

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL**

**MELISSA OSTERLUND FERREIRA**

**A VARIÇÃO DA PREPOSIÇÃO *PARA* NA FALA DE CURITIBA E DE  
FLORIANÓPOLIS PELOS DADOS DO VARSUL**

**PORTO ALEGRE**

**2018**

MELISSA OSTERLUND FERREIRA

**A VARIAÇÃO DA PREPOSIÇÃO *PARA* NA FALA DE CURITIBA E DE  
FLORIANÓPOLIS PELOS DADOS DO VARSUL**

Dissertação apresentada ao Programa de Pós-graduação do Instituto de Letras da Universidade Federal do Rio Grande do Sul como requisito parcial para a obtenção do título de Mestra em Estudos da Linguagem: Fonologia e Morfologia.

Prof<sup>ª</sup> Dra. Valéria Neto de Oliveira Monaretto  
Orientadora

**PORTO ALEGRE**

**2018**

## CIP - Catalogação na Publicação

Ferreira, Melissa Osterlund

A Variação da Preposição PARA na Fala de Curitiba e de Florianópolis pelos Dados do VARSUL / Melissa Osterlund Ferreira. -- 2018.

114 f.

Orientadora: Valéria Neto de Oliveira Monaretto.

Dissertação (Mestrado) -- Universidade Federal do Rio Grande do Sul, Instituto de Letras, Programa de Pós-Graduação em Letras, Porto Alegre, BR-RS, 2018.

1. Preposição PARA. 2. Fonologia. 3. Teoria da Variação. 4. VARSUL. I. Monaretto, Valéria Neto de Oliveira, orient. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

MELISSA OSTERLUND FERREIRA

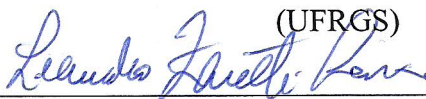
**A VARIAÇÃO DA PREPOSIÇÃO *PARA* NA FALA DE CURITIBA E DE  
FLORIANÓPOLIS PELOS DADOS DO VARSUL**

Dissertação apresentada ao Programa de Pós-graduação do Instituto de Letras da Universidade Federal do Rio Grande do Sul como requisito parcial para a obtenção do título de Mestra em Estudos da Linguagem: Fonologia e Morfologia.

APROVADA: Porto Alegre, 18 de Maio de 2018.

Prof. Dr. Leandro Zanetti Lara

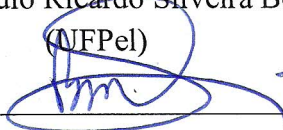
(UFRGS)



---

Prof. Dr. Paulo Ricardo Silveira Borges

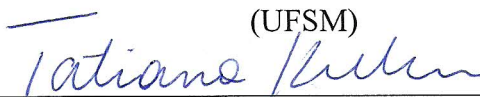
(UFPel)



---

Prof<sup>ª</sup> Dr<sup>ª</sup> Tatiana Keller

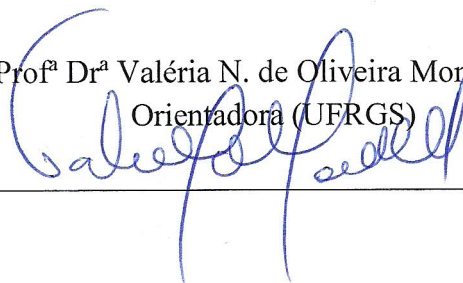
(UFSM)



---

Prof<sup>ª</sup> Dr<sup>ª</sup> Valéria N. de Oliveira Monaretto

Orientadora (UFRGS)



---

## AGRADECIMENTOS

Aos Governos de Luiz Inácio Lula da Silva e de Dilma Rousseff, pela política de cotas para egressos do ensino público, que fez a diferença para o meu ingresso na Letras, em 2008.

À CAPES, pela bolsa concedida. Ao Projeto VARSUL, por disponibilizar os dados.

À UFRGS, pela oportunidade de cursar Graduação e Mestrado em uma instituição pública que possibilitou grande amadurecimento para minha vida nos últimos dez anos.

À minha orientadora, Prof<sup>a</sup> Valéria Monaretto, pela orientação sempre tão dedicada.

À minha mãe, Lígia, pelo apoio incondicional e por ter viabilizado toda a minha trajetória acadêmica até hoje.

Ao meu irmão, Juliano, por ter me ensinado as letras na infância e pelas trocas de ideias na vida adulta.

Ao meu pai, Sérgio, pela ajuda com traduções de textos lidos para o Mestrado e pelo apoio de sempre.

Aos meus colegas e amigos do Mestrado, Bruna, Giuliano, Luana, Samanta, César e Mônica, pelo companheirismo dos momentos felizes, apesar de difíceis, que passamos juntos.

Às minhas amigas da alma, Carol, Maurin e Kétina, também professoras, por existirem na minha vida. Ao amigo Víctor, pela parceria e por todo o apoio tecnológico.

À Rê, pela amizade cada vez maior e pelo yoga cada vez melhor.

À Fernanda e à Carol Formigoni, por terem entrado rápido e fundo na minha vida.

Ao Walter, por todo o apoio e companheirismo.

À Júlia, pela amizade que surgiu na IC e por agora ser também minha aluna.

Ao Roberto, pelas aulas de inglês, que me salvaram no início do Mestrado.

À *FM Cultura – 107.7*, rádio pública em ameaça de extinção, pelas incontáveis horas de Música Popular Brasileira de qualidade, que acompanharam o meu trabalho.

A todos que compreenderam a necessidade do meu distanciamento nestes dois anos e meio de pesquisa.

*“Cambia el rumbo el caminante  
Aunque esto le cause daño  
Y así como todo cambia  
Que yo cambie no es extraño [...]”*

*(Mercedes Sosa)*

## RESUMO

Esta pesquisa trata de descrever e analisar a variação da preposição *para* na fala de Curitiba (PR) e de Florianópolis (SC) a partir de 32 entrevistas que integram o Banco de Dados do Projeto Variação Linguística no Sul do Brasil (VARISUL). A análise e quantificação dos dados basearam-se na Teoria da Variação e Mudança Linguística de Weinreich, Labov e Herzog. Verificou-se que há influência de variáveis linguísticas e sociais na ocorrência desta variação na fala destes indivíduos, comparando-se estes resultados com outras investigações que abordaram o mesmo fenômeno. Observou-se que *pra* é a variante favorita na cidade catarinense e na paranaense, *pa* apresenta poucas realizações nas duas localidades, e a forma padrão *para* mostrou-se insignificante em termos de ocorrência, de modo geral. Por isso, *pra* se tornou a variável dependente do trabalho, sendo examinada conforme grupos de fatores linguísticos e sociais. Os resultados indicaram favorecimento da forma reduzida *pra* e baixo uso da forma padrão *para*, como ocorreu na maioria das pesquisas realizadas em outras cidades brasileiras, o que pode caracterizar uma situação de mudança em curso no português brasileiro. Os informantes mais velhos com escolaridade mais alta e da cidade de Curitiba (PR) manifestaram preferência pela variante *pra*. A hipótese de que a preposição reduzida realizaria processo de juntura (sândi) com a palavra seguinte não se confirmou, pois a maior parte dos dados nesse contexto constitui-se de palavras iniciadas por vogais coronais tônicas, características que impedem a realização de sândi, conforme Bisol (2002).

**Palavras-chave:** Preposição *para*; Fonologia; Teoria da Variação; VARISUL.

## RESUMEN

Esta investigación trata de describir y analizar la variación de la preposición *para* en el habla de Curitiba (PR) y de Florianópolis (SC) a partir de 32 entrevistas que integran el Banco de Datos del Proyecto Variação Linguística no Sul do Brasil (VARISUL). El análisis y la cuantificación de los datos se basaron en la Teoría de Variación y Cambio Lingüístico de Weinreich, Labov y Herzog. Se verificó que hay influencia de variables lingüísticas y sociales en la ocurrencia de esta variación en el habla de estos individuos, al compararse estos resultados con otras investigaciones que abordaron el mismo fenómeno. Se observó que *pra* es la variante favorita en la ciudad catarinense y en la paranaense, *pa* presenta pocas realizaciones en las dos localidades, y la forma patrón *para* se mostró insignificante en términos de ocurrencia, de modo general. Por esa razón, *pra* se convirtió en la variable dependiente del trabajo, siendo examinada conforme grupos de factores lingüísticos y sociales. Los resultados indicaron favorecimiento de la forma reducida *pra* y poco uso de la forma patrón *para*, como ocurrió en la mayoría de las pesquisas realizadas en otras ciudades brasileñas, lo que puede caracterizar una situación de cambio lingüístico en curso en el portugués brasileño. Los informantes mayores con nivel más alto de escolaridad y de la ciudad de Curitiba (PR) manifestaron preferencia por la variante *pra*. La hipótesis de que la preposición reducida realizaría proceso de juntura (*sândi*) con la palabra siguiente no se confirmó, pues la mayor parte de los datos en ese contexto se constituye de palabras iniciadas por vocales coronales tónicas, características que impiden la realización de *sândi*, según Bisol (2002).

**Palabras-clave:** Preposición *para*; Fonología; Teoría de la Variación; VARISUL.



## LISTA DE GRÁFICOS

GRÁFICO 1 – Distribuição Geral dos Dados. ....	55
GRÁFICO 2 – Comparação dos Dados entre Pesquisas Variacionistas da variável <i>para</i> . ....	57
GRÁFICO 3 – Aplicações de <i>pra</i> por informante na cidade de Curitiba. ....	75
GRÁFICO 4 – Aplicações de <i>pra</i> por informante na cidade de Florianópolis. ....	76

## LISTA DE TABELAS

TABELA 1 – A influência da escolaridade na escolha de <i>pa</i> em relação a <i>pra</i> na fala de Porto Alegre. Fonte: Maya (2004). .....	26
TABELA 2 – A influência do contexto morfológico seguinte na escolha de <i>pa</i> em relação a <i>pra</i> na fala de Porto Alegre. Fonte: Maya (2004). .....	27
TABELA 3 – Ilustração de células sociais conforme nossa amostra. ....	44
TABELA 4 – Estratificação dos informantes por sexo, idade e escolaridade. ....	50
TABELA 5 – Distribuição de dados por variante em cada trabalho. ....	55
TABELA 6 – Aplicação Geral de <i>pra</i> – primeira rodada. ....	58
TABELA 7 – Aplicação Geral de <i>pra</i> – com amalgamações e exclusões. ....	58
TABELA 8 – Variante <i>pra</i> considerando <i>Contexto Morfológico Seguinte</i> – com amalgamações. ....	60
TABELA 9 – Variante <i>pra</i> considerando <i>Contexto Fonológico Seguinte</i> – com amalgamações. ....	63
TABELA 10 – Cruzamento entre as variáveis <i>Contexto Fonológico Seguinte</i> e <i>Tonicidade da Sílabas Seguinte</i> . ....	65
TABELA 11 – Variante <i>pra</i> considerando <i>Paralelismo</i> – com amalgamações. ....	67
TABELA 12 – Variante <i>pra</i> considerando <i>Número de Sílabas do Item Seguinte</i> – com amalgamações. ....	68
TABELA 13 – Cruzamento entre as variáveis <i>Número de Sílabas do Item Seguinte</i> e <i>Tonicidade da Sílabas Seguinte</i> . ....	69
TABELA 14 – Aplicações de <i>pra</i> considerando a <i>Idade</i> . ....	71
TABELA 15 – Aplicações de <i>pra</i> considerando a <i>Escolaridade</i> . ....	73
TABELA 16 – Aplicações de <i>pra</i> considerando a <i>Cidade</i> . ....	74
TABELA 17 – Cruzamento entre as variáveis <i>Idade</i> e <i>Escolaridade</i> . ....	78
TABELA 18 – Cruzamento entre as variáveis <i>Idade</i> e <i>Cidade</i> . ....	79
TABELA 19 – Cruzamento entre as variáveis <i>Escolaridade</i> e <i>Cidade</i> . ....	80

## LISTA DE ILUSTRAÇÕES

ILUSTRAÇÃO 1 – Diagrama arbóreo da hierarquia prosódica. Fonte: Bisol (2010). ....	24
ILUSTRAÇÃO 2 – Transcrição de entrevista. Fonte: Projeto VARSUL. ....	45
ILUSTRAÇÃO 3 – Análise acústica da preposição <i>para</i> . Fonte: Toneli (2009). ....	82

## SUMÁRIO

<b>1 INTRODUÇÃO .....</b>	<b>12</b>
<b>2 A PREPOSIÇÃO “PARA”: DO LATIM AO PORTUGUÊS .....</b>	<b>14</b>
<b>3 O ESTATUTO PROSÓDICO DA PREPOSIÇÃO “PARA” .....</b>	<b>20</b>
<b>4 ESTUDOS VARIACIONISTAS SOBRE A PREPOSIÇÃO “PARA” NO PORTUGUÊS BRASILEIRO .....</b>	<b>25</b>
4.1 O ESTUDO DE MAYA (2004) .....	25
4.2 O ESTUDO DE LUCENA (2006) .....	28
4.3 O ESTUDO DE SILVA (2010) .....	31
4.4 O ESTUDO DE GAZOLA (2008) .....	33
4.5 O ESTUDO DE FELGUEIRAS ( <i>apud</i> MAYA, 2004, p. 16) .....	35
4.6 O ESTUDO DE FERREIRA (2014) .....	36
4.7 BREVES COMENTÁRIOS SOBRE A METODOLOGIA DOS ESTUDOS SOBRE A PREPOSIÇÃO .....	37
<b>5 METODOLOGIA .....</b>	<b>40</b>
5.1 TEORIA DA VARIAÇÃO E MUDANÇA LINGUÍSTICA .....	40
<b>5.1.1 Operacionalização do Modelo .....</b>	<b>43</b>
5.2 AMOSTRA .....	49
<b>5.2.1 Variáveis Controladas .....</b>	<b>51</b>
5.2.1.1 Variável Dependente .....	51
5.2.1.2 Variáveis Independentes Linguísticas .....	51
5.2.1.3 Variáveis Independentes Extralinguísticas .....	53
5.3 OBJETIVOS .....	54
5.4 HIPÓTESES .....	54
<b>6 RESULTADOS DA ANÁLISE DE DADOS .....</b>	<b>55</b>
6.1 DISTRIBUIÇÃO DAS VARIANTES DA PREPOSIÇÃO <i>PARA</i> .....	55
6.2 ANÁLISE ESTATÍSTICA DA VARIANTE <i>PRA</i> E DISCUSSÃO DOS RESULTADOS .....	57

6.3 VARIÁVEIS LINGUÍSTICAS SIGNIFICATIVAS .....	59
6.3.1 <i>Contexto Morfológico Seguinte</i> .....	59
6.3.2 <i>Contexto Fonológico Seguinte</i> .....	62
6.3.3 <i>Paralelismo</i> .....	66
6.3.4 <i>Número de Sílabas do Item Seguinte</i> .....	68
6.4 VARIÁVEIS EXTRALINGUÍSTICAS OU SOCIAIS .....	71
6.4.1 <i>Idade</i> .....	71
6.4.2 <i>Escolaridade</i> .....	72
6.4.3 <i>Cidade</i> .....	74
6.4.4 <i>Quantidade de PRA por informante</i> .....	75
6.5 CRUZAMENTOS ENTRE FATORES EXTRALINGUÍSTICOS OU SOCIAIS .....	77
6.5.1 <i>Idade X Escolaridade</i> .....	78
6.5.2 <i>Idade X Cidade</i> .....	79
6.5.3 <i>Escolaridade X Cidade</i> .....	80
6.6 BREVE DISCUSSÃO DOS RESULTADOS .....	81
<b>7 CONCLUSÃO</b> .....	<b>84</b>
<b>REFERÊNCIAS</b> .....	<b>86</b>
<b>ANEXO A – Rodada Eneária</b> .....	<b>89</b>
<b>ANEXO B – Rodada Binária sem Amalgamações</b> .....	<b>99</b>
<b>ANEXO C – Rodada Binária com Amalgamações</b> .....	<b>109</b>
<b>ANEXO D – Quadro de codificação de variáveis</b> .....	<b>114</b>

## 1 INTRODUÇÃO

Este trabalho pretende investigar o comportamento variável da preposição *para* na fala das cidades de Curitiba (PR) e de Florianópolis (SC). Com base na Teoria da Variação e Mudança Linguística, nos dados coletados pelo Projeto Variação Linguística do Sul do Brasil (VARSUL) e na Teoria Fonológica, serão analisados possíveis fatores linguísticos e extralinguísticos que possam condicionar a realização das variantes *para*, *pra* ou *pa*. Os resultados serão comparados com os estudos sobre o fenômeno em outras cidades brasileiras. Espera-se que os dados possam contribuir para a compreensão de fenômenos de variação, inerentes a todas as línguas.

A escolha por Curitiba e Florianópolis, dentre as demais cidades do corpus do VARSUL, deveu-se ao fato de serem capitais da região sul do Brasil ainda não contempladas com estudos sobre a preposição *para*. A amostra desta pesquisa está composta por 32 entrevistas, que foram estratificadas pelas variáveis sociais *idade*, *escolaridade*, *sexo* e *cidade*. Também foram controladas algumas variáveis linguísticas: *contexto morfológico seguinte*, *contexto fonológico seguinte*, *tonicidade da sílaba seguinte*, *paralelismo*, *número de sílabas do item seguinte*, *posição em relação a pausas* e *processo de sândi com a sílaba seguinte*. Juntamente com o estudo de Maya (2004), sobre Porto Alegre (RS), espera-se poder contribuir para a pesquisa sobre este fenômeno nessa região. Já a escolha da variável *para* deu-se por parecer que receba a interferência de aspectos prosódicos, segundo Bisol (2000). Anteriormente, investigamos a mesma regra variável na cidade de Londrina (PR), em Ferreira (2014). O presente estudo busca ampliar, em relação à pesquisa anterior, a amostra e o aprofundamento da análise.

Considerando-se os dados de alguns estudos descritivos sobre a variação da preposição *para/pra/pa*, espera-se que a forma reduzida *pra* seja mais utilizada do que as variantes *para* e *pa* na fala das cidades de Curitiba (PR) e de Florianópolis (SC). Além disso, espera-se que a escolha de *pra*, variante típica da língua falada coloquial, possa motivar-se por um grau de escolaridade mais baixo do indivíduo. Por fim, espera-se confirmar que a variante *para*, com acento na penúltima sílaba, altera seu estatuto prosódico quando é reduzida para *pra* ou *pa*, tornando-se uma sílaba sem acento e uma forma dependente. Como tal, a preposição reduzida sofreria um processo de juntura (chamado sândi) com a palavra seguinte que inicie com vogal átona.

Este trabalho está organizado como segue. No primeiro capítulo, apresentamos o fenômeno estudado: a preposição *para*, suas variantes e um percurso histórico, desde sua origem no latim até a passagem para o português. No segundo capítulo, fazemos algumas considerações sobre o estatuto prosódico da preposição *para*. No terceiro capítulo, constam alguns estudos variacionistas sobre a preposição *para* no português falado em cidades brasileiras. No quarto capítulo, apresentamos a metodologia adotada para esta pesquisa e descrevemos a amostra, os grupos de variáveis controladas e a variável dependente. O quinto capítulo apresenta os resultados das três variantes *para/prá/pá* e de *prá* como variável dependente. A seguir, trazemos o percentual de aplicação dessa variante em cada grupo de fatores e as comparações com outros trabalhos. O sexto e último capítulo traz as conclusões sobre os resultados e as hipóteses da pesquisa. Por fim, anunciamos as próximas etapas da investigação.

## 2 A PREPOSIÇÃO “PARA”: DO LATIM AO PORTUGUÊS

Camara Jr. (1979, p. 117) explica que a categoria e o mecanismo das preposições do latim surgiu de uma associação do advérbio ao complemento, o que provocou uma redundância com a desinência de caso acusativo ou ablativo para expressar a relação do complemento com o verbo. A “preposição”, como sugere o nome, era definida como uma partícula adverbial que precedia o complemento. Krischke (1939, p. 7) afirma que as preposições, inicialmente, eram apenas “advérbios especializados”, já que eram usadas em compostos verbais para expressar noções de tempo, lugar, modo etc. A diferença, de acordo com o autor, estaria na característica de partícula conectiva da preposição.

Almeida (1960, p. 290) define as preposições como palavras invariáveis que têm a função de ligar o complemento à palavra completada. A origem de sua nomenclatura, conforme o autor, vem do latim *prae* (diante de) mais *positionem* (posição), que significa “pôr na frente de uma palavra outra que a completa”. Para ele, a diferença da preposição para a conjunção – que também tem função de conectivo – é que a primeira liga *palavras*, enquanto a segunda liga *orações*. O autor ilustra essa definição com os exemplos comparativos que seguem: em “livro *de* Pedro”, *de* é uma preposição, pois liga a palavra *Pedro* à palavra *livro*. Já em “Pedro foi *mas* não voltou”, *mas* é uma conjunção, pois estabelece uma ligação entre a oração “Pedro foi” e a oração “não voltou”.

Para Krischke (1939, p. 7), “preposição” significa, etimologicamente, “colocar antes”, “antepor”. A origem da palavra seria a partir de dois vocábulos latinos: *prae* (*por*, *per*), que dá a ideia de “antes”, e *ponere*, verbo que significa “colocar”. Para o autor, a preposição não apresenta “ideia alguma apreciável”, apenas relações entre nomes e pronomes.

Uma característica das línguas românicas, aos moldes do latim, segundo Camara Jr. (1979, p. 176), é utilizar um sistema de preposições que estabeleça relações de subordinação entre os constituintes de uma oração. O latim já possuía um sistema parecido, em que certos complementos eram subordinados a um verbo. A indicação de subordinação se dava por meio de desinências de caso acusativo ou ablativo, que vinham ligadas ao nome complemento, como em *ire ad forum* (“ir ao foro”), *ire in silvam* (“ir pela selva adentro”) e *ire sub freta* (“ir pelas ondas abaixo”), exemplos em que se acumulam as desinências de caso e as preposições.



Ainda de acordo com Camara Jr. (1979, p. 176), também é frequente o uso de preposições para estabelecer relações nominais de subordinação de um substantivo a outro nas línguas românicas. No latim, isso era possível devido à flexão de genitivo no substantivo subordinado, em referência ao substantivo núcleo do sintagma. Com a remodelação do sistema latino, o português passou a cumprir essa função com a preposição *de*. Camara Jr. ilustra a situação com os seguintes exemplos:

Assim, a uma frase latina como – “*Historia est uita memoriae*”, em que *memoriae* está com a flexão de genitivo (-ae, dos nomes tema em -a), para indicar a sua subordinação ao substantivo *uita* “vida”, corresponde em português – “(A) *história é (a) vida d(a) memória*”, com a mesma subordinação indicada pela preposição *de*. (CAMARA JR., 1979, p. 176).

Assim, diz Camara Jr. (1979), criou-se, no latim, uma duplicidade de representação dessa relação de subordinação, que foi dissolvida no português, e nas línguas românicas em geral, pela “vitória” da preposição. O sistema de casos morfológicamente marcados, por conseguinte, foi eliminado das línguas românicas, devido à força da preposição em delimitar o elo de dependência. A preposição, então, assumiu a marca de subordinação ao verbo de forma exclusiva e pôs fim à redundância vinda do latim. Ao mesmo tempo, passou a funcionar também junto aos complementos verbais.

Krischke (1939, p. 9) explica que o valor das preposições cresceu devido à queda gradativa das flexões gramaticais. Com a importância que as preposições adquiriram, chegaram ao ponto de determinar relações de casos. Segundo Krischke (1939, p. 10), no latim clássico, elas limitavam-se a reforçar as relações de acusativo e ablativo; entretanto, no latim popular da idade média, passaram a suprir a falta de desinências nos casos oblíquos, com exceção do caso reto – equivalente ao nominativo do latim.

Apesar de as preposições terem começado a exercer um papel mais importante, segundo Camara Jr. (1979, p. 177), grande parte delas se perdeu do latim ao português após um processo de simplificação e economia. Houve uma redistribuição de emprego de algumas dessas partículas, que depois foram substituídas por outras.

Conforme Camara Jr. (1979, p. 177), o sistema de preposições em português funciona em dois planos: o locativo e o dinâmico. O locativo refere-se à localização no espaço e no tempo. Preposições como *em*, *entre* e *sobre*, por exemplo, indicam uma localização estática. Já o plano dinâmico funciona com preposições como *de*, que expressa noção de

“afastamento”, *por* (do latim *pro*, que surgiu a partir de *per*), que representa “percurso” e *a* (do latim *ad*), que expressa “direção”.

A preposição *para* (*pera*, em português arcaico), assunto deste trabalho, é fruto da aglutinação de *per* e *ad*, ocorrida no latim vulgar imperial, quando até então marcava um percurso com direção definida (CAMARA JR., 1979). Já no português, essa indicação de direção se torna mais complexa, com as noções complementares de chegada e permanência. O sentido das preposições “a” e “para” nos enunciados *ir a Paris* e *ir para Paris* são diferentes, pois, no primeiro caso, há uma significação geral de direção, enquanto no segundo caso acrescenta-se o sentido de se estabelecer no local.

A. G. da Cunha (1991, *apud* POGGIO, 2002, p. 239) destaca que a variante antiga *pera* é muito comum em textos medievais. Apenas por volta do século XVII é que essa variante foi substituída pela forma atual *para*.

Krischke, com base no que chama de “autoridades filológicas” de sua época, trata da força atrativa das preposições:

Pertencem as preposições, à semelhança dos advérbios e das conjunções, à categoria de partículas invariáveis que, por via de regra, influem na colocação dos pronomes pessoais tônicos *me, te, se, o, lhe, nos, vos, os, lhes*, e os atraem {sic}. (KRISCHKE, 1939, p. 27).

Para Krischke, algumas preposições teriam mais poder atrativo do que outras. Sobre a preposição “para”, o autor afirma que ela faz parte de um grupo de partículas em relação às quais “varia a prática dos bons escritores, mas na maioria dos casos, nota-se a tendência de atraírem {sic} as variações pronominais átonas.” (KRISCHKE, 1939, p. 27).

Almeida (1960, p. 292) inclui *para* no grupo de preposições essenciais, aquelas que não exercem outra função senão a de preposições. Ao tratar da função da preposição *a*, o autor comenta que *para*, muitas vezes, desempenha papel idêntico ao daquele conectivo, exemplificando a situação com as frases “Disse *a* (*para*) você” e “Dei *ao* (*para o*) irmão”.

Para Almeida (1960, p. 291), as preposições não têm significação intrínseca e sim relativa, dependendo do verbo que a acompanhe. Os verbos *Ir* e *vir*, combinados com a preposição *a*, expressam transitoriedade de movimento, enquanto indicam permanência ou destino quando utilizados com a preposição *para*. Na frase *Vamos à Argentina*, tem-se a ideia

de “ir a passeio, ir para voltar”, ao passo que *Carlos foi para os Estados Unidos* denota que “foi fixar residência” neste país. (ALMEIDA, 1960, p. 292).

