados obtidos, além de uma análise das limitações da pesquisa e dos possíveis passos seguintes a partir deste trabalho.

1 FUNDAMENTAÇÃO TEÓRICA

O objetivo deste capítulo é apresentar os conceitos que subjazem esta pesquisa. Uma vez que nossa intenção é estudar os "graus de saliência das variáveis e variantes linguísticas" por meio da "percepção da variação por falantes cultos do português", temos que definir o que são *variáveis* e *variantes* (cf. 1.1), quem são esses tais *falantes cultos* (cf. 1.2) e o que entendemos por *saliência* (cf. 1.3).

1.1 Teoria da Variação Linguística

Esta pesquisa se subscrive teoricamente à Teoria da Variação Linguística ou Sociolinguística Variacionista, subárea da Linguística que estuda “a língua em uso no seio das comunidades de fala, voltando a atenção para um tipo de investigação que correlaciona aspectos linguísticos e sociais” (MOLLICA, 2013, p. 9). Essa escola de pesquisa linguística surgiu a partir do trabalho de William Labov e outros linguistas americanos na década de 1960, como uma reação ao Gerativismo de Noam Chomsky. Na época, as pesquisas linguísticas seguiam basicamente os pressupostos teóricos da Linguística Gerativa, que tem por objeto a **competência linguística** dos falantes, isto é, os processos cognitivos *subjacentes* à produção linguística. Por ser uma teoria de base formalista, podendo inclusive ser encarada como uma continuação do Estruturalismo bloomfieldiano (cf. LYONS, 1987, cap. 7), o Gerativismo deixa de fora o *componente social* da linguagem: não importa a essa corrente teórica o uso concreto da língua em interações sociais reais, vividas por indivíduos histórica, política e culturalmente definidos, mas tão somente as operações mentais que dão origem às orações e textos da língua. Por isso se diz que o Gerativismo é uma teoria mentalista. Para Labov, que buscava uma teoria mais social que mentalista, essa omissão é inadmissível.

Em Weinreich, Labov e Herzog (2006 [1968]), trabalho fundador da Sociolinguística Variacionista, os autores já criticavam longamente a postura homogeneizadora das escolas linguísticas surgidas até então: desde o Estruturalismo saussuriano, passando pelo Descritivismo de Bloomfield até o Gerativismo chomskyano, buscava-se *relativizar* a variação linguística, abstraindo a heterogeneidade da língua, tratando o fenômeno da variação como marginal e desconsiderável face as regularidades do sistema linguístico. Labov e seus companheiros, por outro lado, insistiram na possibilidade de sistematizar essa variação, de “pôr ordem no caos” aparente das línguas naturais (cf. TARALLO, 1986).

A Sociolinguística Variacionista⁴ surgiu, dessa forma, assentada em duas premissas básicas: a **heterogeneidade inerente e sistemática** e o **relativismo cultural**.

O conceito de heterogeneidade diz respeito exatamente ao fato de que todas as línguas naturais exibem variação: nenhuma língua se apresenta, em nenhum momento histórico, como homogênea e invariável. A heterogeneidade é uma característica universal e necessária das línguas – é *inerente*, no sentido de que “o sistema linguístico que serve a uma comunidade heterogênea e plural deve ser também heterogêneo e plural para desempenhar plenamente as suas funções” (LUCCHESI e ARAÚJO, 2016). Além disso, a heterogeneidade é *sistemática*, uma vez que pode ser sistematicamente relacionada a fatores *sociais*, externos ao sistema, e *linguísticos*, internos ao sistema: pressupõe-se a existência de uma *ordem* por trás da heterogeneidade da língua falada. Segundo Bortoni-Ricardo (2014, p. 157), “os recursos da variação, que toda língua natural oferece, estão sistematicamente organizados em sua estrutura e contribuem para tornar a comunicação entre os falantes mais produtiva e adequada”.

É aqui que se dá o primeiro ponto de contato, promovido pela Sociolinguística, entre língua e sociedade. Segundo o modelo teórico-metodológico da Sociolinguística Variacionista, a variação linguística está intimamente ligada aos *estratos* sociais, refletindo nas regras da língua a estratificação da sociedade. Na introdução deste trabalho, já demos o exemplo da regra de concordância verbal e citamos pesquisas que indicam que a escolaridade do falante pode interferir em suas escolhas linguísticas. Outros aspectos sociais também podem ser interferentes na língua: as mulheres, em geral, usam mais formas prestigiadas e menos palavrões que os homens (cf. PAIVA, 2013); as pessoas mais velhas tendem a usar formas linguísticas mais antigas que os mais jovens, que por sua vez preferem formas inovadoras (cf. NARO, 2013); é possível perceber claramente as diferenças de pronúncias entre pessoas que vivem em estados diferentes (aquilo que chamamos de "sotaque"); etc.

O relativismo cultural, por sua vez, é um conceito advindo das ciências sociais, em especial da Antropologia Cultural, segundo o qual uma manifestação cultural prestigiada pela sociedade não é intrinsecamente superior às demais. Aplicando-se este princípio à Linguística, temos que nenhuma língua ou variedades de uma mesma língua podem ser consideradas inferiores – pelo menos não do ponto de vista linguístico. Com efeito, algumas línguas de povos que considerávamos “primitivos” (como as línguas indígenas e africanas), segundo o obsoleto

⁴ Há outros tipos de Sociolinguística, como a Sociolinguística Interacional (cf. BORTONI-RICARDO, 2014), e todas tratam de algum modo da variação linguística. No entanto, reserva-se tradicionalmente o termo “variacionista” à escola iniciada por Labov, que também pode ser chamada de “quantitativa” ou “laboviana”.

conceito de cultura como “civilização”, mostram-se extremamente complexas, ricas e produtivas do ponto de vista morfológico e sintático. Da mesma forma, todas as variedades de uma língua de civilização, como o português ou o francês, têm um alto grau de regularidade e seguem regras internas coerentes e sistemáticas, tendo suas próprias normas de correção – normas essas que todos os falantes seguem naturalmente.

Nos próximos itens discutiremos alguns conceitos centrais na teoria Sociolinguística, corolários dos princípios gerais da heterogeneidade inerente e do relativismo cultural: as variáveis linguísticas e suas variantes, a relação da variação com mudança linguística e a teoria dos contínuos de variação.

1.1.1 Variação, Variante, Variável, Variedade

A **variação** linguística, como fenômeno universal, pressupõe a existência de regras linguísticas que podem se superficializar por meio de mais de uma forma linguística. Essas formas linguísticas alternativas são chamadas de **variantes**. Por exemplo: no Brasil, há duas possibilidades de pronúncia do primeiro fonema consonantal da palavra *tia*. A pronúncia mais difundida nas regiões Sul, Sudeste e Centro-Oeste é a pronúncia *palatizada*, chamada tecnicamente de "africada palatal surda": [tʃi.ɐ]. Já em boa parte da região Nordeste e nas variedades europeias do português, a forma mais comum é a "oclusiva alveolar surda": [ti.ɐ]. Nesse caso, ambas as possibilidades de pronúncia são variantes.

Tarallo (1986, p. 8) afirma que "variantes linguísticas são diversas maneiras de se dizer a mesma coisa em um mesmo contexto e com o mesmo valor de verdade". Essas diversas maneiras de se dizer a mesma coisa configuram um fenômeno variável, ou regra variável, ou simplesmente **variável** linguística. Assim, dizemos que a regra de superficialização do fonema /t/ diante de vogal anterior alta (como o /i/) é uma variável *binária* (tem duas possibilidades). Já a regra de realização do objeto direto anafórico, que comentamos na Introdução, é uma regra variável *ternária* (tem três possibilidades).

As variáveis linguísticas são também chamadas de variáveis dependentes, pois a escolha entre variantes *depende* de outras variáveis (que são *independentes*), que aqui chamaremos de **fatores condicionadores**. Por exemplo: a variável dependente “superficialização do fonema /t/ diante de vogal anterior alta” (como em *tia*) é afetada por fatores de ordem social, uma vez que o uso da variante africana ou oclusiva depende da origem geográfica do falante. Além disso, esse fenômeno também tem motivação linguística, pois é o contexto linguístico que determina

o uso da variante africada: ela só ocorre diante de vogal anterior alta. Como se pode observar, o uso de uma ou outra variante é influenciado por **fatores linguísticos** (estruturais, internos ao sistema da língua) ou **sociais** (extralinguísticos, ligados ao contexto sociocultural da interação).

Sendo assim, a Sociolinguística Variacionista tem por objeto a língua em seu contexto sociocultural, uma vez que parte da explicação para a heterogeneidade inerente e sistemática pode ser encontrada em fatores externos ao sistema linguístico. Como observa Mollica (2013, p. 10), “os usos de estruturas linguísticas são motivados e as alternâncias configuram-se por isso sistemáticas e estatisticamente previsíveis”. Um estudo sociolinguístico, portanto, tem por objetivo descrever fenômenos variáveis, analisando e sistematizando as variantes linguísticas usadas em uma comunidade de fala e relacionando-as aos fatores sociais e linguísticos que as desencadeiam. Isso é feito calculando-se a influência que cada fator interno ou externo ao sistema linguístico tem na realização de cada variante.

Um conjunto de variáveis pode identificar uma **variedade** linguística. Esse termo é similar a “dialeto” e, assim como este, tem problemas de definição. Monteiro (2008) comenta que há muita dificuldade em definir o conceito de variedade, que foi cunhado na literatura linguística exatamente para substituir “dialeto”, devido à carga negativa que essa expressão guarda em certos contextos⁵. Wardhaugh (2010) define *variedade* em termos de conjuntos específicos de itens linguísticos que podem ser biunivocamente associados a algum fator externo (geográfico ou social). Dessa forma, podemos falar em “variedade campo-grandense do português” ou “variedade culta do português” ou ainda em “variedade do português falada por médicos”. Em última análise, esse termo está intimamente ligado ao conceito de *norma linguística*, que discutiremos adiante.

1.1.2 Variação e Mudança

Por mais que a Sociolinguística Variacionista tenha em sua história oferecido ferramentas importantes para realizar “fotografias sociolinguísticas” das variedades de uma língua, isto é, descrições *sincrônicas* dos fenômenos variáveis correntes na língua, a análise sociolinguística tem por objetivo original estabelecer uma relação entre os processos de

⁵ Usa-se, na linguagem cotidiana, o termo “dialeto” para se referir a variedades menos prestigiadas de uma língua. Diz-se, por exemplo, que o português brasileiro é “apenas um dialeto” do português europeu. Outro uso do termo proveniente do senso comum é o que imagina a existência de uma variedade padrão da língua, mais pura e original, que teria vários “dialetos” (suas corrupções), indo de encontro ao pensamento científico contemporâneo, que considera a língua como um *conjunto de dialetos* (ou variedades, ou normas), sem que nenhum seja o “original”.

variação que se observam num determinado momento (sincronia) e os processos de **mudança linguística**, que tomam parte na história da língua em sua evolução (diacronia).

Ao investigar os fatores sociais que influenciam as escolhas das variantes, busca-se delinear o quadro de variação observado na variedade linguística em análise em termos da dicotomia entre **variação estável** e **mudança em progresso**. A variação estável ocorre quando não se percebe a tendência de predominância de uma variante sobre a(s) outra(s), do que se depreende que o quadro de variação tende a se manter ainda por algum tempo. Já quando os resultados estatísticos indicam que uma das variantes está se sobrepondo às demais, o diagnóstico é de mudança em progresso. Isso indica que a variante mais utilizada em pouco tempo poderá se tornar a forma categórica e se generalizar, fazendo com que as outras variantes caiam em desuso.

Essa análise é feita correlacionando-se a estrutura linguística variável com fatores da estrutura social e observando-se como uma determinada variante estaria se difundindo entre os diversos segmentos sociais. Para levar a cabo esse estudo empírico das mudanças linguísticas em andamento, Labov (2008 [1972], p. 193) propõe que essa tarefa seja subdividida em três problemas: o **problema da transição**, que diz respeito ao caminho da mudança e se pergunta como uma língua muda de um estágio para outro sem interferir na comunicação entre os falantes; o **problema da avaliação**, que salienta que os falantes podem ter atitudes negativas ou positivas em relação a determinadas variantes, o que poderia definir a tendência de mudança que essa avaliação social favorece; e o **problema do encaixamento**, que se propõe a encontrar as relações entre os fatores linguísticos e sociais que interferem na mudança. Essas informações teriam um papel fundamental para o esclarecimento de como a mudança linguística se completa (o que por si só constitui um outro problema, o “problema da implementação”; cf. WEINREICH, LABOV e HERZOG, 2006 [1968]).

O “problema da avaliação” é de extrema importância para esta pesquisa. Como veremos (cf. 1.2.3), é a partir da avaliação que os falantes cultos fazem das variáveis linguísticas que discutiremos o *status* da saliência das variáveis e variantes em português.

1.1.3 Contínuos de Variação

Bortoni-Ricardo (2004) apresenta um modelo de descrição da variação linguística que distribui as variedades de uma língua em três “contínuos”: o contínuo rural-urbano, o contínuo oralidade-letramento e o contínuo da monitoração estilística. Para a autora, cada contínuo deve

ser imaginado como uma reta: em cada ponto da reta estão os usos linguísticos considerados opostos (os falares mais rurais de um lado e os mais urbanos de outro, os textos mais orais de um lado e os mais letrados de outro, etc.). Por toda a extensão da reta há fases intermediárias entre um extremo e outro, de forma que as diversas variedades se sobrepõem e se entrecruzam, sem divisões estanques, em um degradê linguístico.

No **contínuo rural-urbano** temos, em uma das pontas da reta, os falares rurais mais isolados e sem nenhuma influência da urbanização; e, na ponta oposta, os falares mais urbanos, cuja principal característica é a influência da padronização linguística e da codificação escrita. No meio desse contínuo encontra-se um grupo relativamente novo na história da urbanização do Brasil, que é formado por pessoas que migraram das áreas rurais e preservam parte de seus antecedentes culturais, sobretudo os linguísticos. Bortoni-Ricardo (2004) denomina os falares desses grupos de “rurbanos”. Incluem-se ainda nesse grupo as variedades de regiões interioranas semirurais, que sofrem forte influência da mídia urbana.

Bortoni-Ricardo (2004, p. 53) fala de traços descontínuos e traços graduais: há um conjunto de itens típicos dos falares rurais que não são usados nos falares urbanos (a autora cita exemplos retirados de uma história com o personagem Chico Bento: *inté, prantei, ponhei*, etc.), isto é, são descontinuados à medida em que se aproxima do perímetro urbano; enquanto que outros traços (*ocê, dexei, dibaxo...*) têm uma distribuição gradual entre os vários pontos do contínuo rural-urbano. Essa distinção nos será útil ao tratarmos dos *níveis* de saliência linguística (cf. 1.3.3).

O **contínuo oralidade-letramento** apresenta em um dos seus polos a cultura da oralidade (preponderante na zona rural) e no outro a cultura do letramento (preponderante na zona urbana). Os eventos de letramento são baseados em textos escritos ao passo que os eventos de oralidade são apoiados em textos falados. Apesar de cada tipo de evento estar situado em uma das pontas do contínuo, é comum que eles apareçam juntos: a aula, por exemplo, é um evento de letramento permeado de intervenções orais.

O **contínuo da monitoração estilística** situa em um de seus extremos as situações de fala totalmente espontâneas e sem planejamento, e do outro as situações com a maior carga de planejamento e antecipação, que exigem uma maior atenção do falante. Os falantes se ajustam a situações mais ou menos monitoradas automaticamente, como parte de sua competência linguística. Fatores que influenciam a monitoração incluem o ambiente de interação, o interlocutor ou o tópico da conversa. Por exemplo: prestamos muito mais atenção à forma com

que falamos quando estamos em uma situação formal, como quando falamos em um tribunal, diante de um juiz, do que quando nos encontramos em uma situação informal, como quando conversamos com amigos no bar.

Considerando as características da urbanização do Brasil e o alcance atual dos meios de comunicação, é possível dizer que as variedades que exercem, hoje, maior imposição sobre as demais, são aquelas faladas pelas populações urbanas, com uma renda média-alta, e que, por isso, garantem para si bons níveis de escolaridade e acesso aos bens da cultura escrita (cf. FARACO, 2008, p. 46). Utilizando o modelo dos contínuos, podemos situar esses falares mais próximos do polo urbano no eixo rural-urbano e do polo letramento no eixo oralidade-letramento.

1.2 O Conceito de Norma e os Falantes Cultos

Como já vimos, nenhuma língua é uma entidade unitária e homogênea – a heterogeneidade faz parte da essência das línguas naturais. O conceito de **norma**, dentro dos estudos linguísticos, surgiu para captar essa diversidade em uma época em que a Sociolinguística ainda não havia definido termos como “variação” e “variedade”. Segundo Faraco (2008), podemos entender *norma* como

Determinado conjunto de fenômenos linguísticos (fonológicos, morfológicos, sintáticos e lexicais) que são correntes, costumeiros, habituais numa dada comunidade de fala. Norma nesse sentido se identifica com normalidade, ou seja, com o que é corriqueiro, usual, habitual, recorrente (“normal”) numa certa comunidade de fala (FARACO, 2008, p. 37).

Dessa forma, podemos entender norma linguística como um conjunto de usos e atitudes linguísticas comuns a determinados grupos sociais e que os pode identificar: as normas urbanas, as normas rurais, as normas dos jovens, dos empresários, dos *rappers*, etc. Em outras palavras, norma pode ser entendida como um sinônimo de variedade (cf. 1.1.1).

Podemos dizer, assim, que uma língua é um conjunto de normas. Faraco (2008, p. 33) explica que “não existe língua para além ou acima do conjunto das suas variedades constitutivas, nem existe língua de um lado e as variedades do outro”. A língua é a própria soma de suas normas (ou variedades). Além disso, é importante ter em mente que essas normas são hibridizadas, pois não há separação distinta entre elas: as normas se sobrepõem na medida em que são a expressão dos diversos grupos sociais, e estes são também sobrepostos (uma mesma pessoa pode ser ao mesmo tempo proveniente da região rural, jovem, empresário e *rapper*,

dominando todas as normas desses grupos). Esse contato entre normas poderia também desencadear mudanças linguísticas (“problema do encaixamento”; cf. 1.1.2).

O conceito de norma surgiu a partir das concepções introduzidas pelo linguista Eugenio Coseriu. Esse autor, em seu texto clássico "Sistema, norma y habla" (cf. COSERIU, 1967), critica a dicotomia saussuriana entre língua e fala. Para Saussure, o objeto de estudo da Linguística é a **língua** (em francês, *langue*) tomada em si mesma e por si mesma, isto é, o sistema de signos linguísticos que estabelecem relações entre si e formam uma estrutura autônoma, desvinculada de fatores externos (históricos e sociais); e a **fala** (*parole*), que seria a atualização desse sistema, o uso concreto das regras e signos da língua em situações reais de enunciação. Para Coseriu, entre a língua e a fala, há a norma. Borba (1991, p. 49) explica que, para Coseriu, “a norma é um primeiro grau de abstração. Sendo a língua um conjunto de possibilidades, a norma aparece como um dado número de realizações dela num dado momento e lugar”. Faraco (2008, p. 36), retomando o caráter social do conceito, identifica *norma* para Coseriu como “cada um dos diferentes modos sociais de realizar os grandes esquemas de relações do sistema. Nesse sentido, cada norma se organiza como um certo arranjo de possibilidades admitidas pelo sistema”. Vale frizar que esse conceito de norma não se liga à ideia de “lei”, de “normativa”, e sim à ideia de “comum” ou “normal”, de estatisticamente comprovado.

Neste item, apresentamos os conceitos de norma culta e norma padrão para, posteriormente, explicar o que vem a ser o *falante culto*, um dos personagens principais da nossa pesquisa. Para tanto, é preciso considerar o conceito de (a)gramaticalidade em Linguística (sentenças gramaticais em uma norma podem ser agramaticais em outras), assim como as noções de comunidade de fala (*grosso modo*, cada norma pertence a uma comunidade) e de atitude linguística (a atitude dos falantes cultos em relação às variantes e variáveis).

1.2.1 Norma culta e Norma padrão

Em todas as línguas de civilização encontramos aquilo que Lopes (1995) chamou de “problema da normatividade”. O normativismo é um fenômeno histórico-político que considera a existência de uma norma linguística que se identifica com padrões de correção e pureza, em detrimento de várias outras normas que, ao se distanciarem da norma padrão, não podem ser consideradas corretas e não podem servir de modelo a ser imitado. Enquanto o conceito de norma que vimos no item anterior se liga à ideia de *normal*, isto é, aquilo que acontece com frequência, que é uma tendência geral e habitual, esse outro conceito de norma se liga à ideia

de *normativo*, isto é, aquilo que é ideal, definido por juízos de valor, por um conjunto de regras que se devem seguir.

Como narra Lyons (1979), a ideia da existência de um “padrão de correção” linguístico se iniciou, no Ocidente, a partir das reflexões linguísticas dos gregos em Alexandria, por volta do século III a.C. Naquela época, havia um grande interesse pelos poemas homéricos, mas esses textos encontravam-se bastante corrompidos. Ao restaurar os textos originais, os filólogos gregos perceberam que a língua falada por eles próprios e seus contemporâneos era muito diferente do grego de Homero. A admiração por este e outros autores do passado levou à crença de que a própria língua em que escreveram era mais “pura” e “correta” que a língua coloquial de Alexandria. Assim surgiu o costume de escrever **gramáticas**⁶: tratados linguísticos que tinham os objetivos de elucidar dificuldades linguísticas que pudessem perturbar a leitura dos textos antigos e, ao mesmo tempo, recomendar o uso das regras da linguagem literária de Homero como ideais de correção. Lyons (1979, p. 9-10) chamou essa atitude grega de “erro clássico” no estudo da língua: primeiro, consideraram a língua escrita (literária e antiga) como modelo de correção para a língua falada; segundo, supunham que a língua dos escritores antigos era mais correta que a fala coloquial do seu tempo e que a pureza de uma língua era mantida pelas pessoas cultas e corrompida pelos iletrados. Como afirma Lopes (1995),

Guardadas as devidas proporções, pode-se ver, na intransigência com que certos gramáticos se esforçam por incutir no povo as regras da norma culta de suas línguas, uma sobrevivência daquele mesmo preconceito linguístico dos primeiros gregos (LOPES, 1995, p. 27).

Como já vimos (cf. 1.1), as línguas são compostas de conjuntos de regras variáveis, sendo que cada regra variável pode se substanciar por meio de duas ou mais variantes. Segundo Tarallo (1986), as variantes de uma comunidade de fala encontram-se sempre em relação de concorrência: padrão *versus* não padrão, conservadoras *versus* inovadoras e de prestígio *versus* estigmatizadas. “Em geral, a variante considerada padrão é, ao mesmo tempo, conservadora e aquela que goza do prestígio sociolinguístico na comunidade. As variantes inovadoras, por outro lado, são quase sempre não-padrão e estigmatizadas pelos membros da comunidade” (TARALLO, 1986, p. 12). Desde os gregos, o conjunto de variantes mais conservadoras⁷ perfazem aquilo que se convencionou chamar de “norma culta” ou “norma padrão”. Essa norma é prestigiada em nossa sociedade (é ensinada nas escolas e cobrada nos vestibulares e

⁶ Palavra de origem grega que, no original, significava “a arte de escrever”.

⁷ Ou pelo menos é isso que se supõe. Há estudiosos (cf., por exemplo, BAGNO, 2003; TARALLO, 1983) que mostram que as variantes populares (não padrão, estigmatizadas) não têm nada de inovadoras, pois fazem parte da língua há muitos séculos.

concursos). A “norma culta” se opõe às normas não padrão ou populares, que não gozam do mesmo prestígio e nem da mesma aceitação em ambientes formais e acadêmicos.

Esses dois termos, “norma culta” e “norma padrão”, são geralmente usados como sinônimos. No entanto, nos últimos tempos tem havido um movimento para diferenciar esses dois conceitos. Vejamos a história de cada um deles e quais são suas definições mais modernas, oferecidas pelos estudos linguísticos.

Uma vez que as regras da gramática normativa são derivadas de escritos literários antigos⁸, especialmente daqueles considerados “clássicos”, usava-se tradicionalmente o termo *norma culta* para designar o modelo de língua que essas gramáticas tentam preservar. É o que dizem, por exemplo, Cunha e Cintra (2001, p. xxiv) no prefácio de sua gramática: “Trata-se de uma tentativa de descrição do português atual na sua forma culta, isto é, da língua como a têm utilizado os escritores portugueses, brasileiros e africanos do Romantismo para cá”. Cegalla (1990, p. xix) não cita a literatura, mas apresenta sua gramática dizendo que sua obra “pretende ser uma Gramática Normativa da Língua Portuguesa, conforme a falam e escrevem as pessoas cultas na época atual”. Curiosamente, apesar de citar a língua falada, o autor não usa exemplos dessa modalidade para abonar suas prescrições.

É possível concluir que esses autores associam o adjetivo “culto” à língua literária dos grandes escritores clássicos, seguindo a tradição grega. Nesse sentido, a “norma culta” seria uma língua ideal, abstrata, “que não corresponde a nenhum conjunto real das regras que governam a atividade linguística por parte dos falantes de carne e osso” (BAGNO, 2003, p. 50). Essa atitude normativista diante dos fatos da língua criou uma série de critérios dicotômicos usados para qualificar as produções linguísticas: “certo” *versus* “errado”, “belo” *versus* “feio”, “culto” *versus* “inculto”.

No entanto, para a Linguística moderna, que tem a *língua falada* como principal objeto de estudo (cf. LYONS, 1987, p. 24-29) e que se propõe ser *científica* (cf. LOBATO, 1986, p. 21-34), não é possível adotar o mesmo conceito de língua culta da tradição gramatical ocidental. Em primeiro lugar porque seu objeto não pode ser tão abstrato e ideal: é preciso colher dados *reais* de uso da língua a partir de interações reais entre seus falantes. Em segundo lugar porque, por adotar um caráter científico, a Linguística não pode fazer julgamentos de valor de seu objeto

⁸ Nas principais gramáticas do gênero atuais, é possível encontrar prescrições abonadas por exemplos de autores dos sécs. XVIII e XIX. Cf., por exemplo, Bechara (2006) e Cunha e Cintra (2001).

como fazem os estudos tradicionais⁹. Sendo assim, foi necessário criar um outro conjunto de noções para o uso do rótulo “norma culta”.

Para a Linguística moderna (em especial para os estudos de Sociolinguística), a **norma culta** “se refere à linguagem concretamente empregada pelos cidadãos que pertencem aos segmentos mais favorecidos da nossa população” (BAGNO, 2003, p. 51) ou a “o conjunto de fenômenos linguísticos que ocorrem habitualmente no uso dos falantes letrados em situações mais monitoradas de fala e escrita” (FARACO, 2008, p. 73). Esse é o conceito adotado por grandes projetos de descrição da norma culta do português, como, por exemplo, o projeto NURC (Norma Urbana Culta; cf. SILVA, 1996). Nesse caso, os **falantes cultos** são definidos a partir de dois critérios: escolaridade superior completa¹⁰ e antecedentes biográfico-culturais urbanos. Essa norma culta não é ideal, mas *real*: os falantes cultos a utilizam no seu dia a dia, em situações de comunicação formal. Sendo assim, a norma culta para a Linguística está ligada ao sentido de *normal*, isto é, comportamento que pode ser observado e medido. Utilizando o modelo dos contínuos (cf. 1.1.3 Contínuos de Variação), podemos identificar o lugar ocupado por essa nova noção de norma culta num ponto próximo do polo urbano, embora se saiba que muitos traços da norma culta sejam graduais, de forma que até mesmo os falares mais rurais utilizam praticamente as mesmas regras das normas urbanas (aqueles que não são graduais são exatamente os que sofrem maior estigmatização). No eixo da monitoração estilística, essas variedades cultas conhecem, como as demais, diferentes estilos, desde os menos até os mais monitorados.

Já para designar o modelo ideal de língua “correta”, derivado da tradição gramatical grega, muitos linguistas têm proposto o termo **norma padrão**. A norma padrão não é real, mas *ideal*: é um construto sócio-histórico que serve de *referência* para o uso linguístico em situações de alto monitoramento estilístico. Ela é descrita nas gramáticas normativas e se consolida como um conjunto de *prescrições*, e não como usos *reais* da língua. É por isso que a norma padrão se aproxima do sentido “normativo” do termo norma.

⁹ Da mesma forma que não faz sentido um botânico achar que uma planta é “errada” ou um astrônomo não querer estudar um planeta porque o considera “feio”. Se fenômenos novos aparecem, é dever do cientista explicá-los, e não os julgar. Cabe aqui trazer a reflexão de Lopes (1995, p. 27), segundo o qual “não cabe ao linguista ser contra a normatividade, ou a favor dela: o que lhe compete é insistir no fato de que a problemática da gramaticalidade (não confundir com o problema da correção ou do purismo linguístico) é matéria legitimamente linguística”.

¹⁰ “Chama-se de culto o falante que detém grau de escolarização formal classificável de universitário completo, isto é, um indivíduo que fez todo o percurso dos três níveis de educação: o fundamental (8 anos), o médio (3 anos) e o superior (4-5 anos em média)” (BAGNO, 2000, p. 179-180).

Utilizando o contínuo da monitoração estilística, podemos dizer que a norma padrão se situa acima do extremo do alto monitoramento, e por isso ela é ideal: trata-se de uma *intenção* de se aproximar ao máximo desse extremo. Os falantes cultos usam a língua tentando utilizar o máximo de regras da norma padrão que aprenderam em sua formação escolar. Isso acontece porque, segundo Milroy (2007), esses falantes seguem uma "ideologia da padronização linguística":

Observa-se que muitas das línguas usadas em larga escala, tais como o inglês, o francês ou o espanhol, têm uma variedade **padrão**, e isso afeta a maneira pela qual os falantes pensam sobre sua própria língua e sobre a linguagem em geral. Pode-se dizer que os falantes dessas línguas, ao contrário de falantes de línguas menos difundidas, vivem em **culturas de língua padrão**. Nessas culturas, as **atitudes linguísticas** [cf. 1.2.3] são dominadas por poderosas posições ideológicas que são em grande parte baseadas na suposta existência de uma forma padronizada; essas posições, em seu conjunto, constituem uma **ideologia da padronização linguística** ou uma 'ideologia da língua padrão'. Os falantes geralmente não têm consciência de que são condicionados por essas posições ideológicas: em geral, eles acreditam que suas atitudes em relação à língua são o senso comum e assumem que virtualmente todos os outros falantes compartilham das mesmas posições (MILROY, 2007, p. 133, grifos do autor, tradução nossa).

Sendo assim, a norma culta (realmente falada pelos falantes cultos) é uma *tentativa* de se aproximar da norma padrão, sem contudo se sobrepor a ela. No capítulo 3, quando faremos a análise dos dados obtidos dos informantes dessa pesquisa, faremos referência à norma padrão, que é aquela que esses falantes *tentam* alcançar, e não à norma culta, que não consideramos nesta pesquisa (afinal, não acessamos a real variedade usada por esses falantes).

Nas últimas décadas, alguns linguistas têm proposto gramáticas do português de cunho descritivo, sendo que a maioria delas se concentra na descrição da língua culta, isto é, da língua falada efetivamente usada por falantes cultos (geralmente essas obras são baseadas em pesquisas sociolinguísticas; cf., por exemplo, BAGNO, 2011 ou CASTILHO, 2010). Perini (2005), por outro lado, é uma gramática descritiva da língua padrão. Mas, ao contrário das gramáticas tradicionais que descrevem e prescrevem a norma padrão, Perini não baseia sua descrição em obras de literatura. Segundo a proposta do autor, a norma padrão que deve ser descrita e ensinada nas escolas é aquela encontrada em textos técnicos e jornalísticos, "língua essa que apresenta uma grande uniformidade gramatical, e mesmo estilística, em todo o Brasil" (PERINI, 2005, p. 26).

1.2.2 Gramaticalidade e Agramaticalidade

Sendo a língua composta de diversas normas (ou variedades), cada uma com suas regras e fenômenos variáveis, é de se esperar que a aceitação de determinadas construções gramaticais

por falantes nativos difira na mesma medida em que são diferentes as normas que esses falantes utilizam. Isso quer dizer que determinada construção pode ser **gramatical** para os falantes de uma determinada norma, mas **agramatical** para falantes de outra.

O conceito de *gramaticalidade* é usado extensivamente em diversas escolas da Linguística para identificar construções possíveis e impossíveis em determinada língua ou variedade de língua. Lopes (1995, p. 197) diz que “uma frase é gramatical (ou aceitável) quando está bem formada, fonológica e sintaticamente; em caso contrário, dizemos que ela é agramatical”. Essa definição é típica da escola gerativista em sua "versão padrão" (cf. LOBATO, 1986), quando se acreditava que a sintaxe deveria ser isolada da semântica e estudada por si só (cf. CHOMSKY, 2002 [1957], p. 15). Essa posição dos gerativistas os levava a considerar gramaticais frases como a famosa oração criada por Chomsky, “Ideias verdes incolores dormem furiosamente”. Afinal, afora seja um absurdo semântico, essa oração está gramaticalmente bem formada. Com o advento de versões posteriores ao modelo padrão (cf. MIOTO, SILVA e LOPES, 2005), os gerativistas passaram a incluir fenômenos semânticos na explicação da linguagem humana e frases como a de Chomsky passaram a ser consideradas agramaticais.

Da forma que era usado pela Gramática Gerativa, o conceito de agramaticalidade serve apenas para demonstrar a impossibilidade de certas construções e provar a existência de regras de formação de sentenças na língua em questão. Por exemplo:

(3) *Menino o pedra uma janela a quebrou com

Essa sentença artificialmente criada pode ser usada para demonstrar, dentre outras coisas, que no português não é possível pospor os artigos aos substantivos (“menino o”, “pedra uma”), sob pena de tornar a frase inaceitável para os falantes.

Dentro dos estudos sociolinguísticos, a agramaticalidade é vista sob outro foco: é importante observar que a agramaticalidade das orações é suscetível de ser apreciada em termos de *graus*. É agramatical uma sentença que não é bem formada para uma norma específica – o que não quer dizer que a mesma sentença não possa ser perfeitamente gramatical para outra norma da mesma língua. Tomemos as seguintes sentenças como exemplos:

- (4) a. As meninas gostaram dos livros
 b. Aquela é a menina que eu falei
 c. Esses motorista não sabe dirigir

Tomando como régua as regras prescritas pela norma padrão do português, apenas a sentença (4a) é gramatical. As outras duas são agramaticais: (4b) viola a regra de relativização, que deve levar em conta a regência do verbo da oração encaixada (cf. 2.1.1). Como o correto é “Eu falei da menina”, a relativização padrão seria “Aqueela é a menina de que eu falei”. (4c) viola as regras de concordância nominal e verbal, segundo as quais os itens não nucleares de um sintagma nominal devem concordar com seu núcleo (“Esses motoristas”) e o predicador de uma oração deve concordar com seu argumento externo sujeito (“Esses motoristas não sabemm”).

Já para outras normas do português, as construções (4b) e (4c) podem ser perfeitamente aceitáveis, isto é, *gramaticais*. A relativa em (4b), chamada de “cortadora”, é extensivamente observada na fala de pessoas cultas (cf. BAGNO, 2004), o que significa que ela não é agramatical para os usuários da norma culta: não causa estranheza como (3) nem é estigmatizada como (4c). Orações com as propriedades de (4c), por sua vez, são evitadas por falantes cultos, mas são comuns nas normas populares do PB. Essa variação no julgamento dos falantes certamente é derivada do fato de as várias normas do português serem dispostas no contínuo rural-urbano de que vimos falando por todo este capítulo. Nesta pesquisa, é esse julgamento em relação à (a)gramaticalidade das orações em português que usamos para avaliar os graus de saliência das variáveis e variantes linguísticas.

1.2.3 Comunidade de Fala e Atitudes Linguísticas

O conceito de **comunidade de fala** é um dos mais frequentes dentro da teoria Sociolinguística (cf. MONTEIRO, 2008) e, talvez por isso, seja um dos mais complexos e problemáticos: não é fácil estabelecer os limites geográficos ou sociais de uma comunidade. Ao descrever uma variedade linguística, o sociolinguista colhe dados a partir da fala (a *parole* saussuriana) de seus informantes para, em seguida, proceder à análise dos fenômenos da língua (a *langue*) utilizada por aquele grupo. Sendo assim, sempre que inicia uma nova pesquisa sociolinguística, o investigador se depara com um problema: onde ou como se delimitam as fronteiras de uma variedade linguística?

Uma boa maneira de considerar o que constitui uma comunidade de fala é definindo o que uma comunidade de fala não é (MULLANY, 2007, p. 84). Uma comunidade de fala não existe simplesmente porque indivíduos falam uma mesma língua ou dialeto. Ainda que essa ideia tenha vindo à tona nas primeiras definições do termo, essa é uma visão facilmente refutável. Como aponta Wardhaugh (2010, p. 119-120), ainda que o inglês seja falado em vários lugares ao redor do mundo, os falantes de inglês desses lugares não constituem uma

comunidade de fala, uma vez que eles utilizam a língua de várias formas diferentes e são isolados uns dos outros. O mesmo pode ser dito para o português: ainda que a língua falada nos dois lados do Atlântico possa ser considerada a mesma (há linguistas que discordam dessa afirmação; cf. BAGNO, 2007; GALVES, 2001), não se pode entender os falantes portugueses, africanos e brasileiros como pertencentes a uma mesma comunidade de fala, devido às diferenças culturais e linguísticas existentes e à descontinuidade geográfica das áreas lusófonas. Para se chegar a uma definição sociolinguística de uma comunidade de fala, portanto, outras categorias além da língua devem ser consideradas.

Uma das primeiras definições de uma comunidade de fala foi dada por Labov (2008 [1972]), baseando-se em suas pesquisas feitas na cidade de Nova Iorque. Sua classificação tira o foco dos problemas associados a uma definição de cunho puramente linguístico:

A comunidade de fala não é definida por nenhuma concordância marcada pelo uso de elementos linguísticos, mas sim pela participação num conjunto de normas partilhadas; estas normas podem ser observadas em tipos de comportamento avaliativo explícito e pela uniformidade de padrões abstratos de variação que são invariantes no tocante a níveis particulares de uso (LABOV, 2008 [1972], p. 150, grifos nossos).

Labov salienta que, para que uma comunidade de fala exista, os falantes não precisam concordar que falam a mesma língua ou que falam da mesma forma, mas eles devem concordar a respeito das *normas* de avaliação. Ele descobriu que, enquanto certas variáveis fonológicas eram pronunciadas diferentemente por membros de classes sociais distintas, ao examinar diferentes estilos de conversação os falantes de todas as classes convergiam para um mesmo padrão de uso, utilizando mais variantes não padrão em situações não formais, e vice-versa. Sendo assim, ainda que os falantes utilizassem a língua de diferentes formas, havia evidências de *atitudes* compartilhadas, com falantes de todas as diferentes classes sociais avaliando as formas padrões da mesma maneira, utilizando as formas mais prestigiadas com maior frequência nas situações mais formais e, portanto, mais monitoradas.

O problema da **atitude linguística** foi abordado já no trabalho pioneiro de Weireich, Labov e Herzog (2006). Como comentamos em 1.1.2, a teoria laboviana coloca o problema da avaliação dentre os “princípios empíricos para a teoria da mudança linguística”. Segundo os autores,

A teoria da mudança linguística deve estabelecer empiricamente os correlatos subjetivos dos diversos estratos e variáveis numa estrutura heterogênea. Esses correlatos subjetivos das avaliações não podem ser deduzidos a partir do lugar das variáveis dentro da estrutura linguística. Além disso, o nível de consciência social é uma propriedade importante da mudança linguística que tem de ser determinada diretamente. Correlatos subjetivos da mudança são por natureza mais categóricos que

os padrões cambiantes do comportamento: a investigação destes correlatos aprofunda nosso entendimento dos modos como a categorização discreta é imposta ao processo contínuo da mudança (WEINREICH, LABOV e HERZOG, 2006, p. 124).

Os correlatos subjetivos devem ser preferencialmente acessados, portanto, de forma *direta*. Hora (2012) comenta que, dentre as três abordagens típicas usadas em estudos de atitude – análise de conteúdo, abordagem direta e abordagem indireta –, a abordagem direta é a mais comum. Labov, por exemplo, utilizou “testes subjetivos de reação” como técnica para aferir a atitude linguística dos falantes nova-iorquinos em relação ao uso dos róticos. Labov visitou várias lojas de um *shopping* em Nova Iorque, solicitando aos atendentes a informação de onde ele poderia encontrar lâmpadas, pergunta cuja resposta ele já sabia: no quarto andar, em inglês *fourth floor*. Labov queria com isso observar a maneira como cada atendente pronunciava o fonema /r/ pós-vocálico, e detectou duas maneiras distintas de se pronunciar essa unidade: com a presença do segmento fônico [r], variante considerada padrão pelos falantes americanos, e com a ausência [Ø] desse segmento nos mesmos contextos, variante comum na fala de Nova Iorque, mas estigmatizada socialmente. Os resultados mostraram que, quanto maior fosse o nível socioeconômico da loja visitada (e de seus frequentadores), mais o [r] pós-vocálico era pronunciado¹¹.

