

CAMILA DE BONA

**O PAPEL DA FREQUÊNCIA LEXICAL EM FENÔMENOS FONOLÓGICOS
CONDICIONADOS MORFOLOGICAMENTE DO PORTUGUÊS BRASILEIRO**

PORTO ALEGRE

2018

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS
ÁREA: ESTUDOS DA LINGUAGEM
ESPECIALIDADE: FONOLOGIA E MORFOLOGIA

**O PAPEL DA FREQUÊNCIA LEXICAL EM FENÔMENOS FONOLÓGICOS
CONDICIONADOS MORFOLOGICAMENTE DO PORTUGUÊS BRASILEIRO**

CAMILA DE BONA
ORIENTADOR: PROF. DR. LUIZ CARLOS DA SILVA SCHWINDT

Tese de Doutorado em Fonologia e Morfologia,
apresentada como requisito parcial para a
obtenção do título de Doutora pelo Programa de
Pós-Graduação em Letras da Universidade
Federal do Rio Grande do Sul.

Porto Alegre
2018

CIP - Catalogação na Publicação

DE BONA, CAMILA
O PAPEL DA FREQUÊNCIA LEXICAL EM FENÔMENOS
FONOLÓGICOS CONDICIONADOS MORFOLOGICAMENTE DO
PORTUGUÊS BRASILEIRO / CAMILA DE BONA. -- 2018.
151 f.
Orientador: LUIZ CARLOS DA SILVA SCHWINDT.

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

1. Frequência Lexical. 2. Redução da Nasalidade.
3. Apagamento de /r/. I. SCHWINDT, LUIZ CARLOS DA
SILVA, orient. II. Título.

AGRADECIMENTOS

À minha família e aos meus amigos, pelo apoio.

À direção e aos colegas do Cursão Pré-Vestibular, do Colégio Nossa Senhora de Lourdes e do Colégio La Salle Carmo, por permitirem que eu me ausentasse por vezes para participar de cursos, orientações e eventos da área – por entenderem minhas ausências e me apoiarem durante os três primeiros anos do doutorado.

À direção do IFSul Câmpus Novo Hamburgo, em especial a Gelson Corrêa e Richard Martins, por aprovarem meu afastamento mesmo eu estando em estágio probatório. Aos colegas Adriana Braun, Carla Vianna, Daniel Meyer e Juneor Brehm, por, cada um à sua maneira, terem me ajudado sempre e muito.

Aos membros do Círculo Linguístico – Fonologia e Morfologia. Em especial, às queridas Camila Ülich, Luciana Morales e Raquel Chaves, pela nossa bonita sintonia.

À Athany Gutierrez e ao Pablo Ribeiro, por serem importantes parcerias.

À professora Sabrina Pereira de Abreu, por todos os ensinamentos desde o início de tudo.

À professora Gisela Collischonn, pelas inúmeras contribuições na qualificação do projeto desta tese, pelas aulas inquietantes, pela inspiração.

Às professoras Cláudia Brescancini e Valéria Monaretto, pelas contribuições preciosas na banca de qualificação, as quais foram determinantes para o desenvolvimento deste trabalho. Agradeço por terem aceito serem, mais uma vez, membros de minha banca. À professora Taís Bopp da Silva, agradeço por também ter aceito nosso convite. As contribuições certamente serão inúmeras.

À professora Mirjam Ernestus e ao seu grupo de pesquisa na Universidade Radboud Nijmegen e no Instituto Max Plack de Psicolinguística, pela acolhida, pelos aprendizados, pelos desafios.

À professora Leda Bisol e aos membros do grupo de estudos em Fonologia da PUCRS.

À Júlia Ricardo, por ter sido uma parceira fundamental no desenvolvimento desta tese. Por saber conciliar dedicação e leveza na condução da pesquisa.

Ao meu orientador, professor Luiz Carlos Schwindt, pela generosidade de ter tornado tão rica minha experiência com esta pesquisa. Por ter me dado a oportunidade de fazer parte do grupo de estudos em Fonologia na PUCRS; por ter me incentivado a fazer o doutorado sanduíche; por ter possibilitado minha parceria com a Júlia Ricardo; por ter dividido comigo seu rigor metodológico, seu conhecimento, suas dúvidas e tanto do seu tempo. Por ter me tornado parte de um grupo que sempre admirei. Obrigada pelo privilégio.

À CAPES e ao Programa de Pós-Graduação em Letras da Universidade Federal do Rio Grande do Sul.

RESUMO

Este trabalho teve por objetivo analisar o papel da frequência lexical em dois fenômenos fonológicos variáveis que apresentam condicionamentos morfológicos, quais sejam Redução da Nasalidade e Apagamento de /r/ Final, no intuito de verificar: 1) a interação de frequência lexical com variáveis linguísticas; 2) a interação de frequência lexical com variáveis sociais; 3) o melhor modelo, se abstracionista ou exemplarista, para tratar dos resultados aqui obtidos. Para a redução da nasalidade, reanalisamos dados de Schwindt e Bopp da Silva (2010), adicionando a informação de frequência lexical para a análise dos dados; para o fenômeno de apagamento de /r/ final, novos dados foram coletados de entrevistas com informantes do Rio de Janeiro. A informação de frequência lexical dos dados foi obtida do *corpus* de referência ASPA (Avaliação Sonora do Português Atual). Nossos resultados apontam que frequência lexical apresenta correlação positiva com a aplicação dos fenômenos, estando condicionada à classe gramatical, não à estrutura morfológica interna da palavra. Idade apresenta uma interação bastante significativa em dados de redução da nasalidade, principalmente na classe de não verbos sem *gem*. A mesma interação com idade não foi encontrada no fenômeno de apagamento de /r/. Uma explicação para isso talvez esteja relacionada com a caracterização de frequência de *type* e frequência de *token*. Tendo em vista que, na análise de frequência lexical, regras relativas à morfologia interna à palavra e à fonologia parecem não ser determinantes na aplicação dos fenômenos, defendemos neste trabalho a superioridade das abordagens exemplaristas.

Palavras-chave: Frequência Lexical; Redução da Nasalidade; Apagamento de /r/.

ABSTRACT

The purpose of this study was to analyze the role of lexical frequency in two variable phonological phenomena which present morphological conditioning, such as Nasality Reduction and Final /r/ Deletion, in order to verify: 1) lexical frequency interaction with linguistic variables ; 2) lexical frequency interaction with social variables; 3) the best model, if abstractionist or exemplarist, to deal with the results obtained here. For nasality reduction, we reanalyzed data from Schwindt and Bopp da Silva (2010), adding lexical frequency information for data analysis; for the final /r/ deletion phenomenon, new data were collected from interviews with informants from Rio de Janeiro. Lexical frequency information of the data was obtained from the reference corpus ASPA (Sound Evaluation of Current Portuguese). Our results point out that lexical frequency has positive correlation with the application of the phenomena, being conditioned by grammatical class, not by the internal morphological structure of the word. Age has a very significant interaction in nasality reduction data, especially in the class of non-verbs without *gem*. The same interaction with age was not found in the final /r/ deletion phenomenon. An explanation for this may be related to the characterization of type and token frequency. In lexical frequency analysis, considering that rules regarding the internal morphology of words and regarding phonology do not seem to be decisive in the application of both phenomena, we argue for the superiority of the exemplarist approaches.

Keywords: Lexical Frequency; Nasal Reduction; Final / r / deletion.

LISTA DE GRÁFICOS

Gráfico 1: Coeficiente de correlação em não verbos	76
Gráfico 2: Coeficiente de correlação em verbos	76
Gráfico 3: Coeficiente de Correlação em não verbos – 25 a 50 anos	77
Gráfico 4: Coeficiente de correlação em não verbos – 51 anos ou mais	78
Gráfico 5: Apagamento de /r/ final.....	79
Gráfico 6: Apagamento de /r/ final em verbos	83
Gráfico 7: Apagamento de /r/ final em não verbos	88

LISTA DE FIGURA

Figura 1: Regiões estudadas e processo de apagamento do /r/ (SERRA e CALLOU, 2013, p. 587)	46
--	----

LISTA DE TABELAS

Tabela 1: Classe de palavra e RN no sul do Brasil (adaptado de SCHWINDT e BOPP DA SILVA, 2010)	57
Tabela 2: Frequência Lexical e RN – verbos e não verbos	59
Tabela 3: Frequência Lexical e RN – não verbos	59
Tabela 4: Frequência Lexical e RN – verbos	59
Tabela 5: Frequência lexical e RN – não verbos	60
Tabela 6: Frequência lexical e RN – contexto precedente em não verbos (não palatal)	61
Tabela 7: Frequência lexical e RN – palatal em contexto precedente em não verbos.....	61
Tabela 8: Frequência lexical e RN – tempos verbais	62
Tabela 9: Frequência lexical e RN – contexto precedente em verbos	62

LISTA DE QUADROS

Quadro 1: Análise comparativa (SILVEIRA, 2010, p. 82)	45
Quadro 2: Regras de caráter regional (SERRA e CALLOU, 2013, p. 592)	47
Quadro 3: RN e variáveis linguísticas (SCHWINDT e BOPP DA SILVA, 2010, p.19).....	51
Quadro 4: RN e variáveis sociais (SCHWINDT e BOPP DA SILVA, 2010, p. 18)	52
Quadro 5: Informantes e idade – Amostra 1.....	55
Quadro 6: Informantes e idade – Amostra 2.....	55
Quadro 7: Palavras comuns sujeitas a RN no ASPAVARSUL entre as 20 mais frequentes.....	58
Quadro 8: Classe de palavra em RN.....	64
Quadro 9: Frequência contínua em RN.....	64
Quadro 10: Idade em RN.....	65
Quadro 11: Escolaridade em RN.....	65
Quadro 12: Contexto Precedente 1 em RN.....	65
Quadro 13: Contexto Precedente 2 em RN.....	66
Quadro 14 – Resultados para verbos e não verbos em RN.....	67
Quadro 15: Resultados para informantes de 25 a 50 anos.....	69
Quadro 16: Resultados para informantes de 51 anos ou mais.....	70
Quadro 17: Variável classe de palavra.....	80
Quadro 18: Variável contexto precedente.....	80
Quadro 19: Variável contexto seguinte.....	81
Quadro 20: Variável tamanho do vocábulo.....	81
Quadro 21: Variável gênero.....	81
Quadro 22 – Resultados para verbos e não verbos.....	82
Quadro 23 – Resultados para verbos.....	84
Quadro 24 – Resultados para verbos II.....	86
Quadro 25: Percentual de apagamento para as vogais /a/, /e/, /i/, /o/ em contexto precedente.....	87
Quadro 26 – Resultados para não verbos.....	88

Quadro 27: Percentual de apagamento para as vogais /a/, /e/, /ɛ/, /o/ e /ɔ/ em contexto precedente de não verbos	90
Quadro 28: não verbos em /-er/.....	91
Quadro 29: não verbos em /-ɛr/.....	92
Quadro 30: não verbos em /-ɔr/.....	92
Quadro 31: não verbos em /-ar/.....	93
Quadro 32: não verbos em /-or/.....	93
Quadro 33 – Redução da nasalidade em não verbos mais frequentes ASPA/VARSUL (Adaptado de DE BONA e SCHWINDT, 2017)	98
Quadro 34 – Redução da nasalidade em verbos mais frequentes ASPA/VARSUL..	99
Quadro 35 – Apagamento de /r/ em não verbos mais frequentes ASPA/Amostra /r/ 2018	100
Quadro 36 – Apagamento de /r/ em verbos mais frequentes ASPA/Amostra /r/ 2018.....	101

SUMÁRIO

INTRODUÇÃO: Variação Fonológica, Frequência Lexical e Morfologia.....	14
CAPÍTULO 1: SOBRE EFEITOS DE FREQUÊNCIA.....	20
1.1 <i>Efeitos de Frequência: dos modelos Neogramático e Difusionista à Fonologia de Uso</i>	<i>20</i>
1.2 <i>Redução acústica: modelos abstracionistas versus modelos exemplaristas.....</i>	<i>24</i>
1.3 <i>Implicações para a Morfologia.....</i>	<i>29</i>
CAPÍTULO 2: ESTUDOS SOBRE VARIAÇÃO NO PB.....	32
2.1 <i>Fenômenos fonológicos variáveis no PB e frequência lexical.....</i>	<i>32</i>
2.2 <i>Estudos sobre Redução da Nasalidade em PB.....</i>	<i>36</i>
2.3 <i>Estudos sobre Apagamento de /r/ final em PB.....</i>	<i>41</i>
CAPÍTULO 3: METODOLOGIA DE TRABALHO.....	50
3.1 <i>Redução da Nasalidade.....</i>	<i>50</i>
3.1.1 <i>Sobre o corpus.....</i>	<i>50</i>
3.1.2 <i>Perguntas norteadoras e procedimentos para a análise.....</i>	<i>43</i>
3.2 <i>Apagamento de /r/ final.....</i>	<i>53</i>
3.2.1 <i>Sobre o corpus.....</i>	<i>54</i>
3.2.2 <i>Perguntas norteadoras e procedimentos para a análise.....</i>	<i>56</i>
CAPÍTULO 4: RESULTADOS E DISCUSSÃO.....	57
4.1 <i>Redução da Nasalidade.....</i>	<i>57</i>
4.1.1 <i>Resultados do GoldVarb X.....</i>	<i>57</i>
4.1.2 <i>Resultados da Plataforma R.....</i>	<i>64</i>
4.2 <i>Apagamento de /r/ final.....</i>	<i>75</i>
4.2.1 <i>Análise preliminar – Resultados coeficiente de correlação.....</i>	<i>75</i>
4.2.2 <i>Resultados do Programa R.....</i>	<i>78</i>
4.3 <i>Sobre Léxico – Comparação entre os corpora.....</i>	<i>97</i>
CONSIDERAÇÕES FINAIS.....	103
REFERÊNCIAS.....	110
Anexo I: <i>Detalhamento da Amostra 2 – Apagamento de /r/.....</i>	<i>115</i>
Anexo II: <i>Lista de palavras analisadas na tese.....</i>	<i>118</i>

INTRODUÇÃO: Variação Fonológica, Frequência Lexical e Morfologia

Por meio de diversos estudos já publicados na área linguística, temos comprovada a importância que a frequência de determinadas unidades léxicas e padrões linguísticos exerce no que tange à linguagem de um modo geral. Clements (2009) aponta que a frequência é há muito tomada como reveladora do que é ou não tendencialmente universal e do que é ou não marcado nas línguas do mundo ou numa língua particular. Processos como redução, assimilação e regularização de formas irregulares podem ser potencializados através da elevada frequência de certas palavras.

Nas propostas teóricas de base gerativa, tendo em mente a visão neogramática em que a mudança afeta os sons gradualmente, fenômenos fonológicos variáveis foram encarados como sendo de superfície, limitados aos últimos estágios de derivação, sem, com isso, ter a possibilidade de acessar componentes que figuravam nos estágios primeiros dos modelos formalistas, assim como a morfologia. A concepção de gramática, nesse sentido, é unidirecional: cada componente seria alimentado pelo componente precedente sem poder, depois disso, retornar ao componente anterior (COETZEE, 2008; SCHWINDT, 2012).

Os estudos variacionistas encararam a heterogeneidade como parte do sistema, o que também motivou a inclusão de variáveis morfológicas nas suas análises. Essas variáveis, então, poderiam ser de vários tipos, tais como morfemas específicos caracterizando alvos ou gatilhos de processos, classes de palavras mais ou menos suscetíveis a certos processos ou até mesmo itens lexicais mais ou menos frequentes (COETZEE, 2008; SCHWINDT, 2012).

Guy (1991), em estudos acerca da redução das consoantes finais t/d do inglês, aponta que palavras não derivadas ou monomorfêmicas (como *mist* e *pact*) apresentam uma maior incidência do fenômeno de apagamento, seguidas pelas flexionadas que fazem parte do sufixo de passado em verbos irregulares (como *kept*, *left*) e, com o menor índice de apagamento, seguem-se as flexionadas no passado regular (*missed*, *packed*), em que -t e -d são os únicos marcadores da noção semântica de passado.

Com a descoberta de mais exemplos análogos aos de Guy, surge, então, a necessidade de lidar com a variação como uma parte integrante da gramática fonológica, tendo em vista que nem toda variação poderia ser relegada à implementação fonética. Surgem, com isso, vários modelos que incorporam formalmente a variação à sua gramática (LABOV, 1969; ANTILLA, 1997; BOERSMA, 1998; BOERSMA & HAYES, 2001; COETZEE, 2006). Apesar de esses modelos lidarem bem com a influência da gramática na aplicação de fenômenos variáveis, há, segundo Coetzee (2008), outro fator também conhecido por influenciar a variação, com o qual os modelos existentes não são capazes de lidar adequadamente, qual seja o léxico. Processos de apagamento, como o *t/d deletion*, são tipicamente mais aplicados em palavras com maior frequência de uso (COETZEE, 2008, p. 2-3):

Unfortunately, we do not have detailed information on how usage frequency impacts *t/d deletion*. In the literature, there is some informal acknowledgement that usage frequency is relevant to the application of this process. Many studies exclude forms like *and*, *n't*, *went* and *just* from their data since these forms typically show anomalously high rates of deletion.

Coetzee (2008) afirma que é possível projetar modelos gramaticais que deem conta das influências da gramática formal e também das influências extragramaticais, como frequência de uso lexical, simultaneamente. Tendo em vista a forte evidência de que esses fatores influenciam na variação e de que o usuário da língua tem acesso a esses dois tipos de informação, não é somente possível projetar esses modelos, mas também absolutamente necessário.

Em artigo intitulado *Condicionamento Morfológico em Fenômenos Fonológicos Variáveis do Português Brasileiro*, Schwindt (2012) retoma uma pergunta clássica em linguística, qual seja: processos fonológicos variáveis podem acessar informações lexicais? Para tentar responder a essa pergunta, o autor reexamina três análises em que fatores morfológicos se mostraram supostos condicionadores de variação fonológica: harmonia vocálica, (SCHWINDT, 1995), redução da nasalidade de ditongos finais átonos (SCHWINDT e BOPP DA SILVA, 2010) e vocalização da lateral pós-vocálica (COLLISCHONN e QUEDNAU, 2010).

Através dessas três reanálises, o pesquisador aponta que, relativamente à harmonia vocálica, não se pode sustentar a hipótese de influência de fatores morfológicos, tendo em mente que expoentes prosódicos (tais como contiguidade

fonológica de alvo e gatilho e tonicidade do gatilho) seriam suficientes para se obter os mesmos resultados. No que diz respeito à redução da nasalidade, Schwindt aponta que os dados sugerem que o processo é capaz de acessar a distinção entre nomes e verbos, sustentando, portanto, a hipótese de influência da variável 'classe de palavra'. Finalmente, no que atine à vocalização da lateral, parece haver, como na harmonia vocálica, uma motivação principalmente prosódica, considerando que o processo prefere a fronteira de palavra fonológica.

Com esses fenômenos em vista, pode-se dizer que há, sim, a possibilidade de algum tipo de influência morfológica na aplicação do fenômeno, principalmente, como vimos, em relação à redução da nasalidade em ditongos finais átonos. O autor aponta, pois, que essa constatação direcionar-nos-ia a outra questão, qual seja: se de fato há influência morfológica sobre a variação fonológica, que tipo de fator morfológico pode exercê-la? De acordo com os dados analisados no artigo, o fator *classe gramatical* foi o mais determinante. No entanto, o autor menciona que outra variável, dentre as morfológicas, ainda precisa ser também averiguada, qual seja a da frequência lexical.

Levando em conta as constatações anteriores, percebe-se que temos uma necessidade premente de ampliação e aprofundamento dos estudos acerca da frequência lexical e sua relação com fenômenos fonológicos variáveis que apresentam condicionamento morfológico. Ao pensarmos no português brasileiro, vemos que essa necessidade se faz ainda mais urgente, tendo em vista a maior complexidade morfológica do português comparativamente ao inglês, principalmente no que tange aos processos flexionais.

Como advoga Bybee (2007), a descoberta de tipologias como ordem de palavra, marcação de caso ou forma morfológica seria impossível sem um foco nos padrões gerais da linguagem. No entanto, já é tempo de examinar o papel que determinadas expressões e estruturas em particular exercem na gramática da língua e, para tanto, o estudo da frequência lexical se faz primordial. A pesquisa aqui proposta, pois, tem o intuito de contribuir com as análises já realizadas até o presente momento, buscando outras informações através do estudo de outros fenômenos do português brasileiro (doravante PB) que possam nos auxiliar na ampliação e aprofundamento de nosso conhecimento acerca do papel da frequência lexical e da sua relação com a variação fonológica condicionada morfológicamente.

Os fenômenos analisados relativamente à frequência serão os seguintes:

- a) A redução da nasalidade de ditongos finais átonos: dando continuidade aos resultados da pesquisa realizada por Battisti (2002) e Schwindt e Bopp da Silva (2010), fizemos uma reanálise dos dados, com o intuito de verificar o papel da variável frequência de uso lexical na aplicação do processo, e a relação disso com o status morfológico dos ditongos finais (se forem nomes terminados por *gem* no sufixo, como em *malandragem*; se nomes terminados por *gem* na raiz, como em *viagem*; se forem apenas nomes sem complexidade morfológica, como *homem*, ou se forem verbos flexionados, como *comem*, *ligaram*).
- b) O apagamento do /r/ final em verbos e em não verbos: dando continuidade à pesquisa realizada por Huback (2003, 2006), tencionamos averiguar, em novos dados coletados do falar carioca, novamente a relação entre frequência lexical e o status morfológico do /r/ final, se apenas constituinte final da lexia (como em *lugar*, *mulher*, *amor*), se parte de sufixo (como em *popular*, *professor*, *escritor*), se morfema de infinitivo verbal (*chamar*, *querer*, *pôr*).

OBJETIVOS

O objetivo desta pesquisa, de um modo geral, é estudar o papel da frequência lexical na implementação da mudança sonora e sua implicação relativamente aos domínios morfológicos dos itens lexicais. Para tanto, pretende-se:

1. Verificar, concernente aos fenômenos descritos e analisados, qual seria a relação existente entre frequência lexical e fatores linguísticos, principalmente entre i) frequência lexical e morfologia interna à palavra e ii) frequência lexical e classe gramatical, ou seja, se fenômenos sujeitos à influência da frequência também podem estar sujeitos à influência morfológica e/ou morfossintática.
2. Averiguar a relação existente entre i) frequência lexical e fatores sociais, tais como idade, sexo e escolaridade; e ii) frequência lexical e informante. Nosso intuito é verificar se a influência da frequência está no indivíduo, com seu

conjunto particular de itens léxicos, ou na comunidade, principalmente representada pelo fator social idade, tendo em vista que o léxico sofre influências geracionais.

3. Analisar quais modelos linguísticos, se abstracionistas ou se exemplaristas, apresentam maior adequação explanatória considerando os resultados dos fenômenos aqui apresentados.

HIPÓTESES

Nossas hipóteses são as seguintes:

1) A frequência lexical apresenta correlação positiva com a aplicação dos fenômenos de redução e de apagamento, ou seja: quanto mais frequente o item lexical, mais suscetível à aplicação dos processos estudados ele é. No entanto, essa sujeição está condicionada à morfologia e à classe gramatical: morfemas serão mais resistentes à redução ou ao apagamento, mesmo em itens lexicais mais frequentes.

2) A relação entre frequência lexical e aspectos extragramaticais é significativa. A variável social que mostra mais interação com frequência é idade, tendo em vista que léxico é geracional.

3) Os modelos com maior adequação explanatória são os modelos baseados em regras, ou seja, os modelos abstracionistas, tendo em vista que, em decorrência da hipótese 1, frequência lexical é uma variável que, mesmo podendo exercer papel na aplicação dos fenômenos, não opera significativamente na presença de regras fonológicas e/ou morfológicas.

A organização deste trabalho se dá da seguinte maneira: no capítulo 1, apresentamos nossa revisão de literatura sobre frequência lexical em fenômenos fonológicos variáveis condicionados morfologicamente, com a discussão das duas correntes de modelos teóricos: abstracionistas e exemplaristas. No capítulo 2, apresentamos trabalhos realizados no PB com foco na frequência lexical; também revisamos pesquisas sobre os dois fenômenos que são foco de análise deste trabalho, quais sejam Redução da Nasalidade e Apagamento de /R/. O capítulo 3 apresenta

nossa metodologia de trabalho, e o capítulo 4 traz nossos resultados e discussão. Em seguida, apresentamos nossas considerações finais.

1. SOBRE EFEITOS DE FREQUÊNCIA

1.1 Efeitos de Frequência: dos modelos Neogramático e Difusionista à Fonologia de Uso

Para os neogramáticos—uma escola alemã do final do século XIX—uma mudança sonora afeta todas as palavras que satisfazem as condições estruturais que governam a implementação da mudança, isto é, todas palavras que apresentam o contexto esperado em um determinado processo. Disso segue que as mudanças sonoras são foneticamente graduais e lexicalmente abruptas. Essas mudanças sonoras são condicionadas linguisticamente, configurando-se como regras fonológicas-entendidas, portanto, como regulares. Mudanças irregulares, quando existirem, são explicadas apenas por empréstimo ou analogia.

Schuchardt (1885), em contraste às ideias neogramáticas, foi pioneiro ao defender que palavras com diferentes faixas de frequência tendem a ser afetadas de forma diferente na mudança linguística: “rarely-used words drag behind; very frequently used ones hurry ahead. Exceptions to the sound laws are formed in both groups” (p. 58).

Wang (1969)—o principal representante da tese conhecida como difusão lexical—propôs que a palavra é a principal unidade da mudança e, como consequência, as mudanças sonoras são entendidas como lexicalmente graduais e foneticamente abruptas, já que certos itens lexicais aderem às mudanças gradualmente. Nessa abordagem, a ideia é a de que as mudanças não são mais condicionadas por regras fonológicas, mas por características lexicais como frequência e previsibilidade de ocorrência. Fatores como a familiaridade do item lexical, por exemplo, estariam entre os aspectos extra-gramaticais responsáveis pela maior ou menor vulnerabilidade de uma palavra à mudança sonora. Nessa perspectiva, uma mudança sonora ocorre inicialmente em uma palavra e é propagada para outras que apresentam uma estrutura fonológica similar. Pode haver palavras que não estão sujeitas à difusão lexical, assim como pode haver a possibilidade de que a mudança atinja todas as palavras potenciais. Irregularidades neste cenário são justificadamente esperadas, tendo em vista que a mudança sonora não precisa ser generalizada.

Labov (1981), ao analisar /æ/ tenso e frouxo norte-americano, defende uma distribuição paramétrica entre a mudança neogramática e a difusionista. A regularidade neogramática estaria em regras cujos outputs sejam de nível mais baixo (*low-level output rules*) e a difusão lexical na redistribuição de uma classe de palavra abstrata em outras classes abstratas (1981, p.304). No entanto, como argumenta Oliveira (1991, p.97), uma questão remanescente seria: quais palavras são afetadas antes na difusão lexical?

Phillips (1984) defendeu que mudanças fisiologicamente motivadas tendem a afetar palavras mais frequentes primeiro e, do contrário, mudanças não fisiologicamente motivadas afetam palavras menos frequentes. Redução vocálica, assimilação e apagamentos em geral estão entre os fenômenos fisiológicos/articulatórios levados em conta por Phillips. Com isso, a autora sustenta que fatores fonéticos não precisam ser necessariamente os únicos responsáveis pela mudança, mas a influência deles, certamente, precisa ser considerada no estudo das mudanças sonoras.

I do not, of course, mean to create the impression that phonetic factors alone CAUSED this sound changes. If that were true, there would indeed be no dialectal differences based on such changes, for all potential dialects would undergo the same changes at the same time, and hence never diverge from one another. The choice of which potential changes will actually prevail in a speech community is no doubt dependent on many things. But surely surface phonetic influence is a *sine qua non* of the sound changes (...). Hence, I will call these 'physiologically motivated' sound changes (PHILLIPS, 1984, p. 323).

Ao examinar as evidências históricas de Phillips (1984), Oliveira (1991) atenta para o fato de que alguns casos analisados por Labov como neogramáticos foram casos de difusão lexical nos seus estágios iniciais e defende, analogamente a Chen e Wang (1975), que todas as mudanças sonoras são implementadas lexicalmente, ou seja, que não há mudanças sonoras neogramáticas, apesar de podermos ter mudanças que, a longo prazo, pareçam neogramáticas.

Em artigo de 1995, Oliveira trata das características dos itens lexicais mais propensos a serem atingidos primeiramente por uma mudança sonora. O autor argumenta em prol de uma análise do léxico inserido nos enunciados em que ocorre, com sua informação contextual e não em termos de suas características virtuais. Oliveira (1995) propõe que traços como [+/-Frequente] e [+/-Formal] sejam traços atribuídos aos itens lexicais a partir do contexto em que eles se encontram, que

certamente varia de falante para falante, de comunidade para comunidade. Na sua proposta, o léxico é “um conjunto de traços que são construídos caso a caso, nas situações concretas de interação verbal” (p.88).