Krischke (1939, p. 126) menciona alguns adjetivos, substantivos e verbos que seriam utilizados com a preposição *para*. Dentre os verbos, o autor lista os que seguem:

Apelar	Entrar	Preparar-se
Aprender	Escolher	Relegar
Arremeter	Estudar	Remeter
Colher	Forcejar	Retirar-se
Comprar	Habilitar	Seguir
Concorrer	Investir	Servir
Contribuir	Ir	Talhar
Correr	Navegar	Tender
Deitar	Nomear	Tomar
Descambar	Oferecer-se	Traduzir
Designar	Olhar	Transportar
Emigrar	Pender	Ungir (em)
Encaminhar	Predispor	Voltar-se

Talvez o uso de alguns verbos citados por Krischke (1939) seja de uma variedade linguística específica, pois é estranho, no português brasileiro, que se use “aprender para...” ou “deitar para...”.

Segundo Camara Jr. (1979, p. 178), as principais diferenças entre o sistema latino de preposições e o atual são oriundas da riqueza do quadro de partículas que havia em latim. No caso da noção de direção, o emprego de *a* foi ampliado: *ire in silvam* passou a *ir à floresta*, em português moderno – apesar de a língua coloquial brasileira preferir a construção nos moldes do latim: *ir na floresta*, sendo *na* equivalente a *em + a*. No entanto, *a* teve sua difusão um pouco limitada pelo surgimento de *para*, preposição de criação românica.

Camara Jr. (1979, p. 179) explica que a noção de percurso em latim era expressa pela preposição *per*. Por muito tempo, essa forma coexistiu com *pro*, que marcava “posição dianteira” – e de cuja forma é possível que tenha surgido *por*, em português. Por fim, *por* acabou substituindo totalmente a forma *per*.

Krischke (1939, p. 123) traz uma lista das principais relações estabelecidas pela preposição *para*, com algumas frases de exemplo para cada tipo de relação, das quais retiramos o que segue exposto abaixo:

- Direção: *Olhar para as dificuldades;*
- Movimento: *Partir para a Europa;*
- Fim: *Estudo para saber;*
- Tempo futuro: *Das duas para as três horas;*
- Proporcionalidade: *2 está para 4, assim como 4 está para dois;*
- Avaliação aproximada: *Vai para três dias;*
- Afirmação válida: *O casamento é para os Romanos um contrato puramente civil;*
- Capacidade: *Tal trabalho não é para êle.*

De acordo com Camara Jr. (1979, p. 23), algumas preposições, na passagem do latim para o português, transformaram-se devido ao desaparecimento das declinações. Como seria esperado, com a remodelação morfossintática, também houve uma mudança nos padrões sintáticos da língua, o que veio a fixar uma nova tipologia frasal. As preposições, com sua função de partículas conectivas, e a ordem dos vocábulos na frase, passaram a cumprir o papel de estabelecer novas relações sintáticas, função antes exercida pelas flexões nominais. Assim, desinências de *nominativo* e *acusativo*, que até o momento complementavam palavras sem ordem fixa na oração, deram lugar à ordem gramatical românica, com os papéis de sujeito e objeto direto posicionalmente fixos. Camara Jr. (1979, p. 24) traz o exemplo de *O menino viu o lobo*, frase que no latim ocorria de forma variável, como *puer vidit lupum*, *lupum puer vidit*, *lupum vidit puer* ou *vidit lupum puer*.

No latim, as preposições já eram utilizadas junto às desinências de acusativo e ablativo para caracterizarem alguns complementos verbais, conforme destaca Camara Jr. (1979, p. 24). Então, elas acabaram gerando um desgaste fonético dessas desinências e atraíram para si a verdadeira indicação sintática. Além disso, também assumiram o lugar de complementos que eram indicados por dativo ou genitivo. No caso do dativo, a desinência de objeto indireto deu lugar à construção românica *dar a (o) menino*, devido ao uso da preposição *ad* com o nome no acusativo, que funcionava exclusivamente como complemento de direção no latim clássico. Camara Jr. ilustra esse caso com a frase latina *ire ad templum*, em português, *ir a (o)*

*templo*. No caso do genitivo, a subordinação de um nome a outro foi substituída, aos poucos, pela construção de preposição com ablativo: *de Deo munus* (dádiva de Deus) em vez de *Dei munus*, do latim clássico.

Grande parte das gramáticas históricas e demais obras de referência (ALI, COUTINHO, WILLIAMS etc.) consultadas para embasar esta investigação sobre a preposição *para* apresentou maior dedicação a outras preposições ditas essenciais da língua portuguesa, como *de*, *a*, *em* etc. Espera-se que as informações expostas anteriormente possam contribuir para pesquisas sobre esta preposição.

Neste capítulo, tratamos da evolução da preposição *para* do latim ao português. No próximo capítulo, procederemos à exposição do estatuto prosódico da preposição *para* de forma a contextualizar o tema escolhido para este trabalho.

### 3 O ESTATUTO PROSÓDICO DA PREPOSIÇÃO “PARA”

As preposições são consideradas palavras invariáveis que relacionam dois termos de uma oração, sendo o primeiro explicado ou completado pelo segundo. A relação que se estabelece entre essas palavras ligadas por uma preposição pode implicar movimento ou uma situação daí resultante (Vou *a* Roma/ Todos saíram *de* casa). As preposições com ideia de movimento levam em conta um ponto limite A em referência ao qual o movimento será de aproximação (B→A), como nas situações Vou *a* Roma e Foi *para* o Norte, ou de afastamento (A→C), como em Venho *de* Roma/Saíram *pela* porta (CUNHA; CINTRA, 2014, p. 571).

Ambas as preposições *a* e *para* apresentam o valor de movimento. A diferença entre elas é que *para* se distingue de *a* por comportar um traço significativo que implica maior destaque do ponto de partida, com predominância da ideia de direção sobre o término do movimento (CUNHA; CINTRA, 2014, p. 587). Essa perspectiva, contudo, não é clara no uso e compreensão atual que se faz das duas preposições em questão, pois fala-se tanto Vou *ao* cinema como Vou *no* cinema, com o mesmo sentido.

Em termos fonéticos e fonológicos, as preposições *a* e *para* distinguem-se por número de sílabas e pelo acento. A preposição *a* é formada por uma só unidade sonora de uma sílaba sem acento. Já a preposição *para* tem duas sílabas, recaindo o acento na primeira sílaba.

Nos estudos estruturalistas, consideram-se as palavras como vocábulos formais ou fonológicos. As unidades formais de uma língua são as *formas livres*, que compõem uma sequência possível de funcionar isoladamente (*proscrever*, *lei*), e as formas presas, que só funcionam ligadas a outras (*pro-* em *proscrever*).

Camara Jr. (1970, p. 69) postula um terceiro conceito para os vocábulos formais, o de *forma dependente*: uma forma que não é livre, pois não estabelece comunicação suficiente, mas também não é presa, porque pode ocorrer tanto à direita quanto à esquerda da forma livre à qual esteja ligada. Exemplos de formas dependentes em português são as partículas proclíticas átonas que podem se unir a verbos em posição de próclise ou ênclise, como em *se fala* ou *fala-se*.

Segundo Camara Jr. (1970, p. 70), as formas *dependentes* são as partículas átonas como o artigo e as preposições, a partícula *que* e outras mais. A forma dependente é um vocábulo formal que faz parte de um vocábulo fonológico. Este tem como característica

possuir acento proeminente. No caso de *guarda-roupa*, a sequência de estruturas forma um só vocábulo fonológico por ser “*rou*” a sílaba proeminente.

Os vocábulos dependentes, que têm a função de se relacionar uns com os outros, estabelecem uma conexão. Por isso são chamados, por Camara Jr. (1970, p. 79), de conectivos. As preposições seriam do tipo subordinativo.

Na teoria gerativista, a relação entre vocábulos é examinada em termos de constituintes e de relação do tipo dominante e dominado. No caso das partículas átonas, referidas por Camara Jr. como formas dependentes, são comumente denominadas como *clíticos*, que são elementos de proeminência acentual fraca, dependentes do acento primário da palavra adjacente à qual se associam (SILVA, 2011, p. 74).

Para Guzzo (2015, p. 103), clíticos são palavras funcionais monossilábicas inacentuadas que sofrem processos fonológicos, como elevação vocálica e sândi externo, característicos de sílabas não proeminentes na língua. Em português brasileiro, segundo a autora, os clíticos apresentam comportamento fonológico e morfossintático aproximado, mantêm-se à esquerda de seu hospedeiro e podem pertencer a várias classes gramaticais, como preposições (*de* e *com*), conjunções (*que* e *se* em *se ele for*), artigos (*o*, *a*) e pronomes (*me* e *se*, em *se chama*). Podem ser pronominais, quando selecionam um único tipo de hospedeiro (o verbo principal) ou não pronominais, quando podem selecionar hospedeiros de qualquer classe gramatical, além de poderem, exclusivamente, formar sequências de clíticos.

Guzzo (2015, p. 105) salienta que alguns clíticos não pronominais podem ser o resultado de cliticização de palavras funcionais dissilábicas proeminentes, desde que não estejam em posição de foco. A forma reduzida da preposição *para*, assunto deste trabalho, parece encaixar-se nessa situação, já que pode sofrer redução silábica, assumindo as formas *pra* ou *pa*. Ou seja, para a autora, *para* não é um clítico, pois apresenta proeminência por ser dissilábico, mas *pra/pa*, sim. (GUZZO, 2015, p. 104).

Outro processo possível em clíticos não pronominais, segundo Guzzo (2015), é a fusão clítica, um processo morfossintático (com consequência fonológica), já que é registrado na ortografia para alguns clíticos não pronominais (*de* e *em*), mas não para outros (*com* e *pra*). Pode ocorrer entre clíticos e artigos definidos e indefinidos (*pra + o = pro*, *pra + a = pra*) e, no caso de *de* e *em*, entre clíticos e alguns pronomes demonstrativos (*de + este = deste*, *em + aquele = naquele*).

Os clíticos do PB devem ser diferenciados de palavras lexicais e de afixos. Para Simioni (2008, p. 1), palavras lexicais são aquelas que pertencem às categorias de nome, verbo e adjetivo, diferentemente de palavras funcionais. Entretanto, de acordo com Bisol (2000), não é tão simples fazer essa diferenciação, já que os clíticos possuem comportamento híbrido, com propriedades de palavras lexicais e de afixos ao mesmo tempo.

Simioni acredita, como Bisol (2000), que clíticos não portam acento; por isso, apoiam-se no acento de palavras adjacentes, além de não ocorrerem isoladamente. Os clíticos têm em comum com os afixos a falta de acento. No entanto, apresentam certa mobilidade na frase, caso dos clíticos pronominais, que podem ser proclíticos ou enclíticos. Essa mobilidade do clítico, para Simioni (2008), foi a razão pela qual Camara Jr. (1970) inseriu os clíticos entre as formas dependentes, que se opõem às formas livres e às formas presas. Segue exemplo dessa mobilidade dos clíticos pronominais, em Bisol (2000):

*Comprei-te um livro.*

*Te comprei um livro.*

Simioni (2008, p. 2) referencia a pesquisa de Bisol (2010) para tratar de propriedades universais dos clíticos: (i) são átonos, (ii) são formas dependentes e (iii) pertencem a diferentes classes morfológicas. Clíticos podem ser artigos definidos, algumas preposições e alguns pronomes pessoais. Todos são palavras funcionais, mas nem todas as palavras funcionais são clíticos, como no caso de palavras que possuem acento.

Segundo Bisol (2000, p. 10), clíticos são formativos difíceis de classificar como palavras independentes ou afixos. Diferem da palavra independente por não serem candidatos a receber acento; de afixos flexionais, por serem formas livres; e de afixos derivativos, por serem sempre periféricos.

O clítico e seu hospedeiro mantêm entre si a relação de dominância que define um constituinte prosódico: o cabeça é a palavra de conteúdo e o dominado é um clítico ou mais de um. Mas é um grupo que pressupõe uma origem sintática, como qualquer frase fonológica. É o menor constituinte frasal. (BISOL, 2000, p. 19).



Os clíticos são identificados prosodicamente, segundo Bisol (2000, p. 10-11) como elementos fracos e são encontrados em diferentes categorias. Podem receber acento enfático como qualquer sílaba de uma cadeia de fala. No entanto, não se trataria de acento primário, lexical, que identifica a palavra fonológica. Bisol (*op. cit.* p. 10) cita alguns exemplos:

*Vou **para** casa* (preposição)

*Pedro toca piano, **mas** João toca violino* (conjunção)

*Ganhou **uma** bicicleta* (determinativo)

*Vejo-**te** amanhã* (pronome)

Conforme Bisol (2010, p. 264), existem dois tipos de clíticos. Os que formam uma só unidade fonológica junto à palavra de conteúdo, como no caso da variante dialetal [te kõnsideru]ω, em que o pronome *te* (clítico) forma uma só *palavra fonológica*<sup>1</sup> com o vocábulo *considero*, não possibilitando a atuação de uma regra de neutralização de átona final. O clítico, nesse caso, constitui, com a unidade adjacente, um só vocábulo.

Já no caso de [tʃi]ω [kõnsideru]ω, o clítico *te* se comporta com certa independência, seguindo as mesmas regras da palavra fonológica, sujeitando-se a regras de neutralização. Bisol assume, assim como Nespor e Vogel (1986, apud Bisol, 2010), que esse segundo tipo de clítico é chamado de *grupo clítico* (locução), pois é uma unidade prosódica que contém um ou mais clíticos e uma só palavra de conteúdo. Esse último parece ser o caso representativo do português, segundo a autora, porque a neutralização é uma forma geral em muitas cidades brasileiras.

---

<sup>1</sup> Palavra Fonológica é uma unidade prosódica assumida ser um domínio de aplicação de fenômenos fonológicos. (CRISTÓFARO-SILVA, 2011, p. 176).

Segundo Bisol (2000, p. 19), o Grupo Clítico (representado pelo símbolo C) é o menor constituinte pós-lexical derivado da sintaxe, formado por um clítico e uma palavra de conteúdo adjacente. Esse elemento está, na hierarquia prosódica, acima dos constituintes *palavra fonológica*, *pé* e *sílaba*, conforme diagrama arbóreo a seguir.

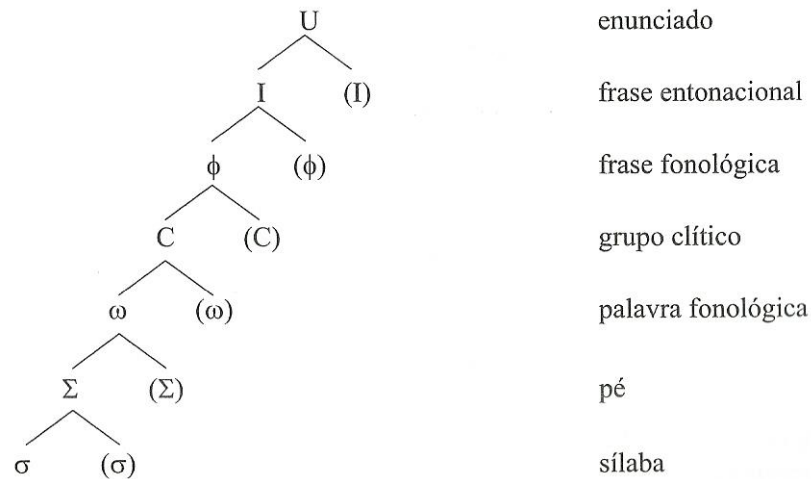


Ilustração 1 – Diagrama arbóreo da hierarquia prosódica. Fonte: Bisol (2010).

Simioni (2008) questiona a existência do grupo clítico, defendida por Bisol (2000, 2010). Segundo a autora, para aceitar a existência deste constituinte, seria necessário identificar um processo que ocorresse apenas no interior de um grupo clítico. Para Bisol (2010), esse processo é a elisão da vogal *e*.

Na próxima seção procederemos à apresentação de outros trabalhos que tratam da variação da preposição *para* no português falado do Brasil.

## 4 ESTUDOS VARIACIONISTAS SOBRE A PREPOSIÇÃO “PARA” NO PORTUGUÊS BRASILEIRO

Alguns estudos variacionistas, baseados em dados da língua falada, têm investigado este caso de variação da preposição *para*. Segundo o modelo teórico da Teoria da Variação, exposto na seção 5 deste trabalho, a variação linguística é considerada inerente à língua, ordenada, sistematizada e condicionada por fatores linguísticos e sociais. No caso da preposição *para*, pode-se ver, nos estudos descritos a seguir, alguns destes fatores condicionadores.

### 4.1 O ESTUDO DE MAYA (2004)

O estudo de Maya (2004), intitulado “*A Variação da Preposição para na Fala de Porto Alegre/RS*”, propõe investigar a variação da preposição *para* na fala portoalegrense a partir da perspectiva da Teoria da Variação. Maya (2004) apresenta uma pesquisa feita com 24 informantes de Porto Alegre entrevistados pelo Projeto Variação Linguística da Região Sul (VARSUL), na década de 1990.

As variáveis sociais controladas foram *sexo*, *escolaridade* e *idade*, e as linguísticas, *contexto morfológico precedente e seguinte*, *contexto fonológico seguinte*, *tonicidade da sílaba seguinte*, *número de sílabas do item seguinte*, *posição em relação a pausas e paralelismo formal*. As hipóteses levantadas por Maya encontram-se nas seções de exposição dos resultados de cada variável estudada pelo autor.

Nos resultados, foram obtidos 2034 dados, sendo 62% de ocorrências da forma *pra*, 36% de *pa* e apenas 2% de *para*. A baixa ocorrência da forma padrão *para* levou Maya (2004) à decisão de se deter à análise binária *pa* em relação a *pra*.

Considerando-se *pa* como variável dependente, o grupo selecionado pelo programa GOLDVARB como mais influente foi a variável social *Escolaridade*. Esta pesquisa adotou a divisão do banco de dados do Projeto VARSUL, adaptando-se às categorias *Primário* (de 1 a 4 anos de estudo), *Secundário* (até 11 anos de estudo) e *Superior* (acima de 11 anos de estudo).

Na Tabela 1, pode-se verificar o papel da escolaridade no uso de *pa* em relação a *pra*:

	N	%	Peso relativo
Primário	253/699	36	0.49
Secundário	212/697	48	<b>0.61</b>
Superior	196/684	29	0.40
<b>Total</b>	741/1992	37	

Input: 0.38                      Significância: 0,004

Tabela 1 – A influência da escolaridade na escolha de *pa* em relação a *pra* na fala de Porto Alegre. Fonte: Maya (2004).

Analisando os dados da Tabela 1, percebe-se que a forma inovadora *pa* é mais recorrente em informantes menos escolarizados, com ensino primário e secundário (36% e 48% respectivamente) do que nos mais escolarizados (29%). Entretanto, analisando-se o peso relativo, apenas o nível secundário (0.61) mostra-se mais favorável ao uso da variante *pa*. A hipótese de que a forma *pa* seria favorecida pelos indivíduos menos escolarizados, portanto, não é confirmada por Maya (2004).

O segundo grupo selecionado pelo programa como mais favorável ao uso da variante *pa* foi a variável linguística *Contexto Morfológico Seguinte*. Primeiramente, Maya havia dividido este grupo de fatores em: substantivo, adjetivo, pronomes pessoais, demais pronomes, verbo e demais classes gramaticais. No entanto, devido a influências semelhantes que algumas dessas classes de palavras exerciam, decidiu amalgamar numerais, advérbios, pronomes pessoais e demais classes gramaticais (artigos, conjunções etc.), passando-se a denominar este fator como “Outros”, e mantendo os fatores “Verbo” e “Nome”. Ao concluir que tanto “Verbo” quanto “Nome” exerceriam igual influência na escolha de *pa* em relação a *pra*, devido ao mesmo percentual (43%) e peso relativo praticamente idênticos, o autor resolveu amalgamar os dois fatores em um só.

Na Tabela 2 a seguir, constam o percentual e o peso relativo de verbos e nomes amalgamados e de outras classes de palavras (artigos, conjunções etc.) para a realização da variante inovadora *pa*:

	N	%	Peso relativo
Verbos e nomes	518/1203	43	<b>0.55</b>
Outros	212/697	30	0.42
<b>Total</b>	730/1900	38	

Input: 0.38      Significância: 0,009

Tabela 2 – A influência do contexto morfológico seguinte na escolha de *pa* em relação a *pra* na fala de Porto Alegre. Fonte: Maya (2004).

Maya (2004, p. 36) conclui, em relação à influência do Contexto Morfológico Seguinte, que a forma reduzida *pa* é favorecida quando é seguida de uma palavra lexical (nomes e verbos, 0.55). Já quando é seguida de uma palavra gramatical, essa variante é desfavorecida (0.42). O autor, portanto, confirma sua hipótese de que a redução de *para* seria favorecida quando um nome ou um verbo seguissem a palavra prosódica resultante da junção entre a preposição e a palavra seguinte.

Outras variáveis que se destacaram na influência do uso de *pa* foram *Sexo*, com os homens favorecendo (0.57) e as mulheres desfavorecendo (0.44) a escolha dessa variante; *Contexto Fonológico Seguinte*, com destaque para as consoantes dentais (0.63), palatais (0.57) e bilabiais (0.57) e *Tonicidade da Sílabas Seguinte*, em que a categoria de sílabas átonas mostrou favorecimento para o uso da preposição reduzida *pa*, com peso relativo de 0.56.

Maya (2004) constata que a forma padrão *para* é muito pouco utilizada em Porto Alegre (RS), sendo *pra* a variante preferida e *pa*, restrita à fala e desconhecida pelas gramáticas tradicionais. Além disso, a preferência dos homens pela forma *pa* poderia indicar que essa variante é desprestigiada pela sociedade. Quanto à idade, variável não selecionada pelo programa estatístico como influente no uso de *pa*, o autor conclui que apenas os informantes mais velhos utilizaram a forma padrão *para*.

#### 4.2 O ESTUDO DE LUCENA (2006)

A pesquisa de Lucena (2006), intitulada “*Elementos para o Estudo da Variação Lingüística na Paraíba*”, propõe analisar o comportamento variável da preposição *para* e suas variantes *pra* e *pa* na cidade de João Pessoa, capital da Paraíba. Para o autor, a variável ternária *para~pra~pa* corresponderia, respectivamente, a forma padrão, forma não padrão e forma inovadora. O modelo teórico-metodológico utilizado, conforme o autor, é a sociolinguística quantitativa.

Para esta investigação, Lucena (2006) utilizou dados do corpus do Projeto VALPB (Variação Linguística no Estado da Paraíba). Foram selecionadas 24 entrevistas de informantes estratificados por sexo, faixa etária e nível de escolarização. As variáveis linguísticas controladas foram *contexto fonológico seguinte*, *posição em relação a pausas*, *paralelismo formal*, *número de sílabas do item seguinte* e *presença de vibrante no item seguinte*.

Lucena (2006, p. 43) afirma que os manuais de gramática tradicionais brasileiros tratam de preposições apenas em relação a sua forma, conteúdo significativo e função relacional, mas poucas incluem suas formas não padrão. O seu estudo busca evidências sobre a influência de aspectos fonológicos, discursivos e sociais na escolha de uma das variantes da preposição *para*. Além disso, procura identificar se alguma dessas variantes pode ter desaparecido do português brasileiro ou se este fenômeno seria apenas uma característica da linguagem oral.

De acordo com Lucena, Faraco & Moura e Bechara seriam mais inovadores do que outros gramáticos por terem registrado a forma não padrão *pra*; no entanto, ignoraram a forma inovadora *pa* – a mais estigmatizada das variantes. O autor também afirma: “Como nem todas as diferenças são sinais de mudança, algumas delas decorrem de características próprias da oralidade em oposição àquelas próprias da escrita.”. (LUCENA, 2006, p. 18).

Lucena (2006, p. 48) levanta três hipóteses sobre o condicionamento de aspectos sociais em seu corpus: falantes mais escolarizados usariam mais a variante padrão *para*, enquanto as variantes não padrão e inovadoras *pra* e *pa* seriam mais usadas pelos analfabetos; mulheres utilizariam mais a forma padrão *para*, devido à tendência do gênero feminino de escolher a forma de prestígio com mais frequência do que os homens; e a última hipótese, de que adultos

mais velhos prefeririam a variante padrão *para* em detrimento das variantes inovadoras *pra* e *pa*, em relação aos mais jovens.

Sobre a influência de fatores linguísticos, Lucena (2006, p. 50) levanta cinco hipóteses. A primeira é que as formas *pra* e *pa* seriam favorecidas por vogal central e por consoantes bilabiais, dentais e alveolares no contexto seguinte. Já a forma padrão *para* seria favorecida por segmentos mais distantes do ponto labial. Para o autor, poderia haver uma tendência de se repetir sons de natureza articulatória semelhante.

A segunda hipótese trata da variável *Paralelismo Formal*. A forma *pa* tenderia a ocorrer mais quando precedida por outra forma semelhante em uma frase, bem como as formas *para* e *pra* também apresentariam a mesma tendência. A terceira hipótese diz que, se a palavra seguinte carrega uma consoante vibrante, haveria maior tendência ao uso de *pa*. Na ausência de vibrante, seria mais provável a ocorrência de *para* ou *pra*. A explicação seria que pode haver certa dificuldade em pronunciar o som vibrante em duas palavras seguidas. A quarta hipótese seria que, quanto maior a palavra seguinte, mais probabilidade de que segmentos fossem apagados – gerando formas não padrão. A quinta e última hipótese defende que a presença de pausas favoreceria a ocorrência de *para*, devido a processos fonológicos condicionados pela velocidade da fala.

Para Lucena (2006, p. 45), as variantes não padrão *pra* e *pa* resultariam de dois processos de supressão de segmentos. O primeiro seria a síncope da vogal central /a/, gerando a queda de uma sílaba. Com essa supressão, a estrutura silábica seria modificada de consoante/vogal (CV) para consoante/consoante/vogal (CCV), como segue:

[para]	----->	[pra]
pa-ra (2 sílabas)		pra (1 sílaba)
(padrão silábico: CV-CV)		(padrão silábico: CCV)

O segundo processo seria uma síncope da vibrante /r/, também resultando em mudança da estrutura silábica:

[pra]	----->	[pa]
pra (1 sílaba)		pa (1 sílaba)
(padrão silábico: CCV)		(padrão silábico: CV)

No último processo, o padrão silábico volta à estrutura CV, que é a mesma de *para*, considerada a mais comum e simples das línguas naturais.

De acordo com Lucena:

A forma padrão *para* é um dissílabo átono e, por isto, o /a/ átono é sujeito à supressão, ocorrendo a queda da vogal central e mudando o padrão silábico CVCV para CCV, resultando na forma não padrão *pra*. A forma *pra*, por sua vez, está cada vez mais sujeita a uma simplificação, com a supressão da vibrante, resultando na forma inovadora *pa*, que se evidencia como específica da fala. (LUCENA, 2006, p. 62).