Atualmente, as pesquisas sobre atitudes linguísticas utilizam técnicas mais sofisticadas para atestar as avaliações subjetivas dos usuários da língua em relação às variáveis linguísticas. Garrett (2007) comenta que, para minizar os efeitos negativos dessa abordagem direta, os pesquisadores têm utilizado a técnica do *matched-guise*: gravam-se várias falas de um ator que consegue imitar diversas vozes e reproduzir diferentes sotaques ou variedades de uma língua e apresentam-se essas gravações aos informantes. Os informantes não sabem que se trata sempre da mesma pessoa na gravação, e são levados a avaliar as diversas formas de falar, indicando o que pensam sobre a classe social, aparência, personalidade, etc. dos “diferentes” falantes. Os resultados geralmente indicam que os informantes prestigiam certas formas de falar a partir de como imaginam ser a pessoa que as utilizam, atribuindo aos usos padrão adjetivos como “inteligente”, “competente” ou “atraente”.

Outro tipo de teste bastante utilizado para avaliar as atitudes linguísticas é o teste de percepção e produção (cf. TARALLO, 1986, p. 55; COELHO, GÖRSKI, *et al.*, 2010, p. 120).

¹¹ Interessantemente, essa é a situação oposta da Inglaterra. No inglês britânico, a ausência do [r] pós-vocálico é prestigiada e considerada padrão (faz parte do “inglês da rainha”), enquanto que sua presença é estigmatizada. Isso corrobora o postulado da Sociolinguística que afirma que não há nada intrinsecamente correto ou errado em uma variante linguística (princípio do relativismo cultural; cf. 1.1) – o que faz uma determinada variante ser considerada “correta” em detrimento de outra(s) é um acidente histórico (cf. LYONS, 1987, p. 24-25).

Em testes de percepção, é apresentada aos informantes uma bateira de construções, sem uma ordem pré-determinada, para que eles digam quais são as mais aceitáveis e as menos aceitáveis. Para cada variável e variante, há várias opções de oração. A partir dos dados colhidos, é possível identificar como os informantes percebem subjetivamente cada variante, assim como observar os padrões de uso da comunidade estudada: caso os falantes não percebam certos tipos de variação como não padrão, isso pode significar que uma mudança linguística está se processando na direção dessa variante em detrimento das outras ("problema da implementação"; cf. 1.1.2). Os testes de produção, por sua vez, trazem mecanismos que levam os informantes a construir sentenças com a variável estudada. Na tentativa de produção, eles utilizarão uma ou outra variável, e é essa escolha que o pesquisador considera para quantificar os usos e relacioná-los aos aspectos sociais da comunidade de fala estudada.

Para fechar essa breve discussão sobre o conceito de atitude linguística, retomamos o item anterior sobre norma padrão e citamos Hora (2012), que diz que

O fato de pensarmos na existência de uma língua padrão afeta o modo como os falantes pensam sua própria língua e a língua em geral. Podemos dizer que os falantes dessas línguas, diferentes dos falantes de algumas menos conhecidas, vivem em culturas de língua padrão. Em tais culturas, as atitudes linguísticas são dominadas por posições ideológicas de poder que são amplamente baseadas na existência suposta dessa forma padrão, e essas, tomadas juntas, constituem a "ideologia da língua padrão". Os falantes, normalmente, não têm consciência de que eles estejam condicionados por essas posições ideológicas: eles, normalmente, acreditam que suas atitudes em relação à língua sejam de senso comum e assumem que, virtualmente, todos concordam com eles (HORA, 2012, p. 372-373).

É a atitude linguística dos falantes cultos em relação aos usos variáveis do português brasileiro que buscamos atestar por meio desta pesquisa, e é a partir dela que poderemos identificar quais variantes e variáveis são mais salientes pra esses mesmos falantes. Na próxima seção, traremos uma discussão sobre o conceito de saliência linguística e como ele se relaciona com os outros conceitos já abordados neste capítulo.

1.3 Saliência Linguística

Neste trabalho, estudamos a percepção e a atitude linguística dos falantes cultos buscando identificar quais variáveis e variantes da variedade brasileira do português são mais salientes para esses falantes, além de verificar quais contextos linguísticos (isto é, quais fatores condicionadores) mais favorecem essa saliência. "Saliência" não é um termo comum dentro da Sociolinguística; por isso, nesta seção, veremos com quais sentidos esse termo já foi usado dentro da pesquisa linguística. Ao final, traremos uma breve explicação de como esse conceito é utilizado nesta pesquisa.

Um dos objetivos centrais da Sociolinguística é identificar como as estruturas sociais podem se refletir no uso linguístico (pois, como vimos em 1.1, a língua pode ser sistematicamente relacionada às estruturas sociais). Parte desse conhecimento pode ser obtido observando-se como os usuários da língua reagem às diversas variantes linguísticas, e como se comportam linguisticamente a partir dessas reações ("atitude linguística"; cf. 1.2.3). Nesse contexto, é possível perceber que, enquanto várias das características das línguas podem superficializar a estratificação social, apenas algumas parecem ser notadas pelos falantes.

Dessa forma, “saliência”, em Sociolinguística, se refere a uma propriedade ou série de propriedades que leva uma variável linguística a ser mais proeminente, isto é, mais *evidente* aos usuários da língua. No entanto, a teoria linguística ainda não se lançou no empreendimento de busca dessas propriedades: ainda não há concordância geral sobre a definição desse termo, e ele tem sido usado como um “termo guarda-chuva” para designar uma quantidade enorme de fenômenos que não necessariamente guardam relações entre si (cf. RÁCZ, 2013). A maioria dos trabalhos que lida com esse conceito tenta identificar *quais seriam* essas variáveis mais salientes (como dissemos na Introdução, ao citar a fala de Bortoni-Ricardo (2014) que motivou esta pesquisa, nosso trabalho segue essa tendência), o que de certa forma é o primeiro passo de um empreendimento de pesquisa que busca identificar as *razões* da saliência.

Mesmo que o termo “saliência” não seja de uso frequente e não haja uma definição operacionalizável e única para ele na literatura linguística, é possível estabelecer as raízes desse conceito nos primeiros trabalhos em Sociolinguística. Labov (2008) observa que há dois possíveis comportamentos tipicamente associados a uma variável linguística: algumas variáveis podem ser o objeto de vários processos sociolinguísticos, tais como a hipercorreção ou a estigmatização, ao passo que outras parecem ser invisíveis a esses processos. Isso acontece porque há julgamentos sociais conscientes e inconscientes sobre a língua: com base no nível de consciência que os falantes têm sobre determinada variável, Labov distingue dois¹² tipos de elementos: *indicadores* e *marcadores*.

Indicadores são elementos linguísticos sobre os quais há pouca força de avaliação. Eles podem variar junto com a estratificação social, mas não têm interpretação social – em outras palavras, indicadores são “invisíveis” ao julgamento dos falantes. Se há uma norma culta e uma

¹² Há uma terceira categoria proposta pelo autor, a dos estereótipos, que são traços socialmente marcados de forma consciente, mas não necessariamente prestigiados ou estigmatizados. Esse tipo de elemento é bastante explorado em programas de humor, ao compor personagens característicos de determinada região, como por exemplo os mineiros com o fenômeno da elisão (“popopó” por “pode pôr pó”) ou os caipiras com seu /r/ retroflexo.

norma popular em uma determinada língua, uma variável indicadora vai ser diferente nas duas; no entanto, os falantes da norma popular não tentarão usar a variante culta ao falarem a norma culta, e tampouco isso será notado pelos falantes cultos. Por exemplo: no PB, há um fenômeno conhecido como “monotongação dos ditongos /ej/ e /ow/”, que atinge palavras como *peixe* e *couro*, pronunciadas como *pexe* e *coro*. Esse é um traço socialmente estratificado, mas não é sujeito à variação estilística e não tem força avaliativa: poucos sequer percebem que esse fenômeno existe.

Marcadores por sua vez correlacionam-se fortemente com as estratificações sociais e estilísticas e podem ser facilmente diagnosticadas em testes de percepção (cf. COELHO, GÖRSKI, *et al.*, 2010, p. 33). Se um marcador está associado a uma norma popular, os falantes mais escolarizados tentarão evitá-lo em situações mais formais e vão identificá-lo como “feio” ou “errado”. Um exemplo de marcador no PB é a variável Concordância Verbal: a variante popular, isto é, a não conformação do verbo com o o sujeito da oração (“Eles vai pescar”) é evitada pelos falantes cultos e é facilmente perceptível, sendo também uma construção estigmatizada¹³ na consciência sociolinguística dos usuários da língua. Devemos ter em mente que falar de “rejeição” de uma variável não implica que esse seja um ato consciente. A percepção dos falantes sobre as variáveis linguísticas é um campo bastante complexo: em geral, os falantes são capazes de identificar o uso de um marcador, mas são incapazes de apontar o marcador em si (cf. RÁCZ, 2012, p. 59).

A saliência pode ser considerada um requisito para a existência de um marcador, ou pode ser vista como apenas mais uma de suas propriedades. Também não se sabe exatamente se variantes mais salientes são eleitas como marcadores ou se as variantes se tornam salientes *depois* de serem eleitas como marcadores. Veremos, a seguir, dois trabalhos que tratam de saliência a partir da dicotomia laboviana entre indicadores e marcadores.

1.3.1 Duarte e Tarallo (1988): Saliência e Mudança Linguística

Os trabalhos de Duarte (1989) e Tarallo (1983) trouxeram uma inovação aos estudos sociolinguísticos brasileiros, que estavam ainda em fase inicial na época: além de recolher amostras da fala de seus informantes, esses pesquisadores recolheram também a *opinião* dos informantes sobre as variáveis linguísticas. Como vimos em 1.2.3, esse tipo de pesquisa se relaciona ao estudo de atitudes linguísticas dos falantes, e estão diretamente ligados aos estudos

¹³ O tão discutido preconceito linguístico deriva, segundo essa teoria, do uso inadvertido de marcadores por falantes não escolarizados. Sobre a teoria do preconceito linguístico, cf. Bagno (2007).

de saliência, como bem demonstra o artigo conjuntamente escrito pelos dois pesquisadores citados, Duarte e Tarallo (1988), que resume e traz novas reflexões sobre seus trabalhos anteriores.

Duarte e Tarallo (1988) utilizam os conceitos de indicadores e marcadores de Labov para explicar a mudança linguística:

O encaixamento sociolinguístico da variável estudada, ao lado da avaliação de suas variantes pelos informantes da comunidade, permite, por sua vez, escalonar as variáveis em: indicadores, estereótipos e marcadores. Destes, os últimos, ao cruzarem “preferências” sócio-econômicas e, por extensão, sócio-estilísticas, predizem a possibilidade (ou não) de alguma das variantes estar em processo de transição e/ou implementação dentro da comunidade, ou seja, em uma situação de mudança (DUARTE e TARALLO, 1988, p. 44).

Uma questão que os autores colocam no início do trabalho diz respeito à saliência das variáveis internas de uma variante: estariam aqueles fatores que são menos salientes enveredando a fala por caminhos totalmente despercebidos? Isto é, a inobservância de variantes não salientes não “facilitaria” a entrada e a consolidação dessas variantes no sistema da língua? Os autores argumentam que *sim*: com base em testes de reação subjetiva sobre o uso da variável anáfora zero (uma das variáveis que estudaremos neste trabalho; cf. 2.1), demonstram que uma mesma variável pode “variar” seu grau de saliência, de forma que suas variantes estão controladas pela ideologia da gramática somente em seu lado saliente, enquanto que a não saliência de determinadas variantes garante sua implementação no sistema.

O trabalho de Duarte e Tarallo (1988) traz uma importante constatação que usaremos em nosso conceito de saliência (cf. 1.3.3). Segundo os autores,

Uma determinada variante estigmatizada pela comunidade pode não ganhar em prestígio (i.e. avaliação positiva), mas, sem dúvida alguma, a mesma variante pode ser reduzida em seu estigma (i.e. avaliação negativa) em que pesem diferentes configurações internas (DUARTE E TARALLO, 1988, p. 55).

No caso específico da variável analisada por eles, o objeto direto anafórico, foi constatado que a variante em que o objeto direto é constituído pelo pronome lexical é condenada pela comunidade; no entanto, essa mesma comunidade tolera essa variante em contextos sintáticos mais complexos. Ou seja: os fatores linguísticos (variáveis independentes) podem “camuflar” o uso de uma determinada variante, ao passo que outros contextos podem “evidenciar” esse uso. Sendo assim, o contexto pode aumentar ou diminuir o nível de saliência das variantes.

1.3.2 Rácz (2013): Buscando um Modelo de Operacionalização da Saliência

O trabalho de Rácz (2013) traz uma discussão aprofundada e inédita sobre o conceito de saliência dentro de uma teoria variacionista da Sociolinguística. Seu objetivo é desenvolver esse conceito ao propor uma definição empiricamente testável do termo, para que este deixe de ser um termo “guarda-chuva” e passe a ser um rigoroso construto teórico. A teoria proposta pelo autor serve em especial para analisar variáveis fonológicas, mas a discussão que se realiza nos primeiros capítulos traz *insights* importantes para o nosso trabalho.

Inicialmente, Rácz se faz duas perguntas: que tipos de variação linguística são salientes para os usuários e como essas diferenças são utilizadas pela comunidade linguística? Isso o leva a definir dois tipos de saliência: a *cognitiva* e a *social*. A saliência cognitiva diz respeito às propriedades objetivas da variação linguística que a fazem perceptível para o falante; a saliência social está ligada ao conjunto de atitudes, estereótipos culturais e valores sociais associados às variantes. A saliência cognitiva emerge dos padrões de variação linguística e serve de *input* à saliência social, que podemos observar por meio das atitudes demonstradas pela comunidade linguística.

O autor começa por revisar os usos do termo “saliência” na literatura linguística. A saliência é uma propriedade da *diferença* linguística, de modo que duas formas diferentes de dizer a mesma coisa (uma definição de *variável* linguística) têm *significados sociais* distintos, isto é, os falantes costumam associar certas formas linguísticas a categorias sociais, tais como origem geográfica ou classe social. As formas que mais chamam a atenção para esses significados sociais são mais *salientes* para os falantes. Dois dialetos podem se diferenciar em *n* variáveis, mas apenas uma parte dessas variáveis é reconhecida pelos falantes (não necessariamente de forma consciente). Um falante reconhece uma variável porque ele reconhece uma de suas variantes que ele nunca usaria, ou que nunca usaria em determinado contexto. O fato de uma variável ser facilmente reconhecida é o que os estudiosos de Sociolinguística geralmente chamam de “saliência”. Um exemplo desse fato em português é o já citado problema das estratégias de relativização (cf. 2.1 e 3.4): essa variável tem três variantes, sendo uma padrão e duas não padrão. O que as pesquisas sociolinguísticas nos dizem é que uma dessas variantes não padrão (a relativa copiadora) é sistematicamente evitada pelos falantes mais escolarizados, ao passo que a outra (a relativa cortadora), apesar de também ser “errada” para a norma padrão, é extensivamente usada por não ser percebida como uma impropriedade gramatical. Dessa forma, a relativa copiadora é *mais saliente* que a cortadora.

Essa ideia de saliência é a mesma utilizada na dicotomia entre indicadores e marcadores de Labov, que já vimos. RácZ comenta que essa distinção, no entanto, é uma simplificação: evoluções posteriores da teoria linguística do uso, beneficiando-se das ideias da psicologia social, consideraram que indicadores e marcadores não são “rótulos” que servem para identificar a proveniência social de um indivíduo, mas, antes, ferramentas de que os falantes dispõem para se posicionar e demonstrar sua atitude (cf. 1.2.3): “um indicador é uma variável que o falante descarta ao se posicionar socialmente; um marcador é uma variável que ele usa para expressar seu prestígio social ou atitude” (RÁCZ, 2013, p. 5, tradução nossa).

No entanto, reduzir o conjunto de variáveis a indicadores e marcadores é “perigoso”, ao implicar uma completa falta de *gradiência*. A percepção linguística teria muitos níveis, e poucos são categóricos. O trabalho de RácZ (2013) visa expandir o conceito binário de Labov, tratando os marcadores como uma variável que afeta as atitudes e comportamentos dos falantes, invocando atitudes negativas ou positivas e motivando mudanças estilísticas. A saliência é, assim, interpretada como uma propriedade que permite a uma variável linguística ser um marcador. O autor traz vários capítulos com estudos de caso, nos quais ele utiliza testes de atitude para determinar quais marcadores são utilizados em várias comunidades linguísticas com propósitos de identificação.

Esse modo de enxergar a saliência, ligado às variáveis linguísticas, coloca uma série de questões: a saliência de uma variável está ligada às suas propriedades estruturais e extralinguísticas? A saliência tem um papel na propagação da mudança linguística? E, mais importante: como podemos dizer com certeza que uma variante é saliente? Para responder essas perguntas, RácZ (2013) traz contribuições de teorias não linguísticas, como o conceito de saliência dentro das ciências cognitivas da percepção visual, e traz a hipótese de que a saliência em Sociolinguística deriva da “baixa probabilidade de ocorrência”, no sentido de que a saliência deriva da surpresa causada pela variante que está sendo usada. Dessa forma, uma variável alcança saliência cognitiva quando suas variantes têm um alto valor de surpresa quando comparadas entre si. Essas variantes com alta saliência cognitiva podem carregar interpretações sociais, e assim se tornam socialmente salientes para a comunidade de fala.

1.3.3 Níveis de Saliência de Variantes e Variáveis

Labov (2008), ao propor os conceitos de indicadores e marcadores, postula uma *dicotomia* para definir a saliência das variáveis linguísticas. Neste trabalho, usaremos as mesmas ideias propostas pelo autor, mas distinguiremos graus ou **níveis de saliência**,

expandindo a visão dicotômica de Labov e adotando a posição gradiente defendida por Rącz (2013). Propomos aqui que não só as variáveis, mas também as variantes podem ser mais salientes que outras. Propomos ainda que há contextos linguísticos (isto é, variantes independentes estruturais, ou fatores linguísticos; cf. 1.1.1) que podem salientar uma variante e “esconder” outra, como demonstraram Duarte e Tarallo (1988). Explicaremos essas categorias nos próximos parágrafos.

O PB apresenta em suas variedades uma enorme quantidade de regras variáveis, mas algumas delas se revestem de maior importância sociolinguística. Isso acontece porque algumas dessas variáveis têm *maior saliência* no uso linguístico, o que leva a um maior nível de *prestígio* ou *estigmatização* de suas variantes, segundo a percepção e o julgamento dos falantes. Por exemplo: é comum no PB o fenômeno conhecido como "alçamento das vogais pretônicas". Segundo esse fenômeno fonológico, a pronúncia de vogais que são núcleo de sílabas que antecedem a sílaba tônica é “alçada” uma categoria acima: se a vogal é média, como /e/, ela passa a ser pronunciada alta, como /i/. É o que acontece em palavras como *menino*, *perigo*, *tomate* ou *mosquito*, que são geralmente pronunciadas como *minino*, *p̄rigo*, *tumate* e *musquito*. Outro exemplo de fenômeno fonológico bastante comum no português é o rotacismo: sob sua influência, palavras como *placa* ou *flecha* são pronunciadas como *praca* e *frecha*. Esse é um fenômeno antigo e sistemático no português, que vem influenciando a pronúncia de várias palavras: o que, em latim, era pronunciado com um encontro consonantal com L (*ecclesia-*, do grego *ekklēsia*, *plaga-*, *sclavu-*, *fluxu-*) passou a ser pronunciado com um R em português (*igreja*, *praga*, *escravo*, *frouxo*) e continua com L em espanhol (*iglesia*, *playa*, *sclavo*, *flojo*), conforme Bagno (2006, p. 43)¹⁴.

Notemos que, apesar de ambos os fenômenos serem exatamente iguais, isto é, de ambos se relacionarem a *um som que é sistematicamente substituído por outro som em determinados contextos*, o fenômeno do rotacismo é muito mais estigmatizado que o fenômeno do alçamento das vogais pretônicas. Uma vez que o rotacismo é um fenômeno típico de variantes que se aproximam da extremidade rural do contínuo rural-urbano (cf. 1.1.3), os falantes urbanizados e escolarizados tendem a rotulá-lo negativamente. Já o fenômeno do alçamento das vogais é bastante difundido tanto nas variedades rurais quanto nas urbanas, sendo comum na fala até mesmo das pessoas mais escolarizadas. Isso faz com que esse fenômeno seja, de certa forma, “invisível” ou “neutro” aos ouvidos dos falantes cultos. Nos termos de Bortoni-Ricardo (2004),

¹⁴ É preciso não confundir o fenômeno linguístico do rotacismo com a patologia rotacismo, que deve ser tratada por fonoaudiólogos.

trata-se de um traço gradual no contínuo rural-urbano. Nesse caso, podemos dizer que o rotacismo é uma **variável mais saliente** que o alçamento das vogais pretônicas, pois é mais facilmente percebido e mais amplamente estigmatizado.

Esses dois fenômenos de que falamos são variáveis binárias, para as quais há apenas duas possibilidades de superficialização (duas variantes). Há alguns tipos de variáveis que apresentam três ou mais variantes. Um bom exemplo no PB é a pronúncia do /r/ pós-vocálico, onde são percebidas pelo menos cinco possibilidades de pronúncia: *tepe*, *retroflexo*, *vibrante alveolar*, *fricativa velar* e *apagamento* (cf. SILVA, 2012). Nenhuma dessas variantes, entretanto, apresenta maior saliência – talvez a variante retroflexa “salte aos ouvidos” de falantes de certas partes do país, por ser uma pronúncia típica da zona “caipira” do Brasil (interior de São Paulo e Centro-Oeste). É possível também que os falantes reconheçam que o apagamento do /r/ final, que gera pronúncias como *comê* ou *falá*, seja imprópria aos padrões cultos da língua; no entanto, essa variante é altamente utilizada por falantes cultos e também passa despercebida (é um exemplo de *indicador*).

Outra variável fonológica com mais de duas possibilidades de superficialização é a realização da vogal pós-tônica não final em proparoxítonas. Uma palavra como *fósforo* teria três possibilidades de pronúncia: com a vogal média, que caracteriza a pronúncia mais próxima do ideal da norma culta: *fósforo*; com o alçamento da vogal: *fósfuro*; e com a síncope (apagamento) da vogal e conseqüente rearranjo silábico da palavra, que passa a ser paroxítona: *fósfro*. A primeira pronúncia é prestigiada e a segunda, como já vimos, é neutra. A última é estigmatizada, talvez por ser típica da fala de pessoas pouco escolarizadas (cf. AMARAL, 2002). Podemos dizer, assim, que a síncope da vogal é a **variante mais saliente** dentre as três possibilidades.

Um outro fenômeno fonológico menos frequente, mas também muito esclarecedor, é a palatização dos fonemas /t/ e /d/ em certos contextos. Como já salientamos em 1.1.1, esse fenômeno é bastante comum no contexto “anteposto a vogal alta anterior”, ou seja, em palavras como *tia* e *dia*, que são pronunciadas [tʃiɐ] e [dʒiɐ]. No entanto, em algumas regiões do Norte e Nordeste brasileiros, a palatização desses fonemas ocorre também no contexto “posposto a ditongo”, em palavras como *muito* e *deitado*, que são pronunciadas como [mũjtʃu], “muitcho”, e [dejtʃadu], “deitchado”. O interessante desse exemplo é que tanto a palatização no contexto “anteposto a vogal alta anterior” quanto no contexto “posposto a ditongo” são frequentes nos dialetos das regiões citadas, mas apenas o contexto “posposto a ditongo” é estigmatizado (cf.

DIAS, 2009)¹⁵. Sendo assim, não é a variável ou a variante que se mostra mais saliente: temos o mesmo fenômeno variável e a mesma variante em jogo, mas um dos contextos “camufla” o que seria uma pronúncia diferente da padrão, enquanto que o outro salienta essa pronúncia “imprópria”. Nesse caso, dizemos que **o contexto salienta a variante**. Em geral, uma variante salientada é mais facilmente percebida como imprópria, e por isso ela sofre mais estigma.

Chegamos, assim, às categorias que pautam nossa pesquisa: dependendo do julgamento dos falantes (no nosso caso, os cultos), uma variável pode ser mais saliente que outra; dentre as variantes de uma variável, é também possível que uma seja mais saliente que outra; e dentre os contextos provocadores das variantes, um pode salientar mais a mesma variante, deixando-a mais visível ou invisível aos julgamentos dos falantes. Vimos exemplos de fenômenos fonológicos que comprovam a existência dessas categorias. Nesta pesquisa, utilizamos fenômenos sintáticos, conforme descrevemos no próximo capítulo.

¹⁵ A pronúncia palatizada antes de vogal alta anterior, inclusive, é abonada pela gramática normativa de Bechara (2006, p. 79), “evitando-se, entretanto, o exagero”.

2 PROCEDIMENTOS DE PESQUISA

O objetivo deste capítulo é apresentar a metodologia utilizada para colher os dados que analisaremos mais à frente. Os trabalhos em Sociolinguística sempre se referem a uma ou mais variáveis da língua em questão, cujos usos são registrados por meio de técnicas específicas a partir da colaboração de informantes falantes nativos e analisadas por métodos quantitativos que buscam aferir as estatísticas de uso dos fenômenos em questão. Aqui narraremos como foram escolhidas as *variáveis* deste trabalho (cf. 2.1), quais os *procedimentos* metodológicos e qual foi a *técnica* utilizada para registrar os dados que nos interessam (cf. 2.2) e como essa técnica foi *aplicada* aos informantes (cf. 2.3).

2.1 A Escolha das Regras Variáveis

Como dissemos na Introdução a este trabalho, o objetivo desta pesquisa é identificar, considerando as regras variáveis escolhidas para análise, quais delas são mais salientes para os falantes cultos, quais de suas variantes são mais estigmatizadas e quais contextos mais afetam a percepção dos falantes cultos para os usos estigmatizados. Como não seria possível fazer esses testes para todas as regras variáveis do PB, escolhemos para análise três das variáveis mais estudadas na história das pesquisas sociolinguísticas no país: a "concordância verbal", a "retomada anafórica por objeto direto" e as "estratégias de relativização".

A **concordância verbal** foi um dos primeiros fenômenos variáveis do PB abordados em pesquisas sociolinguísticas (cf. LEMLE e NARO, 1977) e tem sido desde então objeto de diversas pesquisas variacionistas. Esse destaque a esse fenômeno se dá possivelmente porque a falta de concordância entre verbo e sujeito (uma das variantes possíveis para essa regra variável) é altamente estigmatizada entre os falantes do português, e é uma das principais características que distinguem o português popular do português culto ou padrão, e o português brasileiro do português europeu. É também um assunto bastante valorizado nas aulas de língua portuguesa, exatamente por ter uma caracterização sociolinguística tão saliente (cf. VIEIRA, 2007). A despeito de ter sido já amplamente estudada e debatida, a concordância ainda suscita grande interesse dos pesquisadores, que refazem os passos de trabalhos anteriores com dados inéditos (cf., por exemplo, trabalhos recentes como o de MONGUILHOTT, 2001; OLIVEIRA, 2005; ou de MONTE, 2007).

Há duas possibilidades de aplicação da regra de concordância verbal no PB, como já comentamos na Introdução: a subjugação morfológica do verbo, que copia as informações de

número e *pessoa* do sujeito e se flexiona consoante essas informações; e a não subjugação morfológica do verbo, que nesse caso se mantém na forma *singular*, mas ainda carregando traços da pessoa gramatical do sujeito. Trata-se, portanto, de uma regra de caráter binário. Veja:

- (5) a. Os meninos comeram o bolo
 b. Os menino comeu o bolo¹⁶

A primeira forma é padrão, segundo as gramáticas tradicionais normativas do português: o verbo se encontra flexionado na 3ª pessoa do plural, pois seu sujeito “os meninos” equivale gramaticalmente ao pronome reto “eles”, representante da 3ª pessoa do plural. A segunda forma é não padrão: o verbo concorda com a pessoa gramatical (3ª pessoa: *ele*), mas não copia a informação de “plural” do sujeito, mantendo-se no singular. Essa forma é estigmatizada e tratada como agramatical pelas gramáticas normativas e pelos falantes cultos do português.

A **retomada anafórica por objeto direto de 3ª pessoa** diz respeito às possibilidades de superficialização de um objeto direto *correferente*¹⁷ com um sintagma nominal anteriormente mencionado (nos casos que nos interessam, esse sintagma nominal é de terceira pessoa). Trata-se de uma regra variável ternária, pois tem, basicamente, três possibilidades de realização. Veja:

- (6) a. A dissertação está pronta, mas ainda não a revisei
 b. A dissertação está pronta, mas ainda não revisei \emptyset
 c. A dissertação está pronta, mas ainda não revisei ela

Em (6) podemos ver as três estratégias possíveis de realização anafórica por meio de objeto direto do sintagma nominal “a dissertação”: em (6a) temos a versão tradicional, com o pronome acusativo (que é o “caso” do objeto direto; cf. MIOTO, SILVA e LOPES, 2005), chamado pela gramática tradicional de *pronome oblíquo átono* – essa é a forma conservadora e prestigiada; em (6b) há um exemplo da retomada por categoria vazia, isto é, por um “pronome” sem realização fonética; e em (6c) temos a retomada por pronome sujeito (chamado de reto pela tradição gramatical), eminentemente condenada pelas gramáticas normativas. Essas três possibilidades de retomada por objeto direto anafórico foram detectadas e medidas em

¹⁶ Neste trabalho, focaremos apenas a regra variável da concordância verbal na terceira pessoa do plural, não envolvendo na análise a variação na concordância de outras pessoas gramaticais.

¹⁷ Duas expressões linguísticas são *correferentes* quando têm o mesmo referente semântico no mundo. Por exemplo: na sentença “eu gosto do João, mas ele é grosso”, os SNs [o João] e [ele] se referem ao mesmo ser no mundo, e são, portanto, correferentes. A retomada de [o João] pelo pronome [ele] é chamada de *anáfora* pela teoria semântica.

trabalhos clássicos como o de Duarte (1989) e de Cyrino (1994), e continuam sendo objeto de pesquisas sociolinguísticas (cf., por exemplo, CORRÊA, 2014; PEREIRA e COELHO, 2013).

As **estratégias de relativização**, por sua vez, dizem respeito às três possibilidades de construção das chamadas orações subordinadas adjetivas, ou *relativas*, como são mais conhecidas na literatura linguística. Essas três possibilidades podem ser identificadas quando a oração relativa está ligada a um verbo cuja regência exige a existência de uma preposição, como é o caso de “gostar (de)” ou “conversar (com)”. Veja:

- (7) a. Aquela é a moça com quem conversei ontem
 b. Aquela é a moça que eu conversei Ø ontem
 c. Aquela é a moça que eu conversei com ela ontem

A estratégia (7a) é a única aceita pela gramática normativa e é a única opção gramatical dentro da língua padrão. A estratégia (7b) é chamada de relativa cortadora, pois apresenta uma categoria vazia no espaço onde deveria haver uma preposição. A estratégia (7c) é chamada de relativa copiadora, pois há a repetição pleonástica do objeto indireto¹⁸ da oração subordinada, que nesse caso se vê representado por duas expressões linguísticas: *que*, pronome relativo com valor funcional de objeto indireto, e *ela*, pronome pessoal correferente ao pronome relativo. Essa é a variante mais estigmatizada.

Veremos, antes de explicar a elaboração dos testes de percepção, o que dizem as gramáticas tradicionais normativas sobre os fenômenos variáveis aqui tratados, e como essas prescrições normativas são seguidas ou não pelos falantes do português a partir da revisão de trabalhos variacionistas.

2.1.1 O Que Dizem as Gramáticas Normativas

Como parâmetro para identificar os usos considerados padrão, visitamos as prescrições de gramáticas normativas tradicionais do português. Dessa forma, para cada uma das regras variáveis que temos em foco, revisamos o que três das principais gramáticas normativas do mercado brasileiro têm a dizer: Bechara (2006), Cunha e Cintra (2001) e Rocha Lima (2011). Concordamos com Bagno (2004) quando o autor diz que essas gramáticas representam obras

¹⁸ Embora saibamos que é impróprio chamar de objeto indireto um complemento regido de preposição “com” sem correspondência com o pronome “lhe”, sendo mais apropriado rotulá-lo de *complemento relativo* (cf. DUARTE, 2007, p. 188; ROCHA LIMA, 2011, p. 311), manteremos essa classificação para fins de simplificação, tendo em vista que muitas gramáticas tradicionais não fazem a diferenciação entre essas duas funções sintáticas.

escritas por gramáticos competentes, com amplo conhecimento da história e da constituição da norma padrão do português, e que são boas fontes das regras da norma padrão.

Usamos gramáticas tradicionais como parâmetro de comparação porque, como comentamos em 1.2.1, a ideologia da padronização linguística está presente no imaginário coletivo da comunidade de fala de uma língua de civilização, como é o português. Milroy (2007) identifica quatro características dessa ideologia: a noção de correção, a importância da autoridade, a relevância do prestígio e a ideia de legitimidade. A primeira consequência da padronização é a ideia de **correção** linguística: desenvolveu-se uma norma padrão, considerada inerentemente correta e canônica na língua. Todos os falantes subscrevem a essa ideia de correção, e faz parte do senso comum a crença de que haja fontes acreditadas de divulgação e normalização da língua padrão. A manutenção da norma padrão depende da obediência a essas fontes, que são as **autoridades** sobre a língua: dicionários, gramáticas, manuais de uso, etc. As formas padrão são as mais **prestigiadas**, enquanto que as não padrão são **estigmatizadas**, e sumariamente evitadas pelos falantes mais escolarizados (cf. 1.2.1). As formas padrão são também protegidas pela **legitimidade** que lhes é atribuída.

Além disso, há uma tradição popular de queixa sobre a língua, segundo a qual os usos não padrão estariam degenerando a língua, e que cabe às autoridades citadas frear essa deturpação da língua correta. Como vemos, o processo de padronização é o principal responsável pelas atitudes positivas ou negativas que os falantes têm em relação à língua (cf. HORA, 2012). Se isso é verdade, e se é verdade que as gramáticas tradicionais são as autoridades responsáveis pela proteção da norma padrão, justificamos assim o uso dessas gramáticas como parâmetro para avaliar as percepções e atitudes linguísticas dos falantes cultos.

2.1.1.1 Sobre a Concordância Verbal

No que diz respeito à concordância verbal, as gramáticas tradicionais apontam que a regra geral é o verbo concordar com o sujeito em *número* e *pessoa*: se o sujeito tiver um só núcleo, o verbo concorda no singular ou no plural, a depender da pessoa gramatical que representa o sujeito; se o sujeito tiver mais de um núcleo, a concordância é feita no plural (cf. BECHARA, 2006, p. 554; CUNHA e CINTRA, 2001, p. 497-498; ROCHA LIMA, 2011, p. 472-473). Este é um fato *morfossintático*, pois a relação entre o verbo e o sintagma nominal sujeito se concretiza por meio das desinências verbais número-pessoais.

As gramáticas tradicionais normativas também concordam que essa regra geral nem sempre é respeitada, e que há exceções à regra. Essas gramáticas buscam recursos para explicar a variabilidade que está presente na concordância verbal, legitimando essa variação, embora não de forma explícita. Postulam, para isso, a ideia de "concordância ideológica", que diz respeito aos casos em que a concordância não se dá da forma esperada. Esse tipo de concordância é rotulado de "silepse de número", e está presente em sentenças como "Os brasileiros somos fortes", em que a concordância se dá com a ideia de pertencimento do falante ao grupo dos brasileiros (por isso o uso da primeira pessoa). Outro caso de variabilidade da concordância verbal abonado pelas gramáticas tradicionais é a concordância facultativa com o sujeito mais próximo: em sentenças como "No carro entraram o professor, a aluna e o pai" é facultado ao falante fazer o verbo concordar apenas com o substantivo mais próximo, "No carro entrou o professor, a aluna e o pai".

Somam-se ainda à regra geral diversas regras complementares que são tratadas como "exceções" ou "particularidades". É o caso de sujeitos com expressões partitivas (*parte de, o resto de, metade de, um dos que...*), com pronomes relativos ou interrogativos (*que, quem, quando...*), com conectivos (*e, ou, ou... ou, nem... nem*), etc. São também casos particulares as orações com o verbo ser ("Tudo é flores" ou "Tudo são flores"?), com a partícula "se" ("Alugam-se casas" ou "Alugam-se casas"?), com a locução *é que*, etc. Todos esses casos demonstram que, por um lado, as gramáticas tradicionais são inconsistentes no tratamento do assunto, chegando a admitir a concordância com elementos da sentença que não são o núcleo do sujeito; e, por outro, que as gramáticas legitimam a variabilidade da regra de concordância, sem, no entanto, admitir de fato a existência de variáveis linguísticas.

2.1.1.2 Sobre a Retomada Anafórica por Objeto Direto

Em relação à retomada anafórica por objeto direto, buscamos nas gramáticas os capítulos referentes aos objetos diretos e aos pronomes pessoais (oblíquos). As gramáticas tradicionais consultadas explicam o objeto direto dentro da seção "complementos verbais". Cunha e Cintra (2001, p. 140) dizem que o objeto direto é o complemento de um verbo transitivo direto, que normalmente vem ligado a esse verbo sem preposição. A seguir, passam a listar as formas que pode assumir o objeto direto: substantivo¹⁹ ("Não recebo dinheiro nenhum"),

¹⁹ Erro conceitual clássico das gramáticas normativas: não é apenas o substantivo que constitui o objeto direto, mas também todas as outras palavras a ele relacionadas e que formam um sintagma nominal completo (no exemplo, "nenhum" faz parte do objeto direto e não é substantivo).

pronome substantivo (“Nunca o interrompi”), numeral (“Já tenho seis lá em casa”) e oração substantiva (“Não quero que fiqueis tristes”). Rocha Lima (2011, p. 299) dá menos atenção à *forma* do objeto direto e se concentra nas *noções semânticas* que esse complemento pode indicar: o ser sobre o qual recai a ação (“Castigar o filho”), o resultado da ação (“Construir uma casa”) ou o conteúdo da ação (“Discutir política”). Tanto Rocha Lima (2011) quanto Bechara (2006, p. 416) indicam que uma das formas de identificar um objeto direto é por meio da possibilidade de esse complemento ser comutado, na 3ª pessoa, com os pronomes átonos *o*, *a*, *os*, *as*: “Não o recebo”, “Já as tenho lá em casa”, “Castigá-lo”, “Construí-la”.

Sobre os pronomes pessoais oblíquos de 3ª pessoa, Rocha Lima (2011, p. 157) indica que “as formas *o*, *a*, *os*, *as* empregam-se em substituição a um substantivo que, sem vir precedido de preposição, completa o regime de um verbo”. Em outras palavras, essas formas pronominais ocupam o mesmo lugar funcional de um objeto direto (complemento sem preposição), sendo por isso chamados pelo autor também de “pronomes objetivos”. No capítulo “Emprego do pronome” (p. 386 a 415) há a menção a diversos casos de usos populares dos pronomes, mas nenhuma diz respeito ao uso de formas retas em detrimento das formas oblíquas ou da omissão desses pronomes (as duas variantes que, como veremos, são as mais produtivas na fala dos brasileiros; cf. 2.1.2). Bechara (2006, p. 175) comenta que “o pronome *ele*, no português moderno, só aparece como objeto direto quando precedido de *todo* ou *só* (adjetivo) ou se dotado de acentuação gráfica” (grifos do autor). Bechara também ignora, portanto, a imensa quantidade de usos não padrão na retomada por objeto direto anafórico.