Phillips (1999), por outro lado, defende que o léxico mental, além de ser rico em detalhamento fonético, deve conter informação sobre a categoria gramatical e, também, fornecer informação sobre frequência lexical. Em trabalho publicado em 2001 (p. 123-124), a autora argumenta que “sound changes which require analysis—whether syntactic, morphological, or phonological—during their implementation affect the least frequent words first; others affect the most frequent words first.” Além disso, ela reconhece classe de palavra como um fator independente na mudança sonora, que tem precedência em relação à frequência lexical: palavras funcionais apresentam frequências elevadas e geralmente são pioneiras na mudança em contextos de redução ou assimilação. No domínio da classe de palavra, mudanças sonoras que demandam análises mais detalhadas têm como alvo palavras menos frequentes antes.

Mais recentemente, outra corrente teórica tem conquistado importantes avanços na discussão a respeito do papel da frequência lexical sobre a mudança, a Fonologia de Uso, defendida, entre outros, por Bybee (2001, 2002). Trata-se de um modelo de estocagem das palavras na mente dos falantes concebido a partir de algumas premissas básicas. A primeira delas se refere à representação das unidades linguísticas do léxico mental: tanto as unidades linguísticas como as não linguísticas são representadas da mesma forma, sendo que essa representação será moldada através da experiência de cada indivíduo.

The reason frequency or repetition plays a role in Grammar formation is that the mind is sensitive to repetition. This is a domain-general principle; that is, it does not apply just to language but to other cognitive domains as well (BYBEE, 2007, p. 8).

Com essa ideia, temos que palavras mais frequentes no vocabulário cotidiano de determinadas pessoas estão mais ativas e, por isso, serão acessadas com maior facilidade. Outra característica importante presente no modelo está na determinação das unidades de categorização e estocagem mental: palavras e/ou frases inteiras. Não é através de fonemas ou de traços fonológicos, portanto, que os indivíduos farão,

por exemplo, generalizações morfológicas, mas através da associação fonética e semântica promovida entre os itens estocados no léxico mental.

A Fonologia do Uso visa a abarcar todos os subsistemas (fonologia, morfologia, sintaxe, semântica), ainda que não os tome como primitivos representacionais. Além disso, o modelo não se restringe ao estudo das estruturas linguísticas, já que leva em consideração o fato de que o uso da língua, ao incluir todo o processamento cognitivo e as interações sociais, exerce um forte impacto sobre sua substância. Com isso, a representação de um item lexical, constituída basicamente por conteúdo fonético e semântico, pode ser constantemente moldada relativamente à sua experiência de uso em diferentes contextos (BYBEE, 2001).

A frequência, de acordo com Bybee, seria, então, um dos recursos de que o léxico mental se vale para categorizar itens léxicos. A pesquisadora propõe que o conceito de frequência seja dividido em duas medidas diferentes, quais sejam **frequência de ocorrência e frequência de tipo** (*token frequency* e *type frequency*, respectivamente). A frequência de ocorrência diz respeito a quantas vezes uma determinada palavra ou expressão ocorre em determinado corpus linguístico. Procura-se, por exemplo, quantas vezes (ocorrências) a palavra 'você' (um tipo) é registrada em determinado banco de dados; o número de casos encontrados será a frequência de ocorrência dessa palavra. Já a frequência de tipo pode corresponder à frequência de um determinado padrão ou estrutura linguística. Podemos procurar em certo corpus o número de palavras que apresenta a sequência *gem* (por exemplo *viagem*, *vantagem*, *camaradagem*) – esse número será a frequência do conjunto desses nomes, um tipo. No que diz respeito à produtividade, a frequência de tipo é de extrema relevância, tendo em mente que, se determinados padrões tendem a ser bastante frequentes, eles serão aplicados a outros itens que se enquadrarem em estrutura similar.

High-frequency words and phrases grow strong with repetition and loom large, forming looser connections with other items, while low-frequency words and expressions are less prominent but gain stability by conforming to patterns used by other items. General patterns dominate networks where more specific patterns can be overpowered unless represented by high-frequency items. Words that have phonological similarities cluster together; constructions are connected if they have properties in common. Instances of constructions that grow to high-frequency slowly disengage from the more general pattern to become independent constructions. Thus the phonetic and semantic substance of language is ever being shaped by the effects of usage (BYBEE, 2007, p. 9).

A Fonologia do Uso (BYBEE, 2001, 2002), portanto, entende que as representações linguísticas são múltiplas, redundantes. A Teoria dos Exemplos, primeiramente apresentada por Johnson (1997) e, após, desenvolvida por Pierrehumbert (2001, 2003), é uma teoria que acolhe essas múltiplas representações linguísticas, propondo-se a lidar com a variação, a gradiência, a difusão lexical, os efeitos de frequência lexical, os contextos favorecedores e os efeitos probabilísticos. Segundo a Teoria dos Exemplos, os falantes retêm na memória os exemplos das palavras que eles já ouviram com alto detalhamento fonético. Essa pretensa nuvem de exemplos fornece ao falante as informações sobre a pronúncia de itens lexicais individuais, os detalhes da realização fonética, os padrões de variação evidentes na comunidade e a distribuição quantitativa desses fatos.

1.2 Redução acústica: modelos abstracionistas versus modelos exemplaristas

Ernestus (2014) afirma que o fenômeno de redução linguística tem recebido pouca atenção da literatura até então: na fala espontânea, mais do que uma variante para determinada palavra é antes a regra que a exceção. A autora discute como determinados dados desafiam as duas linhas teóricas que são populares tanto na linguística quanto na psicolinguística, quais sejam os modelos de representações abstratas e os baseados em exemplos. Ambos modelos, segundo ela, têm características que parecem necessárias para dar conta do processo de redução.

No que diz respeito aos modelos abstracionistas, Ernestus (2014) aponta que, em sua versão mais básica, o léxico mental contém apenas uma representação lexical para cada palavra (ou morfema), que consiste de símbolos abstratos: os fonemas. As variantes na pronúncia são derivadas desta única representação, por meio de processos gerais, que se aplicam a várias palavras. Nesses modelos, a produção de uma palavra envolve a ativação da sua representação lexical e a aplicação de regras fonológicas (ou restrições, como na OT), que adapta a palavra ao seu contexto fonológico. A representação fonológica abstrata resultante é traduzida por mecanismos fonéticos (regras ou restrições) em eventos articulatórios/acústicos. Os modelos abstracionistas dão conta da produção de redução acústica tanto pelo componente fonológico quanto pelo fonético. Já que a redução acústica é opcional,

esses mecanismos também o são. Regras fonológicas opcionais têm a sua frequência de aplicação especificada nas descrições estruturais; já na Teoria da Otimidade, a opcionalidade dos processos fonológicos pode ser obtida ao assumir que os falantes escolhem entre diferentes rankings de restrições.

No que tange aos modelos exemplaristas, o léxico mental contém muitos exemplares de cada palavra, que juntos formam uma nuvem. Esses exemplares representam os diferentes tokens encontrados pelos usuários da língua, tanto na sua produção quanto na sua percepção. As representações armazenadas não consistem de símbolos abstratos nem se tornam abstratos. A suposição de representações lexicais altamente especificadas é embasada por um número considerável de estudos experimentais sobre produção e compreensão do discurso. Os modelos de exemplares tratam da redução acústica da mesma forma que tratam das diferenças de pronúncia entre falantes e situações: todas as variantes estão armazenadas no léxico mental. A produção de uma palavra envolve a ativação de todos os seus exemplares (ERNESTUS, 2014)

Muitos estudos defendem a ideia de que a redução acústica resulta, acima de tudo, da implementação fonética, mostrando que a velocidade do discurso e a previsibilidade das palavras e das palavras vizinhas afetam o grau de redução. Outros estudos sugerem que a produção da redução acústica é um processo mais complexo, que só pode ser abarcado com o armazenamento lexical de pelo menos algumas variantes. A autora aponta que a redução acústica é resultado de várias características do fluxo do discurso, assim como do seu falante e de suas palavras.

A conclusão a que Ernestus (2014) chega é a de que tanto modelos abstracionistas quanto exemplaristas precisam ser aperfeiçoados para explicar os dados de redução acústica. O modelo final deve permitir o armazenamento de diferentes variantes de pronúncia, mas também reservar um papel importante para a implementação fonética. Isso, segundo a pesquisadora, é mais bem explicado por modelos híbridos, que assumem tanto representações abstratas quanto exemplares. No entanto, nenhum modelo híbrido formulado até então é capaz de explicar os dados de redução.

Em artigo de 2008, Coetzee expõe seu interesse pela interação estabelecida entre fatores de cunho gramatical e extra-gramatical na variação fonológica, já que

pesquisas realizadas sobre frequência lexical, como a de Bybee (2000), não mostram se a contribuição da frequência e da gramática são independentes ou se há uma interação mais complexa entre esses dois fatores. Ao estudar o apagamento de t/d finais, Coetzee se pergunta: a frequência é responsável por uma maior incidência de apagamento em todos os três contextos gramaticalmente definidos no estudo, quais sejam posição pré-consonantal (*west bank*), posição pré-vocálica (*west end*) e posição final (*west###*)? Através de uma análise em Teoria da Otimidade, com base em restrições, o autor afirma que duas predições podem ser feitas sobre como a gramática, ou seja, o ranking de restrições, e a frequência de uso, isto é, as funções de distribuição lexical, interagem.

De acordo com o pesquisador, dada qualquer palavra, sem atentar para sua frequência de uso, o apagamento ocorre mais em posições pré-consonantais e menos em posições finais. Essa constatação segue-se do ranking estabelecido entre as três restrições de marcação (Pre-C >> Pre-V >> Pre-##) e é, pois, resultado da gramática. Após, vê-se que, dados dois itens lexicais lex_1 e lex_2 , sendo lex_1 mais frequente que lex_2 , lex_1 terá maior incidência de apagamento em todos os contextos gramaticais. Então, quanto mais frequente for lex_1 , mais ele estará propenso a ser avaliado por baixos rankings de restrições de fidelidade, tendo em vista a forma da sua função de distribuição.

As constatações do estudo também implicam que não é muito informativo comparar as incidências de apagamento de dois itens lexicais aleatoriamente. Um item lexical altamente frequente como *just* pode apresentar uma incidência de apagamento no contexto mais conservador (posição final) do que um item lexical menos frequente como *bust* no contexto mais liberal (pré-consonantal). Dados interpretados fora de contexto podem dar a impressão de que a gramática (contextos definidos fonologicamente) não contribui para a probabilidade de apagamento. É, portanto, muito importante sempre atentarmos para a forma através da qual um processo variável influencia um item lexical específico como parte de um sistema maior, e não apenas como um dado individual.

Segundo Coetzee, a gramática ainda é o fator primordial na decisão de ocorrência ou não da variação. Se as condições estabelecidas pela gramática não forem atendidas, a variação não será passível de ocorrência e a frequência de uso

não terá influência alguma no *output* final. No entanto, se as condições estabelecidas pela gramática forem atendidas, a variação será possível e só então a frequência de uso será capaz de influenciar a forma pela qual processos variáveis influenciam itens lexicais específicos.

O pesquisador aponta que não tem, até a fase de estudo em questão, informação detalhadamente suficiente acerca do fenômeno *t/d deletion* para testar a validade dessas constatações; no entanto, essas previsões advêm da arquitetura básica do modelo de gramática por ele proposto e tem-se a oportunidade e a necessidade de testá-lo com outros fenômenos.

Como afirmou Guy (2014), não é trivial conciliar modelos de gramática baseados em regras e modelos baseados no uso, tendo em vista que os primeiros primam pela abstração, pela regularidade e pela generalização, enquanto os últimos buscam justamente os fenômenos fonológicos lexicalmente diferenciados, incluindo variabilidade, gradiência e propriedades probabilísticas. O autor assinala que as abordagens mais tradicionais no que tange à fonologia têm sido, de fato, abordagens baseadas em regras, com a postulação de representações mentais abstratas das palavras, as quais são submetidas a operações fonológicas capazes de capturar os padrões de sons mais generalizáveis de determinada língua.

Guy (2014) destaca, ainda, que a regularidade apontada pelos neogramáticos é bem justificada, uma vez que a maioria das mudanças fonológicas não deixou resíduos históricos de segmentos que não sofreram o processo de mudança em palavras excepcionais. Segundo o autor, o modelo baseado em regras, desenvolvido pelos neogramáticos, apresenta uma adequação explanatória bastante satisfatória, considerando que prevê produtividade, ou seja, prevê a capacidade de o falante saber pronunciar neologismos e empréstimos lexicais para os quais não se tem modelos prévios de pronúncia, além de ser capaz de realizar operações fonológicas abstratas em determinadas classes de sons por todo o léxico. No entanto, apesar de alcançar adequação explanatória, o modelo de regras não está completamente aparelhado para dar conta de alguns tipos de fatos fonológicos, especialmente os que envolvem itens lexicais em específico. Para remediar essa insuficiência, surge a Fonologia do Uso (BYBEE, 2001, 2002) e a Teoria dos Exemplares (PIERREHUMBERT, 2001).

Guy (2014), discutindo o tratamento de determinados fenômenos na Teoria de Exemplares, levanta alguns questionamentos para o modelo: 1) como alguém pode

produzir determinado item lexical para o qual não há exemplares prévios?; 2) como explicar operações abstratas realizadas regularmente em todo o léxico?; 3) podemos prever de forma correta todos os fatos prescindindo das regras?; 4) a fonologia pode sobreviver sem uma dieta balanceada, que inclui regras e abstração juntamente com o uso?. Diante do que entende como lacunas na abordagem estrita de exemplares, o autor sugere que a alternativa ideal, que incorpora tanto a produtividade gerativa quanto a precisão quantitativa, é ainda o modelo de Regra Variável (RV), da Sociolinguística laboviana. Para Guy, a análise de RV preserva as vantagens dos modelos baseados em regras, tais como a abstração e a capacidade de representação dos processos categóricos, mas é capaz de resolver muitas das limitações desses modelos formais ao fazer uso da quantificação probabilística (qualquer processo ou restrição fonológica deve associar-se a uma probabilidade, o que permite o tratamento da variação e da gradiência). Por outro lado, reconhece que um dos atributos fundamentais da Teoria dos Exemplares, os efeitos lexicais, não está incorporado no modelo de RV. Isso exige, por exemplo, que, para lidar com exceções lexicais de processos fonológicos variáveis, devam ser postuladas múltiplas representações subjacentes. Para ilustrar, considere-se um vocábulo como *and* (*e*), do inglês: a Teoria de Exemplares explica o fato de ser a palavra que apresenta maior índice de apagamento da consoante final por se tratar também da palavra mais frequente na língua com oclusiva coronal final; o modelo de RV também pode explicitar essa constatação por meio do pressuposto de que *and* apresenta duas representações subjacentes, uma com e outra sem o /d/ final. Guy defende que a abstração fonológica não deve ser abandonada simplesmente porque a sua implementação tradicional não inclui quantificação. Para o autor, uma teoria fonológica adequada precisa tanto da gramática e das regras quanto da memória para dar conta de todos os fatos. Apenas um modelo de RV enriquecido por múltiplas representações subjacentes, defende Guy, é capaz de fornecer uma versão da fonologia mais equilibrada, a qual apresenta abstração e, ao mesmo tempo, também presta atenção nas quantidades.

Seguindo o debate presente em Myers & Guy (1997), Walker (2012) investigou algumas afirmações sobre a influência dos efeitos formais, funcionais e de frequência no apagamento de t/d final do inglês de Toronto. Apesar de os resultados iniciais sugerirem uma clara correlação entre frequência lexical e apagamento, uma vez que

a interação e os efeitos lexicais foram levados em conta, apenas os grupos de fatores concernentes à fonologia e à morfologia se mostraram significativos. Alguns relatos prévios de efeitos de frequência podem resultar de diferentes medidas de frequência, juntamente com a contribuição da sobreposição de alguns grupos de fatores. Ao assinalar que restrições formais apresentam o maior efeito na variação, com restrições fonológicas em primeiro plano e morfológicas em segundo, Walker (2012) conclui que a hipótese baseada no uso, de que a frequência lexical influencia o apagamento de t/d, não apresenta suporte consistente depois de se considerar a interação com o *status* morfológico e com um pequeno grupo de itens lexicais. O autor defende, pois, que precisamos reconhecer que a frequência não opera monotonicamente, tendo uma interação mais dinâmica com o léxico.

1.3 Implicações para a Morfologia

Bybee (1988, 1995) propõe um modelo que permite o entendimento da estrutura interna da palavra como um conjunto de relações com outras palavras. O modelo proposto permite a identificação de parte de uma palavra como uma unidade de significado recorrente sem a necessidade de atribuir significado ao restante da palavra; além disso, permite a identificação de relações fonológicas mesmo na falta de uma relação semântica mais clara. O modelo também fornece variados graus de parentesco entre palavras como uma consequência de características fonológicas e semânticas. A noção de força lexical, inclusive, está no modelo para dar conta do efeito de frequência no acesso lexical. Padrões morfológicos recorrentes emergem como acúmulos de conjuntos similares ou idênticos de conexões e são descritos como esquemas: conexões semânticas e fonológicas paralelas constituem relações morfológicas. A diferença mais importante nesse modelo é a de que as generalizações, que em teorias abstracionistas são chamadas de regras, são parte das representações. Elas emergem da organização da substância fonética e semântica da língua e não são independentes dessa substância, ou seja, não são independentes do uso.

Losiewicz (1992, 2009), no que diz respeito ao efeito da frequência na morfologia, apresenta evidências de que um segmento fonético morfêmico tem duração acústica maior que um segmento fonético idêntico que não seja morfema. Isso sugere, segundo a pesquisadora, que a carga semântica, refletida na duração

acústica, pode ser a principal variável que afeta a natureza do acesso lexical em palavras multimorfêmicas. Além disso, a autora aponta que palavras de baixa frequência têm maior duração acústica que palavras de alta frequência; também, o morfema ED, foco de seu estudo, apresentou maior duração acústica em verbos de baixa frequência que em verbos de alta frequência. Essa é uma evidência, segundo a autora, de que a frequência da palavra afeta a representação lexical e o acesso a palavras multimorfêmicas.

Hanique e Ernestus (2012) examinam o papel da morfologia na redução acústica ao discutir e reanalisar dados da literatura, tais como os de Losiewicz (1992) e os de Hay (2003). As pesquisadoras não encontraram evidência de que o status morfológico de um segmento afeta seu grau de redução, o que vai de encontro a premissas de modelos tradicionais de processamento morfológico que assumem que palavras complexas não são armazenadas no léxico mental, mas computadas por meio de seus morfemas através de regras (CHOMSKY e HALLE, 1968; PINKER, 1991). As autoras defendem, portanto, que as palavras são armazenadas como unidades completas no léxico mental, sendo acessadas diretamente. Essa visão sustenta modelos psicolinguísticos como os de Bybee (2001), em que a estrutura morfológica não é altamente relevante para a produção da fala, tendo em vista que é através da conexão entre palavras relacionadas que a estrutura morfológica emerge.

As autoras defendem que um segmento que identifica uma palavra (ou seja, a carga de informação da palavra) tem um papel mais determinante na redução que um segmento que identifica a estrutura morfológica de tal palavra. Isso, segundo as autoras, sugere que as unidades importantes na produção do discurso são palavras, não morfemas. A revisão de trabalhos prévios feita pelas autoras demonstra que não há evidência de que a estrutura morfológica apresenta um papel substancial na redução; além disso, muitos estudos que apresentam papel da estrutura morfológica, segundo as pesquisadoras, parecem fazê-lo porque o efeito reportado não é relacionado apenas com a estrutura morfológica, mas também com a carga de informatividade da palavra.

RESUMO DO CAPÍTULO: Neste capítulo, apresentamos de forma sucinta as ideias dos Neogramáticos, dos Difusionistas, da Fonologia do Uso e da Teoria dos

Exemplares no que diz respeito à variação e mudança na sua relação com frequência lexical. Também tratamos de duas abordagens de análise linguística, quais sejam as abstracionistas e as exemplaristas e, ao final, apresentamos alguns estudos que abordaram a relação entre frequência lexical e morfologia.

Tendo em vista nosso objetivo de número 3, qual seja analisar quais modelos linguísticos, se abstracionistas ou se exemplaristas, apresentam maior adequação explanatória, verificaremos nos dados deste trabalho se regras fonológicas e morfológicas se apresentam mais determinantes para a aplicação dos fenômenos de redução e apagamento, como defendido nos trabalhos de Walker (2012), ou se exemplares, como defendido por Hanique e Ernestus (2012)

Na próxima seção, relatamos alguns estudos realizados no PB acerca dos efeitos de frequência. Também fazemos uma revisão da literatura dos fenômenos que são foco deste estudo, quais sejam Redução da Nasalidade e Apagamento de /r/.

2 ESTUDOS SOBRE VARIAÇÃO NO PB

Nesta seção, em um primeiro momento, apresentamos trabalhos realizados no PB com foco na frequência lexical; após, revisamos pesquisas sobre os dois fenômenos que são foco de análise deste trabalho, quais sejam Redução da Nasalidade e Apagamento de /r/.

2.1 Fenômenos fonológicos variáveis no PB e frequência lexical

Em artigo de 2001, levando em conta a proposta da Difusão Lexical, Cristóforo-Silva analisa três casos de variação/mudança no português brasileiro, quais sejam: 1) a vocalização da lateral em final de sílaba, 2) a palatalização de oclusivas alveolares antes de [i] e 3) a quebra de encontros consonantais (mais especificamente: /Cr/ → [C] seguido de vogal). Esses três casos refletem processos de variação que, segundo a autora, foram ou estão sendo implementados lexicalmente.

No caso em 1, a autora aponta que a mudança está praticamente concluída em quase todas as variedades do português; já em 2, apesar de a aplicação da regra ocorrer significativamente, ainda há variedades do PB em que o uso da oclusiva [t] ao invés de [tʲ] é preponderante. No caso em 3, a mudança ainda é incipiente. Cristóforo-Silva defende que, nos três casos, a aplicação dos processos de vocalização, palatalização e quebra de encontro consonantal se deu por difusão lexical, ou seja, algumas palavras começaram a ser pronunciadas com a forma inovadora e, em 1, a mudança alcançou praticamente todas as palavras da língua (evidência para isso seria a forma lexicalizada em 'gol/gous'); em 2, como já apontado, o fenômeno é relativo a dialetos, mas também há evidência a favor da teoria difusionista, já que se encontram formas lexicalizadas como "Pa[dʒi]Cícero" e temos um comportamento alternante em termos das restrições que regulam a boa formação da estrutura no PB (ele[ti]cista). Acerca desse fenômeno, a pesquisadora aponta que seria interessante avaliá-lo em comunidades em que ainda ocorrem com frequência formas concorrentes com [ti, tʲi] e [di, dʒi]. Para finalizar, relativamente ao fenômeno em 3, a pesquisadora advoga que a proposta da difusão lexical é a mais adequada, tendo em vista que a mudança sonora se dá no nível da palavra, podendo não se propagar para todo o léxico, justificando a ainda incipiência de aplicação desse fenômeno.

Em estudo acerca da monotongação de ditongos decrescentes [al, el, ol, ul] em sílabas abertas e fechadas, Haupt (2011) analisou quantitativamente as ocorrências das entrevistas do banco de dados do VARSUL Florianópolis, com o objetivo de verificar os efeitos de frequência de uso no fenômeno. Os resultados apontaram que a monotongação ocorre com itens lexicais com alta frequência de ocorrência, mesmo em contextos não considerados propícios fonologicamente. O efeito foi averiguado tanto em sílabas fechadas (*mais*) quanto em abertas finais (*comecei*). Ademais, Haupt aponta que palavras pouco frequentes que carregam informação morfossintática (como as palavras no plural) monotongaram menos e a mesma tendência de preservação do ditongo que carrega informação morfossintática foi observada nos monossílabos com alta frequência de ocorrência, tais como *vai, sei, foi e fui*. Na pesquisa em questão, houve também monotongação em itens lexicais pouco frequentes; nesses casos, Haupt considera-os itens lexicais suscetíveis à reanálise ou que o fenômeno foi condicionado pela frequência de tipo, já que os ditongos naquelas estruturas são pouco frequentes na língua, deixando prevalecer o padrão mais recorrente nesses contextos, qual seja o monotongo. Haupt defende, através de sua análise, que é possível considerar o item lexical como lócus da mudança, assim como preconiza o modelo da Fonologia de Uso.

Huback (2013) divide sua análise de casos concretos do PB em dois momentos: inicialmente, no estudo de fenômenos que afetam palavras mais frequentes primeiro e, depois, no estudo de fenômenos que afetam palavras menos frequentes primeiro. Dentre os fenômenos que afetam palavras mais frequentes analisados pela pesquisadora, temos o apagamento do /R/ final em nominais (apresentado na seção 2.3 deste trabalho), a palatalização de /S/ e o apagamento de /t/. No que tange aos fenômenos que atingem palavras menos frequentes, temos análises acerca do plural das palavras terminadas em /-ão/ no singular e o plural das palavras terminadas em ditongo em /-u/ no singular.

Baseada no referencial da Fonologia de Uso, Huback (2013, p. 82-83) aponta algumas consequências da interação entre as medidas de frequência:

- palavras com alta frequência de ocorrência estão mais ativas no léxico mental e, por isso, são acessadas mais rapidamente, ao passo que palavras de baixa

frequência necessitam estabelecer conexões com outros itens lexicais semelhantes para serem ativados;

- palavras irregulares e frequentes apresentam muitas conexões no léxico mental e, por isso, resistem a mudanças analógicas, ao contrário de palavras irregulares pouco frequentes, que, por não apresentarem força suficiente para manter sua irregularidade, estão mais propensas a adotarem os paradigmas mais frequentes da língua;

- palavras derivadas que apresentam frequência de ocorrência baixa ou média dependem da sua classe para serem lembradas e, através disso, estabelecem conexões mais fortes entre si do que palavras derivadas de alta frequência – itens pouco frequentes, a cada vez que são acionados, reforçam a coesão de sua rede; em contrapartida, itens com alto índice de frequência de ocorrência são ativados por si próprios autônoma e automaticamente, atribuindo à frequência de tipo a tarefa de garantir a produtividade de seu grupo.

Acerca do segundo fenômeno que afeta palavras mais frequentes, Cristóforo-Silva e Oliveira (2004) fizeram uma análise sobre o apagamento de [tʃ] na cidade de Belo Horizonte. Na pronúncia padrão, não ocorre a palatalização de /s/ em fronteira de sílaba. Apesar disso, por causa da assimilação de gestos articulatórios, quando o /s/ é seguido por [tʃ], ele é realizado foneticamente como ([ʃ]): “destino” – [deʃʃinʊ], “castigo” – [kaʃʃigʊ], etc. Seguido a isso, as autoras analisaram um desdobramento desse fenômeno de palatalização: em algumas sequências de –sti– (como nos exemplos acima), às vezes o [tʃ] é cancelado, e permanece apenas a sequência [ʃi], como em [deʃʃinʊ] e [kaʃʃigʊ]. Com esse quadro em vista, a análise acerca da frequência de ocorrência se fez pertinente e os resultados mostraram que palavras mais frequentes de fato foram mais suscetíveis ao apagamento da africada [tʃ]:

A alta frequência de ocorrência desses itens faz com que os gestos articulatórios ocorram mais rapidamente, em menos tempo e, conseqüentemente, a redução fonética ocorra. A princípio, parece que a hipótese de Phillips (1984), de que mudanças fisiologicamente motivadas afetam as palavras mais frequentes primeiro, se corrobora a partir dos dois fenômenos analisados (HUBACK, 2013, p. 87).

Huback (2013) analisa ainda as propriedades relativas às palavras menos frequentes no léxico. Um dos fenômenos destacados é o plural de palavras que

terminam em [-ão] no singular, as quais apresentam três formas distintas de pluralização, quais sejam: [-ões] (“balão” – “balões”), [-ãos] (“irmão” – “irmãos”) e [-ães] (“capitão” – “capitães”). Como se sabe, no latim essas formas eram diferentes tanto no singular quanto no plural e, por isso, não havia problema relativamente à forma de plural que se aplicava a cada tipo de palavra. Porém, no PB atual, como todas as formas de singular restaram idênticas, não há correlato fonético (ou ortográfico) para que se opte por um dos três morfemas de plural: “por causa disso, em corpora do PB já encontramos formas como *cidadões* ou *capitões* em vez de *cidadãos* ou *capitães*” (HUBACK, 2013, p.88).

Ao fazer uma busca pela frequência de tipo dessas terminações no Dicionário Eletrônico Houaiss, Huback (2013) percebe que o morfema [-ões] apresenta frequência de tipo muito mais alta que as demais terminações. Por causa disso, a hipótese adotada pela pesquisadora é a de que palavras etimologicamente pluralizadas em [-ães] ou [-ãos] estão adotando plurais em [-ões] devido à alta frequência de tipo dessa classe:

Se cada ocorrência de um item lexical (especialmente de frequência baixa ou média) fortalece sua rede, podemos supor que a rede em [-ões] é mais forte que as duas outras e, por isso, pode atrair membros em [-ãos] ou [-ães]. Além disso, nossa hipótese de trabalho também era de que palavras pouco frequentes poderiam ser as primeiras a sofrer essa variação. Se um item é pouco usado, depende de sua rede para que seja lembrado. Já que sua rede (em [-ãos] ou [-ães]) tem baixa frequência de tipo, não garante a preservação da forma de plural. Sendo assim, palavras de baixa frequência deveriam ser as primeiras a adotar formas variantes em [-ões], ao passo que palavras de alta frequência, por sua representação autônoma no léxico mental, deveriam manter seu plural de baixo tipo (em [-ãos] ou [-ães]) (HUBACK, 2013, p. 88-89).