Nos resultados, foram obtidos 1445 dados, sendo 87 (6%) aplicações de *para*, 655 (45%) de *pra* e 703 (49%) de *pa*. Devido à baixa ocorrência da forma padrão *para*, o autor optou por realizar uma primeira rodada binária. Para isso, amalgamou as variantes não padrão *pra* e *pa* e confrontou-as com *para*.

As variáveis linguísticas que se mostraram mais relevantes para o comportamento variável da preposição *para*, conforme Lucena (2006, p. 80), foram *Paralelismo Formal* e *Contexto Fonológico Seguinte*. No caso do primeiro grupo, as variáveis *pra* e *pa*, quando precedidas por outras formas semelhantes em uma frase, foram escolhidas em 97% (*pra*, com peso relativo 0.68) e 99% (*pa*, com peso relativo 0.84) dos casos. A segunda hipótese linguística do autor, portanto, foi confirmada.

O outro grupo selecionado, *Contexto Fonológico Seguinte*, também apresentou mais influência para a aplicação das formas reduzidas *pra* e *pa*, sendo 99% (0.81) para labiodentais, 96% (0.59) para alveolares e 96% (0.59) para dentais. Lucena (2006, p. 85) sugere que os resultados da variável *Contexto Fonológico Seguinte* parecem indicar uma preferência pelas formas reduzidas *pra* e *pa* quando a palavra seguinte comece por sons próximos ao ponto labial. Como as duas variantes iniciam pela consoante bilabial /p/, que está próxima a esse ponto, pode ser que o contexto fonológico seguinte, quando também inicie por sons próximos ao ponto labial, venha a influenciar uma repetição de sons semelhantes, fenômeno que se chama de aliteração.

As vogais, no entanto, mostraram-se desfavoráveis para a aplicação de *pra* e *pa*: a vogal central obteve peso relativo de 0.31, as vogais posteriores, 0.33 e as vogais anteriores, 0.43. O autor sugere que o fenômeno de aliteração possa influenciar também o comportamento das



vogais, já que a presença de vogais no item seguinte parece favorecer a manutenção da vogal central /a/ na forma *para*. A primeira hipótese linguística, portanto, foi confirmada parcialmente.

Como variáveis sociais que se destacaram, Lucena (2006, p. 96) menciona *Anos de escolarização e Faixa etária*. Segundo o autor, a escolarização teve papel decisivo na escolha das variantes: enquanto os universitários apresentaram peso relativo de 0.91 (13%) para a variante padrão *para*, os analfabetos escolheram as formas reduzidas *pra* e *pa* de maneira quase categórica (0.86 de peso relativo) dos casos. Portanto, Lucena (2006) confirma sua primeira hipótese social.

No caso da faixa etária, o autor observou preferência dos mais jovens (15 a 25 anos) pela variante padrão (0.63, 9%). Para ele, essa situação poderia ser explicada pela imposição da norma padrão feita pela escola. Os informantes de 26 a 49 anos foram os que mais utilizaram *pra/pa*: 0.58 (96%). Para Lucena (2006), tal fato poderia estar relacionado com a influência do que Labov (1966) chamou de “prestígio encoberto” (*covert prestige*): o falante quer usar certas variantes linguísticas, pois assim será aceito em um determinado grupo social. Nessa faixa etária intermediária, de acordo com Lucena (2006), seria maior a pressão social sofrida pelos indivíduos para se identificarem a um grupo, visto que é um momento de maior potencial de trabalho em suas vidas. Após os 50 anos, por fim, a tendência de utilizar a forma padrão retorna (0.48), o que demonstra, segundo o autor, um processo de variação estável, contrariando sua terceira hipótese social.

Lucena (2006) conclui que as variantes linguísticas preferidas do português falado em João Pessoa (PB) são *pra* e *pa*, e que a forma padrão *para* está cada vez mais restrita à língua escrita.

#### 4.3 O ESTUDO DE SILVA (2010)

Na pesquisa “*A Preposição para e suas Variantes no Falar Araguaetinense*”, Silva (2010) tratou de investigar se a variação de *para/pra/pa* ocorreria por motivações linguísticas e/ou sociais na fala de Araguaetins/TO.

A autora utilizou um corpus de 36 entrevistas que formam parte do Projeto VALTINS (Variação Linguística no Estado do Tocantins), cujos informantes foram estratificados

igualmente por sexo, idade e escolaridade. As variáveis linguísticas controladas foram *contexto fonológico seguinte*, *paralelismo formal*, *pausa* e *presença de vibrante no item seguinte*. A principal hipótese do estudo de Silva (2010) é de que a variante padrão *para* seria mais utilizada entre falantes mais escolarizados, enquanto as variantes não padrão *pra* e *pa* seriam mais frequentes entre informantes de média ou baixa escolaridade.

Silva (2010), assim como Lucena (2006), acredita que há preferência de uso das formas *pra/pa* devido à reestruturação silábica de um padrão CVCV, passando por CCV, até novamente chegar em CV – no caso da variante inovadora *pa*.

Em um total de 3210 dados, a forma padrão *para* teve 35 ocorrências (1%); *pra* teve 1852 (54%) e *pa*, 1323 (45%). Devido à insignificância do percentual de *para*, foi feita uma nova rodada binária, considerando-se a variante *pra* como variável dependente, confrontando-a com *pa*. Desta vez, os resultados indicaram, como influentes para a realização da variante *pra*, os seguintes fatores, em ordem de relevância: 1º: *Contexto Fonológico Seguinte* -- dorsal (0.53), vogais anteriores (0.60) e vogais posteriores (0,66); 2º: *Escolaridade* --alta (0.61); 3º: *Faixa Etária* -- + de 49 anos (0.56); 4º: *Sexo* --mulher (0.53).

Na análise da variável *Contexto Fonológico Seguinte*, foram consideradas alterações de velocidade, especialmente em relação a pausas, que seriam um fator inibidor de supressão, por diminuírem a velocidade da fala. Diferentemente das vogais anteriores e posteriores, a vogal central mostrou-se mais favorável à realização da variante *pa*, com peso 0.34.

Pode-se observar uma questão interessante sobre a variável *Faixa Etária* no estudo realizado por Silva (2010). Mesmo que a faixa mais velha (+ de 49 anos) tenha realizado mais a forma *pra*, em segundo lugar está a faixa mais nova, de 15 a 25 anos, e não a intermediária, de 26 a 49 anos. Para a autora, a explicação seria que os mais jovens sofrem influência do ambiente escolar, que costuma apresentar a forma mais padrão (neste caso, trata-se de *pra* em oposição a *pa*). Assim, os mais jovens realizam menos a forma *pa* (0.45) do que os informantes de idade intermediária (0.54), que apresentam o percentual mais alto desta variante.

#### 4.4 O ESTUDO DE GAZOLA (2008)

A pesquisa “*A Estrutura Prosódica da Preposição para*” foi publicada por Gazola em 2008 na revista *Uniletras*, da Universidade Estadual de Ponta Grossa (UEPG). Este trabalho não se desenvolve nos mesmos moldes dos anteriores, tendo por base a variação e o modelo quantitativo de análise. Trata-se de analisar o uso de *para* ou *pra* em alguns informantes quanto à aceitabilidade ou não de emprego dessas formas, por meio de exposição de algumas frases, e de verificar se há algum papel prosódico. O objetivo geral desse estudo foi demonstrar que as palavras funcionais (neste caso específico, as preposições) apresentam características fonológicas diferentes das palavras lexicais. O objetivo específico de Gazola (2008) foi o de estudar a preposição *para*. A autora apresenta a hipótese de que a forma reduzida *pra* não apareceria em final de sentença, mais especificamente, em sentenças do tipo interrogativa completa; ou seja, sem complemento, conforme exemplo abaixo:

Ele foi *para*?

E não

\* Ele foi *pra*?

Para tentar confirmar sua hipótese, Gazola realizou, com seis informantes, um experimento que pretendia descobrir as preferências desses falantes em relação aos dois tipos de enunciados acima. Além disso, a autora se propôs a fazer uma descrição da estrutura prosódica de *para* no contexto acima, a fim de verificar como e em que circunstâncias esta preposição se organizaria dentro das palavras prosódicas da sentença. A partir daí, o objetivo foi o de comprovar que tal palavra funcional poderia aparecer tanto em uma forma fraca de acento quanto em uma forma forte, dependendo de sua posição na frase.

A metodologia utilizada por Gazola (2008) tem por base propostas de Guglielmo Cinque (1993), Elisabeth Selkirk (1995) e Marina Vigário (1999), que tratam de descrever estratégias de focalização na língua, a partir da interação entre Sintaxe e Fonologia.

Para a realização do experimento, foram levantadas oito sentenças escritas em que a preposição *para* aparece em final de frase, como pode ser visto a seguir.

(1) O valor do frango vai diminuir *para*?

(2) Vamos *para*?

- (3) Maria voltou para?
- (4) As crianças saíram para?
- (5) Meu filho viajou para?
- (6) Pedro foi transferido para?
- (7) Queriam partir para?
- (8) O progresso caminha para?

As sentenças escritas foram apresentadas aos seis informantes, todos falantes de língua portuguesa, para que eles dissessem se aceitavam ou rejeitavam a forma *para* na posição final. Se rejeitassem, seria considerado que aceitariam automaticamente as mesmas sentenças com a preposição reduzida *pra*. O próximo passo, então, seria a apresentação de pares de sentenças que se opunham pela forma da preposição *para*: uma frase com sua forma reduzida e outra com sua forma não reduzida.

O resultado do experimento indicou a preferência pela forma padrão *para* no final de sentença, e não por sua forma reduzida. Em um total de 48 tentativas, as oito sentenças terminadas em *para*, apresentadas aos seis informantes, foram rejeitadas apenas 10 vezes, gerando um total de 20% de rejeição, percentual relativamente baixo. Especificamente sobre cada uma das frases, a sentença (1) foi rejeitada por apenas um falante; a (2), por dois falantes; a (3), por nenhum falante; a (4), por um falante; a (5), por três falantes; a (6), por três falantes e a (7) e a (8), por nenhum falante.

Dessa forma, a hipótese inicial de Gazola é confirmada: a forma reduzida *pra* não ocorre – ou ocorre pouco, neste experimento – em final de sentença. Segundo a autora, pode-se concluir que a forma *pra* não é aceita em final de frase pelo fato de ser átona, por não se apresentar na forma de foco na sentença e, conseqüentemente, por não ser o constituinte mais à direita na sentença. Assim, *pra* pediria um complemento, que não existe no caso de final de sentença.

Mesmo tendo confirmado suas hipóteses, a autora reconhece que são necessárias outras pesquisas mais aprofundadas para avançar no estudo desse fenômeno, já que seu experimento contou com uma quantidade limitada de dados.

Apesar de não se tratar de um trabalho dedicado à análise de fala, como o nosso e os demais analisados, pensamos que seria relevante mencionar a pesquisa de Gazola pelo fato de tal autora ter analisado a preposição *para* em um contexto de prosódia. Entretanto, os

resultados não serão comparados com os nossos, visto que ela utiliza uma metodologia distinta para realizar sua análise.

#### 4.5 O ESTUDO DE FELGUEIRAS (*apud* MAYA, 2004, p. 16)<sup>2</sup>

A pesquisa de Felgueiras, “*Análise da Variação no uso da preposição para*”, foi realizada na UFRJ, em 1993. A autora investiga possíveis condicionamentos para a escolha de variantes da preposição *para* na fala do Rio de Janeiro, como aspectos estilístico, prosódico, fonológico e discursivo.

Os dados do estudo de Felgueiras são dos Projetos NURC (Norma Urbana Culta), com informantes universitários, e PEUL (Programa de Estudos sobre o Uso da Língua), com informantes de nível primário. Foram utilizadas 14 entrevistas do NURC e 12 do PEUL, estratificadas por sexo, idade, nível de formalidade e instrução.

Maya (2004, p. 16) afirma que a autora, inicialmente, pretendia realizar a pesquisa apenas com informantes universitários do NURC, ideia abandonada por razão da baixa ocorrência da forma *pa* nessa amostra. A análise de condicionamentos para o uso dessa variante tornou-se possível, portanto, a partir da inclusão dos dados do PEUL.

A amostra de Felgueiras contou com a seguinte distribuição:

1) Informantes do NURC:

- a) sexo: 7 homens e 7 mulheres;
- b) faixa etária: 8 com 26 a 49 anos; 6 com mais de 50 anos;
- c) nível de formalidade da fala: 6 de Elocução Formal (fala normal) e 8 de Diálogo entre Informante e Documentador (fala informal);
- d) nível de instrução: 14 de nível universitário;

2) Informantes do PEUL:

- a) sexo: 6 homens e 6 mulheres;
- b) faixa etária: 2 de 15 a 25 anos; 2 de 26 a 49 anos; 2 com mais de 50 anos;

---

<sup>2</sup> Devido à indisponibilidade do trabalho – possivelmente por ser um estudo antigo –, tanto em meio eletrônico quanto impresso, utilizaremos a referência feita pela pesquisa de Maya (2004).

- c) nível de formalidade da fala: 12 de fala informal;
- d) nível de instrução: 12 de nível primário.

Maya (2004) observa que a amostra do NURC é dividida em apenas duas faixas etárias, enquanto a do PEUL divide-se em três. Além disso, todos os informantes com nível primário apresentam fala informal, enquanto os universitários estão divididos por fala formal e informal. O autor também critica a existência de 20 informantes de fala informal e somente 6 de fala formal, desequilíbrio que gera problemas nos resultados da pesquisa.

Com um total de 1954 dados, os resultados da análise de Felgueiras apontam 176 (9%) ocorrências da forma padrão *para*, 1436 (73%) da forma *pra* e 342 (18%) da forma *pa*. Maya (2004) observa que a autora, apesar de ter investigado a variável *pa*, não a incluiu em seu texto por não ter obtido resultados conclusivos.

Como não tivemos acesso ao trabalho original de Felgueiras, já que se trata de um estudo antigo (1993), não compararemos seus resultados com os nossos e os dos demais trabalhos, justamente por estarem incompletas as informações que pudemos retirar da pesquisa de Maya (2004), na seção de resultados. No entanto, o levantamento das variantes faz-se pertinente.

#### 4.6 O ESTUDO DE FERREIRA (2014)

A pesquisa “*A Variação da Preposição para na Fala de Londrina pelos Dados do VARSUL*”, de Ferreira (2014), analisou aspectos linguísticos e sociais que poderiam influenciar a variação de *para/prá/pa* na fala de Londrina (PR). A autora utilizou um corpus de oito entrevistas do Banco VARSUL, estratificando-as por idade, sexo e escolaridade. As variáveis linguísticas controladas neste estudo foram *contexto morfológico seguinte, contexto fonológico seguinte, tonicidade da sílaba seguinte, paralelismo, número de sílabas do item seguinte, posição em relação a pausas e processo de sândi com a sílaba seguinte*.

As hipóteses levantadas por Ferreira (2014) são as que seguem: a forma padrão *para* ocorreria em proporção menor do que a variante *pra* e *pa*; as formas reduzidas *pra* e *pa* estariam motivadas por um grau de escolaridade mais baixo do indivíduo; a variante *para*, com acento na penúltima sílaba, quando reduzida para *pra* ou *pa* altera seu estatuto prosódico, tornando-se uma sílaba sem acento. A preposição em sua forma reduzida unir-se-ia a uma palavra que iniciasse com vogal átona.

De um total de 349 dados, Ferreira (2014) obteve 257 (73.6%) ocorrências de *pra*, 90 (25.8%) de *pa* e apenas duas (0.6%) de *para*. O percentual insignificante de ocorrências de *para* e a preferência dos informantes de Londrina pela variante *pra* (confirmando sua primeira hipótese) levou a autora à decisão de realizar uma análise de *pra* como variável dependente.

Os resultados indicaram que o papel da escolaridade deve ser relativizado em relação ao sexo do informante, pois são as mulheres com menos estudo (88%) e os homens com mais estudo (88%) que demonstram preferência pela variante *pra*, o que não confirma totalmente a segunda hipótese da investigação de Ferreira (2014). A terceira hipótese também não foi confirmada devido ao relativo equilíbrio entre os fatores Tônica (0.54) e Átona (0.45), da variável *Tonicidade da Sílabas Seguinte*.

Ademais, outras variáveis linguísticas mostraram-se relevantes, como os contextos seguintes com pausa (0.72) e com palavras não lexicais (0.57). Também se destacou a variável social *Idade*, sendo que os mais velhos, em geral (82,7%, peso 0.64), e as mulheres mais jovens (79%) utilizaram mais a forma reduzida *pra*.

#### 4.7 BREVES COMENTÁRIOS SOBRE A METODOLOGIA DOS ESTUDOS SOBRE A PREPOSIÇÃO

Feita a exposição das investigações sobre o comportamento variável da preposição *para* em diferentes cidades brasileiras, procedemos à síntese dos resultados obtidos em cada trabalho. Serão feitas algumas considerações a respeito de como essas pesquisas foram conduzidas a fim de se comparar em que diferem e em que se assemelham.

Primeiramente, chama a atenção o fato de Maya (2004) ter definido como variável dependente binária a forma reduzida *pa* em vez de *pra*. O autor, que examinou a fala da cidade de Porto Alegre, obteve 62% de aplicação da variante *pra*, contra apenas 36% de *pa*. Assim, talvez o foco da análise pudesse estar em *pra*, devido à maior frequência dessa variante.

Ainda sobre a questão da escolha da forma de aplicação da variável dependente, outro trabalho também difere em relação a sua definição. Lucena (2006) opta por confrontar os resultados das formas reduzidas *pra* e *pa* com a variante padrão *para*. O autor apresenta resultados das duas variantes juntas, de modo que não se sabe o percentual de aplicação de cada uma. Apesar de carregarem a mesma característica de preposição reduzida, seria

recomendável analisá-las separadamente, já que apresentam estruturas distintas de padrão silábico (*pra*, CCV e *pa*, CV).

Outro aspecto merecedor de comentário é o trabalho de Silva (2010), que compara seus resultados apenas com a pesquisa de Lucena (2006), ainda que já estivessem disponíveis, na época, os estudos de Felgueiras (1993), Maya (2004) e Gazola (2008). Outra consideração a respeito dessa pesquisa é que a autora restringe sua análise a poucas variáveis independentes linguísticas: Contexto Fonológico Seguinte, Presença de Vibrante no Item Seguinte e Paralelismo Formal, sem considerar outras variáveis que já se mostraram relevantes em outras pesquisas, como a classe morfológica, o número de sílabas ou o acento do item seguinte.

Lucena (2006) utilizou uma quantidade maior de variáveis linguísticas. No entanto, não controlou a classe morfológica e a tonicidade da sílaba seguinte. O autor, na exposição de seus resultados, faz descrições bastante detalhadas, inclusive em relação aos grupos de fatores que foram considerados irrelevantes em sua análise, assim como Maya (2004). Há também comentários críticos a respeito de outros trabalhos sobre a preposição *para* que são sucintos em seu texto, bem como em Silva (2010).

Silva (2010, p. 51), ao expor os resultados para a variável social *Idade*, salienta uma questão interessante a respeito da variante *pra*. Ela poderia ser considerada uma forma padrão já que seria mais utilizada pela faixa etária de 15 a 25 anos, sujeita à influência da escola e de alguns livros didáticos que trazem registros da preposição nessa forma.

Em Silva (2010), é interessante observar que os dados de *para* tiveram apenas 1% (35 ocorrências) de aplicação nos dados, sendo que alguns desses dizem respeito à expressão de um programa social do governo federal, chamado “*luz para todos*”, utilizada em contextos de língua escrita. A autora não menciona, entretanto, o percentual dessa variante entre os informantes mais jovens. Apenas diz que 18 dessas 35 ocorrências referem-se a um homem adulto (não se sabe se da faixa etária intermediária, de 26 a 49 anos, ou da mais velha, acima de 49 anos).

Maya (2004), Lucena (2006) e Silva (2010) não realizaram cruzamentos entre fatores linguísticos e sociais em seus estudos. Para esclarecer melhor todos esses pontos relativos ao possível condicionamento de cada fator, acredita-se que esse procedimento é essencial em estudos variacionistas. Assim, é possível determinar com mais clareza o papel dos fatores condicionadores, principalmente no caso das variáveis sociais.



No próximo capítulo, procederemos à exposição da metodologia adotada, bem como as hipóteses e os objetivos de nossa investigação.

## 5 METODOLOGIA

### 5.1 TEORIA DA VARIAÇÃO E MUDANÇA LINGUÍSTICA

A Teoria da Variação e Mudança Linguística consiste em um modelo teórico-metodológico proposto por Weinreich, Labov e Herzog na década de 1960. Este modelo analisa a variação linguística tendo como objeto de estudo a língua falada. De acordo com essa proposta, a variação faz parte do sistema da língua e pode sofrer interferências linguísticas e/ou sociais que condicionariam a escolha de uma variante linguística em detrimento de outra.

A questão central que motivou o surgimento da Teoria da Variação e Mudança Linguística foi o que Weinreich, Labov e Herzog nomearam de “heterogeneidade ordenada”. “(...) se uma língua tem de ser estruturada, a fim de funcionar eficientemente, como ela funciona enquanto a estrutura muda?” (Weinreich, Labov e Herzog, 2006, p. 87). Como resposta a esse questionamento, os autores afirmam que a língua é variável e sistematizada ao mesmo tempo; ou seja, ela continua funcionando de maneira eficiente enquanto as mudanças ocorrem na sua estrutura. Para chegarem a essa proposta, foi preciso romper a crença em um sistema linguístico homogêneo, característica defendida pela tradição neogramática e por autores como Hermann Paul, Saussure e Chomsky, que desconsideravam a coexistência de variabilidade e sistematicidade.

Faraco (2005, p. 31) também afirma que toda língua humana é sempre um conjunto heterogêneo de variedades. O autor ressalta a importância de se abandonar de forma definitiva a imagem de língua homogênea, cultivada pela tradição gramatical e sustentada pela escola, que trata algumas variantes linguísticas como “incorretas”. Pesquisas dialetológicas e sociolinguísticas têm contribuído empiricamente para mostrar que não existe língua homogênea, já que existe variação inclusive dentro de cada sistema linguístico, os quais nomeamos por rótulos como *português*, *japonês*, *espanhol* etc. (FARACO, 2005, p. 31).

Para Weinreich, Labov e Herzog (2006, p. 121), os problemas que devem ser resolvidos por uma Teoria de Mudança Linguística são os que seguem:

**Fatores condicionantes:** conjunto de mudanças possíveis e condições possíveis para a mudança.

**Transição:** 1) surge a nova variante; 2) as duas variantes coexistem; 3) uma das formas desaparece.

**Encaixamento:** 1) na estrutura linguística; 2) na estrutura social.

**Avaliação:** modo como os falantes avaliam uma mudança linguística.

**Implementação:** elevação da consciência social sobre a mudança linguística na comunidade de fala.

Weinreich, Labov e Herzog (2006, p. 125) defendem a importância de se estabelecer alguns princípios gerais para o estudo da mudança linguística. São eles os que seguem:

1. A mudança não deve ser identificada como variação aleatória; ela começa quando a generalização de determinada alternância em uma comunidade de fala toma uma direção e assume o papel de diferenciação ordenada.
2. A estrutura linguística é heterogênea e ordenada.
3. Nem toda variação implica mudança, mas toda mudança pressupõe variação.
4. A generalização da mudança não é instantânea nem uniforme, pois requer uma covariação de formas durante um período de tempo.
5. As gramáticas são da Comunidade de Fala, não de idioletos.
6. A mudança linguística ocorre dentro da comunidade como um todo.
7. O desenvolvimento da mudança depende de fatores linguísticos e sociais.

Os objetivos da Teoria da Variação são: analisar e sistematizar a heterogeneidade da língua, que pode apresentar variação condicionada e coerente; descrever o funcionamento da língua para se descobrir interferências da fala na escrita; reconhecer as variantes que tenham prestígio e as que tenham estigma em determinadas sociedades. Como forma de viabilizar tais propósitos, essa teoria postulou a necessidade de uma pesquisa profunda e metódica, com combinações de dados que possibilitassem descobrir por que determinada pessoa de determinada idade, escolaridade e região falava diferentemente de outra(s) com características opostas.

A nova concepção gerada pela Teoria da Variação Linguística, de Weinreich, Labov e Herzog, confirmou que as escolhas entre sons, palavras ou estruturas distintas não ocorreriam aleatoriamente, por pura vontade do falante, mas sim obedecendo a um padrão sistemático,

com *regras variáveis*, que são processos de escolha entre duas ou mais estruturas linguísticas em diversos contextos. De acordo com Mollica (2004, p. 11), as diferentes formas de se dizer a mesma coisa, em um mesmo contexto, constituem *variantes linguísticas*. Um conjunto de variantes forma, então, uma *variável linguística*.

Para executar análises de regras variáveis, Labov baseava-se no método da “Sociolinguística Quantitativa”. Esse método opera por meio de análise estatística de dados de fala e seus possíveis condicionamentos, quantificando e testando os efeitos que um fator contextual pode ter para uma variável linguística. O conjunto de escolhas possíveis forma uma *variável dependente* e os possíveis condicionamentos são as *variáveis independentes*.

Quando um fenômeno de variação apresenta uma variante liderando a “competição”, a situação de variação dá lugar a uma *mudança linguística*. A mudança linguística tem algumas características principais, conforme Faraco (2005, p. 44): é contínua, pois se dá em todas as línguas; lenta e gradual, já que não ocorre de forma abrupta e vai atingindo gradativamente partes da língua; relativamente regular, ou seja, um elemento em processo de mudança vai alcançar, de forma progressiva, o estágio completo de mudança, em todas as suas ocorrências.

Para se realizar um estudo de mudança linguística, é necessário controlar a idade do informante, que costuma manter os padrões de língua da sua época de formação. Ao chegar à fase adulta, o falante tende a adequar sua fala aos padrões desta faixa etária.

Segundo Faraco (2005, p. 21), pode-se verificar mudança linguística em duas situações: *tempo aparente* e *tempo real*. Para uma análise em tempo aparente, deve-se realizar um recorte transversal da amostra sincrônica em função da faixa etária dos informantes (jovens, meia idade e idosos, por exemplo). Feita a análise dos dados, é preciso correlacionar as variantes à faixa etária. Se os mais jovens utilizam mais e os mais velhos utilizam menos a forma inovadora, formando uma linha decrescente de uso conforme a idade, é possível que essa seja uma situação de mudança em progresso (FARACO, 2005, p. 186).