Um ponto sobre os pronomes oblíquos que recebe muita atenção nas gramáticas tradicionais normativas é a questão da *colocação pronominal*. Essa variável independente não será considerada para efeitos de nossa análise, mas é importante identificar as possibilidades de posicionamento do pronome: quando vem antes do verbo, ocorre o fenômeno da *próclise* (“Me dá um cafezinho”); quando vem após o verbo, ocorre a *ênclise* (“Dá-me um cafezinho”); e quando o pronome vem intercalado entre os morfemas do verbo, temos a *mesóclise* (“Dar-me-ia um cafezinho”). As regras de colocação brasileiras divergem das portuguesas – os brasileiros preferem a próclise, ao passo que a regra geral em Portugal é a ênclise. As gramáticas normativas prescrevem as regras de colocação a partir da variedade portuguesa: “sendo o pronome átono objeto direto ou indireto do verbo, a sua posição lógica, normal, é a ênclise” (CUNHA e CINTRA, 2001, p. 309). É importante notar, por fim, que quando vêm na posição enclítica, os pronomes oblíquos de 3ª pessoa têm sua forma alterada: para *lo(s)*, *la(s)* quando a forma verbal terminar em consoante (que desaparece: *comprar* > *comprá-la*; *fez* > *fê-lo*;

encontramos > encontramos-la); e para *no(s)*, *na(s)* quando terminar em ditongo nasal (*dão-no*, *tem-na*).

2.1.1.3 Sobre as Estratégias de Relativização

Sobre as estratégias de relativização, buscamos os capítulos referentes aos pronomes relativos e às orações adjetivas. Segundo Cunha e Cintra (2001, p. 342), os pronomes relativos recebem esse nome porque se referem, em geral, a um termo anterior, *relacionando* dois termos que são correferentes na oração. Rocha Lima (2011, p. 162) também dá destaque ao fato de que pronomes relativos representam outros termos de uma oração anterior e de que não têm significação própria: “são palavras que reproduzem, numa oração, o sentido de um termo ou da totalidade de uma oração anterior”. Esses pronomes, em verdade, sempre aparecem quando há orações subordinadas adjetivas ocupando a função de adjunto adnominal, e exercem função dupla: relacionam a oração subordinada à principal e estabelecem uma relação entre o substantivo a que se referem e a subordinada. Dessa forma, eles podem ter quase todas as funções que um sintagma nominal pode ter (sujeito, objeto direto, indireto, adjuntos, etc.).

A lista de pronomes relativos trazida por essas gramáticas incluem: *que*, *quem*; *quanto* e flexões (*quanta*, *quantos*, *quantas*); *cujo* e flexões (*cuja*, *cujos*, *cujas*); *o qual* e flexões (*a qual*, *os quais*, *as quais*); e *onde*. Classificam-se todos esses pronomes, com a exceção de *cujo* e suas flexões, como “pronomes substantivos”. Em relação ao pronome relativo *que*, Rocha Lima (2011, p. 403) salienta que, pelo fato de esse pronome representar diversas funções sintáticas, ele acaba se correspondendo com diversas preposições: “A casa *a que* vou. Os elementos *com que* conto. Os recursos *de que* disponho. Não há *por que* desistir do intento. As razões *em que* se estribam” (grifos do autor).

Como já dissemos, os pronomes relativos encabeçam as chamadas orações adjetivas. Sobre esse tipo de oração subordinada, Rocha Lima (2011, p. 333) explica que são orações que “valem por adjetivos”, funcionando como adjunto adnominal. Dentro do período, ligam-se a um termo da oração principal cujo núcleo seja substantivo. Veja:

- (8) a. O aluno estudioso vence na vida
 b. O aluno que estuda vence na vida

Em (8a), o adjetivo *estudioso* se liga ao substantivo *aluno*, funcionando como adjunto adnominal. Em (8b), em vez de um simples adjetivo, temos uma oração inteira funcionando como adjunto adnominal e ocupando o espaço do adjetivo, sendo por isso chamada de oração

adjetiva. Bechara (2006, p. 465), de onde extraímos o exemplo em (8), propõe a existência de um nível profundo em que há uma oração independente *O aluno estuda*, que é transposta pelo pronome relativo *que* para a função de adjunto do sujeito da oração principal *O aluno [...] vence na vida*. Esse transpositor reintroduz o antecedente a que se refere na oração subordinada, acumulando uma função de acordo com a estrutura sintática dessa oração: em *O aluno estuda*, “o aluno” é sujeito; sendo assim, em “O aluno que estuda vence na vida”, o pronome relativo funciona como sujeito da oração subordinada.

Bechara (2006, p. 466) dedica um espaço destacado em sua gramática para salientar a importância de utilizar o “índice preposicional” para introduzir o relativo *que* quando esse pronome está ligado a um verbo com regência indireta. O exemplo que o autor utiliza é construído com o verbo *gostar* (“O livro de que gostas está esgotado”), que, como observa Bagno (2004, p. 81), é o verbo que mais frequentemente se encontra em orações relativas preposicionadas.

Também em Bechara (2006, p. 201), encontramos uma curiosa observação feita pelo gramático em relação às orações relativas que aqui estamos chamando de *copiadoras*, como o exemplo dado pelo próprio autor: “Ali vai o homem que eu falei com ele”. Bechara chama o pronome nessas orações de “relativo universal”, que aparece “despido de qualquer função sinática, como simples transpositor oracional”. Segundo o autor, trata-se de uma construção que desfaz “uma complicada contextura gramatical”, tornando-se um “elemento linguístico extremamente prático”. Apesar de reconhecer a praticidade da relativa copiadora, Bechara deixa claro que se trata de uma opção sintática típica da linguagem coloquial e popular.

2.1.2 O Que Dizem as Pesquisas Linguísticas

As descrições e prescrições feitas pelas gramáticas tradicionais normativas não são as únicas possibilidades de explicação para os fenômenos linguísticos apresentados²⁰. A Linguística moderna e as várias subdisciplinas dessa ciência dispõem de instrumental teórico capaz de fornecer descrições ou explicações muito mais profundas e exatas dos fenômenos linguísticos que a doutrina tradicional. Diversos autores já propuseram “revisões” das descrições da gramática tradicional, apontando seus erros conceituais e inconsistências descritivas (cf., por exemplo, PERINI, 1985; HAUY, 1983) ou até mesmo produzindo

²⁰ De fato, as *descrições* e *explicações* da gramática tradicional não têm a intenção de ser completas ou cientificamente coerentes. Elas estão ali como pano de fundo para as *prescrições* gramaticais, que não podem ser afixadas sem uma nomenclatura metagramatical.

gramáticas completas seguindo teorias linguísticas modernas e dando especial atenção à descrição do PB (cf., por exemplo, AZEREDO, 2002; BAGNO, 2011; CASTILHO, 2010; NEVES, 2000; PERINI, 2005).

Nesta seção, não trazemos as descrições feitas pelas obras citadas, mas sim aquelas propostas por pesquisas de cunho variacionista. Estudos dos fenômenos variáveis sob a ótica da Teoria da Variação apresentam a caracterização da regra variável em termos de variantes mais ou menos utilizadas de acordo com uma série de fatores tanto linguísticos quanto sociais. Usamos essas análises como ponto de partida para a elaboração dos testes de percepção a serem aplicados em nossa pesquisa (cf. 2.2).

2.1.2.1 Sobre a Concordância Verbal

Sobre a concordância verbal, estudos linguísticos revelam que ela constitui um fato variável, podendo ser concretizada ou não pelo falante, a depender de fatores de natureza linguística e extralinguística. Os pioneiros na delimitação desses fatores foram Lemle e Naro (1977), que abordaram a regra de concordância do verbo na linguagem oral empregada por alunos adultos em alfabetização do projeto Mobral no Rio de Janeiro. Os autores comprovaram a produtividade da não concordância nas variedades populares do PB, em comparação com a realização da concordância nas variedades cultas. Isso significa que quanto menor a escolaridade do falante, maiores são as chances de ele não aplicar a regra de concordância canônica. Esse mesmo resultado foi obtido por pesquisas em outras regiões do país: Monguilhott (2001, p. 59) na região de Florianópolis, Rodrigues (1997, p. 146) em Rio Branco e Oliveira (2005, p. 148) em Vitória da Conquista.

Para investigar a opção dos falantes em relação à concordância verbal, os trabalhos que tratam desse fenômeno controlam a influência de vários fatores (também chamados de variáveis independentes; cf. 1.1.1) que podem favorecer ou desfavorecer o aparecimento da variável dependente. A maioria desses trabalhos (como exemplos, cf. os trabalhos citados no parágrafo anterior) costuma testar os mesmos fatores, para fins de comparação. Dessa forma, alguns fatores estão sempre presentes e sempre se mostram relevantes para a aplicação da regra de concordância verbal, tais como: *posição do sujeito em relação ao verbo*, *distância entre o núcleo do sujeito e o verbo*, *paralelismo no nível oracional*, *animacidade do sujeito*, *saliência fônica do verbo*, *a constituição do sujeito*, dentre outros. Discutiremos brevemente esses fatores e os resultados obtidos em vários trabalhos sobre Concordância Verbal, já separando aqueles fatores que nos serão úteis para a elaboração dos testes de percepção.

a) **Posição do sujeito em relação ao verbo:** como adiantamos na Introdução, a posição que o sujeito ocupa em relação ao verbo da oração pode afetar a concordância verbal. O PB classifica-se como uma língua SVO, pois a ordem natural em que os constituintes frasais se dispõem é Sujeito-Verbo-Objeto. No entanto, essa ordem pode ser alterada por razões estilísticas, e o sujeito pode aparecer posposto ao verbo:

(9) Fica dois guardas de vigília na escola²¹

Quando o sujeito está posposto ao verbo, a tendência é que não haja a concordância de número, como no exemplo (9). Segundo Bagno (2003, p. 27), “no português brasileiro, tudo o que se colocar depois do verbo é intuitivamente analisado pelo falante como objeto e, desse modo, mantido fora da esfera da concordância verbal”. Vários estudos concordam com Bagno: Lemle e Naro (1977) testaram ao mesmo tempo a posição e a distância do sujeito em relação ao verbo, e detectaram que o sujeito anteposto seria a categoria mais saliente, pois favorece a concordância verbal, ao passo que o sujeito posposto apresenta baixo índice de probabilidade de aplicação da regra. Outro fator considerado pelos autores é a existência de sujeito explícito: sujeitos ocultos tendem a influenciar a ocorrência da concordância. Outros trabalhos que chegaram aos mesmos resultados sobre a posição do sujeito em relação ao verbo foram: Rodrigues (1997, p. 95 a 113), que atestou apenas 21% de frequência de concordância para sujeitos pospostos ao verbo e até 66% para sujeito preposto; Monguilhott (2001, p. 44 a 46), que chegou à frequência de 52% de aplicação da concordância para sujeito posposto e 84% para preposto; e, talvez o resultado mais impressionante, o de Monte (2007, p. 86 a 92), que, estudando 20 entrevistas de falantes da cidade de São Carlos – SP, não encontrou concordância em nenhuma das ocorrências de sujeito posposto (0%), e 25% com sujeito preposto.

b) **Distância entre o núcleo do sujeito e o verbo:** outro fator que comprovadamente exerce algum nível de influência na regra de concordância verbal é a distância entre o sujeito e o verbo da oração. Alguns estudos (cf. LEMLE e NARO, 1977; RODRIGUES, 1997; MONTE, 2007; OLIVEIRA, 2005) preferem misturar este critério com o anterior, criando categorias que vão, por exemplo, do “sujeito preposto separado do verbo por um conjunto de palavras tônicas”, passando pelo “sujeito imediatamente anterior ao verbo” até chegar em “sujeito posposto ao verbo”. Outros estudos separam esses dois fatores (cf. MONGUILHOTT, 2001). De qualquer forma, tem-se detectado que quanto maior a distância entre o sujeito e o verbo, menor é a probabilidade de haver concordância.

²¹ Todos os exemplos desta seção foram adaptados dos próprios trabalhos citados.

O trabalho de Rodrigues (1997, p. 95 a 113) revelou que se o sujeito está imediatamente antes do verbo (“Eles começaram a fazer parte da igreja”), a probabilidade de haver concordância é de 60%; já se o sujeito está mais ou menos próximo (“Meus pais mesmo já morreram”) ou distante (“As seguranças hoje em dia em lugar mais nenhum *existe*, né?”) do verbo, as chances de haver concordância caem para 51% e 49%, respectivamente. Monte (2007) dividiu esse fator em duas categorias: “sujeito anteposto imediatamente” e “sujeito anteposto distante” e concluiu que a frequência de concordância com a primeira categoria é de 25%, ao passo que com a segunda é de 18%. Monguilhott (2001, p. 57 a 58) estudou como aquilo que chama de “material interveniente” afeta a aplicação da regra de concordância. A autora utilizou 10 categorias, cada uma representando um número de sílabas que vai de zero a “oito ou mais”, e testou como cada quantidade de material entre o sujeito e o verbo influencia a concordância. Os resultados indicaram que, pelo menos nos dados usados pela pesquisadora, esse fator não tem nenhuma influência na concordância verbal.

c) **Paralelismo formal:** o próximo fator linguístico amplamente citado nas pesquisas variacionistas sobre a concordância verbal de 3ª pessoa do plural é o paralelismo oracional. Segundo esse princípio, há uma tendência a que zeros linguísticos levem a zeros linguísticos, e marcas levem a marcas. Aplicado à concordância verbal, esse princípio se mostra na tendência observada de marcas explícitas de plural presentes no último elemento do sujeito levarem a marcas explícitas de plural no verbo. De modo análogo, a ausência de plural no sujeito leva à ausência de plural no verbo.

Há várias possibilidades de construção do sujeito, pois a concordância nominal também é uma regra variável em português. A norma padrão preconiza o uso do plural em todos os elementos do sintagma nominal que exerçam a função de adjuntos adnominais (exceto aqueles que se apresentam na forma de um sintagma preposicionado ou de oração subordinada adjetival): “Os meninos travessos”. Por essa concordância ser também redundante, variedades populares do PB acabam por apagá-la de outros elementos do SN que não sejam seu primeiro elemento: “Os menino travesso” seria uma possibilidade. Também é possível atestar concordâncias parciais dentro do SN, em que alguns elementos vão para o plural, mas não todos aqueles que deveriam ir: “Os meninos travesso” (cf. TARALLO, 1986).

Monguilhott (2001, p. 47 a 48) observou que a presença de forma plural no último elemento do sujeito leva a uma probabilidade de 84% de aplicação da concordância, e que a ausência de forma plural no último elemento do sujeito favorece a não concordância, com 61% de chances de a regra ser aplicada. Monte (2007, p. 73 a 77) testou exatamente as mesmas

categorias e obteve o seguinte resultado: 28% de aplicação da regra de concordância quando o último elemento do sujeito apresenta presença de plural; 11% de aplicação quando apresenta ausência de plural.

d) **Animacidade do sujeito:** suspeita-se que sujeitos de referência animada (*o peixe, as criaturas, etc.*) ou humana (*as pessoas, o advogado, etc.*) favoreçam a realização da concordância verbal, enquanto que sujeitos de natureza inanimada ou não humana (*barco, lápis, céu*) não a favoreçam. Monguilhott (2001, p. 49 a 50) controlou se o traço [+/- humano] teria alguma relação com a concordância, e chegou à conclusão de que quando sujeito tem o traço [+ humano] a probabilidade de concordância verbal é de 84%, contra apenas 58% dos sujeitos [- humano].

e) **Saliência fônica do verbo:** a saliência fônica do verbo diz respeito à diferença fonética entre a forma singular e a plural de 3ª pessoa do verbo analisado: quanto maior a diferença fonética, maior a possibilidade de haver concordância. Todos os estudos que consideraram esse fator chegaram à conclusão de que ele é o mais relevante para a existência ou não de concordância, ainda que diferentes pesquisas tenham trabalhado com níveis de saliência diversos. Rodrigues (1997, p. 113 a 130), seguindo proposta de Lemle e Naro (1977), utilizou uma hierarquia de categorias que vão das diferenças menos salientes às mais salientes entre os pares singular/plural dos verbos:

- 1ª categoria: verbos com desinência átona e oposição singular/plural realizada pelo acréscimo de um travamento nasal (*come/comem*);
- 2ª categoria: verbos com desinência átona e oposição singular/ plural realizada pela adição de segmentos nasalizados (*faz/fazem*);
- 3ª categoria: uma vogal oral tônica no singular se opõe a um ditongoônico nasal no plural (*dá/dão, está/estão*);
- 4ª categoria: formas em que o acento recai na desinência no singular e na vogal temática no plural (*falou/falaram, vendeu/venderam*);
- 5ª categoria: formas em que o acento recai no radical no singular e na vogal temática no plural (*fez/fizeram, trouxe/trouxeram*);
- 6ª categoria: formas totalmente diferentes de singular e plural (*é/são*).

Os resultados obtidos por Rodrigues (1997) indicam que, quanto menor a diferença entre as formas singular e plural do verbo, maior será a probabilidade de não haver concordância: a 1ª categoria de verbos é a que mais condiciona a não concordância (46% de aplicação da regra);

a 5ª categoria, por outro lado, é a que mais condiciona a concordância (89% de aplicação da regra), ao lado da 4ª categoria (75%) e da 3ª (77%); a 6ª categoria, apesar de ter duas formas totalmente diferentes para a oposição singular/ plural leva a 74% de aplicação da regra, menos do que a 3ª, 4ª e 5ª categorias. Isso pode se explicar pelo fato de essas formas serem foneticamente leves (são monossílabos) e serem extremamente frequentes na fala. Resultados similares foram encontrados por Monguilhott (2001, p. 40 a 44), Monte (2007, p. 69 a 73) e Oliveira (2005, p. 115 a 121).

f) **Constituição do sujeito:** esta variável morfossintática reúne fatores que dizem respeito à constituição formal do sujeito de 3ª pessoa do plural. Rodrigues (1997) encontrou sete categorias que podem representar os sujeitos:

- Sujeito oculto ou subentendido: “Machuraram a cabeça dele”
- Pronome pessoal explícito: “Eles não querem nem saber”
- Núcleo plural: “Minhas irmãs todas já eram casadas”
- Núcleo sem marca de plural, mas com determinante plural: “Essas criança que vive passando fome”
- Outros pronomes ou numeral: “Dois queriam bater nele”
- Pronome relativo “que”: “(Pessoas) que estão precisando de apoio”
- Dois ou mais núcleos no singular ou no plural: “Morreram minha mãe e meu pai”

No decorrer da pesquisa, a autora resolveu amalgamar as categorias “sujeito oculto ou subentendido” e “pronome relativo ‘que’”, pois ambos apresentaram resultados muito similares. A mesma coisa foi feita com as categorias “outros pronomes ou numeral” e “núcleo plural”. Como resultado final, Rodrigues (1997, p. 132) observou que, quando o pronome pessoal está explícito, a aplicação da regra de concordância chega a 73%; quando o sujeito está oculto ou subentendido, ou quando é retomado pelo relativo *que*, a concordância aparece em 58% dos casos; no caso dos núcleos plurais, juntamente a outros pronomes ou numeral, a concordância ocorre em 49% dos casos; SNs com núcleo sem plural mas com determinante no plural levam a 33% de concordância; e, por último, sujeitos com dois ou mais núcleos foram os que menos condicionaram a concordância: apenas 15%.

Monguilhott (2001, p. 54 a 57) chegou aos seguintes resultados para esse fator: quando o sujeito é composto por pronome pessoal ou pronome demonstrativo (*esses, aquelas*), a chance de ocorrer concordância é de 88%; quando é constituído por SN + pronome relativo *que*, a aplicação da regra de concordância cai para 83%; com o sujeito pleno simples (igual à categoria

“núcleo plural” de Rodrigues (1997)) ou nu (quando aparece sozinho na sentença, sem adjuntos adnominais: “Navios entram aqui”), a concordância ocorre em 71% dos casos; com quantificador (*todos, muitos*) ou pronome indefinido (*alguns, poucos, outros*), a aplicação da regra é de 60%; e, com SNs compostos (dois ou mais núcleos), a concordância ocorre em 40% dos casos. Como podemos ver, os resultados de Monguilhott (2001) concordam em boa parte com os de Rodrigues (1997), apesar de este trabalho ter sido realizado na região Norte e aquele na região Sul do país: em ambos, o sujeito com pronome lexical explícito é o tipo de sujeito que mais condiciona a concordância verbal e os sujeitos compostos são os que menos o fazem. Também se igualam em certa medida os resultados para sujeitos retomados pelo relativo *que* e os sujeitos plenos simples com núcleo plural.

Dos fatores que descrevemos nesta seção, utilizamos os seguintes para a elaboração dos testes de percepção:

- I. Posição e distância do sujeito em relação ao verbo, com três categorias: 1. Sujeito preposto distante do verbo (= com pelo menos duas palavras tônicas separando o sujeito do verbo); 2. Sujeito preposto imediatamente ao verbo; e 3. Sujeito posposto ao verbo;
- II. Animacidade do sujeito, com duas categorias: 1. Sujeito animado; e 2. Sujeito inanimado;
- III. Constituição do sujeito, com duas categorias: 1. Sujeito simples (com pronome lexical ou sintagma nominal simples); e 2. Sujeito composto.

As categorias que utilizamos foram escolhidas de forma arbitrária, a partir dos resultados mais significativos das pesquisas descritivas que aqui resenhamos. O fator “paralelismo formal” não foi utilizado porque ele depende da existência de sujeitos com variação na concordância nominal, o que poderia levar a conflitos de análise. O fator “saliência fônica” obviamente não pôde ser usado por ser uma variável independente de natureza fonológica, e nossos testes são escritos (cf. 2.2.3). Dentro do fator “constituição do sujeito”, não utilizamos a categoria “sujeito retomado pelo relativo *que*” por conta da dificuldade que tem esse fator de interagir com as outras categorias (cf. 2.2); também amalgamamos as categorias “sujeito com pronome lexical” e “sujeito pleno com núcleo plural”, pois pelo menos um dos trabalhos analisados mostrou que a influência de cada um é praticamente a mesma. A elaboração das orações-teste a partir desses fatores e categorias será explicada no capítulo 2.2.

2.1.2.2 Sobre a Retomada Anafórica por Objeto Direto

Como vimos em 2.1.1, as gramáticas tradicionais em geral não citam a possibilidade de o objeto direto, no PB, aparecer na sentença sob a forma de pronome pessoal reto de 3ª pessoa (*ele, ela*) ou sob a forma de uma categoria vazia, que em Linguística é chamado de *objeto nulo*. Mas, já na década de 50, Câmara Jr. afirmava que “um dos traços mais característicos do português do Brasil é o uso de *ele* (e suas variantes de feminino e plural), como um acusativo; ex: *vejo ele*, em lugar de *vejo-o*” (CÂMARA JÚNIOR, 2004, p. 96, grifos do autor). Desde então, vários pesquisadores²², tanto da Sociolinguística quanto de outras correntes teóricas, em especial da Linguística Gerativa, têm se debruçado sobre o fenômeno de perda do sistema de clíticos de 3ª pessoa no PB. Segundo Roberts (1993),

Um desenvolvimento surpreendente no PB atual, que não é compartilhado (até onde sei) por outra língua românica, é a perda do sistema de clíticos de complemento, ainda em processo. Enquanto que o português europeu possui um paradigma de clíticos objetos que é, deixando de lado a mudança fonológica e a posição sintática, essencialmente o mesmo encontrado no francês, espanhol e italiano, esses clíticos parecem estar saindo de uso no PB. Este é particularmente o caso de clíticos de 3ª pessoa. Os clíticos são substituídos de várias formas: por um pronome pleno (potencialmente tônico), por um SN repetido ou, mais interessante do ponto de vista das questões teóricas, por uma categoria vazia (ROBERTS, 1993, p. 417-8).

As pesquisas de Duarte (1989) e Cyrino (1994) indicam que, em geral, o clítico acusativo é a forma menos usada para realizar objetos diretos anafóricos no PB. Cyrino (1994, p. 171) constata que os pronomes tônicos e objetos nulos passaram a substituir o clítico acusativo desde meados do século XIX. Galves (2001, p. 13-14) argumenta que o clítico acusativo tem um estatuto marcado no PB, “que se explica perfeitamente se admitirmos que não é o produto da aquisição natural da linguagem, mas de uma aprendizagem posterior no contexto escolar e no convívio com a escrita culta”.

Dentro das pesquisas variacionistas, vários autores têm se dedicado a decifrar quais seriam os fatores linguísticos e sociais que mais influenciariam o uso de cada uma das três possibilidades de realização anafórica do objeto direto. O trabalho variacionista clássico sobre o assunto é o de Duarte (1989)²³. A autora utilizou um *corpus* que continha gravações da fala natural de 50 informantes nativos da cidade de São Paulo, além de um corpus da linguagem

²² Por exemplo: Duarte (1989), Corrêa (1991), Silva (2004) e Corrêa (2014), para pesquisas sociolinguísticas; Cyrino (1994) e Galves (2001) para pesquisas gerativistas.

²³ Além das três estratégias de realização anafórica do objeto direto que vimos comentando por todo este trabalho, Duarte (1989) atestou também a possibilidade de o objeto direto estar representado na forma de um sintagma nominal pleno ou de ser retomado pelo pronome “isso”. Como essa variante não é necessariamente estigmatizada nem considerada padrão, uma vez que não se realiza por meio de um pronome, não levaremos em consideração essa possibilidade.

televisiva, com gravações de episódios de novelas e de entrevistas. A escolha dos informantes levou em consideração três níveis de escolaridade e três níveis de faixa etária, totalizando nove células de perfis sociais, com uma célula extra composta por jovens de 15 a 17 anos. Nessa pesquisa ficou constatado que, de todas as formas variantes do objeto direto anafórico, a menos utilizada é o clítico acusativo (4,9%), seguindo-se o pronome lexical (15,4%), os SNs lexicais plenos e o demonstrativo *isso* (17,1%). A categoria vazia alcançou o maior índice (62,6%). Resultados similares foram atestados pelos trabalhos de Corrêa (1991, p. 55), que analisou textos orais e escritos de alunos do ensino fundamental de Campinas, SP; Silva (2004, p. 112), com 12 informantes de uma comunidade rural afro-brasileira na Bahia; e Corrêa (2014, p. 60), com 12 informantes de uma comunidade rural do Alto Araguaia, MT. Dentre os trabalhos que estudaram somente a língua escrita, Pereira e Coelho (2013, p. 10) obtiveram resultados parecidos com os encontrados na língua falada: 30% de uso do pronome reto, 24% do oblíquo e 46% da categoria vazia, o que mostra que até mesmo em textos mais monitorados a alternativa padrão tem perdido terreno.

Em busca dos condicionamentos linguísticos atuantes na realização da variável, Duarte (1989) levantou fatores de natureza morfológica (a forma verbal), sintáticos (a estrutura projetada pelo verbo) e semânticos (traço [+/- animado] do objeto). Comentaremos esses três fatores, cotejando ao mesmo tempo outros estudos que se valeram das mesmas categorias nas pesquisas sobre o objeto direto anafórico.

a) **Forma verbal:** segundo Duarte (1989, p. 21), os clíticos acusativos, quando aparecem na fala dos informantes, são proclíticos em 40,2% dos casos e enclíticos em 59,8%. Nas construções proclíticas, o verbo está num tempo simples do indicativo: “o conhecia”, “a via”; nas enclíticas, a grande maioria segue verbos no infinitivo, adotando a forma “lo”: “informá-lo”, “chamá-lo”. Os pronomes lexicais também aparecem quase sempre com tempos simples e as categorias vazias supera todas as demais variantes, não importante a forma verbal. A autora comenta o interessante fato de que o clítico nunca aparece com o imperativo, em tempos compostos ou em locuções verbais com gerúndio.

b) **Estrutura projetada pelo verbo:** Duarte (1989, p. 22) chamou de “condicionamento sintático” o fato de que a estrutura projetada pelo verbo (em termos gerativos, a grade temática do predicador; cf. DUARTE, 2007) parece também interferir na escolha entre as possíveis variantes desta variável. A autora testou as seguintes estruturas sintáticas:

- Estrutura simples: verbo + objeto direto [SN ou oração]: “Na volta, o carro matou [ele]”, “Eu sei que ele roubou, mas ninguém mais sabe [Ø]”
- Estrutura com predicativo: verbo + objeto direto [SN ou oração] + [predicativo]: “Eu acho [ele] [sensacional]”, “Eu queria ter uma irmã. Eu acho [Ø] [tão bom]”
- Estrutura complexa: verbo + objeto direto [SN] + [oração no infinitivo ou gerúndio]: “Quando nós estávamos saindo da loja, nós vimos [eles] [quase parando o carro]”.

Observou-se que, no caso das estruturas simples com objeto direto SN, o uso da categoria vazia é de 62,3%, o do pronome lexical é de 14,1% e o do clítico é de 5,4%²⁴. Em estruturas simples com objetos oracionais, a categoria vazia aparece em 79,7% dos casos, o pronome não aparece (0%) e os clíticos aparecem em 1,3%. Nas estruturas com predicativo, quando o objeto direto é uma SN, a categoria vazia se apresenta em 46,6% dos casos, o pronome lexical em 35,6% e os clíticos em 5,9%; quando a estrutura com predicativo tem um objeto direto oracional, a categoria vazia está presente em 84,6% das orações, e as outras duas possibilidades não aparecem. Por fim, nas estruturas complexas, o pronome lexical ganha da categoria vazia e aparece em 71,9% dos casos, contra 19,3% de apagamento e apenas 1,8% de aparecimento dos clíticos. A hipótese da autora para essa mudança de cenário é que o objeto direto dessas construções exerce a função de sujeito da oração subordinada, o que determinaria uma tendência maior à lexicalização do objeto anafórico.

c) **Traço [+/- animado] do objeto:** não só Duarte (1989), mas vários outros estudos que levaram em consideração a variável “traço semântico [+/- animado] do objeto”, demonstraram que esse é o fator que mais influencia o uso das possibilidades de retomada por objeto direto anafórico (cf. CORRÊA, 1991, p. 76; SILVA, 2004, p. 114; PEREIRA e COELHO, 2013, p. 314; CORRÊA, 2014, p. 71). Duarte (1989, p. 24) observou que o objeto que apresenta traço semântico [+ animado] condiciona fortemente o uso do objeto preenchido, isto é, das variantes pronome lexical e clítico acusativo (92,4% e 78,4%, respectivamente). Já objetos com traço semântico [- animado] condicionam o aparecimento da categoria vazia, que é a variante escolhida em 76, 3% dos casos. Silva (2004) atestou o uso do pronome lexical em 65% dos casos com objetos [+ animado] e o uso do clítico em 97% dos casos com objetos [- animado]. Os dados de Pereira e Coelho (2013) se revelaram ainda mais expressivos: em 99%

²⁴ As porcentagens não somam 100% porque, como dissemos na nota 26, a autora incluiu uma outra variante, a do sintagma nominal simples, que aqui não estamos considerando.

dos casos de categoria vazia, o objeto tinha traço semântico [- animado] e em 94% dos usos do pronome lexical, o traço do objeto era [+ animado].

Dos fatores relevantes para a escolha entre as três opções de retomada por objeto direto anafórico, escolhemos os seguintes para subsidiar a elaboração do teste de percepção:

- I. Estrutura projetada pelo verbo, com três categorias: 1. Estrutura simples; 2. Estrutura com predicativo; e 3. Estrutura complexa.
- II. Animacidade do objeto, com duas categorias: 1. Objeto animado; e 2. Objeto inanimado.

A variante “forma verbal” não foi utilizada, pois o condicionamento por ela operado não é tão decisivo quanto o dos outros fatores.

2.1.2.3 Sobre as Estratégias de Relativização

Antes de proceder à análise de estudos anteriores sobre as estratégias de relativização, convém fazer uma distinção. Como vimos em 2.1.1, as orações adjetivas/ relativas podem exercer várias funções sintáticas:

- (10) a. O cachorro que morreu não era meu
 b. O leite que o João derramou estava fedendo
 c. A menina de que eu gosto é aquela morena
 d. O professor com que eu falei se chama Márcio
 e. Esse é o amigo de cuja casa ela não sai

Em (10a), o pronome relativo *que* exerce a função de sujeito do verbo *morrer* ("O cachorro morreu") e em (10b), é objeto direto do verbo *derramar* ("O João derramou o leite"). Percebemos que, nessas orações, não há diferença entre a relativa padrão e a cortadora, pois não há preposição para ser cortada, uma vez que sujeitos e objetos diretos não são sintagmas preposicionais²⁵. Em (10c) e (10d), temos a versão padrão de orações com o pronome relativo atuando como objeto indireto ("Eu gosto da menina") e adjunto adverbial ("Eu falei com o professor"). Essas orações têm versões copiadoras e cortadoras:

- (10) c'. A menina que eu gosto dela é aquela morena
 c". A menina que eu gosto é aquela morena
 d'. O professor que eu falei com ele se chama Márcio

²⁵ Ou, em termos tradicionais, não são regidos por preposição.

d". O professor que eu falei se chama Márcio

Em (10e), temos um caso de uso do pronome *cujó* em uma função que podemos chamar de "genitiva" ou de adjunto adnominal ("Ela não sai da casa dele"). Esse pronome está praticamente extinto da gramática dos falantes brasileiros (cf. BAGNO, 2004, p. 84), mas ainda aparece em textos escritos (cf. SANTOS, 2015). Orações com o pronome *cujó* também têm versões copiadoras e cortadoras:

- (10) e'. Esse é o amigo que ela não sai da casa dele
 e". Esse é o amigo que ela não sai da casa

Como nosso objetivo é testar a percepção dos falantes sobre as três estratégias em situações similares de uso, optamos por tratar neste trabalho especificamente das orações relativas que estão em posições sintáticas preposicionadas, também chamadas de *oblíquas*: objeto indireto, adjunto adverbial e adjunto adnominal/ genitivo, também incluindo os poucos casos de existência de complemento nominal ("A prova de que eu tinha medo...").

O trabalho pioneiro sobre as construções relativas em português brasileiro é o de Mollica (1977, apud TARALLO, 1983, p. 20). A autora utilizou como *corpus* gravações com falantes adultos em fase de alfabetização do projeto Mobral. Foram controlados fatores semânticos (animacidade, especificidade e coletividade do sintagma antecedente relativo) e um fator psicolinguístico (a distância entre o relativo e seu antecedente), ligado a um fator sintático (função sintática do elemento relativizado). Como não foram encontrados casos de relativas padrão com o movimento da preposição (apenas cortadoras), a autora resolveu agrupar as relativas padrão de sujeito e objeto com as relativas oblíquas não padrão cortadoras de um lado e as relativas copiadoras de outro. Os resultados de Mollica (1977) mostraram que a relativa cortadora é a predominante em todas as funções sintáticas, até mesmo naquelas regidas de preposição, com 61,4% de cortadoras e 38,6% de copiadoras. O trabalho de Mollica foi especialmente importante para mostrar que as relativas padrão com funções oblíquas estão ausentes na fala de indivíduos adultos pouco escolarizados, o que sugere que esse tipo de construção não faça mais parte da gramática brasileira, só sendo usada pelos falantes após a fase de escolarização.

Embora Mollica (1977) seja o trabalho pioneiro no estudo das orações relativas, é a tese de doutorado de Fernando Tarallo (TARALLO, 1983) que traz a descrição mais pormenorizada, tanto no aspecto sincrônico quanto no diacrônico, das estratégias de relativização no PB. Esse pesquisador também constatou a baixa ocorrência das relativas padrão nas funções regidas de

preposição e a alta frequência das relativas cortadoras. Diacronicamente, Tarallo considerou esse fenômeno como a derivação de uma mudança pela qual o PB está passando: seu estudo demonstrou que a relativa copiadora é um fenômeno muito antigo em português, ao passo que a relativa cortadora se constitui em uma inovação do PB. A explicação para essa inovação está no apagamento da cópia pronominal em todas as posições, juntamente com o apagamento da preposição em funções oblíquas, decorrentes de mudança no sistema pronominal da nossa língua.

Tarallo (1983) e outros pesquisadores posteriores (cf. CORRÊA, 1998; SILVA, 2005; VALE, 2014) levantaram um conjunto de fatores linguísticos que podem influenciar a escolha entre uma das variantes dessa variável. Alguns exemplos são:

a) **Função sintática do pronome relativo:** nesse fator, levantam-se as diferentes funções sintáticas que podem ser exercidas pelo pronome relativo. Tarallo (1983, p. 174) observou que, em funções oblíquas e de objeto indireto, a estratégia cortadora predominou com 82,2% e 75%, respectivamente. Já na função genitiva (adjunto adnominal de posse), a estratégia copiadora predominou com 52,9% das ocorrências. Ao cruzar esses resultados com a variável classe social, o autor concluiu que as classes mais altas utilizam a relativa padrão em apenas 18,2% dos casos e a copiadora em 7,8% dos casos, sendo o uso da relativa cortadora predominante não só nessa, mas em todas as classes sociais. Silva (2005), trabalhando com 93 entrevistas de informantes de vários níveis de escolaridade, observou que a estratégia cortadora é predominante em estruturas com adjunto adverbial (68%), complemento relativo (91%) e complemento nominal (100%), cedendo sua liderança à copiadora em casos de adjuntos adnominais (67% para copiadora, 33% para padrão e 0% para cortadora). Vale (2014) dividiu seus resultados em duas tabelas: uma para os falantes mais letrados e outra para os menos letrados. Na fala dos mais letrados, a relativa padrão predomina nas funções de complemento oblíquo (72%), predicador locativo (86%), complemento nominal (75%) e adjunto adverbial (77%); no genitivo, ocorrem 15,5% de padrão, 38,5% de copiadoras e 46% de cortadoras. A autora atribui a alta ocorrência da forma padrão ao uso do pronome "onde". Nos falantes com menos grau de letramento, em todas as funções, exceto na de predicador locativo, prefere-se a relativa cortadora: adjunto adverbial (93%), complemento oblíquo (81%) e genitivo (56%). Para o predicador locativo, 60% das ocorrências foi de relativas padrão.

b) **Animacidade do antecedente:** supõe-se que o traço [+ animado] do antecedente do pronome relativo pudesse favorecer a escolha da relativa copiadora, ao passo que o traço [- animado] pudesse favorecer a estratégia padrão. Tarallo (1983, p. 92), por exemplo mostrou

que das 162 relativas copiadoras encontradas em seu corpus, em 124 o antecedente era [+animado].

c) **Preposição regida pelo verbo:** Corrêa (1998, p. 116) propõe dividir as preposições requeridas pelo verbo e que acompanham o pronome relativo em "fortes", que resistem ao apagamento pelo fato de sua supressão prejudicar o conteúdo semântico da oração (*sobre, para, entre, etc.*), e "fracas", que são apagadas com naturalidade (*em, de, com, etc.*). Segundo os resultados da autora, as preposições fracas foram as que mais condicionaram o uso das estratégias não padrão: *de* (87%), *a* (71%) e *com* (75%), ao passo que as preposições fortes levaram ao uso da relativa padrão em 75% dos casos. A preposição *em* aparece em grande quantidade com relativas não padrão (56%), mas também ocorre muito com relativas padrão (44%). Corrêa (1998) justifica esse comportamento argumentando que a preposição *em* é bastante frequente em relativas com a função de adjunto adverbial, as mais numerosas em seu corpus e que mais apresentam a estratégia padrão. Bagno (2000, p. 194-195) também notou a alta incidência da preposição *em* e seu alto grau de apagamento (41% dos casos), embora a forma pronominal "em + que" apresente maior ocorrência por fazer referência a tempo e lugar.

Vários outros fatores foram testados nos mais diversos trabalhos sobre relativização no PB e seria improdutivo listar todos aqui. Apenas a título de exemplo, listamos outros fatores usados no trabalho de Tarallo (1983): *função sintática do antecedente, posição do relativo em relação à oração principal e em relação ao antecedente, posição do pronome dentro da relativa, grau de recursividade da relativa, existencialidade e restritividade da relativa, pluralidade e definitude da relativa.*

Dos grupos de fatores condicionadores das estratégias de relativização mencionados, escolhemos dois para utilizar na elaboração dos testes de percepção:

- I. Função sintática do pronome relativo, com três categorias: 1. Objeto indireto; 2. Adjunto adverbial; e 3. Adjunto adnominal (genitivo).
- II. Animacidade do antecedente com duas categorias: 1. Antecedente animado; e 2. Antecedente inanimado.

Não selecionamos o fator "preposição regida pelo verbo" pela dificuldade de encontrar verbos transitivos indiretos que regem as preposições fortes – essas preposições são muito mais frequentes na função sintática de adjunto adverbial.

2.2 A Elaboração do Teste de Percepção

Para avaliar as atitudes linguísticas dos falantes cultos em relação às regras variáveis escolhidas para análise, utilizamos testes de percepção. Segundo Tarallo (1986, p. 55), “como o próprio nome indica, no teste de percepção você solicitará a seu informante que se manifeste em relação à aceitabilidade ou não de certas variantes”. O pesquisador cria uma bateria de orações (que, apesar de não conter exatamente *questões*, chamaremos de “questionário”²⁶) contendo exemplos da regra variável sendo aplicada e da regra variável não sendo aplicada, isto é, com cada uma das variantes *padrão* e *não padrão*. As sentenças não devem seguir qualquer ordem para evitar que o informante reconheça o padrão de organização do teste.