Os dados presentes na análise da autora demonstram a tendência de que palavras mais frequentes resistam mais à variação analógica que palavras pouco frequentes.

Um outro fenômeno analisado por Huback (2013) foi o plural das palavras terminadas em ditongo em /-u/ no singular. Ao termos em mente que o [-l] em posição de coda é vocalizado na maioria dos dialetos do PB, não existe diferença fonética entre as letras L e U nos itens: “carnaval” / “berimbau”, “confortável” / “museu”. A partir disso, classes de palavras terminadas em ditongo em [-u] e em [-l] começaram a mostrar variações na realização pluralizada (Huback, 2013, p. 90): “Itens como *degrais*

e *troféis* (em vez de *degraus* e *troféus*) já são encontrados em corpora do PB.” Em termos de frequência de tipo, o Dicionário Houaiss aponta que os itens terminados em [-l] compõem um percentual mais significativo no léxico do PB que os itens terminados em ditongo em [-u]. Por causa dessa diferença de frequência de tipo, as mesmas hipóteses feitas para os plurais em [-ão] foram levantadas:

Uma vez que não existe distinção fonética entre [-l] e [-u] em fim de palavra, essas duas classes de plural sofreram um certo grau de fusão que permite que, em princípio, qualquer um dos dois plurais (em [-is] ou [-us]) possa ser aplicado às palavras. Sendo a frequência de tipo de [-l] em fim de palavra mais alta que a de [-u], supõe-se que itens em ditongo em [-u] adotem o plural em [-l] (HUBACK, 2013, p. 90)

É claro que essa questão abordada pela autora é passível de discussão, haja vista haver muitas formas derivadas que recuperam o /l/; o caso de *chapéis*, por exemplo, pode ser mera hipercorreção. Ainda assim, através dos fenômenos analisados, vê-se que, de acordo com a argumentação de Phillips (1984, 2001), há uma tendência de as palavras mais frequentes estarem mais propensas a fenômenos articulatorios de assimilação/apagamento e em resistir à variação analógica, por terem representação autônoma no léxico mental. Diferentemente, as palavras de baixa frequência podem ser as primeiras a sofrer as variações de cunho analógico.

As próximas duas seções apresentam a revisão da literatura dos dois fenômenos aqui estudados, quais sejam Redução da Nasalidade (seção 2.2) e Apagamento de /R/ (seção 2.3).

2.2 Estudos sobre Redução da Nasalidade em PB

Nesta seção apresentamos de forma sucinta os principais estudos sobre o primeiro dos dois fenômenos fonológicos variáveis que analisamos nesta tese, a redução da nasalidade em português brasileiro, fixando-nos particularmente no papel de classe de palavra, haja vista a relação dessa variável com frequência lexical – objeto que nos propusemos a discutir inicialmente.

Os primeiros estudos sobre redução da nasalidade em PB são de Votre (1978) e Guy (1981). A análise de Votre revelou que as variáveis linguísticas têm um papel

preponderante no condicionamento do fenômeno, comparativamente às variáveis extralinguísticas. Para o pesquisador, tendo em vista que a redução se estende a todas as categorias presentes em cada uma das variáveis, esse seria um fenômeno de variação em mudança. Guy estudou o fenômeno na fala de cariocas alfabetizando do MOBRAL¹. O autor atestou, no que diz respeito às variáveis linguísticas, que consoantes palatais precedentes favorecem a aplicação do processo de redução. Por isso, palavras terminadas em *gem* também tiveram grande aplicação em seu estudo.

Battisti (2002) separou em sua pesquisa nomes em geral de nomes terminados em *gem*. Vocábulos com essa terminação podem estar sujeitos a desnasalização por diferentes razões. Uma delas é sua instabilidade na história das línguas da península ibérica (como em espanhol *pasaje*, *viaje*, mas em português *passagem*, *viagem*), que resultou, inclusive, em formas duplas no registro padrão do português, como *garagem/garage*. No estudo de Battisti, de fato nomes em *gem* foram os campeões de aplicação do processo de redução, seguidos de nomes em geral. Além disso, dentre as variáveis linguísticas, classe morfológica foi a que demonstrou maior influência na aplicação do fenômeno, com aplicação maior em nomes do que em verbos.

Schwindt e Bopp da Silva (2010) apresentaram um panorama da redução da nasalidade em ditongos átonos finais no português do sul do Brasil, que incluiu as amostras de Battisti (2002) e de Bopp da Silva (2005); os autores ampliaram as amostras para todas as cidades que compõem o Projeto VARSUL, como já abordado na seção de Metodologia deste trabalho. Nos seus resultados, no que diz respeito à variável *classe de palavra*, os autores observaram que o fenômeno da redução é favorecido em nomes – primeiramente os sufixados em *-gem* e, em seguida, nos nomes em geral. Já a classe dos verbos mostrou-se, quase que completamente, inibidora do processo. Tendo em vista esses resultados, os pesquisadores especializaram a análise para os não verbos que mais se repetiram (*homem*, *jovem*, *viagem* e *origem*) e para os diferentes tempos verbais (pretérito perfeito do indicativo, pretérito imperfeito do subjuntivo, presente do indicativo, presente do subjuntivo, futuro do pretérito do indicativo, imperativo, pretérito imperfeito do indicativo, infinitivo

¹Os dados de Guy provêm do projeto Competências Básicas do Português, coordenado pela professora Miriam Lemle, da UFRJ, que era alimentado por dados de informantes do Movimento Brasileiro de Alfabetização. O MOBRAL foi criado pela Lei nº 5.379 e funcionou de 1967 a 1985.

peçoal e futuro do subjuntivo). O objetivo, com esta subdiviso dos verbos, era verificar se aqueles tempos que so passveis de neutralizao com outro tempo resistiam mais ao processo (ex. *eles pegu∅ versus eu pego; ele pode versus eles pode∅*). Em relao aos no verbos repetidos, foi observado que, em se retirando da rodada esses vocbulos, a ordem dos resultados se mantinha sem qualquer alterao. No que concerne  subdiviso dos tempos verbais, os dados no permitiram confirmar a hiptese de que formas no neutralizveis estariam mais sujeitas  reduo do que formas neutralizveis.

Schwindt, Bopp da Silva e Quadros (2012) retomaram a anlise sobre a influncia da morfologia na realizao do fenmeno. Seus resultados corroboraram os de estudos anteriores: as taxas de reduo da nasalidade mostraram-se mais altas em formas nominais do que em formas verbais e nomes que possuem a terminao *gem* apresentaram-se como mais sujeitos ao processo, com o incremento da observao da influncia dessa terminao independentemente de fazer parte da raiz (ex. *vantagem*) ou de se constituir como sufixo da lngua (ex. *pilantragem*). A reanlise do contexto fonolgico precedente realizada por Schwindt (2012) aponta que palatais nessa posio, tanto em nomes quanto em verbos (ex. *ga[n]am, dei[[j]am* etc.), favorecem o processo. Tambm o desmembramento dos tempos verbais mostrou-se pouco significativo, ainda que se sobressaia o papel do pretrito perfeito do indicativo, com aplicao acima do ponto neutro. Em relao a esse tempo, o principal argumento  o de ser a nica forma verbal em que a nasal final poderia no carregar sozinha a informao de nmero e pessoa (ex. *eles chega+∅+ram*, em oposio a *eles chega+∅+m* ou *eles chega+va+m*). Isso vai ao encontro de um princpio proposto por Kiparsky (1982[1972]) para a diacronia lingustica, conhecido como "distintividade contrastiva", segundo o qual estruturas gramaticais redundantes so menos resistentes  mudana.

Os autores concluem seu estudo com a hiptese de que a reduo da nasalidade tende a preservar a expresso de morfemas, podendo se aplicar mais livremente quando esta expresso no est em risco.

Ao estudar a redução do ditongo nasal [ãw]² postônico na morfologia verbal do português brasileiro, Cristófar-Silva, Fonseca e Cantoni (2012) propuseram uma análise pautada na perspectiva teórica de modelos multirrepresentacionais (BYBEE, 2001, 2002; PIERREHUMBERT, 2001, 2003). A pesquisa foi realizada com 12 sujeitos naturais e residentes de Belo Horizonte: 6 com idade entre 20 e 25 anos e 6 com idade acima de 35 anos. Desses 12 sujeitos, 6 eram homens e 6 eram mulheres – todos com nível superior completo ou em andamento. Seus dados foram coletados em laboratório com tratamento acústico e englobaram verbos da primeira conjugação nos seguintes tempos verbais: presente do indicativo, pretérito perfeito do indicativo e pretérito imperfeito do indicativo. Foi solicitado aos sujeitos que lessem frases-veículo dispostas em uma apresentação de *Power Point*. Nos slides também havia uma figura relacionada ao tema da frase em questão. Os dados foram controlados quanto aos valores de frequência lexical, os quais foram obtidos no banco do Projeto Avaliação Sonora do Português Atual (ASPA).

Os pesquisadores observaram que a frequência de tipo é bastante semelhante para todos os verbos, porém a frequência de ocorrência é consideravelmente maior para os verbos do presente do indicativo, sendo que o pretérito imperfeito apresenta o menor índice de frequência de ocorrência entre os três tempos. Para cada um dos tempos verbais analisados, foram selecionados 5 verbos de alta frequência de ocorrência e 5 de baixa. Inicialmente, pediu-se aos participantes que lessem as frases cuidadosamente em voz alta; em um segundo momento, solicitou-se a leitura das mesmas sentenças em uma velocidade de fala mais rápida, tendo em vista que a redução do ditongo é um fenômeno de fala coloquial e mais rápida (cf. BATTISTI, 2002). Após a coleta de dados, os estímulos foram editados acusticamente por meio do software Praat (BOERSMA e WEENINK, 2011). A análise estatística foi feita através do Programa R.

Os resultados da pesquisa indicam que a redução do ditongo não difere quantitativamente para os três tempos verbais. Segundo os autores, pode-se, pois, dizer que a redução desse ditongo expressa um fenômeno generalizado de redução de ditongos nasais postônicos, figurando como um mecanismo amplo de redução segmental vocálica em posição prosódica fraca ou não acentuada. Ainda assim, verbos com frequência de ocorrência alta apresentaram maiores índices de redução

²Transcrição fonética para o ditongo adotada pelos autores.

do que verbos com frequência de ocorrência baixa. A redução do ditongo nas terminações do presente do indicativo e do pretérito perfeito do indicativo ocorre de forma distinta para verbos de alta frequência e para verbos de baixa frequência. No entanto, para verbos no pretérito imperfeito do indicativo, a diferença entre verbos de alta e baixa frequência não foi significativa. Os autores sugerem, por esse resultado, um conflito entre a motivação fonética do fenômeno e o que entendem por nivelamento analógico.

A redução do ditongo ocorre de forma diferenciada nos tempos investigados: no presente e no pretérito imperfeito do indicativo, o ditongo é reduzido para uma vogal central *a*, enquanto no pretérito perfeito do indicativo, o ditongo é reduzido para uma vogal posterior *u*. De acordo com os pesquisadores, uma possível explicação para esse percurso diferenciado reside no fato de que a redução nos dois primeiros tempos opere para nivelar as formas verbais e por redução fonética; já no pretérito perfeito, a terminação *u* átona final é adotada no nivelamento analógico de regularização das formas verbais de 2ª e 3ª pessoas do plural em todas as conjugações verbais. Esse quadro sugeriria, na visão de Cristófaros-Silva, Fonseca e Cantoni, que a organização gramatical opera em redes interconectadas entre a semântica e a fonética, contribuindo para a simplificação do paradigma verbal do PB. A redução do ditongo [ãw] átono na morfologia verbal se contextualizaria na fonologia do PB através da redução segmental em posição prosodicamente fraca, afetando, em maiores índices, os verbos mais frequentes.

Percebemos que, no que diz respeito aos tempos verbais, esses resultados diferem dos de Schwindt, Bopp da Silva e Quadros (2012), pois o único tempo que se destacou na aplicação do processo no estudo destes foi o pretérito perfeito do indicativo. Os tipos de dados utilizados em ambas as pesquisas diferem, tendo em vista que o experimento de Cristófaros-Silva, Fonseca e Cantoni (2012) gera dados mais controlados que os de Schwindt, Bopp da Silva e Quadros, que apresentam dados vindos de entrevistas sociolinguísticas. No entanto, é interessante termos a oportunidade de averiguar, nesta tese, o papel da frequência lexical na aplicação do fenômeno em diferentes tempos verbais, tendo em vista que isso não foi realizado na pesquisa de Schwindt, Bopp da Silva e Quadros, para ver se os resultados obtidos em Cristófaros-Silva, Fonseca e Cantoni (2012) também se confirmam aqui, em dados pouco controlados, advindos de entrevistas sociolinguísticas.

2.3 Estudos sobre Apagamento de /r/³ final em PB

O apagamento da vibrante em coda está entre os processos considerados variáveis no português brasileiro (ex. lugar ~ luga∅; cirurgia ~ ciru∅gia; amar ~ ama∅; perder ~ pe∅der), conforme relataram Monaretto (1992,1997) e Callou, Leite e Moraes (1996), entre outros. Esse processo atinge verbos e não verbos não uniformemente: enquanto em verbos o apagamento tem aplicação quase categórica em todo o país, em não verbos o processo, além de menos preponderante de modo geral, é marcadamente mais comum nos dialetos que possuem também uma realização aspirada da vibrante (r > h > ∅), como o falado no Rio de Janeiro (cf. Callou, Serra e Cunha, 2015).

Schwindt (2015, 2016), numa tentativa de formalização do fenômeno em termos de arquitetura da gramática, sugere tratar-se de dois processos distintos: um mais específico, que tem acesso a informação de natureza morfológica; outro mais geral, que se aplica *across the board*, satisfazendo as exigências de uma variação tipicamente neogramática (não ter condicionamento morfológico e ser foneticamente gradual). Desse modo, verbos estariam sujeitos ao primeiro processo (sendo o expoente morfológico do segmento apagado realizado pelo acento da vogal temática, de acordo com Schwindt, 2014) e ao segundo, mas não-verbos só se submeteriam ao segundo, o que aumentaria consideravelmente as chances de verbos sofrerem o processo.

O fato é que alguns estudos mostram que o apagamento, especialmente em não-verbos, parece demonstrar predileção por algumas palavras (cf. Oliveira, 2012). Na perspectiva de regra variável, palavras com maior ou menor aplicação do processo não teriam, em princípio, papel especial na explicação do fenômeno, podendo ser justificadas por contexto ou mesmo por reestruturação de formas subjacentes.

Callou, Moraes e Leite (1998) apontam que a perda do /r/ é mais frequente em final de vocábulo e, no que diz respeito à classe de palavra, nos verbos: o infinitivo e a

³ Na ausência de evidências contrárias à tese de Câmara Jr. (2008[1953]) quanto à representação básica da classe dos róticos em português, escolhemos representar a vibrante ao longo do trabalho por meio do arquifonema /r/, no entendimento de que essa é a representação de uma classe de sons. Essa escolha, contudo, não tem consequências cruciais para o fenômeno que investigamos, já que está em foco neste trabalho o apagamento de um segmento realizado mais do que uma pretensa estrutura subjacente.

primeira e terceira pessoas do futuro do subjuntivo são marcados em português tanto pela presença do /r/ final quanto pela tonicidade da sílaba que contém o segmento. Nos não verbos, em que o /r/ não carrega informação morfológica, o peso relativo é baixo. Para os nomes, o tamanho do vocábulo é um fator significativo, a perda do /r/ sendo praticamente bloqueada em vocábulos monossilábicos. Já para os verbos, a variável tamanho do vocábulo tem um comportamento neutro. Os autores apontam que o apagamento do /r/ final tem sido considerado um caso de mudança de baixo para cima que já atingiu seu limite, sendo uma variação estável, sem marca de classe social. Os autores também defendem a necessidade de se analisar verbos e não verbos em separado, tendo em vista que a seleção dos grupos de fatores significativos pode não refletir corretamente os ambientes que condicionam o apagamento se a análise for conjunta.

Oliveira (1997) analisa o apagamento da vibrante no final de nominais com os seguintes fatores: estilo, tonicidade da sílaba, vogal precedente, classe de palavra, contexto seguinte, faixa etária, sexo, classe social e informante. Tonicidade da sílaba final, vogal precedente e classe de palavra não foram selecionados como significativos em seu estudo. O pesquisador defende que, não obstante a necessidade de um maior número de dados para desenvolver uma análise mais satisfatória, não se pode sustentar uma explicação neogramática, tendo em vista que os nominais que terminam em vibrante não tiveram o mesmo destino em termos da aplicação da regra de cancelamento de /r/, fato esse que se enquadraria melhor em um modelo difusionista. A análise do autor também sugere que não são necessariamente as palavras mais frequentes as mais propensas ao apagamento, ou seja, frequência lexical não é determinante para a aplicação do processo (como já apontado em Oliveira, 1991), e que o indivíduo teria seu papel próprio na variação, o qual independe do comportamento do grupo social ao qual pertence.

Nos estudos de Monaretto (1997, 2002), no que diz respeito ao apagamento da vibrante no português do sul do Brasil, fatores como posição, classe morfológica e idade foram selecionados como significativos. Sobre o primeiro fator, a autora aponta que a aplicação da regra ocorre basicamente em final de palavra; no que tange à classe morfológica, a perda da vibrante acontece mais em verbos, tendo em vista a redundância da marcação do infinitivo e da primeira e terceira pessoas do subjuntivo tanto pela presença da vibrante final quanto pela tonicidade; no que diz respeito à

idade, a pesquisadora defende a possibilidade de uma mudança em progresso, tendo em vista que informantes mais jovens apagam mais do que os mais velhos. Ao analisar o apagamento exclusivamente em verbos, fatores selecionados como significativos em seu estudo foram sexo, dimensão da palavra e contexto precedente. O resultado para sexo apontou a mulher como mais propensa à realização do apagamento; sobre o fator dimensão da palavra, o resultado, contrariando a expectativa inicial da pesquisadora, aponta que palavras com até duas sílabas favorecem mais o pagamento que palavras com três ou mais sílabas; sobre contexto precedente, a preferência do apagamento se dá com vogais posteriores, o que, para a autora, é justificado tendo em vista o paralelo com as vibrantes – vibrantes posteriores, diferentemente das anteriores, tendem a sofrer mais apagamento.

Brescancini e Monaretto (2008) apresentam comparações e generalizações de resultados de pesquisas realizadas sobre os róticos com dados de fala do Projeto VARSUL. As autoras apontam que, de forma geral, a classe morfológica dos verbos e a vibrante com função morfêmica, como em *escolar*, apresentaram-se como favoráveis ao apagamento. No que diz respeito ao contexto precedente, as autoras concluem que os resultados de Monaretto (2000) e Monguilhot (1997) são concordantes, tendo em vista que as vogais anteriores são as mais favorecedoras no primeiro estudo e, no segundo, a vogal [E]. O resultado de Gregis (2001) aponta a preferência pelo apagamento após vogais [-arredondado]. No que tange à influência dos anos de escolarização, o apagamento é favorecido em informantes com baixa escolaridade segundo Monguilhot (1997) e Monaretto (2000). Os mais jovens são os maiores produtores de apagamento em Monaretto (2000) e Pimentel (2002).

Huback (2003, 2006) analisa o apagamento do /r/ final em nominais em entrevistas de 30 falantes de Belo Horizonte. Nesse estudo, constatou-se que o /R/ final em nominais foi apagado em 22% dos casos. Os fatores analisados foram: classe de palavra, número de sílabas, acento, estrutura morfológica, segmento seguinte, vogal precedente, gênero, idade, classe social e nível de escolaridade. Os fatores selecionados como determinantes para o fenômeno foram: ambiente seguinte (consoante, vogal, pausa), vogal precedente, número de sílabas (monossilábicos, polissilábicos), estrutura morfológica (-dor agentivo, -dor não agentivo e -tor, outros sufixos, não-sufixo, conectivos) e tipo de informante. No que diz respeito ao ambiente seguinte, a vibrante tem taxas mais altas de apagamento quando está diante de

consoante. Ao analisar o efeito da vogal precedente, a aplicação do fenômeno parece ser favorecida diante da vogal /i/; no entanto, por meio de análise mais minuciosa, a autora aponta que esse resultado dizia respeito a apenas uma expressão, qual seja 'a partir'. Sobre o número de sílabas, itens polissilábicos se mostraram mais suscetíveis ao apagamento. No que tange à estrutura morfológica, sufixos -dor, agentivo ou não, e -tor se mostraram favorecedores do apagamento, juntamente com os conectivos (qualquer, apesar, a partir). Sobre o indivíduo, a autora afirma que embora exista um certo padrão recorrente dentro de um mesmo grupo social, indivíduos diferentes apresentam comportamentos linguísticos diferentes.

Ao comparar os itens lexicais mais favoráveis ao apagamento e suas taxas de frequência em um corpus de referência, Huback (2006) aponta que apesar de não haver uma perfeita correlação entre frequência do item léxico e apagamento da vibrante, todas as palavras líderes em cancelamento são altamente frequentes no corpus analisado. A autora finaliza sua argumentação chamando a atenção para dois fatos relativos aos seus resultados: 1) itens frequentes como 'a partir' e 'qualquer' estão liderando a mudança e 2) palavras com o sufixo -dor favorecem o cancelamento da vibrante. Segundo a autora (p. 25), estes fatos confirmam a teoria proposta por Bybee (1995, 2001), no que tange à frequência de *token* e à frequência de *type*, respectivamente. Pertinentemente, no que diz respeito a esses dois tipos de frequência, Huback (2013, p. 86) analisa o fato de o apagamento do /r/ final ser mais frequente em verbos que em formas nominais:

Todos os infinitivos verbais do PB terminam em (r), então a frequência do (R) final em verbos é possivelmente maior que a do (r) final em nomes. Além disso, certos verbos, como *ser, estar, ter, fazer*, etc., têm, também, alta frequência de ocorrência. Possivelmente, uma interação entre a alta frequência de tipo e de ocorrência faz com que o (r) final seja mais apagado em verbos que em nomes.

A pesquisa de Silveira (2010) sobre o apagamento da vibrante final em nominais apresentou como fatores linguísticos significativos: contexto precedente, classe morfológica, contexto seguinte. Na análise desses fatores, a autora aponta a possível influência de um desequilíbrio de dados na amostra, defendendo que esse fenômeno apresenta evidências que sustentam a hipótese de que o apagamento ocorre em determinadas palavras, ou seja, pode estar sendo implementado lexicalmente na

língua, não sendo produto de regras. Ao estabelecer uma análise comparativa entre os dados de sua pesquisa e os de Oliveira (1997), Huback (2003) e Monaretto (2000; 2002), a pesquisadora aponta que o apagamento do /r/ ocorre, na maioria das vezes, nas mesmas palavras em todos os corpora.

DADOS	HUBACK	OLIVEIRA	MONARETTO	SILVEIRA
1. lugar	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
2. mulher	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
3. qualquer	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
4. melhor	<i>X</i>	<i>X</i>	<i>X</i>	
5. maior	<i>X</i>	<i>X</i>	<i>X</i>	
6. pior	<i>X</i>	<i>X</i>		
7. menor	<i>X</i>	<i>X</i>		
8. professor	<i>X</i>	<i>X</i>		
9. amor	<i>X</i>		<i>X</i>	
10. bar	<i>X</i>		<i>X</i>	
11. interior	<i>X</i>		<i>X</i>	
12. apesar	<i>X</i>		<i>X</i>	
13. vestibular	<i>X</i>			<i>X</i>

Quadro 1: Análise comparativa (SILVEIRA, 2010, p. 82)

Silveira (2010) aponta que embora essas palavras tenham sido extraídas de corpora diferentes, em diferentes cidades e com informantes diversos, os itens que sofrem apagamento são os mesmos, com taxas de aplicação bastante similares. Além disso, o fato de parecer não haver contexto fonológico que justifique o apagamento final em não verbos sugere que a hipótese da difusão lexical é mais apropriada para explicar o fenômeno, segundo a autora.

Em artigo de 2013, Serra e Callou apresentam dados de apagamento de /r/ em coda silábica final em três cidades brasileiras, quais sejam Salvador, Rio de Janeiro e Porto Alegre. Segundo as autoras (p. 587), a primeira realiza o /r/ como fricativa laríngea, apresentando índices significativos de apagamento; no Rio, a norma de pronúncia é a fricativa posterior, local intermediário de variação na aplicação do

processo; já em Porto Alegre, cuja realização da vibrante é anterior, o cancelamento apresenta índices mais baixos.

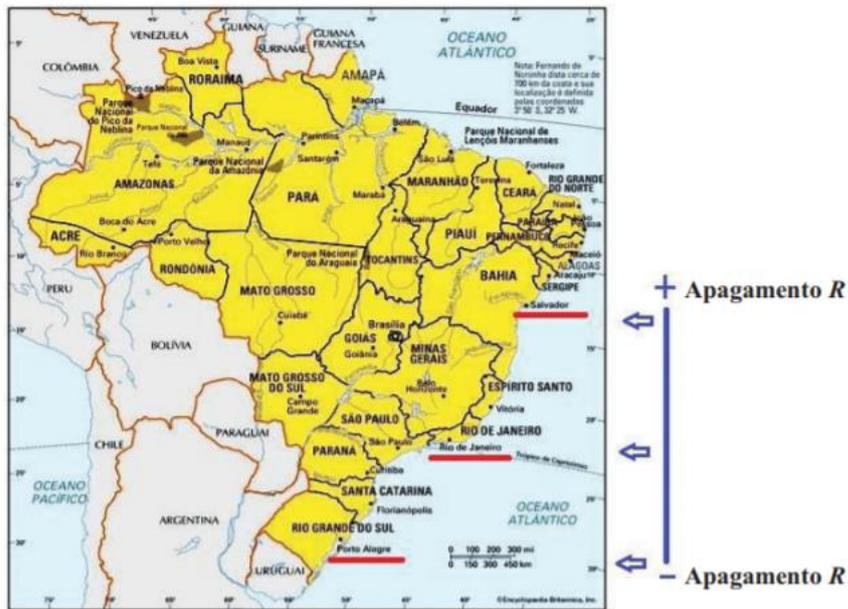


Figura 1: Regiões estudadas e processo de apagamento do /r/ (SERRA e CALLOU, 2013, p. 587)

As autoras apontam para a escalaridade geográfica na realização do /r/ e defendem que a retenção do segmento acontece nos falares que conservam a articulação [+anterior] da vibrante. Em suas conclusões, Serra e Callou (2013) postulam três regras de caráter regional para o fenômeno:

Tipo de regra	Caraterísticas
<ul style="list-style-type: none"> Regra categórica de apagamento sensível à classe morfológica 	Há variedades que aplicam a regra de apagamento só em verbos, e, coincidentemente, têm como norma de pronúncia, em coda silábica interna, uma vibrante anterior (Porto Alegre – Sul).
<ul style="list-style-type: none"> Regra variável de apagamento 	Há variedades em que a regra é variável em verbos e não-verbos -- sensível ao tipo de fronteira prosódica -- em variedades em que a norma é uma fricativa posterior, velar (Rio de Janeiro - Sudeste).
<ul style="list-style-type: none"> Regra categórica de apagamento 	Há variedades em que a regra atua, independentemente de classe morfológica e tipo de fronteira prosódica, nas quais a norma de pronúncia é uma fricativa posterior, laringea – aspiração (Salvador - Nordeste).

Quadro 2: Regras de caráter regional (SERRA e CALLOU, 2013, p. 592)

As autoras apontam que o fenômeno ainda carece de análise mais aprofundada, pois uma pergunta pertinente seria: cada variante do /r/ representa um passo na escala ordenada de enfraquecimento ou o apagamento corresponde a um único processo, sem etapas intermediárias? Ainda é necessário, segundo as pesquisadoras, explorar evidências de condicionamentos fonológicos, morfológicos, sociais e prosódicos, além de verificar se o apagamento em coda final poderia representar uma regra lexical ou pós-lexical, dependendo da classe morfológica do item.

Em artigo de 2015, Callou, Serra e Cunha analisam amostras de fala culta do Projeto ALiB de falantes de nove capitais do Nordeste, quais sejam São Luís, Teresina, Fortaleza, Natal, João Pessoa, Recife, Maceió, Aracaju e Salvador. No que diz respeito aos não verbos, as capitais não têm comportamentos tão uniformes como nos verbos: em João Pessoa, Salvador, Maceió e Fortaleza, o apagamento em não verbos é quase categórico como nos verbos; em Natal, Recife e São Luís, a frequência de apagamento é um pouco menor nos não verbos; já em Aracaju e Teresina, o processo de apagamento nos não verbos está apenas a meio-termo. Para essas últimas, a dimensão do vocábulo se mostrou um condicionador importante: o apagamento em não verbos é menos frequente em monossílabos. No que diz respeito

às variáveis sociais, os mais jovens parecem ser propulsores da mudança. Gênero foi um fator selecionado em Maceió, Recife e Fortaleza, cidades onde as mulheres lideram a aplicação do fenômeno.