No caso de uma análise em tempo real, de acordo com Faraco (2005), deve-se proceder ao encaixamento linguístico da variável em estudo no tempo real por meio de pesquisas em fontes como Atlas Linguístico, textos em prosa, cartas pessoais etc., a fim de se obter dados que podem ser valiosos para atestar a mudança. Uma possibilidade também é fazer uma comparação entre falas de diferentes épocas, ainda que gravações de fala, por seu caráter mais

recente, nem sempre tragam formas suficientemente antigas para se constatar um caso de mudança linguística.

### 5.1.1 Operacionalização do Modelo

Brescancini (2002, p. 15) recomenda alguns procedimentos para se fazer um estudo de regra variável. Em primeiro lugar, deve-se definir a *variável dependente*, que é o fenômeno linguístico variável a ser investigado. Logo após, faz-se um levantamento de todas as variantes possíveis para a variável em questão. As variáveis dependentes podem ser binárias, que são formadas por duas variantes, ou eneárias, com três ou mais variantes.

No caso da preposição *para*, há formas variantes como *pra* e *pa*, que são utilizadas de forma variável na fala do português brasileiro atual, como *Vou para/pra/pa casa*, por exemplo. De acordo com a Teoria da Variação, *para* é uma regra variável que possui diferentes formas de manifestação fonética, que seriam condicionadas por algum aspecto linguístico e/ou social.

O passo seguinte para um estudo de regra variável é definir as *variáveis independentes*, que são as hipóteses iniciais formuladas pelo pesquisador sobre os condicionamentos que se espera encontrar. Esses condicionamentos podem ser linguísticos, como o contexto morfológico seguinte ao fenômeno analisado, ou sociais, como a idade do informante. Os fatores condicionantes das variáveis independentes podem se mostrar favoráveis para a atuação de uma variante ou não.

O próximo passo de preparação para o estudo da regra variável, segundo Brescancini (2002), é delimitar a amostra e obter os resultados. Para isso, pode-se utilizar material já coletado em bancos de dados de fala ou realizar as próprias entrevistas. Independentemente da decisão, é essencial utilizar alguns critérios para formar uma amostra representativa, tais como informações sobre fronteiras geográficas e sociais da comunidade de fala, relevância da idade, sexo ou presença de grupos étnicos que apresentem determinadas variantes de fala.

Também é importante selecionar pessoas que tenham nascido ou vivido a maior parte de suas vidas na cidade estudada, para que sejam representativas étnico/culturalmente da comunidade onde vivem. Assim, é possível generalizar os resultados obtidos para uma pequena quantidade de informantes possam representar toda a sua comunidade de fala.

Para saber onde e como selecionar indivíduos para formar um banco de dados, deve-se contatar a prefeitura da cidade ou o IBGE, que podem fornecer informações como perfil social recente dos indivíduos e estrutura dos bairros. Neste momento, é aconselhável o apoio de sociólogo ou antropólogo que conheça bem a comunidade. (SILVA, 2004, p. 124)

Para definir uma amostra que seja representativa de uma comunidade de interesse, em termos de quantidade de informantes, o pesquisador pode adotar a técnica de multiplicar entre si o número total de fatores de cada grupo social. Na Tabela 3 abaixo, têm-se as variáveis sociais e os fatores delimitados para a nossa pesquisa (BRESCANCINI, 2002, p. 17):

Variáveis Sociais	Fatores de cada variável
Sexo	Feminino
	Masculino
Idade	- de 50 anos
	+ de 50 anos
Escolaridade	Até 4 anos de estudo
	Até 11 anos de estudo

Tabela 3 – Ilustração de células sociais conforme nossa amostra. Fonte: a autora.

O cálculo baseado nos dados da ilustração acima é  $2 \times 2 \times 2 = 8$ . O resultado representa o número de células que serão preenchidas de acordo com as características dos falantes pertencentes à amostra: neste caso, um indivíduo por célula. Ou seja, na amostra haverá uma célula composta por uma mulher com - de 50 anos e até quatro anos de estudo, outra mulher com - de 50 anos, porém, com até onze anos de estudo etc.

Entretanto, é importante observar a distribuição de informantes por células sociais: se houver apenas uma pessoa com determinado grupo de características para cada célula, talvez a amostra não represente bem o comportamento de todos os indivíduos de determinada localidade, alerta Brescancini (2002, p. 18).

O método de coleta de dados mais habitual e vantajoso para a maioria dos fenômenos costuma ser a realização de entrevistas, segundo Mollica e Braga (2004, p. 125). Trata-se de uma conversa informal, pois deseja-se uma fala casual, habitual, de acordo com o **paradoxo do observador**, conceito de Labov trazido pelas autoras, que consiste em “observar a fala do falante quando ele não é observado.”.

Para fenômenos mais frequentes, Mollica e Braga (2004) afirmam que são necessárias menos horas de gravação. É recomendado que se conheça o melhor possível o falante antes da entrevista, momento em que se aplica um questionário social com informações de idade, profissão, interesses pessoais etc. Na hora de criar a entrevista, deve-se planejar uma quantidade suficiente de assuntos. Além disso, é aconselhável utilizar formas menos usuais em entrevistas, em uma tentativa de garantir uma diversidade de estruturas na fala do entrevistador. Deve-se, também, provocar outros gêneros no entrevistado, como o narrativo e o argumentativo, por exemplo. Após a gravação, os dados devem ser transcritos.

A transcrição de dados pressupõe um conjunto de decisões norteadas pelos objetivos que o pesquisador tenha. Entretanto, um objetivo geral deve ser a fidelidade aos dados orais. Após a gravação, deve-se transcrever atentando para questões do discurso falado como hesitações, gaguejos, interrupções, falsos começos etc., já que não têm representação possível no sistema ortográfico. Além disso, é importante numerar as linhas a fim de facilitar buscas.

Antes de começar a transcrição, deve ser definido o grau de detalhamento desejado pelo pesquisador. Para descrições fonéticas mais detalhadas, a recomendação é escutar a gravação, além de utilizar a transcrição. A maioria dos sistemas de transcrição de dados de fala, conforme pode ser visto na ilustração a seguir, parte do sistema ortográfico.

Segue exemplo de transcrição retirado de Mollica e Braga (2004):

```

que eles dizem [o]- a favela
s v n n n v d d
lá embaixo. *A F[do]- do José Pinto
pd pd s a --a d o + 0 1
Ffavela ficou com o nome.
*Agora o bairro aqui 0 00+ o e0
x' s v p d s a d s a
Fa prefeitura que classificou como
o bairro da e 0 x'
d s n v c d s pd
FRebouças (est) que era um nome [bem
mais]- x 0+ e :
s n v d s a a
F(hes) mais bem parecido, né? (est) *[E]- e
o0 1 : a
a j av,pi c e
Fagora com essa, tínhamos o trilho do trem aí
1 o o a p n
v d s pd s a
F[que]- que cortava a metade [da]-[da]-[da]- do e e
r e o n n v d
s pd pd pd pd
Fbairro, né? (est) agora, com a idéia do0 #' 0 0
0 : o: s av,pi a
p d s pd

```

Ilustração 2 – Transcrição de entrevista. Fonte: Projeto VARSUL.

A ilustração 2 traz a transcrição da gravação de uma entrevista, que procura representar a oralidade, ao registrar marcas como correções (“[o]- a favela...”), hesitações (“bem mais]- (hes) mais bem parecido, né?”) e gaguejo (“que cortava a metade [da]- [da]- [da]- do...”). A primeira letra da imagem, “F”, introduz a frase do falante/entrevistado, ao passo que a fala do entrevistador é precedida de uma letra “E”. Já as letras que constam abaixo de cada linha de texto representam códigos, padronizados pelo Projeto VARSUL, para classificar como o informante realizou a pronúncia, a entonação etc.

O próximo passo para o estudo de regra variável é a audição das entrevistas e anotação de como o informante realiza o dado alvo. Dependendo do fenômeno escolhido, deve-se extrair apenas uma palavra ou mesmo um enunciado inteiro. No fenômeno analisado neste trabalho, é essencial extrair enunciados inteiros com a preposição *para*, incluindo contexto precedente e seguinte, e marcar qual foi a variante eleita pelo entrevistado, se *para*, *pra* ou *pa*.

Segue um exemplo de enunciado utilizado neste trabalho:

*Depois tinha que vim a pé de lá **pra**, trazer **pra** cá... uns trequinho de carroça*

Para sistematizar todos os dados coletados e relacioná-los a cada grupo de variáveis independentes, deve-se atribuir um código para cada fator, a fim de facilitar a quantificação: uma letra, um número ou mesmo um símbolo. Esses códigos, utilizados no programa estatístico GOLDVARB, são digitados em ordem preestabelecida, conforme os grupos de fatores. Com esse programa de análise estatística de regra variável, pode-se quantificar os dados e se obter um produto final, a *rodada*, com resultados preliminares que devem ser analisados posteriormente pelo pesquisador. A codificação completa, criada para a análise da variável *para*, encontra-se em uma tabela na seção ANEXO C deste trabalho.

A seguir, exemplificamos essa codificação a partir de um grupo de fatores de nosso *corpus*:

## **GRUPO 1: CONTEXTO MORFOLÓGICO SEGUINTE**

**Código n:** Nome → 2nkwi8sCy-4Z *Se eu contá o lugar, eu vô **pa** cadeia.*

**Código v:** Verbo → 1vfwi7sFy+5A *Às vezes, quando a planta tá **pra** florir,*

**Código w:** Numeral → 1wd@i6Fy+5O *Só levar um bolo **pra** cem pessoas,*

**Código q:** Pronome → 1qdti7sFy-5R *Não tem mais espaço **pra** nada,*

**Código g:** Artigo → 0gu@i6sFx+5M *Ou então mudou da noite **para** o dia,*



**Código c:** Conjunção → 1ck@i6sFy+5V *até hoje a Capitania no- Ministério da Marinha não- não liberou aquela área **pra** que a Prefeitura possa abrir.*

**Código ?:** Advérbio → 1?k@i6sFy-4Z *da rodoviária nova **pra** cá,*

**Código !:** Nada (pausa) → 1!3XrX[Fy+4I *Mas eu ia com amigos **pra-** pra escola*

No exemplo de codificação acima, o primeiro código de cada enunciado representa a variável dependente da análise, ou seja, se o indivíduo entrevistado utilizou *para* (código 0), *pra* (código 1) ou *pa* (código 2). O segundo código, destacado em negrito em cada frase, é referente ao grupo 1 (*Contexto Morfológico Seguinte*) e os códigos que seguem, da esquerda para a direita, referem-se às demais variáveis linguísticas e sociais controladas nesta pesquisa.

No momento de realizar a digitação dos códigos, deve-se fixar a atenção diretamente na palavra alvo; no caso de nossa pesquisa, a preposição *para* ou suas variantes. Cada enunciado com a preposição é considerado um dado. Este dado é codificado com base nos fatores levantados como possíveis condicionadores, como se vê no enunciado a seguir:

**1vfw7sFy+5A** *Às vezes, quando a planta tá **pra** florir,*

O código “**1**” marca que a variante aplicada foi: *Pra*;

O código “**f**” marca o contexto fonológico seguinte: *Consoante Labiodental [f]*;

O código “**+**” marca a idade do informante: *+de 50 anos*;

O código “**5**” marca a escolaridade da informante: *Até 11 anos*.

Os resultados numéricos obtidos nas rodadas por meio do programa GOLDVARB têm valor apenas estatístico. O valor linguístico real deve ser interpretado pelo linguista. De acordo com Guy e Zilles,

Os números não são a resposta a nenhuma de nossas perguntas; eles são apenas estatísticas inferenciais adicionais que podemos usar como indicadores empíricos na nossa busca por respostas. Talvez o tipo mais básico de inferência a ser extraída dos resultados da regra variável seja estimar a direção e a magnitude dos efeitos dos fatores. (GUY e ZILLES, 2007, p. 42).

Segundo Guy e Zilles (2007, p. 41), para explicação e interpretação mais detalhada dos dados, várias ações podem ser tomadas ao longo da análise. É possível calcular medidas de

resumo dos dados, realizar cruzamentos de grupos de fatores, recodificar e/ou eliminar fatores ou grupos inteiros, e inclusive separar os dados em arquivos diferentes, no caso de uma análise sobre duas ou mais regiões, por exemplo. A razão para esse tipo de alteração durante o processo, para Tarallo (2012), pode ser justificada pelo surgimento de contextos de fala imprevisíveis, que exigiriam modificações, a fim de possibilitar resultados mais precisos e completos.

Após a classificação de todos os fatores linguísticos e sociais, utiliza-se o programa computacional GOLDVARB para se obter a frequência e a distribuição de cada variante para cada fator linguístico e extralinguístico. Esta análise quantitativa permite que o pesquisador obtenha uma estatística das variáveis independentes mais influentes na aplicação do fenômeno, com percentual de uso e *peso relativo*, que é uma estimativa dos efeitos restritivos e sua significância, nas palavras de Guy e Zilles (2007, p. 41).

O peso relativo é um valor entre 0 e 1, que resulta de um algoritmo realizado pelo programa e indica em que medida e direção o fator analisado influencia a taxa de aplicação da regra. Esse valor não varia conforme a quantidade de dados, de acordo com os autores, sendo considerado significativo um peso relativo a partir de 0.50. Dessa maneira, um peso abaixo de 0.50 desfavorece, enquanto um peso exatamente igual a 0.50, chamado “ponto neutro”, segundo Guy e Zilles (2007, p. 239), não representa nenhum efeito na aplicação da regra. Já um valor próximo a 0 indica que a regra nunca se aplica, enquanto um valor próximo a 1 representa que ela sempre será aplicada no contexto do fator analisado.

Depois que a digitação de todos os códigos no programa GOLDVARB esteja completa, realiza-se a rodada estatística e se inicia a interpretação dos resultados. Esta é a etapa mais importante, pois exige capacidade de analisar a relevância dos percentuais e pesos relativos apresentados pelo programa, possivelmente de reorganizar fatores e hipóteses e, por fim, de realizar novas rodadas.

Por se tratar de um programa quantitativo, o GOLDVARB requer uma interpretação detalhada de cada rodada estatística para corrigir possíveis problemas gerados por ele. Um desses problemas pode ser a existência de *células vazias*, que são as combinações de dados não coletados ou impossíveis de ocorrerem determinada regra variável, conforme atestam Guy e Zilles (2007, p. 52). No caso de variáveis linguísticas, é possível que não se obtenha nenhuma ocorrência da variante *para* diante de uma conjunção, por exemplo. Já no caso de variáveis sociais, um informante pode pertencer apenas a um sexo e a uma faixa etária.

Outro problema que pode ser causado pelo programa estatístico GOLDVARB é a irrelevância de certos grupos de fatores. Neste caso, os grupos com poucos dados ou nenhum dado (os *knockouts*) podem ser excluídos da análise (GUY e ZILLES, 2007, p. 57). Também existe a possibilidade de refinar a análise a partir da junção de alguns fatores com poucas ocorrências e que possam apresentar influência semelhante para a realização da variante analisada. Esse processo se chama *amalgamação*. Feitas as amalgamações necessárias, realiza-se uma nova rodada.

Os diversos níveis de análise permitem verificar interferência entre as variáveis, que pode ser causada por sobreposição ou enviesamento de fatores. Os resultados estatísticos, além de quantificarem hipóteses sobre o fenômeno investigado, trazem novas informações sobre ele, o que pode levar o pesquisador a reavaliar os dados de forma mais rigorosa.

## 5.2 AMOSTRA

A amostra que compõe este trabalho utiliza dados do Projeto VARSUL (Variação Linguística do Sul do Brasil), que é um banco de dados de língua falada que reúne gravações de fala espontânea de quatro cidades representativas étnico e culturalmente de cada estado da região Sul: Paraná, Santa Catarina e Rio Grande do Sul. Tem por objetivo geral a descrição do português falado com a parceria de quatro universidades brasileiras: Universidade Federal do Rio Grande do Sul, Pontifícia Universidade Católica do Rio Grande do Sul, Universidade Federal de Santa Catarina e Universidade Federal Tecnológica do Paraná.

O Projeto VARSUL também tem como propósito oferecer subsídios para teste e desenvolvimento de teorias linguísticas, condições para formação de novos pesquisadores e para programas educacionais, promovendo o conhecimento e o respeito às variedades linguísticas. Maiores detalhes podem ser obtidos no site <http://varsul.org.br>.

Os dados utilizados foram obtidos das cidades de Curitiba (PR) e Florianópolis (SC), coletados nos anos 1990. Foram selecionados 32 informantes, sendo 16 de cada cidade, estratificados por sexo (16 homens e 16 mulheres), idade (- de 50 e + de 50 anos) e escolaridade (primário, ginásio e 2º grau, adaptados por nós para os dois fatores “até 4 anos” e “até 11 anos” de estudo).

Na Tabela 4, constam as características da estratificação dos informantes selecionados.

<b>Informantes</b>	<b>Sexo</b>	<b>Idade</b>		<b>Escolaridade</b>
<b>CTB 01</b>	Masculino	- de 50	40 anos	Até 4 anos
<b>CTB 02</b>	Masculino	+ de 50	+ de 55 anos	Até 11 anos
<b>CTB 05</b>	Masculino	- de 50	35 anos	Até 4 anos
<b>CTB 07</b>	Masculino		36 anos	Até 11 anos
<b>CTB 09</b>	Masculino	- de 50	26 anos	Até 4 anos
<b>CTB 13</b>	Masculino		+ - 50 anos	Até 11 anos
<b>CTB 15</b>	Masculino	+ de 50	+ de 70 anos	
<b>CTB 21</b>	Masculino		58 anos	Até 4 anos
<b>CTB 06</b>	Feminino		70 anos	Até 11 anos
<b>CTB 08</b>	Feminino	- de 50	39 anos	Até 4 anos
<b>CTB 10</b>	Feminino		+ - 40 anos	
<b>CTB 12</b>	Feminino		+ - 30 anos	Até 11 anos
<b>CTB 14</b>	Feminino	+ de 50	77 anos	Até 4 anos
<b>CTB 19</b>	Feminino	- de 50	+ - 30 anos	Até 11 anos
<b>CTB 20</b>	Feminino	+ de 50	+ de 50 anos	Até 4 anos
<b>CTB 22</b>	Feminino		+ de 50 anos	Até 11 anos
<b>FLP 01</b>	Feminino	- de 50	25 anos	Até 4 anos
<b>FLP 03</b>	Feminino		34 anos	
<b>FLP 07</b>	Feminino	+ de 50	55 anos	
<b>FLP 08</b>	Feminino		76 anos	
<b>FLP 11</b>	Feminino	- de 50	42 anos	Até 11 anos
<b>FLP 15</b>	Feminino	+ de 50	62 anos	
<b>FLP 20</b>	Feminino	- de 50	27 anos	
<b>FLP 22</b>	Feminino	+ de 50	52 anos	
<b>FLP 02</b>	Masculino	- de 50	39 anos	Até 4 anos
<b>FLP 04</b>	Masculino		45 anos	
<b>FLP 05</b>	Masculino	+ de 50	59 anos	
<b>FLP 06</b>	Masculino		76 anos	
<b>FLP 10</b>	Masculino	- de 50	35 anos	Até 11 anos
<b>FLP 14</b>	Masculino	+ de 50	52 anos	
<b>FLP 18</b>	Masculino	- de 50	42 anos	Até 11 anos
<b>FLP 21</b>	Masculino	+ de 50	55 anos	

Tabela 4 – Estratificação dos informantes por sexo, idade e escolaridade. Fonte: a autora.

## 5.2.1 Variáveis Controladas

### 5.2.1.1 Variável Dependente

A variável dependente controlada é a preposição *para*, que pode ser realizada como *para*, *pra* ou *pa* no português brasileiro. Por ter três variantes, é considerada uma variável eneária. A delimitação da variável a ser estudada requer um levantamento de todo o grupo de variantes possíveis, já que os objetivos de análise de certos pesquisadores podem ser totalmente distintos, ainda que estejam investigando fenômeno semelhante. Neste trabalho, a variante escolhida para a análise de condicionamento linguístico e/ou social é *pra*. Esta forma será controlada, na análise estatística, como a variante de aplicação da regra variável da preposição *para*.

### 5.2.1.2 Variáveis Independentes Linguísticas

Quando estudamos uma Regra Variável pela perspectiva da Teoria da Variação, é necessário conduzir a pesquisa buscando possíveis motivações para a realização de uma variante em detrimento de outra. Para isso, faz-se necessário estabelecer alguns aspectos de análise tanto de ordem linguística, quanto de ordem social, que podem servir como fatores condicionadores ou não da aplicação de regras. A possibilidade de uma classe gramatical ou um som específico – que preceda ou suceda a variável em questão – condicionarem uma forma em detrimento de outra pode ser medida pelo seu número de ocorrências nos dados utilizados.

Foram controladas as variáveis linguísticas *contexto morfológico seguinte*, *contexto fonológico seguinte*, *tonicidade da sílaba seguinte*, *paralelismo*, *número de sílabas do item seguinte*, *posição em relação a pausas e processo de sândi<sup>3</sup> com a sílaba seguinte*; e as variáveis extralinguísticas *sexo*, *idade*, *escolaridade* e *cidade*.

A seguir, descreveremos cada variável independente utilizada em nossa pesquisa.

---

<sup>3</sup> De acordo com Bisol (2002, p. 232), sândi é o processo que ocorre quando entram em contato duas palavras, uma terminada por vogal e outra iniciada por vogal. Essa junção causa um choque dos núcleos silábicos, o que desfaz a sílaba final da primeira palavra e gera uma ressilabificação entre as duas palavras.

A variável *contexto morfológico seguinte* foi subdividida em oito fatores: *nome (pra casa)*, *verbo (pra dizer)*, *numeral (pra uma hora)*, *pronome (pra mim)*, *artigo (para a mãe)*, *conjunção (pra que fosse)*, *advérbio (pra cá)* e *nada* (quando não há contexto morfológico seguinte, como, por exemplo, uma interrupção na fala). Foi examinada com o objetivo de se localizar um possível condicionamento para a realização de *para*, *pra* ou *pa* conforme a classe gramatical da palavra seguinte. Pode ser que uma sequência de verbo ou substantivo motive mais significativamente a variação do fenômeno em questão do que um advérbio ou um numeral, como apontam os resultados de Maya (2004).

A variável *contexto fonológico seguinte* tem nove subdivisões: *vogal posterior (pra uma, pro ônibus, pra ópera)*, *vogal central (pra andar)*, *vogal anterior (pra ir, pra ela, pro elevador)*, *consoante labiodental (pra fora)*, *consoante velar (pra confeccionar)*, *consoante bilabial (pra movimentar)*, *consoante dental (pra ser)*, *consoante palatal (pra tia)* e *pausa* (quando não há contexto fonológico seguinte, como uma interrupção ou pausa na fala, por exemplo). Visto que nosso estudo se concentra em um fenômeno de ordem não apenas morfológica, mas também fonológica, consideramos essencial observar o som imediatamente posterior à preposição *para/pra/pa*. Fenômenos como sândi e elisão, por exemplo, muitas vezes estão relacionados a uma motivação prosódica para a variação, de acordo com Bisol (2000).

A *tonicidade da sílaba seguinte* foi subdividida em quatro fatores: *tônica (pra praia)*, *átônica (pra valer)*, *clítica (pra te falar)* e *nada* (quando inexistia qualquer sequência de fala). Acredita-se que esses fatores devem ser analisados para se descobrir uma possível influência prosódica na ocorrência do fenômeno de variação.

A variável *paralelismo* apresenta, para a preposição estudada, as classificações *ocorrência isolada*, *primeira da série*, *antecedida de pra*, *antecedida de para* e *antecedida de pa*. É investigada a possibilidade de uma ou mais de uma ocorrência no mesmo período – considerando também a variante precedente e a seguinte – exercer influência na escolha da variante em questão, assim como testou Maya (2004) em seu estudo.

Em *número de sílabas do item seguinte*, temos *1 sílaba*, *2 sílabas*, *3 sílabas*, *4 ou mais sílabas* e *nada* (quando inexistia qualquer sequência de fala). Assim como a *tonicidade da sílaba seguinte*, acredita-se que a quantidade de sílabas da palavra que segue a preposição estudada possa influenciar prosodicamente a sua variação, uma vez que a passagem de *para* a *pra* modifica a estrutura de duas sílabas para uma sílaba, conforme Bisol (2000).

Na variável *posição em relação a pausas*, marcamos os fatores *sem pausas*, *depois de pausa*, *antes de pausa* e *entre pausas*. Espera-se descobrir se determinadas variantes são realizadas dependendo de pausas que elas precedam, sucedam ou mesmo quando ocorram entre duas.

A última das variáveis linguísticas definidas, *processo de sândi com a sílaba seguinte* funciona no contexto de sândi entre a preposição *para/prá/pa* e o artigo que a sucede (*a*, *o*, *as*, *os*). Surgiu pela necessidade de uma análise mais precisa, e foi dividida em oito casos: *Com artigo “a” no plural (pra + as = pras)*, *Com artigo “o” no plural (pra + os = prus, pus)*, *Com artigo “a” no singular (para/prá + a = pra, pa)*, *Com artigo “o” no singular (pra/pa + o = pru, pu)*, *Com artigo “um” no singular (pra + um = pr’um)*, *Pra + palavra com duas sílabas (pra + você = p’cê)*, *Pra + palavra com três sílabas (pra + aquilo = pr’aquilo)* e *Nada (sem sândi)* – porque não há artigo após a preposição.

### 5.2.1.3 Variáveis Independentes Extralinguísticas

As variáveis extralinguísticas examinadas são aquelas disponíveis pelo VARSUL: *sexo*, *idade* e *escolaridade*. No caso da variável *sexo*, o objetivo é verificar se o fato de um informante ser homem ou mulher teria relevância para a escolha de uma ou outra variante. Sabe-se que as mulheres costumam utilizar as variantes inovadoras, como reforça Labov (2008).

A variável *Faixa etária* (- de 50 anos e + de 50 anos) foi controlada com o fim de se descobrir se o fato de um informante ser mais jovem ou mais velho teria relação com a eleição de uma e não de outra forma. De acordo com Naro (*in* Mollica e Braga, 2004, p. 47), a idade pode ser um critério de análise importante para a questão da mudança linguística. Os falantes mais jovens, ao escolherem formas mais inovadoras, podem indicar mudança em progresso. No entanto, conforme o autor, são os jovens e os velhos que apresentam o mesmo comportamento linguístico, em contraste com a população de meia-idade, que costuma viver sob as pressões do mercado de trabalho.

A *escolaridade* (*Até 4 anos* e *Até 11 anos*, adaptação feita da terminologia *Primário* e *2º grau*, do Projeto VARSUL) foi analisada para verificarmos se a quantidade de anos de escolarização de cada informante exerceria influência em sua escolha por uma das variantes.

Consideramos a hipótese de que este fator possa influenciar a escolha da forma mais inovadora *pa* em falantes com menor escolaridade, já que essa variante está mais distante da ortografia padrão, *para*.