O uso do questionário pode ser relacionado à hipótese de que a percepção da saliência linguística se manifesta por meio da *surpresa* (cf. RÁCZ, 2013, cap. 2). Os informantes são levados a imaginarem um cenário em que as orações testadas seriam usadas, e isso afeta a percepção que eles têm dessas orações: caso esteja presente um elemento que o falante não reconhece como próprio do tipo de interação imaginado, ele o marcará como inadequado. A surpresa causada pela presença do elemento estranho poderá nos dizer que se trata de um fator saliente dentro das variáveis estudadas. A depender das opções de opinião disponibilizadas aos informantes, podemos ainda levantar qual a atitude dos informantes em relação às variáveis e variantes testados.

Nosso questionário foi precedido da seguinte introdução:

Imagine que as frases a seguir tenham sido ditas em contextos formais de conversação, entre pessoas que não se conhecem muito bem. Leia cada frase avaliando sua correção linguística e analise-a por alguns segundos. Não dê muita importância a possíveis problemas de ortografia ou pontuação (vírgulas, pontos, etc.). Depois de refletir sobre o tipo de erro que você encontrou ou se não encontrou erro nenhum, marque uma das seguintes opções para cada frase:

Não há erro: a frase está bem construída. O produtor da frase provavelmente é bem escolarizado.

Errada, mas aceitável: a frase contém um erro, mas é um erro pouco grave na sua opinião. O produtor da frase provavelmente é bem escolarizado, mas cometeu um lapso que muitos cometem.

²⁶ Rosa (2015, p. 152-153) prefere chamar esses tipos de questionários “que têm por objetivo medir as atitudes ou crenças de um indivíduo” de *opinários* ou *escalas de atitudes*.

Inaceitável: a frase contém um erro grave. O produtor da frase provavelmente tem escolaridade baixa ou mediana.

Obs.: se assim desejar, leia as frases em voz alta antes de se decidir. Isso pode facilitar bastante.

Convém comentar alguns dos termos utilizados nessa introdução. Utilizamos o termo “frase”, no lugar de outros termos pertinentes como “oração” ou “sentença”, por ser o mais difundido entre leigos em Linguística. Também utilizamos o termo “correção linguística” em detrimento de outros como “correção gramatical” ou “impropriedade linguística/ gramatical” por ser mais genérico e usual: os falantes em geral têm uma noção maior sobre o que seja língua e menor sobre o que venha a ser gramática (é mais comum, por exemplo, ouvirmos certos tipos de julgamento sobre a fraca habilidade linguística de um falante com frases do tipo “ele não sabe português” em vez de frases como “ele não sabe gramática”). Utilizamos também livremente o termo “erro” por imaginarmos que os informantes não necessariamente tenham conhecimento sobre as polêmicas noções de “certo” e “errado” em língua, que a Linguística moderna já concorda em substituir por “adequado/ inadequado” e similares (cf. POSSENTI, 1996).

A situação que deveria ser imaginada pelo informante ao avaliar as questões é uma conversa formal entre duas pessoas que pouco se conhecem. Isso deveria levá-los a entender que a *norma culta* deveria ser utilizada na interação, e julgar as orações a partir desse entendimento. Para facilitar esse intento, as questões traziam itens lexicais relativos a um contexto empresarial, onde ocorrem conversas sobre trabalho, negócios e política, assuntos que tendem a ser discutidos em um tom mais formal.

Para computar a percepção dos falantes cultos de modo objetivo, criamos três “níveis” de (in)adequação das orações apresentadas. Nosso objetivo foi medir a opinião dos informantes sobre (i) a **aceitabilidade** da sentença e (ii) sua **atitude** em relação a ela. A aceitabilidade diz respeito ao julgamento dos falantes sobre a (a)gramaticalidade de cada sentença (cf. 1.2.2), de modo que o informante teria duas opções: se considerasse a oração *gramatical*, ele deveria marcar o primeiro nível: “Não há erro”; e, caso considerasse a oração *agramatical* (isto é, fugindo fortemente da língua culta), deveria marcar um dos dois outros níveis: “Errada, mas aceitável” ou “Inaceitável”. Esses dois últimos níveis dão margem a diferentes graus de atitude linguística do falante (cf. 1.2.3) e podem ajudar a identificar o nível de estigmatização das variáveis: o falante pode considerar certos erros “mais errados” que outros, e identificá-los como sendo “inaceitáveis” ou “aceitáveis”. Os níveis de inadequação foram misturados a

descrições de natureza social (nível de escolaridade) de modo a guiar o informante em seu julgamento das orações. Optamos por essa descrição dos níveis porque as pesquisas sociolinguísticas apontam que, de fato, a escolaridade dos falantes é um dos fatores mais importantes no uso de formas padrão e não padrão (cf. VOTRE, 2013).

Cada variável escolhida para análise tem uma quantidade X de variantes e uma quantidade Y de fatores condicionadores; cada fator condicionador se subdivide em uma quantidade Z de categorias (cf. 1.1.1). Como as diversas categorias interagem nas orações, cada questão elaborada deve conter uma das alternativas possíveis de cada fator Y. Para encontrar o número de questões que devem ser elaboradas para cada variável, precisamos multiplicar a quantidade Z de categorias de cada fator Y, e multiplicar o resultado pela quantidade X de variantes. Por exemplo: a variável "concordância verbal" tem 2 variantes (*padrão* e *não padrão*), e 3 fatores ("posição e distância do sujeito", com 3 categorias; "animacidade do sujeito", com 2 categorias; e "constituição do sujeito", com 2 categorias). Multiplicando as categorias de cada fator, temos que $3*2*2 = 12$. Multiplicando esse resultado pelo número de variantes, temos que $12*2 = 24$. Dessa forma, tínhamos 24 interações possíveis entre as categorias e variantes dessa variável, e portanto tínhamos de formular 24 orações para o teste de percepção.

2.2.1 A Confecção das Questões

Nesta seção, descreveremos o processo de criação de cada uma das orações que constam do teste de percepção. A ordem em que as orações aqui aparecem **não é** a mesma em que estão dispostas no questionário. Cada oração é identificada por um código que indica: 1. Qual variável ela representa; 2. Qual variante ela representa; 3. Quais categorias dos fatores condicionadores ela contém. A codificação usada está descrita no Quadro 1 abaixo:

CONCORDÂNCIA VERBAL = CV

Variantes: padrão = **1**; não padrão = **2**

Fatores:

- Posição e distância do sujeito em relação ao verbo

Categorias:

- ✓ Sujeito preposto distante do verbo = **A**
- ✓ Sujeito preposto imediatamente ao verbo = **B**

✓ Sujeito posposto ao verbo = **C**

- Animacidade do sujeito

Categorias:

- ✓ Sujeito animado = **D**
- ✓ Sujeito inanimado = **E**

- Constituição do sujeito

Categorias:

- ✓ Sujeito simples (pronomes ou SN simples) = **F**
- ✓ Sujeito composto = **G**

RETOMADA POR OBJETO DIRETO ANAFÓRICO = OD

Variantes: clítico acusativo (padrão) = **1**; pronome lexical (não padrão) = **2**; categoria vazia (não padrão) = **3**

Fatores:

- Estrutura projetada pelo verbo

Categorias:

- ✓ Estrutura simples = **H**
- ✓ Estrutura com predicativo = **I**
- ✓ Estrutura complexa = **J**

- Animacidade do objeto

- ✓ Objeto animado = **L**
- ✓ Objeto inanimado = **M**

ESTRATÉGIAS DE RELATIVIZAÇÃO = ER

Variantes: padrão = **1**; cortadora (não padrão) = **2**; copiadora (não padrão) = **3**

Fatores:

- Função sintática do pronome relativo

- ✓ Objeto indireto = **N**
- ✓ Adjunto adverbial = **O**
- ✓ Adjunto adnominal (genitivo) = **P**

- Animacidade do antecedente

- ✓ Antecedente animado = **Q**
- ✓ Antecedente inanimado = **R**

Por exemplo: há uma questão com o código CV-1-ADF, o que indica que ela traz um exemplo da variável "concordância verbal" (CV) em sua variante *padrão* (1), que está testando o fator “posição e distância do sujeito em relação ao verbo” em um contexto de *sujeito preposto* distante do verbo (A), o fator “animacidade do sujeito” em um contexto de sujeito [+ *animado*] (D) e o fator “constituição do sujeito” em um contexto de *sujeito simples* (F).

Começamos pelas questões que testaram a percepção sobre a variável "concordância verbal". Aplicando-se a propriedade distributiva da multiplicação às diversas categorias de cada um dos *fatores* dessa variável, e levando-se em consideração o fato de que ela tem duas *variantes*, temos os seguintes códigos de questão:

| | | | |
|------|--------------|--------------|--------------|
| (11) | 01. CV-1-ADF | 09. CV-1-BDF | 17. CV-1-CDF |
| | 02. CV-2-ADF | 10. CV-2-BDF | 18. CV-2-CDF |
| | 03. CV-1-ADG | 11. CV-1-BDG | 19. CV-1-CDG |
| | 04. CV-2-ADG | 12. CV-2-BDG | 20. CV-2-CDG |
| | 05. CV-1-AEF | 13. CV-1-BEF | 21. CV-1-CEF |
| | 06. CV-2-AEF | 14. CV-2-BEF | 22. CV-2-CEF |
| | 07. CV-1-AEG | 15. CV-1-BEG | 23. CV-1-CEG |
| | 08. CV-2-AEG | 16. CV-2-BEG | 24. CV-2-CEG |

As orações elaboradas estão nos quadros seguintes. No Quadro 2, a seguir, apresentamos as questões relativas aos 8 primeiros códigos, que fazem parte do grupo de orações com a categoria “sujeito preposto distante do verbo” (A). Resolvemos dividir a apresentação das orações em blocos para facilitar a posterior explicação de detalhes da elaboração do questionário. O sujeito e o verbo estão destacados com itálico e sublinhado, respectivamente.

| | |
|----------|--|
| CV-1-ADF | <i>Esses políticos</i> aqui no Brasil não <u>entendem</u> que quem manda é o povo |
| CV-2-ADF | <i>Eles</i> sem nenhum pudor <u>leva</u> materiais de escritório para casa |
| CV-1-ADG | <i>Os clientes e os vendedores</i> neste nosso contexto <u>precisam</u> falar a mesma língua |
| CV-2-ADG | <i>Os chefes e os funcionários</i> por incrível que pareça se <u>comunica</u> muito bem por aqui |
| CV-1-AEF | Sobre os malotes, <i>eles</i> em pouco tempo <u>chegam</u> para o cliente |
| CV-2-AEF | <i>As impressoras</i> na pior hora sempre <u>para</u> de funcionar |
| CV-1-AEG | <i>O caminhão e a van</i> quando eu dirijo não <u>levam</u> nenhuma multa |
| CV-2-AEG | <i>O computador e a impressora</i> pelo que vi ontem <u>precisa</u> de conserto |

Quadro 2: Questões da variável "concordância verbal", bloco 1 (categoria "sujeito preposto distante do verbo")

As orações que se referem à categoria “sujeito preposto distante do verbo” foram elaboradas com pelo menos duas palavras tônicas e não monossilábicas separando o sujeito do verbo. Em geral, optamos por inserir adjuntos adverbiais entre sujeito e verbo, por ser essa função sintática a mais maleável e natural fora de seu lugar original. Os tempos verbais foram mantidos no presente do indicativo, pois, apesar de o fator “saliência fônica” (cf. 2.1.2) não estar sendo testado, utilizar verbos com diferentes graus de saliência entre singular e plural poderia influenciar o resultado da análise.

O Quadro 3, abaixo, traz as orações que apresentam a categoria “sujeito preposto imediatamente ao verbo” (B):

| | |
|----------|--|
| CV-1-BDF | <i>Meus amigos <u>conhecem</u> esse novo estoquista</i> |
| CV-2-BDF | <i>Esses atendentes <u>acredita</u> mesmo nessa abordagem?</i> |
| CV-1-BDG | <i>Os familiares e amigos <u>podem</u> usar as instalações</i> |
| CV-2-BDG | <i>Os colegas e clientes <u>considera</u> sua saída uma pena</i> |
| CV-1-BEF | <i>Gosto desses novos produtos, mas <i>eles</i> <u>precisam</u> de uma boa publicidade</i> |
| CV-2-BEF | <i>As peças <u>chega</u> em fevereiro</i> |
| CV-1-BEG | <i>Aposto que <i>o carro e a moto</i> <u>pertencem</u> ao chefe</i> |
| CV-2-BEG | <i>A pista e o depoimento <u>completa</u> a solução do problema</i> |

Quadro 3: Questões da variável "concordância verbal", bloco 2 (categoria "sujeito preposto imediatamente ao verbo")

As mesmas observações que já fizemos para as orações do bloco 1 se aplicam a essas do bloco 2. As questões do bloco 3, que traz orações com a categoria “sujeito posposto ao verbo” (C), estão no Quadro 4:

| | |
|----------|--|
| CV-1-CDF | <i><u>Nascem</u> muitos bebês nessa época do ano</i> |
| CV-2-CDF | <i>Todos os dias <u>chega</u> vários compradores de São Paulo</i> |
| CV-1-CDG | <i><u>Existem</u> um gerente e um contador exatamente para isso</i> |
| CV-2-CDG | <i>Se alguém fica sabendo desse erro, <u>morre</u> eu e você</i> |
| CV-1-CEF | <i><u>Existem</u> muitos problemas com essa nova abordagem</i> |
| CV-2-CEF | <i>Antigamente <u>saía</u> várias vendas por dia. Hoje está mais difícil</i> |
| CV-1-CEG | <i>É possível que <u>cresçam</u> o interesse e a procura por nossos produtos</i> |
| CV-2-CEG | <i><u>Causou-me</u> estranheza o acidente e a morte numa rua tão calma</i> |

Quadro 4: Questões da variável "concordância verbal", bloco 3 (categoria "sujeito posposto ao verbo")

No caso das orações com a categoria “sujeito posposto ao verbo”, utilizamos apenas os verbos conhecidos como **inacusativos** (cf. MIOTO, SILVA e LOPES, 2005; DUARTE, 2007). Esses verbos, segundo a Gramática Gerativa, têm seu único argumento gerado na posição interna, isto é, à direita do predicador verbal; no entanto, esse argumento não é marcado pelo caso acusativo, recebendo assim o caso nominativo, o que o licencia como sujeito da oração. Por esse motivo, sujeitos ligados a verbos inacusativos têm maior maleabilidade, podendo ser prepostos e pospostos com facilidade, o que não acontece com verbos intransitivos ou transitivos²⁷ (cf. CIRÍACO e CANÇADO, 2004).

A variável retomada anafórica por objeto direto (**OD**) tem 3 variantes (*clítico acusativo* (1), *pronomes lexical* (2) e *categoria vazia* (3)) e escolhemos dois fatores para serem testados: "estrutura projetada pelo verbo", com 3 categorias (*estrutura simples* (H), *estrutura com predicativo* (I) e *estrutura complexa* (J)) e "animacidade do sujeito", com 2 categorias (*sujeito animado* (L) e *sujeito inanimado* (M)). Sendo assim, temos $3*3*2 = 18$ perfis de oração:

| | | | |
|------|-------------|-------------|-------------|
| (12) | 01. OD-1-HL | 07. OD-1-IL | 13. OD-1-JL |
| | 02. OD-2-HL | 08. OD-2-IL | 14. OD-2-JL |
| | 03. OD-3-HL | 09. OD-3-IL | 15. OD-3-JL |
| | 04. OD-1-HM | 10. OD-1-IM | 16. OD-1-JM |
| | 05. OD-2-HM | 11. OD-2-IM | 17. OD-2-JM |
| | 06. OD-3-HM | 12. OD-3-IM | 18. OD-3-JM |

O Quadro 5, a seguir, traz as orações elaboradas para esta variável. O objeto direto anafórico está destacado por meio de sublinhado e seu antecedente, por meio de itálico:

| | |
|---------|---|
| OD-1-HL | Quando <i>a secretária</i> voltou do almoço, eu <u>a</u> cumprimentei pelo bom trabalho |
| OD-2-HL | <i>Os clientes</i> chegaram e levei <u>eles</u> até a sala de espera |
| OD-3-HL | Onde está <i>o office-boy</i> ? Faz muito tempo que não ___ vejo. |
| OD-1-HM | Comprei <i>os materiais</i> e já <u>os</u> deixei no depósito |
| OD-2-HM | Sempre que <i>uma cadeira</i> quebra, eu jogo <u>ela</u> fora |
| OD-3-HM | <i>Um dos celulares</i> ainda funciona. Quer que eu deixe ___ aqui com você? |
| OD-1-IL | <i>O estagiário</i> andava muito distraído, mas eu <u>o</u> deixei em estado de alerta |
| OD-2-IL | Acho <i>a Maria</i> uma pessoa difícil, mas o João acha <u>ela</u> simpática |

²⁷ Confira como exemplos as seguintes orações com verbos inacusativos: “Apareceu uma flor no jardim”, “Caiu um menino do balanço”, “Nasceu um bebê”, em comparação com orações com verbos intransitivos: “?Falou um menino”, “?Andou uma criança” e transitivos diretos: “?Comeu o menino o bolo”, “?Salvou o bombeiro um gato”.

| | |
|---------|---|
| OD-3-IL | Conheço <i>o Carlos</i> , mas não considero ___ um bom administrador |
| OD-1-IM | Mostramos <i>os três currículos</i> para o chefe, mas nenhum <u>o</u> deixou satisfeito |
| OD-2-IM | <i>Esta sala</i> costumava ser horrível, mas a Sílvia tornou <u>ela</u> agradável |
| OD-3-IM | A <i>nova campanha</i> não me agradou muito, mas o cliente achou ___ ótima |
| OD-1-JL | O <i>novo atendente</i> já está bem treinado. Eu já <u>o</u> deixo fazer tudo sozinho |
| OD-2-JL | Eu não confio <i>nesses políticos</i> porque eu já vi <u>eles</u> falando mal do povo |
| OD-3-JL | O chefe brigou com <i>o office-boy</i> e mandou ___ sair da sala |
| OD-1-JM | Não sei para que serve <i>essa máquina</i> , eu nunca <u>a</u> vi funcionando |
| OD-2-JM | A <i>impressora</i> estava travando, mas eu fiz <u>ela</u> funcionar de novo |
| OD-3-JM | <i>Essa sua cadeira</i> precisa de óleo! Você não ouviu ___ rangendo? |

Quadro 5: Questões da variável "retomada anafórica por objeto direto"

Para os objetos diretos [+ animado], foram usados apenas sintagmas nominais contendo referências a seres humanos. Para as orações com estrutura sintática simples, utilizamos apenas objetos diretos constituídos de SN simples, e não orações.

A variável estratégias de relativização (**ER**) tem três variantes (*padrão (1)*, *cortadora (2)* e *copiadora (3)*) e escolhemos para testá-la dois fatores: função sintática do pronome relativo, com três categorias (*objeto indireto (N)*, *adjunto adverbial (O)* e *adjunto adnominal genitivo (P)*), e animacidade do antecedente, com duas categorias (*antecedente animado (Q)* e *antecedente inanimado (R)*). Sendo assim, tivemos de elaborar um total de $3*3*2 = 18$ perfis de oração:

- | | | | |
|------|-------------|-------------|-------------|
| (13) | 01. ER-1-NQ | 07. ER-1-OQ | 13. ER-1-PQ |
| | 02. ER-2-NQ | 08. ER-2-OQ | 14. ER-2-PQ |
| | 03. ER-3-NQ | 09. ER-3-OQ | 15. ER-3-PQ |
| | 04. ER-1-NR | 10. ER-1-OR | 16. ER-1-PR |
| | 05. ER-2-NR | 11. ER-2-OR | 17. ER-2-PR |
| | 06. ER-3-NR | 12. ER-3-OR | 18. ER-3-PR |

O Quadro 6, a seguir, traz as orações elaboradas para a variável estratégias de relativização. O pronome relativo e a preposição estão destacados por meio de sublinhado e seu antecedente (assim como a cópia, no caso das copiadoras), por meio de itálico:

| | |
|---------|---|
| ER-1-NQ | <i>O distribuidor</i> <u>em quem</u> eu confiei está nos passando para trás |
| ER-2-NQ | O Messias é <i>o vendedor</i> <u>que</u> eu mais gosto ___! |

| | |
|---------|---|
| ER-3-NQ | <i>O diretor <u>que</u> eu discordei <u>dele</u> estava muito seguro de sua opinião</i> |
| ER-1-NR | <i>As regras <u>com as quais</u> eu concordei não são essas</i> |
| ER-2-NR | <i>Sinto muito, mas esse não é o modelo <u>que</u> eu gosto ____</i> |
| ER-3-NR | <i>Amanhã já é a reunião <u>que</u> você precisa participar <u>dela</u></i> |
| ER-1-OQ | <i>O rapaz <u>com que</u> eu falei é loiro e meio gordinho</i> |
| ER-2-OQ | <i>O nome do cliente <u>que</u> eu falei ____ é Marcelo</i> |
| ER-3-OQ | <i>A menina <u>que</u> nós estudamos <u>com ela</u> já está fazendo doutorado</i> |
| ER-1-OR | <i>O dia <u>em que</u> você chegar na hora será um dia chuvoso</i> |
| ER-2-OR | <i>A cidade <u>que</u> eu nasci ____ é a 100 km daqui</i> |
| ER-3-OR | <i>Esse não é o carro <u>que</u> eu vim ontem <u>com ele</u></i> |
| ER-1-PQ | <i>O cliente <u>cuja</u> compra eu fechei gastou mais de 10 mil reais!</i> |
| ER-2-PQ | <i>Os dois candidatos <u>que</u> eu olhei os currículos ____ são ótimos</i> |
| ER-3-PQ | <i>Os diretores <u>que</u> a reunião <u>deles</u> foi cancelada já tinham embarcado</i> |
| ER-1-PR | <i>Aquele é o telefone <u>cuja</u> linha está com defeito</i> |
| ER-2-PR | <i>O caminhão <u>que</u> eu consertei a roda ____ já estragou de novo</i> |
| ER-3-PR | <i>Os produtos <u>que</u> a validade <u>deles</u> já venceu foram jogados fora</i> |

Quadro 6: Questões da variável "estratégias de relativização"

Em todas as orações para essa variável, utilizamos estruturas com verbos de ligação e com a oração relativa subordinada ao núcleo do sujeito ou do predicativo do sujeito. Usamos na maioria das sentenças o pronome *que*, por ser o mais frequente relativo, ao lado de um uso de *quem* e de *as quais* e de dois usos de *cujo* para as orações relativas padrão da função de adjunto adnominal genitivo. Para as relativas de adjunto adverbial, utilizamos a preposição *com* para as que têm referente [+ animado].

Ainda para as relativas, resolvemos incluir duas questões adicionais nas quais a preposição está em seu lugar natural, após o verbo, e longe do relativo. Esse tipo de oração, para as quais se diz que a preposição é "órfã", é comum na língua inglesa ("That's *the girl that* I was talking about"), mas não faz parte da gramática padrão da língua portuguesa. Apesar disso, temos notado informalmente que há alguns contextos, sobretudo aqueles que envolvem preposição "fortes" (cf. 2.1.2) e adjuntos adverbiais, em que esse tipo de construção está passando despercebido por falantes cultos. Por esse motivo, adicionamos as seguintes orações ao questionário:

| | |
|---------|--|
| ER-QE-1 | Esse é o livro <u>que</u> eu te falei <u>sobre</u> |
| ER-QE-2 | Foi <i>essa ideia</i> <u>que</u> nós votamos <u>contra</u> |

Quadro 7: Questões extras da variável "estratégias de relativização"

2.2.2 Se a Linguística Estuda a Fala, por que Testes Escritos?

Dissemos em 1.1 que a Linguística difere dos estudos gramaticais tradicionais por, dentre outros motivos, dar muito mais importância à língua falada que à língua escrita. Isso é verdade para a maioria das subdisciplinas dessa ciência, e sobretudo para a Sociolinguística, que tem por objeto de estudo os usos reais da língua em situações de interação face a face. Sendo assim, por que optamos por utilizar testes escritos nesta pesquisa?

Em primeiro lugar porque acreditamos, como foi referido em 1.1.3, que a variação linguística pode ser apreendida em contínuos, tais como os teorizados por Bortoni-Ricardo (2004). As regras da norma padrão influenciam quaisquer dos pontos desse contínuo: no contínuo rural-urbano, quanto mais próximo da extremidade "urbana", maior a influência dos expedientes normatizadores, mas isso não quer dizer que haja a total ausência desses expedientes fora dessa extremidade. Exatamente por ser um contínuo, a influência só vai diminuindo à medida que nos aproximamos da extremidade rural. Seguindo a terminologia da própria autora, as regras normativas são *traços graduais*, pois ocorrem em todo o contínuo das variedades, em maior ou menor frequência. O mesmo acontece com o contínuo oralidade-letramento: em um dos extremos estão os eventos mediados pela língua escrita, e do outro os eventos de oralidade, em que não há influência da língua escrita. Mas não há fronteiras bem marcadas: elas são fluidas e há muitas sobreposições. E também nesse contínuo o conjunto de prescrições normativas não se situa em apenas um dos extremos: ele perpassa todo o contínuo, e pode portanto ser detectado e medido em qualquer ponto desse contínuo.

Em segundo lugar, há muitos trabalhos de Sociolinguística que analisam dados escritos. Todos aqueles que utilizam testes de percepção ou produção (cf. TARALLO, 1986, p. 55) empregam de algum modo questionários escritos para avaliar a percepção e a atitude dos informantes. Outras pesquisas, sobretudo as de cunho diacrônico, não têm outro modo de apreender a variação na história da língua a não ser recorrendo a gêneros textuais escritos (cf., por exemplo, TARALLO, 1983). E, mesmo em estudos sincrônicos, é comum comparar *corpora* de língua falada com *corpora* escritos (cf., por exemplo, BAGNO, 2000) ou até mesmo

utilizar apenas dados escritos, como redações de alunos (cf., por exemplo, PEREIRA e COELHO, 2013) para estudar a variação linguística.

Dessa forma, acreditamos que o uso de dados advindos da língua escrita não representa uma contradição aos postulados da Sociolinguística, nem diminui a qualidade da pesquisa ou desautoriza os resultados obtidos, sobretudo no nosso caso em que o que se busca é a percepção dos informantes em relação a entidades tão abstratas quanto as regras da norma padrão.

2.3 Informantes e Dados

Nesta seção, descreveremos brevemente os métodos utilizados para a escolha dos informantes, a aplicação do teste de percepção e a tabulação e tratamento dos dados coletados.

2.3.1 A Escolha dos Informantes

Em geral, as pesquisas sociolinguísticas selecionam uma amostra da comunidade em estudo por meio do método *aleatório estratificado*: "divide-se a população em 'células' ('casas', 'estratos') compostas, cada uma, de indivíduos com as mesmas características sociais, procedendo-se posteriormente, para preencher cada casa, a uma seleção aleatória" (SILVA, 2013, p. 121). A quantidade de células da amostra depende, assim, da quantidade de variáveis sociais escolhidas: se escolhêssemos sexo/ gênero (duas variantes: "masculino" e "feminino"), dois níveis de escolaridade ("ensino fundamental completo" e "ensino médio completo", por exemplo) e duas faixas etárias ("entre 18 e 34 anos", e "entre 35 e 59 anos", por exemplo), teríamos 8 células, pois $2*2*2 = 8$.

A nossa amostra de falantes do PB foi composta apenas por falantes cultos, segundo a definição tradicional dada pela Sociolinguística (cf. 1.2.1); logo, não usamos a variável social "escolaridade" como fator para compor a amostra, uma vez que todos os informantes devem ter ao menos ensino superior completo. Nossa amostra foi estratificada apenas por sexo/ gênero ("masculino", "feminino")²⁸ e faixa etária ("de 20 a 34 anos" e "de 35 a 59 anos")²⁹, o que equivale a um total de 4 células sociais.

²⁸ "A menos que os propósitos de uma certa pesquisa variacionista fossem especificamente direcionados a uma comunidade de práticas constituída por indivíduos de apenas um dos dois sexos, não faria sentido que uma amostra não incluísse sujeitos-informantes de ambos" (MENDES, 2011, p. 10).

²⁹ Seguimos, nessa variável, as ideias de Mendes (2011, p. 11), cuja classificação "toma por base, principalmente, a posição relativa das faixas etárias no mercado de trabalho, mas também leva em conta modos/estilos de vida de cada grupo, caracterizados aqui *in abstracto*. O grupo de falantes mais novos, aqueles entre 20 e 34 anos de idade, engloba jovens adultos que, em geral, encontram-se relativamente menos estabilizados [...].O grupo de falantes

Para cada célula social, estipulamos o número de 10 informantes. Sendo assim, tivemos um total de 40 informantes, assim distribuídos:

Tabela 1: Células sociais e número de informantes

| | Entre 20 e 34 anos de idade | Entre 35 e 59 anos de idade |
|------------------------|--|--|
| Sexo/ gênero masculino | 10 | 10 |
| Sexo/ gênero feminino | 10 | 10 |

Uma vez que nossa pesquisa não se restringe a nenhuma variedade específica do português brasileiro, não foi exigido que o informante tivesse nascido ou morado em alguma região determinada do país, mas ele devia ter passado a maior parte de sua infância e adolescência em uma região urbana. Dos 40 informantes para os quais o teste de percepção foi aplicado, 32 viveram a maior parte da infância/ adolescência na região Centro-Oeste, 5 na região Sudeste, 2 na região Sul e 1 na região Nordeste. Dessa forma, podemos dizer que esta pesquisa é representativa apenas da percepção de falantes cultos do centro-sul brasileiro, com destaque para as variedades do Centro-Oeste.

Essa estratificação em fatores *sociais* foi pensada apenas para fins de organização dos informantes. A análise levada a cabo no Capítulo 3 foca primordialmente nos fatores de natureza *linguística*. Faremos uma breve consideração sobre a influência dos fatores sociais na percepção da variação pelos falantes cultos, mas esta será uma análise subsidiária.

Os informantes foram selecionados e contatados pelo método "amigo de amigo": o próprio pesquisador aplicou o questionário a pessoas próximas e pediu a indicação de *amigos* que pudessem se encaixar em um dos perfis sociais. Optamos por não incluir nenhum egresso de cursos de Letras como informante, pois é muito provável que esses falantes tenham tido contato com pesquisas sociolinguísticas que tratam dos fenômenos aqui explorados, e isso poderia influenciar suas respostas. Também tentamos balancear a amostra, na medida do possível, tanto com informantes cuja formação superior tenha sido na área de ciências humanas quanto com aqueles com formação em ciências exatas e naturais. Ao final, conseguimos apenas 15 informantes advindos de áreas das ciências exatas e naturais.

entre 35 e 59 anos, por sua vez, pretende abarcar pessoas mais plenamente inseridas no mercado de trabalho e relativamente mais estáveis (ou que buscam mais seriamente uma certa estabilidade)."

2.3.2 Aplicação dos Questionários e Tratamento dos Dados

Para a aplicação do questionário, resolvemos utilizar a plataforma virtual gratuita Google Docs, por meio da qual é possível criar formulários virtuais com questões de preenchimento textual, múltipla escolha, etc.³⁰ (cf. Figura 1, pág. 79). As 62 questões cujo processo de criação foi explicado na seção 2.2.1 foram embaralhadas utilizando-se um programa "randomizador" *on-line*³¹ (cf. Figura 2, pág. 79). As orações foram copiadas para o programa, uma por linha, e randomizadas pelo processo de "ruído atmosférico" (cf. Figura 3, pág. 80). O questionário foi precedido, no formulário, pela introdução (cf. 2.2) e sucedido por outro formulário onde o informante deveria colocar seus dados.

Após contatar o informante, enviamos o *link* do questionário por *email*, para que ele pudesse responder como e quando quisesse. Pedimos que, durante as respostas, o informante não utilizasse fontes externas de conhecimento gramatical (como gramáticas normativas, guias de redação, *sites* de tira-dúvidas, etc.) e que tentasse responder ao questionário completo de uma só vez. Não fizemos questão de omitir do informante a informação de que se tratava de uma pesquisa sobre língua e variação linguística, como é de praxe nas coletas de dados em Sociolinguística. O próprio teor do questionário, que pedia que frases fossem analisadas por sua correção linguística, já deixava isso claro.

Ao final do questionário, solicitamos que os informantes assinalassem uma declaração de ciência de que seus dados seriam utilizados em uma pesquisa de pós-graduação (cf. Figura 4, p. 80). Não utilizamos termos de consentimento livre e esclarecido mais formais por orientação do Conselho de Ética da UEMS (comunicação pessoal).

³⁰ Acessível pelo endereço <https://docs.google.com/forms>.

³¹ Acessível pelo endereço <https://www.random.org/lists/>.

Projeto de pesquisa "Graus de
saliência das variáveis e variantes
linguísticas" - UEMS, 2016

Lembre-se das opções:

Não há erro: a frase está bem construída. O produtor da frase provavelmente é bem escolarizado.

Errada, mas aceitável: a frase contém um erro, mas é um erro pouco grave na sua opinião. O produtor da frase provavelmente é bem escolarizado, mas cometeu um lapso que muitos cometem.

Inaceitável: a frase contém um erro grave. O produtor da frase provavelmente tem escolaridade baixa ou mediana.

As frases são as seguintes:

Foi essa ideia que nós votamos contra

Não há erro

Errada, mas aceitável

Inaceitável

Comprei os materiais e já os deixei no depósito

Figura 1: Questionário no Google Docs

Home Games Numbers Lists & More Drawings Web Tools Statistics Testimonials Learn More Login

RANDOM.ORG

Search RANDOM.ORG
Google Custom Search Search

True Random Number Service

Do you own an iOS or Android device? [Check out our app!](#)

List Randomizer

This form allows you to arrange the items of a list in random order. The randomness comes from atmospheric noise, which for many purposes is better than the pseudo-random number algorithms typically used in computer programs.

Part 1: Enter List Items

Enter your items in the field below, each on a separate line. Items can be numbers, names, email addresses, etc. A maximum of 10,000 items are allowed. Please don't enter anything you would consider confidential ([here's why](#)).

```

passada
O chefe deixou essas planilhas para eu revisar até amanhã
Comprei esses dois livros para mim estudar o assunto
Nessa loja, vendem-se vários equipamentos de última ponta
Cria-se muitas novas leis inúteis todos os dias!
Aqui não se ouvem muitas histórias como essa
Nunca se viu tantos modelos novos assim no mercado
A gente recebeu essa remessa mês passado
A gente elaboramos um novo método de entrega mais
eficiente

```

(you're viewing this form securely)

Part 2: Go!

Be patient! It may take a little while to randomize your list...

Figura 2: "Randomizador" on-line

Home Games Numbers Lists & More Drawings Web Tools Statistics Testimonials Learn More Login

RANDOM.ORG

Search RANDOM.ORG
Google™ Custom Search Search

True Random Number Service

Do you own an iOS or Android device? [Check out our app!](#)

List Randomizer

There were 80 items in your list. Here they are in random order:

1. Foi essa ideia que nós votamos contra
2. Comprei os materiais e já os deixei no depósito
3. Já fazem 2 dias que eu liguei reclamando da demora
4. Nasceram muitos bebês nessa época do ano
5. Quando a secretária voltou do almoço, eu a cumprimentei pelo bom trabalho
6. As peças de reposição chega em fevereiro
7. Conheço o Carlos, mas não considero um bom administrador
8. O caminhão e a van quando eu dirijo não levam nenhuma multa
9. O diretor que eu discordei dele estava muito seguro de sua opinião
10. O Messias é o vendedor que eu mais gosto!
11. Já revisei os relatórios e não havia erros
12. Aquele é o telefone cuja linha está com defeito
13. Os clientes chegaram e levei eles até a sala de espera
14. O chefe deixou essas planilhas para eu revisar até amanhã
15. Um dos celulares ainda funciona. Quer que eu deixe aqui com você?
16. O rapaz de que eu falei é loiro e meio gordinho
17. Acho a Maria uma pessoa difícil, mas o João acha ela simpática
18. Os dois candidatos que eu olhei os currículos são ótimos
19. Esses políticos aqui no Brasil não entendem que quem manda é o povo
20. Mostramos os três currículos para o chefe, mas nenhum o deixou satisfeito
21. As regras com as quais eu concordei não são essas
22. Faz 5 anos que trabalho nesse setor e ainda estou aprendendo!
23. Nós chegamos no aeroporto atrasados e perdemos o voo
24. O cliente cuja compra eu fechei gastou mais de 10 mil reais!
25. Se alguém fica sabendo desse erro, morre eu e você
26. Criei muitas novas leis antes de sair do dia!

Figura 3: Questões já embaralhadas

Entre 20 e 34 anos
 Entre 35 e 59 anos

Em que cidade você viveu a maior parte da sua infância e adolescência?

Your answer _____

Estou ciente de que minhas respostas a este questionário serão usadas na pesquisa "Graus de saliência das variáveis e variantes linguísticas" do acadêmico Luiz Leandro Gomes de Lima (luiz.gomes@ufms.br) do programa de pós-graduação em Letras da UEMS e que meus dados pessoais não serão tornados públicos na divulgação dos resultados da pesquisa.

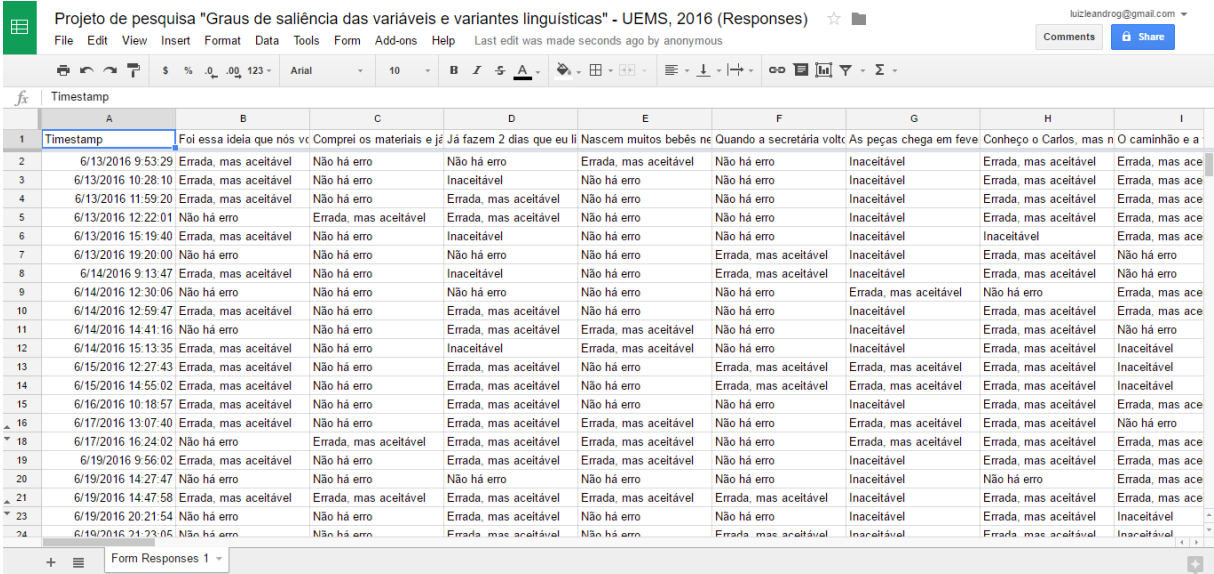
SIM, estou ciente
 NÃO, desisto de participar

Never submit passwords through Google Forms.

Figura 4: Declaração de ciência

As respostas foram reunidas pelo próprio sistema do Google Docs em uma planilha de cálculos, a partir da qual nos foi possível manipular esses dados e realizar os cálculos que apresentamos no capítulo seguinte (cf. Figura 5). Uma vez que trabalhamos com conceitos estatísticos simples, não foi necessário utilizar nenhuma ferramenta computacional mais sofisticada, como os pacotes de programas estatísticos comumente usados em pesquisas sociolinguísticas (como o Varbrul ou, mais recentemente, o RBrul). Um editor de planilhas simples foi suficiente para realizar os levantamentos de que precisávamos para alcançar os objetivos desta pesquisa, calcular as porcentagens que refletem as percepções e atitudes dos falantes cultos e traduzir tudo isso em gráficos e tabelas, conforme veremos no Capítulo 3.

O número de opiniões para cada questão foi calculado por meio da função CONT.SE aplicada na fórmula padrão =CONT.SE(X:Y;"valor"), em que X e Y equivalem ao intervalo de células onde estão as opiniões sobre a questão em análise e "valor" equivale a cada uma das opiniões possíveis ("Não há erro", "Errada, mas aceitável" e "Inaceitável"). Após encontrarmos os números de cada questão, foram gerados gráficos de pizza que, automaticamente, calculam as porcentagens equivalentes.