Na análise dos verbos, as autoras (p. 204) apontam que a regra de apagamento se configura como semi-categórica, sem apresentar, praticamente, restrições de natureza estrutural e/ou social. Para Aracaju, capital que apresenta percentual mais baixo de apagamento, a única variável selecionada foi contexto subsequente (vogal, consoante, pausa), em que a presença da pausa propicia menos o cancelamento do rótico, que, segundo as autoras, pode estar relacionado à fronteira prosódica.

Em estudos de aquisição de /r/ em coda, Menezes e Gomes (2012) demonstraram que as crianças não incorporam um padrão fonológico invariável (CVC), mas incorporam distribuições de tokens como um reflexo da dinâmica sociolinguística da comunidade de fala a que pertencem. As autoras apontam que o comportamento das crianças no que diz respeito às codas finais de verbos pode ser visto como o final de um processo de enfraquecimento de /r/, a conclusão de uma mudança, tendo em vista que no input recebido a ausência da coda é praticamente categórica.

Gomes (2017), estudou o papel do item lexical na direcionalidade da mudança sonora em grupos sociais distintos no Rio de Janeiro – falantes de classe média, adolescentes moradores de favela com algum tipo de inserção em programa social e adolescentes moradores de favela sem inserção em programa social. Seus resultados apontam a importância tanto do condicionamento fonético quanto do item lexical na variação. No que diz respeito ao comportamento dos falantes, os dois primeiros grupos apresentaram resultados semelhantes, o que indica que o grau de inserção social é relevante para o comportamento dos falantes.

Ao estudar padrões de covariação entre seis variáveis, dentre elas apagamento de /r/ e concordância verbal na terceira pessoa do plural, Oushiro (2016) identificou restrições internas e externas subjacentes à simultaneidade do uso de diferentes variáveis. A autora defende que a coesão dialetal é uma função direta da densidade da comunicação entre falantes que apresentam maior interação social. Oushiro aponta que a covariação é promovida por restrições linguísticas mais gerais que se

correlacionam com múltiplas variáveis, tais como saliência fônica, contexto precedente e seguinte, frequência e classe de palavra.

RESUMO DO CAPÍTULO: Apresentamos a revisão da literatura no que diz respeito a estudos já realizados em PB que, de certa forma, consideram frequência lexical em suas análises. Após, fizemos uma breve revisão da literatura que aborda os dois fenômenos analisados nesta tese: redução da nasalidade e apagamento de /r/.

3 METODOLOGIA DE TRABALHO

Nesta seção apresentamos nossa metodologia de trabalho no que diz respeito à análise de dois fenômenos fonológicos variáveis, quais sejam Redução da Nasalidade e Apagamento de /r/. Apresentamos a composição dos nossos *corpora*, as perguntas norteadoras e as ferramentas estatísticas para os procedimentos de análise.

No que diz respeito à análise de frequência, variável crucial para esta pesquisa, trabalhamos com o *corpus* de referência do projeto ASPA (Avaliação Sonora do Português Atual), que apresenta 607.392 types de palavras, com 228.766.402 ocorrências, dados esses provindos de *corpus* escrito (CRISTÓFARO-SILVA, ALMEIDA, FRAGA, 2005). Trabalhar com um *corpus* significativamente maior que os nossos para obter informação de índices de frequência é importante, tendo em vista que o ASPA tem uma maior representatividade no que diz respeito aos dados do PB, considerando a abrangência das fontes que o alimentam.⁴

Passemos agora aos procedimentos metodológicos no que diz respeito a cada um dos fenômenos estudados.

3.1 Redução da Nasalidade

No português brasileiro (PB), ditongos nasais em sílabas finais átonas, diferentemente se em sílabas tônicas, estão sujeitos à variação, pois se realizam com ou sem nasalidade (ex. falaram ~ falaru; homem ~ homi). A realização sem nasalidade implica, também, a redução do ditongo.

3.1.1 Sobre o *corpus*

Reanalisamos os dados de verbos e não verbos com os ditongos finais /eN/ e /aN/ da amostra estudada por Schwindt e Bopp da Silva (2010). A pesquisa dos autores,

⁴ O acesso ao Projeto ASPA se dá em www.projetoaspa.org

com o intuito de traçar um panorama da redução da nasalidade na Região Sul do país, considerou todas as cidades que compõem o Projeto VARSUL⁵.

Os autores analisaram, como aplicação do processo, apenas os casos em que a nasalidade se perde junto com a redução do ditongo. De acordo com resultados de estudos precedentes (tais como o de Battisti (2002) e o de Bopp da Silva (2005)), as variáveis linguísticas consideradas por eles foram as seguintes:

Grupos de fatores	Exemplos
<i>Classe de palavra</i>	
Nomes com <i>gem</i> na raiz	<i>homenagem</i>
Nomes com sufixo <i>-gem</i>	<i>reciclagem</i>
Nomes	<i>homem</i>
Verbos	<i>cantaram</i>
<i>Consoante do onset</i>	
Consoante nasal	<i>am<u>a</u>m</i>
Consoante não-nasal posterior	<i>fi<u>ç</u>am</i>
Consoante não-nasal anterior	<i>estud<u>a</u>m</i>
Onset vazio	<i>sa<u>e</u>m</i>
<i>Tonicidade do contexto seguinte</i>	
Átona	<i>vieram <u>m</u>orar</i>
Tônica	<i>sabiam <u>d</u>isso</i>
<i>Contexto fonológico seguinte</i>	
Vogal	<i>falam <u>a</u>lemão</i>
Consoante não-nasal	<i>falam <u>d</u>iferente</i>
Pausa	<i>falam#</i>
Consoante nasal	<i>falam <u>m</u>ais</i>

Quadro 3: RN e variáveis linguísticas (SCHWINDT e BOPP DA SILVA, 2010, p.19)

E as variáveis extralinguísticas:

⁵ Projeto Variação Linguística na Região Sul do Brasil: <http://www.varsul.org.br>.

<i>Localização geográfica</i> Porto Alegre, São Borja, Panambi, Flores da Cunha (RS) Florianópolis, Blumenau, Chapecó, Lages (SC) Curitiba, Londrina, Pato Branco, Irati (PR)
<i>Idade</i> De 25 a 50 anos 51 anos ou mais
<i>Escolaridade</i> 0-4 anos (primário) 9-12 anos (segundo grau)

Quadro 4: RN e variáveis sociais (SCHWINDT e BOPP DA SILVA, 2010, p. 18)

Os autores selecionaram 3 entrevistas para preencher cada célula social, que resultou em um total de 144 informantes, sendo 12 de cada uma das 12 cidades⁶ estudadas (cf. SCHWINDT e BOPP DA SILVA, 2010, p.18).

Tendo em vista que nossa análise toma como base os arquivos de dados dos autores, achamos conveniente apresentar as variáveis, tanto linguísticas quanto sociais, com as quais eles trabalharam. No entanto, nas análises deste trabalho, quando necessário e possível, algumas modificações nos grupos de fatores foram feitas. No que diz respeito, portanto, às variáveis e aos grupos de fatores analisados nesta tese, apresentamo-los em nosso capítulo de Resultados e Discussão.

O trabalho aqui apresentado retoma os dados de Schwindt & Bopp da Silva (2010) e adiciona a variável *frequência lexical*, extraída do projeto ASPA. Importante ressaltar que foi necessário excluir dados da amostra dos autores que não apresentavam claramente o item lexical junto da codificação no arquivo disponibilizado, tendo em vista a impossibilidade de atribuímos a informação de frequência lexical.

A amostra deste trabalho, então, consiste de 1,515 tokens para /eN/ e 3,817 tokens para /aN/. Tendo em vista que havia apenas um não verbo com ditongo final /aN/, restringimos nossa análise dessa classe de palavra para /eN/. Para os verbos, no entanto, os dois ditongos são analisados.

⁶ Neste trabalho, não analisamos a variável Localização Geográfica.

3.1.2 Perguntas norteadoras e procedimentos para a análise

Pretendemos problematizar a hipótese de redução na nasalidade enquanto regra variável, considerando os efeitos de frequência lexical com foco na sua relação com condicionamento morfológico. Tendo em vista nossos três objetivos gerais apresentados na introdução desta tese, lançamos as seguintes perguntas para este fenômeno:

Pode-se falar em efeitos de frequência lexical nos dados do sul do Brasil? Se sim:

- 1) palavras mais frequentes estão mais sujeitas ao processo?
- 2) palavras menos frequentes estão menos sujeitas ao processo?

Sub-questões:

- a) esses efeitos são diferentes para não verbos e para verbos?
- b) qual a relação entre frequência lexical e fatores linguísticos, tais como contexto precedente e morfologia interna à palavra?
- c) qual a relação entre frequência lexical e fatores sociais, tais como idade e escolaridade?
- d) nos nomes, qual a relação entre frequência lexical e terminação em -gem e outras?
- e) nos verbos, qual a relação entre frequência lexical e cada tempo verbal?

A análise foi feita por meio de regressão logística, em dois programas estatísticos, quais sejam o GoldVarb X (Sankoff, Tagliamonte e Smith, 2005) e a Plataforma R (R Core Team, 2016).

3.2 Apagamento de /r/ final

Tendo em vista que uma realização aspirada da vibrante ($r > h > \emptyset$), como a falada no Rio de Janeiro (cf. Callou, Serra e Cunha, 2015), tende a ser mais suscetível ao

apagamento, nossos dados são relativos a esse variedade. No que diz respeito à variável dependente, em um primeiro momento analisamos quatro variantes: tepe, velar, glotal e apagamento. Na análise dos dados, para evitar problemas de convergência, juntamos as primeiras duas variantes na categoria de ‘realização do /r/’ e as duas últimas na categoria de ‘apagamento’.

Na análise de dados, consideramos apenas palavras oxítonas e monossílabos tônicos.

3.2.1 Sobre o corpus

Nosso corpus é composto por 16 entrevistas de informantes do Rio de Janeiro. Os dados são relativos a duas amostras: a 1ª, que engloba entrevistas do Programa da Marília Gabriela, e a 2ª, que é relativa a entrevistas de depoimentos pessoais dos entrevistados sobre o *rap* carioca e sobre suas experiências na época da ditadura.⁷ Os critérios para a seleção dos entrevistados foram os seguintes:

- Ter nascido na cidade do Rio de Janeiro
- Morar no RJ
- Fricativização do R na coda
- Palatalização do S na coda

Nesta seção, apresentaremos os informantes considerados na análise das duas amostras.⁸

Amostra 1

No que diz respeito à primeira amostra, temos 8 entrevistas, cujo tempo médio de duração é de 45 minutos. A seguir, uma tabela com os informantes e suas

⁷ Maiores detalhamentos sobre a segunda amostra está no Anexo I desta tese. No que diz respeito a esse anexo e ao levantamento dos dados de apagamento de /r/, contamos com a incansável colaboração da estudante Júlia Ricardo, bolsista de Iniciação Científica do Projeto “Exponência morfológica na fonologia do português brasileiro”, coordenado por Luiz Carlos Schwindt e do qual participo como membro de pesquisa.

⁸ Em um primeiro momento da pesquisa, tínhamos a intenção de trabalhar com diferenças de formalidade e, por isso, temos duas amostras diferentes. No entanto, após análise das entrevistas, percebemos a impossibilidade e a incoerência de tentar contrastar formalidade nas duas amostras. Neste capítulo apresentamos a constituição das duas amostras mas, em nossa análise de dados, ambas são consideradas conjuntamente.

respectivas idades no momento da entrevista – que serão analisadas de forma contínua pela Plataforma R.

Informante	Idade
Joel Santana	67
Lulu Santos	63
Naldo Benny	35
Dudu Nobre	40
Christina Rocha	55
Sandra de Sá	59
Kelly Key	31
Tatá Werneck	33

Quadro 5: Informantes e idade – Amostra 1

Amostra 2

No que diz respeito à segunda amostra, temos 8 entrevistas, cujo tempo médio de duração é de 26 minutos. A seguir, uma tabela com os informantes e suas respectivas idades no momento da entrevista – que serão analisadas de forma contínua pela Plataforma R.

Informante	Idade
Rôssi Alves	42
Edd Wheeler	45
Maurício	25
DJ Tamenpi	35
Vera Sílvia Magalhães	55
Beatriz Ryff	91
Mauro César Pimentel	51
Célio Borja	75

Quadro 6: Informantes e idade – Amostra 2

A apresentação das variáveis linguísticas e sociais com as quais trabalhamos, e de seus respectivos grupos de fatores, está no capítulo Resultados e Discussão.

O trabalho aqui apresentado analisa *frequência lexical* com índices extraídos do *corpus* de referência ASPA.

A amostra conta com 2.610 *tokens*, sendo 2.129 verbos e 481 não verbos.

3.2.1 Perguntas norteadoras e procedimentos para a análise

No que diz respeito ao apagamento de /r/, pretendemos problematizar a hipótese de apagamento da vibrante enquanto regra variável, explorando o papel do léxico nesse processo. Para isso, tendo em vista nossos três objetivos gerais apresentados na introdução desta tese, lançamos as seguintes perguntas para este fenômeno:

Pode-se falar em efeitos de frequência lexical nos dados do sul do Brasil? Se sim:

- 1) palavras mais frequentes estão mais sujeitas ao processo?
- 2) palavras menos frequentes estão menos sujeitas ao processo?

Sub-questões:

- a) esses efeitos são diferentes para não verbos e para verbos?
- b) qual a relação entre frequência lexical e fatores linguísticos, tais como contexto precedente e morfologia interna à palavra?
- c) qual a relação entre frequência lexical e fatores sociais, tais como idade e sexo?

A análise foi feita, em um primeiro momento, por coeficiente de correlação e, após, por regressão logística, na Plataforma R (R Core Team, 2016).

4 RESULTADOS E DISCUSSÃO

Nesta seção, apresentamos nossos resultados para os dois fenômenos analisados nesta tese: redução da nasalidade e apagamento de /r/.

4.1 Redução da Nasalidade

No que diz respeito a esse fenômeno, apresentamos o resultado por meio de duas seções: a primeira apresenta resultados advindos de regressão logística feita pelo programa estatístico GoldVarb X, cuja análise apresentada aqui toma como base o artigo de Schwindt e De Bona (2017); a segunda seção apresenta resultados de regressão logística por meio da plataforma R. O uso desses dois programas reflete nosso percurso de pesquisa: em um primeiro momento nossa análise foi feita com o GoldVarb X, tratando a variável *frequência* como uma variável nominal; com os resultados dessa primeira análise, procedemos à análise na plataforma R, em uma regressão logística que nos possibilitou tratar *frequência* como uma variável contínua.

4.1.1 Resultados do GoldVarb X

Como vários outros fenômenos variáveis no PB caracterizados como uma mudança abaixo do nível de consciência (Labov, 1966), RN tem uma aplicação moderada no sul do Brasil – 34% em nossa amostra.

De acordo com diferentes estudos, classe de palavra tem um papel crucial no processo, distinguindo verbos de não verbos.

Classe de palavra	%	N
Não verbo <i>ontem</i>	56	426/826
Verbo <i>pedem</i>	32	2678/8487

Tabela 1: Classe de palavra e RN no sul do Brasil
(adaptado de SCHWINDT e BOPP DA SILVA, 2010)

A seguir, trazemos nossa análise do papel da frequência lexical, com foco especialmente no papel da classe de palavra. Essa análise, como já mencionamos, usa informações de frequência do Projeto ASPA, considerado aqui como *corpus* de referência. Das palavras do nosso corpus, 98% foram encontradas no ASPA; além disso, entre não verbos e verbos, 8 das 20 palavras mais frequentes no ASPA também estão entre as 20 mais frequentes no nosso corpus. Em se tratando de apenas não verbos, este número sobe para 12; entre verbos, cai para 8. Uma comparação mais detalhada entre os *corpora* analisados nesta tese e o ASPA é feita na seção 4.3.

Classe de Palavra	Palavras Comuns
não verbos e verbos	<i>foram, homem, eram, estavam, viagem, querem, dizem, fizeram</i>
não verbos	<i>reportagem, homem, ordem, viagem, origem, jovem, vantagem, passagem, homenagem, margem, mensagem, coragem</i>
Verbos	<i>foram, eram, estavam, querem, dizem, fizeram, tinham, ficam</i>

Quadro 7: Palavras comuns sujeitas a RN no ASPA/VARSUL entre as 20 mais frequentes

Ao considerarmos a análise de toda a amostra em três níveis de frequência⁹, não há diferenças significativas entre os níveis, como podemos ver na Tabela 2.

⁹ Tendo em vista que o GoldVarb X não opera com variáveis contínuas, para a análise de frequência construímos uma escala logarítmica de três níveis (alta, média e baixa frequência) a partir do *corpus* de referência ASPA e codificamos os dados da nossa amostra com esses índices, submetendo a rodadas com diferentes cruzamentos, de modo a obter respostas para as questões apresentadas em nossa metodologia. O ideal, no que diz respeito ao tratamento da frequência lexical, seria atribuir a cada item sua respectiva frequência, tendo em vista que é difícil estabelecermos limites precisos para cada faixa de frequência, e esses limites, quando impostos arbitrariamente, podem obscurecer a análise do fenômeno.

Nossos níveis de frequência foram construídos da seguinte maneira, de acordo com Chaves (2017, p. 230): a subtração entre o logaritmo do item mais frequente e o logaritmo do item menos frequente resultou em uma escala. Para construir os níveis, somamos a escala obtida ao logaritmo do item menos frequente e, então, obtivemos o nível de baixa frequência. O limite do nível de baixa frequência mais a escala resultou no nível de média frequência e, então, o limite do nível de média frequência mais a escala resultou no nível de alta frequência.

Frequência	%	N
Alta <i>viagem, estavam</i>	32,5	719/2214
Média <i>garagem, mataram</i>	31,9	816/2554
Baixa <i>nuvem, ajoelharam</i>	29,3	165/564

Tabela 2: Frequência Lexical e RN – verbos e não verbos

No caso de não verbos, no entanto, há uma prevalência de RN no contexto de palavras de alta frequência.

Frequência	%	N
Alta <i>homem, origem</i>	64,3	182/283
Média <i>ordem, chantagem</i>	50,0	65/130
Baixa <i>nuvem, ladroagem</i>	43,8	14/32

Tabela 3: Frequência Lexical e RN – não verbos

O mesmo padrão não é observado nos verbos, em que a RN apresenta um papel relativamente balanceado entre os três níveis de frequência, com alguma proeminência no nível de média frequência.

Frequência	%	N
Alta <i>eram, dizem</i>	27,8	537/1931
Média <i>pensam, bebem</i>	31,0	751/2424
Baixa <i>arrumaram, tecem</i>	28,4	151/532

Tabela 4: Frequência Lexical e RN – verbos

Com isso, podemos concluir que o papel da frequência lexical na RN está limitado a determinadas classes de palavras. Como esperávamos, a análise sugere que não verbos são mais sensíveis à frequência que verbos. No entanto, subgrupos dessas classes devem ainda ser analisados separadamente com o intuito de verificar se frequência lexical exerce algum papel dentro dessas classes. Esse é o caso de 1) nomes terminados pelo sufixo *-gem*; 2) nomes em que *gem* é parte da raiz e 3) outros não verbos sem *gem*. No que diz respeito aos verbos, diferentes tempos podem ser examinados separadamente.

Como pode ser visto na Tabela 5, frequência lexical não parece ter papel determinante nos não verbos sufixados, sugerindo que o processo seja, em princípio, cego à estrutura interna dessa classe de palavra.

Contexto Morfológico	Alta	Média	Baixa
-gem como sufixo <i>ladroagem, pilotagem</i>	61%	72%	65%
gem como parte da raiz <i>viagem, origem</i>	58%	65%	--
Outras palavras não sufixadas <i>homem, nuvem</i>	73%	21%	8%

Tabela 5: Frequência lexical e RN – não verbos

Guy (1981) sugere que o sufixo *-gem* apresenta um comportamento irregular na sua evolução do latim, resultando em palavras com e sem o ditongo nasal em português. No que tange à sincronia, como já apontado, percebemos que a aplicação do fenômeno não enxerga condicionamento morfológico, ou seja, não diferencia se *gem* é parte da raiz ou do sufixo. Não temos como sustentar, neste caso, a atuação de uma regra morfológica na aplicação do fenômeno.

A questão que se mantém é por que razão, apesar de a RN ser cega à estrutura interna dos não verbos, as palavras terminadas em *gem* são mais sujeitas à aplicação do fenômeno. De acordo com Schwindt (2012), isso se deve provavelmente ao contexto fonético que precede a nasal do ditongo, particularmente por causa do efeito das consoantes palatais, como já apontado por Guy (1981). No entanto, precisa-se verificar se essa alta taxa de aplicação do fenômeno se mantém também em verbos

precedidos por palatal para poder sustentar a atuação da regra fonológica nesse caso. Caso essa alta taxa de aplicação não se mantenha, pode-se ter um argumento a favor das abordagens exemplaristas, considerando não o efeito da frequência do item lexical, mas o efeito da frequência do segmento *gem*.

Interessantemente, frequência lexical é relevante para RN apenas em contextos de não verbos sem terminação em *gem*. Isso é apresentado nas Tabelas 6 e 7 a seguir.

Frequência	%	N
Alta <i>ontem, homem</i>	72,8	83/114
Média <i>ordem, jovem</i>	20,8	11/53
Baixa <i>nuvem</i>	8,3	1/12

Tabela 6: Frequência lexical e RN – contexto precedente em não verbos (não palatal)

Frequência	%	N
Alta <i>origem, linguagem</i>	58,6	99/169
Média <i>chantagem, garagem</i>	70,1	54/77
Baixa <i>ladroagem, serragem</i>	65,0	13/20

Tabela 7: Frequência lexical e RN – palatal em contexto precedente em não verbos

Entre os não verbos, apenas identificamos papel da frequência lexical em não verbos que não apresentam a terminação *gem*. Em nomes com essa terminação, como já sugerido, parece que é a atuação do *type gem* que condiciona a alta aplicação do fenômeno.

No que diz respeito aos diferentes tempos verbais, não parece haver um papel notável da frequência lexical na aplicação da RN. No passado do indicativo, o único

tempo que favorece o processo de acordo com Schwindt, Bopp da Silva, e Quadros (2012), as taxas de frequência são distribuídas de forma bastante semelhante nos três níveis. Com taxas mais baixas de aplicação da RN, o mesmo padrão é encontrado em outros tempos, se os analisarmos conjuntamente.

Tempo verbal	Alto	Médio	Baixo
Pretérito Perfeito do Indicativo <i>foram</i>	37%	43%	45%
Outros tempos <i>comem</i>	24%	26%	24%

Tabela 8: Frequência lexical e RN – tempos verbais

Os resultados apresentados na Tabela 8 diferem em alguma medida dos encontrados em Cristófaró Silva, Fonseca e Cantoni (2012), que encontraram papel da frequência lexical no presente do indicativo e no pretérito perfeito do indicativo.

Como mencionado, Schwindt (2012) defende que consoantes palatais em contexto que precede ditongos nasais favorecem RN. No entanto, isso não parece proceder para itens verbais. Além disso, a aplicação do fenômeno não parece estar relacionada à frequência lexical, como mostra a Tabela 9 a seguir.

Contexto precedente	Alta	Média	Baixa
Palatal <i>beijam, tinham</i>	25%	31%	14%
Outros contextos precedentes <i>chamam</i>	28%	31%	29%

Tabela 9: Frequência lexical e RN – contexto precedente em verbos

Podemos perceber que, no que tange aos dados verbais, contexto precedente palatal apresenta uma aplicação do fenômeno relativamente baixa (23,3%, em média) se compararmos com contexto palatal em nomes (64,2%, em média). Com isso em vista, não é possível sustentar a hipótese de atuação de uma regra fonológica que condicione o apagamento. Parece que temos evidências, portanto, a favor das

abordagens exemplaristas, considerando, como já argumentado, não o efeito da frequência do item lexical, mas o efeito da frequência do segmento *gem*.

No que diz respeito aos fatores sociais, Schwindt e Bopp da Silva (2010) mostraram que RN é condicionado por idade, com uma pequena preferência de aplicação do fenômeno pelos informantes jovens e com baixa escolaridade.

No domínio dos verbos, nenhuma correlação entre fatores sociais e frequência lexical foi observada. No que diz respeito aos não verbos, por outro lado, a análise conjunta de idade e escolaridade em contraste com frequência lexical nos leva a duas considerações: (i) informantes mais velhos e de baixa escolaridade apresentam taxas reduzidas de RN nas palavras de baixa frequência, apontando para a preservação do ditongo em palavras menos frequentes na língua; (ii) informantes mais jovens e de mais alta escolaridade parecem ignorar a informação de frequência no que diz respeito a RN.

Esses resultados já nos apontam para alguns direcionamentos de análise: efeitos de frequência lexical parecem atuar mais em não verbos que em verbos; em nomes com a terminação *gem*, um efeito de frequência de tipo, antes de um efeito de frequência lexical, parece ser responsável pela alta aplicação do fenômeno, tendo em vista que RN, nesses casos, não enxerga estrutura morfológica interna nem é passível de ser abarcado por uma regra fonológica – considerando a baixa aplicação em verbos com palatal em contexto precedente; fatores sociais, tais como idade e escolaridade, parecem interagir de alguma forma com a informação de frequência.

Tendo em vista 1) a impossibilidade de o programa em questão analisar frequência como uma variável contínua e 2) a variável frequência não ter sido selecionada na melhor rodada dos não verbos, o que nos impede de apresentarmos inputs e significâncias, partimos para a análise dos dados na plataforma R, com o intuito de obtermos mais confiabilidade estatística.

4.1.2 Resultados plataforma R¹⁰

Ao analisar os dados em regressão logística binária no pacote estatístico R (R Core Team, 2016), trabalhamos com as seguintes variáveis: classe de palavra, frequência contínua (por meio de logaritmo da informação de frequência presente no ASPA), idade, escolaridade, contexto precedente. Vejamos mais detalhadamente os fatores presentes em cada variável a ser analisada.

Na variável classe de palavra, temos os seguintes fatores:

Não verbos em geral (homem)
Nomes em gem – raiz (viagem)
Nomes em gem – sufixo (passagem)
Presente do indicativo (cantam)
Pretérito perfeito do indicativo (cantaram)
Pretérito imperfeito do indicativo (cantavam)
Presente do subjuntivo (partam)
Futuro do pretérito do indicativo (partiriam)
Pretérito do subjuntivo (cantassem)
Futuro do subjuntivo ou Infinitivo flexionado (cantarem)

Quadro 8: Classe de palavra em RN

Exemplo do tratamento de frequência lexical – frequência no ASPA/Logaritmo:

Palavra	Frequência no ASPA	Logaritmo da Frequência
podem	87088	4.939
contagem	3499	3.543
acomodavam	12	1.079

Quadro 9: Frequência contínua em RN

¹⁰ Meus agradecimentos à professora Mirjam Ernestus, da Radboud Universiteit Nijmegen e do Max Planck Institute for Psycholinguistics, pela ajuda na rodada dos dados e pela discussão dos resultados.

Idade:

De 25 a 50 anos
51 anos ou mais

Quadro 10: Idade em RN

Escolaridade:

Primário (0 a 4 anos de escolaridade)
Ensino Médio (9 a 12 anos de escolaridade)

Quadro 11: Escolaridade em RN

Contexto Precedente ao ditongo:

Lábio-dental (falavam, nuvem)
Bilabial (deturpam, roubam)
Alveolar (responderam, acordam)
Palatal e palato-alveolar (colhem, viagem)
Velar (chegam, aplicam)
Nasal labial (chamam, homem)
Nasal alveolar (abandonam, aprisionam)
Nasal palatal (acompanham, ganham)
Zero (queriam, mereciam)

Quadro 12: Contexto Precedente 1 em RN

Em função de problemas de convergência no programa e guiados pelos resultados apresentados na seção anterior, decidimos analisar palatal em contexto precedente e amalgamar os outros contextos. Isso nos deixou com os seguintes fatores para a variável contexto precedente:

Palatal
Outros contextos precedentes

Quadro 13: Contexto Precedente 2 em RN

Para refinar a análise feita no GoldVarb X, trabalhamos com um modelo que apresentasse interações entre as variáveis classe de palavra, frequência e idade, e que nos desse resultados no que diz respeito a escolaridade e a contexto precedente. Para verificar o ajuste desse modelo comparativamente a outros, recorreremos ao AIC (*Akaike information criterion*), um critério que avalia a qualidade do modelo estatístico para o conjunto de dados. O modelo de melhor ajuste apresenta um AIC menor. Ao comparar o modelo que apresenta idade interagindo com frequência e classe de palavra com outras interações¹¹, o modelo com idade apresenta um AIC menor, ou seja, de melhor ajuste. Os resultados das variáveis significativas estatisticamente estão no quadro a seguir:

Variável	Nível	β	z	p-valor ¹²
Intercept		0.60016		
Verbo	Futuro do subjuntivo ou infinitivo flexionado	-3.34029	-2.414	<0.1
Não verbos	Não verbos sem -gem	-33.06468	-5.104	<0.001
Frequência	+ Frequente	-0.86116	-2.159	<0.1
Escolaridade	Primário	0.25657	4.066	<0.001
Interação	Pretérito do Subjuntivo * Frequência	1.37577	2.587	<0.01

¹¹ Na construção do modelo, tentamos interações entre frequência e classe de palavra com escolaridade e, após, com contexto precedente. No entanto, o valor do AIC com essas duas interações é maior que o apresentado com idade (AIC = 6308.4 com interação com escolaridade, AIC = 6293.1 com interação com contexto precedente), o que sugere que o modelo de melhor ajuste seja o que apresenta interação com idade.