### 5.3 OBJETIVOS

Esta pesquisa pretende investigar a variação da preposição *para* nas cidades de Curitiba (PR) e de Florianópolis (SC) com base na Teoria da Variação e Mudança Linguística, de Weinreich, Labov e Herzog. Mais especificamente, nosso objetivo é verificar se existem motivações linguísticas e/ou sociais para a ocorrência desse fenômeno variável, comparando-se resultados alcançados em trabalhos similares.

### 5.4 HIPÓTESES

As hipóteses que norteiam este trabalho são as que seguem.

1. A variante *pra* seria utilizada em proporção maior do que as variantes *para* e *pa*, na fala de indivíduos de Curitiba e de Florianópolis, semelhantemente ao que ocorre em outras regiões do Brasil. Pesquisas, em outras regiões do País, apontam para o uso mais frequente das formas reduzidas, caracterizando um processo de mudança em curso no português brasileiro.
2. A forma reduzida *pra* pode estar motivada por um grau de escolaridade mais baixo do indivíduo, tendo em vista que é típica da língua falada coloquial e não cuidada.
3. A variante *para*, com acento na penúltima sílaba, quando reduzida para *pra* ou *pa* altera seu estatuto prosódico, tornando-se uma sílaba sem acento e uma forma dependente. Como tal, a preposição reduzida sofre um processo de junção (chamado sândi) com a palavra seguinte que inicia com vogal átona.

A próxima seção apresentará os resultados de nosso estudo sobre a variação da preposição *para* no português brasileiro falado.



## 6 RESULTADOS DA ANÁLISE DE DADOS

### 6.1 DISTRIBUIÇÃO DAS VARIANTES DA PREPOSIÇÃO *PARA*

A análise eneária de nosso estudo, de um total de 2591 dados, obteve 2086 (80,5%) aplicações de *pra*, 456 (17,6%) de *pa* e 49 (1,9%) da variante padrão *para*. A situação é ilustrada no Gráfico 1 a seguir:

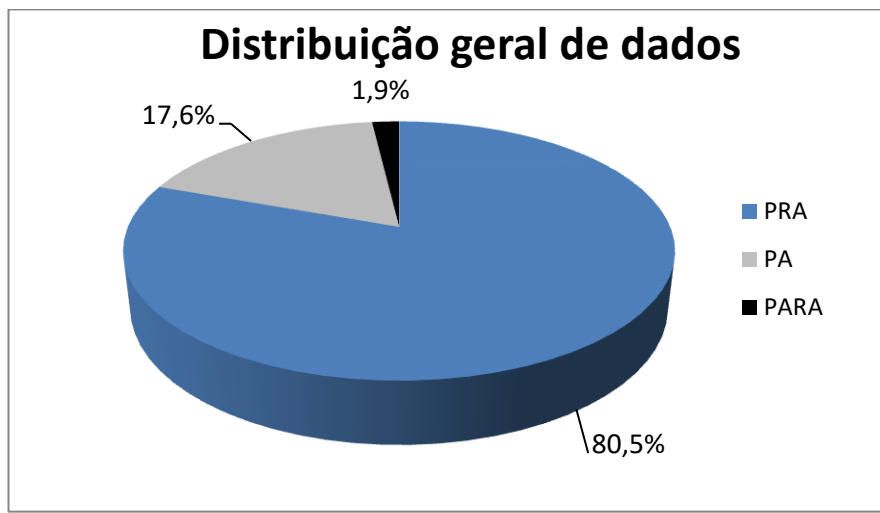


Gráfico 1 – Distribuição Geral dos Dados. Fonte: a autora.

A distribuição das variantes *para/pra/pa* em trabalhos abordados no capítulo 4 pode ser sintetizada, de forma comparativa, pela Tabela 5:

	Cidade	Variável Dependente	<i>Para</i>		<i>Pra</i>		<i>Pa</i>		Nº de Dados
			Apl.	%	Apl.	%	Apl.	%	
Maya (2004)	Porto Alegre (RS)	<i>Pa</i>	42	2	1251	62	741	36	2034
Lucena (2006)	João Pessoa (PB)	<i>Pra/pa</i> (juntos)	87	6	655	45	703	49	1445
Silva (2010)	Araguatins (TO)	<i>Pra</i>	35	1	1852	54	1323	45	3210
Felgueiras (1993)	Rio de Janeiro (RJ)	<i>Pra</i>	176	9	1436	73	342	18	1954
Ferreira (2014)	Londrina (PR)	<i>Pra</i>	2	0,6	257	73,6	90	25,8	349
Ferreira (2018)	Curitiba (PR) e Florianópolis (SC)	<i>Pra</i>	49	1,9	2086	80,5	456	17,6	2591

Tabela 5 – Distribuição de dados por variante em cada trabalho. Fonte: a autora

Na Tabela 5, vemos que todos os trabalhos obtiveram percentual muito baixo de *para*. Já a variante *pra* teve preferência na maioria das pesquisas, exceto em Lucena (2006). Os resultados mais próximos em favor dessa variante foram o de Felgueiras (1993) e Ferreira (2014) (aproximadamente 73%), dos quais este presente estudo se aproxima (80,5%).

A variante *pa* apresenta mais aplicações na cidade de João Pessoa (PB), em Lucena (2006), com 49%, e na cidade de Araguatins (TO), em Silva (2010), com 45%. No trabalho que analisou a cidade tocantinense, obteve-se a maior quantidade de dados entre todas as investigações, com 3210 contextos.

Faz-se necessária uma observação a respeito do estudo de Gazola (2008): a autora trabalhou com uma quantidade consideravelmente menor de dados em relação ao nosso trabalho e aos demais, além de não utilizar dados de fala. Gazola realizou um experimento que se trata de um teste, no qual se apresentam a seis pessoas algumas sentenças terminadas sempre com a preposição *para*. Essas pessoas deveriam dizer se as aceitavam ou as rejeitavam. Ao rejeitarem *para*, automaticamente se concluíu que a forma *pra* seria aceita, conforme a autora.

Além disso, a pesquisa de Gazola (2008) é a única que não se baseia na perspectiva da Teoria da Variação, método de análise adotado pelos outros trabalhos. Gazola, como já foi descrito no capítulo 4, segue as propostas de autores como Cinque (1993), Selkirk (1995) e Vigário (1999), os quais tratam de descrever estratégias de focalização na língua, a partir da interação entre Sintaxe e Fonologia. Portanto, em razão destas diferenças de metodologia e objeto de estudo, optamos por não expor os resultados de Gazola (2008) nas tabelas e gráficos comparativos de nossa análise.

Outra ressalva refere-se ao trabalho de Felgueiras (1993), abordado no estudo de Maya (2004). Mesmo tendo investigado a variante *pa*, a autora não incluiu os dados referentes a essa análise em seu texto, por não ter obtido resultados conclusivos.

De modo geral, é possível observar o que segue: 1) a variante *pra* foi a preferida pelos informantes na maioria dos estudos, exceto em Lucena (2006); 2) a variante *pa* ficou em segundo lugar na maioria das pesquisas, com percentual mais baixo – exceto em Lucena (2006) e Silva (2010); 3) a forma padrão *para* apresenta poucas ocorrências, evidenciando que é uma variante em desuso.

No Gráfico 2 a seguir pode-se visualizar, de outra forma ilustrativa, uma comparação de distribuição das variantes entre todos os trabalhos:

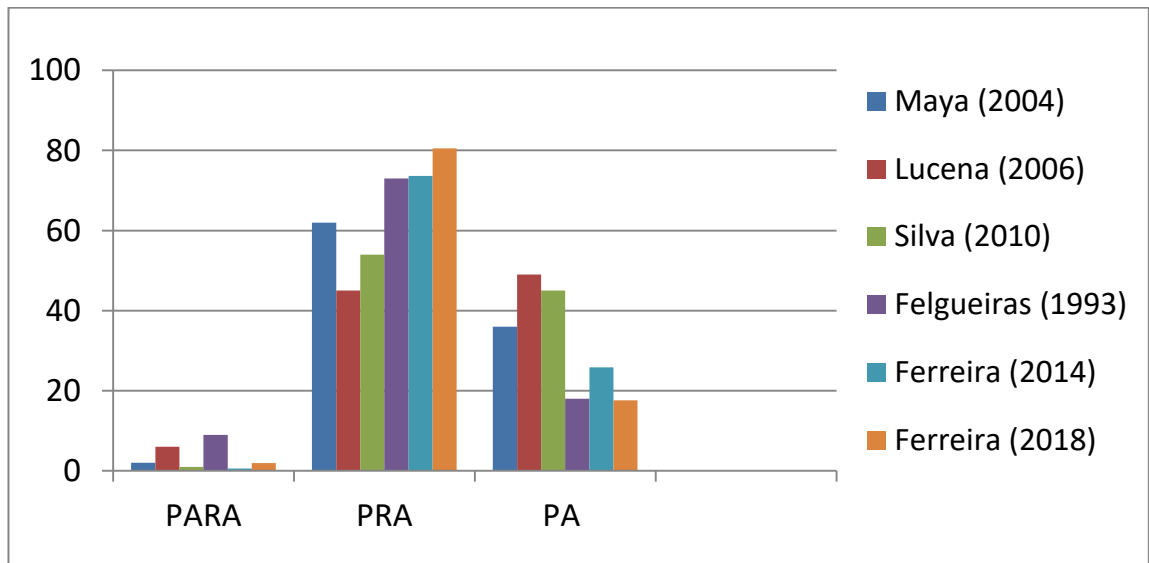


Gráfico 2 – Comparação dos Dados entre Pesquisas Variacionistas da Variável *para*. Fonte: a autora.

Pode-se observar, pelo Gráfico 2, que o percentual de uso das variantes da preposição *para* é semelhante entre todas as pesquisas, sendo a variante *pra* a favorita.

Em suma, dentre as três variantes da preposição *para* examinadas, há o predomínio, na língua falada de diferentes amostras, da forma *pra*, seguida por *pa* e, de modo muito menos frequente, por *para*.

A variante *para* parece dar lugar a formas reduzidas, na língua falada. Como *pra* é de uso mais frequente, essa variante foi escolhida como a forma de aplicação da regra variável (ver capítulo 5 – Metodologia).

## 6.2 ANÁLISE ESTATÍSTICA DA VARIANTE *PRA* E DISCUSSÃO DOS RESULTADOS

Ao confirmarmos nossa hipótese de que o percentual de ocorrências de *pra* seria a forma predominante, optamos por realizar uma análise binária, definindo como variável dependente a variante *pra*, já que o programa estatístico de análise de regra variável GOLDVARB não executa rodadas enéarias na aferição dos condicionamentos das variáveis independentes. Considerando-se os 2591 dados, o resultado da primeira rodada que realizamos tinha apontado para 2086 aplicações dessa variante, com um percentual de 80,5%.

A Tabela 6 a seguir apresenta os dados de forma mais clara.

<b>Aplicação/Total</b>	<b>%</b>
2086/2591	80,5

Tabela 6 – Aplicação Geral de *pra* – primeira rodada. Fonte: a autora.

Do total de 2591 contextos, houve 2086 aplicações de *pra*, totalizando um percentual de 80,5%, índice relativamente alto.

Para se eliminar problemas no tratamento dos dados, possivelmente gerados pelo programa estatístico GOLDVARB, foi necessária uma série de novas rodadas até a conclusão desta análise. Foram retirados alguns *knockouts*<sup>4</sup> e realizadas diversas tentativas de amalgamações e exclusões de fatores, inclusive de grupos inteiros, devido a percentuais de aplicação muito baixos em vários grupos de fatores.

Após as alterações mencionadas, obtivemos os dados que estão na Tabela 7 a seguir:

<b>Aplicação/Total</b>	<b>%</b>
1913/2349	81,4

Tabela 7 – Aplicação Geral de *pra* – com amalgamações e exclusões. Fonte: a autora.

O total de contextos, desta vez, foi reduzido a 2349 (devido à eliminação de alguns dados), dos quais 1913 foram realizados com a forma *pra*, totalizando 81,4%.

A seguir, procedemos à descrição das variáveis linguísticas selecionadas pelo programa GOLDVARB como possíveis condicionantes para a ocorrência da preposição *pra*. Apesar de *para* e *pa* serem variantes muito distintas em termos segmentais e prosódicos, juntamos os dados dessas duas formas, para confrontar com a variante *pra*, em virtude da baixa ocorrência de *para*, o que, estatisticamente, não interferirá nos resultados. No entanto, os poucos casos de *para* foram examinados manualmente. A ordem em que o programa selecionou as variáveis foi *Contexto Morfológico Seguinte*, *Contexto Fonológico Seguinte*, *Paralelismo*, *Idade*, *Cidade*, *Escolaridade* e *Número de Sílabas do Item Seguinte*.

<sup>4</sup> Fatores com poucos dados ou nenhum dado, conforme exposto na seção 5.1.1.

### 6.3 VARIÁVEIS LINGUÍSTICAS SIGNIFICATIVAS

As variáveis linguísticas controladas nesta investigação foram definidas com base em pesquisas anteriores sobre o mesmo fenômeno. Descreveremos a seguir o percurso dos resultados estatísticos de todos os grupos de fatores examinados.

Apesar de Guy e Zilles (2007) sugerirem a descrição inclusive dos resultados não selecionados como significativos pela análise estatística, já que eles também representam uma descoberta (conforme segue), não exporemos tal detalhamento.

[...] a prática de não apresentar resultados negativos tem um efeito negativo para o progresso da ciência: a falta de relato sobre um assunto acaba sendo ambígua: pode indicar que ninguém jamais pensou em investigar tal assunto, ou que sim, alguém investigou e teve resultados sem significância. A prática de pesquisadores em variação linguística deve ser, então, a de sempre descrever os fatores investigados, deixando claro quais deles obtiveram significância, e também quais deles deram resultados sem significância. (GUY E ZILLES, 2007, p. 214-215).

Os grupos de fatores selecionados como significativos pelo programa GOLDVARB, no entanto, serão apresentados mais detalhadamente.

#### 6.3.1 Contexto Morfológico Seguinte

No começo da análise, o Grupo *Contexto Morfológico Seguinte* estava subdividido em oito fatores: *Nome*, *Verbo*, *Numeral*, *Pronome*, *Artigo*, *Conjunção*, *Advérbio* e *Nada* (quando não há contexto morfológico seguinte, como uma pausa na fala, por exemplo). O quadro geral com todos os grupos de fatores encontra-se no Anexo D, ao final deste trabalho.

Ao longo da pesquisa, entretanto, os fatores *Numeral*, *Pronome*, *Artigo*, *Conjunção* e *Advérbio* foram amalgamados entre si e reclassificados como *Palavra Gramatical*, pois essas categorias resultaram em um percentual baixo de ocorrência em contextos seguintes (*pra ele*, *pra cá*). Pela mesma razão, foi eliminado o fator *Nada*. Os fatores *Nome* e *Verbo*, por sua vez, foram renomeados, respectivamente, como *Palavra Lexical Nominal* e *Palavra Lexical Verbal*.

Após uma nova rodada, houve desequilíbrio na distribuição de dados do fator *Palavra Gramatical* (1029) em relação aos fatores *Palavra Lexical Nominal* (642) e *Palavra Lexical*

*Verbal* (589). Por apresentarem, em relação ao fator *Palavra Gramatical*, metade da quantidade de contextos cada um, decidimos amalgamar os dados de nomes e verbos em um novo fator, o qual denominamos simplesmente *Palavra Lexical*.

De uma nova rodada, obtivemos os percentuais e pesos relativos que seguem na Tabela 8:

	<b>Aplicação/Total</b>	<b>%</b>	<b>Peso relativo</b>
Palavra Gramatical <i>pra mim, pra cá</i>	935/1058	<b>88,4</b>	<b>0.65</b>
Palavra Lexical <i>pras alunas, pra pegar</i>	978/1291	75,8	0.37

Input: 0.84

Tabela 8 – Variante *pra* considerando *Contexto Morfológico Seguinte* – com amalgamações. Fonte: a autora.

Com 88,4% de aplicações e peso relativo de 0.65, a categoria *Palavra Gramatical*, que engloba todos os casos de não verbos e não nomes, como pronomes, numerais, advérbios etc., mostrou-se mais significativa do que a *Palavra Lexical* (75,8%, peso relativo 0.37), em termos percentuais, para o uso de *pra*. Esses resultados poderiam indicar que o *pra* não é condicionado significativamente em termos de peso relativo se for seguido por uma palavra lexical.

Maya (2004, p. 36), apesar de investigar variante diferente da nossa como variável dependente, obteve peso relativo praticamente igual para verbos e nomes (0.55 e 0.54, respectivamente), fato que o levou à decisão de amalgamar estes fatores, decisão que nos incentivou a tomar a mesma atitude, tendo em vista que nosso resultado foi similar para a variável dependente que analisamos. O resultado final de Maya foi 43% (peso relativo 0.55) de aplicações de *pa* precedendo verbos e nomes e 30% (peso relativo 0.42) antes do fator *Outros* (numerais, advérbios, pronomes etc.). O autor conclui que a variante *pa* é favorecida quando é seguida de uma palavra lexical (nomes e verbos) e desfavorecida quando a ela se segue uma palavra gramatical.

Maya (2004, p. 37-38) faz uma observação, a respeito da categoria *Outros*, a qual consideramos oportuno destacar: ela representa um percentual de 67% de aplicações após a variante *pra* e 29% seguindo a variante *pa*. Quando apresenta os dados referentes à rodada

eneária de onde retirou esses dados, o autor afirma que houve grande desencontro entre as frequências e os pesos relativos das três variantes. Por esse motivo, esclarece que apenas os percentuais, mas não os pesos, seriam levados em conta na análise desta rodada. Contudo, mesmo desconsiderando os pesos relativos, Maya (2004) conclui que o fator *Outros favoreceu* o uso de *pra*, como é o caso de nossa pesquisa, e *desfavoreceu pa*.

Sabe-se que o favorecimento, segundo Guy e Zilles (2007, p. 211), é determinado não pelos percentuais, mas pelo peso relativo, que é o cálculo dos efeitos dos fatores de cada grupo em relação ao nível geral de ocorrência da variante examinada. Já o percentual, conforme os autores, representa apenas a frequência de ocorrência de cada variante no contexto analisado, sem considerar a distribuição dos dados em relação aos outros grupos de fatores.

Para se indicar favorecimento, portanto, não se pode ter como base apenas os percentuais de aplicação. É necessário haver um peso relativo acima do ponto neutro<sup>5</sup>, que é 0.50, sendo, os valores abaixo deste, considerados desfavorecedores para a aplicação de uma regra variável. Por essa razão, não se pode afirmar que os resultados de Maya (2004) sobre a escolha da variante *pra* coincidem com os nossos, em relação ao condicionamento do *Contexto Morfológico Seguinte*. Assim, devido à opção por examinar uma variante distinta da nossa, não compararemos os resultados das variáveis controladas pelo autor com os nossos.

Em nosso estudo anterior, de 2014, a categoria *Outros* (numerais, advérbios, pronomes etc.) tinha se mostrado mais favorável ao uso de *pra*. Com peso relativo de 0.57 para palavras gramaticais, também se concluiu, em Ferreira (2014), que o *pra* não necessita realizar processo de junção com palavras lexicais – verbais (0.51) e, principalmente, nominais (0.39).

Lucena (2006) e Silva (2010) não analisaram o contexto morfológico seguinte à variável dependente. Portanto, não será possível fazer comparações desta variável com esses estudos.

---

<sup>5</sup> Conforme visto na seção 5.1.1, *ponto neutro* é o valor matemático de um fator que não favorece nem desfavorece o uso da regra variável analisada. (GUY e ZILLES, 2007, p. 239).

### 6.3.2 Contexto Fonológico Seguinte

As vogais posteriores /u, o, ɔ/ e a vogal central /a/ apresentaram uma baixa quantidade de contextos (124 e 118, respectivamente)<sup>6</sup> nos dados em relação a quase todas as consoantes e às vogais anteriores, sendo que estas, com mais casos, mostraram-se mais favoráveis para a realização de *pra*. Portanto, como não apresentavam, separadamente, grande diferença em um possível condicionamento que poderiam exercer, as vogais posteriores e a vogal central foram amalgamadas, criando-se um novo fator chamado de *Vogal Dorsal*, nomeado de acordo com um traço fonético compartilhado por elas. O fator antes classificado como *Vogal Anterior*, então, foi renomeado para *Vogal Coronal*.

As consoantes, de modo geral, apresentaram diferença na quantidade de contextos nos dados. Alguns valores de aplicação, no entanto, mostraram-se aproximados em termos de peso relativo, evidenciando que não haveria favorecimento destes fatores. Em uma análise inicial, as velares (*pra cada*), as dentais (*pra tudo*) e as bilabiais (*pra puxar*) se destacaram, com valores de aplicação de 84%, 79,2% e 73,4%, respectivamente. A consoante labiodental, apesar de um percentual de aplicação de 81%, teve quantidade mais baixa de contextos (279) em relação aos fatores mencionados acima. A consoante palatal (*janela*) apareceu em quantidade ainda menor (94 casos), sendo 87 realizações iniciando uma palavra que sucedia a variante *pra*.

Por não apresentarem influência na escolha da variante *pra*, devido à baixa quantidade de dados, as consoantes labiodental e palatal foram amalgamadas com outros fatores. A primeira juntou-se à consoante labial e a segunda, à dental. Os novos fatores gerados a partir destas junções foram nomeados, respectivamente, como *Consoante Labial* e *Consoante Coronal*. Partindo-se desta alteração, a consoante velar passou a se chamar *Consoante Dorsal*. Após alguns testes, devido à proximidade articulatória, optou-se por amalgamar as consoantes *Labial* e *Coronal*, com o objetivo de melhorar os resultados estatísticos. Assim, originou-se o novo fator *Consoante [- posterior]*.

---

<sup>6</sup> Os dados referentes a esta rodada encontram-se no Anexo B, no final deste trabalho.



O resultado final para este grupo de fatores está exposto na Tabela 9 abaixo:

	<b>Aplicação/Total</b>	<b>%</b>	<b>Peso relativo</b>
Vogal Coronal <i>pra ela, pra isso</i>	309/341	90,6	<b>0.65</b>
Consoante Dorsal <i>Pra casa, pro grupo</i>	305/361	84,5	<b>0.56</b>
Consoante [- posterior] <i>Pra morar, pra frente, pro doutor</i>	1136/1416	80,2	0.48
Vogal Dorsal <i>Pra ajudar, pra ônibus, pra uma</i>	163/231	70,6	0.29

Input: 0.84

Tabela 9 – Variante *pra* considerando *Contexto Fonológico Seguinte* – com amalgamações. Fonte: a autora.

Com 90,6% de aplicações e peso relativo de 0.65, a categoria *Vogal Coronal* apresentou mais favorecimento para a realização da variante *pra*. Em segundo lugar, vem a *Consoante Dorsal*, com percentual de 84,5% e peso relativo de 0.56. Os demais fatores, *Consoante [- posterior]* e *Vogal Dorsal*, apesar de percentuais de 80,2% e 70,6%, respectivamente, obtiveram peso relativo abaixo do ponto neutro: 0.48 e 0.29.

Em Maya (2004, p. 41), verifica-se um favorecimento das consoantes dentais (peso relativo 0.63) e, de forma moderada, das bilabiais e das palatais (0.57 cada fator) na escolha da variante *pa*. Já as consoantes velares (0.44) e labiodentais (0.39) apresentaram desfavorecimento para o uso desta variante. Por fim, as alveolares (0.51) não exercem nenhuma influência na aplicação do fenômeno investigado. Maya (2004) conclui, sobre seus resultados, que existe uma tendência à repetição de segmentos próximos, tanto no caso da escolha de sua variável dependente *pa*, quanto na escolha de *pra*.

Lucena (2006) também menciona a possível tendência à repetição de sons semelhantes, chamada *aliteração*, como responsável pelos resultados encontrados. O autor, no entanto, expõe uma análise baseada no confronto entre a variante padrão *para* e as duas variantes *pra* e *pa*, juntas. Para a escolha das formas reduzidas, houve mais influência da consoante labiodental, assim como em nosso estudo, com 99% de aplicação, peso de 0.81. Em segundo

lugar, estão os fatores *consoante alveolar* e *consoante dental*, com percentuais idênticos de 96% e peso de 0.59 cada um. Devido à junção de *pra* e *pa* que Lucena (2006) realizou, não será possível comparar os resultados específicos para a variante *pra*, desta e das demais variáveis estudadas, com os de nosso estudo.

Silva (2010) também obteve resultados que mostram favorecimento da consoante dorsal (61%, peso 0.53) e da vogal anterior (69%, peso 0.60) para a escolha da forma *pra*. Além disso, a autora destaca o valor de aplicação da vogal posterior, que foi de 73% (peso 0.66). Apesar de ser mais alto do que os demais, é preciso fazer uma observação sobre este resultado: a quantidade de dados deste fator foi baixa, em comparação aos outros mais frequentes, com apenas 71 contextos, sendo que a consoante dorsal e a vogal anterior têm, respectivamente, 499 e 509 ocorrências. Neste caso, seria necessária uma quantidade maior de dados deste fator para se tirar conclusões mais precisas a respeito de sua influência na aplicação da variável dependente em questão.

Em Ferreira (2014), o resultado final mostra pouca diferença entre os fatores *Consoante* e *Vogal*, de modo geral. Com 74,6% para consoantes e 71,9% para vogais, sendo os pesos relativos idênticos de 0.50, concluiu-se que não há significância neste Grupo, já que o peso para os dois fatores encontra-se no ponto neutro.

Resumindo, o contexto fonológico seguinte *Vogal Coronal* (*pra ela, pra isso*) favorece o uso de *pra* (0.65), e o contexto *Vogal Dorsal* (*pra ajudar, pra uma*) desfavorece (0.29). Os demais contextos fonológicos seguintes são neutros em relação ao uso da variante *pra*.

Um aspecto a observar, que talvez explique esse comportamento de vogais terem papéis contrários (coronais favorecendo; dorsais, desfavorecendo) pode estar relacionado ao acento da vogal que segue a preposição *pra*, condição exigida para a realização de sândi vocálico.

A fim de confirmar se a maioria dos dados de vogais dorsais seria de palavras com a sílaba seguinte tônica ou átona, realizamos um cruzamento entre as variáveis *Contexto Fonológico Seguinte* e *Tonicidade da Sílaba Seguinte*, que pode ser visualizado na Tabela 10 a seguir:

	Vogal Dorsal <i>pra ônibus</i>	Consoante Dorsal <i>pro grupo</i>	Consoante [- posterior] <i>pra frente</i>	Vogal Coronal <i>pra isso</i>
Tônica <i>Pra praia</i>	65/76 = 86%	81/92 = 88%	301/357 = 84%	202/213 = <b>95%</b>
Tônica monos. <i>Pra nós</i>	22/53 = 42%	104/107 = <b>97%</b>	419/503 = 83%	33/39 = 85%
Átona (2 síl.) <i>Pra Curitiba</i>	31/50 = 62%	41/54 = 76%	49/75 = 65%	25/29 = 86%
Átona (1 síl.) <i>Pra valer</i>	38/42 = 90%	74/101 = 73%	341/451 = 76%	40/48 = 83%

Tabela 10 – Cruzamento entre as variáveis *Contexto Fonológico Seguinte* e *Tonicidade da Sílaba Seguinte*.  
Fonte: a autora.