The screenshot shows a Google Docs spreadsheet with the following data:

| Timestamp | A | B | C | D | E | F | G | H | I |
|--------------------|---------------------------|---------------------------|---------------------------|------------------------|--------------------------|------------------------|-------------------------|-----------------|---|
| Timestamp | Foi essa ideia que nós vc | Comprei os materiais e já | Já fazem 2 dias que eu li | Nascem muitos bebês ne | Quando a secretária volt | As peças chega em feve | Conheço o Carlos, mas n | O caminhão e a | |
| 6/13/2016 9:53:29 | Errada, mas aceitável | Não há erro | Não há erro | Errada, mas aceitável | Não há erro | Inaceitável | Errada, mas aceitável | Errada, mas ace | |
| 6/13/2016 10:28:10 | Errada, mas aceitável | Não há erro | Inaceitável | Não há erro | Não há erro | Inaceitável | Errada, mas aceitável | Errada, mas ace | |
| 6/13/2016 11:59:20 | Errada, mas aceitável | Não há erro | Errada, mas aceitável | Não há erro | Não há erro | Inaceitável | Errada, mas aceitável | Errada, mas ace | |
| 6/13/2016 12:22:01 | Não há erro | Errada, mas aceitável | Errada, mas aceitável | Não há erro | Não há erro | Inaceitável | Errada, mas aceitável | Errada, mas ace | |
| 6/13/2016 15:19:40 | Errada, mas aceitável | Não há erro | Inaceitável | Não há erro | Não há erro | Inaceitável | Inaceitável | Errada, mas ace | |
| 6/13/2016 19:20:00 | Não há erro | Não há erro | Não há erro | Não há erro | Errada, mas aceitável | Inaceitável | Errada, mas aceitável | Não há erro | |
| 6/14/2016 9:13:47 | Errada, mas aceitável | Não há erro | Inaceitável | Não há erro | Errada, mas aceitável | Inaceitável | Errada, mas aceitável | Não há erro | |
| 6/14/2016 12:30:06 | Não há erro | Não há erro | Não há erro | Não há erro | Não há erro | Errada, mas aceitável | Não há erro | Errada, mas ace | |
| 6/14/2016 12:59:47 | Errada, mas aceitável | Não há erro | Errada, mas aceitável | Não há erro | Não há erro | Inaceitável | Errada, mas aceitável | Errada, mas ace | |
| 6/14/2016 14:41:16 | Não há erro | Não há erro | Errada, mas aceitável | Errada, mas aceitável | Não há erro | Inaceitável | Errada, mas aceitável | Não há erro | |
| 6/14/2016 15:13:35 | Errada, mas aceitável | Não há erro | Inaceitável | Errada, mas aceitável | Não há erro | Inaceitável | Errada, mas aceitável | Inaceitável | |
| 6/15/2016 12:27:43 | Errada, mas aceitável | Não há erro | Errada, mas aceitável | Não há erro | Errada, mas aceitável | Errada, mas aceitável | Errada, mas aceitável | Inaceitável | |
| 6/15/2016 14:55:02 | Errada, mas aceitável | Não há erro | Errada, mas aceitável | Não há erro | Errada, mas aceitável | Errada, mas aceitável | Errada, mas aceitável | Inaceitável | |
| 6/16/2016 10:18:57 | Errada, mas aceitável | Não há erro | Errada, mas aceitável | Não há erro | Não há erro | Inaceitável | Errada, mas aceitável | Errada, mas ace | |
| 6/17/2016 13:07:40 | Errada, mas aceitável | Não há erro | Errada, mas aceitável | Errada, mas aceitável | Não há erro | Errada, mas aceitável | Errada, mas aceitável | Não há erro | |
| 6/17/2016 16:24:02 | Não há erro | Errada, mas aceitável | Errada, mas aceitável | Errada, mas aceitável | Não há erro | Errada, mas aceitável | Errada, mas aceitável | Errada, mas ace | |
| 6/19/2016 9:56:02 | Errada, mas aceitável | Não há erro | Errada, mas aceitável | Errada, mas aceitável | Não há erro | Inaceitável | Errada, mas aceitável | Errada, mas ace | |
| 6/19/2016 14:27:47 | Não há erro | Não há erro | Não há erro | Não há erro | Não há erro | Inaceitável | Não há erro | Errada, mas ace | |
| 6/19/2016 14:47:58 | Errada, mas aceitável | Errada, mas aceitável | Errada, mas aceitável | Errada, mas aceitável | Errada, mas aceitável | Inaceitável | Errada, mas aceitável | Errada, mas ace | |
| 6/19/2016 20:21:54 | Não há erro | Não há erro | Errada, mas aceitável | Não há erro | Não há erro | Inaceitável | Errada, mas aceitável | Inaceitável | |
| 6/19/2016 21:23:05 | Não há erro | Não há erro | Errada, mas aceitável | Não há erro | Errada, mas aceitável | Inaceitável | Errada, mas aceitável | Inaceitável | |

Figura 5: Planilha do Google Docs

2.3.3 Alguns problemas

Durante e depois da aplicação dos testes de percepção, percebemos que algumas das decisões tomadas na fase de elaboração das questões foram equivocadas e poderiam ter um impacto negativo nos resultados obtidos. Passamos a listá-las nesta seção para, em primeiro

lugar, reconhecer que tais problemas podem trazer certas limitações aos resultados (discutidas em pormenores no capítulo seguinte e nas Considerações Finais) e, em segundo lugar, servir de orientação para pesquisas futuras que se enquadrem na mesma linha metodológica da nossa.

O primeiro problema observado foi o da quantidade de questões: por mais que o teste fosse completado em cerca de 20 minutos (muito menos tempo despendido pelo informante do que em uma entrevista sociolinguística tradicional), a quantidade de orações a serem analisadas pode ter levado alguns informantes a completarem o questionário de forma displicente, "apenas para terminar logo e ajudar o amigo". Dessa forma, a avaliação das últimas questões pode ter sofrido efeito do cansaço ou da falta de atenção dos informantes, o que pode levar a resultados distintos dos que seriam obtidos se elas estivessem dispostas, por exemplo, no início ou no meio do questionário. Uma forma de minimizar esse problema e que poderíamos ter utilizado seria o de "sentenças-controle", isto é, um conjunto de orações totalmente agramaticais aos olhos de qualquer falante que seriam dispostas em diferentes pontos do questionário para testar a atenção dos informantes, mas essa ideia só surgiu depois de o questionário já ter sido aplicado.

O segundo problema observado foi que não levamos em conta a possibilidade de os informantes *hiperanalisarem* as orações, encontrando erros onde eles não existem (um fenômeno parecido com a hipercorreção, mas manifestado na percepção em vez de na produção linguística). Por exemplo: muitas orações contêm os pronomes demonstrativos *esse* e *isso* (e suas flexões), que estão em variação no PB com *este* e *isto* (e suas flexões). Em nenhuma das orações do teste esses pronomes estão usados de forma divergente da norma padrão, mas não previmos que os informantes pudessem pensar que sim. Dessa forma, naquelas questões onde esses pronomes aparecem (exemplos: CV-1-ADF, OD-2-IM e ER-1-NR), não é possível ter certeza de se os informantes as classificaram como erradas ou corretas por conta desse fator ou da variável que gostaríamos que eles observassem.

Outro problema parecido foi detectado com as orações em que inserimos adjuntos adverbiais entre o sujeito e o verbo para testar a variável "Posição e distância do sujeito em relação ao verbo" (orações CV-1-ADF a CV-2-AEG). Esse adjunto deslocado soa pouco natural na leitura, apesar de ser bastante comum na fala. Além disso, na escrita adjuntos deslocados normalmente vêm separados do restante da oração por vírgulas, que evitamos utilizar na redação das questões. Apesar de termos pedido na introdução do teste que o informante não se atentasse a possíveis problemas de ortografia e pontuação, imaginamos que a falta de vírgulas somada à estranheza desse tipo de oração possa ter levado os informantes a hiperanalisarem também essas orações.

3 ANÁLISE DOS DADOS

A partir das respostas aos testes de percepção, apresentamos neste capítulo os resultados obtidos, tendo-se em vista os problemas de pesquisa que arrolamos como objetivos de pesquisa na Introdução a este trabalho. Separamos uma seção para discutir os *resultados gerais* (cf. 3.1), uma para cada uma das três *variáveis* escolhidas para análise (cf. 3.2, 3.3 e 3.4).

3.1 Variáveis em Conjunto

Para realizar a análise das respostas aos testes de percepção, apresentaremos gráficos que resumem os dados obtidos para todas as orações do questionário. A seguir, faremos uma série de comparações entre esses dados a fim de observar as tendências estatísticas que cada fator e cada categoria dos fatores têm na saliência das variáveis (e, conseqüentemente, na percepção dos informantes). Trabalharemos com porcentagens e, para os efeitos dessa pesquisa, um resultado "significativo" é aquele cuja diferença entre as duas grandezas comparadas é maior que 10 pontos percentuais.

Os principais dados vêm da percepção dos falantes em relação às orações **não padrão**: é a partir delas que podemos quantificar o quanto cada fator ou categoria de fator é determinante na saliência de uma variante linguística. No entanto, analisaremos também todas as orações **padrão** presentes no questionário, visto que a partir delas podemos ter alguns *insights* sobre as dificuldades que os falantes cultos possam ter em relação ao uso das variáveis estudadas.

O primeiro problema que divisamos na Introdução diz respeito à saliência das três variáveis linguísticas que analisamos, consideradas em seu conjunto: qual das três variáveis recebe mais atenção dos falantes cultos? Isto é: quais variáveis são mais ou menos salientes, considerando o nível de percepção dos informantes?

No caso das orações a partir das quais testamos a variável "**concordância verbal**" e que continham a variante padrão, os informantes marcaram 67,78% das ocorrências como "sem erro" e 32,22% como "erradas, mas aceitáveis" (23,85%) ou como "inaceitáveis" (8,37%). Atribuimos essa razoável quantidade de informantes que consideraram erradas sentenças na norma padrão ao problema que já comentamos no capítulo anterior: das 12 sentenças padrão para a variável concordância verbal presentes no questionário, 6 contêm os pronomes demonstrativos *esse/este*, que, por estarem em variação no PB, podem ter desviado o foco dos

informantes, que, por insegurança com as regras de uso desses pronomes, marcaram como erradas orações sem nenhuma inadequação para a norma padrão.

No caso das orações não padrão de concordância verbal, os informantes consideraram 85,39% das ocorrências como "erradas, mas aceitáveis" (25,68%) ou como inaceitáveis (59,71%). Sendo assim, 14,61% das ocorrências não padrão de concordância verbal **não** foram percebidas pelos falantes como inadequadas segundo a norma padrão.

Em relação às orações que testaram a variável "**objeto direto anafórico**", 83,68% das ocorrências padrão foram marcadas como "sem erro" e 16,32% como "erradas, mas aceitáveis" (14,23%) ou "inaceitáveis" (2,09%). Das 6 questões que testaram a variante padrão dessa variável, 2 contêm os pronomes *esse/ este*, o que, como já dissemos, pode ter influenciado na marcação de orações sem desvios de norma padrão como incorretas. Das ocorrências não padrão da realização anafórica do objeto direto, 70,77% foram marcadas como "erradas, mas aceitáveis" (58,66%) ou como "inaceitáveis" (12,11%) e 29,23% como "sem erro". Ou seja: quase 30% dos informantes **não** perceberam que se tratava de fugas à norma padrão.

Para a variável "**estratégias de relativização**", 71,67% das ocorrências da variante padrão foram consideradas "sem erro" pelos informantes, ao passo que 28,33% foram tachadas como "erradas, mas aceitáveis" (25%) ou "inaceitáveis" (3,33%). Para as ocorrências da variante não padrão, 69,37% foram consideradas "erradas, mas aceitáveis" (42,29%) ou "inaceitáveis", enquanto que 30,63% foram marcadas como "sem erro". Mais uma vez, cerca de 30% dos informantes não perceberam que havia fugas à norma padrão nesse grupo de sentenças.

Para responder à nossa pergunta-problema e tentar atingir o primeiro objetivo deste trabalho, vamos comparar a quantidade de sentenças não padrão das três variáveis que foram marcadas com os rótulos de "erradas, mas aceitáveis" ou "inaceitáveis". Confira a Tabela 2:

Tabela 2: Marcações das sentenças não padrão para as três variáveis

| | Concordância verbal | | Objeto direto anafórico | | Estratégias de relativização | |
|-----------------------|---------------------|--------------|-------------------------|--------------|------------------------------|--------------|
| | Nº | % | Nº | % | Nº | % |
| Não há erro | 70 | 14,58 | 140 | 29,23 | 147 | 30,63 |
| Errada, mas aceitável | 123 | 25,63 | 282 | 58,66 | 203 | 42,29 |
| Inaceitável | 287 | 59,79 | 58 | 12,11 | 130 | 27,08 |
| TOTAL | 480 | 100 | 480 | 100 | 480 | 100 |

Cada variável foi testada por meio de 12 sentenças contendo a(s) variante(s) não padrão, de forma que se multiplicarmos esse valor pela quantidade de informantes, temos um total de 480 opiniões sobre cada uma das referidas sentenças, conforme vemos na Tabela 2. Como se pode ver, os informantes perceberam 85,42% dos casos de desvio da norma padrão com a variável "concordância verbal", 70,77% das ocorrências com a variável "objeto direto anafórico" e 69,37% das ocorrências com a variável "estratégias de relativização".

O que esse resultado nos mostra? Aparentemente, a concordância verbal é a variável mais saliente das três, uma vez que foi a que teve suas sentenças não padrão mais marcadas como "inaceitáveis" ou "erradas, mas aceitáveis". As outras duas variáveis obtiveram resultados praticamente idênticos, com por volta de 70% de suas orações não padrão marcadas como tal. Apesar de a distância entre as três variáveis não ser tão discrepante (apenas cerca de 15 pontos percentuais entre a primeira colocada e as duas outras), é possível perceber que a *atitude* dos informantes é radicalmente diferente em relação à primeira variável, quando comparada às outras duas: em quase 60% dos casos, as orações não padrão com a variável concordância verbal foram consideradas *inaceitáveis*, isto é, totalmente *agramaticais*, contra apenas cerca de 12% para as sentenças com objeto direto anafórico e cerca de 27% para aquelas com estratégias de relativização. Ou seja: a aceitabilidade de "erros" de retomada anafórica ou de relativização de orações é maior, ao passo que os "erros" de concordância verbal são mais estigmatizados pelos falantes cultos.

Esse resultado parece endossar as palavras de Vieira (2007), segundo a qual

A não realização da regra de concordância verbal, no português do Brasil, constitui, sem dúvida, um traço de diferenciação social, de cunho estigmatizante, que se revela, com mais nitidez, no âmbito escolar (VIEIRA, 2007, p. 85).

As pesquisas sociolinguísticas revelam que, quanto maior a escolaridade dos falantes, maior será a aplicação da regra de concordância. De forma análoga, nossos informantes, todos com ensino superior completo, parecem também perceber mais essa variação e estigmatizar mais o não uso da concordância. A falta de concordância verbal de 3ª pessoa é um fenômeno altamente presente na fala das pessoas menos escolarizadas e mais pobres, e talvez por isso o uso dessa variante não padrão seja tão estigmatizado. Como afirma Gnerre (1991, p. 6-7), "uma variedade linguística 'vale' o que 'valem' na sociedade seus falantes, isto é, vale como reflexo do poder e da autoridade que eles têm nas relações econômicas e sociais". Se a concordância verbal é tão estigmatizada, é porque ela está na boca de indivíduos socialmente discriminados. Se compararmos, por exemplo, a rejeição gerada pelo "erro" de concordância com a rejeição

do "erro" de relativização, veremos que esta é muito maior que aquela, pois as variantes não padrão da regra de relativização é usada por todos os falantes, inclusive pelos cultos³².

Vamos analisar agora o que as variáveis externas *sexo/gênero* e *faixa etária* nos dizem sobre a percepção linguística dos falantes cultos. Abaixo, apresentamos a mesma tabela acima, mas com dados separados para os dois sexos/ gêneros que circunscrevemos nesta pesquisa:

Tabela 3: Marcação das sentenças não padrão para as três variáveis, separadas por sexo/ gênero

Tabela 3, seção 1

| | Concordância verbal | | | | Objeto direto anafórico | | | |
|-----------------------|---------------------|--------------|-----------|--------------|-------------------------|--------------|-----------|-----------|
| | Feminino | | Masculino | | Feminino | | Masculino | |
| | Nº | % | Nº | % | Nº | % | Nº | % |
| Não há erro | 34 | 14,7 | 36 | 15 | 81 | 33,75 | 60 | 25 |
| Errada, mas aceitável | 62 | 25,83 | 61 | 25,42 | 137 | 57,08 | 144 | 60 |
| Inaceitável | 144 | 60 | 143 | 59,58 | 22 | 9,17 | 36 | 15 |
| TOTAL | 240 | 100 | 240 | 100 | 240 | 100 | 240 | 100 |

Tabela 3, seção 2

| | Estratégias de relativização | | | |
|-----------------------|------------------------------|--------------|-----------|--------------|
| | Feminino | | Masculino | |
| | Nº | % | Nº | % |
| Não há erro | 74 | 30,83 | 73 | 30,42 |
| Errada, mas aceitável | 109 | 45,42 | 94 | 39,17 |
| Inaceitável | 57 | 23,75 | 73 | 30,42 |
| TOTAL | 240 | 100 | 240 | 100 |

Em relação à variável "concordância verbal", as opiniões de ambos os sexos foram praticamente simétricas. Em relação às outras duas variáveis, não houve tanta simetria, mas mesmo assim as divergências não foram tão significativas: para a variável "objeto direto anafórico", os homens perceberam mais os erros que as mulheres (75% *versus* 66,25%), ao passo que as mulheres consideraram menos sentenças "inaceitáveis" que os homens (9,17% *versus* 15%). Para a variável "estratégias de relativização", a percepção dos homens e das mulheres para a impropriedade gramatical foi praticamente idêntica (69,59% de marcações "errada" para o sexo masculino, contra 69,17% para o sexo feminino); quanto à atitude em relação a essas orações não padrão, os resultados também foram muito parecidos. Tendo em vista esses dados, podemos concluir que a variável sexo/ gênero não condiciona a percepção dos falantes cultos em relação às variáveis linguísticas analisadas.

³² Nossos agradecimentos à profa. Patrícia Graciela da Rocha por sugerir essa análise.

A tabela a seguir traz os mesmos dados das duas últimas tabelas, mas separados por faixa etária:

Tabela 4: Marcação das sentenças não padrão para as três variáveis, separadas por faixa etária

Tabela 4, seção 1

| | Concordância verbal | | | | Objeto direto anafórico | | | |
|-----------------------|----------------------------|--------------|------------|--------------|--------------------------------|--------------|------------|--------------|
| | 20-34 anos | | 35-59 anos | | 20-34 anos | | 35-59 anos | |
| | Nº | % | Nº | % | Nº | % | Nº | % |
| Não há erro | 31 | 12,92 | 39 | 16,32 | 62 | 25,83 | 79 | 32,92 |
| Errada, mas aceitável | 57 | 23,75 | 66 | 27,62 | 162 | 67,50 | 119 | 49,58 |
| Inaceitável | 152 | 63,33 | 134 | 56,07 | 16 | 6,67 | 42 | 17,50 |
| TOTAL | 240 | 100 | 240 | 100 | 240 | 100 | 240 | 100 |

Tabela 4, seção 2

| | Estratégias de relativização | | | |
|-----------------------|-------------------------------------|--------------|------------|--------------|
| | 20-34 anos | | 35-59 anos | |
| | Nº | % | Nº | % |
| Não há erro | 70 | 29,17 | 77 | 32,08 |
| Errada, mas aceitável | 107 | 44,58 | 96 | 40 |
| Inaceitável | 63 | 26,25 | 67 | 27,92 |
| TOTAL | 240 | 100 | 240 | 100 |

Em geral, os resultados também não foram muito discrepantes. Novamente, apenas no caso da variável "objeto direto anafórico" houve uma diferença significativa entre as percepções das duas faixas etárias analisadas: apesar de a percepção em relação ao erro ser muito próxima (os mais jovens marcaram as orações não padrão como "erradas" em 72,17% dos casos, contra 67,08% dos mais velhos), fica evidente que a faixa etária de 20 a 34 anos aceita mais o uso não padrão da retomada anafórica por objeto direto: 67,5% dos mais jovens marcaram as orações não padrão como "erradas, mas aceitáveis", ao passo que apenas 49,58% dos mais velhos o fizeram.

A partir dos dados apresentados, podemos avaliar que as variáveis sociais "sexo/gênero" e "faixa etária" não contribuem significativamente para a percepção dos falantes cultos no que diz respeito às variantes analisadas.

3.2 A Variável Concordância Verbal

Como já vimos, das três variáveis que escolhemos para análise nesta pesquisa, a regra da concordância verbal é a mais saliente, havendo uma maior estigmatização da variante não

padrão por parte dos falantes cultos. Nesta seção, tentaremos responder ao terceiro problema proposto na Introdução deste trabalho: quais contextos mais afetam a percepção dos usos não padrão. O segundo problema, que diz respeito a identificar qual das variantes é mais saliente, não faz sentido para essa variável: das duas variantes, apenas uma é padrão – a conformação morfológica do verbo com o sujeito. A outra – a não conformação do verbo – é naturalmente mais estigmatizada.

A primeira sentença testada para essa variável foi a CV-1-ADF, "*Esses políticos aqui no Brasil não entendem que quem manda é o povo*". Como indica seu código, essa sentença possui um sujeito preposto distante do verbo, com valor [+ animado] e composição simples. As considerações dos informantes sobre essa oração estão resumidas no Gráfico 1:

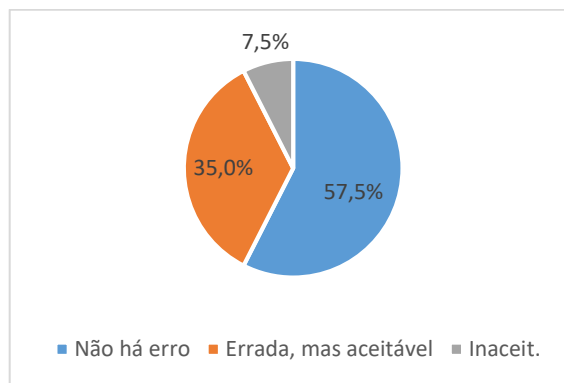


Gráfico 1: Questão CV-1-ADF "Esses políticos aqui no Brasil não entendem que quem manda é o povo"

Vamos comparar esses números aos obtidos para a oração CV-2-ADF, que é sua contraparte não padrão, "*Eles sem nenhum pudor leva materiais de escritório para casa*":

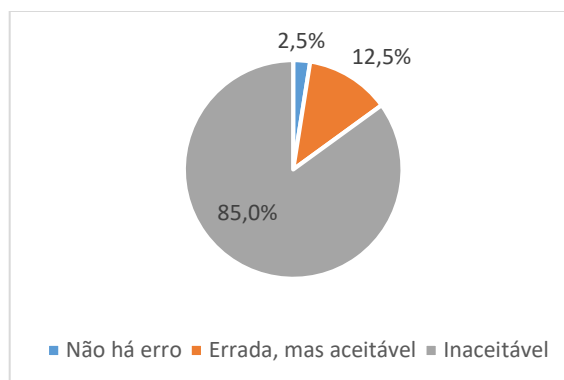


Gráfico 2: Questão CV-2-ADF "Eles sem nenhum pudor leva materiais de escritório para casa"

Como era de se esperar, a variante não padrão foi a mais marcada como "inaceitável", embora 12,5% dos informantes a tenham considerado "errada, mas aceitável" e 2,5% como

"sem erro"³³ (apenas uma ocorrência). Já para a variante padrão, 57,5% dos informantes a consideraram como "sem erro", enquanto que o restante a considerou como "errada". Como já dissemos, esse alto número de informantes considerando uma sentença perfeitamente gramatical para a norma padrão como "errada" pode ter a ver com o uso do pronome demonstrativo *esses*.

Vamos comparar essas duas primeiras orações (com sujeito simples, preposto distante e [+ animado]) com as duas orações, padrão e não padrão, que têm *sujeito composto*: CV-1-ADG, "*Os clientes e os vendedores neste nosso contexto precisam falar a mesma língua*", e CV-2-ADG, "*Os chefes e os funcionários por incrível que pareça se comunica muito bem por aqui*". Seus respectivos gráficos são os seguintes:

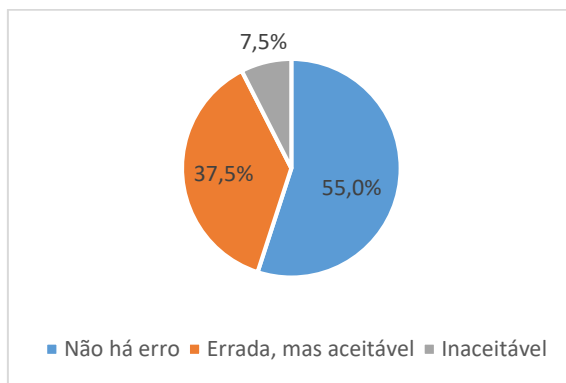


Gráfico 3: Questão CV-1-ADG, "*Os clientes e os vendedores neste nosso contexto precisam falar a mesma língua*"

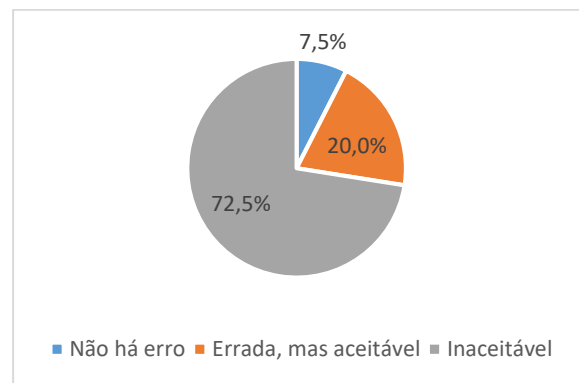


Gráfico 4: Questão CV-2-ADG, "*Os chefes e os funcionários por incrível que pareça se comunica muito bem por aqui*"

A questão com a variante padrão recebeu 55% de marcações como "não há erro", muito próximo dos 57,5% que recebeu sua contraparte com sujeito simples (lembramos que ambas possuem pronomes demonstrativos). Da mesma forma, a oração com a variante não padrão foi marcada por 72,5% dos informantes como "inaceitável", um pouco abaixo dos 85% da oração com o sujeito simples. Esse resultado difere muito dos resultados que encontramos em pesquisas linguísticas e que comentamos em 2.1.2.1: na produção oral dos falantes, os sujeitos compostos levam a um índice de concordância muito baixo. Como vemos, quando o que se testa é a percepção dos falantes cultos, esses dois tipos de sujeito (simples ou composto) praticamente não se diferem no que diz respeito à concordância verbal (no decorrer da análise, voltaremos a essa questão para alterar um pouco nossa conclusão).

³³ Apesar de a expressão presente no questionário ser "não há erro", para facilitar a coesão textual por vezes utilizamos o sinônimo "sem erro".

Todas essas quatro orações que já vimos têm o sujeito do tipo [+ animado]. Vamos compará-las com as orações com sujeito [- animado] presentes no teste. Primeiramente, vejamos as orações CV-1-AEF, "Sobre os malotes, *eles* em pouco tempo chegam para o cliente", e CV-2-AEF, "*As impressoras* na pior hora sempre para de funcionar", que são, respectivamente, as versões padrão e não padrão das orações com sujeito simples, preposto distante e [- animado]. As avaliações dos informantes para essas duas orações estão nos gráficos a seguir:

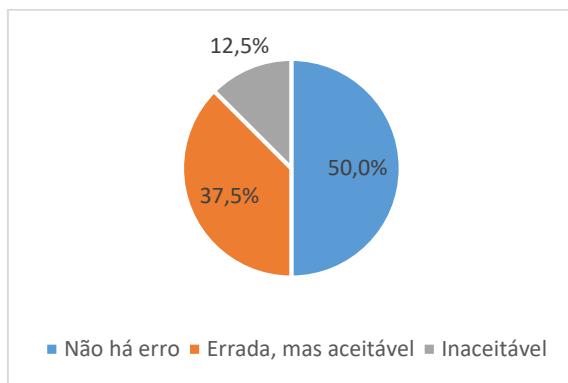


Gráfico 5: Questão CV-1-AEF, "Sobre os malotes, *eles* em pouco tempo chegam para o cliente"

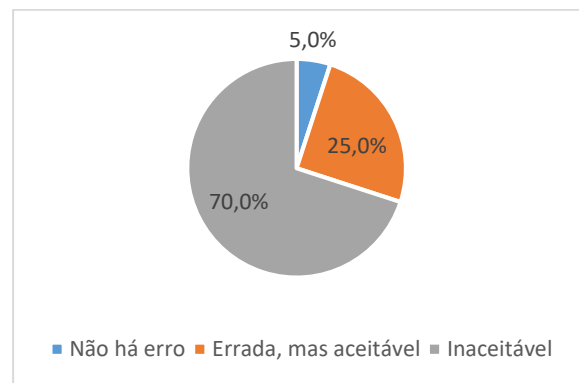


Gráfico 6: Questão CV-2-AEF, "As impressoras na pior hora sempre para de funcionar"

No caso da questão CV-1-AEF, apesar de ser uma oração que segue as prescrições da norma padrão, 50% dos informantes a consideraram errada, sendo que, desses, 37,5% a marcaram como "errada, mas aceitável". Como comentamos em 2.3.3, imaginamos que esse comportamento se deva à presença do material interveniente que separa o sujeito do verbo, um adjunto adverbial. Esse adjunto está fora de seu lugar natural, que seria ao final da oração, fato que pode ter influenciado a percepção dos informantes.

A questão CV-2-AEF, que traz a variante não padrão, foi marcada por 70% dos informantes como "inaceitável" e por 25% como "errada, mas aceitável". Apenas 5% dos informantes não encontraram fugas à norma padrão na oração, o que era de se esperar.

Vejamos agora as questões que contêm sujeito [- animado] e composto, ainda dentro das orações com sujeito preposto distante do verbo:

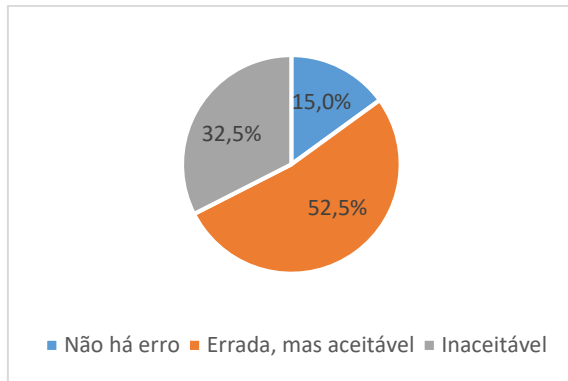


Gráfico 7: Questão CV-1-AEG, "O caminhão e a van quando eu dirijo não levam nenhuma multa"

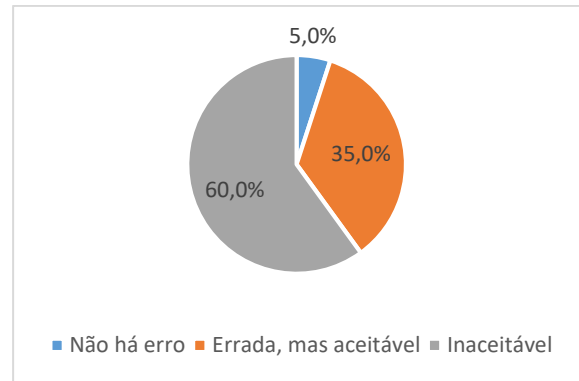


Gráfico 8: Questão CV-2-AEG, "O computador e a impressora pelo que vi ontem precisa de conserto"

A questão CV-1-AEG, que está dentro das normas padrão, foi surpreendentemente mais marcada como "errada, mas aceitável" ou como "inaceitável" do que como "sem erro": ao todo, 85% dos informantes acharam que havia um erro em sua construção. Se formos considerar novamente que o adjunto adverbial deslocado possa ter causado esse tipo de percepção, é possível afirmar que, pelo fato de o adjunto nessa oração ser uma oração completa, isso tenha agravado o quadro. A oração CV-2-AEG, não padrão, recebeu 95% de marcações "errada, mas aceitável" ou "inaceitável". Mais uma vez o adjunto adverbial oracional pode ter influenciado esse resultado.

Vamos comparar as percepções dos informantes em relação às orações não padrão comentadas até agora (com sujeito preposto distante do verbo):

Tabela 5: Comparação das orações não padrão com sujeito preposto distante (variável "concordância verbal")

| | Sujeito [+ animado] | | Sujeito [- animado] | |
|-----------------------|---------------------|----------|---------------------|----------|
| | CV-2-ADF | CV-2-ADG | CV-2-AEF | CV-2-AEG |
| Errada, mas aceitável | 12,5% | 20% | 25% | 35% |
| Inaceitável | 85% | 72,5% | 70% | 60% |
| TOTAL | 97,5% | 92,5% | 95% | 95% |

Como se pode depreender da Tabela 5, não houve grandes diferenças na percepção da agramaticalidade, para a norma padrão, dessas sentenças: em todos os casos, cerca de 95% (+ ou - 2,5%) dos informantes tacharam essas orações como fugas à norma padrão. Já o grau de aceitabilidade variou levemente: a oração com sujeito animado e simples foi marcada por 85% dos informantes como "inaceitável"; a com sujeito animado e composto foi considerada "inaceitável" por 72,5% dos informantes; a com sujeito inanimado e simples, por 70%; e a com

sujeito inanimado e composto, por 60%. Podemos dizer que a animacidade do sujeito ou sua composição não contribuem significativamente para a saliência da variante não padrão, embora a junção desses dois fatores pareça ser importante para essa saliência: de um lado, a concordância não canônica em orações com sujeitos simples e animados é eminentemente reprovada por 85% dos falantes cultos; do outro, sujeitos compostos e inanimados são inaceitáveis para 60% desses falantes. De qualquer forma, a diferença não é tão importante – pelo menos não ao considerarmos apenas as sentenças vistas até este ponto. Até o fim do capítulo, reveremos essa interpretação.

Vejam agora as questões que trazem sujeitos antepostos imediatamente ao verbo. Primeiro, faremos a mesma análise já feita para as orações com sujeitos antepostos distantes, para depois comparar essas duas categorias de orações.

A questão CV-1-BDF, "*Meus amigos conhecem esse novo estoquista*", que tem sujeito simples, [+ animado] e imediatamente preposto ao verbo, foi considerada "errada, mas aceitável" por 7,5% dos informantes e "sem erro" por 92,5% deles. Nenhum dos informantes a considerou "inaceitável". Veja o Gráfico 9:

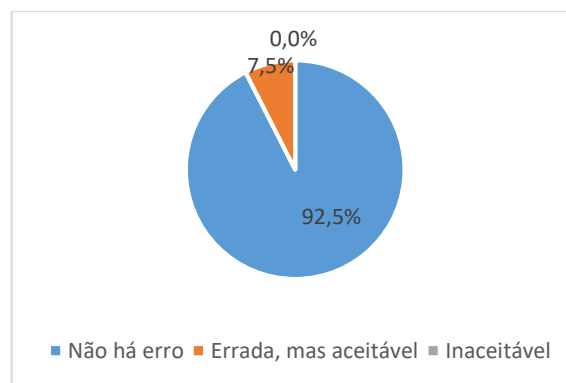


Gráfico 9: *Questão CV-1-BDF, "Meus amigos conhecem esse novo estoquista"*

Sua contraparte não padrão (CV-2-BDF), "*Esses atendentes acredita mesmo nessa abordagem?*", foi marcada como "errada, mas aceitável" por 12,5% dos informantes, "sem erro" por 5% e como "inaceitável" por 82,5% dos informantes. Veja o Gráfico 10:

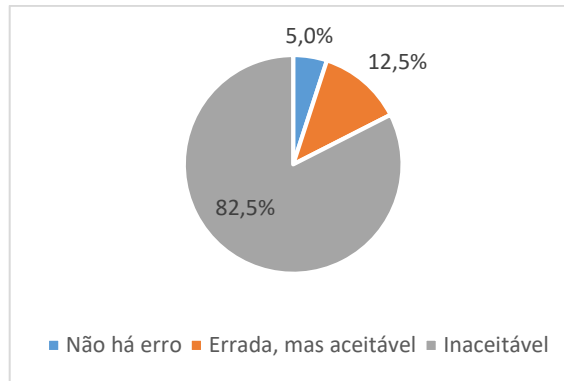


Gráfico 10: Questão CV-2-BDF, "Esses atendentes acredita mesmo nessa abordagem?"

O primeiro fato que se pode notar, a partir dessas duas questões, é que, quando o sujeito está imediatamente preposto ao verbo, parece haver mais segurança dos informantes em marcar as orações como "sem erro" ou "inaceitáveis". Essas orações têm um processamento mais fácil, o que pode facilitar a percepção da obediência ou fuga à norma padrão. Ambas as sentenças foram avaliadas exatamente como se esperava.

As questões CV-1-BDG, "Os familiares e amigos podem usar as instalações", e CV-2-BDG, "Os colegas e clientes considera sua saída uma pena", que contêm sujeitos [+ animados] e compostos, receberam as seguintes avaliações:

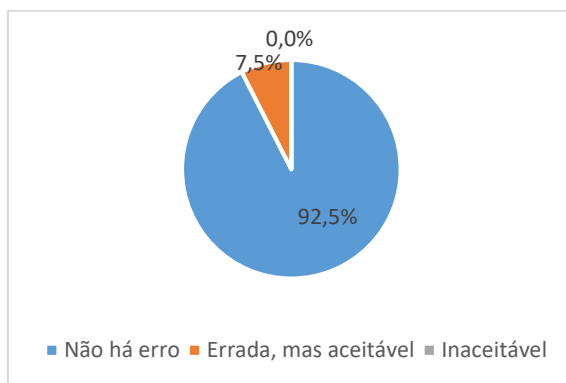


Gráfico 11: Questão CV-1-BDG, "Os familiares e amigos podem usar as instalações"

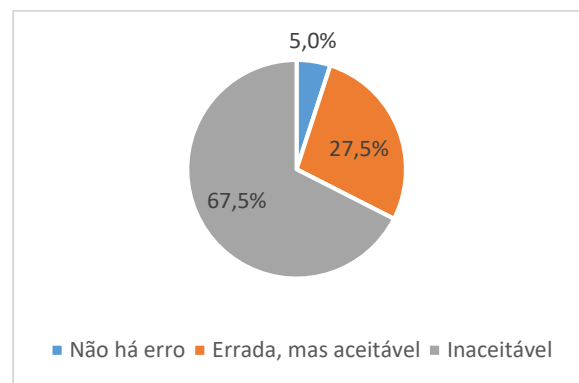


Gráfico 12: Questão CV-2-BDG, "Os colegas e clientes considera sua saída uma pena"

Novamente, no caso da sentença com a variante padrão, a aceitação foi quase unânime: 92,5% dos informantes a marcaram como "sem erro". Já no caso da oração não padrão, houve uma leve diferença: a oração com sujeito composto recebeu mais avaliações de "errada, mas aceitável" que sua versão com sujeito simples.