Interação	Futuro do Subjuntivo ou Infinitivo Flexionado * Frequência	1.20133	2.303	<0.1
Interação	Não verbos sem -gem * Frequência	7.90018	5.535	<0.001
Interação	Presente do Indicativo * Frequência	0.86655	2.100	<0.1
Interação	Não verbos sem -gem * 25 a 50 anos	27.74186	3.806	<0.001
Interação	Futuro do Subjuntivo ou Infinitivo Flexionado * Frequência * 25 a 50 anos	-2.07406	-2.021	<0.1
Interação	Não verbos sem -gem * Frequência * 25 a 50 anos	-6.92079	-4.089	<0.001

Quadro 14 – Resultados para verbos e não verbos em RN. O intercept corresponde a preservação da nasalidade, presente do subjuntivo, menos frequente, mais de 51 anos, ensino médio.

A análise desta tabela se dá da seguinte maneira: o intercept corresponde a preservação da nasalidade, presente do subjuntivo, menos frequente, mais de 51 anos, ensino médio. Todos os fatores que se mostraram significativos são analisados em relação aos fatores presentes no intercept.

Das variáveis analisadas, contexto precedente foi uma que não se mostrou significativa, não tendo sido selecionada como estatisticamente relevante. Dentre as selecionadas, percebemos que futuro do subjuntivo / infinitivo flexionado e não verbos

sem *gem* apagam menos que presente do subjuntivo – que está no intercept. Em relação à frequência, percebemos que, ao analisar todos os dados em conjunto, palavras mais frequentes tendem a apagar menos. No que diz respeito à escolaridade, informantes que apresentam apenas o primário tendem a apagar mais que informantes com 9 a 12 anos de escolaridade.

O que nos interessa particularmente neste quadro são as interações com frequência. Podemos perceber que verbos no pretérito do subjuntivo, no futuro do subjuntivo (ou infinitivo flexionado) e no presente do indicativo são sensíveis à frequência para a aplicação do fenômeno de redução da nasalidade: quanto mais frequentes são essas palavras, mais propensas à redução. No que diz respeito aos não verbos sem *gem*, essa sensibilidade é ainda maior, tendo em vista a significância estatística.

No que diz respeito à interação com idade, temos resultados também significativos estatisticamente. O que podemos ler dessa tabela relativamente às interações classe de palavra e idade é o seguinte: comparativamente ao que está no intercept, não verbos sem *gem* em falantes de 25 anos a 50 anos apagam mais. Relativamente às interações significativas entre classe de palavra, frequência e idade, temos que, comparativamente ao que consta no intercept: dados de futuro do subjuntivo (ou infinitivo flexionado), mais frequentes de falantes de 25 a 50 anos apagam menos. Da mesma forma, dados de não verbos sem *gem*, mais frequentes de falantes de 25 a 50 anos também apagam menos.

O fator idade, assim como analisado na seção anterior, novamente parece ter impacto na aplicação do fenômeno, com interação significativa com a variável frequência lexical. Com base nisso, vejamos como a RN opera em relação à frequência separadamente nos informantes de 25 a 50 anos e nos informantes acima de 51 anos ou mais.

Variável	Nível	β	z	p-valor
Intercept		-0.15949		
Escolaridade	Primário	0.49559	4.770	<0.001

Quadro 15: Resultados para informantes de 25 a 50 anos.

Como já vimos nas análises anteriores, informantes mais jovens tendem a aplicar mais o fenômeno da redução da nasalidade. No entanto, essa redução não é sensível à frequência – não obtivemos resultados significativos no que diz respeito à interação entre frequência com nenhuma classe de palavra. Ao analisar informantes de 25 a 50 anos, a única variável que se mostrou significativa para a aplicação do fenômeno é a escolaridade – quanto menos escolaridade, maior a aplicação do fenômeno.

Vejamos a tabela relativa aos informantes com 51 anos ou mais:

Variável	Nível	β	z	p-valor
Intercept		0.63606		
Verbo	Futuro do subjuntivo ou infinitivo flexionado	-3.24891	-2.351	<0.1
Não verbos	Não verbos sem -gem	-33.19672	-5.138	<0.001
Frequência	+ Frequente	-0.84427	-2.125	<0.1
Interação	Pretérito do Subjuntivo * Frequência	1.37503	2.591	<0.01
Interação	Futuro do Subjuntivo ou Infinitivo Flexionado * Frequência	1.16251	2.233	<0.1
Interação	Não verbos sem -gem * Frequência	7.91494	5.560	<0.001
Interação	Presente do Indicativo * Frequência	0.84787	2.062	<0.1

Quadro 16: Resultados para informantes de 51 anos ou mais.

No que diz respeito aos dados de informantes de 51 anos ou mais, novamente em comparação ao *intercept*, futuro do subjuntivo (ou infinitivo flexionado) e não verbos sem *gem* tendem a aplicar menos o fenômeno. Frequência, quanto mais alta, menos favorece o processo – analisando em conjunto todas as classes de palavra.

Como podemos ver pelas interações, informantes de 51 anos ou mais são sensíveis à frequência no pretérito do subjuntivo, no futuro do subjuntivo (ou infinitivo flexionado) no presente do indicativo e, de forma bastante contundente, nos não verbos sem *gem*. Quanto mais frequentes são essas palavras, mais propensas à redução da nasalidade elas estão para esses falantes.

Esse resultado pode apontar para um possível caso de mudança linguística: os jovens se mostram mais propensos à aplicação do fenômeno da RN, mas essa redução não interage com frequência; em contrapartida, os informantes mais velhos, apesar de apresentarem taxas menores de redução, são bastante sensíveis à frequência lexical na aplicação do fenômeno, principalmente nos não verbos (maior significância estatística nos não verbos que nos verbos).

No pretérito do indicativo, temos 1.677 dados; no pretérito do subjuntivo, temos 64 dados; e no futuro do subjuntivo / infinitivo flexionado, temos 162 dados. No que diz respeito ao último, caracterizamo-lo dessa forma dupla tendo em vista que, em nosso arquivo de dados, muitas vezes só tínhamos a palavra – precisávamos de mais contexto para precisar se era o caso de futuro de subjuntivo ou de infinitivo flexionado. Sabemos que há casos de verbos irregulares em que ambos não convergem para uma mesma forma, há mudança no radical (*tiverem/terem*, *fizerem/fazerem*), diferentemente do que acontece em verbos regulares, em que as formas se mantêm as mesmas (*brincarem/brincarem*, *guardarem/guardarem*); no entanto, optamos por essa classificação dupla para evitar a criação de mais uma categoria com um número pequeno de dados que, provavelmente, nos daria problemas de convergência nas rodadas. De qualquer forma, tanto verbos regulares quanto irregulares no pretérito do subjuntivo e no infinitivo flexionado apresentam a terminação *rem*. A redução da nasalidade, assim como acontece no pretérito perfeito do indicativo, não implica a perda da informação morfológica número-pessoal, tendo em vista a vogal /e/, que não estaria presente em formas do singular.

A perda da informação morfológica acontece, porém, no pretérito do subjuntivo. Em verbos como *cantassem, fizessem, quisessem*, a nasalidade é a responsável pela informação número-pessoal. O mesmo acontece na maioria dos verbos do presente do indicativo, tais como *cantam, vendem, atendem*. Ou seja: mesmo em verbos em que a nasalidade é responsável por uma informação morfológica importante, a frequência apresentou interação estatisticamente significativa na aplicação do fenômeno em nossos dados. Apesar disso, não podemos assegurar que essa sensibilidade seja efetiva – a presença ou não do sujeito e sua proximidade com o verbo em questão parece ser uma variável a ser considerada, capaz de exercer papel determinante na variação em questão. Tendo em vista que essa variável não foi controlada neste estudo, pouco podemos afirmar sobre essa sensibilidade à frequência lexical desses verbos.

Schwindt, Bopp da Silva e Quadros (2012), ao analisarem o papel da expressão do sujeito na redução da nasalidade em um recorte de dados, apontaram que a redução da nasalidade tende a preservar a expressão de informações morfológicas, tendo em vista que o fenômeno parece se aplicar com mais liberdade quando há sujeito exposto. O presente do indicativo, inclusive, foi um dos tempos analisados em que essa ideia se confirmou.

No que diz respeito aos não verbos sem gem, temos 179 ocorrências em nosso corpus, que dizem respeito a cinco *types* de itens lexicais, quais sejam, em ordem crescente de frequência: *nuvem, jovem, ordem, homem, ontem*. A interação com frequência nesses dados de informantes de 51 anos ou mais se mostrou bastante significativa ($p\text{-valor} < 0.001$). Como já apontado anteriormente, apesar de o grupo dos jovens apagar mais, a interação dessa aplicação do fenômeno com frequência não se apresentou significativa. Esse resultado, principalmente no que diz respeito a essa classe de palavra, pode apontar para um possível caso de mudança linguística em que a frequência do item lexical apresenta papel significativo.

Pierrehumbert (2001) aponta que o efeito da frequência lexical na taxa de redução de segmentos deve ser observável tanto sincronicamente – ao se comparar a pronúncia de palavras de diferentes frequências – quanto diacronicamente – ao examinar a evolução das pronúncias das palavras através da passagem do tempo. O modelo de exemplares, segundo a pesquisadora, é o único modelo atual que apresenta essas propriedades.

Retomemos, agora, as perguntas feitas na metodologia sobre esse fenômeno:

Pode-se falar em efeitos de frequência lexical nos dados do sul do Brasil? Se sim:

- 1) palavras mais frequentes estão mais sujeitas ao processo?**
- 2) palavras menos frequentes estão menos sujeitas ao processo?**

Ao considerarmos a análise de toda a amostra em três níveis de frequência no GoldVarb, não há diferenças significativas entre os níveis. No Programa R, palavras mais frequentes parecem inibir o processo. No entanto, é necessário que a análise seja feita considerando diferentes classes de palavra. Passemos, então, à retomada das sub-questões propostas em nossa metodologia para análise do fenômeno de redução da nasalidade.

Sub-questões:

- a) esses efeitos são diferentes para não verbos e para verbos?**

Sim, esses efeitos parecem ser diferentes: para não verbos sem gem, frequência parece ter papel significativo na aplicação do fenômeno; para nomes com gem final, sendo raiz e sufixo, frequência do item lexical parece não ter papel da redução – o que defendemos é que a frequência do segmento gem pode estar atuando nestes casos, ou seja, não a frequência do item lexical, mas a frequência de tipo; para verbos, como analisado na última seção, frequência parece ter papel em dados de informantes de 51 anos ou mais no presente do indicativo, no pretérito do subjuntivo e no futuro do subjuntivo / infinitivo flexionado. No entanto, como não analisamos neste estudo o papel da expressão do sujeito e de sua proximidade com o verbo em questão, pouco podemos afirmar sobre essa interação.

- b) qual a relação entre frequência lexical e fatores linguísticos, tais como contexto precedente e morfologia interna à palavra?**

Contexto precedente não se mostrou significativo na aplicação do fenômeno em nossa análise. Em uma tentativa de interação entre contexto precedente e

frequência, não obtivemos interação significativa com nenhuma classe de palavra. No que diz respeito à morfologia interna à palavra, o fenômeno parece não enxergar estrutura morfológica interna, tendo em vista a análise de dados com *gem* na raiz e no sufixo. Interação com frequência lexical nos nomes terminados em *gem*, seja na raiz, seja no sufixo, não se mostrou significativa. Parece-nos que o que afeta a alta aplicação do fenômeno em nomes terminados por *gem* é a frequência do segmento final, ou seja, do tipo *gem*.

c) qual a relação entre frequência lexical e fatores sociais, tais como idade e escolaridade?

Podemos perceber que há interações bastante significativas entre frequência e idade, o que, de certa forma, corrobora a ideia de que léxico é geracional. Efeitos de frequência significativos foram encontrados em dados de informantes com mais de 51 anos nas seguintes classes de palavra: não verbos sem *gem*, presente do indicativo, pretérito imperfeito do subjuntivo, futuro do subjuntivo / infinitivo flexionado. Apesar de os informantes de 25 a 50 anos aplicarem mais o fenômeno, essa aplicação não está relacionada com efeitos de frequência, diferentemente do que acontece nos dados de informantes de 51 anos ou mais. Esse resultado pode ser o indício de uma mudança linguística. No que diz respeito à interação entre frequência lexical e escolaridade, encontramos interação significativa com a classe de palavra não verbos sem *gem*: quanto mais frequente, maior a aplicação do fenômeno tanto em falantes mais escolarizados quanto em falantes menos escolarizados. Também encontramos correlação positiva entre frequência lexical e aplicação do fenômeno nos dados de informantes de baixa escolaridade no pretérito imperfeito do subjuntivo. Não encontramos interação significativa de frequência lexical com outras classes de palavra.¹³

¹³ Em informantes de alta escolaridade, a interação entre o pretérito imperfeito do subjuntivo e frequência lexical foi positiva, com o valor $p=0.06$. Em informantes de baixa escolaridade, a interação entre frequência lexical e nomes com *gem* no sufixo foi positiva, com valor $p=0.098$. Ambas, no entanto, não figuram como significativas estatisticamente.

d) nos nomes, qual a relação entre frequência lexical e terminação em -gem e outras?

Retomando em parte a resposta dada à pergunta b), interação com frequência lexical nos nomes terminados em *gem*, seja na raiz, seja no sufixo, não se mostrou significativa. Parece-nos que o que afeta a alta aplicação do fenômeno em nomes terminados por *gem* é a frequência do segmento final, ou seja, do tipo *gem*. No que diz respeito aos não verbos sem *gem*, quais sejam *nuvem*, *jovem*, *ordem*, *homem*, *ontem*, a interação é bastante significativa: quanto mais frequente, maior é a aplicação do fenômeno.

d) nos verbos, qual a relação entre frequência lexical e cada tempo verbal?

Os tempos verbais que se mostraram significativos em relação à frequência foram o presente do indicativo, o pretérito imperfeito do subjuntivo e o futuro do subjuntivo / infinitivo flexionado. Nestes casos, assim como nos não verbos sem *gem*, a correlação foi positiva: quanto mais frequente, maior a aplicação do fenômeno. Fazemos uma ressalva em relação a esse resultado, tendo em vista a necessidade de controle da variável expressão do sujeito.

Retomemos a primeira grande pergunta da metodologia, qual seja: Pode-se falar em efeitos de frequência lexical nos dados do sul do Brasil? Após análise dos dados, chegamos à conclusão de que sim, podemos falar de efeitos de frequência nos dados do sul do Brasil ao analisarmos, separadamente, determinadas classes de palavra. Em fazendo isso, nas classes de palavra em que encontramos efeitos de frequência na aplicação do processo (não verbos sem *gem*, presente do indicativo, pretérito imperfeito do subjuntivo e futuro do subjuntivo/infinitivo flexionado), esse efeito é positivo: quanto mais frequente a palavra, maior é a aplicação do fenômeno. Ao analisar todos os dados conjuntamente, não conseguimos ver esse efeito.

No que diz respeito a este fenômeno, variáveis como taxa de elocução e indivíduo, além da expressão do sujeito para o caso dos verbos, deveriam também ser analisados. Tendo em vista nossa metodologia, o fato de termos apenas os

arquivos de dados de RN com o propósito de reanálise, não foi possível proceder ao estudo dessas variáveis, considerando nossa limitação às informações presentes nos arquivos de dados. Estudos futuros devem contemplar também essas variáveis.

4.2 Apagamento de /r/ final

O apagamento de /r/ é analisado preliminarmente por meio de coeficiente de correlação e, após, também pelo Programa R.

4.2.1 Análise preliminar – coeficiente de correlação

No que diz respeito ao fenômeno de apagamento do /r/ final, preliminarmente foram analisadas as 8 entrevistas da amostra do Programa da Marília Gabriela por meio do coeficiente de correlação entre a aplicação do fenômeno e a frequência de cada item lexical, tomando-se o ASPA como corpus de referência.

As análises foram realizadas por coeficiente de correlação tendo em vista a impossibilidade de estabelecermos escalas de frequência para utilizarmos o GoldVarb X (por mais níveis que propuséssemos, a maioria das palavras ficava concentrada no nível de alta frequência). Para a construção dos gráficos, foram excluídos os casos de aplicação categórica e os dados com 50% de aplicação, além dos itens lexicais “for” e “expor”. Os resultados são os que seguem:

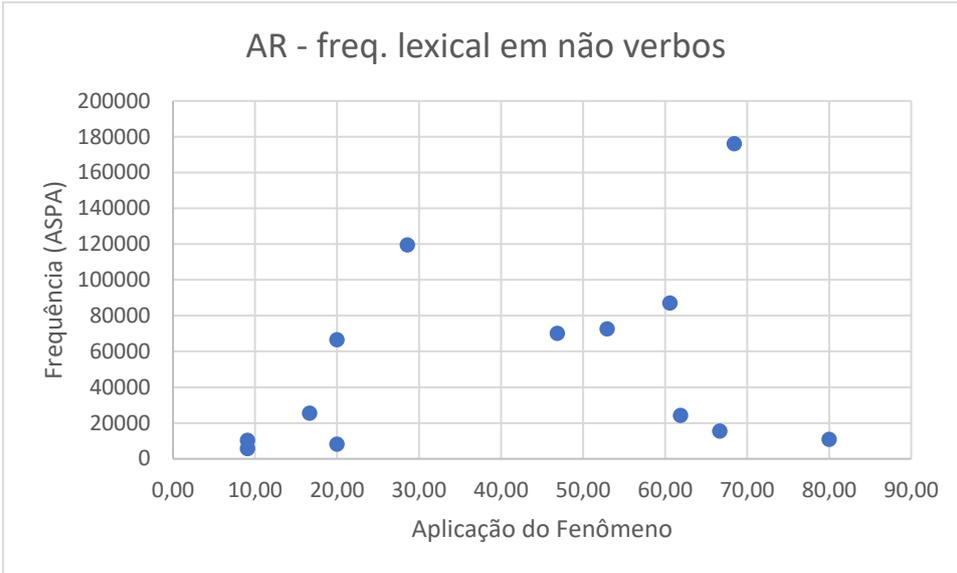


Gráfico 1: Coeficiente de correlação em não verbos

Coeficiente de Correlação de Pearson: 0,28

P = 0,05

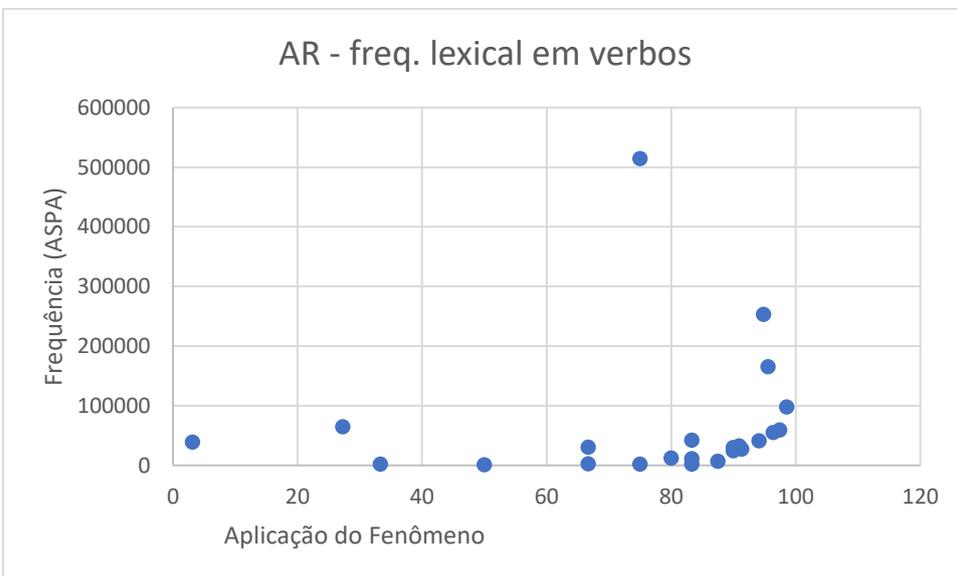


Gráfico 2: Coeficiente de correlação em verbos

Coeficiente de Correlação de Pearson: 0,13

P = 0,05

Tanto para verbos quanto para não verbos, a correlação entre aplicação do fenômeno e frequência lexical é positiva, apesar de pequena (coeficientes de correlação próximos a zero). No entanto, assim como na análise do fenômeno de redução da nasalidade, os dados de não verbos apresentam maior influência da frequência na aplicação do fenômeno (coeficiente de 0,28 para não verbos e de 0,13 para verbos).

No que diz respeito à variável social, trazemos o resultado em relação à idade para os dados de não verbos:

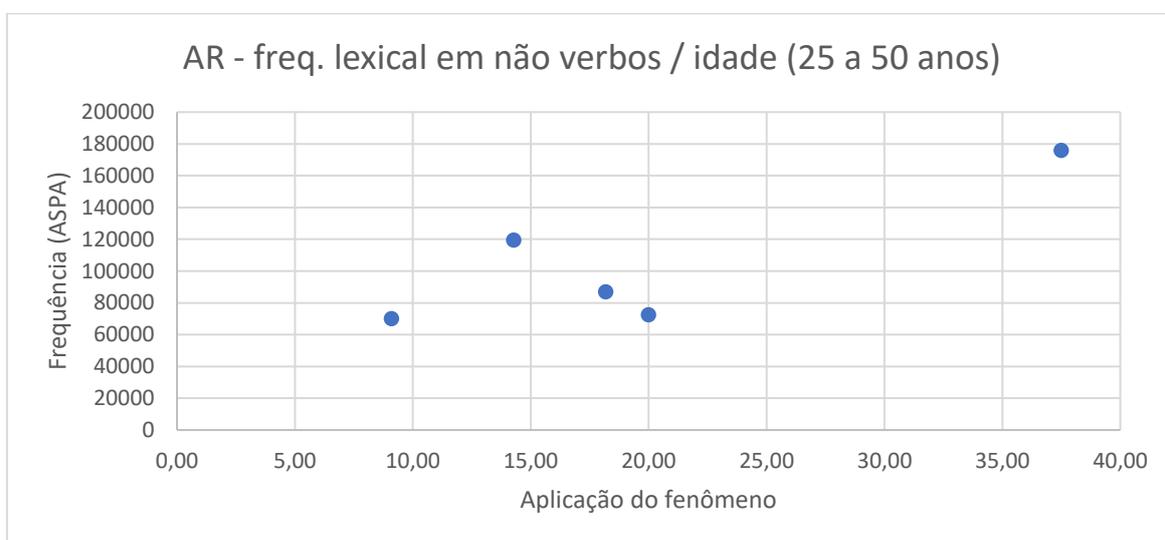


Gráfico 3: Coeficiente de Correlação em não verbos – 25 a 50 anos

Coeficiente de Correlação de Pearson: 0,82

P = 0,05

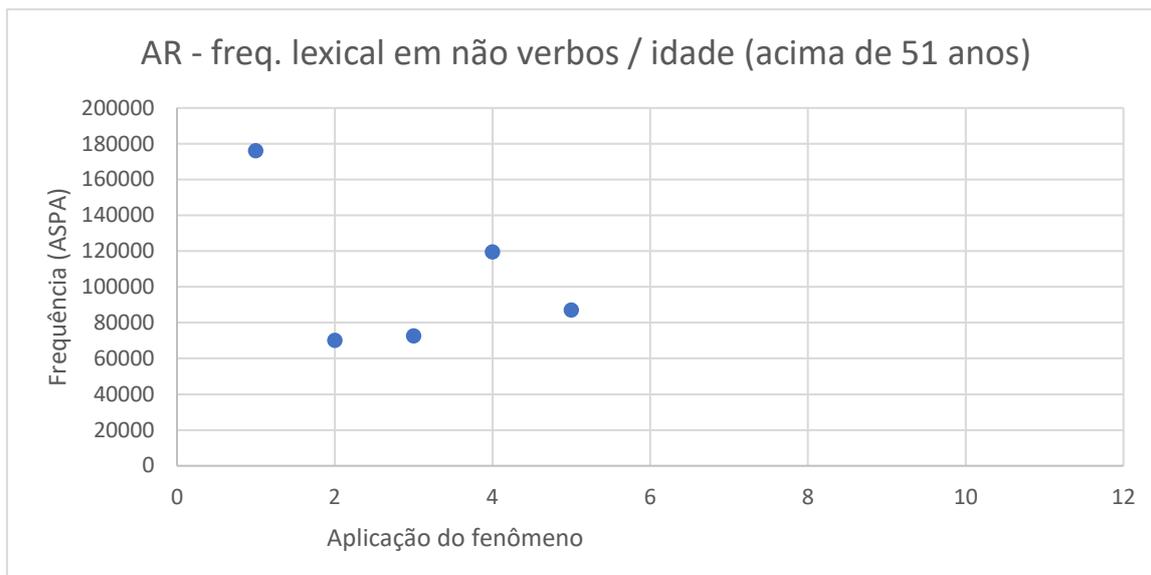


Gráfico 4: Coeficiente de correlação em não verbos – 51 anos ou mais

Coeficiente de Correlação de Pearson: 0,29

P: 0,05

A correlação entre aplicação do fenômeno e frequência lexical é positiva em ambas as faixas etárias. No entanto, percebe-se uma correlação mais forte na faixa etária dos 25 aos 50 anos, ou seja, entre os mais jovens – a variável social idade, novamente, parece ter papel no processo.

Não encontramos dados significativos estatisticamente no que diz respeito à análise de idade em verbos.

Essa análise por meio de coeficiente de correlação foi preliminar e exploratória apenas, tendo em vista que apresenta somente dados da primeira amostra. Passemos, então, à análise de todos os dados.

4.2.2 Resultados Plataforma R

Dos 2.610 dados de nossa amostra, o apagamento acontece em 2.227, o que configura 85,3% dos casos.



Gráfico 5: Apagamento de /r/ final

Ao analisar todos os dados na plataforma R (R Core Team, 2016), criamos um modelo de regressão logística. As variáveis eleitas para a análise do fenômeno de apagamento do /R/, com base na literatura da área, foram: classe de palavra, vogal precedente, dimensão do vocábulo, contexto seguinte, frequência lexical e idade e gênero.

Vejamos os fatores presentes em cada variável. Começamos por classe de palavra. Em um primeiro momento, trabalhamos com a seguinte organização de dados:

Verbos no infinitivo (fazer, querer)
Verbos conjugados (quer, souber)
Substantivo (poder, professor)
Adjetivo (espetacular, popular)
Pronome (qualquer)
Advérbio (melhor, pior)

Conjunção (apesar)

Quadro 17: Variável classe de palavra

Considerando análises prévias e tendo em vista um reduzido número de dados para os não verbos, decidimos por agrupar os dois primeiros fatores na categoria de *verbos* e os demais na de *não verbos*.

No que diz respeito à vogal precedente, analisamos os seis fatores a seguir:

/a/ (militar, gostar)
/e/ (poder, ver)
/ɛ/ (mulher, quisier)
/i/ (partir)
/o/ (professor, expor)
/ɔ/ (melhor, pior)

Quadro 18: Variável contexto precedente

Em relação ao contexto seguinte:¹⁴

¹⁴ Em uma primeira rodada, incluímos os seguintes fatores para contexto seguinte:

oclusiva bilabial
oclusiva alveolar
oclusiva velar
nasal bilabial
nasal alveolar
fricativa labiodental
fricativa alveolar
fricativa palato-alveolar
aproximante lateral alveolar
africada palato-alveolar
tepe alveolar
vibrante alveolar
fricativa velar
anterior baixa
anterior média fechada
anterior média aberta
anterior alta
posterior média fechada

Consoante (maior que)
Vogal (mulher alta)
Pausa (precisar #)

Quadro 19: Variável contexto seguinte

Em relação ao tamanho do vocábulo, trabalhamos com uma dupla distinção:

Monossílabo (pôr, par)
Dissílabo, trissílabo, polissílabo: não monossílabo (mulher, estudar)

Quadro 20: Variável tamanho do vocábulo

E gênero:

Feminino
Masculino

Quadro 21: Variável gênero

A variável idade foi codificada como variável contínua, assim como a variável frequência – por meio de logaritmo. O detalhamento da idade e de cada um dos informantes está apresentado na metodologia deste trabalho.

posterior média aberta
posterior alta
pausa curta
pausa longa

Após algumas análises com problemas de convergência, optamos por deixar apenas três níveis de fatores nesta variável.