Na Tabela 10, pode-se verificar que, ao seguirem a variante *pra*, a maioria das palavras iniciadas por vogais coronais carrega, ao mesmo tempo, a primeira sílaba tônica. Os fatores *Vogal Coronal* e *Tônica*, combinados, representam um percentual de 95% de aplicação, evidenciando que há interação<sup>7</sup> entre as duas variáveis linguísticas. Também chama a atenção, apesar da menor quantidade de contextos, o percentual de 97% de aplicação de *pra* antes de palavras monossílabas tônicas iniciadas por consoantes dorsais.

Bisol (2002, p. 231) verifica, em dados de língua falada, que os processos de ressilabação com degeminação e elisão entre duas palavras são motivados por uma restrição rítmica e segmental. Para Bisol (2010, p. 125): “A elisão se aplica geralmente quando a vogal seguinte for posterior e opcionalmente quando for frontal. (...) A degeminação ocorre quando as

<sup>7</sup> Para Brescancini (2002, p. 55), *interação* é um tipo de interferência nos dados, gerada pela força de atuação de um fator específico de uma variável. Ou seja, em vez de ocorrer a atuação conjunta de todos os grupos de fatores, observa-se apenas a relação entre duas categorias na definição dos resultados. Como exemplos de interação, têm-se a interferência entre as variáveis *Contexto Fonológico Seguinte (Vogal Coronal)* e *Tonicidade da Sílaba Seguinte (Tônica)*, verificada nesta investigação, e entre as variáveis *Idade* e *Escolaridade*, quando os informantes mais velhos e com mais estudo aplicam mais o fenômeno em questão.

duas vogais que se encontram são semelhantes, desde que a segunda vogal não tenha acento primário.”.

Bisol (2010, p. 125) afirma que a elisão afeta a vogal /a/, ainda que também possa ocorrer com outras vogais, conforme ilustra nos exemplos com as junções “*merend[e]scolar*” e “*resistênci[o]rgânica*”. Além disso, esse fenômeno aplica-se somente na fronteira entre palavras, mas não no interior delas, como *gauchada* > \**g[u]chada*<sup>8</sup>, ou entre morfemas, como *paraense* > \**pareense*. Sobre a degeminação, além de se aplicar entre palavras, como nos exemplos “*menin[a]legre*” e “*cami[za]marela*”, Bisol (2010) acrescenta que esse processo pode ocorrer também no interior de uma palavra, como em *coordenador* > *c[o]rdenador*.

Considerando-se que a elisão ocorre apenas se as duas vogais forem átonas e que a degeminação exige que seja átona a segunda vogal, justificam-se os dados expostos na Tabela 10. De acordo com Bisol (2010, p. 126), esses dois processos compartilham o fato de acontecerem entre dois núcleos silábicos que entram em contato. Uma vez que essa configuração costuma ser rejeitada pelo português, o núcleo que for prosodicamente mais fraco acaba desaparecendo – como é o caso da primeira vogal, que normalmente é átona final. Portanto, no caso de expressões como “*pra ela*” e “*pra aula*”, o fato de a primeira sílaba da palavra seguinte ao *pra* ser tônica impede a realização de sândi vocálico.

### 6.3.3 Paralelismo

Esta variável, inicialmente, estava subdividida em cinco fatores. Na primeira rodada enéaria, verificamos valores de ocorrência muito baixos, tanto de contextos quanto de aplicações, para os fatores *Antecedida de para* e *Antecedida de pa* (ou seja, quando a preposição alvo estava sucedendo, na mesma frase, uma forma realizada como *para* ou *pa*). Para este fator, havia 87 contextos; para aquele, apenas 7. Por isso, os dois fatores foram eliminados da análise. A seguir, já na primeira rodada binária de *pra*, observaram-se resultados muito próximos entre os fatores restantes, *Ocorrência isolada* (sem paralelismo), *Primeira da série* e *Antecedida de pra*.

Em outra rodada da análise, decidimos juntar os resultados de *Ocorrência isolada* (quando não há paralelismo) e *Primeira da série*. A explicação para essa amalgamação é que

<sup>8</sup> O símbolo “\*” indica, nesse caso, uma construção impossível de ocorrer na fala.

houve uma baixa quantidade de dados do fator *Primeira da série*. Além disso, esse fator só poderia ser amalgamado aos dados em que não houve paralelismo, devido a questões de semelhança: ao se analisar o primeiro *pra* de uma frase, tecnicamente, o segundo *pra* ainda não existe. Por outro lado, o segundo *pra* pressupõe a existência de um primeiro, motivo por que não seria possível amalgamar o fator *Primeira da série* ao fator *Antecedido de pra*.

A junção entre esses dois fatores gerou os resultados apresentados na Tabela 11 a seguir:

	<b>Aplicação/Total</b>	<b>%</b>	<b>Peso relativo</b>
<i>Antecedido de pra</i>	374/422	88,6	<b>0.64</b>
<i>Ocorrência isolada/primeiro da série</i>	1539/1927	80	0.46

Input: 0.84

Tabela 11 – Variante *pra* considerando *Paralelismo* – com amalgamações. Fonte: a autora.

Os valores expostos na Tabela 11 evidenciam preferência pela variante *pra* quando ela é antecedida por ela mesma em uma frase (“*pra* você, *pra* tua família...”), com um percentual de 88,6% de aplicação e peso relativo de 0.64. Nota-se que, apesar de a maior quantidade de contextos estar em *Ocorrência isolada/primeiro da série* (1927), é o fator *Antecedido de pra* que apresenta um percentual maior (88,6%).

A junção entre *Ocorrência isolada* e *Primeira da série*, realizada também por Maya (2004), justifica-se pela pouca diferença prática entre os dois fatores. Quando o primeiro *pra* da frase é realizado, ainda não se sabe se haverá um segundo *pra* ou não, diferentemente de quando se trata do segundo *pra* desta mesma frase, situação que classificamos como *Antecedida de pra*. No caso desse fator, como já se conhece o seu contexto anterior, não haveria razão para que ele se envolvesse em qualquer amalgamação.

Em Maya (2004, p. 62), a variável dependente *pa* é favorecida quando é antecedida por ela mesma (*Antecedido de pa*). O autor justifica esse favorecimento expondo o peso relativo de 0.57, contra 0.51 para o fator *Sem Antecedente* e 0.46 para *Antecedido de pra*.

Em Lucena (2006), evidencia-se a mesma tendência de repetição de formas semelhantes, já verificada em estudos sobre várias línguas, conforme o autor. Quando há um *pra*, tende-se a utilizar outro *pra* a seguir, o que também ocorre com a forma reduzida *pa*, mais comumente seguida por outro *pa*. Ainda que não esteja muito claro em sua pesquisa, é

possível inferir que esses resultados ilustram o que segue: o fator *Antecedido de pra* gerou um percentual de 97% de aplicação (peso relativo de 0.68) em relação à variante *pra*; o fator *Antecedido de pa* gerou um percentual de 99% de aplicação (peso relativo de 0.84) em relação à variante *pa*.

Ferreira (2014) classificou os fatores desta variável como *Sem Paralelismo* e *Com Paralelismo* (junção entre *Primeira da série* e *Antecedida de pra*). O primeiro fator gerou um peso relativo abaixo do ponto neutro (0.48), bem como em nossa pesquisa atual. O segundo fator, embora com peso acima do ponto neutro (0.55), não se distancia muito do valor do primeiro, diferença que não representa um condicionamento claro do *Paralelismo* em Ferreira (2014). Já em Silva (2010), o grupo *Paralelismo* não foi selecionado como significativo pelo programa GOLDVARB.

A situação não é a mesma em nosso trabalho, em que o peso relativo para o fator *Antecedida de pra* foi de 0.64, mostrando-se, pois, relevante a Variável Independente *Paralelismo*.

#### 6.3.4 Número de Sílabas do Item Seguinte

A variável *Número de Sílabas do Item Seguinte* analisada só se mostrou relevante com a amalgamação de fatores, tendo em vista o desequilíbrio de dados. Foram amalgamados os contextos em que a variável dependente *pra* era seguida por vocábulos com duas ou mais sílabas (*pro centro*; *pra transportar*), deixando-se o fator *Uma sílaba (pra mim)*. Esse procedimento gerou resultados mais expressivos em termos de pesos relativos, conforme pode ser visto na Tabela 12.

	<b>Aplicação/Total</b>	<b>%</b>	<b>Peso relativo</b>
Duas ou mais sílabas <i>Pro lado, pra trabalhar, pra prefeitura</i>	1334/1643	81,2	<b>0.53</b>
Uma sílaba <i>Pra dar, pra nós</i>	579/706	82	0.43

Input: 0.84

Tabela 12 – Variante *pra* considerando *Número de Sílabas do Item Seguinte* – com amalgamações. Fonte: a autora.

O fator *Duas ou mais sílabas* mostrou-se um pouco mais favorável à realização da variante *pra*, com peso relativo de 0.53 (percentual 81,2%), o que pode representar uma preferência moderada de a preposição reduzida *pra* ocorrer antes de uma palavra com duas ou mais sílabas. No entanto, os pesos relativos de ambos fatores encontram-se próximos (diferença de 0.10 pontos) entre si e de um ponto neutro (0.50).

É preciso considerar uma questão relacionada ao grupo *Tonicidade da Sílabas Seguinte*. Esse grupo apresentou sobreposição<sup>9</sup> com o *Número de Sílabas do Item Seguinte*, devido a muitas células vazias, situação em que se faz necessário realizar as rodadas de toda a análise com apenas um dos grupos. Neste caso, optamos por incluir o *Número de Sílabas do Item Seguinte* e retirar a *Tonicidade da Sílabas Seguinte*.

Para ilustrar o problema de células vazias mencionado, expomos o cruzamento entre as variáveis *Número de Sílabas do Item Seguinte* e *Tonicidade da Sílabas Seguinte* na Tabela 13 a seguir:

	<i>Duas sílabas</i>	<i>Uma sílaba</i>	<i>Três sílabas</i>	<i>Pausa</i>	<i>Quatro ou +</i>
<i>Tônica</i>	<b>644/737</b> 87%	1/1 100%	0/1 0%	---	347/471 74%
<i>Tônica monossil.</i>	2/4 50%	<b>586/707</b> 83%	0/1 0%	---	---
<i>Átona (2 sílabas)</i>	9/11 82%	---	111/162 69%	---	144/187 77%
<i>Pausa</i>	---	---	---	120/135 89%	---
<i>Átona (1 sílaba)</i>	---	---	38/53 72%	---	1/1 100%
<i>Átona (3 sílabas)</i>	---	---	---	---	49/62 79%

Tabela 13 – Cruzamento entre as variáveis *Número de Sílabas do Item Seguinte* e *Tonicidade da Sílabas Seguinte*.  
Fonte: a autora.

<sup>9</sup> Segundo Guy e Zilles (2007, p. 60), sobreposição é um tipo de distribuição que frequentemente passa despercebido, mas que pode distorcer os resultados analíticos. No caso da Tabela 13, há concentração de dados entre fatores de duas variáveis diferentes: *Tônica com Duas Sílabas* (644/737); *Tônica Monossilábica com Uma Sílaba* (586/707).

A Tabela 13 mostra células vazias (ausência de dados) em vários fatores. Em alguns casos, tratam-se de combinações impossíveis de ocorrer, como uma palavra de uma sílaba que tivesse duas sílabas átonas, uma palavra tônica monossilábica com quatro ou mais sílabas ou mesmo a combinação de qualquer palavra com uma pausa.

Em uma etapa analítica seguinte, decidimos realizar uma rodada que incluía a *Tonicidade da Sílaba Seguinte* e excluía o *Número de Sílabas do Item Seguinte*. Essa iniciativa visava testar a possibilidade de se melhorar a significância estatística dos dados, o que não foi alcançado. A variável *Tonicidade* foi selecionada como significativa – bem como o *Número de Sílabas do Item Seguinte*, em uma rodada anterior, também em último lugar na ordem de grupos. No entanto, essa alternância não trouxe qualquer benefício para nossa análise, motivo pelo qual mantivemos a rodada que considerava o *Número de Sílabas do Item Seguinte* para a apresentação dos resultados.

Maya (2004, p. 67) encontrou o mesmo problema de sobreposição entre este Grupo e *Tonicidade da Sílaba Seguinte*. O autor, portanto, também executou rodadas com os dois grupos separadamente, para poder interpretar os dados de forma mais eficiente.

Maya (2004, p. 69) obteve resultados que indicam favorecimento de *pa* na presença de palavras prosódicas com três ou mais sílabas (0.56). O fator *2 sílabas* não interfere na escolha desta variante, já que seu peso relativo foi de 0.50. Por fim, palavras de uma sílaba, após a preposição reduzida *pa*, desfavorecem a sua aplicação (0.45).

Em Lucena (2006), este grupo de fatores não foi selecionado pelo programa estatístico VARBRUL. Com exceção do fator *dissílabo* (0.57), todas as categorias mantiveram-se com pesos relativos abaixo do ponto neutro, o que evidencia, segundo o autor, a irrelevância do número de sílabas do item seguinte para a escolha da variante *pra* em seu estudo.

Ferreira (2014), para investigar esta variável, utilizou uma categorização diferente deste estudo: depois de algumas amalgamações, optou por classificar os dados em palavras com *Até 3 sílabas* e palavras com *4 ou mais sílabas*. Dessa forma, torna-se mais difícil comparar os resultados dessa pesquisa com a nossa. De qualquer forma, não foi esse o fator mais favorável no trabalho mencionado, para a aplicação de *pra*, e sim a categoria *4 ou mais sílabas*, com percentual de 81,2%, peso relativo de 0.66. Já em Silva (2010), não foi possível comparar os resultados desta variável, uma vez que a autora não a incluiu em sua análise.

Em suma, os resultados mostram que o uso da preposição reduzida *pra* em nossa pesquisa é favorecido por palavras gramaticais, vogais coronais tônicas e contextos em que



essa variante é antecedida por ela mesma em uma frase. Já o fato de uma palavra ter duas ou mais sílabas exerce pouca influência na escolha de *pra*.

Na próxima seção, apresentaremos os resultados obtidos para as variáveis extralinguísticas/sociais.

#### 6.4 VARIÁVEIS EXTRALINGUÍSTICAS

Descrevemos, a seguir, os resultados das três variáveis extralinguísticas/sociais que foram selecionadas em nossa pesquisa, por ordem de significância.

##### 6.4.1 Idade

A variável *Idade* foi dividida em - de 50 anos e + de 50 anos, a mesma estratificação etária adotada pelo Projeto VARSUL. A Tabela 14 a seguir apresenta a distribuição total de dados para cada fator e seus respectivos percentuais e pesos relativos.

	<b>Aplicação/Total</b>	<b>%</b>	<b>Peso relativo</b>
+ de 50 anos	920/1082	85	<b>0.56</b>
- de 50 anos	993/1267	78,4	0.44

Input: 0.84

Tabela 14 – Aplicações de *pra* considerando a *Idade*. Fonte: a autora.

Apesar de os informantes com menos de 50 anos apresentarem maior quantidade de dados, são os maiores de 50 anos que utilizam mais a forma *pra*. De um total de 1082 contextos, houve 920 aplicações desta variante, com um percentual de 85% (peso relativo de 0.56), o que pode representar um favorecimento relativamente moderado desta faixa etária em optar pela forma reduzida *pra*. Os informantes menores de 50 anos, por sua vez, parecem desfavorecer a sua escolha, devido ao peso relativo abaixo do ponto neutro: 0.44, apesar do percentual de 78,4%.

Maya (2004, p. 94), ao analisar a aplicação de *pa*, diferentemente de nossa análise, encontrou os indivíduos mais jovens como os mais propensos a utilizarem essa variante. Isso

sugere que talvez, no caso da variante *pra*, possam ser os mais velhos os que mais exerçam papel, inferência essa que permite aproximar os resultados de ambas pesquisas.

Em Lucena (2006), houve preferência moderada da faixa etária intermediária, de 26 a 49 anos (0.58), e desfavorecimento dos mais jovens (15 a 25 anos) pela variante *pra*. Os falantes maiores de 50 anos, com peso relativo de 0.52, não representam praticamente nenhuma influência na escolha dessa forma reduzida da preposição.

Silva (2010), que classificou a idade dos informantes em três faixas etárias em sua análise, afirma que os mais jovens (15 a 25 anos) e os mais velhos (+ de 49 anos) apresentaram percentual quase idêntico de realização da variante *pra*, com peso relativo de 0.55 e 0.56, respectivamente. Conforme a autora, a explicação de os mais jovens escolherem *pra* em vez de *pa* pode ser a de que estariam submetidos à interferência da escola, já que alguns livros didáticos registram a variante *pra*. No caso da faixa etária intermediária (26 a 49 anos), a ocorrência diminui (0.46). Ao observar-se o retorno do favorecimento de *pra* entre os mais velhos, Silva indica que o fenômeno representaria uma situação de variação estável.

Em nossa pesquisa anterior (FERREIRA, 2014), a situação é semelhante à desse trabalho, pois os informantes mais velhos também favoreceram a utilização de *pra*. A diferença entre os pesos relativos das duas faixa etárias, no entanto, foi bem maior: maiores de 50 anos representam 82,7% (peso de 0.64) de aplicações, enquanto os menores de 50 anos desfavorecem a aplicação do fenômeno, com um peso relativo de 0.39.

O papel da idade em nosso estudo poderá ser confirmado na próxima seção, com os cruzamentos entre fatores sociais, para se esclarecer que tipo de jovem estaria favorecendo a aplicação de *pra*. Preliminarmente, podemos conjecturar também que o fato de os mais velhos favorecerem o uso de *pra* parece indicar que essa variante não seja a forma mais inovadora.

#### **6.4.2 Escolaridade**

A Escolaridade parece desempenhar um papel moderado, já que os valores dos pesos relativos entre as duas faixas escolares estão relativamente próximos entre si e perto do ponto neutro (0.50). Esta variável indica o favorecimento de *pra* entre os informantes com até 11 anos de estudo.

A seguir, apresenta-se a Tabela 15, que descreve os resultados:

	<b>Aplicação/Total</b>	<b>%</b>	<b>Peso relativo</b>
Até 11 anos de estudo	1132/1350	84	<b>0.53</b>
Até 4 anos de estudo	781/999	78,2	0.45

Input: 0.84

Tabela 15 – Aplicações de *pra* considerando a *Escolaridade*. Fonte: a autora.

Os dados referentes aos informantes com até quatro anos de estudo, além de apresentarem menos contextos de ocorrências (999), indicam um desfavorecimento dessa variante, com peso relativo de 0.45.

Maya (2004, p. 81) realizou sua investigação com três graus de escolaridade: primário (de 1 a 4 anos), secundário (de 9 a 11 anos) e superior, distinguindo-se do presente estudo no que diz respeito a esse último nível escolar. Os resultados do autor mostram que a maior parte dos indivíduos que estudaram por até 11 anos preferiu a variante *pa* (peso relativo 0.61).

No estudo de Lucena (2006), a escolaridade dos informantes destacou-se com bastante significância, de acordo com o autor. Cabe salientar, porém, que Lucena dividiu esta variável em dois fatores extremos: *universitários* e *analfabetos*. Segundo os seus dados, a forma reduzida *pra* foi eleita quase categoricamente entre os analfabetos, com peso relativo de 0.86. Os falantes mais escolarizados, portanto, dariam pouquíssima preferência às formas reduzidas *pra* e *pa*.

Para Silva (2010), o papel da escolaridade foi decisivo em sua investigação. Os informantes de sua amostra estão divididos em três graus de escolaridade. Informantes com um grau mais alto de escolaridade optaram mais pela variante *pra* (peso relativo 0.61). Em segundo lugar, estão os de grau médio (0.50) e, por último, os de grau baixo (0.37). Observa-se, a partir dos resultados da autora, que os informantes do grau intermediário de escolarização não mostraram preferência por nenhuma das duas variantes. Tanto em *pra* quanto em *pa*, há um peso relativo idêntico (0.50). Os menos escolarizados, em compensação, parecem desfavorecer o uso de *pra*, resultado esse divergente de nossa pesquisa.

Em Ferreira (2014), também são os entrevistados com mais anos de estudo que utilizam mais a variante *pra*, com percentual de 79%, peso relativo de 0.57. Por outro lado, os que estudaram menos revelam uma tendência a não usá-la, devido ao peso relativo de 0.45.

Esses resultados parecem indicar que o uso de *pra* ou de *pa* não está sendo muito influenciado pela escolaridade, já que a diferença entre os graus de estudo não é distante (0.53 e 0.45).

### 6.4.3 Cidade

A variável *Cidade* indica preferência pelo uso de *pra* na capital Curitiba. Em Florianópolis, parece haver desfavorecimento para o uso dessa variante. Os resultados para esse Grupo também se mostram moderados, pois os valores de peso relativo estão próximos entre si (0.55 e 0.45) e ao redor do ponto neutro (0.50), conforme pode ser visto na Tabela 16:

	<b>Aplicação/Total</b>	<b>%</b>	<b>Peso relativo</b>
Curitiba	965/1143	84,4	<b>0.55</b>
Florianópolis	948/1206	78,6	0.45

Input: 0.84

Tabela 16 – Aplicações de *pra* considerando a *Cidade*. Fonte: a autora.

Assim, segundo a Tabela 16, os resultados mostram que os informantes da cidade de Curitiba (peso relativo de 0.55) utilizaram mais a forma *pra* do que os informantes de Florianópolis (peso relativo de 0.45). A diferença não é tão representativa, já que o peso referente à capital do Paraná não está muito acima do ponto neutro. Cabe aos cruzamentos entre fatores a tarefa de especificar o condicionamento desta variável, situação que poderá ser confirmada na próxima seção.

Das variáveis sociais selecionadas como significativas para a aplicação do fenômeno de variação investigado, esta é a única que não será comparada com outros trabalhos, pois não foi controlada por nenhum deles. Ainda assim, julgamos que seria relevante verificar se uma

cidade poderia apresentar mais tendência que outra na escolha de uma forma em detrimento de outra.

Em Ferreira (2014), que investigou a cidade de Londrina (PR), a preposição reduzida *pra* foi utilizada em 73,6% dos casos. Em Porto Alegre (RS), na pesquisa de Maya (2004), o uso da variante *pra* também foi mais frequente, com 62% de aplicação. No estudo de Silva (2010), em Araguatins (TO), essa variante foi a preferida, ainda que em menor frequência, com 54% de aplicação. Já em João Pessoa (PB), na pesquisa de Lucena (2006), o *pra* teve um percentual de 45% de uso, em detrimento de *pa*, variante preferida nesse estudo, com 49% de aplicação.

#### 6.4.4 Quantidade de PRA por informante

Para verificar o comportamento de cada informante em relação ao uso da variante *pra*, elaboramos um gráfico para cada cidade investigada, de modo a ilustrar o uso de cada indivíduo.

No Gráfico 4, há a quantidade de ocorrências dessa variante na cidade de Curitiba:

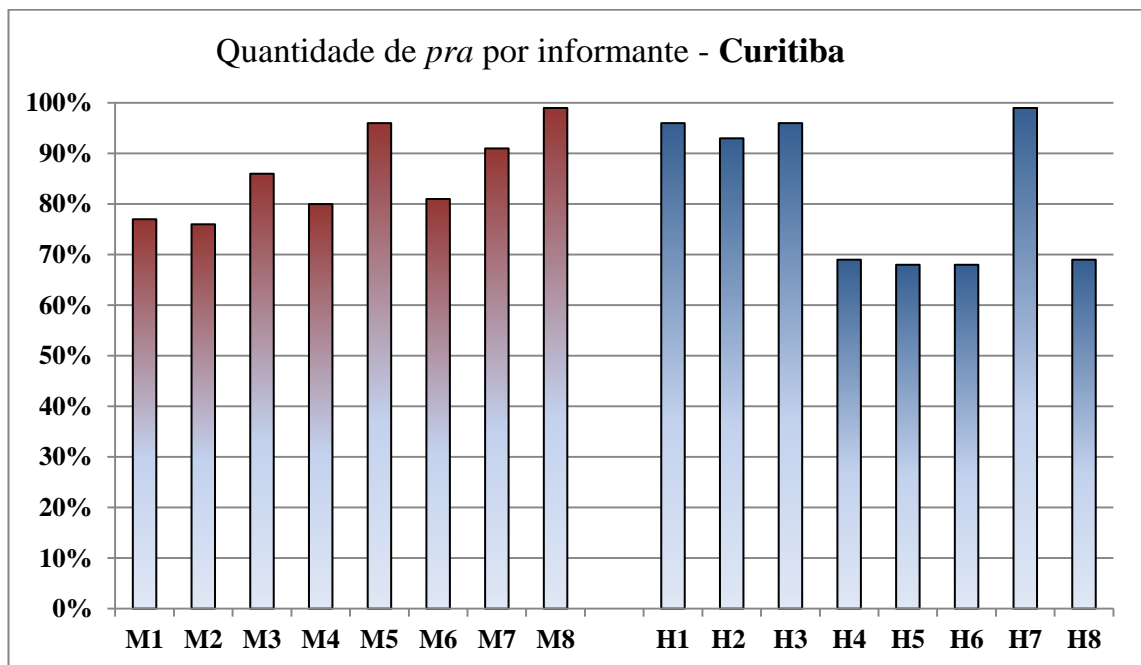


Gráfico 4 – Aplicações de *pra* por informante na cidade de Curitiba. Fonte: a autora.

O Gráfico 4 mostra que não há uma diferença significativa, em termos de percentual, entre os entrevistados do sexo feminino da cidade de Curitiba, de modo geral. É possível observar, no máximo, uma pequena diferença de percentual no caso de alguns informantes específicos do sexo masculino, como H4, H5, H6 e H8, que se sobressaem aos demais por utilizarem em torno de 60% de *pra*.

Esses quatro indivíduos mostram-se de forma equilibrada em relação à idade e à escolaridade: há dois da faixa etária mais jovem e dois da mais velha, além de que dois deles têm mais e dois têm menos escolarização, respectivamente. Tampouco pode-se afirmar que os homens favoreçam a aplicação do fenômeno, visto que a variável *Sexo* gerou percentuais idênticos (81%) para homens e mulheres, não exercendo, portanto, qualquer condicionamento.