As orações CV-1-BEF, "Gosto desses novos produtos, mas *eles* precisam de uma boa publicidade", e CV-2-BEF, "As *peças* chega em fevereiro", contêm sujeitos simples e [-animados] e foram avaliadas como se segue:

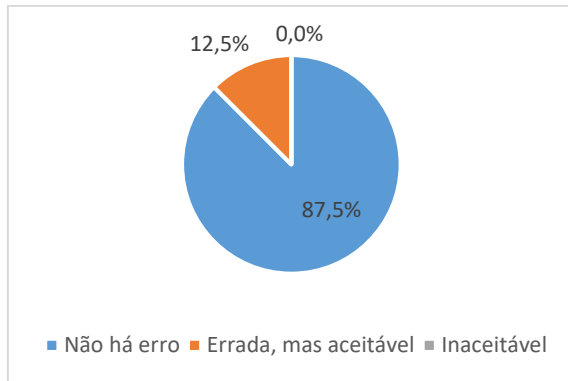


Gráfico 13: Questão CV-1-BEF, "Gosto desses novos produtos, mas eles precisam de uma boa publicidade"

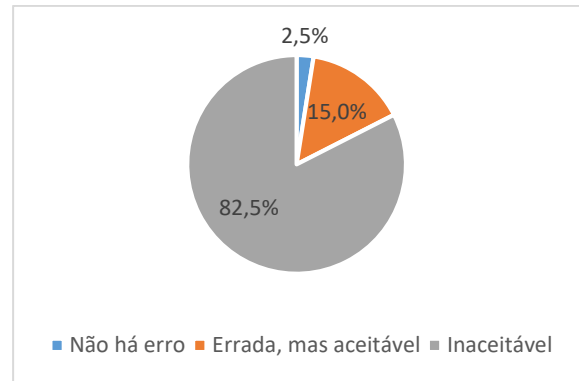


Gráfico 14: Questão CV-2-BEF, "As peças chega em fevereiro"

Os resultados foram praticamente iguais aos encontrados para as orações com sujeitos [+ animado], como se pode ver nos Gráficos 13 e 14 (na verdade, no caso das orações não padrão, foram *exatamente* iguais). Podemos concluir, como já fizemos antes, que o fator "animacidade do sujeito" parece não contribuir significativamente para a avaliação da concordância verbal e, conseqüentemente, para a saliência da variante não padrão. Vamos analisar agora os resultados da avaliação das questões com sujeito composto, [- animados] e imediatamente prepostos para incluir novos dados que possam corroborar essa conclusão:

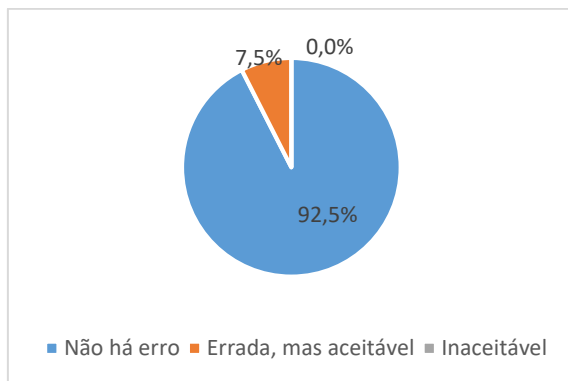


Gráfico 15: Questão CV-1-BEG, "Aposto que o carro e a moto pertencem ao chefe"

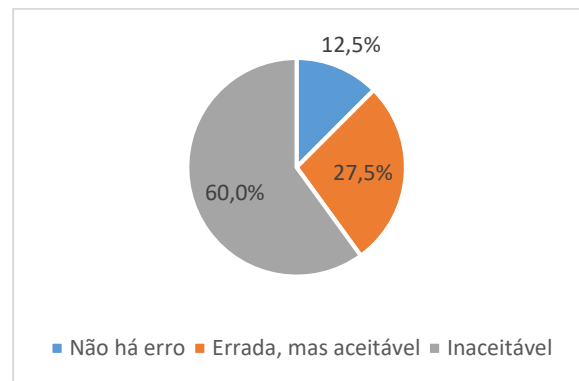


Gráfico 16: Questão CV-2-BEG, "A pista e o depoimento completa a solução do problema"

Como podemos observar, os dados obtidos para a sentença CV-1-BEG, "Aposto que o carro e a moto pertencem ao chefe", são idênticos aos obtidos para as questões CV-1-BDF e CV-1-BDG, que são as versões com sujeito simples/ [+ animado] e composto/ [+ animado]. A oração CV-2-BEG, "A pista e o depoimento completa a solução do problema", que é não padrão, recebeu avaliação negativa de 87,5% dos informantes, sendo que, desses, 60% a consideraram inaceitável. Se compararmos esses números com os das questões já analisadas, veremos que tanto na questão CV-2-BEG quanto na CV-2-BDG, o índice de avaliações "inaceitável" foi próximo: 60% e 67,5%, respectivamente. A mesma coisa acontece ao compararmos as questões

CV-2-BEF e CV-2-BDF, com 82,5% de marcações "inaceitável" para cada. Olhando para os resultados desses dois pares de orações, podemos afirmar com certeza que o fator "animacidade do sujeito", que é o único fator que muda entre os dois pares, de fato não exerce influência no grau de saliência da variável concordância verbal.

Tendo analisado essas primeiras 16 orações, já estamos em condições de construir parcialmente o perfil de comportamento do fator "posição e distância do sujeito em relação ao verbo", comparando suas duas primeiras categorias: "sujeito preposto distante do verbo" e "sujeito preposto imediatamente ao verbo". Confira a Tabela 6:

Tabela 6: Comparação das orações não padrão com sujeito preposto distante e com sujeito preposto imediatamente ao verbo (variável "concordância verbal")

| | Sujeito preposto distante do verbo | | | | Sujeito preposto imediatamente ao verbo | | | |
|-----------------------|------------------------------------|----------|---------------------|----------|---|----------|---------------------|----------|
| | Sujeito [+ animado] | | Sujeito [- animado] | | Sujeito [+ animado] | | Sujeito [- animado] | |
| | CV-2-ADF | CV-2-ADG | CV-2-AEF | CV-2-AEG | CV-2-BDF | CV-2-BDG | CV-2-BEF | CV-2-BEG |
| Errada, mas aceitável | 12,5% | 20% | 25% | 35% | 12,5% | 27,5% | 15% | 27,5% |
| Inaceitável | 85% | 72,5% | 70% | 60% | 82,5% | 67,5% | 82,5% | 60% |
| TOTAL | 97,5% | 92,5% | 95% | 95% | 95% | 95% | 97,5% | 87,5% |

Observando-se os pares de orações "irmãs", vemos que as únicas diferenças significativas entre os resultados das que têm sujeito preposto distante e sujeito preposto imediato se dá entre as questões CV-2-AEF e CV-2-BEF: a primeira foi marcada como "errada, mas aceitável" por 25% dos informantes e a segunda, por 15%; a primeira foi marcada como "inaceitável" por 70% dos informantes e a segunda, por 82,5%. Aparentemente, quando a oração possui um sujeito inanimado e simples, o fator "posição e distância do sujeito e relação ao verbo" exerce uma influência levemente considerável na percepção da variante não padrão.

Podemos comparar também, já com alguma segurança, as orações com sujeito simples e composto: aquelas com sujeito composto sempre recebem menos avaliações negativas: olhando na tabela a linha de porcentagens de marcações da categoria "inaceitável", vemos que entre a CV-2-ADF e a CV-2-ADG houve uma diferença de 12,5%; entre a CV-2-AEF e a CV-2-AEG, de 10%; entre a CV-2-BDF e a CV-2-BDG, de 15%; e entre a CV-2-BEF e a CV-2-BEG, de 22,5%. Isto é: em primeiro lugar, observamos que, como previsto pelas pesquisas linguísticas (cf. 2.1.2), os sujeitos simples são mais salientes para a concordância verbal (isto é, mais deixam transparecer a falta de concordância); em segundo lugar, observamos que a combinação sujeito composto + sujeito [- animado] revelam os maiores índices de marcação

"inaceitável". Basicamente as mesmas proporções foram observadas para as marcações de "errada, mas aceitável".

As últimas orações que testaram a variável concordância verbal são aquelas que trazem a categoria "sujeito posposto" no fator "posição e distância do sujeito em relação ao verbo". Vejamos as duas primeiras, que representam as versões padrão e não padrão com sujeito simples e [+ animado]: CV-1-CDF, "Nascem *muitos bebês* nessa época do ano", e CV-2-CDF, "Todos os dias chega *vários compradores* de São Paulo".

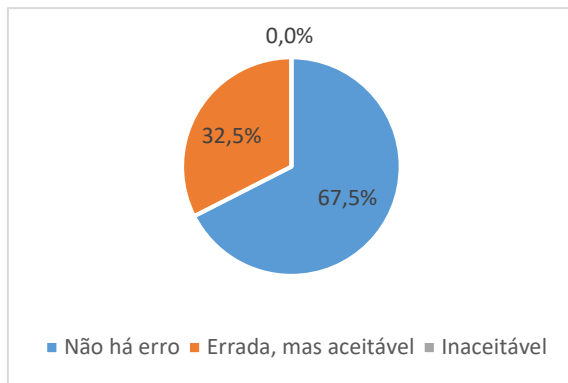


Gráfico 17: Questão CV-1-CDF, "Nascem muitos bebês nessa época do ano"

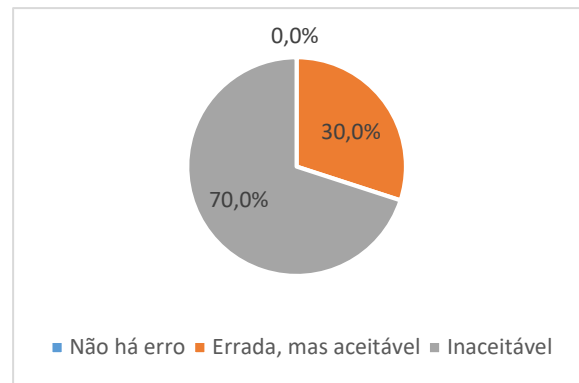


Gráfico 18: Questão CV-2-CDF, "Todos os dias chega vários compradores de São Paulo"

Nenhum informante considerou inaceitável a questão CV-1-CDF, que é padrão, mas 32,5% a consideraram "errada, mas aceitável". Mais uma vez atribuímos esse resultado à presença do pronome demonstrativo. No caso da oração não padrão CV-2-CDF, houve uma surpresa: até o momento, ela foi a única oração não padrão que nenhum informante considerou "sem erro", o que contraria totalmente as expectativas, já que os estudos linguísticos mostram que o sujeito posposto é o que menos salienta a falta de concordância (cf. 2.1.2). 30% dos informantes a consideraram "errada, mas aceitável", número que está dentro da média do que vimos para cada questão até agora.

As próximas questões são a CV-1-CDG, "Existem *um gerente e um contador* exatamente para isso", e CV-2-CDG, "Se alguém fica sabendo desse erro, morre *eu e você*", que são as versões com sujeito composto e [+ animado]:

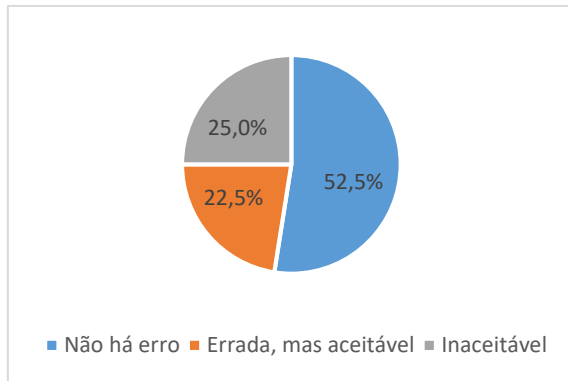


Gráfico 19: Questão CV-1-CDG, "Existem um gerente e um contador exatamente para isso"

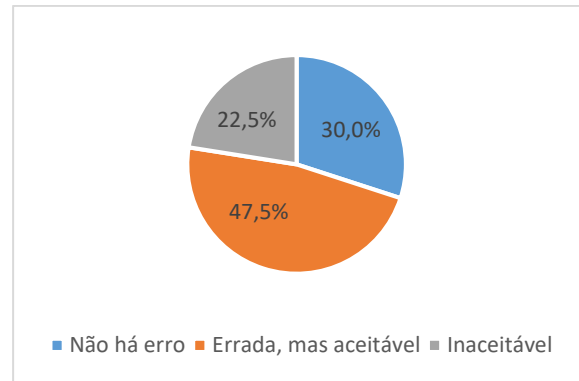


Gráfico 20: Questão CV-2-CDG, "Se alguém fica sabendo desse erro, morre eu e você?"

Essas duas questões tiveram respostas mais diversificadas. A versão padrão levou 52,5% dos informantes a marcarem a opção "não há erro"; 22,5% a consideraram "errada, mas aceitável" e 25%, "inaceitável". Nesse caso, atribuímos o alto índice de marcações negativas a uma oração perfeitamente padrão a dois motivos: em primeiro lugar, à presença do pronome demonstrativo; em segundo lugar, à presença do verbo *existir*, que, ao lado de *haver* (como veremos em 3.5), confunde os usuários da língua quando o assunto é concordância: como os falantes cultos em geral conhecem as particularidades do verbo *haver* quando significa "existir", há a tendência a se hipercorrigir e transferir essas particularidades também para o verbo *existir*, usando-o sempre na terceira pessoa do singular. Reveremos essa suposição mais à frente, quando analisarmos a questão CV-1-CEF.

A oração não padrão foi considerada "sem erro" por 30% dos informantes, o que é revelador. Ao a compararmos com sua contraparte com sujeito simples (CV-2-CDF), vemos que o sujeito composto parece ter agido decisivamente para camuflar a falta de concordância, deixando a variante não padrão menos saliente. Veremos se esse resultado se mantém ao analisarmos as próximas questões. Outro dado revelador desses resultados é que mesmo aqueles informantes que detectaram o erro o consideraram "menos grave", tendo em vista que 47,5% o consideraram aceitável.

As questões com sujeito simples e [- animado] são as seguintes: CV-1-CEF, "Existem muitos problemas com essa nova abordagem", e CV-2-CEF, "Antigamente saía várias vendas por dia. Hoje está mais difícil":

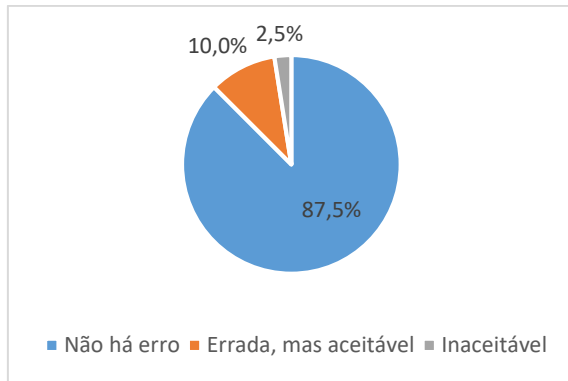


Gráfico 21: Questão CV-1-CEF, "Existem muitos problemas com essa nova abordagem"

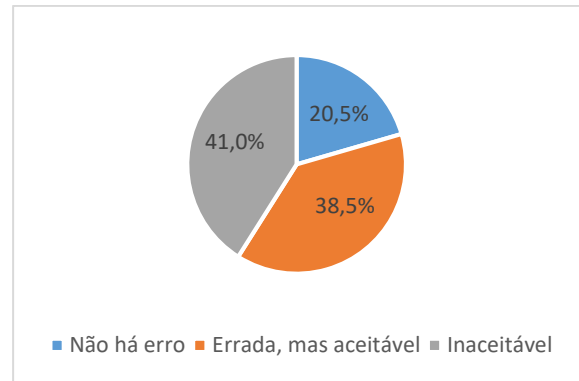


Gráfico 22: Questão CV-2-CEF, "Antigamente saía várias vendas por dia. Hoje está mais difícil"

Os dados relacionados à questão CV-1-CEF nos mostram que, a despeito de o verbo predicador da sentença ser *existir*, os informantes não tiveram a mesma reação observada na análise da questão CV-1-CDG, uma vez que a imensa maioria (87,5%) a marcou como "sem erro". Temos de rever nossa suposição: não foi a presença do verbo *existir* que levou 47,5% dos informantes a considerarem a oração CV-1-CDG "errada", mas sua combinação com o sujeito composto "um gerente e um contador". Como vimos em 2.1.1.1, a concordância excepcional com o núcleo de sujeito mais próximo (no caso, "um gerente") é inclusive prevista na gramática normativa como uma possibilidade viável.

Para a oração não padrão CV-2-CEF, notamos novamente um alto índice de aceitação: 38,5% dos informantes enxergaram o "erro", mas o consideraram aceitável; e 20,5% sequer consideraram essa oração inadequada para a norma padrão. Como dessa vez temos um sujeito simples (ao contrário da oração CV-2-CDG, para a qual também notamos grande aceitação), temos que supor que foi o traço [- animado] do sujeito que camuflou a falta de concordância. Se realmente for esse o caso, a oração com sujeito composto e traço [- animado] seria a que deixa menos evidente a variante não padrão. Vejamos.

As duas últimas orações que utilizamos para testar a percepção sobre a variável concordância verbal são a CV-1-CEG, "É possível que creçam o *interesse e a procura por nossos produtos*", e a CV-2-CEG, "Causou-me estranheza o *acidente e a morte* numa rua tão calma". Seus respectivos gráficos são os seguintes:

Tabela 7, seção 2

| | Sujeito posposto ao verbo | | | |
|-----------------------|---------------------------|-----------|---------------------|-----------|
| | Sujeito [+ animado] | | Sujeito [- animado] | |
| | Sim-ples | Com-posto | Sim-ples | Com-posto |
| | CV-1-CDF | CV-1-CDG | CV-1-CEF | CV-1-CEG |
| Não há erro | 67,5% | 52,5% | 87,5% | 64,1% |
| Errada, mas aceitável | 32,5% | 22,5% | 10% | 23,1% |
| Inaceitável | 0% | 25% | 2,5% | 12,8% |
| TOTAL | 100% | 100% | 100% | 100% |

A partir dessa primeira tabela, que, lembramos, resume os resultados da percepção dos falantes cultos sobre as orações **padrão** com a variável concordância verbal do questionário, podemos extrair algumas conclusões:

Em geral, as orações com sujeito preposto imediatamente ao verbo foram as mais reconhecidas como "sem erro" (média de 91,5%), seguidas pelas orações com sujeito posposto ao verbo (média de 68%) e por aquelas com sujeito preposto distante (média de 43,8%). Aparentemente, os sujeitos distante e posposto ao verbo levam o falante culto a ter dúvidas quanto à correta concordância do verbo, o que parece ser agravado quando o sujeito é composto (como se pode ver pelos relativamente altos índices de não aceitação das orações CV-1-CDG e CV-1-CEG).

As orações com sujeito preposto distante do verbo foram as que mais foram marcadas como "inaceitáveis" pelos falantes cultos (média de 15%), seguida por aquelas com sujeito posposto (média de 10%), apesar de todas serem gramaticais para a norma padrão. As orações com sujeito preposto imediato não foram tidas como inaceitáveis por nenhum informante, apesar de duas delas (CV-1-BDF e CV-1-BEF) terem um pronome demonstrativo *esse* em sua composição, o que de certa forma pede uma revisão da hipótese que vimos defendendo de que foi a presença desse tipo de pronome que levou os falantes cultos a considerarem orações padrão como erradas.

O traço semântico [\pm animado] do sujeito não interferiu significativamente na percepção dos falantes no que diz respeito às orações padrão.

A tabela a seguir traz os resultados da percepção de falantes cultos sobre as orações **não padrão** do questionário:

Tabela 8: Comparação das orações não padrão com a variável "concordância verbal"

Tabela 8, seção 1

| | Sujeito preposto distante do verbo | | | | Sujeito preposto imediatamente ao verbo | | | |
|-----------------------|------------------------------------|----------|---------------------|----------|---|----------|---------------------|----------|
| | Sujeito [+ animado] | | Sujeito [- animado] | | Sujeito [+ animado] | | Sujeito [- animado] | |
| | Simples | Composto | Simples | Composto | Simples | Composto | Simples | Composto |
| | CV-2-ADF | CV-2-ADG | CV-2-AEF | CV-2-AEG | CV-2-BDF | CV-2-BDG | CV-2-BEF | CV-2-BEG |
| Não há erro | 2,5% | 7,5% | 5% | 5% | 5% | 5% | 2,5% | 12,5% |
| Errada, mas aceitável | 12,5% | 20% | 25% | 35% | 12,5% | 27,5% | 15% | 27,5% |
| Inaceitável | 85% | 72,5% | 70% | 60% | 82,5% | 67,5% | 82,5% | 60% |
| TOTAL | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% |

Tabela 8, seção 2

| | Sujeito posposto ao verbo | | | |
|-----------------------|---------------------------|----------|---------------------|----------|
| | Sujeito [+ animado] | | Sujeito [- animado] | |
| | Simples | Composto | Simples | Composto |
| | CV-2-CDF | CV-2-CDG | CV-2-CEF | CV-2-CEG |
| Não há erro | 0% | 30% | 20,5% | 80% |
| Errada, mas aceitável | 30% | 47,5% | 38,5% | 2,5% |
| Inaceitável | 70% | 22,5% | 41% | 17,5% |
| TOTAL | 100% | 100% | 100% | 100% |

Para a categoria "posição e distância do sujeito em relação ao verbo", os fatores "sujeito preposto distante do verbo" e "sujeito preposto imediatamente ao verbo" afetam negativamente a percepção da variante estigmatizada: orações com o sujeito nessa configuração foram marcadas como "inaceitáveis" por uma média de 72,5% dos informantes, contra média de 37,7% para orações com sujeito posposto; já o fator "sujeito posposto ao verbo" parece ser decisivo para "camuflar" a impropriedade gramatical, com média de 32,6% de marcações "sem erro" para as orações com a concordância errada, de acordo com a norma padrão. Dito de outra forma, a preposição do sujeito ao verbo é um contexto sintático que *salienta mais* a variável concordância verbal que a posposição do sujeito.

A animacidade do sujeito novamente parece interferir pouco na percepção dos informantes em relação à variável concordância verbal. Tomando-se, por exemplos, os resultados relativos às orações com sujeito preposto (distante/ imediato), vemos que a média de marcações "inaceitável" foi de 78,7%/ 75% para as orações com sujeito [+ animado] e de 65%/ 71,5% para as de sujeito [- animado], o que representa uma diferença de pouco mais de 10% para o primeiro par e menos de 5% para o segundo. As médias de marcação "errada, mas

aceitável" para essas mesmas orações foram de 16,2%/ 20% para sujeitos [+ animados] e 30%/ 21,2% para sujeitos [- animados]. Já ao olharmos para as médias de marcações "errada, mas aceitável", vemos que as orações com sujeito [- animados] foram sistematicamente mais aceitas que suas contrapartes [+ animadas]. Sendo assim, podemos dizer que a animacidade do sujeito *salienta pouco* a variável concordância verbal.

O nível de aceitabilidade dos "erros" em orações com o sujeito posposto foi sensivelmente maior (média de 29%) do que para as orações com sujeito preposto (média de 21,5%).

A composição do sujeito foi relativamente decisiva na percepção dos informantes, de forma que, olhando para a Tabela 6, podemos perceber que em geral as orações com sujeito simples tiveram um nível de aceitação menor (na maioria dos casos, mais de 10 pontos percentuais menor). Dessa forma, tal como observamos nos trabalhos descritivos que resenhamos em 2.1.2.1, podemos dizer que o sujeito simples *salienta mais* a variável concordância verbal que o sujeito composto (que, por sua vez, "camufla" mais uma possível impropriedade gramatical para a norma padrão).

Vamos retomar o problema-objetivo proposto na Introdução: no que se refere à variável concordância verbal, quais contextos mais afetam a percepção dos usos impróprios segundo a gramática normativa? Ou, em outras palavras, quais fatores linguísticos mais contribuem para a saliência de uma variante? Depois da análise feita, podemos concluir que tanto a **posição preposta ao verbo do sujeito** quanto a **estrutura composta do sujeito** salientam *mais* essa variável que a posição posposta e a estrutura simples.

3.3 A Variável Objeto Direto Anafórico

Como vimos em 3.1, a variável "objeto direto anafórico" fica em segundo lugar quando consideramos o grau de saliência das três variáveis principais escolhidas para este estudo: 70,83% das ocorrências não padrão de retomada por objeto direto anafórico foram percebidas como tais pelos informantes, o que significa que em 29,17% dos casos os falantes cultos não perceberam que se tratava de orações com desvios gramaticais, segundo a norma padrão.

Mais importante que isso: dos 70,83% de casos de ocorrência não padrão dessa variável que foram percebidos pelos informantes, em apenas 12,08% as orações foram tachadas de "inaceitáveis". O restante (58,75%) foi visto como "errada, mas aceitável", dado que nos diz muito sobre a atitude linguística dos informantes. Aparentemente, desvios nas regras de

retomada anafórica não são menos perceptíveis, mas são mais acolhidos como possibilidades não tão estigmatizadas. Esse fato pode refletir a conclusão a que chegou Duarte (1989, p. 30-32), segundo a qual o uso do clítico acusativo, que é uma das estratégias não padrão de retomada por objeto direto, apesar de correto, é considerado pedante pelos falantes cultos, que recorrem em geral à categoria vazia para a retomada anafórica.

Na tabela a seguir comparamos as porcentagens da marcações para cada variante dessa variável:

Tabela 9: Comparação das três variantes da variável "objeto direto anafórico"

| | Clítico acusativo (padrão) | Pronome lexical (não padrão) | Categoria vazia (não padrão) |
|-----------------------|---------------------------------------|---|---|
| Não há erro | 71,13% | 25,83% | 32,92% |
| Errada, mas aceitável | 13,81% | 61,25% | 55,83% |
| Inaceitável | 15,06% | 12,92% | 11,25% |
| TOTAL | 100% | 100% | 100% |

Como vimos em 2.1.2.2, das três variantes possíveis para a variável "objeto direto anafórico", a categoria vazia é a mais frequente na *fala* dos brasileiros, chegando a um índice de $\approx 62\%$ de uso constatado por Duarte (1989) contra apenas $\approx 5\%$ do clítico acusativo e $\approx 15\%$ do pronome lexical. No entanto, a Tabela 7 nos mostra que, quando o que é testado é a *percepção* dos falantes cultos sobre as realizações do objeto direto anafórico, nenhuma das variantes se sobrepõe. A variante padrão foi reconhecida como tal em 71,13% dos casos, o que significa que 28,87% das ocorrências dessa variante foram marcadas como "erradas" pelos falantes cultos (das 6 orações padrão que utilizamos para testar essa variável, apenas duas apresentam os pronomes demonstrativos *esse/este*, o que não justifica esse índice de marcações "erradas" para sentenças perfeitamente gramaticais). No caso das orações não padrão, a variante com pronome lexical foi reconhecida como "errada" em 74,17% das ocorrências e como "sem erro" em 25,83% dos casos; e a variante com categoria vazia foi marcada como "errada" em 67,08% das vezes e como "sem erro" em 32,92% das ocorrências. Não há portanto, diferença significativa entre as três variantes: a padrão é reconhecida como correta em mais de 70% dos casos e as não padrão são reconhecidas como erradas em média de 70% dos casos. Nenhuma das variantes é mais saliente que a outra.

Passemos à análise de cada questão em separado, de modo a averiguar em que medida cada contexto linguístico (fator condicionador) contribui para a saliência das variáveis padrão e não padrão. O Gráfico 25 traz as porcentagens relativas à percepção dos informantes em

relação à oração OD-1-HL, "Quando *a secretária* voltou do almoço, eu a cumprimentei pelo bom trabalho", que, como sabemos, possui a categoria "estrutura simples" no fator "estrutura projetada pelo verbo" e a categoria "objeto animado" no fator "animacidade do objeto":

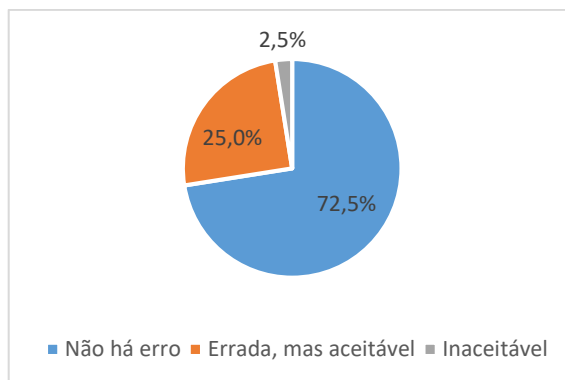


Gráfico 25: Questão OD-1-HL, "Quando a secretária voltou do almoço, eu a cumprimentei pelo bom trabalho"

Como era de se esperar, 72,5% dos informantes perceberam que não havia erro para a norma padrão nessa questão. Mesmo assim, a quantidade de informantes que a consideraram "errada" foi relativamente alta: 27,5%. Não conseguimos imaginar outro motivo para tantos informantes marcarem essa sentença como errada a não ser a pouca familiaridade desses falantes com as regras de retomada anafórica e a insegurança no uso dos pronomes oblíquos. Veremos se esse é mesmo o caso na análise das sentenças seguintes.

A próxima questão é a OD-2-HL, que é a versão não padrão com pronome lexical da oração anterior: "Os clientes chegaram e levei eles até a sala de espera"

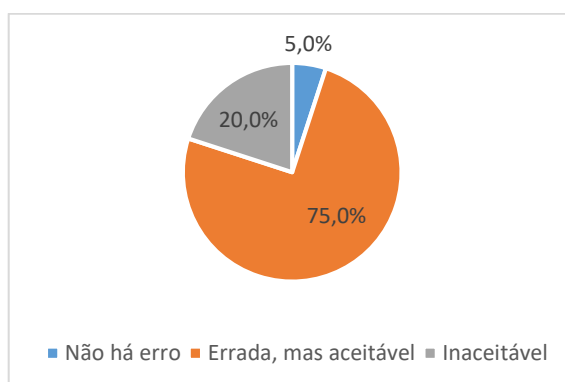


Gráfico 26: Questão OD-2-HL, "Os clientes chegaram e levei eles até a sala de espera"

Apenas 5% dos informantes não enxergaram erro nessa oração. Dos 95% restantes, 75% a consideraram "errada, mas aceitável", o que pode novamente corroborar os achados de Duarte (1989, p. 30-32): os informantes sabem que o uso do pronome lexical não é adequado dentro da norma padrão, mas consideram este um uso aceitável, pois a utilização do clítico soa pedante.

A segunda oração não padrão com estrutura simples e objeto animado é a que traz a categoria vazia, OD-3-HL, "Onde está o *office-boy*? Faz muito tempo que não __ vejo". Confira o Gráfico 27, abaixo:

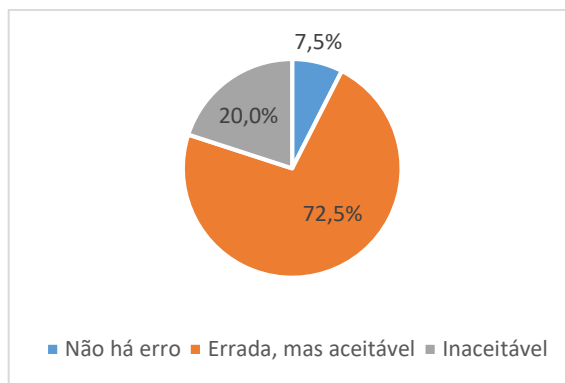


Gráfico 27: Questão OD-3-HL, "Onde está o *office-boy*? Faz muito tempo que não vejo"

Os resultados foram praticamente os mesmos que os da versão com pronome lexical, o que pode apontar que as duas estratégias não padrão têm o mesmo *status* sob o ponto de vista da percepção linguística dos falantes cultos.

Passemos agora às orações do questionário que têm estrutura simples, tais como as três que já vimos, mas que contam com objetos diretos [– animados]. A primeira delas é a OD-1-HM, "Comprei *os materiais* e já os deixei no depósito", que traz o clítico acusativo padrão com estratégia de retomada por objeto direto anafórico:

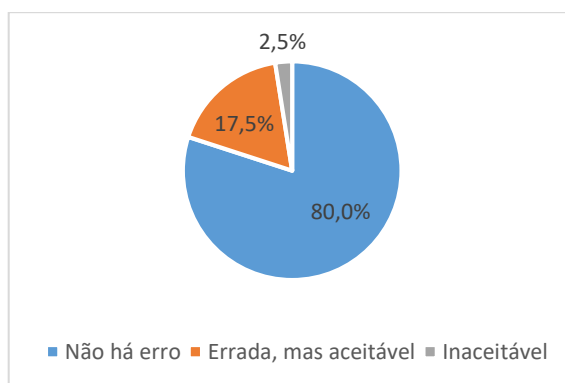


Gráfico 28: Questão OD-1-HM, "Comprei os materiais e já os deixei no depósito"

Um número razoavelmente menor de informantes considerou essa sentença agramatical: ao todo, apenas 20%, contra 27,5% da oração com objeto animado (OD-1-HL). Aparentemente, o fato de o objeto ter o traço [\pm animado] não interfere na percepção dos falantes em relação às possibilidades de retomada anafórica. Vamos conferir mais alguns dados para confirmar ou descartar essa hipótese:

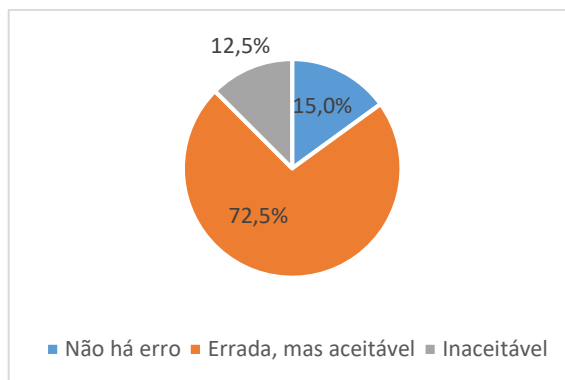


Gráfico 29: Questão OD-2-HM, "Sempre que *uma cadeira* quebra, eu jogo ela fora"

O Gráfico 29 traz os resultados para a questão OD-2-HM, "Sempre que *uma cadeira* quebra, eu jogo ela fora". Novamente os números obtidos são muito parecidos com os de sua contraparte com objeto direto [+ animado]. A única diferença significativa está no número maior de informantes que não enxergaram erro nessa oração: 15%. Vamos comparar esses resultados com os da oração não padrão com categoria vazia, OD-3-HM, "Um dos celulares ainda funciona. Quer que eu deixe ___ aqui com você?":

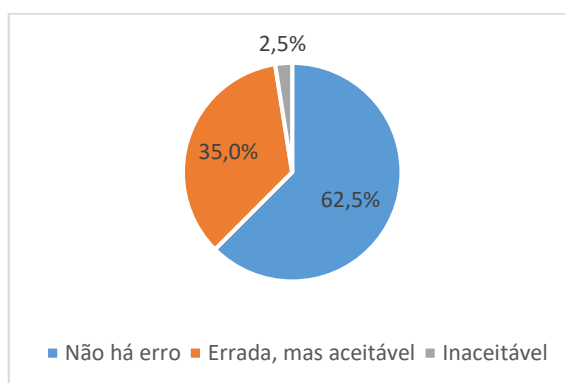


Gráfico 30: Questão OD-3-HM, "Um dos celulares ainda funciona. Quer que eu deixe aqui com você?"

Aparentemente, quando combinamos a variante "categoria vazia" com a categoria "objeto direto [- animado]", a percepção dos falantes cultos é totalmente diferente da dos outros casos: 62,5% dos informantes não viram erro algum na oração OD-3-HM; daqueles que detectaram a impropriedade para a norma padrão, 35% não a consideraram inaceitável. Esse é um dado importante, que nos faz rever as hipóteses levantadas até agora e que corrobora os achados das pesquisas sociolinguísticas que apresentamos em 2.1.2.2: a animacidade do objeto tem, sim, influência sobre a percepção e a atitude dos falantes em relação à retomada por objeto direto anafórico.

A questão OD-1-IL, "O estagiário andava muito distraído, mas eu o deixei em estado de alerta", traz a versão padrão (com clítico acusativo) da oração com objeto animado e estrutura

de predicativo: o pronome *o* é sujeito de uma "oração pequena" (*small clause*; cf. MIOTO, SILVA e LOPES, 2005, p. 107), de forma que "em estado de alerta" assume a função de predicativo do objeto direto. Vejamos as opiniões dos informantes sobre essa sentença:

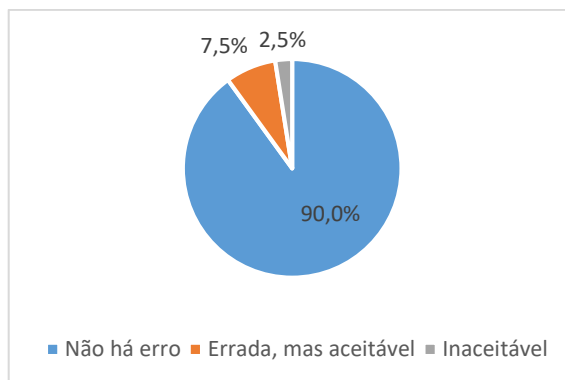


Gráfico 31: Questão OD-1-IL, "O estagiário andava muito distraído, mas eu o deixei em estado de alerta"

90% dos informantes a consideraram correta e, dos poucos que a consideraram incorreta, 7,5% a marcaram como "errada, mas aceitável", o que indica ampla aceitação da gramaticalidade da oração. Vamos comparar esses resultados com suas contrapartes não padrão com pronome lexical (OD-2-IL, "Acho *a Maria* uma pessoa difícil, mas o João acha ela simpática") e com a categoria vazia (OD-3-IL, "Conheço *o Carlos*, mas não ___ considero um bom administrador"):

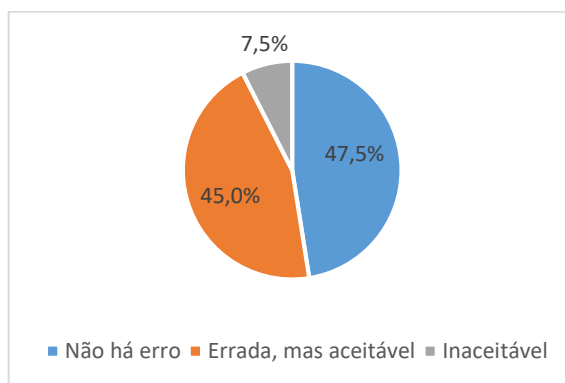


Gráfico 32: Questão OD-2-IL, "Acho *a Maria* uma pessoa difícil, mas o João acha ela simpática"

A oração com pronome lexical recebeu 47,5% de marcações "sem erro" – um número alto ao se considerar que o uso do pronome lexical é estigmatizado. Se as próximas sentenças com essa variante mantiverem esse índice de aceitação, poderemos supor que a estrutura com predicativo *camufla* o uso não padrão, ou melhor, que a estrutura simples (das três orações que vimos anteriormente) *salienta* o uso não padrão. Dos 52,5% de informantes que perceberam a inadequação à norma padrão dessa oração, apenas 7,5% a consideraram inaceitável, o que também é um número significativo, já que, no caso da oração com as mesmas características

mas com estrutura simples (OD-2-HL), 20% dos 95% de informantes que a marcaram como "inaceitável". Ou seja: comparando essas duas orações, OD-2-HL e OD-2-IL, a com estrutura simples recebeu ao mesmo tempo mais marcações como errada (95% *versus* 52,5%) e, dentre essas marcações, a com estrutura simples foi mais estigmatizada (20% *versus* 7,5% de marcações "inaceitável"). Vejamos os resultados para a questão OD-3-IL:

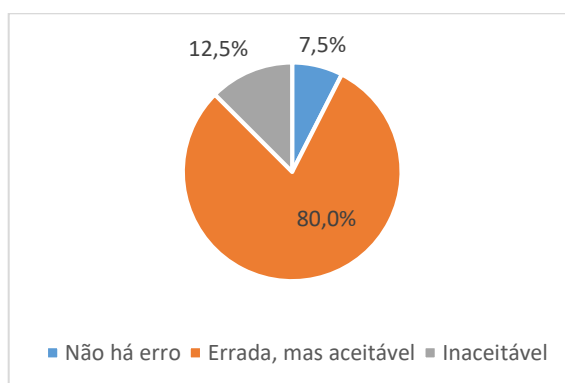


Gráfico 33: Questão OD-3-IL, "Conheço o Carlos, mas não considero um bom administrador"

Ao contrário do que se esperava, a versão com categoria vazia (preferida pelos falantes na produção linguística; cf. 2.1.2.2) foi mais tachada de "errada" que a versão com pronome lexical. 92,5% dos informantes marcaram a sentença OD-3-IL como "errada", ainda que, desses, 80% a consideraram "errada, mas aceitável". Esses números, no entanto, estão consistentes com os resultados obtidos para a versão com estrutura simples, para a qual houve a mesma porcentagem de marcações "errada".