Na tabela a seguir, apresentamos apenas as variáveis que se mostraram significativas estatisticamente:

Variável	Nível	B	z	p-valor
Intercept		0.657314	1.039	0.2987
Gênero	Masculino	1.015706	7.424	<0.001
Classe de Palavra	Verbo	2.610165	13.224	<0.001
Vogal Precedente	/e/	-0.391582	-2.041	<0.1
Vogal Precedente	/i/	-0.580118	-2.236	<0.1
Vogal Precedente	/o/	-0.525187	-2.234	<0.1
Contexto Seguinte	Pausa	0.336308	2.002	<0.1
Contexto Seguinte	Vogal	-0.294828	-1.967	<0.1

Quadro 22 – Resultados para verbos e não verbos. O que está no intercept corresponde à preservação de /r/, não verbos, gênero feminino, vogal precedente /a/, contexto seguinte consoante.

Ao analisar verbos e não verbos conjuntamente, nossos resultados confirmam muito do que já foi dito até então: verbos têm um índice maior de apagamento, assim como palavras de duas ou mais sílabas. Em relação à vogal precedente, em comparação a /a/, que está no intercept, contexto precedente com as vogais /e/, /i/ e /o/ parecem ser menos favoráveis ao apagamento. Comparativamente à consoante em contexto seguinte, pausa parece favorecer o fenômeno e vogal, ao contrário, inibir.

Contexto precedente com as vogais abertas /ε/ e /ɔ/, comparativamente a /a/, Dimensão do vocábulo, idade e frequência lexical não foram variáveis selecionadas como significativas.

De acordo com Callou, Moraes e Leite (1998), também defendemos a necessidade de se analisar verbos e não verbos em separado nesse fenômeno, tendo

em vista que a seleção dos grupos significativos pode não refletir corretamente os ambientes que condicionam o apagamento se a análise for conjunta.

VERBOS

Em nossa amostra, dos 2.129 verbos, o apagamento ocorre em 1.971, ou seja, em 92,5% dos casos.

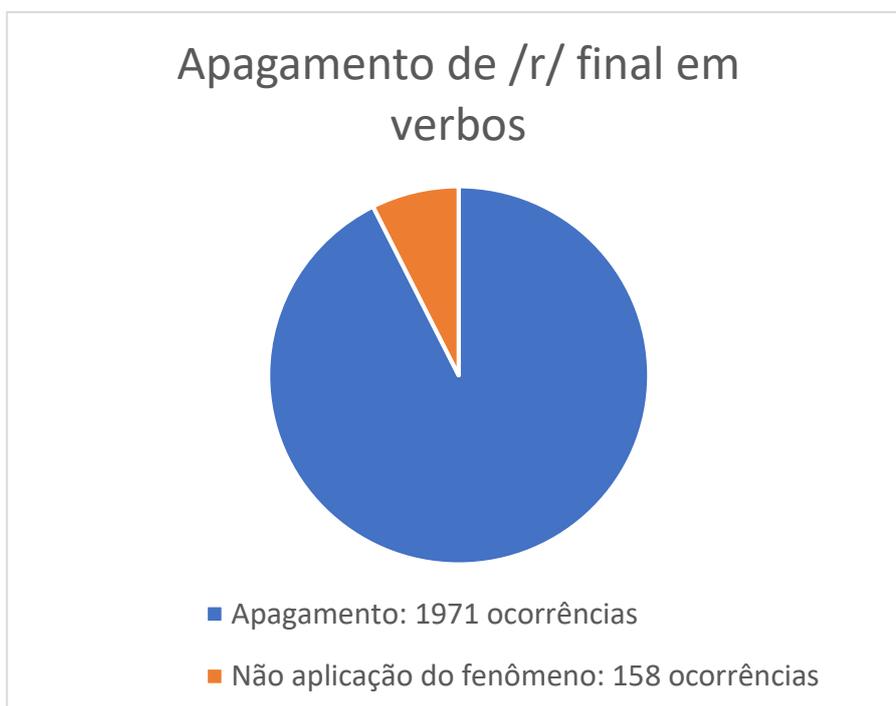


Gráfico 6: Apagamento de /r/ final em verbos

De acordo com a literatura sobre o fenômeno, principalmente com base em Monaretto (2002) e Callou, Serra e Cunha (2015), elegemos as seguintes variáveis para a análise do apagamento de /R/ final em verbos: sexo, dimensão do vocábulo, vogal precedente, contexto seguinte e idade. Além dessas variáveis, seguindo os propósitos deste trabalho, incluímos a variável frequência lexical como variável contínua.

Ao rodar essas variáveis com a variável aleatória informante, tivemos problemas de convergência. Nossa estratégia, então, foi fazer uma rodada em um modelo de regressão logística sem efeitos mistos, ou seja, sem a variável aleatória

informante. Nossos resultados para verbos estão na tabela a seguir, em que constam apenas as variáveis significativas estatisticamente:

Intercept		4.96258		
Gênero	Masculino	0.91170	5.698	<0.001
Idade	Mais idade	-0.02575	-5.233	<0.001
Vogal Precedente	/i/	-0.60843	-2.260	<0.1
Vogal Precedente	/o/	-3.13711	-6.375	<0.001

Quadro 23 – Resultados para verbos. O que está no intercept corresponde a preservação de /R/, gênero feminino, menos idade, vogal precedente /a/.

Ao analisar apenas verbos, nossos resultados apontam três variáveis como significativas: gênero, idade e contexto precedente. No que diz respeito a gênero, diferente dos resultados de Monaretto (2002), concernente aos dados do sul do Brasil, homens parecem favorecer o apagamento em nossa amostra, comparativamente a mulheres, que estão no intercept. Para idade, como já atestado na literatura, jovens parecem ser mais propensos à aplicação do processo. Sobre a única variável linguística que se mostrou significativa, contexto precedente com as vogais /i/ e /o/ parece inibir o apagamento, comparativamente à vogal /a/ que está no intercept.

Contexto precedente com as vogais /ɛ/ e /e/ não mostrou significância estatística comparativamente a /a/, nem as variáveis contexto seguinte (diferentemente da análise de Callou, Serra e Cunha, 2015, para a cidade de Aracaju) e dimensão do vocábulo (diferentemente da análise de Monaretto, 2002, nos dados do sul do Brasil). A análise de Monaretto (2002) sugere que a preferência do apagamento se dá em contexto precedente com vogais posteriores, quais sejam /a/ e /o/, comparativamente a vogais frontais /e/ e /i/. Como analisamos as vogais separadamente, encontramos um contraste entre as vogais /a/ e /o/ no que diz respeito à aplicação do fenômeno.

Apesar de o apagamento de /R/ em final de verbos ser atestado na literatura como um fenômeno praticamente categórico, nossa análise sugere a importância de atentarmos para a vogal precedente. Ao que tudo indica, este apagamento está

bastante relacionado à frequência não do item lexical propriamente dito, ou seja, à frequência de token, mas sim à frequência de um padrão linguístico – a frequência de tipo.

Para refinarmos nossa análise, construímos um outro modelo apenas com as variáveis que se mostraram relevantes estatisticamente, incluindo a variável aleatória informante, em um modelo misto.

Como já apontado, uma das vantagens de analisar o fenômeno na plataforma R é a possibilidade de incluir no modelo não apenas variáveis fixas, as linguísticas e as sociais, mas também variáveis aleatórias, ou seja, variáveis que não podemos controlar como as variáveis fixas, pois têm uma variação própria, idiossincrática. Fatores randômicos, isto é, que não podemos controlar experimentalmente, podem nos afastar das nossas previsões: ao incluir a variável informante como aleatória estamos, de certa forma, fornecendo uma estrutura para essa variação randômica, que é relativa a diferenças individuais (WINTER, 2013). Além de informante, muitos estudos incluem, também como variável aleatória, o próprio item lexical. Nesta análise, optamos por não incluí-lo, tendo em vista problemas de convergência no programa. Acima disso, considerando os propósitos deste trabalho, interessa-nos enxergar a variação de cada item lexical e sua relação com frequência.

Incluir indivíduo como variável aleatória parece fazer com que tenhamos uma análise mais precisa do papel das variáveis fixas – as linguísticas e as sociais. Ao comparar por meio da função Anova dois modelos – um com a variável aleatória informante e o outro apenas com as variáveis fixas –, o modelo misto com variável aleatória que apresentamos a seguir apresenta um AIC menor que o apresentado anteriormente (1010.4 *versus* 1045.6), o que sugere ser de melhor ajuste:

Variável	Nível	B	z	p-valor
Intercept		4.22699		
Gênero	Masculino	0.81786	2.015	<0.1
Idade	Mais velhos	-0.02842	-2.492	<0.1

Vogal Precedente	/e/	-0.63828	-3.231	<0.01
Vogal Precedente	/i/	-0.60368	-2.245	<0.1
Vogal Precedente	/o/	-3.25562	-6.607	<0.001

Quadro 24 – Resultados para verbos II. O que está no intercept corresponde a preservação de /R/, gênero feminino, menos idade, vogal precedente /a/.

Neste modelo, gênero e idade continuam apresentando relevância estatística – homens e informantes mais jovens tendem a apagar mais o /r/ final de verbos. No que diz respeito à vogal precedente, comparativamente a /a/, que está no intercept, as vogais /i/ e /o/ continuam favorecendo menos o apagamento. A novidade neste modelo está em um outro fator da variável contexto precedente que se mostrou significativo - /e/ parece também não favorecer tanto a aplicação do processo, comparativamente a /a/.

Na tentativa de verificar o papel do informante na aplicação do fenômeno, fizemos um teste criando um modelo em que informante não mais constava como uma variável aleatória, mas como uma variável fixa, dependente. Por meio desse modelo, observamos que determinados informantes de fato apresentam diferenças significativas na aplicação do fenômeno. No entanto, ao tentar estabelecer uma interação entre informante e frequência, vimos que essas diferenças não estão relacionadas à frequência do item lexical. Ao incluir coeficientes angulares aleatórios para frequência por informante, com o intuito de verificar se o efeito de frequência variava significativamente entre os informantes, tivemos problemas de convergência; no entanto, mesmo tentando estabelecer uma comparação entre o modelo com coeficiente angular para frequência por informante e o modelo com apenas a variável aleatória para informante, percebemos que o segundo modelo ainda apresentava um AIC menor (1010.4 *versus* 1014.1), ou seja, o efeito de frequência parece não variar por informante.

Como já demonstrado ainda no primeiro modelo para verbos, frequência do item lexical não apresenta papel na aplicação do fenômeno. No entanto, interessante, parece que podemos falar de frequência de tipo no que tange à

aplicação do fenômeno de apagamento em verbos: itens lexicais terminados em -ar, -er, -ir, -or apresentam taxas distintas de apagamento. Vejamos a tabela a seguir com o número de tokens de cada contexto precedente em nossa amostra, juntamente com sua respectiva taxa de aplicação do fenômeno.

Contexto Precedente	Total de Tokens	Total de Apagamentos	Percentual de Apagamento
/a/	1034	979	94,6%
/e/	728	666	91,4%
/i/	240	217	90,4%
/o/	22	10	45,4%

Quadro 25: Percentual de apagamento para as vogais /a/, /e/, /i/, /o/ em contexto precedente.

Na tabela, podemos observar que a aplicação do processo de apagamento é diretamente proporcional ao número de tokens de cada vogal precedente. Ou seja: mesmo se tratando de um fenômeno conhecido como de aplicação praticamente categórica, parece-nos que, antes de qualquer regra fonológica, a frequência de tipo pode ter papel relevante na aplicação ou não do processo.

NÃO VERBOS

No que diz respeito aos não verbos, tivemos um total de 481 ocorrências. A aplicação se deu em 256 casos, ou seja, em 53,2%

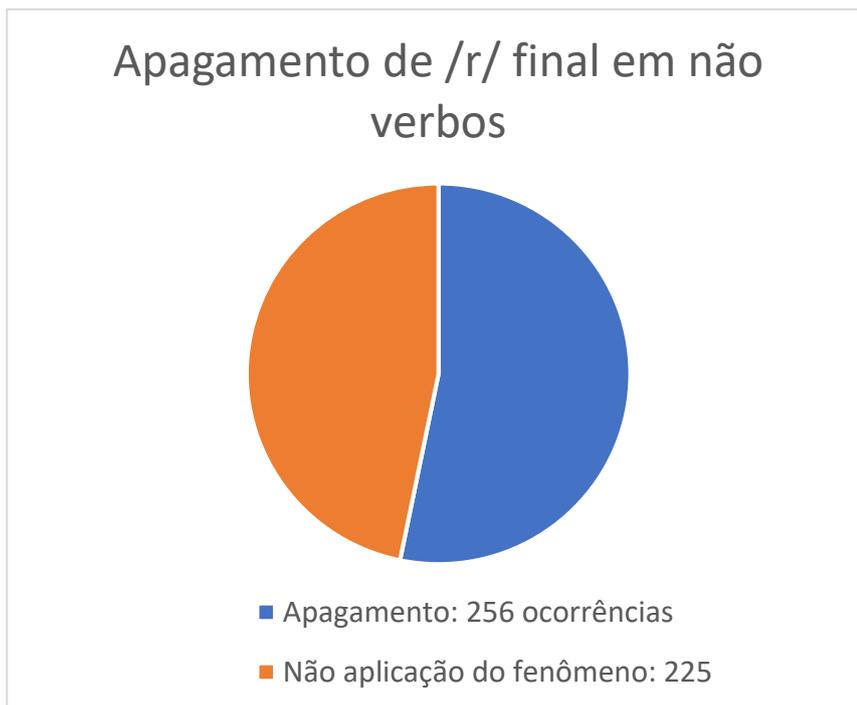


Gráfico 7: Apagamento de /r/ final em não verbos

De acordo com a literatura, resolvemos testar em nossa amostra as seguintes variáveis em um modelo com apenas variáveis fixas: classe de palavra (substantivo, adjetivo, advérbio, pronome, conjunção), vogal precedente (/a/, /e/, /ɛ/, /i/, /o/, /ɔ/), idade (sendo analisada como variável contínua), gênero (feminino ou masculino), dimensão do vocábulo (monossílabo ou não), contexto seguinte (vogal, consoante ou pausa) frequência lexical (sendo analisada como variável contínua em logaritmo) e informação morfológica (se o /r/ estiver no morfema, ou se em parte da raiz). Como ainda não havíamos apresentado essa variável, segue o quadro:

Morfema (estudar, compositor)
Raiz (poder, mulher)

Quadro 26: Variável informação morfológica

Vejamos as variáveis que se mostraram significativas:

Variável	Nível	B	z	<i>p</i> -valor
Intercept		-3.120605	-5.123	<0.001
Gênero	Masculino	0.967829	4.674	<0.001
Idade	Mais idade	0.020442	3.312	<0.001
Vogal Precedente	/a/	1.971413	3.705	<0.001
Vogal Precedente	/ε/	2.310290	4.308	<0.001
Vogal Precedente	/o/	1.631664	3.208	<0.01
Vogal Precedente	/ɔ/	1.722781	3.170	<0.01
Contexto Seguinte	Pausa	0.554690	2.346	<0.1
Contexto Seguinte	Vogal	-0.774054	-2.772	<0.01

Quadro 26 – Resultados para não verbos. O que está no intercept corresponde a preservação de /R/, gênero feminino, menos idade, vogal precedente /e/, contexto seguinte consoante.

Nossos resultados apontaram classe de palavra, dimensão do vocábulo, informação morfológica e, novamente, frequência lexical como variáveis não significativas estatisticamente para a aplicação ou não do processo.

Resultados de Callou, Serra e Cunha (2015) e de Huback (2003, 2006) apontaram dimensão do vocábulo como uma variável significativa; no entanto, isso pode ser relativo ao monossílabo átono ‘por’, não incluído neste estudo, que tende a inibir o apagamento. Antes de decidirmos pela análise apenas de vocábulos oxítonos e monossílabos tônicos, a preposição ‘por’ figurava em nossos dados com 144 ocorrências, dentre as quais apenas 17 aplicavam o processo, ou seja, apenas 11,8% dos casos.

Gênero, idade, contexto precedente e contexto seguinte se mostraram significativos. Assim como apontado nos resultados para não verbos e verbos conjuntamente e nos resultados para verbos, homens tendem a apagar mais o /r/ final

também em não verbos. No que diz respeito à idade, diferentemente do que acontece com verbos em nossa amostra, em que informantes mais jovens tendem a aplicar mais o processo, nos não verbos o apagamento ocorre mais entre os informantes de mais idade. Comparativamente a /e/ em contexto precedente, não verbos com contexto precedente em /a/, /ɛ/, /o/ e /ɔ/ aplicam mais o processo. No que tange ao contexto seguinte, comparativamente a consoante, itens lexicais cujo contexto seguinte é vogal tendem a aplicar menos o processo, e itens cujo contexto seguinte é pausa tendem a apagar mais o /r/ final, em nossa amostra, comparativamente a contexto precedente de consoante.

Tendo em vista o pequeno número de dados, a construção de um modelo misto com informante como variável aleatória não foi possível em função de problemas de convergência. Por esse mesmo motivo, também não foi possível checar coeficientes angulares aleatórios para frequência por informante, para verificar se o papel da frequência varia significativamente por informante. Interações entre frequência e idade, frequência e sexo, frequência e contexto precedente e frequência e contexto seguinte não se mostraram significativas, o que sugere que a variável frequência lexical, como demonstrado pelo nosso modelo, não apresenta papel na aplicação do fenômeno de apagamento de /r/ final em nossa amostra.

Tendo em vista esse resultado, passemos a uma análise que apresente o número de types e tokens, com seus respectivos percentuais de apagamento:

ANÁLISE DE TOKENS E TYPES POR CONTEXTO PRECEDENTE

Contexto Precedente	Número de Types	Total de Tokens	Total de Apagamentos	Percentual de Apagamento
/a/	21	93	55	59,1%
/e/	5	37	6	16,2%
/ɛ/	4	97	58	59,7%
/o/	64	178	100	56,1%
/ɔ/	6	78	38	48,7%

Quadro 27: Percentual de apagamento para as vogais /a/, /e/, /ɛ/, /o/ e /ɔ/ em contexto precedente de não verbos.

Podemos perceber que há um número mais elevado de types com vogal precedente /a/ e, ainda mais, com vogal precedente /o/. Nestes dois casos, a aplicação do fenômeno parece ser favorecida. Em relação a /e/, temos uma aplicação do fenômeno bastante baixa. No que diz respeito a /ɔ/ e a /ɛ/, as taxas de apagamento, principalmente deste último que está acima do ponto neutro, parece ser relativa a determinados itens lexicais. Vejamos mais detalhadamente quais são os itens lexicais, seus números de ocorrência e suas taxas de aplicação para cada contexto precedente.

Palavra	Ocorrências	Aplicação	Percentual de Aplicação
poder	19	3	15,7%
prazer	8	2	25%
ser	7	0	0%
lazer	2	0	0%
nascer	1	1	100%

Quadro 28: não verbos em /-er/

Apesar de os dados serem poucos, pode-se perceber que, no que tange a não verbos, /e/ em contexto precedente parece inibir o apagamento de /r/. Interessantemente, não podemos falar em regra fonológica, pois /e/ em contexto precedente em verbos, como vimos, tem altas taxas de aplicação do fenômeno. Um exemplo disso é a palavra *ser* – temos 119 ocorrências de *ser*¹⁵ como verbos, e o apagamento ocorre 95 vezes, ou seja, em 79,9% dos casos. Quando *ser* é usado como substantivo, apesar de o número de ocorrências ser baixo, percebe-se que há uma inibição do apagamento – não devido a alguma propriedade da vogal em contexto precedente, mas em função da classe gramatical.

¹⁵ Importante destacar que o corpus de referência ASPA não nos fornece informações de classe gramatical. O item 'ser' apresenta apenas um único índice de frequência.

Palavra	Ocorrências	Aplicação	Percentual de Aplicação
mulher	48	28	58,3%
qualquer	46	30	65,2%
colher	2	0	0%
sequer	1	0	0%

Quadro 29: não verbos em /-er/

De fato, no que diz respeito a /ɛ/, os índices de apagamento dizem respeito a duas palavras em específico, quais sejam *mulher* e *qualquer*.

Palavra	Ocorrências	Aplicação	Percentual de Aplicação
Melhor	31	12	38,7%
Maior	25	13	52%
Pior	10	6	60%
Menor	8	4	50%
Major	3	2	66%
Redor	1	1	100%

Quadro 30: não verbos em /-or/

Palavras que terminam em /o/, apesar de serem poucas em nossa amostra, apresentam percentuais de apagamento, em geral, acima do ponto neutro.

No que diz respeito a /a/ e /o/ em contexto precedente, vamos analisar os dados que apresentaram no mínimo três ocorrências. São eles:

Palavra	Ocorrências	Aplicação	Percentual de Aplicação
Lugar	26	16	61,5%
Popular	20	16	80%
Militar	12	6	50%
Apesar	6	2	33%
Patamar	4	3	75%
Ar	3	1	33%
Singular	3	0	0%

Quadro 31: não verbos em /-ar/

Palavra	Ocorrências	Aplicação	Percentual de Aplicação
Amor	13	4	30%
Jogador	13	10	76,9%
Humor	12	2	24%
Senhor	12	8	66,6%
Treinador	10	6	60%
Embaixador	9	5	55%
Produtor	9	0	0%
Governador	8	5	62,5%
Professor	6	5	83%
Compositor	4	3	75%
Motor	4	4	100%
Organizador	4	4	100%
Anterior	3	2	66,6%
Computador	3	1	33,3%
Dor	3	2	66,6%
Valor	3	2	66,6%
Voador	3	2	66,6%

Quadro 32: não verbos em /-or/

No que diz respeito ao apagamento de /r/ em vogal precedente /a/ e /o/, apesar de os dados serem poucos, parece que o apagamento está mais ligado à cadeia final e não a um item lexical em específico.

No entanto, parece-nos que há necessidade de especificar a classe gramatical a que a cadeia pertence. Muitos itens lexicais terminados em -or estão acima do ponto neutro no que diz respeito à porcentagem de aplicação do fenômeno em não verbos; em verbos, essa tendência é oposta. Isto é: a frequência de tipo, ou seja, a frequência do segmento -or, fazendo parte da raiz ou de um sufixo, parece ter papel na aplicação do fenômeno. Se esse segmento estiver em verbos, a redução é inibida; se estiver em não verbos, é favorecida. O que importa na aplicação ou não do fenômeno não está ligado a uma regra morfológica, porque o fenômeno parece não enxergar estrutura interna; também não está ligado a uma regra fonológica relativa a propriedades da vogal em questão, mas parece estar ligado a informação de classe gramatical. A informação de classe gramatical, portanto, parece importar para a aplicação do fenômeno.

Retomemos nossas perguntas em relação a esse fenômeno presentes na metodologia deste trabalho:

Pode-se falar em efeitos de frequência lexical nos dados do falar carioca? Se sim:

1) palavras mais frequentes estão mais sujeitas ao processo?

2) palavras menos frequentes estão menos sujeitas ao processo?

Tendo em vista nossa análise com todos os dados de forma conjunta, tanto com verbos quanto com não verbos, não encontramos significância estatística no que diz respeito ao papel da frequência lexical na aplicação do fenômeno. Parece-nos que há algum papel da frequência de tipo, ou seja, da recorrência de determinado segmento – que envolveria, neste caso, contexto precedente.

Sub-questões:

a) esses efeitos são diferentes para não verbos e para verbos?

Diferentemente do fenômeno de Redução da Nasalidade, em que há papel da frequência lexical em determinadas classes gramaticais, na análise de apagamento de /r/ não obtivemos resultados quanto à influência dessa variável com nenhuma classe gramatical analisada. No entanto, parece-nos que pode haver influência da frequência de tipo.

Com a terminação -or, particularmente, temos poucos tipos de verbos, em geral *pôr* e seus derivados. Essa baixa frequência de tipo em -or parece influenciar na baixa aplicação do processo nesses verbos (45,4%). No entanto, o segmento -or em não verbos apresenta uma taxa maior de aplicação do fenômeno (56,1%), pois, além de apresentar um número maior de dados, também é mais produtivo em PB. Podemos fazer uma relação semelhante com o contexto precedente em /e/: em verbos, temos uma alta taxa de aplicação do fenômeno (91,4%). Em não verbos, no entanto, esse percentual de aplicação cai para 16,2%, tendo em vista o número reduzido de tipos - apenas 5 tipos, com 37 tokens.

Ou seja, não temos efeitos de frequência de item lexical, mas temos um efeito de frequência de tipo, que opera tanto em verbos quanto em não verbos.

b) qual a relação entre frequência lexical e fatores linguísticos, tais como contexto precedente e morfologia interna à palavra?

Frequência lexical, como já observamos, não apresenta papel, por isso, não apresenta interação com as variáveis vogal precedente e morfologia interna à palavra. Entretanto, frequência de tipo parece exercer papel. Interessantemente, esse papel não interage com vogal precedente nem com morfologia interna à palavra – essas informações parecem não importar para a aplicação do fenômeno. Retomemos nossa análise para os dados em /e/ para exemplificar melhor essa ideia: “Apesar de os dados serem poucos, pode-se perceber que, no que tange a não verbos, /e/ em contexto precedente parece inibir o apagamento de /r/. Interessantemente, não podemos falar em regra fonológica, pois /e/ em contexto precedente em verbos, como vimos, tem altas taxas de aplicação do fenômeno. Um exemplo disso é a palavra *ser* – temos 119

ocorrências de *ser* como verbo, e o apagamento ocorre 95 vezes, ou seja, em 79,9% dos casos. Quando *ser* é usado como substantivo, apesar de o número de ocorrências ser baixo, percebe-se que há uma inibição do apagamento – não devido a alguma propriedade da vogal em contexto precedente, mas em função da classe gramatical.”

No que diz respeito à informação morfológica, as análises feitas na plataforma R apontaram essa variável como não significativa estatisticamente para a aplicação ou não do fenômeno, por isso parece que o processo não enxerga morfologia. Esse caso se assemelha ao caso dos nomes terminados em *gem*, analisados no fenômeno de redução da nasalidade, em que a aplicação do fenômeno também não enxerga estrutura interna morfológica: a aplicação do fenômeno de apagamento de /r/ em nomes com um sufixo agentivo, tais como *professor*, *colaborador*, não parecem diferir de nomes como *humor*, *valor*; o mesmo se dá com o sufixo formador de adjetivos, como *espetacular*, *popular*, em que a aplicação do processo não parece diferir da aplicação em palavras com -ar na raiz, tais como *devagar*, *apesar*.

Na linha de argumentação de Hanique e Ernestus (2012), que examinaram o papel da morfologia na redução acústica, parece que na análise de nossos dados também não temos evidência de que o status morfológico de um segmento afeta seu grau de redução, o que vai de encontro a premissas de modelos abstracionistas (CHOMSKY e HALLE, 1968; PINKER, 1991). Parece-nos, portanto, que as palavras são armazenadas como unidades completas no léxico mental, sendo acessadas diretamente. Isso vai ao encontro de modelos exemplaristas (BYBEE, 2001; PIERREHUMBERT, 2001), em que a estrutura morfológica não é altamente relevante para a produção da fala, tendo em vista que é através da conexão entre palavras relacionadas que a estrutura morfológica emerge.

Além disso, a argumentação das autoras de que um segmento que identifica uma palavra (ou seja, a carga de informação da palavra) tem um papel mais determinante na redução que um segmento que identifica a estrutura morfológica de tal palavra também parece proceder para a análise de nossos dados.

c) qual a relação entre frequência lexical e fatores sociais, tais como idade e sexo?

Interações entre frequência e idade e frequência e sexo não se mostraram significativas, o que sugere que a variável frequência lexical, como demonstrado pelo nosso modelo inicial, não apresenta papel na aplicação do fenômeno de apagamento de /r/ final em nossa amostra.

Passemos agora à seção que faz uma análise entre frequência lexical no corpus de referência e nos corpora analisados por nós, para podermos pensar no papel do indivíduo na aplicação dos fenômenos analisados nesta tese.

4.3 Sobre Léxico – Comparação entre os corpora

Tendo em vista que não foi possível analisar o papel do indivíduo nos dados de redução da nasalidade e considerando que, por problemas de convergência, também não foi possível fazer uma análise satisfatória sobre o papel do indivíduo nos dados de apagamento de /r/, nesta seção procedemos a uma comparação entre os *corpora* com os quais trabalhamos.

Uma pergunta caracteriza o exercício que realizamos aqui, qual seja: qual a relação entre verbos e não verbos mais frequentes nas amostras estudadas e a frequência geral de palavras com essa estrutura no PB no ASPA?

Passemos à análise do fenômeno de redução da nasalidade em não verbos e em verbos.

O quadro a seguir posiciona os 15 não verbos que apresentam o maior número de ocorrências, ou seja, que são os mais frequentes em nossa amostra-base, em relação aos 15 itens de mesma natureza mais frequentes no corpus de referência adotado, o Projeto ASPA.