A seguir, apresentamos o Gráfico 5, em que há a quantidade de ocorrências de *pra* na cidade de Florianópolis:

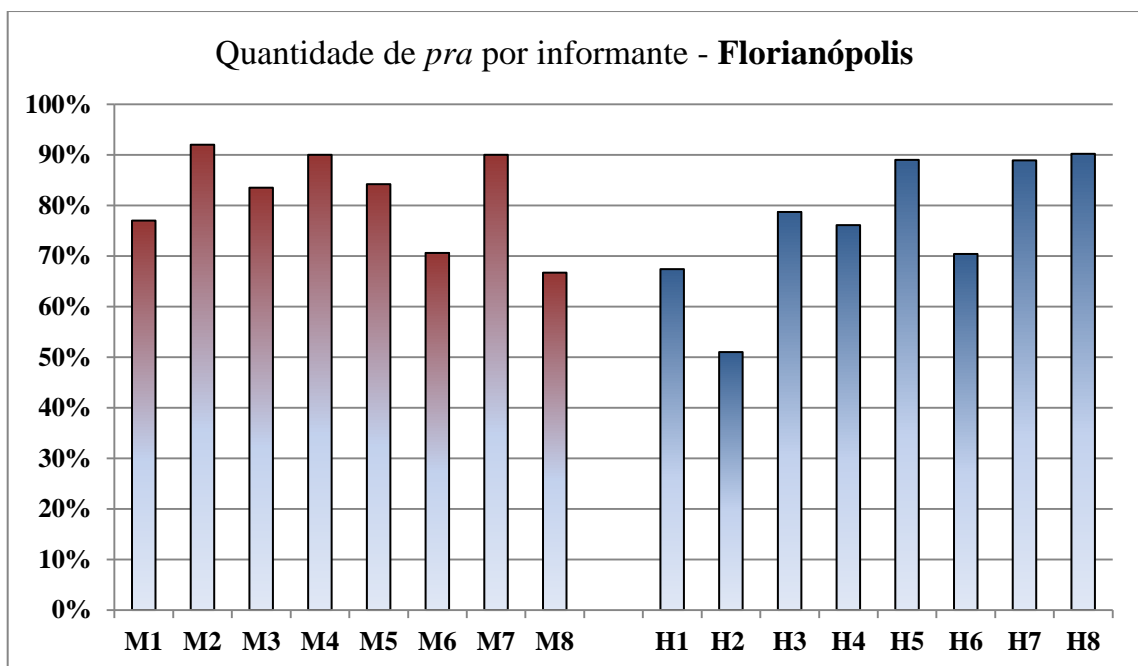


Gráfico 5 – Aplicações de *pra* por informante na cidade de Florianópolis. Fonte: a autora.

No Gráfico 5, pode-se perceber que a situação se repete na cidade de Florianópolis em termos de emprego de *pra* acima de 70% da maioria dos informantes. Com exceção de uma pessoa (H2, com 51% de aplicação), todos os informantes representam um percentual que varia de 70% a 90% de aplicação do fenômeno estudado, não havendo, portanto, diferença

significativa para a escolha da preposição reduzida *pra* entre os entrevistados da capital catarinense. No entanto, comparando-se Florianópolis com Curitiba, a análise do indivíduo sinaliza uma variação maior de frequência de uso em Florianópolis.

## 6.5 CRUZAMENTOS ENTRE FATORES EXTRALINGUÍSTICOS OU SOCIAIS

O cruzamento entre fatores tem o objetivo de investigar interferências entre eles e descobrir possíveis células vazias, geradas por combinações de fatores que não apresentaram nenhuma ocorrência da variável investigada. Conforme Brescancini (2002, p. 51), é comum a ocorrência de células vazias quando se analisa o comportamento de cada informante de uma amostra. Quando se insere o fator *Informante* no programa estatístico, verifica-se que é inevitável o surgimento de célula vazia em “Sexo Masculino”, por exemplo, se a entrevistada for do sexo feminino. Por essa razão, a solução para o problema é retirar esse grupo de fatores dos dados logo depois de analisá-lo, o que foi feito em nossa análise para a apresentação dos resultados nas seções 6.3 e 6.4 deste estudo.

O problema das células vazias também pode ser consequência de uma distribuição desequilibrada dos dados, o que pode ocasionar que a variante controlada não ocorra nenhuma vez antes de fatores mais raros em determinados fenômenos, como um numeral ou uma palavra com três sílabas pretônicas átonas, por exemplo. Se todas as células resultantes do cruzamento apresentam dados, com os grupos de fatores co-ocorrendo livremente e sem serem subcategorias uns dos outros, diz-se que eles são *ortogonais*. Dessa forma, tem-se um âmbito maior de comparações, mesmo que às vezes não seja possível obter dados em todas as células. (GUY e ZILLES, 2007, p. 52).

Em nossa análise, foram realizados cruzamentos entre os três grupos de fatores sociais que se mostraram influentes: *Idade*, *Escolaridade* e *Cidade*. A seguir, expomos tabelas com os resultados desses cruzamentos, os quais nos permitiram confirmar a relevância, apontada pelo programa, de cada um dos grupos.

### 6.5.1 Idade X Escolaridade

A Tabela 17 a seguir apresenta os dados referentes ao cruzamento entre as variáveis *Idade* e *Escolaridade*:

	+ de 50 anos Aplicação/Total	- de 50 anos Aplicação/Total	TOTAL
<i>Até 11 anos de estudo</i>	<b>603/669 = 90%</b>	529/681 = 78%	1132/1350 = 84%
<i>Até 4 anos de estudo</i>	317/413 = 77%	464/586 = 79%	781/999 = 78%
<b>TOTAL</b>	920/1082 = 85%	993/1267 = 78,4%	1913/2349 = 81%

Tabela 17 – Cruzamento entre as variáveis *Idade* e *Escolaridade*. Fonte: a autora.

O cruzamento exposto na Tabela 17 confirma o condicionamento das duas variáveis, conforme anunciamos nas seções 6.4.1 e 6.4.2: analisadas separadamente, a idade e a escolaridade dos indivíduos de nossa amostra pareciam exercer influência moderada no uso da preposição reduzida *pra*. A partir deste cruzamento, no entanto, verifica-se que são os falantes mais velhos e mais escolarizados que demonstram preferência por esta variante, com um percentual de 90% de aplicações.

Os demais fatores trazem resultados muito próximos, consideração que nos leva a concluir que os informantes mais velhos e menos escolarizados, além dos mais jovens em geral, não favorecem o uso da variante *pra*. Dessa forma, nossa hipótese de que o uso de *pra* estaria motivado por um grau de escolaridade mais baixo não se confirma em relação aos informantes mais velhos. A hipótese da escolaridade, portanto, não está totalmente descartada, já que os jovens com até quatro anos de estudo lideram o percentual de aplicação (79%).



### 6.5.2 Idade X Cidade

A Tabela 18 apresenta os dados referentes ao cruzamento entre as variáveis *Idade* e *Cidade*:

	+ de 50 anos Aplicação/Total	- de 50 anos Aplicação/Total	TOTAL
<i>Curitiba</i>	<b>511/567 = 90%</b>	454/576 = 79%	965/1143 = 84%
<i>Florianópolis</i>	409/515 = 79%	539/691 = 78%	948/1206 = 79%
<i>TOTAL</i>	920/1082 = 85%	993/1267 = 78,4%	1913/2349 = 81%

Tabela 18 – Cruzamento entre as variáveis *Idade* e *Cidade*. Fonte: a autora.

O cruzamento exposto na Tabela 18 revela que, além de os informantes mais velhos demonstrarem preferência pelo uso de *pra*, isso se acentua na cidade de Curitiba, onde 90% dos entrevistados desta faixa etária utiliza a variante em questão. Já os florianopolitanos mais velhos apresentam comportamento praticamente idêntico ao dos informantes mais jovens em geral, evidenciando a falta de influência dessas características na aplicação de *pra*. Contudo, a análise feita por indivíduo, conforme o Gráfico 5, indica uma certa instabilidade entre os informantes, sendo que alguns apresentam 90% de aplicação *versus* outros que usam o *pra* em torno de 70%.

### 6.5.3 Escolaridade X Cidade

A Tabela 19 a seguir expõe os dados referentes ao cruzamento entre as variáveis *Escolaridade* e *Cidade*:

	<i>Até 11 anos de estudo</i> Aplicação/Total	<i>Até 4 anos de estudo</i> Aplicação/Total	<i>TOTAL</i>
<i>Curitiba</i>	<b>670/771 = 87%</b>	295/372 = 79%	965/1143 = 84%
<i>Florianópolis</i>	462/579 = 80%	486/627 = 78%	948/1206 = 79%
<i>TOTAL</i>	1132/1350 = 84%	781/999 = 78%	1913/2349 = 81%

Tabela 19 – Cruzamento entre as variáveis *Escolaridade* e *Cidade*. Fonte: a autora.

O cruzamento exposto na Tabela 19 esclarece que os informantes que têm mais anos de estudo, na cidade de Curitiba, demonstram preferência pelo uso da variante *pra*, com 87% de aplicações. Quem vive em Florianópolis, no entanto, apresenta comportamento quase idêntico aos falantes com menos escolaridade, de modo geral. Em síntese, a escolaridade parece exercer papel muito moderado, em favor dos menos escolarizados em Florianópolis. Em Curitiba, no entanto, os informantes com mais escolaridade utilizam mais o *pra*.

A análise das Variáveis Extralinguísticas buscou confirmar prováveis condicionamentos indicados nos resultados de cada variável a partir dos cruzamentos entre fatores sociais, a fim de gerar interações entre cada um. Dessa maneira, foi possível concluir que o papel da *escolaridade* deve ser relacionado às variáveis *Idade* e *Cidade* para que possa ser considerado influente na escolha de *pra*: foram os informantes mais velhos e com mais estudo que optaram por essa variante na cidade de Curitiba.

Em seguida, discutiremos, de forma sucinta, os resultados obtidos em relação às observações sobre a preposição *para* na Teoria Fonológica, apresentadas no capítulo 3.

## 6.6 BREVE DISCUSSÃO DOS RESULTADOS

Segundo os resultados de nossa pesquisa, com base em dados de língua falada das cidades de Curitiba e de Florianópolis pertencentes ao Banco VARSUL, a preposição *para* apresenta três variantes, *para*, *pra* e *pa*, sendo a mais frequente a forma *pra*. Essa forma é utilizada pelos informantes mais velhos e com mais escolaridade. Em Curitiba, há mais uso do que em Florianópolis.

Em relação às variáveis linguísticas, a análise dos dados mostrou que o *pra* foi favorecido quando seguido de palavra gramatical (*pra ela*, *pra nós*, *pra isso*, *pra ti*, etc.), contexto fonológico seguinte *vogal coronal tônica* e pelo fato de ser antecedido por outro *pra*.

A forma *pra* é considerada na literatura um clítico (SIMIONI, 2008; TONELI, 2009; GUZZO, 2015) ou um grupo clítico (BISOL, 2000). A diferença entre essas duas categorias parece estar relacionada ao fato de as formas atuarem como unidades prosódicas independentes ou não.

No caso da forma independente, há um vocábulo fonológico não suscetível à regra de neutralização de vogais átonas. Conforme exposto no capítulo 3 desta pesquisa, Bisol (2010) afirma que o clítico e sua unidade adjacente formam uma só palavra fonológica. Já no caso da unidade prosódica dependente, segundo a autora, o processo de neutralização ocorre. O clítico une-se à palavra de conteúdo para formar um grupo clítico, comportando-se de acordo com as regras da palavra fonológica.

A forma *pra* não é muito transparente para o processo de neutralização de vogais átonas, pois formas alofônicas de /a/ não são muito comuns no português brasileiro, como ocorre em [mas] *versus* [mãs]; [bã'nana] *versus* [ba'nana]. Por isso, fica difícil falar de neutralização em relação à preposição *pra*, além do fato de existir um resíduo fonético acústico da consoante /r/, segundo a Ilustração 3 a seguir, retirada de Toneli (2009, p. 73):

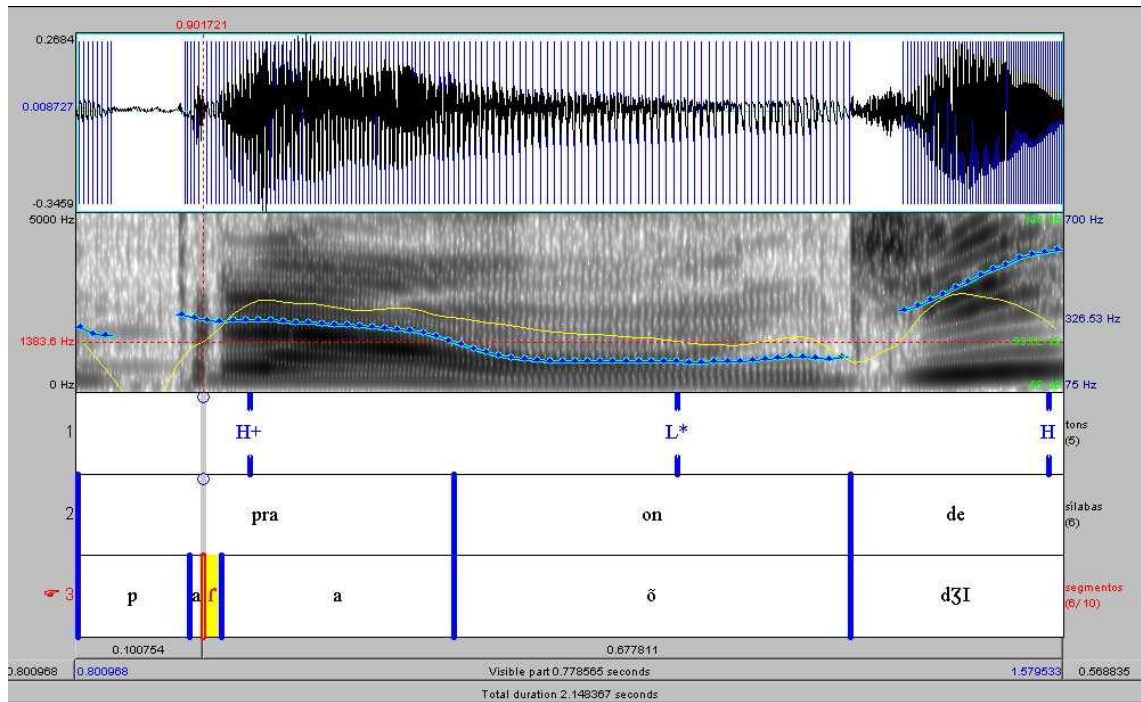


Ilustração 3 – Análise acústica da preposição *para*. Fonte: Toneli (2009).

A análise acústica realizada por Toneli (2009) e exposta na Ilustração 3 mostra que, na redução da preposição *para*, a vogal *-a* da sílaba /pa/ é quase apagada. Contudo, ainda é possível encontrar vestígios de sua presença.

O fato de *pra*, em nossos dados, ser usado preferencialmente em contextos com palavras funcionais iniciadas por vogais coronais tônicas é um impeditivo para a aplicação do processo de sândi vocálico, comum em contextos de contato de vogais, como no caso de *menina elegante* [meninelegante]ω, por exemplo.

Assim, a expressão *pra ela* não é comumente realizada como \* [prɛla] ou \* [pɛla], como um só vocábulo fonológico, mas como [[pra]ω [ɛla]ω ]C, em que o *pra* não sofre o processo de sândi vocálico, formando-se, pois, duas palavras fonológicas pertencentes a um domínio maior, denominado grupo clítico (C).

A restrição rítmica e silábica para o uso da forma *pra* e não *para/pa* ocorre pela natureza da palavra, se gramatical ou lexical, que segue a preposição. No caso de palavras lexicais, é mais comum haver a realização de formas não reduzidas, como o *para*, em frases como, por exemplo, *para o ônibus, para isso*. A variante *pa* é uma forma que parece não possibilitar ressilabação e sândi, como em *pra isso* – [pa iso] e não \* [´piso].

Também cabe frisar o condicionamento pragmático da forma *pra*, que parece ser mais usada quando segue outra forma semelhante no discurso. Assim, quem usa uma vez *pra*, usa outra vez na mesma frase, como em “(...) pra eu usar isso pra viajar”.

Por fim, acreditamos que a forma *pra* seja a variante mais frequente na fala, fato esse embasado por questões linguísticas prosódicas, independentemente da escolaridade, do sexo, da idade e da cidade do indivíduo, apesar de os nossos resultados mostrarem leve diferença entre essas categorias.

O próximo capítulo apresentará todas as conclusões desta pesquisa de forma mais detalhada.

## 7 CONCLUSÃO

Este trabalho investigou a variação da preposição *para* na fala de Curitiba (PR) e Florianópolis (SC), cidades em que ela pode ser realizada também como *pra* e *pa*, assim como outras cidades analisadas por outras pesquisas. A amostra pertence ao *corpus* do Projeto VARSUL e a perspectiva adotada foi a Teoria da Variação e Mudança Linguística, de Weinreich, Labov e Herzog ([1975] 2006).

A seguir, retomamos as hipóteses definidas neste trabalho sobre a variação da preposição *para*: 1) A forma reduzida *pra* seria mais utilizada do que as variantes *para* e *pa* na fala das cidades de Curitiba (PR) e Florianópolis (SC); 2) A escolha de *pra*, variante típica da língua falada coloquial, estaria motivada por um grau de escolaridade mais baixo do indivíduo; 3) A variante *para*, com acento na penúltima sílaba, altera seu estatuto prosódico quando reduzida para *pra* ou *pa*, tornando-se uma sílaba sem acento e uma forma dependente. Como tal, a preposição reduzida sofreria um processo de juntura (chamado *sândi*) com a palavra seguinte que iniciasse com vogal átona.

Os resultados confirmam nossa primeira hipótese, de que a forma *pra* seria mais produtiva em relação a *para* e *pa*, com aplicação de 81%. Uma vez que as formas reduzidas são mais utilizadas também em outras regiões do País, segundo algumas pesquisas, esta regra variável caracteriza uma mudança em curso no português brasileiro, devido a seu crescente uso.

A segunda hipótese, de que o uso de *pra* estaria motivado por um grau mais baixo de escolaridade, não pode ser confirmada nem totalmente descartada. Quem parece favorecer um pouco a escolha por essa variante, ao contrário, são os informantes com mais estudo (*Até 11 anos de estudo*), destacando-se os mais velhos e naturais da cidade de Curitiba (PR). No entanto, é baixa a diferença entre os pesos relativos nas variáveis *Escolaridade*, *Idade* e *Cidade*, razão pela qual não é possível confirmar nem descartar completamente essa hipótese. Portanto, essas variáveis parecem não exercer papel relevante na análise.

A terceira hipótese não pode ser verificada. Nossos dados sobre a variante *pra* mostram que ela foi utilizada, na maior parte das vezes, junto a palavras funcionais iniciadas por vogais coronais tônicas. De acordo com Bisol (2002, p. 231), existem restrições rítmicas e segmentais que podem determinar ou não a ocorrência de processos de *sândi*. Segundo a autora, no caso da elisão, a vogal inicial da palavra seguinte deve ser posterior ou frontal; já no caso da degeminação, as duas vogais que entram em contato devem ser semelhantes, desde

que a segunda seja átona. Portanto, o fato de a vogal inicial da palavra seguinte à preposição reduzida ser coronal e tônica mostra-se um impeditivo para a ocorrência de processo de junção (sândi). Por isso, não é possível confirmar nem descartar essa hipótese.

Algumas variáveis linguísticas mostraram-se relevantes para o uso da preposição reduzida *pra*, como os contextos seguintes com palavras gramaticais, vogais coronais e consoantes dorsais, além dos casos em que havia uma forma *pra* antecedida por um *pra* na mesma frase. Em relação às variáveis sociais, os informantes mais velhos, mais escolarizados e da cidade de Curitiba utilizaram um pouco mais a forma *pra*.

Esperamos que os resultados obtidos neste estudo possam contribuir para outras pesquisas na área, a fim de ampliar o conhecimento sobre a fala no português brasileiro. Uma próxima etapa em pesquisas futuras poderia envolver um maior aprofundamento nos estudos da Teoria Prosódica e o controle de aspectos de velocidade da fala analisada.

## REFERÊNCIAS

- ALMEIDA, Napoleão Mendes de. *Gramática Metódica da Língua Portuguesa: curso único e completo*. São Paulo: Saraiva, 11ª edição, 1960.
- AULETE, Caldas. *Dicionário de Português Aulete Digital*. Disponível em [www.auletedigital.com.br](http://www.auletedigital.com.br).
- BISOL, Leda. *Introdução a Estudos de Fonologia do Português Brasileiro*. Porto Alegre: Edipucrs, 2010.
- \_\_\_\_\_. 2004. Mattoso Câmara Jr. e a Palavra Prosódica. *D.E.L.T.A.*, v. 20 (especial), p. 59-70.
- \_\_\_\_\_. O clítico e seu Status Prosódico. *Revista de Estudos Linguísticos*, v. 9, n.1. Belo Horizonte, 2000, p. 5-30.
- BRESCANCINI, Cláudia Regina. A análise de regra variável e o programa VARBRUL 2S. In: BISOL, L.; BRESCANCINI, C. (Org.). *Fonologia e Variação: Recortes do Português Brasileiro*. Porto Alegre: EDIPUCRS, 2002.
- CÂMARA JR., Joaquim Mattoso. *História e Estrutura da Língua Portuguesa*. Rio de Janeiro: Padrão, 1979.
- \_\_\_\_\_. *Estrutura da língua portuguesa*. Petrópolis: Vozes, 1970.
- CRISTÓFARO-SILVA, Thaís. *Dicionário de fonética e fonologia*. Colaboradoras Daniela Oliveira Guimarães e Maria Mendes Cantoni. São Paulo: Contexto, 2011, 239p.
- CUNHA, Celso; CINTRA, Lindley. *Nova Gramática do português contemporâneo*. 6ª ed. Rio de Janeiro: Lexicon, 2014.
- FARACO, Carlos Alberto. *Linguística Histórica: uma introdução ao estudo da história das línguas*. São Paulo: Parábola Editorial, 2005.
- FERREIRA, Melissa Osterlund. A Variação da Preposição *para* na Fala de Londrina pelos Dados do VARSUL. 2014. 71 f. Monografia (Graduação). Universidade Federal do Rio Grande do Sul, Porto Alegre, 2014. Disponível em: <https://www.lume.ufrgs.br/bitstream/handle/10183/115665/000963805.pdf?sequence=1>  
Acesso em 23 abr. 2018.
- GAZOLA, Adriana. A Estrutura Prosódica da Preposição *para*. *Uniletras*, Ponta Grossa, v. 30, n. 2, p. 367-396, jul./dez. 2008. Disponível em [www.revistas2.uepg.br/index.php/uniletras/article/view/524/526](http://www.revistas2.uepg.br/index.php/uniletras/article/view/524/526) Acesso em 23 abr. 2018.
- GUZZO, Natália Brambatti. A prosodização de clíticos e compostos em português brasileiro. 2015. 232 f. Tese (Doutorado em Letras) – Universidade Federal do Rio Grande do Sul, Programa de Pós-Graduação em Letras, Porto Alegre, BR-RS, 2015. Disponível em:



<http://www.lume.ufrgs.br/bitstream/handle/10183/131569/000973778.pdf?sequence=1>

Acesso em 23 abr. 2018.

GUY, Gregory; ZILLES, Ana. *Sociolinguística quantitativa: instrumental de análise*. São Paulo: Parábola, 2007.

INSTITUTO DE TEORIA LINGUÍSTICA E COMPUTACIONAL. Dicionário de Termos Linguísticos. In: *Portal da Língua Portuguesa*. Disponível em [www.portaldalinguaportuguesa.org/?action=terminology](http://www.portaldalinguaportuguesa.org/?action=terminology) Acesso em 23 abr. 2018.

KRISCHKE, Rev. George Upton. *Do reto uso de preposições em Língua Portuguesa*. Porto Alegre: Gundlach, 1939.

LABOV, William. *Padrões Sociolinguísticos*. São Paulo: Parábola, 2008.

LUCENA, Rubens Marques de. *Elementos para o Estudo da Variação Lingüística na Paraíba*. João Pessoa: Ed. do autor, c 2006. 137 p.

MAYA, Leonardo Zechlinski. A variação da preposição para na fala de Porto Alegre. 2004. 121 f. Dissertação (Mestrado em Letras) – Faculdade de Letras, PUCRS, Porto Alegre, 2004.

MOLLICA, Maria Cecília; BRAGA, Maria Luiza (Orgs.). *Introdução à Sociolinguística: o tratamento da variação*. São Paulo: Contexto, 2004.

POGGIO, Rosauta Maria Galvão Fagundes. *Processos de gramaticalização de preposições do latim ao português: uma abordagem funcionalista*. Salvador: EDUFBA, 2002. 303 p. : il.

PROJETO VARSUL. [www.varsul.org.br](http://www.varsul.org.br).

SILVA, Giselle Machline de Oliveira e. Coleta de dados. In: MOLLICA, M. C.; BRAGA, M. L. (Orgs.). *Introdução à Sociolinguística: o tratamento da variação*. São Paulo: Contexto, 2004.

SILVA, Nahete de Alcântara. A Preposição *para* e suas Variantes no Falar Araguatinense. 2010.74f. Dissertação (Mestrado em Letras) – Centro de Ciências Humanas, Letras e Artes, Universidade Federal da Paraíba, João Pessoa, 2010.

SIMIONI, Taíse. O clítico e seu lugar na estrutura prosódica em português brasileiro. *Alfa*, São Paulo, 52 (2): 431- 446, fev./jun. 2008. Disponível em <http://seer.fclar.unesp.br/alfa/article/view/1526/1233> Acesso em 23 abr. 2018.

TARALLO, Fernando Luiz. *A pesquisa sociolingüística*. São Paulo: Ática, 5ª edição – 1997.

TONELI, Priscila Marques. A palavra prosódica no português brasileiro: o estatuto prosódico das palavras funcionais. 2009. 164 p. Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Estudos da Linguagem, Campinas, SP. Disponível em: [http://repositorio.unicamp.br/jspui/bitstream/REPOSIP/270913/1/Toneli\\_PriscilaMarques\\_M.pdf](http://repositorio.unicamp.br/jspui/bitstream/REPOSIP/270913/1/Toneli_PriscilaMarques_M.pdf) Acesso em 23 abr. 2018.

WEINREICH, Uriel; LABOV, William; HERZOG, Marvin. *Fundamentos Empíricos para uma teoria da mudança lingüística*. Tradução de Marcos Bagno. Revisão técnica de Carlos Alberto Faraco. São Paulo, Parábola – 2006.

WILLIAMS, Edwin Bucher. *Do latim ao português: fonologia e morfologia históricas da língua portuguesa*. Traduzido por Antônio Houaiss. Rio de Janeiro: Instituto Nacional do Livro, 1961. 352 p.

**ANEXO A – Rodada eneária**

CELL CREATION • 05/10/2017 13:24:31 .....