Passemos às orações com objetos diretos [– animados]. A primeira é a questão OD-1-IM, "Mostramos os três currículos para o chefe, mas nenhum o deixou satisfeito":

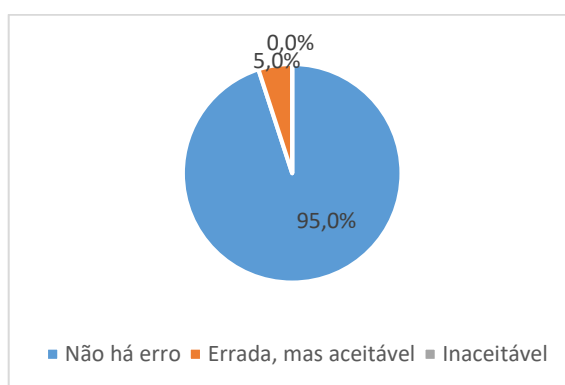


Gráfico 34: Questão OD-1-IM, "Mostramos os três currículos para o chefe, mas nenhum o deixou satisfeito"

Apenas 5% dos informantes consideraram essa oração "errada", e mesmo assim com a ressalva de que ela é aceitável. A segunda oração com objeto direto [– animado] é não padrão

por retomar o objeto direto por meio de um pronome lexical: OD-2-IM, "*Esta sala costumava ser horrível, mas a Sílvia tornou ela agradável*":

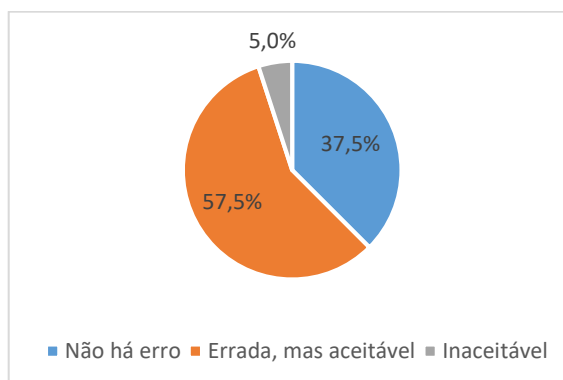


Gráfico 35: Questão OD-2-IM, "*Esta sala costumava ser horrível, mas a Sílvia tornou ela agradável*"

As diferenças entre a versão com objeto direto [- animado] e [+ animado] são sutis: OD-2-IM recebeu 37,5% de avaliações "sem erro", ao passo que OD-2-IL recebeu 47,5%; ambas também tiveram pouquíssimas avaliações de "inaceitável": 5% *versus* 7,5%. A maior diferença, portanto, foi nas marcações "errada, mas aceitável": 57,5% para a versão com objeto inanimado *versus* 45% para a versão com objeto animado.

A última questão com estrutura com predicativo é a OD-3-IM, "*A nova campanha não me agradou muito, mas o cliente ___ achou ótima*":

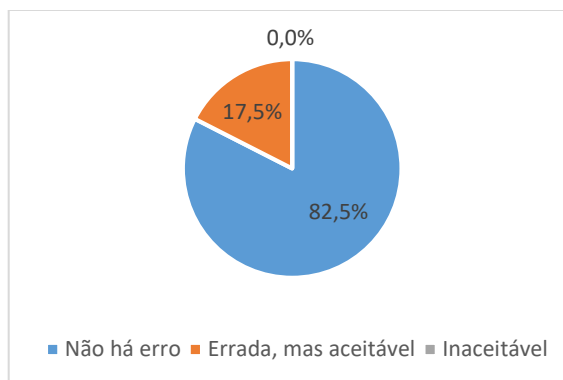


Gráfico 36: Questão OD-3-IM, "*A nova campanha não me agradou muito, mas o cliente achou ótima*"

Novamente a combinação "categoria vazia" + "objeto direto inanimado" camuflou substancialmente a variante não padrão: 82,5% dos informantes não reconheceram nenhuma impropriedade gramatical na oração OD-3-IM, contra 17,5% que enxergaram o erro, mas mesmo assim o consideraram "aceitável". Lembremos que esses resultados se parecem com os obtidos para a questão OD-3-HM, em que 62,5% dos informantes consideraram uma questão com categoria vazia como "sem erro". Se essa tendência se mantiver nas próximas questões,

podemos afirmar com alguma certeza que a categoria [+ animado] é decisiva para salientar a variante "categoria vazia" na retomada por objeto direto anafórico.

Tendo analisado as primeiras 12 questões da variável "objeto direto anafórico", já podemos construir parcialmente o perfil de comportamento dessa variável nos ambientes sintáticos "estrutura simples" e "estrutura com predicativo", no que se refere às duas variantes não padrão:

Tabela 10: Comparação das orações não padrão com estrutura simples e estrutura de predicativo (variável "objeto direto anafórico")

| | Estrutura simples | | | | Estrutura com predicativo | | | |
|-----------------------|--------------------|---------|--------------------|---------|---------------------------|---------|--------------------|---------|
| | Objeto [+ animado] | | Objeto [- animado] | | Objeto [+ animado] | | Objeto [- animado] | |
| | OD-2-HL | OD-3-HL | OD-2-HM | OD-3-HM | OD-2-IL | OD-3-IL | OD-2-IM | OD-2-IM |
| Não há erro | 5% | 7,5% | 15% | 62,5% | 47,5% | 7,5% | 37,5% | 82,5% |
| Errada, mas aceitável | 75% | 72,5% | 72,5% | 35% | 45% | 80% | 57,5% | 17,5% |
| Inaceitável | 20% | 20% | 12,5% | 2,5% | 7,5% | 12,5% | 5% | 0% |
| TOTAL | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% |

Primeiramente, é possível perceber que só há uma diferença significativa entre as avaliações das duas variantes não padrão quando está presente um objeto [- animado]: nesse caso, a variante com pronome lexical desponta como a mais saliente, com uma média de 73,75% dos informantes percebendo-as como "erradas" (contra 27,5% para a variante com categoria vazia), embora a maioria ainda a considere "aceitável".

Comparando-se os resultados para as orações OD-2-HL e OD-2-IL, podemos perceber também que a estrutura simples, nesse caso, salientou mais a variante não padrão: enquanto 95% dos informantes perceberam que havia um "erro" em OD-2-HL, apenas 52,5% perceberam a mesma impropriedade gramatical em OD-2-IL. As outras duas questões com a variante com pronome lexical seguem a mesma tendência: 85% dos informantes avaliaram OD-2-HM como "errada", mas apenas 62,5% deram a mesma avaliação para OD-2-IM. No entanto, esse resultado parece valer apenas para a variante com pronome lexical, pois os resultados para OD-3-HL e OD-3-IL são muito parecidos.

Passemos agora à análise das orações que trazem o fator "estrutura projetada pelo verbo" com a categoria "estrutura complexa". A primeira dessas orações é OD-1-JL, "*O novo atendente já está bem treinado. Eu já o deixo fazer tudo sozinho*":

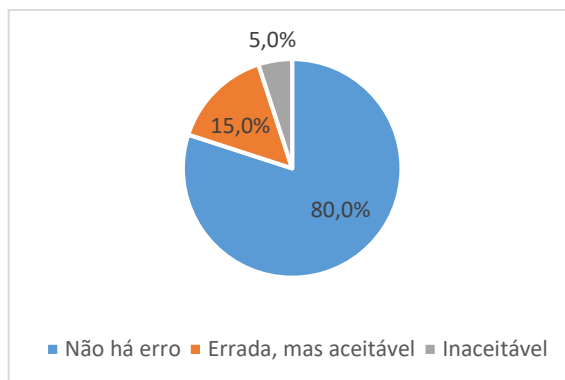
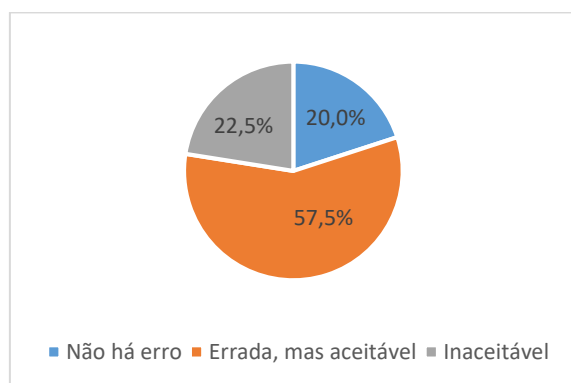


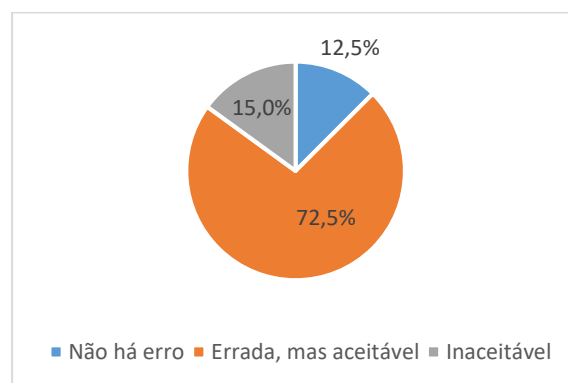
Gráfico 37: Questão OD-1-JL, "O novo atendente já está bem treinado. Eu já o deixo fazer tudo sozinho"

80% dos informantes reconheceram que essa oração não traz nenhum problema gramatical, no que diz respeito à norma padrão. Os 20% restantes a consideraram "errada, mas aceitável" ou "inaceitável". Esse é um número relativamente alto, que só podemos atribuir, novamente, à estranheza causada pelo uso do clítico acusativo, que, apesar de correto segundo a gramática normativa, soa estranho aos ouvidos dos falantes brasileiros.

Abaixo, trazemos as versões com pronome lexical e categoria vazia: OD-2-JL, "Eu não confio *nesses políticos* porque eu já vi eles falando mal do povo", e OD-3-JL, "O chefe brigou com o *office-boy* e mandou ___ sair da sala":



*Gráfico 38: Questão OD-2-JL, "Eu não confio *nesses políticos* porque eu já vi eles falando mal do povo"*



*Gráfico 39: Questão OD-3-JL, "O chefe brigou com o *office-boy* e mandou ___ sair da sala"*

Para ambas as questões não padrão com objeto [+ animado] e estrutura complexa, os informantes marcaram em maior quantidade a opção "errada, mas aceitável": 57,5% para OD-2-JL e 72,5% para OD-3-JL. Aqui podemos observar a tendência descrita pelos trabalhos científicos citados em 2.1.2.2: apesar de ambas as opções de retomada por objeto anafórico serem agramaticais segundo as regras normativas, a categoria vazia é percebida como mais "neutra" que o pronome lexical, sendo *usada* na grande maioria dos casos e, muitas vezes, também *percebida* como mais aceitável.

Passemos às últimas três questões desta variável, que trazem a retomada de objetos [–animados] dentro de uma estrutura frasal complexa. A primeira é OD-1-JM, "Não sei para que serve *essa máquina*, eu nunca a vi funcionando":

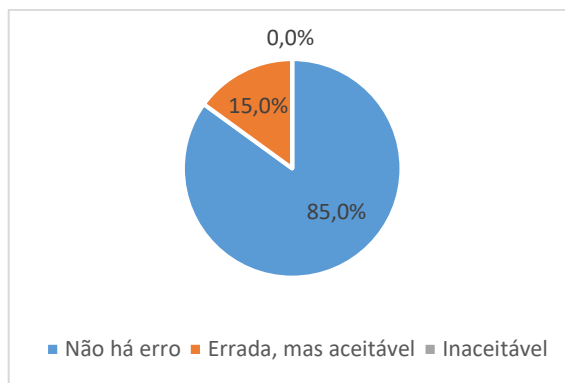


Gráfico 40: Questão OD-1-JM, "Não sei para que serve essa máquina, eu nunca a vi funcionando"

Seguindo a tendência das orações com a variante padrão da variável retomada por objeto direto anafórico, 85% dos informantes consideraram essa oração perfeitamente gramatical e apenas 15% a consideraram "errada, mas aceitável". Tendo analisado essa última oração com a variante padrão, já podemos dizer que, no caso específico dessa variante, nenhum dos fatores ou categorias testados parece interferir decisivamente no julgamento dos falantes cultos.

Comparando as versões com objeto [+ animado] e [– animado], podemos perceber uma ligeira propensão dos informantes a considerarem como "sem erro" mais as retomadas de objetos [– animado]: 80% *versus* 72,5% para OD-1-HL (animado) e OD-1-HM (inanimado); 95% *versus* 90% para OD-1-IL e OD-1-IM; 85% *versus* 80% para OD-1-JL e OD-1-JM. Outro destaque é da leve diferença entre as três estruturas frasais testadas no que diz respeito às marcações de "errada, mas aceitável": as orações com estrutura simples e complexas foram marcadas como "erradas, mas aceitáveis" por uma média de 21,25% e 15% dos informantes, respectivamente, ao passo que aquelas com estrutura de predicativo foram avaliadas dessa forma por uma média de apenas 6,25% dos informantes. Isso pode significar que as estruturas simples e complexas evidenciam o uso do clítico acusativo (torna-o mais saliente), que, como sabemos, é visto com estranheza até mesmo pelos falantes cultos.

A penúltima questão para esta variável é a OD-2-JM, "A impressora estava travando, mas eu fiz ela funcionar de novo":

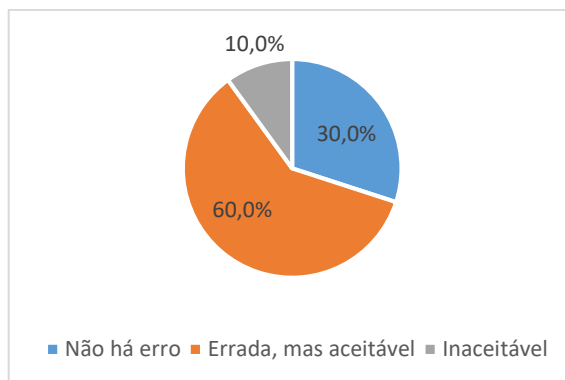


Gráfico 41: Questão OD-2-JM, "A impressora estava travando, mas eu fiz ela funcionar de novo"

A versão com pronome lexical obteve um índice de 30% de marcações "sem erro", 10% a mais que sua contraparte com objeto [+ animado]. Houve ainda 60% dos informantes marcando-a como "errada, mas aceitável" e apenas 10% marcando-a como inaceitável.

Tendo em vista as seis orações com retomada anafórica por pronome lexical, percebemos o seguinte: no caso das orações com estrutura simples e complexa, as orações com objeto [– animado] foram mais marcadas como "sem erro" (em ambos os casos, a diferença foi de exatamente 10%) e as com objeto [+ animado] foram mais marcadas como "inaceitáveis". Já no caso das orações com estrutura de predicativo, o quadro se inverte: a versão com objeto [+ animado] foi mais marcada como "sem erro" (47,5% versus 37,5% da versão inanimada). No que se refere à marcação como "errada, mas aceitável", as estruturas simples, com predicativo e complexas foram assim percebidas por 73,75%, 51,25% e 58,75% dos informantes, respectivamente. Como vemos, as orações simples parecem ser as mais bem aceitas, quando o erro é percebido, ao passo que as orações com predicativo parecem ser as que mais "camuflam" o erro, com média de 42,5% de marcações "sem erro" (contra média de 10% e 25% das orações simples e complexas – ou seja, essas duas estruturas salientam mais o uso não padrão do pronome lexical).

A última oração para a variável retomada por objeto direto anafórico é a OD-3-JM, "Essa sua cadeira precisa de óleo! Você não ___ ouve rangendo?":

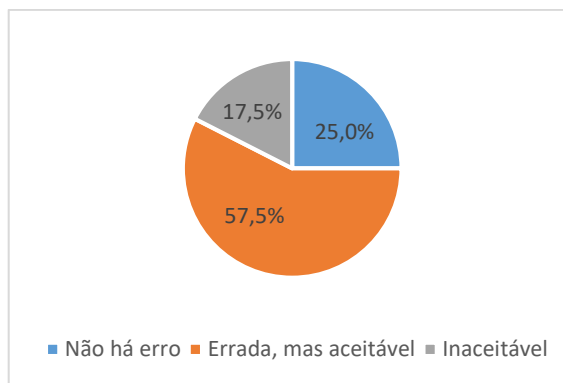


Gráfico 42: Questão OD-3-JM, "Essa sua cadeira precisa de óleo! Você não ouviu rangendo?"

25% dos informantes não encontraram a impropriedade gramatical na oração OD-3-JM, 57,5% a consideraram "errada, mas aceitável" e 17,5% a consideraram "inaceitável". Esses números são consistentes com os resultados já encontrados para as outras orações não padrão com estrutura complexa: em geral, cerca de 50% dos informantes as consideraram "erradas, mas aceitáveis" e o restante se dividiu entre as outras duas opções, com uma leve vantagem para "sem erro".

A tabela a seguir traz os resultados da percepção de falantes cultos sobre as orações **padrão** com a variável objeto direto anafórico do questionário:

Tabela 11: Comparação das orações padrão da variável "objeto direto anafórico"

| | Estrutura simples | | Estrutura com predicativo | | Estrutura complexa | |
|-----------------------|--------------------|--------------------|---------------------------|--------------------|--------------------|--------------------|
| | Objeto [+ animado] | Objeto [- animado] | Objeto [+ animado] | Objeto [- animado] | Objeto [+ animado] | Objeto [- animado] |
| | OD-1-HL | OD-1-HM | OD-1-IL | OD-1-IM | OD-1-JL | OD-1-JM |
| Não há erro | 72,5% | 80% | 90% | 95% | 80% | 85% |
| Errada, mas aceitável | 25% | 17,5% | 7,5% | 5% | 15% | 15% |
| Inaceitável | 2,5% | 2,5% | 2,5% | 0% | 5% | 0% |
| TOTAL | 100% | 100% | 100% | 100% | 100% | 100% |

Podemos extrair desse quadro algumas conclusões:

A média geral de marcações "sem erro" para as orações padrão foi de 83,75%, o que significa que em mais de 80% dos casos os falantes cultos reconheceram que as orações respeitavam a norma padrão da gramática portuguesa. Podemos perceber uma leve vantagem das orações com estrutura de predicativo: no caso delas, 92,5% dos informantes as reconheceram como corretas. Já no caso das estruturas simples e complexas, essa média cai para 76,25% e 82,5%, respectivamente, o que também indica uma média maior de informantes as marcando como "erradas, mas aceitáveis" (21,25% e 15%, respectivamente). Sendo assim,

podemos classificar essas três estruturas em um *ranking*: as orações de estrutura simples são as que mais salientam a variante padrão, seguidas das complexas e das de predicativo.

Ao comparar as três orações com objeto [+ animado] e as três com objeto [– animado], percebemos que as estruturas com objetos inanimados são levemente mais marcadas como "sem erro" e levemente menos marcadas como "erradas, mas aceitáveis", mas a diferença não é significativa.

A tabela a seguir (uma expansão da Tabela 10) traz os resultados da percepção de falantes cultos sobre as orações **não padrão** com a variável objeto direto anafórico:

Tabela 12: Comparação das orações não padrão da variável "objeto direto anafórico"

Tabela 12, seção 1

| | Estrutura simples | | | | Estrutura com predicativo | | | |
|-----------------------|--------------------|-------------|--------------------|-------------|---------------------------|-------------|--------------------|-------------|
| | Objeto [+ animado] | | Objeto [– animado] | | Objeto [+ animado] | | Objeto [– animado] | |
| | OD-2- HL | OD-3- HL | OD-2- HM | OD-3- HM | OD-2- IL | OD-3- IL | OD-2- IM | OD-3- IM |
| Não há erro | 5% | 7,5% | 15% | 62,5% | 47,5% | 7,5% | 37,5% | 82,5% |
| Errada, mas aceitável | 75% | 72,5% | 72,5% | 35% | 45% | 80% | 57,5% | 17,5% |
| Inaceitável | 20% | 20% | 12,5% | 2,5% | 7,5% | 12,5% | 5% | 0% |
| TOTAL | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% |

Tabela 12, seção 2

| | Estrutura complexa | | | |
|-----------------------|--------------------|-------------|--------------------|-------------|
| | Objeto [+ animado] | | Objeto [– animado] | |
| | OD-2- JL | OD-3- JL | OD-2- JM | OD-3- JM |
| Não há erro | 20% | 12,5% | 30% | 25% |
| Errada, mas aceitável | 57,5% | 72,5% | 60% | 57,5% |
| Inaceitável | 22,5% | 15% | 10% | 17,5% |
| TOTAL | 100% | 100% | 100% | 100% |

Podemos extrair desse quadro algumas conclusões:

As estruturas simples e com predicativo, quando combinadas com o fator objeto direto [– animado], parecem camuflar o uso da variante categoria vazia, mas o mesmo não acontece com a estrutura oracional complexa. Ou seja, a estrutura complexa *salienta mais* a categoria vazia de objetos diretos anafóricos inanimados.

Exceto no caso indicado acima, a estrutura simples parece permitir uma avaliação mais positiva em relação aos usos não padrão: pouco mais de 70% dos informantes consideraram as orações não padrão como "erradas, mas aceitáveis" e pouco menos de 20% como "inaceitáveis".

Esse quadro se inverte, como já dito, no caso da oração OD-3-HM, que teve um índice de marcações "sem erro" de 62,5%.

Já no caso das estruturas de predicativo o quadro não foi tão uniforme: no caso das estruturas com predicativo e objeto [+ animado], por exemplo, a oração com pronome lexical teve um índice de marcações "sem erro" seis vezes maior que sua contraparte com categoria vazia e um índice de marcações "errada, mas aceitável" duas vezes menor. Já para as orações com objeto [- animado] esse quadro se inverte totalmente: a oração com pronome lexical foi muito menos marcada como "sem erro" (37,5% *versus* 82,5% da categoria vazia) e muito mais marcada como "errada, mas aceitável" (57,5% *versus* 17,5%). Para as estruturas complexas, o único resultado mais discrepante foi o da oração OD-3-JL, que despontou com 72,5% dos informantes a considerando "errada, mas aceitável" contra uma média de 58,3%.

De modo geral, e respondendo à pergunta que fizemos no início desse capítulo, sobre quais fatores condicionadores mais contribuem para a saliência das variantes não padrão, podemos observar pela tabela que as orações com **objeto direto [- animado]** foram as que mais receberam avaliações de "sem erro", o que pode nos dizer que o objeto direto [+ animado] *salienta mais* os usos impróprios da retomada por objeto direto anafórico. É possível perceber também que as porcentagens relativas à opção "não há erro" são razoavelmente maiores para as **orações com predicativo** (média de 43,75%) que para as orações com estrutura complexa (média de 21,87%), o que pode querer dizer que a estrutura complexa *salienta mais* as variantes não padrão.

3.4 A Variável Estratégias de Relativização

A variável "estratégias de relativização" fica em terceiro lugar quando consideramos o grau de saliência das três variáveis principais escolhidas para este estudo: 69,37% das ocorrências não padrão de relativização foram percebidas como tais pelos informantes, o que significa que em 30,63% dos casos os falantes cultos não perceberam que se tratava de orações com desvios gramaticais, segundo a norma padrão. Dos 69,37% de marcações como "errada" para essas orações, 42,29% foram de marcações "errada, mas aceitável", o que significa que apenas 27,08% dos casos foram vistos pelos informantes como "inaceitáveis".

Na tabela a seguir comparamos as porcentagens da marcações para cada variante dessa variável:

Tabela 13: Comparação das três variantes da variável "estratégias de relativização"

| | Variante padrão | Cortadora | Copiadora |
|-----------------------|-----------------|--------------|--------------|
| Não há erro | 71,67% | 50% | 11,3% |
| Errada, mas aceitável | 25% | 44,2% | 40,4% |
| Inaceitável | 3,33% | 5,8% | 48,3% |
| TOTAL | 100% | 100% | 100% |

Como vimos em 2.1.2.3, das três estratégias de relativização disponíveis aos falantes do PB, a variante padrão é a menos usada pelos falantes, inclusive os cultos. Dentre as duas estratégias não padrão, a relativa cortadora e a copiadora, a cortadora é a mais usada (61,4% de ocorrências, segundo MOLLICA, 1977) e a copiadora a menos usada (38,6%). Segundo Bagno (2004),

Estamos lidando aqui, portanto, com questões de atitudes dos falantes em relação a regras específicas da língua. Para os falantes cultos, a relativa padrão é pouco aceita porque parece ser "certa demais", ao passo que a relativa copiadora, identificada (sem razão, porém) com os falantes menos instruídos, é menos aceita porque parece "errada demais". Assim, para manter uma posição equidistante entre esses dois extremos, o brasileiro culto opta por empregar a relativa cortadora (BAGNO, 2004, p. 89).

E foi exatamente a esse resultado que nossa pesquisa também chegou. Se na produção linguística os falantes tendem a utilizar mais a relativa cortadora em detrimento da copiadora, o mesmo acontece quando o que se testa é a percepção desses falantes. Os informantes não enxergaram erro em 50% dos casos de relativas cortadoras presentes no nosso questionário, mas perceberam que havia uma impropriedade gramatical em 88,7% das ocorrências de relativas copiadoras. E mesmo dentre aqueles que reconheceram que essas duas estratégias fogem às regras da norma padrão, tanto nos casos de relativa cortadora quanto copiadora houve uma média de 42,3% de marcações "errada, mas aceitável". Dessa forma, e respondendo à segunda pergunta-problema que fizemos na Introdução, podemos afirmar com segurança que a relativa copiadora é a variante não padrão *mais saliente* da variável "estratégias de relativização", pois é a que é mais reconhecida como "errada" pelos falantes cultos.

Passemos à análise de cada oração utilizada no questionário para testar esta variável. A primeira é ER-1-NQ, que traz uma oração com a variante padrão de relativização, cuja função sintática do pronome relativo é de objeto indireto e que tem como antecedente do pronome relativo um sintagma nominal [+ animado], "*O distribuidor em quem eu confiei está nos passando para trás*":

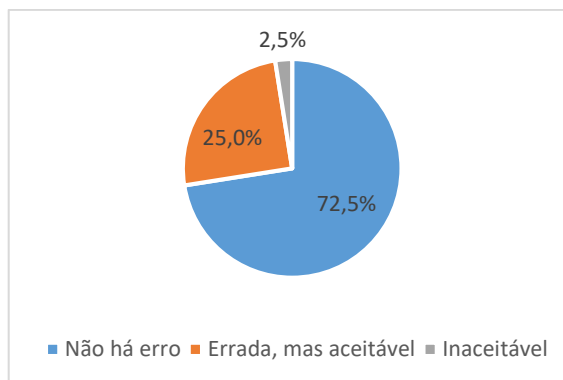


Gráfico 43: Questão ER-1-NQ, "O distribuidor em quem eu confiei está nos passando para trás"

Como esperado, a maioria dos informantes – 72,5% – não considerou essa oração como errada. Mesmo assim, 27,5% deles reconheceram alguma impropriedade gramatical na frase. Talvez possamos atribuir esse número ao fato de as pessoas estranharem o uso padrão da relativização, com a movimentação da preposição para uma posição não natural. Segundo Bagno (2004, p. 87), a relativa padrão é uma estrutura de *síntese* gramatical, estranha ao português, que é uma língua *analítica*. Isso explicaria a relutância dos falantes brasileiros em utilizarem essa estrutura, e pode explicar o porquê de tantos informantes terem enxergado ER-1-NQ como errada em nossa pesquisa.

A versão não padrão com relativa cortadora é a oração ER-2-NQ, "O Messias é o vendedor que eu mais gosto ___!":

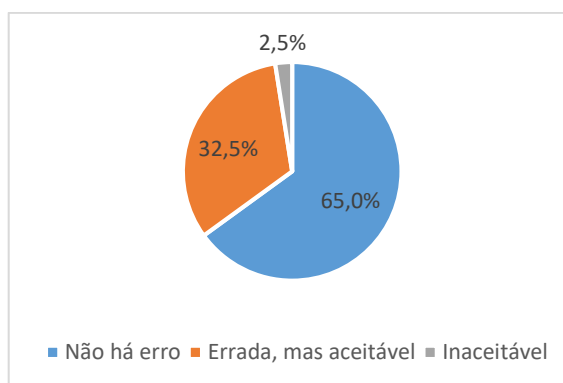


Gráfico 44: Questão ER-2-NQ, "O Messias é o vendedor que eu mais gosto!"

65% dos informantes não perceberam que nessa oração há uma fuga à norma padrão. Dos 35% restantes, que enxergaram a impropriedade, 32,5% não consideram esse erro de grande importância e o consideraram "aceitável". Esse será o quadro para a maioria das orações com a variante cortadora, como já vimos no início desta seção. Já para a oração com a variante copiadora, o resultado é bastante diferente. Veja o gráfico a seguir, que traz os resultados para a oração ER-3-NQ, "O diretor que eu discordei dele estava muito seguro de sua opinião":

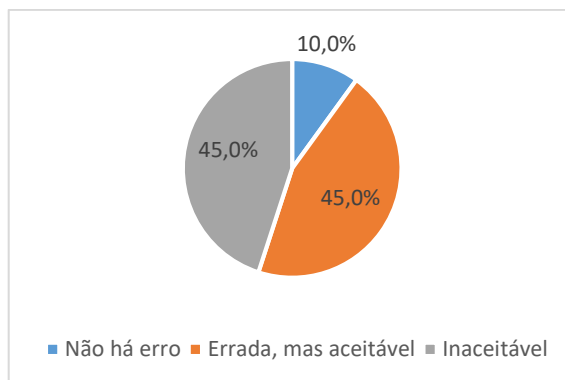


Gráfico 45: Questão ER-3-NQ, "O diretor que eu discordo dele estava muito seguro de sua opinião"

Nesse caso, o número de pessoas que considerou a oração "errada" é muito maior: 90% dos informantes, sendo que, desses, 45% considera a oração "errada, mas aceitável".

Vejam agora as três sentenças com o mesmo tipo de função sintática do pronome relativo que as anteriores, mas com o antecedente [- animado]:

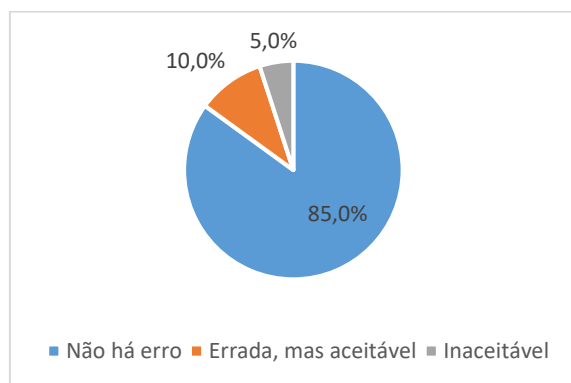


Gráfico 46: Questão ER-1-NR, "As regras com as quais eu concordei não são essas"

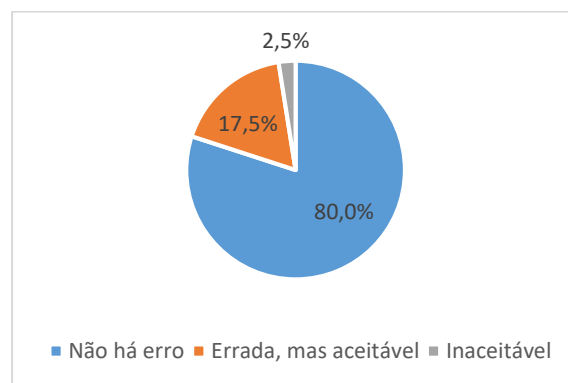


Gráfico 47: Questão ER-2-NR, "Sinto muito, mas esse não é o modelo que eu gosto"

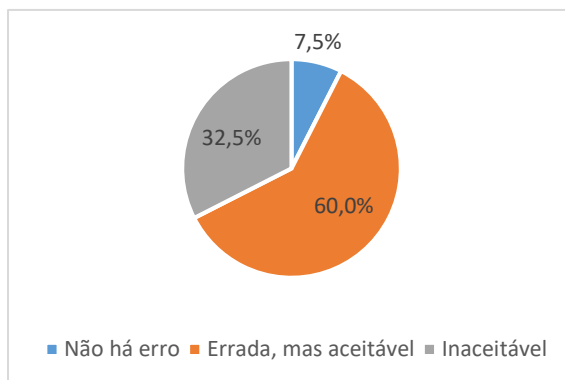


Gráfico 48: Questão ER-3-NR, "Amanhã já é a reunião que você precisa participar dela"

Assim como para sua contraparte com antecedente [+ animado], a grande maioria dos informantes considerou que não havia erro na oração ER-1-NR, "As regras com as quais eu concordei não são essas". Mas um número razoavelmente menor (15%) enxergou erro, o que

pode significar que a animacidade do antecedente pode exercer influência na percepção dos falantes cultos.

A questão com a variante relativa cortadora é ER-2-NR, "Sinto muito, mas esse não é o *modelo* que eu gosto ____". Mais uma vez a relativa cortadora teve um alto índice de marcações "não há erro", e um número menor de informantes que o da oração ER-2-NQ (com antecedente [+ animado]) a considerou "errada". Isso corrobora a hipótese que acabamos de propor de que a falta de animacidade do antecedente age na percepção das propriedades da oração pelos falantes cultos. Observemos, para acrescentar mais dados à discussão, o gráfico com os resultados da última oração com pronome relativo na função de objeto indireto e relativa copiadora, ER-3-NR, "Amanhã já é a *reunião* que você precisa participar *dela*". Primeiramente é preciso registrar que, seguindo a tendência que vimos observando, essa versão teve uma rejeição muito maior que sua contraparte cortadora: apenas 7,5% dos informantes não enxergaram erro em ER-3-NR. Dos 92,5% restantes, 32,5% a consideraram "inaceitável", número menor que aquele obtido pela oração com antecedente [+ animado], que foi de 45%. Nesse caso, o antecedente [- animado] não foi suficiente para mascarar o fato de a oração não seguir as regras da norma padrão.

Passemos às questões com a categoria "adjunto adverbial" no fator "função sintática do pronome relativo":

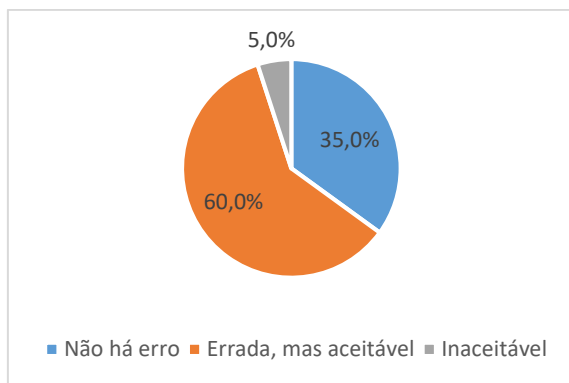


Gráfico 49: Questão ER-1-OQ, "O rapaz com que eu falei é loiro e meio gordinho"

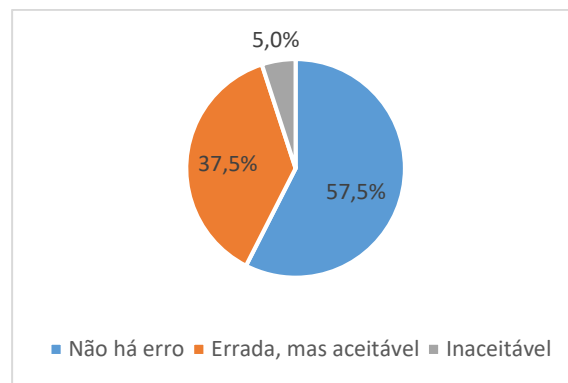


Gráfico 50: Questão ER-2-OQ, "O nome do cliente que eu falei é Marcelo"

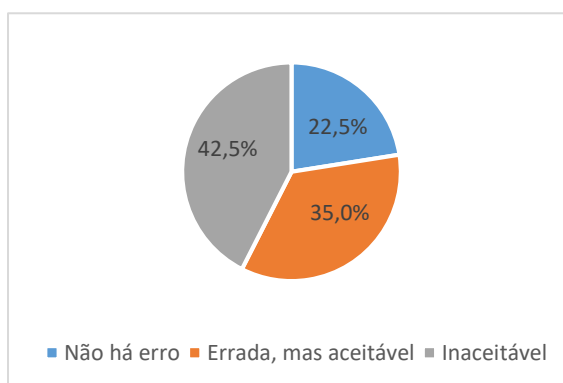


Gráfico 51: Questão ER-3-OQ, "A menina que nós estudamos com ela já está fazendo doutorado"

A primeira delas é ER-1-OQ, "O rapaz com que eu falei é loiro e meio gordinho". Essa oração, apesar de estar na norma padrão, recebeu muitas opiniões negativas: 65% dos informantes enxergaram algum erro em sua estrutura. Imaginamos que esse número se deva à presença do pronome relativo *que* em vez de *quem*, que seria mais natural com um antecedente [+ animado], apesar de ambos estarem corretos, segundo a norma padrão. A oração ER-2-OQ traz a versão cortadora, "O nome do cliente que eu falei ___ é Marcelo". Essa oração traz um problema de interpretação: como falta a preposição, ela pode tanto ser interpretada como contendo uma oração relativa de objeto indireto (*O nome do cliente de que eu falei...*) quanto de adjunto adverbial (*O nome do cliente com que eu falei...*). Dessa forma, não podemos confiar seus resultados inteiramente à percepção dos informantes sobre orações com adjunto adverbial, já que essa não é a única interpretação possível. De qualquer forma, falta-lhe a preposição, o que a classifica como uma oração não padrão, e mesmo assim houve um índice de aprovação de 57,5%, além de um índice de marcações "errada, mas aceitável" de 37,5%. Comparando com a versão com objeto indireto, ER-2-NQ, vemos que a quantidade de informantes que marcaram "não há erro" nas duas é bastante próximo: 65% versus 57,5%.

A oração seguinte, ER-3-OQ, tem uma estrutura com uma relativa copiadora na função de adjunto adverbial e antecedente [+ animado]. Seguindo a tendência que vimos até agora, o antecedente [+ animado] salienta o uso do pronome copiador, deixando mais visível aos informantes que se trata de uma oração não padrão: 42,5% dos informantes a consideraram "inaceitável", número parecido com ER-3-NQ, que também tem um antecedente animado e relativa copiadora.

Passemos às orações com orações relativas na função de adjunto adverbial e com antecedente [- animado].

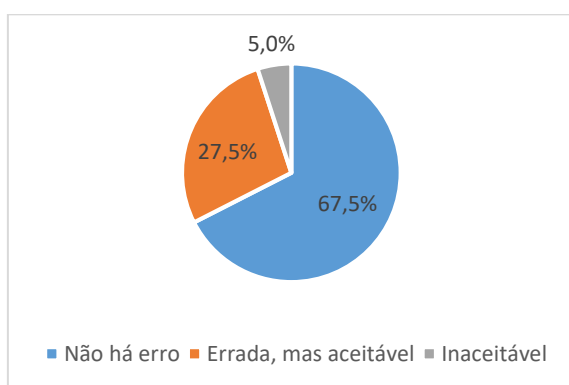


Gráfico 52: Questão ER-1-OR, "O dia em que você chegar na hora será um dia chuvoso"

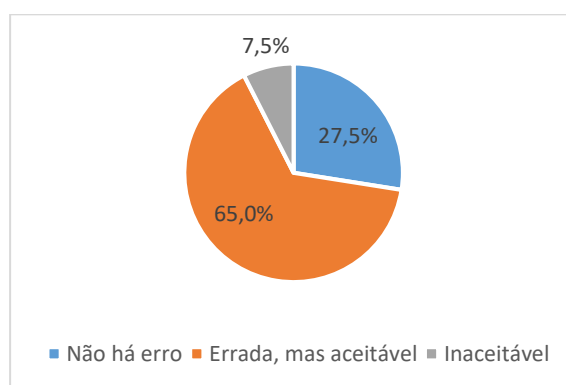


Gráfico 53: Questão ER-2-OR, "A cidade que eu nasci é a 100 km daqui"

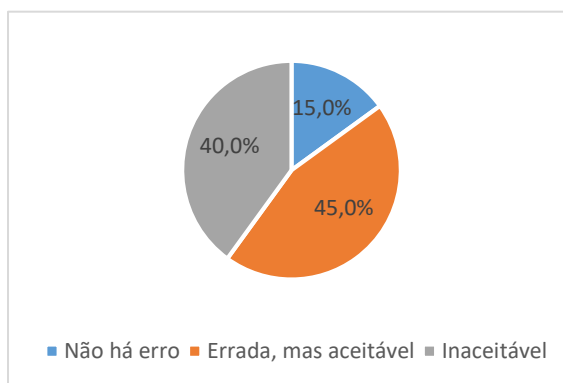


Gráfico 54: Questão ER-3-OR, "Esse não é o carro que eu vim ontem com ele"

A primeira delas é ER-1-OR, "O dia em que você chegar na hora será um dia chuvoso". A maioria dos informantes reconheceu não haver problemas de norma padrão nessa oração, mas uma boa quantidade (32,5%) enxergou algum índice de agramaticalidade. A contraparte cortadora dessa estrutura é a questão ER-2-OR, "A cidade que eu nasci ___ é a 100 km daqui". Pela primeira vez, uma oração cortadora recebe mais avaliações negativas que positivas: 72,5% dos informantes consideram essa oração "errada", contra apenas 27,5% que a marcaram como "sem erro". Aparentemente, a combinação da estrutura relativa em função de adjunto adverbial

com um antecedente [– animado] *salienta* a estrutura cortadora e deixa mais visível a fuga à norma padrão desse expediente gramatical. Já no caso da oração ER-3-OR, "Esse não é o *carro que eu vim ontem com ele*", os resultados foram bem próximos dos de sua contraparte com antecedente [+ animado], embora esta versão com antecedente [– animado] tenha um pouco mais de aceitação.