Vocábulos	Frequência ASPA	Ranking ASPA	Ranking VAR SUL	Frequência VAR SUL
ontem	355407	1 ^o		
reportagem	187071	2 ^o	9 ^o	12
anteontem	63975	3 ^o		
homem	56807	4 ^o	1 ^o	239
imagem	36950	5 ^o		
ordem	32339	6 ^o	6 ^o	32
viagem	27461	7 ^o	3 ^o	65
origem	20647	8 ^o	4 ^o	56
jovem	20338	9 ^o	2 ^o	78
personagem	19363	10 ^o		
vantagem	18513	11 ^o	7 ^o	19
passagem	17259	12 ^o	5 ^o	43
linguagem	13550	13 ^o		
homenagem	11131	14 ^o	9 ^o	12
margem	10304	15 ^o	10 ^o	11
coragem			7 ^o	19
bobagem			8 ^o	15
serragem			9 ^o	12
lobisomem			11 ^o	10
mensagem			11 ^o	10

Quadro 33 – Redução da nasalidade em não verbos mais frequentes ASPA/VARSUL (Adaptado de DE BONA e SCHWINDT, 2017)

O que chama particularmente nossa atenção nessa comparação é o fato de 10 desses itens serem idênticos (linhas sombreadas) nessas duas amostras, de naturezas e porte tão distintos. Vejamos nossa comparação para verbos a seguir.

Tokens	Frequência ASPA	Ranking ASPA	Ranking VAR SUL	Frequência VAR SUL
foram	233914	1 ^o	3 ^o	188
podem	87088	2 ^o		
devem	64216	3 ^o		
eram	41718	4 ^o	2 ^o	221
estavam	40594	5 ^o	6 ^o	104
fazem	39545	6 ^o	1 ^o	249
tiveram	27544	7 ^o		
querem	26541	8 ^o	11 ^o	77
sejam	24922	9 ^o		
dizem	24363	10 ^o	12 ^o	62
ficaram	22944	11 ^o		
existem	21865	12 ^o		
fizeram	21521	13 ^o	10 ^o	83
tinham	21118	14 ^o	4 ^o	153
serem	20593	15 ^o		
vieram			5	109
diziam			7	91
iam			8	87
faziam			9	85
falam			13	56
gostam			14	51
ficam			15	48

Quadro 34 – Redução da nasalidade em verbos mais frequentes ASPA/VARSUL

No que diz respeito a verbos, 8 desses itens são idênticos (linhas sombreadas) nessas duas amostras; aqui, a comparação poderia se estabelecer entre os 14 itens mais frequentes de cada amostra.

Passemos agora à análise de verbos e não verbos do fenômeno apagamento de /r/.

Vocábulos	Frequência ASPA	Ranking ASPA	Ranking Amostra /R/ 2018	Frequência Amostra /R/ 2018
maior	175984	1 ^o	5 ^o	25
melhor	119542	2 ^o	3 ^o	31
diretor	103199	3 ^o		
qualquer	87041	4 ^o	2 ^o	46
valor	78034	5 ^o		
poder	74263	6 ^o	7 ^o	19
lugar	72529	7 ^o	4 ^o	26
setor	72048	8 ^o		
mulher	70112	9 ^o	1 ^o	48
apesar	66524	10 ^o		
governador	63020	11 ^o		
jogador	47664	12 ^o	8 ^o	13
militar	47052	13 ^o	9 ^o	12
professor	43261	14 ^o		
menor	42615	15 ^o		
popular			6 ^o	20
amor			8 ^o	13
humor			9 ^o	12
senhor			9 ^o	12
pior			10 ^o	10
treinador			10 ^o	10
embaixador			11 ^o	9
produtor			11 ^o	9

Quadro 35 – Apagamento de /r/ em não verbos mais frequentes ASPA/Amostra /r/ 2018

No que diz respeito a não verbos, 8 desses itens são idênticos (linhas sombreadas) nessas duas amostras; aqui, a comparação poderia se estabelecer entre os 13 itens mais frequentes de cada amostra.

Passemos agora à análise de verbos.

Vocábulos	Frequência ASPA	Ranking ASPA	Ranking Amostra /R/ 2018	Frequência Amostra /R/ 2018
ser	514408	1º	2º	120
ter	253046	2º	3º	100
fazer	165380	3º	1º	173
partir	99533	4º		
quer	97843	5º	5º	73
poder	74263	6º		
Dar	69635	7º	9º	38
estar	64700	8º		
Ver	59204	9º	7º	52
ficar	55638	10º	8º	46
dizer	54977	11º	6º	63
For	44823	12º		
saber	44723	13º		
pagar	41888	14º		
chegar	41233	15º	15º	23
falar			4º	89
ir			10º	37
Sair			11º	33
tocar			12º	29
voltar			13º	25
trabalhar			14º	24

Quadro 36 – Apagamento de /r/ em verbos mais frequentes ASPA/Amostra /r/ 2018

No que diz respeito a não verbos, 9 desses itens são idênticos (linhas sombreadas) nessas duas amostras.

O resultado dessas quatro tabelas, ainda que fruto de um recorte, aponta para a ideia de que a frequência lexical em uma língua, apesar de ser evidentemente influenciada pelas características do informante, da comunidade, do estilo e mesmo do tipo de entrevista utilizada para se obterem os dados, preserva uma certa uniformidade (DE BONA e SCHWINDT, 2017, p.)

Mais estudos precisam ser feitos para se obter uma análise mais acurada do papel do indivíduo e de sua relação com frequência lexical.

RESUMO DO CAPÍTULO: Neste capítulo, apresentamos nossas análises para os fenômenos de redução da nasalidade e apagamento de /r/ e sua relação com o papel da frequência lexical. Trouxemos, ao final, considerações sobre a comparação entre nossas amostras e nosso *corpus* de referência.

CONSIDERAÇÕES FINAIS

Este trabalho teve por objetivo analisar o papel da frequência lexical em dois fenômenos fonológicos variáveis que apresentam condicionamentos morfológicos, quais sejam redução da nasalidade e apagamento de /r/.

No capítulo 1, apresentamos nossa revisão de literatura sobre frequência lexical em fenômenos fonológicos variáveis condicionados morfológicamente, com a discussão de duas correntes de modelos teóricos: abstracionistas e exemplaristas. No capítulo 2, apresentamos trabalhos realizados no PB com foco na frequência lexical; também revisamos pesquisas sobre os dois fenômenos que são foco de análise deste trabalho. O capítulo 3 apresentou nossa metodologia de trabalho, e o capítulo 4 trouxe nossos resultados e análises. Apresentamos novamente nossos objetivos de trabalho, que constam na introdução, para contextualizá-los com os nossos resultados.

Objetivo 1. Verificar, concernente aos fenômenos descritos e analisados, qual seria a relação existente entre frequência lexical e fatores linguísticos, principalmente entre i) frequência lexical e morfologia interna à palavra e ii) frequência lexical e classe gramatical, ou seja, se fenômenos sujeitos à influência da frequência também podem estar sujeitos à influência morfológica e/ou morfossintática.

Redução da Nasalidade: Efeitos de frequência parecem ser diferentes para diferentes classes de palavra: para não verbos sem gem, frequência parece ter papel significativo na aplicação do fenômeno; para nomes com gem final, sendo raiz e sufixo, frequência do item lexical parece não ter papel da redução – o que defendemos é que a frequência do segmento gem pode estar atuando nestes casos, ou seja, não a frequência do item lexical, mas a frequência de tipo; para verbos, como analisado na última seção, frequência parece ter papel em dados de informantes de 51 anos ou mais no presente do indicativo, no pretérito do subjuntivo e no futuro do subjuntivo / infinitivo flexionado. No entanto, como não analisamos neste estudo o papel da expressão do sujeito e de sua proximidade com o verbo em questão, pouco podemos afirmar sobre essa interação.

Contexto precedente não se mostrou significativo na aplicação do fenômeno em nossa análise. Em uma tentativa de interação entre contexto precedente e frequência, não obtivemos interação significativa com nenhuma classe de palavra. No que diz respeito à morfologia interna à palavra, o fenômeno parece não enxergar estrutura morfológica interna, tendo em vista a análise de dados com *gem* na raiz e no sufixo. Interação com frequência lexical nos nomes terminados em *gem*, seja na raiz, seja no sufixo, não se mostrou significativa. Parece-nos que o que afeta a alta aplicação do fenômeno em nomes terminados por *gem* é a frequência do segmento final, ou seja, do tipo *gem*. No que diz respeito aos não verbos sem *gem*, quais sejam *nuvem*, *jovem*, *ordem*, *homem*, *ontem*, a interação é bastante significativa: quanto mais frequente, maior é a aplicação do fenômeno.

Os tempos verbais que se mostraram significativos em relação à frequência foram o presente do indicativo, o pretérito imperfeito do subjuntivo e o futuro do subjuntivo / infinitivo flexionado. Nestes casos, assim como nos não verbos sem *gem*, a correlação foi positiva: quanto mais frequente, maior a aplicação do fenômeno. Fazemos uma ressalva em relação a esse resultado, tendo em vista a necessidade de controle da variável expressão do sujeito.

Após análise dos dados, chegamos à conclusão de que podemos falar de efeitos de frequência nos dados do sul do Brasil ao analisarmos, separadamente, determinadas classes de palavra. Em fazendo isso, nas classes de palavra em que encontramos efeitos de frequência na aplicação do processo (não verbos sem *gem*, presente do indicativo, pretérito imperfeito do subjuntivo e futuro do subjuntivo/infinitivo flexionado), esse efeito é positivo: quanto mais frequente a palavra, maior é a aplicação do fenômeno. Ao analisar todos os dados conjuntamente, não conseguimos ver esse efeito.

Apagamento de /r/: Diferentemente do fenômeno de Redução da Nasalidade, em que há papel da frequência lexical em determinadas classes gramaticais, na análise de apagamento de /r/ não obtivemos resultados quanto à influência dessa variável com nenhuma classe gramatical analisada. No entanto, parece-nos que pode haver influência da frequência de tipo.

Com a terminação -or, por exemplo, temos poucos tipos de verbos, em geral *pôr* e seus derivados. Essa baixa frequência de tipo em -or parece influenciar na baixa

aplicação do processo nesses verbos (45,4%). No entanto, o segmento -or em não verbos apresenta uma taxa maior de aplicação do fenômeno (56,1%), pois, além de apresentar um número maior de dados, também é mais produtivo em PB. Podemos fazer uma relação semelhante com o contexto precedente em /e/: em verbos, temos uma alta taxa de aplicação do fenômeno (91,4%). Em não verbos, no entanto, esse percentual de aplicação cai para 16,2%, tendo em vista o número reduzido de types - apenas 5 types, com 37 tokens.

Ou seja, não temos efeitos de frequência de item lexical, mas temos um efeito de frequência de tipo, que opera tanto em verbos quanto em não verbos.

Frequência lexical, como já observamos não apresenta papel, por isso, não apresenta interação com as variáveis contexto precedente e morfologia interna à palavra. Entretanto, frequência de tipo parece exercer papel. Interessantemente, esse papel não interage com contexto precedente nem com morfologia interna à palavra – essas informações parecem não importar para a aplicação do fenômeno.

No que diz respeito à informação morfológica, as análises feitas na plataforma R apontaram essa variável como não significativa estatisticamente para a aplicação ou não do fenômeno, por isso parece que o processo não enxerga morfologia. Esse caso se assemelha ao caso dos nomes terminados em *gem*, analisados no fenômeno de redução da nasalidade, em que a aplicação do fenômeno também não enxerga estrutura interna morfológica: a aplicação do fenômeno de apagamento de /r/ em nomes com um pretense sufixo agentivo, tais como *professor*, *colaborador*, não parecem diferir de nomes como *humor*, *valor*; o mesmo se dá com o sufixo formador de adjetivos, como *espetacular*, *popular*, em que a aplicação do processo não parece diferir da aplicação em palavras com -ar na raiz, tais como *devagar*, *apesar*.

OBJETIVO 2: Averiguar a relação existente entre i) frequência lexical e fatores sociais, tais como idade, sexo e escolaridade; e ii) frequência lexical e informante. Nosso intuito é verificar se a influência da frequência está no indivíduo, com seu conjunto particular de itens léxicos, ou na comunidade, principalmente representada pelo fator social idade, tendo em vista que o léxico sofre influências geracionais.

Redução da Nasalidade: Podemos perceber que há interações bastante significativas entre frequência e idade, o que, de certa forma, corrobora a ideia de que léxico é geracional. Efeitos de frequência significativos foram encontrados em dados de informantes com mais de 51 anos nas seguintes classes de palavra: não verbos sem gem, presente do indicativo, pretérito imperfeito do subjuntivo, futuro do subjuntivo / infinitivo flexionado. Apesar de os informantes de 25 a 50 anos aplicarem mais o fenômeno, essa aplicação não está relacionada com efeitos de frequência, diferentemente do que acontece nos dados de informantes de 51 anos ou mais. Esse resultado pode ser o indício de uma mudança linguística. No que diz respeito à interação entre frequência lexical e escolaridade, encontramos interação significativa com a classe de palavra não verbos sem gem: quanto mais frequente, maior a aplicação do fenômeno tanto em falantes mais escolarizados quanto em falantes menos escolarizados. Também encontramos correlação positiva entre frequência lexical e aplicação do fenômeno nos dados de informantes de baixa escolaridade no pretérito imperfeito do subjuntivo. Não encontramos interação significativa de frequência lexical com outras classes de palavra.

Apagamento de /r/: Interações entre frequência e idade e frequência e sexo não se mostraram significativas, o que sugere que a variável frequência lexical, como demonstrado pelo nosso modelo inicial, não apresenta papel na aplicação do fenômeno de apagamento de /r/ final em nossa amostra.

OBJETIVO 3: Analisar quais modelos linguísticos, se abstracionistas ou se exemplaristas, apresentam maior adequação explanatória considerando os resultados dos fenômenos aqui apresentados.

Na linha de argumentação de Hanique e Ernestus (2012), que examinaram o papel da morfologia na redução acústica, parece que na análise de nossos dados também não temos evidência de que o status morfológico de um segmento afeta seu grau de redução. Parece-nos, portanto, que as palavras são armazenadas como unidades completas no léxico mental, sendo acessadas diretamente. Isso vai ao encontro de modelos exemplaristas, em que a estrutura morfológica não é altamente

relevante para a produção da fala, tendo em vista que é através da conexão entre palavras relacionadas que a estrutura morfológica emerge.

Além disso, a argumentação das autoras de que um segmento que identifica uma palavra (ou seja, a carga de informação da palavra) tem um papel mais determinante na redução que um segmento que identifica a estrutura morfológica de tal palavra também parece proceder para a análise de nossos dados.

Portanto, morfologia interna à palavra parece não importar para a aplicação do fenômeno. Regra fonológica relativa ao contexto precedente também não. Retomemos nossa análise para os dados de apagamento de /r/ com contexto precedente em /e/ para exemplificar melhor essa ideia. Apesar de os dados serem poucos, pode-se perceber que, no que tange a não verbos, /e/ em contexto precedente parece inibir o apagamento de /r/. Interessantemente, não podemos falar em regra fonológica, pois /e/ em contexto que antecede o /r/ em verbos, como vimos, tem altas taxas de aplicação do fenômeno. Um exemplo disso é a palavra *ser* – temos 119 ocorrências de *ser* como verbo, e o apagamento ocorre 95 vezes, ou seja, em 79,9% dos casos. Quando *ser* é usado como substantivo, apesar de o número de ocorrências ser baixo, percebe-se que há uma inibição do apagamento – não devido a alguma propriedade da vogal que antecede o /r/, mas em função da classe gramatical.

Tendo em vista as análises relativas aos nossos dados, defendemos neste trabalho a superioridade de abordagens exemplaristas frente a abordagens abstracionistas no que diz respeito à produção. Regras morfológicas e fonológicas não se mostraram determinantes na análise da variação dos fenômenos aqui estudados. Ou seja: os processos aqui analisados não enxergam estrutura interna morfológica, tampouco são passíveis de generalizações por meio de regras fonológicas.

No entanto, é importante ressaltar que a informação de classe de palavra é determinante para a aplicação dos fenômenos aqui analisados e, portanto, se faz relevante nas nuvens de exemplares.

Retomemos, agora, nossas hipóteses, no intuito de corroborá-las ou refutá-las, tendo em vista as análises aqui apresentadas:

HIPÓTESES

1. A frequência lexical apresenta correlação positiva com a aplicação dos fenômenos de redução e de apagamento, ou seja: quanto mais frequente o item lexical, mais suscetível à aplicação dos processos estudados ele é. No entanto, essa sujeição está condicionada à morfologia e à classe gramatical: morfemas serão mais resistentes à redução ou ao apagamento, mesmo em itens lexicais mais frequentes.

Essa hipótese se confirma em parte: frequência lexical apresenta correlação positiva com a aplicação dos fenômenos, estando condicionada à classe gramatical apenas, não à estrutura morfológica interna da palavra.

2. A relação entre frequência lexical e aspectos extragramaticais é significativa. A variável social que mostra mais interação com frequência é idade, tendo em vista que léxico é geracional.

Essa hipótese também se confirma em parte. Idade apresenta uma interação bastante significativa em dados de redução da nasalidade, principalmente na classe de não verbos sem *gem*. A mesma interação com idade não foi encontrada no fenômeno de apagamento de /r/. Uma explicação para isso talvez esteja relacionada com a caracterização de frequência de type e frequência de token. Algumas questões cujas respostas ainda precisam ser mais detalhadas na literatura são as seguintes: 1) qual é o impacto da frequência de type *versus* o impacto da frequência de token para variação/mudança linguística? 2) como melhor enxergar o efeito dessas duas frequências – por meio de pesquisas sincrônicas ou diacrônicas, preferencialmente?

Parece-nos que o efeito de frequência pode ser visto de forma mais precisa por meio de pesquisas diacrônicas. Quando se trata de frequência de item lexical, ou seja, frequência de token, parece que até conseguimos ver seu efeito sincronicamente, ao comparar itens de frequências lexicais distintas – como foi o caso de não verbos sem *gem*: *nuvem*, *jovem*, *ordem*, *homem*, *ontem*. Por outro lado, quando se trata de frequência de type, ou seja, frequência de determinada estrutura linguística, tais como os segmentos -*gem*, -*or*, -*er*, não é simples encontrarmos em análises sincrônicas um efeito de determinado item lexical, um pretense item mais frequente que tenha desencadeado o processo. Neste caso, apenas pesquisas diacrônicas podem nos dar

uma ideia do efeito da frequência lexical no processo. Essa ideia, de certa forma, é análoga à de Oliveira (1991), no que tange ao debate entre neogramáticos e difusionistas: podemos ter mudanças que, a longo prazo, pareçam neogramáticas, mas precisamos de análises diacrônicas para confirmar sua implementação via léxico.

3) Os modelos com maior adequação explanatória são os modelos baseados em regras, ou seja, os modelos abstracionistas, tendo em vista que, em decorrência da hipótese 1, frequência lexical é uma variável que, mesmo podendo exercer papel na aplicação dos fenômenos, não opera significativamente na presença de regras fonológicas e/ou morfológicas.

Essa foi uma hipótese que parece não ter sido corroborada pela nossa análise. Regras relativas à morfologia interna à palavra e à fonologia parecem não ser determinantes à aplicação dos fenômenos e, também, parecem não interagir com a variável frequência lexical.

Pesquisas futuras devem ampliar o número de dados para confirmar ou refutar as evidências aqui apresentadas. Especialmente, pesquisas diacrônicas sobre fenômenos que aparentam ser de aplicação categórica, além da análise de informante e de sua interação com frequência lexical, podem nos fornecer direcionamentos importantes acerca dos mecanismos que subjazem a variação e a mudança linguística.

REFERÊNCIAS

- BATTISTI, E. A redução dos ditongos nasais átonos. In: BISOL, Leda.; BRESCANCINI, Cláudia. (Org.). *Fonologia e Variação: recortes do português brasileiro*. Porto Alegre: EDIPUCRS, 2002. P. 183-202.
- BRESCANCINI, C. R.; MONARETTO, V. O.; Os róticos no Sul do Brasil: panorama e generalizações. *Signum. Estudos de Linguagem*, v. 11/2, p. 49-64, 2008.
- BYBEE, J. Morphology as lexical organization. *Theoretical morphology*, ed. by M. Hammond and M. Noonan, 119–41. San Diego, CA: Academic Press. 1988.
- BYBEE, J. Regular morphology and the lexicon. *Language and Cognitive Processes* 10.425–55. 1995.
- BYBEE, J. *Phonology and Language Use*. Cambridge: Cambridge University Press, 2001.
- BYBEE, J. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, 14: 261-290, 2002.
- BYBEE, J. From usage to grammar: the mind's response to repetition. *Language*, Washington, 82: 529-551, 2006.
- BYBEE, J. *Frequency of use and the organization of language*. New York: Oxford University Press, 2007.
- BOPP DA SILVA, T. *A redução da nasalidade em ditongos de sílaba átona em final de vocábulo entre falantes bilíngües e monolíngües do Rio Grande do Sul*. Porto Alegre: UFRGS (Dissertação de mestrado), 2005.
- CALLOU, D.; LEITE, Y. & MORAES, J. Variação e diferenciação dialetal: a pronúncia do /R/ no português do Brasil. In: Ingedore G. Villaça KOCH. Ed. *Gramática do Português Falado*. 2. ed. Campinas, S.P.: Editora da Unicamp. 463- 489, 1996.
- CALLOU, D. M. I; SERRA, C. R.; CUNHA, C. Mudança em curso no português brasileiro: o apagamento do r no dialeto nordestino. *Revista da ABRALIN*, v. 14, p. 195-219, 2015.
- CHAVES, R. G. *A redução/desnasalização de ditongos nasais átonos finais e a marcação explícita de CVP6*. Tese (Doutorado em Linguística). Universidade Federal de Santa Catarina: Florianópolis, 2017.

- CLEMENTS, G. N. The role of features in phonological inventories. In E. Raimy & C. Cairns (orgs.) *Contemporary Views on Architecture and Representations in Phonology*. Cambridge, Ma.: The MIT Press, 19-68, 2009.
- COETZEE, A. W. Phonological Variation and lexical frequency. In: Anisa Schardl, Martin Walkow & Muhammad Abdurrahman, eds. NELS 38. Volume 1. Amherst: GLSA, p. 189-202, 2008.
- CRISTÓFARO-SILVA, T. Difusão lexical: estudos de caso do português brasileiro. In: MENDES, E. A. M.; OLIVEIRA, O. M.; BENN-IBLER, V. (orgs.) *O Novo Milênio: interfaces linguísticas e literárias*. Faculdade de Letras. Belo Horizonte: 209–218, 2001.
- CRISTÓFARO SILVA, T; ALMEIDA, L. S., FRAGA, Thiago. ASPA: a Formulação de um Banco de Dados de Referência da Estrutura Sonora do Português Contemporâneo. In: XXV Congresso da Sociedade Brasileira de Computação, 2005, São Leopoldo. *Anais do XXV Congresso da Sociedade Brasileira de Computação* (CD-Room). São Leopoldo: Sociedade Brasileira de Computação, v. 1. p. 2268-2277. 2005.
- CRISTÓFARO-SILVA, T.; FONSECA, M; CANTONI, M. A redução do ditongo [ãw] postônico na morfologia verbal do português brasileiro: uma abordagem baseada no uso. *Letras de Hoje*, v. 47, p. 283-292. 2012.
- CRISTÓFARO-SILVA, T.; FONSECA, M.; CANTONI, M. Conflict in patterns of diffusion in diphthong reduction in Brazilian Portuguese. *Phonological studies*, v. 16. Tokyo: Kaitakusha Publishing Co. Ltd., p. 3-7. 2013.
- DE BONA, C. SCHWINDT, L. C. O papel da frequência lexical na desnasalização do ditongo final átono [~ein] em não verbos no português do sul do Brasil. *Cadernos do IL (UFRGS)*, v. 54, p. 27-46, 2017.
- ERNESTUS, M. Acoustic reduction and the roles of abstractions and exemplars in speech processing. *Lingua* 142, 27-41, 2014.
- GOMES, C. A.. Para além das ondas: um ponto de partida sobre o significado social da variação entre itongo nasal átono final e vogal oral no Português Brasileiro. *DIACRÍTICA* (BRAGA), v. 31, p. 5-24, 2017.
- GUY, G. R. *Linguistic variation in Brazilian Portuguese: aspects of the phonology, syntax and language history*. Universidade da Pennsylvania (Tese de doutorado). 1981.

- GUY, G. Contextual conditioning in variable lexical phonology. *Language Variation and Change*, 3: 223-239, 1991.
- GUY, G. R. Linking usage and grammar: Generative phonology, exemplar theory and variable rules. *Lingua*, 142: 57-65, 2014.
- HANIQUE, I. & ERNESTUS, M. The role of morphology in acoustic reduction. *Lingue e linguaggio* 11(2), 147-164, 2012.
- HAUPT, C. Contribuição da fonologia do uso e da teoria dos exemplares no estudo da monotongação. *Revista de Estudos da Linguagem*, v.19, p. 667 – 189, 2011.
- HUBACK, A. P. *Cancelamento do (R) final em nominais na cidade de Belo Horizonte: uma abordagem difusionista*. Dissertação de mestrado. Universidade Federal de Minas Gerais, 2003.
- HUBACK, A. P. Cancelamento do /r/ final em nominais: uma abordagem difusionista. *Scripta (PUCMG)*, v. 9, p. 13-30, 2006.
- HUBACK, A. P. A interferência da frequência em fenômenos linguísticos. *Revista DELTA*, v. 29, p. 79-94, 2013.
- LABOV, W. *The social stratification of English in New York*. Washignton: Center of Applied Linguistics, 1966.
- LABOV, W. Resolving the Neogrammarian controversy. *Language* 57.267–308. 1981.
- LOSIEWICZ, B. L. *The effect of frequency on linguistic morphology*. Dissertation. University of Texas, Austin, Texas, 1992.
- MENEZES, V. C. F. ; GOMES, C. A. . The Acquisition of Variable Coda (R) in the Speech Community of Rio de Janeiro. University of Pennsylvania. Working Papers in Linguistics (Online), v. 18, p. 58-64, 2012.
- MONARETTO, V. Análise sociolinguística da vibrante no sul do país. *Graphos - Revista da Pós Graduação em Letras/UFPB*, 1997.
- MONARETTO, V. A vibrante pós-vocálica em Porto Alegre. In: BISOL, Leda; BRESCANCINI, Cláudia (Org.). *Fonologia e variação: recortes do português brasileiro*. Porto Alegre: EDIPUCRS, p. 253-268, 2002.
- MONARETTO, L. C.; SCHWINDT, L. C.; DE BONA, C. O papel do léxico no apagamento da vibrante em final de palavra. Projeto discutido na PUCRS em 22/junho/2016.
- MYERS, J; GUY, G. R. Frequency effects in Variable Lexical Phonology. *University of Pennsylvania Working Papers in Linguistics*, 4: 215-227, 1997.

HUBACK, A. P. A interferência da frequência em fenômenos linguísticos. *Revista Delta*, 29:1, p. 79–94, 2013.

OLIVEIRA, M. A. de. The Neogrammarian Controversy Revisited. *International Journal Of The Sociology Of Language*, BERLIN - ALEMANHA, v. 89, n.1, p. 93-105, 1991.

OLIVEIRA, M. A. O léxico como controlador de mudanças sonoras. *Revista de Estudos da Linguagem*, Belo Horizonte: Faculdade de Letras da UFMG, v. 36, p. 75-92, 1995.

OLIVEIRA, M. A. Reanalizando o processo de cancelamento de (r) em final de sílaba. *Revista de Estudos da Linguagem*. Belo Horizonte: Faculdade de Letras da UFMG, nº2, p.32-58, 1997.

OUSHIRO, L.. Social and structural constraints in lectal cohesion. *Lingua* (Haarlem. Print), v. 172-3, p. 116-130, 2016.

PHILLIPS, B. Word frequency and the Actuation of Sound Change. *Language*, v. 60, n. 2: 320-342, 1984.

PHILLIPS, B. The mental lexicon: Evidence from lexical diffusion. *Brain and Language*, v. 68, p. 104-109. Academic Press: 1999.

PHILLIPS, B. Lexical diffusion, lexical frequency, and lexical analysis. In: BYBEE, Joan; HOPPER, Paul. (ed.). *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins, 2001.

PIERREHUMBERT, J. B. Exemplar dynamics: Word frequency, lenition, and contrast. In: BYBEE, J. L. and HOPPER, P. *Frequency and the Emergence of Linguistic Structure*. John Benjamins Publishing Company, 137-157, 2001.

PIERREHUMBERT, J. Probabilistic Phonology: discrimination and robustness. BOD, R., HAY, J., JANNEDY, S. (Ed.) *Probabilistic Linguistics*. Cambridge: MIT Press, 2003. p. 177-228.

SCHWINDT, L.C. Condicionamento morfológico em fenômenos fonológicos variáveis no português brasileiro. *Letras & Letras*. v. 28, n.1, p. 115-127, 2012.

SCHWINDT, L.C.; BOPP DA SILVA, T. Panorama da redução da nasalidade em ditongos átonos finais no português do sul do Brasil. In: BISOL, L.; COLLISCHONN, Gisela. (ed.). *Português do Brasil: variação fonológica*. 1 ed. Porto Alegre: EDIPUCRS, 2010.