NAME OF TOKEN FILE: DADOS CURITIBA E FLORIANÓPOLIS.TKN

NAME OF CONDITION FILE: UNTITLED.CND

(

; IDENTITY RECODE: ALL GROUPS INCLUDED AS IS.

(1)

(2)

(3)

(4)

(5)

(6)

(7)

(8)

(9)

(10)

(11)

(12)

(13)

)

NUMBER OF CELLS: 1815

APPLICATION VALUE(S): 120

TOTAL NO. OF FACTORS: 85

GROUP	1	2	0	TOTAL	%
-------	---	---	---	-------	---

1 (2)	1	2	0		
Q N	671	76	5	752	29.0
%	89.2	10.1	0.7		

N N	528	177	3	708	27.3
%	74.6	25.0	0.4		

V N 470 172 4 646 24.9  
 % 72.8 26.6 0.6

! N 138 11 6 155 6.0  
 % 89.0 7.1 3.9

? N 243 15 0 258 10.0  
 % 94.2 5.8 0.0 \* KNOCKOUT \*

W N 26 4 1 31 1.2  
 % 83.9 12.9 3.2

C N 3 0 3 6 0.2  
 % 50.0 0.0 50.0 \* KNOCKOUT \*

G N 7 1 27 35 1.4  
 % 20.0 2.9 77.1

TOTAL N 2086 456 49 2591  
 % 80.5 17.6 1.9

-----  
 2 (3) 1 2 0

U N 80 25 19 124 4.8  
 % 64.5 20.2 15.3

K N 309 55 4 368 14.2  
 % 84.0 14.9 1.1

D N 468 119 4 591 22.8  
 % 79.2 20.1 0.7

E N 313 34 0 347 13.4  
 % 90.2 9.8 0.0 \* KNOCKOUT \*

3 N 138 11 6 155 6.0  
 % 89.0 7.1 3.9

F N 226 53 0 279 10.8  
 % 81.0 19.0 0.0 \* KNOCKOUT \*

M N 378 132 5 515 19.9  
 % 73.4 25.6 1.0

H N 87 7 0 94 3.6  
 % 92.6 7.4 0.0 \* KNOCKOUT \*

A N 87 20 11 118 4.6  
 % 73.7 16.9 9.3

TOTAL N 2086 456 49 2591  
 % 80.5 17.6 1.9

-----  
 3(4) 1 2 0  
 T N 661 100 2 763 29.4  
 % 86.6 13.1 0.3

@ N 589 101 35 725 28.0  
 % 81.2 13.9 4.8

O N 149 66 2 217 8.4  
 % 68.7 30.4 0.9

X N 138 11 6 155 6.0  
 % 89.0 7.1 3.9

W N 499 166 3 668 25.8  
 % 74.7 24.9 0.4

Y N	50	12	1	63	2.4
%	79.4	19.0	1.6		

TOTAL N	2086	456	49	2591	
%	80.5	17.6	1.9		

-----

4 (5)	1	2	0		
I N	1300	291	34	1625	62.7
%	80.0	17.9	2.1		

R N	345	68	7	420	16.2
%	82.1	16.2	1.7		

J N	404	43	6	453	17.5
%	89.2	9.5	1.3		

P N	4	1	2	7	0.3
%	57.1	14.3	28.6		

L N	33	53	0	86	3.3
%	38.4	61.6	0.0		* KNOCKOUT *

TOTAL N	2086	456	49	2591	
%	80.5	17.6	1.9		

-----

5 (6)	1	2	0		
7 N	998	222	3	1223	47.2
%	81.6	18.2	0.2		

6 N	590	104	35	729	28.1
%	80.9	14.3	4.8		

8 N	271	92	4	367	14.2
-----	-----	----	---	-----	------

% 73.8 25.1 1.1

X N 138 11 6 155 6.0

% 89.0 7.1 3.9

9 N 89 27 1 117 4.5

% 76.1 23.1 0.9

TOTAL N 2086 456 49 2591

% 80.5 17.6 1.9

-----  
6(7) 1 2 0

S N 1608 394 29 2031 78.4

% 79.2 19.4 1.4

[ N 120 10 5 135 5.2

% 88.9 7.4 3.7

Z N 18 1 1 20 0.8

% 90.0 5.0 5.0

] N 340 51 14 405 15.6

% 84.0 12.6 3.5

TOTAL N 2086 456 49 2591

% 80.5 17.6 1.9

-----  
7(8) 1 2 0

F N 1588 269 49 1906 73.6

% 83.3 14.1 2.6

G N 13 6 0 19 0.7

% 68.4 31.6 0.0 \* KNOCKOUT \*

E N 6 19 0 25 1.0  
 % 24.0 76.0 0.0 \* KNOCKOUT \*

D N 160 89 0 249 9.6  
 % 64.3 35.7 0.0 \* KNOCKOUT \*

C N 250 63 0 313 12.1  
 % 79.9 20.1 0.0 \* KNOCKOUT \*

A N 25 2 0 27 1.0  
 % 92.6 7.4 0.0 \* KNOCKOUT \*

B N 38 8 0 46 1.8  
 % 82.6 17.4 0.0 \* KNOCKOUT \*

H N 6 0 0 6 0.2  
 % 100.0 0.0 0.0 \* KNOCKOUT \*

TOTAL N 2086 456 49 2591  
 % 80.5 17.6 1.9

-----  
 8 (9) 1 2 0  
 Y N 1036 224 27 1287 49.7  
 % 80.5 17.4 2.1

X N 1050 232 22 1304 50.3  
 % 80.5 17.8 1.7

TOTAL N 2086 456 49 2591  
 % 80.5 17.6 1.9

-----  
 9 (10) 1 2 0  
 + N 1018 150 44 1212 46.8  
 % 84.0 12.4 3.6



- N 1068 306 5 1379 53.2  
 % 77.4 22.2 0.4

TOTAL N 2086 456 49 2591  
 % 80.5 17.6 1.9

-----  
 10 (11) 1 2 0  
 5 N 1258 209 38 1505 58.1  
 % 83.6 13.9 2.5

4 N 828 247 11 1086 41.9  
 % 76.2 22.7 1.0

TOTAL N 2086 456 49 2591  
 % 80.5 17.6 1.9

-----  
 11 (12) 1 2 0  
 A N 117 4 1 122 4.7  
 % 95.9 3.3 0.8

O N 84 5 0 89 3.4  
 % 94.4 5.6 0.0 \* KNOCKOUT \*

R N 77 4 0 81 3.1  
 % 95.1 4.9 0.0 \* KNOCKOUT \*

M N 60 3 16 79 3.0  
 % 75.9 3.8 20.3

Z N 47 27 0 74 2.9  
 % 63.5 36.5 0.0 \* KNOCKOUT \*

C N 78 27 0 105 4.1

	%	74.3	25.7	0.0		* KNOCKOUT *
N N		78	35	0	113	4.4
	%	69.0	31.0	0.0		* KNOCKOUT *
D N		72	14	0	86	3.3
	%	83.7	16.3	0.0		* KNOCKOUT *
J N		69	17	0	86	3.3
	%	80.2	19.8	0.0		* KNOCKOUT *
T N		44	2	0	46	1.8
	%	95.7	4.3	0.0		* KNOCKOUT *
I N		43	19	0	62	2.4
	%	69.4	30.6	0.0		* KNOCKOUT *
X N		81	18	0	99	3.8
	%	81.8	18.2	0.0		* KNOCKOUT *
G N		107	1	1	109	4.2
	%	98.2	0.9	0.9		
L N		31	3	0	34	1.3
	%	91.2	8.8	0.0		* KNOCKOUT *
Y N		71	1	0	72	2.8
	%	98.6	1.4	0.0		* KNOCKOUT *
U N		26	5	6	37	1.4
	%	70.3	13.5	16.2		
Q N		30	9	0	39	1.5
	%	76.9	23.1	0.0		* KNOCKOUT *

P N 69 37 0 106 4.1  
 % 65.1 34.9 0.0 \* KNOCKOUT \*

E N 68 3 3 74 2.9  
 % 91.9 4.1 4.1

B N 73 14 2 89 3.4  
 % 82.0 15.7 2.2

F N 28 28 3 59 2.3  
 % 47.5 47.5 5.1

H N 62 16 0 78 3.0  
 % 79.5 20.5 0.0 \* KNOCKOUT \*

S N 36 4 0 40 1.5  
 % 90.0 10.0 0.0 \* KNOCKOUT \*

V N 57 9 9 75 2.9  
 % 76.0 12.0 12.0

W N 88 17 0 105 4.1  
 % 83.8 16.2 0.0 \* KNOCKOUT \*

K N 109 12 1 122 4.7  
 % 89.3 9.8 0.8

1 N 74 37 0 111 4.3  
 % 66.7 33.3 0.0 \* KNOCKOUT \*

2 N 41 17 0 58 2.2  
 % 70.7 29.3 0.0 \* KNOCKOUT \*

3 N 76 10 0 86 3.3  
 % 88.4 11.6 0.0 \* KNOCKOUT \*

4 N 99 53 1 153 5.9  
 % 64.7 34.6 0.7

5 N 42 1 5 48 1.9  
 % 87.5 2.1 10.4

6 N 49 4 1 54 2.1  
 % 90.7 7.4 1.9

TOTAL N 2086 456 49 2591  
 % 80.5 17.6 1.9

-----  
 12 (13) 1 2 0  
 C N 1085 185 24 1294 49.9  
 % 83.8 14.3 1.9

F N 1001 271 25 1297 50.1  
 % 77.2 20.9 1.9

TOTAL N 2086 456 49 2591  
 % 80.5 17.6 1.9

-----  
 TOTAL N 2086 456 49 2591  
 % 80.5 17.6 1.9

**ANEXO B – Rodada Binária sem Amalgamações**

• CELL CREATION • 19/01/2018 17:03:00 .....

NAME OF TOKEN FILE: DADOS CURITIBA E FLORIANÓPOLIS.TKN

NAME OF CONDITION FILE: UNTITLED.CND

(

; IDENTITY RECODE: ALL GROUPS INCLUDED AS IS.

(1)

(2)

(3)

(4)

(5)

(6)

(7)

(8)

(9)

(10)

(11)

(12)

(13)

)

NUMBER OF CELLS: 1815

APPLICATION VALUE(S): 1

TOTAL NO. OF FACTORS: 85

NON-

GROUP	APPS	APPS	TOTAL	%
-------	------	------	-------	---

-----

1 (2)

Q N	671	81	752	29.0
-----	-----	----	-----	------

%	89.2	10.8		
---	------	------	--	--

N N 528 180 708 27.3  
 % 74.6 25.4

V N 470 176 646 24.9  
 % 72.8 27.2

! N 138 17 155 6.0  
 % 89.0 11.0

? N 243 15 258 10.0  
 % 94.2 5.8

W N 26 5 31 1.2  
 % 83.9 16.1

C N 3 3 6 0.2  
 % 50.0 50.0

G N 7 28 35 1.4  
 % 20.0 80.0

TOTAL N 2086 505 2591  
 % 80.5 19.5

-----  
 2 (3)

U N 80 44 124 4.8  
 % 64.5 35.5

K N 309 59 368 14.2  
 % 84.0 16.0

D N 468 123 591 22.8  
 % 79.2 20.8

E N 313 34 347 13.4  
% 90.2 9.8

3 N 138 17 155 6.0  
% 89.0 11.0

F N 226 53 279 10.8  
% 81.0 19.0

M N 378 137 515 19.9  
% 73.4 26.6

H N 87 7 94 3.6  
% 92.6 7.4

A N 87 31 118 4.6  
% 73.7 26.3

TOTAL N 2086 505 2591  
% 80.5 19.5

-----  
3 (4)

T N 661 102 763 29.4  
% 86.6 13.4

@ N 589 136 725 28.0  
% 81.2 18.8

O N 149 68 217 8.4  
% 68.7 31.3

X N 138 17 155 6.0  
% 89.0 11.0

W N 499 169 668 25.8  
 % 74.7 25.3

Y N 50 13 63 2.4  
 % 79.4 20.6

TOTAL N 2086 505 2591  
 % 80.5 19.5

-----  
 4 (5)

I N 1300 325 1625 62.7  
 % 80.0 20.0

R N 345 75 420 16.2  
 % 82.1 17.9

J N 404 49 453 17.5  
 % 89.2 10.8

P N 4 3 7 0.3  
 % 57.1 42.9

L N 33 53 86 3.3  
 % 38.4 61.6

TOTAL N 2086 505 2591  
 % 80.5 19.5

-----  
 5 (6)

7 N 998 225 1223 47.2  
 % 81.6 18.4

6 N 590 139 729 28.1  
 % 80.9 19.1



8 N 271 96 367 14.2

% 73.8 26.2

X N 138 17 155 6.0

% 89.0 11.0

9 N 89 28 117 4.5

% 76.1 23.9

TOTAL N 2086 505 2591

% 80.5 19.5

-----  
6 (7)

S N 1608 423 2031 78.4

% 79.2 20.8

[ N 120 15 135 5.2

% 88.9 11.1

Z N 18 2 20 0.8

% 90.0 10.0

] N 340 65 405 15.6

% 84.0 16.0

TOTAL N 2086 505 2591

% 80.5 19.5

-----  
7 (8)

F N 1588 318 1906 73.6

% 83.3 16.7

G N 13 6 19 0.7

% 68.4 31.6

E N 6 19 25 1.0

% 24.0 76.0

D N 160 89 249 9.6

% 64.3 35.7

C N 250 63 313 12.1

% 79.9 20.1

A N 25 2 27 1.0

% 92.6 7.4

B N 38 8 46 1.8

% 82.6 17.4

H N 6 0 6 0.2

% 100.0 0.0 \* KNOCKOUT \*

TOTAL N 2086 505 2591

% 80.5 19.5

-----  
8 (9)

Y N 1036 251 1287 49.7

% 80.5 19.5

X N 1050 254 1304 50.3

% 80.5 19.5

TOTAL N 2086 505 2591

% 80.5 19.5  
-----

9 (10)

+ N 1018 194 1212 46.8  
 % 84.0 16.0

- N 1068 311 1379 53.2  
 % 77.4 22.6

TOTAL N 2086 505 2591  
 % 80.5 19.5

-----

10 (11)

5 N 1258 247 1505 58.1  
 % 83.6 16.4

4 N 828 258 1086 41.9  
 % 76.2 23.8

TOTAL N 2086 505 2591  
 % 80.5 19.5

-----

11 (12)

A N 117 5 122 4.7  
 % 95.9 4.1

O N 84 5 89 3.4  
 % 94.4 5.6

R N 77 4 81 3.1  
 % 95.1 4.9

M N 60 19 79 3.0  
 % 75.9 24.1

Z N 47 27 74 2.9  
 % 63.5 36.5

C N 78 27 105 4.1  
 % 74.3 25.7

N N 78 35 113 4.4  
 % 69.0 31.0

D N 72 14 86 3.3  
 % 83.7 16.3

J N 69 17 86 3.3  
 % 80.2 19.8

T N 44 2 46 1.8  
 % 95.7 4.3

I N 43 19 62 2.4  
 % 69.4 30.6

X N 81 18 99 3.8  
 % 81.8 18.2

G N 107 2 109 4.2  
 % 98.2 1.8

L N 31 3 34 1.3  
 % 91.2 8.8

Y N 71 1 72 2.8  
 % 98.6 1.4

U N 26 11 37 1.4  
 % 70.3 29.7

Q	N	30	9	39	1.5
	%	76.9	23.1		
P	N	69	37	106	4.1
	%	65.1	34.9		
E	N	68	6	74	2.9
	%	91.9	8.1		
B	N	73	16	89	3.4
	%	82.0	18.0		
F	N	28	31	59	2.3
	%	47.5	52.5		
H	N	62	16	78	3.0
	%	79.5	20.5		
S	N	36	4	40	1.5
	%	90.0	10.0		
V	N	57	18	75	2.9
	%	76.0	24.0		
W	N	88	17	105	4.1
	%	83.8	16.2		
K	N	109	13	122	4.7
	%	89.3	10.7		
1	N	74	37	111	4.3
	%	66.7	33.3		
2	N	41	17	58	2.2

% 70.7 29.3

3 N 76 10 86 3.3

% 88.4 11.6

4 N 99 54 153 5.9

% 64.7 35.3

5 N 42 6 48 1.9

% 87.5 12.5

6 N 49 5 54 2.1

% 90.7 9.3

TOTAL N 2086 505 2591

% 80.5 19.5

-----

12 (13)

C N 1085 209 1294 49.9

% 83.8 16.2

F N 1001 296 1297 50.1

% 77.2 22.8

TOTAL N 2086 505 2591

% 80.5 19.5

-----

TOTAL N 2086 505 2591

% 80.5 19.5

## ANEXO C – Rodada Binária com Amalgamações

• CELL CREATION • 19/01/2018 18:20:01 .....

Name of token file: Dados Curitiba e Florianópolis.tkn

Name of condition file: Rodada COM os fatores sociais 19.01

(

; Identity recode: All groups included as is.

(1)

(2 (g (col 2 g))

(g (col 2 q))

(g (col 2 c))

(g (col 2 w))

(g (col 2 ?))

(v (col 2 v))

(v (col 2 n))

(nil (col 2 !)))

(3 (a (col 3 a))

(a (col 3 u))

(m (col 3 m))

(m (col 3 f))

(m (col 3 d))

(m (col 3 h))

(nil (col 3 3)))

;(4)

(5 (i (col 5 i))

(i (col 5 r))

(nil (col 5 p))

(nil (col 5 1)))  
 (6 (nil (col 6 X))  
 (8 (col 6 8))  
 (8 (col 6 7))  
 (8 (col 6 9)))  
 ;(7)  
 ;(8)  
 (9)  
 (10)  
 (11)  
 ;(12)  
 (13)  
 )

Number of cells: 319

Application value(s): 1

Total no. of factors: 18

Non-

Group	Apps	apps	Total	%
-------	------	------	-------	---

-----  
 1 (2)

g	N	935	123	1058	45.0
---	---	-----	-----	------	------

%	88.4	11.6			
---	------	------	--	--	--

v	N	978	313	1291	55.0
---	---	-----	-----	------	------

%	75.8	24.2			
---	------	------	--	--	--



Total N 1913 436 2349

% 81.4 18.6

---

2 (3)

a N 163 68 231 9.8

% 70.6 29.4

k N 305 56 361 15.4

% 84.5 15.5

m N 1136 280 1416 60.3

% 80.2 19.8

e N 309 32 341 14.5

% 90.6 9.4

Total N 1913 436 2349

% 81.4 18.6

---

3 (5)

i N 1539 388 1927 82.0

% 79.9 20.1

j N 374 48 422 18.0

% 88.6 11.4

Total N 1913 436 2349

% 81.4 18.6

-----  
4 (6)

8 N 1334 309 1643 69.9

% 81.2 18.8

6 N 579 127 706 30.1

% 82.0 18.0

Total N 1913 436 2349

% 81.4 18.6

-----  
5 (9)

y N 925 217 1142 48.6

% 81.0 19.0

x N 988 219 1207 51.4

% 81.9 18.1

Total N 1913 436 2349

% 81.4 18.6

-----  
6 (10)

+ N 920 162 1082 46.1

% 85.0 15.0

- N 993 274 1267 53.9

% 78.4 21.6

Total N 1913 436 2349

% 81.4 18.6

-----  
7 (11)

5 N 1132 218 1350 57.5

% 83.9 16.1

4 N 781 218 999 42.5

% 78.2 21.8

Total N 1913 436 2349

% 81.4 18.6

-----  
8 (13)

C N 965 178 1143 48.7

% 84.4 15.6

F N 948 258 1206 51.3

% 78.6 21.4

Total N 1913 436 2349

% 81.4 18.6

-----  
TOTAL N 1913 436 2349

% 81.4 18.6

## ANEXO D – Quadro de codificação de variáveis

CODIFICAÇÃO DE VARIÁVEIS LINGUÍSTICAS A VARIAÇÃO DA PREPOSIÇÃO PARA (PRA/PA) NA FALA DE FLORIANÓPOLIS E DE CURITIBA PELOS DADOS DO VARSUL							
Variáveis Linguísticas			Variáveis Extralinguísticas				
(1) VARIÁVEL DEPENDENTE			(9) SEXO				
0	Para: <i>para</i> modificar Curitiba,	x	Informante Feminino				
1	Pra (pros, pras): <i>Filme bom pras</i> crianças.	y	Informante Masculino				
2	Pa (pu, po): <i>Pega e leva um pacote de arroz lá pa</i> mulher!						
(2) CONTEXTO MORFOLÓGICO SEGUINTE			(10) IDADE				
n	Nome: <i>Por que que não dá prus pobre, que não tem?</i>	-	- de 50				
v	Verbo: <i>Sai de lá pra casar.</i>	+	+ de 50				
w	Numeral: <i>Eu acho que eu tinha dezesseis pa dezessete</i> anos.						
(2) CONTEXTO MORFOLÓGICO SEGUINTE			(11) ESCOLARIDADE				
q	Pronome: <i>Pra mim</i> é bom, né?	4	Até 4 anos				
g	Artigo definido: <i>Depois daqui mudei para a Rua</i> lapó.	5	Até 11 anos				
c	Conjunção: <i>para que</i> o vereador participe						
(2) CONTEXTO MORFOLÓGICO SEGUINTE			(12) DISTRIBUIÇÃO DA AMOSTRA				
?	Advérbio: <i>de uns vinte ano pra cá,</i>		C Curitiba				
!	Nada (pausa): <i>não tem outro lugar fora pra-</i>	Inf.	Idade	Fem	Masc	Esc.	
(3) CONTEXTO FONOLÓGICO SEGUINTE			A	- 50		CTB 01	4 anos
u	Vogal posterior ([u, o, ɔ]): <i>mandava pra outra; sacudindo ele p'um</i> lado,	O	+ 50			CTB 02	11 anos
a	Vogal Central ([a]): <i>Todo dia tem gente diferente pa</i> almoçá,	R	- 50			CTB 05	11 anos
e	Vogal Anterior ([i, e, ɛ]): <i>Mas tava muito puxado pra ela; Nós viemos pru</i> IBDF,	M	+ 50	CTB 06			11 anos
f	Consoante Labiodental ([f, v]): <i>Então era pra dá uma força prus</i> filhos.	Z	+ - 35			CTB 07	4 anos
k	Consoante velar ([k, g, x]): <i>pra cada série; Dezesseis anos pra</i> robá	C	39	CTB 08			4 anos
m	Consoante bilabial ([b, p, m]): <i>Transtorno pu</i> padre, né?; <i>pra</i> minha esposa,	N	- 50			CTB 09	11 anos
d	Consoante dental ([t, d, s, n, l, z]): <i>Lá ia eu pru</i> cinema; <i>daí</i> ele fugiu pra tráis,	D	- 50	CTB 10			4 anos
h	Consoante palatal ([ɲ, ʃ, ʒ, ʎ, dʒ, tʃ]): <i>Então ele fazia</i> banquinho pra gente,	J	- 50	CTB 19			11 anos
3	Nada (pausa): <i>tinha que vim a pé de lá pra,</i>	T	77	CTB 14			4 anos
(4) TONICIDADE DA SÍLABA SEGUINTE			I	+ 50		CTB 13	4 anos
t	Tônica: <i>E a gente ia pa</i> praia	@	Tônica monossílaba: <i>pra nós</i>				
o	Átona (2 sílabas): <i>pa</i> trabalhá	y	Átona (3 síl. ou +): <i>voltar pa</i> repartição				
w	Átona (1 sílaba): <i>para</i> servir o Exército,	L	+ 50	CTB 20			4 anos
X	Nada (pausa): <i>aquele tipo assim só pra-</i>	Y	+ 50	CTB 22			11 anos
		U	+ 50			CTB 21	11 anos
(5) PARALELISMO			F Florianópolis				
i	Ocorrência isolada (sem paralelismo): <i>ela fala pra</i> gente assim (...)	Q	- 50	FLP 01			4 anos
r	Primeira da série: <i>Já preparou pru</i> pai corrigir, <i>pru</i> pai ir fazer-	P	- 50			FLP 02	
j	Antecedida de <i>pra/pro</i> : <i>pra</i> dançar, <i>pra</i> í num barzinho curtir música,	E	- 50	FLP 11			11 anos
p	Antecedida de <i>para</i> : <i>de um dia para u</i> outro. <i>O</i> peixe, <i>de um dia para u</i> outro.	B	+ 50	FLP 07			4 anos
l	Antecedida de <i>pa/pu</i> : <i>pa</i> trazê até nos painéis <i>pa</i> fazê a carne na frigideira,	F	+ 50			FLP 06	4 anos
(6) NÚMERO DE SÍLABAS DO ITEM SEGUINTE			H	42		FLP 18	11 anos
6	1 sílaba: <i>eu fui para o</i> Rio de Janeiro.	S	+ 50	FLP 22			
7	2 sílabas: <i>Existem</i> pessoas que não têm dinheiro <i>pa</i> cumê,	V	+ 50			FLP 21	11 anos
8	3 sílabas: <i>Tinha</i> mais trabalho p' pessoal.	W	- 50	FLP 03			4 anos
9	4 ou mais sílabas: <i>e</i> mesmo <i>pra</i> informação,	K	- 50			FLP 04	4 anos
X	Nada (pausa): <i>Daí</i> tem a caldeira <i>pra-</i>	1	76	FLP 08			4 anos
(7) POSIÇÃO EM RELAÇÃO A PAUSAS			2	35		FLP 10	11 anos
s	Sem pausas: <i>que a família é que está aí pa</i> prová isso,	3	+ 50	FLP 15			11 anos
]	Depois de pausa: <i>e</i> como deveriam se alterar, <i>pra</i> que nós pudéssemos,	4	- 50	FLP 20			11 anos
[	Antes de pausa: <i>e</i> outro pessoal da área social <i>pra-</i> participá também.	5	+ 50			FLP 05	4 anos
z	Entre pausas: <i>do horto do Guabirotuba, né? Pra-</i> pras plantas,	6	+ 50			FLP 14	11 anos
(8) Processo de Sândi com a Sílabas Seguinte			(12) CIDADES				
A	Com Artigo /a/ no Plural (Pra + as = pras); (Pa+as = pas): <i>pras</i> festinhas						
B	Com Artigo /o/ no Plural (Pra + os = prus, pus): <i>vão pros</i> seus afazeres						
C	Com Artigo /a/ no Singular (Pra + a = pra, pa): <i>Se eu contá o lugar, eu vô pa</i> cadeia.						
D	Com Artigo /o/ no Singular (Pra + o = pru); (pa + o = pu): <i>cê ia pru</i> mato pegá gavirova,						
E	Pra + palavra c/ 2 síl. (pra + você = p'cê): <i>não dá p'cê</i> entendê	G	Com artigo /um/ (pra + um = prum)				
F	Nada (sem sândi): <i>quando a planta tá pra</i> florir,	H	Pra+palavra c/ 3 síl. (pra+aquilo = pr'aquilo)				