Após a apresentação dessas 12 primeiras orações, podemos comparar os resultados das orações não padrão com relativas na função de objeto indireto e adjunto adverbial. Confira a tabela a seguir:

Tabela 14: Comparação das orações não padrão com relativa objeto direto e relativa adjunto adverbial (variável "estratégias de relativização")

| | Relativa Objeto Indireto | | | | Relativa Adjunto Adverbial | | | |
|-----------------------|--------------------------|---------|-------------------------|---------|----------------------------|---------|-------------------------|---------|
| | Antecedente [+ animado] | | Antecedente [– animado] | | Antecedente [+ animado] | | Antecedente [– animado] | |
| | ER-2-NQ | ER-3-NQ | ER-2-NR | ER-3-NR | ER-2-OQ | ER-3-OQ | ER-2-OR | ER-3-OR |
| Não há erro | 65 | 10 | 80 | 7,5 | 57,5 | 22,5 | 27,5 | 15 |
| Errada, mas aceitável | 32,5 | 45 | 17,5 | 60 | 37,5 | 35 | 65 | 45 |
| Inaceitável | 2,5 | 45 | 2,5 | 32,5 | 5 | 42,5 | 7,5 | 40 |
| TOTAL | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% |

Como vemos, as orações com a variante copidora foram as que mais tiveram avaliações negativas em qualquer configuração de fatores. Isso já demonstra que essa variante de fato é a mais estigmatizada, como vimos no início desta seção. Parece não haver um padrão que diferencie as orações com relativas objeto direto das com relativas adjunto adverbial: ER-2-NQ e ER-2-OQ têm números muito próximos; no caso de ER-3-NQ e ER-3-OQ, a primeira teve um índice de marcações "sem erro" menor, o que pode significar que a relativa objeto direto *salienta* mais o uso do pronome-cópia. O mesmo acontece entre ER-3-NR e ER-3-OR. Por fim, tendo em vista esse quadro geral, podemos observar também que a animacidade do antecedente exerce uma influência não facilmente padronizável na percepção dos informantes.

Continuando com a análise das orações, passemos àquelas que trazem orações relativas na forma de um adjunto adnominal genitivo. Os gráficos a seguir trazem os resultados para as orações ER-1-PQ, ER-2-PQ e ER-3-PQ:

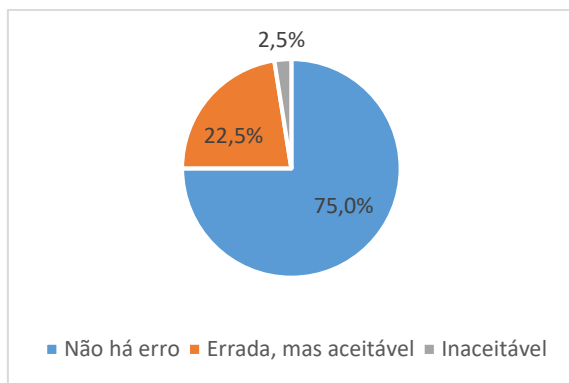


Gráfico 55: Questão ER-1-PQ, "O cliente cuja compra eu fechei gastou mais de 10 mil reais!"

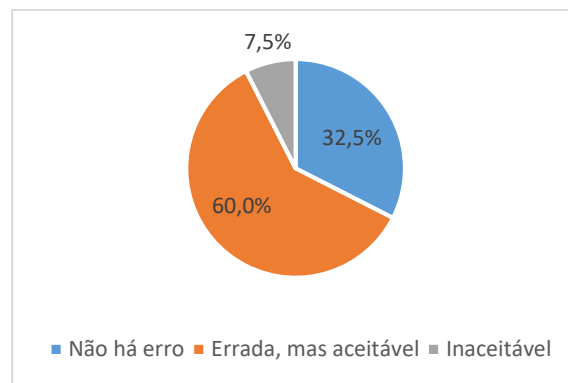


Gráfico 56: Questão ER-2-PQ, "Os dois candidatos que eu olhei os currículos são ótimos"

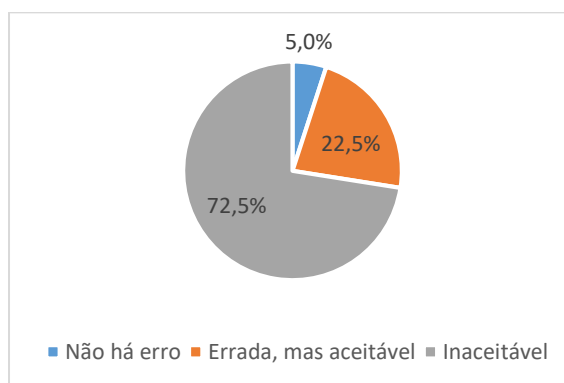


Gráfico 57: Questão ER-3-PQ, "Os diretores que a reunião deles foi cancelada já tinham embarcado"

Ao observarmos esses três gráficos em conjunto, podemos identificar claramente o padrão de percepção dos informantes em relação às três variantes da relativização: no caso da variante padrão em ER-1-PQ, "O cliente cuja compra eu fechei gastou mais de 10 mil reais!", 75% dos informantes pensaram se tratar de uma oração "sem erro". Esse número cai para 32,5% em ER-2-PQ, "Os dois candidatos que eu olhei os currículos ___ são ótimos", e para apenas 5% no caso de ER-3-PQ, "Os diretores que a reunião deles foi cancelada já tinham embarcado". Por outro lado, o índice de marcações como "inaceitável" é de apenas 2,5% para a primeira questão e de 72,5% para a última. Até o momento, esse foi o maior índice de reprovação entre as questões com a variável "estratégia de relativização".

Os gráficos seguintes se referem às orações ER-1-PR, ER-2-PR e ER-3-PR, as contrapartes [– animadas] das orações anteriores:

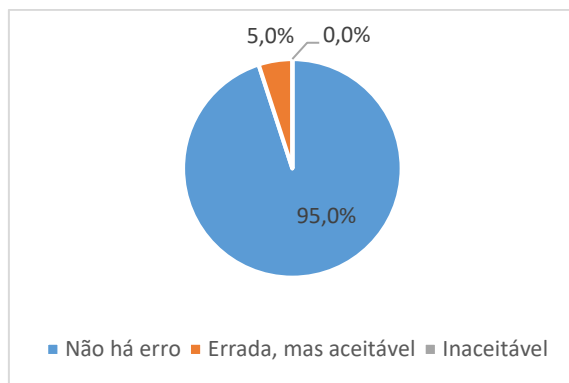


Gráfico 58: Questão ER-1-PR, "Aquele é o telefone cuja linha está com defeito"

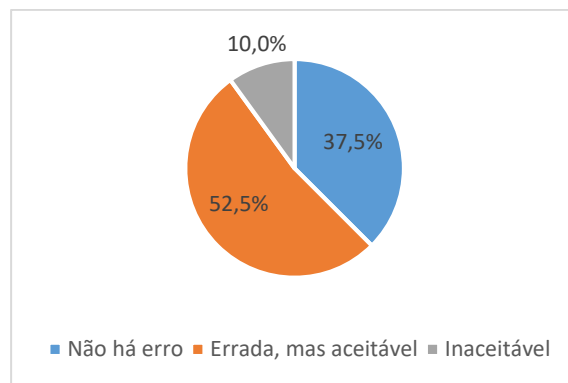


Gráfico 59: Questão ER-2-PR, "O caminhão que eu consertei a roda já estragou de novo"

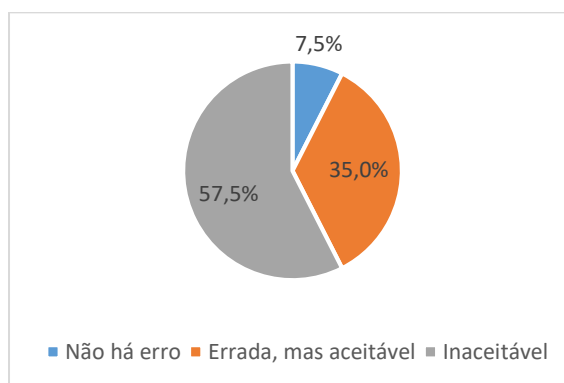


Gráfico 60: Questão ER-3-PR, "Os produtos que a validade deles já venceu foram jogados fora"

O mesmo padrão se verifica. Comparando as orações anteriores com essas, vemos que o índice de aceitação de ER-1-PR, "Aquele é o telefone cuja linha está com defeito", foi muito maior que o de ER-1-PQ, tanto que nenhum informante a considerou "inaceitável". Comparando as questões ER-2-PR, "O caminhão que eu consertei a roda já estragou de novo", e ER-2-PQ, vemos que seus resultados foram muito próximos. Já no caso de ER-3-PR, "Os produtos que a validade deles já venceu foram jogados fora", a versão com antecedente [–animado] teve um índice de marcações "inaceitável" menor, embora o número total de informantes que a consideraram "errada" seja praticamente o mesmo.

Depois de analisar todas as questões com a variável "estratégias de relativização" em separado, podemos reunir os resultados encontrados em um quadro para comparar os números com mais facilidade. Começamos pelas orações **padrão**:

Tabela 15: Comparação das orações padrão da variável "estratégias de relativização"

| | Relativa Objeto Indireto | | Relativa Adjunto Adverbial | | Relativa Adjunto Adnominal Genitivo | |
|-----------------------|--------------------------|---------------------|----------------------------|---------------------|-------------------------------------|---------------------|
| | Anteced. [+animado] | Anteced. [-animado] | Anteced. [+animado] | Anteced. [-animado] | Anteced. [+animado] | Anteced. [-animado] |
| | ER-1-NQ | ER-1-NR | ER-1-OQ | ER-1-OR | ER-1-PQ | ER-1-PR |
| Não há erro | 72,5% | 85% | 35% | 67,5% | 75% | 95% |
| Errada, mas aceitável | 25% | 10% | 60% | 27,5% | 22,5% | 5% |
| Inaceitável | 2,5% | 5% | 5% | 5% | 2,5% | 0% |
| TOTAL | 100% | 100% | 100% | 100% | 100% | 100% |

Em geral, as orações com antecedentes [+ animados] foram as que mais receberam percepções negativas. Isso nos faz supor que a variante padrão (com movimento da preposição) aparenta ser errada para parte dos falantes cultos, quando ela se encaixa em uma relativa com um antecedente [+ animado]. Também de forma geral, as relativas de adjunto adverbial tiveram os maiores índices de reprovação, com a ER-1-OQ ([+ animada]) atingindo 65% de marcações "errada", apesar de ser uma oração que segue as regras da norma padrão. Nossa interpretação desse fato é que alguns falantes, apesar de altamente escolarizados, ainda não adquiriram o algoritmo padrão de criação de orações relativas.

A tabela a seguir traz os resultados para as orações **não padrão** que já analisamos:

Tabela 16: Comparação das orações não padrão da variável "estratégias de relativização"

Tabela 16, seção 1

| | Relativa Objeto Indireto | | | | Relativa Adjunto Adverbial | | | |
|-----------------------|--------------------------|---------|-------------------------|---------|----------------------------|---------|-------------------------|---------|
| | Antecedente [+ animado] | | Antecedente [- animado] | | Antecedente [+ animado] | | Antecedente [- animado] | |
| | ER-2-NQ | ER-3-NQ | ER-2-NR | ER-3-NR | ER-2-OQ | ER-3-OQ | ER-2-OR | ER-3-OR |
| Não há erro | 65% | 10% | 80% | 7,5% | 57,5% | 22,5% | 27,5% | 15% |
| Errada, mas aceitável | 32,5% | 45% | 17,5% | 60% | 37,5% | 35% | 65% | 45% |
| Inaceitável | 2,5% | 45% | 2,5% | 32,5% | 5% | 42,5% | 7,5% | 40% |
| TOTAL | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% |

Tabela 16, seção 2

| | Relativa Adjunto Adnominal Genitivo | | | |
|-----------------------|-------------------------------------|---------|-------------------------|---------|
| | Antecedente [+ animado] | | Antecedente [- animado] | |
| | ER-2-PQ | ER-3-PQ | ER-2-PR | ER-3-PR |
| Não há erro | 32,5% | 5% | 37,5% | 7,5% |
| Errada, mas aceitável | 60% | 22,5% | 52,5% | 35% |
| Inaceitável | 7,5% | 72,5% | 10% | 57,5% |
| TOTAL | 100% | 100% | 100% | 100% |

Podemos perceber primeiramente que, das duas variantes não padrão, a relativa cortadora em todos os casos supera sua concorrente em índice de aprovação e marcações "sem erro". Ou seja: a **relativa copiadora** é, de fato, a *mais saliente*. No entanto, há diferenças significativas nas opiniões quando analisamos as três funções sintáticas testadas: nas relativas de objeto direto a reprovação das cortadoras foi muito baixa: apenas 35% e 20% consideraram as versões [+ animada] e [- animada], respectivamente, como "erradas"; esses números saltam para 42,5% e 72,5% no caso das relativas de adjunto adverbial; e para 67% e 62,5% no caso das relativas genitivas. Em outras palavras, no caso da variante cortadora, o contexto de **adjunto adnominal genitivo** é o que *mais salienta* sua natureza não padrão.

Essa discrepância entre as três funções de relativas que analisamos não se aplica à variante copiadora: sua reprovação foi constante e girou em torno de 90%, exceto para a versão [+ animada] da relativa de adjunto adverbial, cujo índice de marcações "erradas" foi de 77,5%.

Por fim, analisemos as duas questões extras incluídas no questionário para a variável "estratégias de relativização":

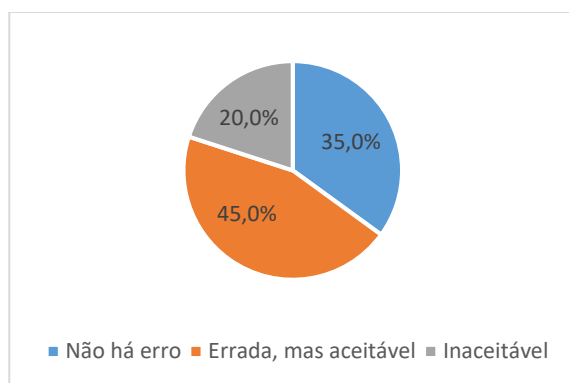


Gráfico 61: Questão ER-QE-1, "Esse é o livro que eu te falei sobre"

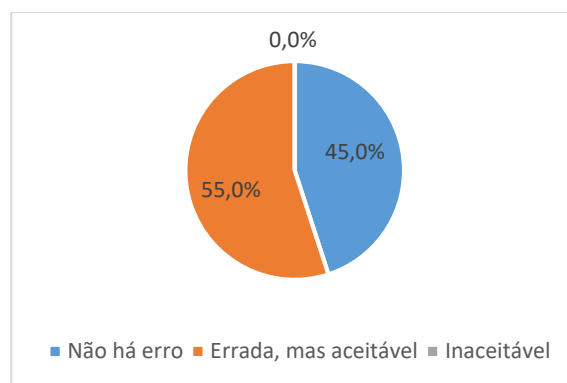


Gráfico 62: Questão ER-QE-2, "Foi essa ideia que nós votamos contra"

Como esperávamos, boa parte dos informantes não considerou essas sentenças "erradas": 35% no primeiro caso, 45% no segundo. Dos que de fato enxergaram uma impropriedade gramatical, 45% consideraram a primeira oração "aceitável" e 55% deram a mesma marcação para a segunda, que não obteve sequer um voto de "inaceitável". Esses números demonstram que essa construção não padrão em que não há movimento da preposição nem seu apagamento está de fato dentro das possibilidades de relativização do PB. É preciso realizar pesquisas sobre em quais condições essa relativa aparece e como seu uso se distribui entre os falantes.

4 CONSIDERAÇÕES FINAIS

O objetivo deste trabalho foi o de analisar a percepção dos falantes cultos do PB no que diz respeito a uma série de regras variáveis da língua portuguesa, com o intuito de descobrir quais elementos dessas regras variáveis são mais salientes para esses falantes, isto é, mais se sobressaem e podem causar atitudes de estigmatização e prestígio. Divisamos três objetivos específicos: identificar e discutir qual variável do PB (dentre as três que escolhemos para análise) é a mais saliente; identificar e discutir qual(is) variante(s) dessas variáveis são mais salientes; e identificar e discutir quais contextos linguísticos mais influenciam na saliência dessas variantes. A hipótese que nos guiou foi a de que certas construções não padrão, por uma série de motivos que se referem a fenômenos estruturais da língua, não seriam percebidas pelos falantes cultos como "erros", enquanto que as mesmas construções, em contextos estruturais diferentes, seriam prontamente apontadas como impropriedades gramaticais. Ao mesmo tempo, buscamos identificar as atitudes de aceitação ou não dos falantes cultos a respeito dessas variáveis e variantes.

Para testar a percepção dos falantes cultos, aplicamos questionários com "testes de percepção" a 40 informantes altamente escolarizados. Os questionários traziam dezenas de orações, algumas respeitando as regras da norma padrão e outras não, e pedimos que os informantes as marcassem como "sem erro", "erradas, mas aceitáveis" ou "inaceitáveis". As variáveis testadas principais foram a "concordância verbal", a "retomada anafórica por objeto direto" e as "estratégias de relativização".

Os resultados obtidos demonstraram que: primeiramente, e comprovando nossa hipótese, de fato há certas construções *não padrão* em português que não são percebidas pelos falantes cultos como "erradas", isto é, destoantes das regras da norma padrão. Podemos destacar nesse sentido a concordância com sujeito composto e [- animado] *posposto* ao verbo (80% dos informantes não reconheceram o "erro"), a retomada anafórica por objeto direto [- animado] por meio da categoria vazia (62,5% dos informantes não enxergaram o "erro" em estruturas simples e 82,5% em estruturas de predicativo) e a relativização de objetos indiretos por meio de relativas *cortadoras* (65% dos informantes não enxergaram o "erro" com antecedentes [+ animados] e 80% com antecedentes [- animados]). Esses contextos parecem "camuflar" as variantes não padrão. Por outro lado, outros contextos parecem *salientar* essas mesmas variantes: para citar exemplos que se relacionam aos que já apresentamos, a concordância com sujeito composto e [- animado] *preposto distante* ao verbo (95% dos informantes perceberam

o "erro"), a retomada anafórica por objeto direto [+ animado] por meio da categoria vazia (92,5% dos informantes perceberam o "erro" tanto em estruturas simples quanto de predicativo) e a relativização de objetos indiretos por meio de relativas copadoras (90% dos informantes enxergaram o "erro" com antecedentes [+ animados] e 92,5% com antecedentes [- animados]).

Destacamos também o nível de aceitação das impropriedades gramaticais pelos falantes cultos: mesmo ao serem capazes de identificar "erros" nas orações, de acordo com a norma padrão, foi alta a aceitação desses erros. Citamos como exemplos a concordância verbal com sujeitos compostos [+ animados] pospostos ao verbo (47,5% de aceitação), *todos* os tipos de retomada anafórica por objeto direto (para todas as orações, a porcentagem de marcações "errada, mas aceitável" foi muito maior que a de "inaceitável", com média de 63,75% para estrutura simples, 61,87% para estruturas complexas e 50% para estrutura com predicativo) e a relativização por meio de relativas genitivas cortadoras (60% de aceitação para as [+ animadas] e 52,5% para as [- animadas]). Por outro lado, houve baixo índice de aceitação para todos os casos não padrão de concordância verbal (média geral de 24,4%), para a retomada anafórica por objeto direto [- animado] em estrutura de predicativo (17,5%) e para relativização de objeto direto por relativa cortadora com antecedente [- animado] (17,5%). Isso nos leva a crer que, em geral, os falantes cultos estão conscientes da variação linguística e de que há diversas possibilidades de expressão que, mesmo condenadas pelos comandos paragramaticais, são legítimas e aceitáveis até mesmo em interações formais, não sofrendo estigmatização.

Respondendo ao primeiro objetivo proposto (*identificar e discutir qual variável do PB é a mais saliente*), vimos em 3.1 que as três principais variáveis receberam as seguintes porcentagens de marcações "errada" para suas orações não padrão: 85,39% para a concordância verbal, 70,77% para o objeto direto anafórico e 69,37% para as estratégias de relativização. Primeiramente, devemos notar que os falantes cultos estão muito atentos aos problemas gramaticais da língua portuguesa e conseguem identificar impropriedades para a norma culta com alguma facilidade. Em segundo lugar, vemos que essa atenção está voltada principalmente à variável da concordância verbal, sendo esta, portanto, a variável mais saliente dentre as três escolhidas para análise. Vimos também em 3.1 que houve um grande grau de rejeição às orações não padrão com essa variável (59,79%), o que não se verifica com as outras variáveis (12,11% para objeto direto anafórico e 27,08% para estratégias de relativização). Isso nos autoriza a dizer que a concordância verbal, além de ser a mais saliente, é também a variável mais estigmatizada pelos falantes cultos. Como dissemos em 3.1, uma possível explicação para esse fato é que a falta de concordância é um fenômeno que ocorre tipicamente com falantes menos

escolarizados, ao passo que em outros fenômenos variáveis, como as estratégias de relativização, as variantes padrão são usadas também por falantes cultos.

Respondendo ao segundo objetivo proposto (*identificar e discutir qual(is) variante(s) dessas variáveis são mais salientes*), vimos que, das duas variáveis estudadas que têm mais de duas variantes, isto é, as variáveis ternárias "retomada anafórica por objeto direto" e "estratégias de relativização", apenas no caso das estratégias de relativização temos o caso de uma variante muito mais saliente que as outras. A variante "relativa cortadora" não foi percebida como um "erro" por uma média de 72,5% dos informantes nas relativas de objeto direto, 42,5% nas relativas de adjunto adverbial e 35% nas relativas genitivas, contra média de 8,75%, 18,75% e 6,25% das relativas copiadoras, respectivamente. Da mesma forma, dos informantes que perceberam que se tratava de uma impropriedade gramatical, a grande maioria conceituou os casos de relativas cortadoras como "aceitáveis", ao passo que a variante copiadora teve mais rejeição que aceitação em quase todos os contextos (cf. Tabela 16). Dessa forma, podemos dizer que os falantes cultos elegeram a relativa cortadora como uma variante "neutra", que, apesar de agramatical pela norma padrão, é aceitável. Já a variante "relativa copiadora" é estigmatizada e tratada como inaceitável na maioria dos casos, sendo assim a variante mais saliente dessa variável.

Esse resultado parece concordar com a hipótese de Tarallo (1983). Segundo o autor, em meados do século XIX, o PB passou a aceitar pronomes nulos (categorias vazias) em orações principais (como exemplo disso temos o objeto direto nulo, que estudamos nesta pesquisa), fenômeno que se estendeu às orações relativas e outras subordinadas. Essa mudança permitiu o aparecimento da estratégia de relativização cortadora, que substituiu abruptamente a relativa padrão e predominou no século XX. Já a relativa copiadora existe há muito tempo no sistema da língua, e sempre teve uma posição marginal, sendo usada em último caso. Sendo assim, e considerando que muitos falantes atuais consideram a relativa padrão pernóstica, é natural que a relativa cortadora se apresente como alternativa viável e "neutra" do ponto de vista sociolinguístico, sendo muito mais aceita que sua contraparte copiadora.

As duas variantes não padrão da variável "objeto direto anafórico" tiveram um comportamento muito instável. Como vimos em 2.1.2.2, na produção linguística dos falantes brasileiros, a variante "anáfora zero" tem uma posição similar à ocupada pela relativa cortadora: é uma solução elegante para quem não deseja usar a retomada padrão por pronome clítico nem a retomada por pronome lexical. Nos nossos dados, a anáfora zero foi menos apontada como "errada" na estrutura simples (com uma vitória esmagadora de 62% contra 15% no caso de

objetos [– animados]), mas foi menos apontada como "errada" no caso da estrutura complexa. Esse quadro se desajusta mais ainda se olharmos para as orações com estrutura de predicativo: para as que têm objeto [+ animado], o pronome lexical não foi visto como erro por 47,5% dos informantes, *versus* 7,5% da categoria vazia; já para as orações com objeto [– animado], a categoria vazia recebeu 82,5% de marcações "sem erro" contra 37,5% do pronome lexical. Sendo assim, não há nada nos dados obtidos que nos autorize a eleger uma ou outra variante como mais saliente.

Para responder ao terceiro problema proposto (*identificar e discutir quais contextos linguísticos mais influenciam na saliência dessas variantes*), vamos analisar os dados das três variáveis em conjunto. Para as três variáveis, testamos a influência que antecedentes [± animados] teriam na percepção dos informantes. As únicas variáveis em que esse fator demonstrou influenciar sistematicamente todos os tipos de orações foram a concordância verbal e as estratégias de relativização. Uma hipótese de Vieira (1997) para o caso da concordância verbal diz que o traço [+ animado] favorece a concordância por conta da possibilidade de o sujeito constituir agente da oração e se correlacionar com o valor semântico típico dos verbos, que é o de expressão de uma ação. Dessa forma, ao se buscar a percepção dos informantes, a falta de concordância com sujeitos [+ animados] fica mais saliente, por desconsiderar essa associação semântica básica. Para a variante "estratégias de relativização", o fator "animacidade" só tem relevância quando combinado a outros fatores.

Para a variável "concordância verbal", vimos ainda que os fatores "sujeito preposto" e "sujeito simples" salientam mais a variante não padrão, não havendo muita diferença entre o fato de o sujeito preposto estar distante ou não do verbo. Como comentamos em 2.1.2.1, os falantes brasileiros tendem a considerar como objeto tudo que está posposto ao verbo, o que explicaria a baixa percepção da falta de concordância no caso de sujeitos pospostos. No caso da variável "retomada anafórica por objeto direto", percebemos que a estrutura com predicativo tem um poder maior de "camuflar" as variantes não padrão, o que significa que as estruturas simples e complexas salientam mais essas variantes. Por último, referimos à Tabela 16 para concluir que as relativas de adjunto adverbial e adjunto adnominal genitivas são as que menos receberam marcações de "sem erro", o que significa que elas são as que mais salientam as estruturas não padrão. De modo geral, esses fatores não tiveram tanta influência na percepção dos informantes, mas mostram-se importantes quando divisados junto a outros fatores ou quando analisamos cada variante em separado.

Em relação às limitações desta pesquisa, temos em mente que as análises feitas não exauriram todos os fenômenos observados nem podem ser consideradas definitivas devido a limitações do *corpus* utilizado. Ao optarmos por trabalhar com três grandes variáveis e diversas variáveis menores, não previmos a quantidade enorme de dados que poderiam ser extraídos do *corpus*; mesmo assim, não podemos dizer com certeza que o *corpus* representa definitivamente a percepção dos falantes cultos, tendo em vista que utilizamos apenas uma questão para cada combinação de fatores condicionadores, e que os próprios fatores condicionadores em si já foram limitados a partir de uma lista muito maior de possibilidades. Talvez se mais orações ou orações diferentes fossem utilizadas, os resultados tivessem sido distintos dos nossos.

Mesmo assim, consideramos que várias informações importantes foram extraídas dos dados que reunimos. Várias hipóteses se confirmaram, mas mais importante foram as não confirmadas, pois os resultados diversos apontam que o senso comum pode estar equivocado. Como possibilidades de continuação desta pesquisa, destacamos o aprofundamento da análise com os dados já conseguidos, talvez com a obtenção de novos dados para fins de comparação.

REFERÊNCIAS

- AMARAL, M. P. D. A síncope em proparoxítonas: uma regra variável. In: BISOL, L.; BRESCANCINI, C. R. **Fonologia e Variação: recortes do português brasileiro**. Porto Alegre: EDIPUCRS, 2002.
- AZEREDO, J. C. S. D. **Fundamentos de gramática do português**. São Paulo: Jorge Zahar, 2002.
- BAGNO, M. **Dramática da língua portuguesa: tradição gramatical, mídia e exclusão social**. São Paulo: Edições Loyola, 2000.
- BAGNO, M. **A norma oculta: língua e poder na sociedade brasileira**. São Paulo: Parábola Editorial, 2003.
- BAGNO, M. **Português ou Brasileiro?: um convite à pesquisa**. São Paulo: Parábola Editorial, 2004.
- BAGNO, M. **A língua de Eulália: novela sociolinguística**. São Paulo: Contexto, 2006.
- BAGNO, M. **Preconceito linguístico: o que é, como se faz**. São Paulo: Edições Loyola, 2007.
- BAGNO, M. **Não é errado falar assim!: em defesa do português brasileiro**. São Paulo: Parábola Editorial, 2009.
- BAGNO, M. **Gramática Pedagógica do Português Brasileiro**. São Paulo: Parábola, 2011.
- BECHARA, E. **Moderna Gramática Portuguesa**. Rio de Janeiro: Lucerna, 2006.
- BORBA, F. D. S. **Introdução aos estudos linguísticos**. Campinas: Pontes, 1991.
- BORTONI-RICARDO, S. M. **Educação em língua materna: a Sociolinguística na sala de aula**. São Paulo: Parábola Editorial, 2004.
- BORTONI-RICARDO, S. M. **Manual de Sociolinguística**. São Paulo: Contexto, 2014.
- CÂMARA JÚNIOR, J. M. **Dispersos**. Rio de Janeiro: Lucerna, 2004.
- CASTILHO, A. T. D. **Nova gramática do português brasileiro**. São Paulo: Contexto, 2010.
- CEGALLA, D. P. **Novíssima gramática da língua portuguesa**. São Paulo: Cia. Editora Nacional, 1990.
- CHOMSKY, N. **Syntactic Structures**. Nova Iorque: Mouton de Gruyter, 2002 [1957].
- CIRÍACO, L.; CANÇADO, M. Inacusatividade e inergatividade no PB. **Cadernos de Estudos Linguísticos**, Campinas, v. 46, n. 2, jul./dez. 2004.
- CLOPPER, ; PISONI,. Perception of Dialect Variation: Some Implications for Current Research and Theory in Speech Perception. **Research on Spoken Language Processing**, Indiana, v. 25, p. 270-289, 2002.
- COELHO, I. et al. **Sociolinguística**. Florianópolis: LLV/CCE/UFSC, 2010.

COELHO, I. L.; MONGUILHOTT, I. D. O. E. S.; SEVERO, C. G. **Norma linguística do português no Brasil**. Florianópolis: LLV/CCE/UFSC, 2014.

CORRÊA, L. D. J. D. D. **A realização do objeto direto anafórico na fala do araguiense: mudança em progresso**. Cáceres: UFMT [dissertação de mestrado], 2014.

CORRÊA, V. R. **O objeto direto nulo no português brasileiro**. Campinas: Unicamp [dissertação de mestrado], 1991.

CORRÊA, V. R. **Oração relativa: o que se fala e o que se aprende no português do Brasil**. Campinas: Unicamp [tese de doutorado], 1998.

COSERIU, E. **Teoría del lenguaje y lingüística general**. Madrid: Editorial Gredos, S. A., 1967.

CUNHA, C.; CINTRA, L. **Nova gramática do português contemporâneo**. Rio de Janeiro: Nova Fronteira, 2001.

CYRINO, S. M. L. **O objeto nulo no português do Brasil: um estudo sintático-diacrônico**. Campinas: Unicamp [tese de doutorado], 1994.

DIAS, A. L. C. **Processo de palatização no português: Lagoa da Pedra e Canabrava - TO**. Goiânia: UFG [Dissertação de Mestrado], 2009.

DUARTE, E. L.; TARALLO, F. Processos de mudança linguística em progresso: a saliência vs. não saliência de variantes. **Ilha do Desterro**, Florianópolis, v. 20, p. 44 a 58, 1988.

DUARTE, M. E. Termos da oração. In: VIEIRA, S. R.; BRANDÃO, S. F. **Ensino de gramática: descrição e uso**. São Paulo: Contexto, 2007.

DUARTE, M. E. L. Clítico acusativo, pronome lexical e categoria vazia no português do Brasil. In: TARALLO, F. **Fotografias sociolinguísticas**. Campinas: Pontes: Editora da Universidade Estadual de Campinas, 1989.

FARACO, C. A. **Norma culta brasileira: desatando alguns nós**. São Paulo: Parábola Editorial, 2008.

GALVES, C. M. C. **Ensaio sobre as gramáticas do português**. Campinas: Editora da Unicamp, 2001.

GARRETT, P. Language attitude. In: LLAMAS, C.; MULLANY, L.; STOCKWELL, P. **The Routledge companion to Sociolinguistics**. Londres: Routledge, 2007.

GNERRE, M. **Linguagem, escrita e poder**. São Paulo: Livraria Martins Fontes Editora, 1991.

HAUY, A. B. **Da Necessidade de uma Gramática-padrão da Língua Portuguesa**. São Paulo: Ática, 1983.

HORA, D. D. Atitude: um conceito teórico, um conceito de vida. **Revista do GELNE**, Natal/RN, v. 14, n. Número especial, p. 367-386, 2012.

ILARI, R.; BASSO, R. **O português da gente: a língua que estudamos, a língua que falamos**. São Paulo: Contexto, 2006.

- LABOV, W. **Padrões Sociolinguísticos**. Tradução de Marcos Bagno; Maria Marta Pereira Scherre e Caroline Rodrigues Cardoso. São Paulo: Parábola Editorial, 2008 [1972].
- LEMLE, M.; NARO, A. J. **Competências básicas do português**. Rio de Janeiro: MOBRAL: Fundação Ford, 1977.
- LEMLE, M.; NARO, A. J. A concordância de número no português do Brasil: um caso típico de variação inerente. In: HORA, D. D. **Diversidade linguística no Brasil**. João Pessoa: Ideia, 1997.
- LOBATO, L. M. P. **Sintaxe gerativa do português**: da teoria padrão à teoria da regência e ligação. Belo Horizonte: Vigília, 1986.
- LOPES, E. **Fundamentos da Linguística contemporânea**. São Paulo: Cultrix, 1995.
- LOPES, L. W. **Preferências e atitudes dos ouvintes em relação à variação linguística regional no telejornalismo**. João Pessoa: UFPB [Tese de doutorado], 2012.
- LUCCHESI, D.; ARAÚJO, S. A teoria da variação Linguística. **Vertentes do português popular do estado da Bahia**, 2016. Disponível em: <<http://www.vertentes.ufba.br/a-teoria-da-variacao-linguistica>>. Acesso em: 25 Janeiro 2016.
- LUFT, C. P. **Ensino e aprendizado de língua materna**. São Paulo: Globo, 2007.
- LYONS, J. **Introdução à linguística teórica**. São Paulo: Cultrix, 1979.
- LYONS, J. **Lingua(gem) e Linguística**: uma introdução. Rio de Janeiro: Guanabara S.A., 1987.
- MENDES, R. B. SP2010 – Construção de uma amostra da fala paulistana. Projeto regular apresentado à FAPESP (Processo FAPESP 2011/09278-6). **SP2010 – Construção de uma amostra da fala paulistana**, 2011. Disponível em: <http://projetosp2010.fflch.usp.br/sites/projetosp2010.fflch.usp.br/files/MENDES_2011-SP2010-ProjRegularFAPESP.pdf>. Acesso em: 01 Junho 2016.
- MILROY, J. The ideology of the standard language. In: LLAMAS, C.; MULLANY, L.; STOCKWELL, P. **The Routledge companion to Sociolinguistics**. Londres: Routledge, 2007.
- MIOTO, C.; SILVA, M. C. F.; LOPES, R. E. V. **Novo manual de sintaxe**. Florianópolis: Insular, 2005.
- MOLLICA, M. C. Fundamentação teórica: conceituação e delimitação. In: MOLLICA, M. C.; BRAGA, M. L. **Introdução à Sociolinguística**: o tratamento da variação. São Paulo: Contexto, 2013.
- MOLLICA, M. C.; BRAGA, L. M. **Introdução à Sociolinguística**: o tratamento da variação. São Paulo: Contexto, 2013.
- MONGUILHOTT, I. D. O. E. S. **Variação na concordância verbal de terceira pessoa do plural na fala dos florianopolitanos**. Florianópolis: UFSC [Dissertação de Mestrado], 2001.
- MONTE, A. **Concordância verbal e variação**: uma fotografia sociolinguística da cidade de São Carlos. Araraquara: UNESP [dissertação de mestrado], 2007.

- MONTEIRO, J. L. **Para compreender Labov**. Petrópolis: Vozes, 2008.
- MULLANY, L. Speech communities. In: LLAMAS, C.; MULLANY, L.; STOCKWELL, P. **The Routledge companion to Sociolinguistics**. Londres: Routledge, 2007.
- NARO, A. J. O dinamismo das línguas. In: MOLLICA, M. C.; BRAGA, M. L. **Introdução à Sociolinguística: o tratamento da variação**. São Paulo: Contexto, 2013.
- NEVES, M. H. D. M. **Gramática de Usos do Português**. São Paulo: Editora da UNESP, 2000.
- OLIVEIRA, M. D. S. **Concordância verbal de terceira pessoa do plural em Vitória da Conquista: variação estável ou mudança em progresso**. Salvador: UFBA [dissertação de mestrado], 2005.
- PAIVA, M. D. C. A variável gênero/ sexo. In: MOLLICA, M. C.; BRAGA, M. L. **Introdução à Sociolinguística: o tratamento da variação**. São Paulo: Contexto, 2013.
- PEREIRA, ; COELHO, I. O uso variável das formas anafóricas no acusativo. **Letrônica**, Porto Alegre, v. 6, n. 1, Jan./Jun. 2013.
- PERINI, M. A. **Para uma nova gramática do português**. São Paulo: Ática, 1985.
- PERINI, M. A. **Gramática descritiva do português**. São Paulo: Ática, 2005.
- PERINI, M. A. **Gramática do Português Brasileiro**. São Paulo: Parábola, 2010.
- RÁCZ, P. **Salience in Sociolinguistics: a quantitative approach**. Berlin/New York: De Gruyter Mouton, 2013.
- ROBERTS, I. O português brasileiro no contexto das línguas românicas. In: ROBERTS, I.; KATO, M. A. **Português brasileiro: uma viagem diacrônica**. Campinas: Editora da Unicamp, 1993.
- ROCHA LIMA, C. H. D. **Gramática normativa da língua portuguesa**. Rio de Janeiro: José Olympio, 2011.
- RODRIGUES, D. D. A. **A concordância verbal na fala urbana de Rio Branco**. Campinas: Unicamp [dissertação de mestrado], 1997.
- ROSA, P. R. D. S. **Uma introdução à pesquisa qualitativa em ensino**. Campo Grande: Editora da UFMS, 2015.
- SANTOS, J. D. C. **O comportamento das estratégias de relativização na escrita culta jornalística brasileira**. Rio de Janeiro: UFRJ [dissertação de mestrado], 2015.
- SCHERRE, M. M. P. **Doa-se lindos filhotes de poodle**. São Paulo: Parábola Editorial, 2005.
- SILVA, B. G. S. G. D. **"O caminhão que eu trabalhava com ele subia qualquer ladeira": um estudo sobre a gramaticalização do que**. Rio de Janeiro: UFRJ [dissertação de mestrado], 2005.
- SILVA, G. M. D. O. E. Coleta de dados. In: MOLLICA, M. C.; BRAGA, M. L. **Introdução à Sociolinguística: o tratamento da variação**. São Paulo: Contexto, 2013.

- SILVA, L. A. D. Projeto NURC: histórico. **Linha d'Água**, São Paulo, n. 10, p. 83-90, julho 1996.
- SILVA, M. C. V. D. F. **O objeto direto anafórico no dialeto rural afro-brasileiro**. Salvador: UFBA [dissertação de mestrado], 2004.
- SILVA, T. C. **Fonética e Fonologia do Português**: roteiro de estudos e guia de exercícios. São Paulo: Contexto, 2012.
- TARALLO, F. **Relativization Strategies in Brazilian Portuguese**. Philadelphia: Univ. of Pennsylvania [tese de doutorado], 1983.
- TARALLO, F. **A pesquisa sociolinguística**. São Paulo: Ática, 1986.
- VALE, M. J. Q. **Estratégias de relativização na fala de adultos maranhenses**. Rio de Janeiro: UFRJ [tese de doutorado], 2014.
- VIEIRA, S. R. A não-concordância em dialetos populares: uma regra variável. **Graphos**, João Pessoa, v. 1, n. 2, p. 115-133, 1997.
- VIEIRA, S. R. Concordância verbal. In: VIEIRA, S. R.; BRANDÃO, S. F. **Ensino de gramática**: descrição e uso. São Paulo: Contexto, 2007.
- VOTRE, S. J. Relevância da variável escolaridade. In: MOLLICA, M. C.; BRAGA, M. L. **Introdução à Sociolinguística**: o tratamento da variação. São Paulo: Contexto, 2013.
- WARDHAUGH, R. **An introduction to Sociolinguistics**. Nova Iorque: Blackwell, 2010.
- WEINREICH, U.; LABOV, W.; HERZOG, M. I. **Fundamentos empíricos para uma teoria da mudança linguística**. Tradução de M. Bagno. São Paulo: Parábola Editorial, 2006 [1968].