- SCHWINDT, L.C.; BOPP DA SILVA, T.; QUADROS, E.S. O papel da morfologia na redução da nasalidade em ditongos finais no português do sul do Brasil. In: LEE, S-H (ED.) *Vogais além de Belo Horizonte*, UFMG, p. 355-365, 2012.
- SCHWINDT, L. C. O acesso à morfologia por processos fonológicos variáveis e a arquitetura da gramática. *Cadernos de Estudos Linguísticos* (UNICAMP), v. 56, p. 2343, 2014.
- SCHWINDT, L. C. Um output, dois processos. *Revista da ABRALIN*, v. 14, n. 1, p. 554-570, 2015.
- SCHWINDT, L. C. Sobre a preservação de expoentes morfológicos na fonologia variável do português brasileiro. *Domínios de Linguagem*. Universidade Federal de Uberlândia, 2016.
- SCHWINDT, L. C.; DE BONA, C. Lexical frequency effects on reduction of final nasal diphthongs in Brazilian Portuguese. *Revista Virtual de Estudos da Linguagem*. Ed. especial, n. 14, p. 168-189, 2017.
- SERRA, C.; CALLOU, D. A interrelação de fenômenos segmentais e prosódicos: confrontando três comunidades. 26 Textos Selecionados, XXVIII Encontro Nacional da Associação Portuguesa de Linguística, Coimbra, APL, pp. 585-594, 2013.
- SILVEIRA, G. *O apagamento da vibrante na fala do sul do Brasil sob a ótica da palavra. Dissertação de mestrado*. Universidade Federal do Rio Grande do Sul, 2010.
- VOTRE, S. J. *Aspectos da variação fonológica na fala do Rio de Janeiro*. Pontifícia Universidade Católica do Rio de Janeiro (tese de doutorado), 1978.
- WALKER, J. A. Form, function and frequency in phonological variation. *Language Variation and Change*, 24: 397-415, 2012.
- WANG, W. S-Y. Competing changes as a cause of residue. *Language*, v. 45, n.1, 1969.
- WANG and CHENG, C-C. Implementation of phonological change: the Shuang-Feng Chinese case. *In the lexicon in Phonological Change*. W. S.-Y. Wang (ed.). 148-158. The Hague: Mouton, 1977.

Anexo I

Detalhamento da Amostra 2 – Apagamento de /r/

- 8 Entrevistas;
- Tempo de duração médio: 26 minutos;

Idade	Sexo	Informante	Fonte	Ano	Link	Duração amostra
42	F	Rôssi Alves	202 Filmes	2014	https://www.youtube.com/watch?v=J1TFdbrGuWY&list=PL6hbyalZ5Hr9Ej20kctcGhZd6EX_83bp4V&index=29	Todo o vídeo (23:55)
45	F	Edd Wheeler	202 Filmes	2016	https://www.youtube.com/watch?v=nIRWc8VGPLs&t=135s	Todo o vídeo (26:55)
25	M	Maurício	202 Filmes	2013	https://www.youtube.com/watch?v=jYsxSeGVwP4&index=22&list=PL6hbyalZ5Hr9Ej20kctcGhZd6EX_83bp4V	05:00 - 30:00 (25:00)
35	M	DJ Tamenpi	202 Filmes	2016	https://www.youtube.com/watch?v=2ybwRlvwZXw&list=PL6hbyalZ5Hr9Ej20kctcGhZd6EX_83bp4V&index=38	Todo o vídeo (29:44)
55	F	Vera Sílvia Magalhães	Memória Política	2003	https://www.youtube.com/watch?v=q8fUe7vsj2s	15:00 – 45:00 (30:00)
91	F	Beatriz Ryff	Memória Política	2000	https://www.youtube.com/watch?v=7ftuYu4N_1w	10:00 – 40:00 (30:00)
51	M	Mauro César Pimentel	Comissão Nacional da Verdade	2014	https://www.youtube.com/watch?v=30oFxKjCV08	Todo o vídeo (22:30)
75	M	Célio Borja	Memória Política	2003	http://www2.camara.leg.br/a-camara/conheca/historia/historia/historiaoral/Memoria%20Politica/Depoimentos/celio-borja.html/ficha-tecnica.html	Todo o vídeo (32:26)

Sobre os vídeos:

- 202 Filmes: Os depoimentos foram selecionados de um documentário chamado “O som do tempo”, que tem como tema os 20 anos do rap carioca. Os depoimentos são de cunho pessoal e tem como tema principal a relação do entrevistado com o rap.

- Comissão Nacional da Verdade: Os vídeos dessa fonte são depoimentos de pessoas que viveram na época da ditadura e relatam suas experiências como presos políticos.
- Memória Política: Os vídeos selecionados são episódios de uma série produzida pela TV Câmara, em que foram coletados depoimentos de personalidades que se destacaram no cenário da história política recente do país. Assim como a fonte acima, os vídeos têm como característica o relato pessoal do entrevistado sobre suas experiências da época da ditadura.

Resumo dos entrevistados:

Rôssi Alves

- Profissão: Professora da UFRJ
- Formação acadêmica: Pós-doutorado
- Cidade natal: Rio de Janeiro (RJ)

Edd Wheeler

- Profissão: Rapper e Advogada
- Formação acadêmica: Pós-graduação em direito
- Cidade natal: Rio de Janeiro (RJ)

Maurício

- Profissão: Rapper, compositor
- Formação acadêmica: Não informado
- Cidade natal: Rio de Janeiro (RJ)

DJ Tamenpi

- Profissão: músico, pesquisador musical e jornalista
- Formação acadêmica: Graduação em Jornalismo
- Cidade natal: Rio de Janeiro (RJ)

Vera Sílvia Magalhães

- Profissão: economista e socióloga
- Formação acadêmica: Não informado
- Cidade natal: Rio de Janeiro (RJ)

Beatriz Ryff

- Profissão: Escritora
- Formação acadêmica: Não informado
- Cidade natal: Rio de Janeiro (RJ)

Mauro César Pimentel

- Profissão: Diretor de empresa.
- Formação acadêmica: Graduação em direito
- Cidade natal: Rio de Janeiro (RJ)

Célio Borja

- Profissão: advogado, professor, jurista e político
- Formação acadêmica: Ensino superior (várias graduações)
- Cidade natal: Rio de Janeiro (RJ)

ANEXO II

Lista de palavras analisadas na tese e número de ocorrências

Fenômeno 1 – Redução da Nasalidade

abandonaram	2
abrem	3
abriam	2
abriram	9
abrirem	1
acabam	3
acabaram	11
acabavam	1
acalmaram	2
acampam	1
acampavam	1
aceitam	3
aceitaram	2
aceitavam	2
acertaram	2
acertarem	1
acham	21
acharam	12
acharem	1
achassem	2
achavam	12
acomodavam	1
acompanham	1
acompanharam	1
acompanhavam	1
aconselhavam	1
acontecem	7
aconteceram	5
aconteciam	1
acordam	1
acordarem	1
acreditam	4
acreditavam	2
acusam	3
adaptaram	1
administram	1
admitem	2

adoram	4
adoraram	1
adoravam	3
adquiriam	1
afastaram	1
agarraram	1
agem	2
agradecem	1
ajoelharam	1
ajudam	11
ajudaram	2
ajudarem	1
ajudavam	7
alcançaram	2
alegam	2
almoçam	1
almoçavam	1
alugam	2
alugavam	1
amadurecem	2
amanheceram	1
amarram	1
amarravam	3
ameaçaram	1
analisam	1
analisem	1
andam	3
andaram	8
andarem	1
andassem	1
andavam	6
apagaram	2
apanharam	1
apanhavam	1
aparecem	5
apareceram	2
apareciam	3
apavoram	1
apelidaram	1
aperfeiçoaram	2
apertaram	1
apertavam	2
aplicam	2
apodreciam	1

apoiavam	1
apreciam	1
aprendem	3
aprenderam	4
aprenderem	1
aprendiam	2
aprendizagem	1
apresentam	3
apresentaram	1
aprisionam	1
aprontam	1
aprontaram	1
aprontavam	4
aproveitam	10
aproveitavam	1
aproveitem	3
armaram	2
arrancarem	1
arrancavam	1
arranjaram	1
arrasaram	1
arrecadavam	1
arrependem	1
arrombam	1
arrombaram	2
arrumam	4
arrumaram	1
arrumarem	1
arrumavam	3
assaltam	2
assaltaram	1
assam	1
asseguraram	1
assem	1
assinam	1
assinarem	2
assinassem	1
assistem	1
assistiam	1
assistiram	1
assistirem	1
assumiram	1
assustavam	1
atacam	2

atacaram	3
atacavam	2
atendem	10
atenderam	3
atingissem	1
atiram	1
atiraram	2
atiravam	2
atracavam	2
atrapalham	1
atravessaram	1
atravessavam	1
atropelaram	3
atualizaram	1
augmentam	2
autorizam	1
avançam	2
avisam	1
avisaram	2
bagagem	1
baixaram	1
baixassem	1
baixavam	1
barragem	7
barram	1
batam	1
batiam	3
batizaram	1
bebem	3
beberam	1
bebiã	1
beijam	1
bobagem	10
bobagens	1
boicotaram	1
botam	5
botaram	7
botavam	4
brigam	5
brigaram	7
brigarem	1
brigavam	5
brincam	1
brincarem	1

brincavam	2
buscavam	2
caem	4
caiam	2
caíssem	1
calaram	1
calçavam	1
caminharem	1
caminhavam	1
candidatam	1
candidatem	1
cansam	1
cantam	3
cantarem	2
cantavam	4
carregam	1
carregarem	1
carregavam	2
casam	4
casaram	30
casarem	1
casem	2
cederam	2
chamam	33
chamaram	5
chamarem	1
chamavam	19
chantagem	1
chegam	19
chegaram	26
chegarem	4
chegavam	6
cheguem	1
choravam	2
chutam	1
chutarem	1
cobram	3
cobravam	3
cobriam	1
colhem	2
colhiam	2
colocam	6
colocaram	10
colocavam	9

colonizaram	1
combateram	1
combinaram	2
combinavam	1
começam	13
começaram	31
começavam	10
comem	7
comemoram	1
comentaram	1
comentavam	1
comiam	7
competem	1
competiam	1
compram	9
compraram	8
comprarem	1
compravam	6
comprometiam	1
comprovaram	1
comungam	1
comunicam	1
concorrem	2
confeccionam	1
confessaram	1
confiam	1
confiaram	1
conhecem	16
conheceram	5
conhecerem	1
conheciam	4
conseguem	11
consequiam	2
consequiram	11
consequirem	1
conservam	4
consideram	1
consigam	2
constituíam	1
constituíssem	1
construíam	1
construíram	8
contagem	1
contam	10

contaram	8
contarem	3
contavam	4
contentam	1
continuam	9
continuaram	5
continuem	1
contornaram	1
contratam	4
contratavam	1
contribuem	1
contribuíram	2
controlam	1
convencem	2
conversam	1
conversaram	1
conversarem	2
convidaram	8
convidavam	2
convivem	1
conviveram	1
convocaram	1
convoquem	1
coragem	16
correm	4
correram	4
corresse	1
corretagem	1
corriam	4
cortam	2
cortaram	10
cortavam	5
costumam	4
cozinham	1
crecem	2
creceram	1
crecerem	2
criaram	8
criarem	1
criavam	3
criticam	2
cruzam	1
cuidam	2
cuidaram	2

cuidarem	1
cuidavam	7
cultivam	1
cumprimentam	1
cumpriram	1
dançam	16
dançaram	1
dançavam	4
darem	5
davam	34
decidiram	1
declamam	2
declararam	1
decoram	1
dedicam	1
dedicaram	1
dedicassem	1
dedicavam	1
defasagem	1
defendam	1
defenderam	1
degolaram	1
deixam	14
deixaram	12
deixavam	21
deleitavam	1
demonstram	1
demonstraram	1
demoraram	3
dependem	2
deram	27
derrubam	1
derrubaram	5
derrubavam	5
descarregavam	2
descobrem	1
descobriam	1
descobriram	2
desejam	1
desembarcaram	1
desenrolavam	1
desentenderam	2
desfilam	1
desgrudaram	1

deslocarem	1
desmancharam	1
desmataram	1
despejavam	2
dessem	3
destroem	1
desviarem	1
desvirtuaram	1
detonavam	1
deturpam	1
devam	1
devem	19
deveriam	6
deviam	5
devolveram	1
digam	2
diriam	1
discordam	1
discutem	3
discutiram	1
disputavam	1
disseram	21
disserem	1
divertem	7
dividem	1
dividiram	2
divirtam	1
divulgam	1
dizem	62
diziam	91
doam	1
doaram	1
dobravam	2
dominam	1
dominavam	2
dormem	3
dormiam	4
dragagem	4
elegeram	1
embarcaram	2
embarcavam	1
empatassem	1
emprestavam	1
empurram	1

encaminharem	1
encarecem	1
encarregam	1
encheram	2
enchiam	1
encontram	6
encontraram	5
encontravam	7
encostavam	8
enfermagem	10
enfrentaram	3
enfrentassem	1
engoliram	1
engrenagem	1
enroscaram	1
ensaíam	1
ensinam	5
ensinaram	5
ensinarem	1
ensinavam	3
entendam	1
entendem	8
entendiam	3
enterraram	1
entram	5
entraram	11
entrarem	1
entravam	4
entregam	1
entregavam	3
envolvem	1
envolviam	1
enxergam	1
enxerguem	1
eram	221
escolheram	1
escolhiam	3
escondiam	2
escrevem	3
escreveram	3
escreviam	1
escutaram	1
esforcem	1
espalharam	1

esperam	1
esperaram	1
esperavam	3
esperem	2
espionagem	1
esquecem	7
esqueceram	1
esqueciam	1
estabeleceram	3
estarem	1
estariam	1
estavam	104
estiveram	5
estiverem	1
estivessem	1
estouravam	1
estragaram	2
estudam	11
estudaram	15
estudavam	5
evitam	1
exigiam	1
exijam	2
existam	2
existem	28
existiam	5
existiram	2
existirem	1
existissem	1
explicaram	1
exploram	1
exportavam	1
expulsavam	1
extinguiram	1
fabricavam	1
façam	3
facilitavam	1
falam	56
falaram	21
falarem	3
falassem	3
falavam	26
faleceram	13
faltam	2

faltavam	1
fariam	1
fazem	249
fazerem	3
faziam	85
fecham	1
fecharam	3
fechavam	2
ficam	48
ficaram	37
ficarem	4
ficariam	1
ficassem	1
ficavam	17
fiquem	3
fizeram	83
fizessem	5
fogem	2
foram	188
forçaram	1
forem	2
formaram	3
formarem	1
formavam	1
forneçam	1
fossem	8
frequentam	1
frequentavam	2
fritam	1
fugiram	3
fumavam	2
funcionam	1
fundaram	1
furavam	1
fuzilaram	1
ganham	22
ganharam	6
ganharem	1
ganhavam	4
garagem	7
gastam	1
gostam	51
gostaram	5
gostavam	17

gravaram	1
gravem	1
grirem	1
guardaram	1
guardavam	1
hidro-	1
massagem	1
homem	98
homenagem	5
homens	11
hospedagem	1
hospedam	1
hospedavam	1
houveram	1
iam	87
imagem	3
imaginam	3
imaginavam	1
imitam	2
importam	1
impunham	1
incentivavam	1
incomodam	2
incomodaram	1
indicaram	2
induziram	1
infiltraram	1
influenciam	1
instalam	1
interessam	5
interviessem	1
invadem	2
invadiram	4
inventam	2
inventaram	6
investem	2
invistam	1
irem	3
jardinagem	1
jogam	17
jogaram	1
jogarem	1
jogavam	7
jovem	32
jovens	4

julgam	1
juntam	2
juntaram	3
juntavam	4
ladroagem	2
lamentaram	1
largam	3
largaram	3
largavam	2
lavavam	2
lembram	4
lembravam	1
lembrem	2
lerem	1
levam	10
levantam	2
levantaram	1
levantavam	1
levaram	28
levarem	1
levavam	11
levem	1
liam	2
liberavam	2
libertaram	1
lidam	2
ligam	6
ligavam	1
limitavam	1
limpavam	1
linguagem	1
linhagem	2
lotearam	1
malandragem	6
mandam	10
mandaram	9
mandarem	1
mandavam	7
mandem	1
manipularam	1
manterem	1
marcam	3
marcavam	1
margem	5

massagem	4
matam	8
mataram	11
matarem	3
matavam	11
medirem	1
melhoraram	2
melhorem	6
mencionaram	1
mensagem	10
merecem	1
mereceram	1
mereciam	2
metem	1
metragem	1
mexem	3
mexeram	2
modificam	1
modificaram	1
moem	1
molecagem	2
montam	1
montaram	6
montavam	2
moram	47
moraram	6
morassem	1
moravam	28
morem	1
morrem	1
morreram	26
mostram	1
mostraram	1
mostravam	1
mudam	3
mudaram	6
mudassem	1
mudavam	1
namoraram	1
namoravam	3
nasceram	3
necessitam	1
negavam	1
notam	2

notavam	1
nuvem	7
nuvens	5
obedecem	1
ofenderam	1
oferecem	1
olham	2
olhavam	1
olhem	2
omitiam	1
ontem	5
operaram	3
operarem	1
optam	1
ordem	17
organizam	2
organizaram	1
origem	37
origens	4
ouvem	1
ouviram	2
pagam	3
pagaram	1
pagarem	2
pagassem	1
pagavam	8
paisagem	1
param	1
pararam	8
paravam	3
parcelaram	2
pareciam	1
participam	10
participaram	2
participarem	1
participavam	3
passagem	21
passagem	21
passagens	1
passaram	8
passassem	1
passavam	8
pastam	1
patrocinam	1

pedem	8
pediam	2
pediram	8
pegam	22
pegaram	15
pegarem	1
pegassem	2
pegavam	16
peguem	2
pensam	14
pensaram	7
pensem	1
percebem	1
perceberem	1
percebessem	1
perdem	7
perderam	16
perderem	1
perdiam	1
perguntam	3
perguntavam	2
permaneceram	1
pernitiam	1
pernoitam	1
persequiam	3
personagem	1
pertencem	1
pescam	1
pilotagem	1
pintam	1
pintavam	2
pisam	1
pisavam	1
plantam	10
plantaram	3
plantavam	10
podem	29
poderiam	2
podiam	21
põem	2
portam	1
posam	2
possam	4
possibilitassem	1

poupam	1
pousavam	1
povoaram	1
praticam	1
praticaram	1
precisam	7
precisassem	1
precisavam	2
preferem	5
preferiam	1
prenderam	6
prendiam	3
preocupam	1
preocuparam	2
preocuparem	1
preparam	1
prepararam	1
preparavam	2
preservam	2
prestam	1
pretendem	1
previram	2
procuram	12
procuraram	4
procurarem	1
procurassem	1
procuravam	3
procurem	1
produzam	1
produzem	4
produzirem	1
prometem	6
promovem	3
provaram	1
providenciem	1
provocam	1
puderam	1
puderem	2
pulam	1
pulavam	1
puseram	5
puxam	4
puxaram	3
puxavam	4

quebraram	2
quebravam	1
queimaram	3
queiram	1
queixam	1
queixavam	1
querem	77
queriam	41
quiseram	8
quiserem	18
quisessem	1
rasgaram	1
realizassem	1
recebem	4
receberam	2
recebiam	1
reciclagem	6
reciclam	1
reclamam	2
reclamaram	1
reclamavam	2
recolhem	3
recuperam	1
recuperaram	1
regridem	1
remontagem	1
reparem	2
reportagem	4
representam	1
requeriam	1
requisitavam	3
residem	1
resistiam	1
resolvam	1
resolveram	1
respeitam	3
respeitavam	2
respondiam	2
retiraram	1
reúnem	5
reuniam	3
reviverem	1
revoltaram	1
rezam	1

rezavam	1
riem	1
riram	1
rodearam	1
roubam	12
roubaram	9
roubarem	1
roubavam	4
roubem	4
sabem	46
saberem	4
sabiam	18
sacrificavam	1
saem	18
saiam	2
saíam	9
saibam	2
saíram	9
saírem	2
saltavam	1
secagem	1
secam	1
seguem	2
seguiram	2
seguravam	1
sejam	6
selvagem	2
sentavam	1
sentem	3
sentiam	2
sentirem	2
separaram	2
serem	8
seriam	2
serragem	10
serraram	1
servem	4
sigam	1
sintonizavam	1
sobem	2
sobraram	1
socorreram	1
sofrem	2
sofreram	11

sofrerem	1
sofressem	1
sofriam	1
solicitavam	1
soltam	4
soltarem	1
soltavam	1
souberam	3
souberem	2
soubessem	1
subiram	2
sugam	1
sujeitam	1
surgem	3
surgiram	2
tecelagem	1
tecem	1
teimam	1
telefonam	1
telefonaram	1
tenham	4
tentaram	6
terem	3
teriam	4
terminam	1
terminaram	2
terraplanagem	1
tenham	153
tiram	7
tiraram	22
tirarem	1
tiravam	8
tirem	2
tiveram	24
tiverem	5
tivessem	12
tocaram	1
tocassem	1
tocavam	2
tomam	9
tomaram	4
tomarem	2
tomavam	6
tombaram	1

torçam	1
torram	1
torturaram	2
torturavam	2
trabalham	35
trabalharam	5
trabalharem	2
trabalhavam	11
trabalhem	2
trancam	1
trancavam	1
transformam	2
transformaram	1
tratam	4
tratavam	5
trazem	12
traziam	14
treinam	1
triagem	1
trocam	1
trocaram	3
trocavam	1
trouxeram	8
unem	1
uniam	1
unirem	1
urinam	1
usam	7
usavam	8
vadiagem	1
valorizam	1
valorizavam	1
vantagem	16
vejam	3
vençam	1
vendem	19
venderam	4
venderem	2
vendiam	7
veneram	2
venham	7
verifiquem	2
vestem	1
vestiram	1

viagem	45
viajam	1
viajavam	1
viam	4
vieram	109
viessem	1
vinham	37
viram	15
viraram	3
virem	6
virgem	3
virgens	1
visam	2
visitavam	1
vivam	1
vivem	18
vivessem	1
viviam	2
voltam	8
voltarem	5
voltavam	2
votaram	2
votarem	2
votem	1

Fenômeno 2 – Apagamento de /r/ final

abaixar	2
abrilhantar	1
abrir	6
absorver	1
acabar	1
aceitar	3
achar	5
acompanhar	3
acontecer	9
acordar	1
acreditar	5
acusar	1
adaptador	2
adequar	1

administrador	2
adorar	1
adotar	3
afirmar	1
agarrar	1
agir	2
agradar	1
agradecer	4
ajudar	6
alegrar	1
alimentar	1
almocar	2
alugar	2
amanhecer	1
amor	13
amortecedor	1
analisar	1
andar	5
anotar	1
anterior	3
apagador	1
apagar	1
aparecer	3
aperfeicoar	1
apertar	1
apesar	6
aplaudir	1
aprender	3
apresentador	1
apresentar	6
aprimorar	1
aprofundar	1
aproximar	1
ar	3
arrancar	1
arrebentar	1
arrumar	3
assaltar	1
assessor	2
assistir	4
atingir	2
atirador	1
ator	1
atrapalhar	1

atrelar	1
atuar	1
autor	2
avisar	2
baixar	2
bar	1
batalhar	2
bater	6
beber	1
beijar	6
bombar	2
botar	9
brincar	3
buscar	4
cair	4
calar	3
calor	1
caminhar	3
canalizar	1
cantar	8
cantarolar	1
cantor	2
caracterizar	1
casar	2
catar	1
categorizar	1
ceder	2
celular	1
cercar	1
chamar	7
chegar	23
chorar	2
cobrar	3
coibir	2
colar	5
coleccionar	1
colher	2
colocar	16
combinar	1
comecar	10
comer	3
comparar	1
compartilhar	1
competir	1

compor	1
compositor	4
comprar	4
computador	3
concentrar	2
conciliar	1
condicionar	2
condutor	1
conduzir	1
conferir	2
conhecer	14
conquistador	1
conseguir	11
construir	6
contar	8
continuar	5
controlar	1
conversar	12
convidar	2
conviver	3
cor	1
correr	5
cozinhar	1
crer	1
crescer	3
criar	7
criticar	3
cuidar	12
cultivar	1
cumprir	1
curtir	2
dancar	4
dar	38
debochar	1
declarar	1
dedicar	4
definir	2
deitar	1
deixar	11
delimitar	2
demorar	1
denominador	1
depende	1
derrubar	1

descansar	1
descer	2
descobridor	1
descobrir	5
desconstruir	1
descrever	1
descriminalizar	1
desfilar	1
desligar	1
deslocar	2
desmoralizar	1
despedir	2
despistar	1
desvirtuar	2
devagar	2
dialogar	2
diminuir	1
diretor	2
dirigir	3
discutir	3
disponibilizar	2
dispor	1
disputar	2
disser	2
distribuir	3
divergir	1
divertir	2
dividir	2
divulgador	2
divulgar	1
dizer	63
doar	1
dor	3
dormir	11
doutor	1
doutrinar	1
edificar	1
educar	1
elaborar	1
eleitor	1
elevador	1
emancipar	1
embaixador	9
emocionar	1

empreendedor	1
empregar	1
encaminhar	1
encarar	1
encomendar	1
encontrar	5
endurecer	1
enfiar	1
engordar	2
engravidar	1
enlouquecer	1
enroscar	1
ensinar	3
entender	11
enterrar	1
entrar	22
entregar	2
entrevistar	1
envolver	1
enxergar	1
equilibrar	1
errar	1
escolher	1
escrever	1
esculachar	1
esforçar	1
espairecer	1
espetacular	2
esquecer	3
estabelecer	1
estar	17
estender	1
estiver	2
estourar	1
estrear	2
estudar	8
evitar	1
evoluir	1
existir	2
experimentar	1
explicar	2
explodir	1
explorar	1
expor	3

expressar	2
exterior	1
falar	89
fator	2
favor	1
fazer	173
fechar	1
ficar	46
financiar	1
fingir	2
fizer	5
focar	1
for	12
forçar	1
formar	2
fortalecer	1
frear	1
frequentar	2
fugir	3
funcionar	4
ganhar	15
garantir	4
gargalhar	1
gastar	1
generalizar	1
gerir	1
gostar	6
governador	8
gravar	9
guardar	1
haver	2
honrar	1
horror	2
houver	2
humor	12
imaginar	3
implantar	1
impor	1
incentivador	1
incitar	1
inserir	1
instrumentalizar	1
interessar	1
interior	1

invadir	1
inventar	1
inventor	1
investigar	1
ir	37
jantar	1
jogador	13
jogar	4
juntar	2
lançar	5
lazer	2
legitimador	2
lembrar	4
ler	4
levantar	1
levar	10
liberar	2
ligar	1
localizar	1
lugar	26
magoar	1
maior	25
major	3
malhar	3
mandar	3
manipular	1
manter	7
mar	2
marcador	1
matar	7
melhor	31
melhorar	1
menor	8
microempreendedor	1
militar	13
modificar	1
montar	4
morar	6
morder	2
morrer	6
mostrar	7
motivador	1
motor	4
movimentar	2

mudar	14
mulher	48
namorar	5
nascer	1
negar	2
negociar	4
ocupar	2
oferecer	1
olhar	8
opressor	1
organizador	4
organizar	7
ouvir	9
pactuar	1
pagar	13
par	1
parar	5
particular	8
partir	17
passar	13
passear	1
patamar	4
patrocinador	1
pavor	1
pedir	3
pegar	21
pensar	11
perceber	6
perder	10
perguntar	5
permanecer	1
permitir	1
pesquisar	3
piar	1
pintar	2
pior	10
pirar	1
poder	40
popular	20
popularizar	2
por	2
prazer	8
precipitar	1
precisar	1

prejudicar	1
preocupar	2
preparador	1
preparar	1
preservar	1
processar	1
procurador	1
produtor	15
produzir	3
professor	6
profissionalizar	1
progredir	2
promotor	1
propor	2
proteger	1
prover	1
puder	4
pular	3
puxador	1
qualquer	46
quer	73
querer	12
quiser	8
radar	2
raptar	2
reagir	1
realizar	2
recair	2
receber	3
reclamar	2
reconhecer	5
recorrer	1
recusar	4
redor	1
reeducar	1
refletir	1
registrar	1
regular	2
reinvindicar	1
render	2
renunciar	4
repercutir	1
repetir	1
representar	1

reprodutor	1
resgatar	1
resolver	1
respeitar	4
responder	2
retomar	1
reunir	2
revezar	1
rimar	2
rir	6
rodar	1
rolar	1
roubar	1
saber	19
sacanear	1
sacrificar	1
sair	33
salvador	1
salvaguardar	1
seduzir	1
seguir	1
segurador	1
segurar	2
selecionar	1
senhor	12
sentar	5
sentir	5
sequer	1
ser	120
singular	3
soar	1
sobreviver	3
sofrer	2
souber	2
subestimar	1
subir	5
suceder	1
sucessor	1
suportar	1
surgir	1
surpreender	3
surtir	1
tacar	1
tentar	3

ter	100
terminar	2
testemunhar	1
tirar	11
titular	1
tiver	6
tocar	29
tomar	11
torcedor	2
tornar	1
trabalhador	1
trabalhar	24
tracar	1
traduzir	2
trair	1
transformar	2
transitar	2
travar	6
trazer	4
treinador	10
treinar	5
trocar	5
usar	8
vacilar	1
valor	3
vencedor	1
vencer	1
vender	7
ver	52
verificar	3
vestibular	1
vestir	2
viajar	9
vibrar	1
vir	2
virar	4
visitar	2
viver	13
voador	3
voltar	25
zoar	2