

**O PAPEL DA FREQUÊNCIA LEXICAL NA VARIAÇÃO  
FONOLÓGICA CONDICIONADA MORFOLOGICAMENTE:  
REVISITANDO ESTUDOS PRECEDENTES**

**THE ROLE OF THE LEXICAL FREQUENCY IN THE PHONOLOGICAL  
VARIATION CONDITIONED MORPHOLOGICALLY:  
REVISITING PREVIOUS STUDIES**

Camila De Bona\*

**Resumo:** Este artigo objetiva realizar uma revisão teórica dos estudos que apresentam a frequência lexical como fator relevante na variação fonológica, principalmente a condicionada morfológicamente. Para tanto, retomamos alguns dos principais expoentes teóricos acerca do estudo da frequência lexical, incluindo os modelos neogramático e difusionista, a Fonologia do Uso e a Teoria da Otimidade. Ao final, relatamos alguns estudos realizados no Português Brasileiro até então acerca dos efeitos de frequência.

**Palavras-chave:** Frequência lexical; Variação fonológica; Condicionamento morfológico.

**Abstract:** This paper aims to conduct a literature review of studies that have lexical frequency as a relevant factor in phonological variation, mainly the ones that are morphologically conditioned. To do so, we return to some of the leading exponents on the theoretical study of lexical frequency, including neogrammatical and diffusionist models, Phonology of Use and Optimality Theory. Finally, we report some studies in Brazilian Portuguese on the effects of frequency.

**Keywords:** Lexical frequency; Phonological variation; Morphological conditioning.

## **Introdução**

Tendo em mente a relevância de estudarmos a língua e seus falantes inseridos socioculturalmente, vemos que a homogeneidade linguística de longo prazo é uma ilusão desencadeada pela existência de um léxico e de uma morfofonologia relativamente estáveis. No entanto, parece haver um ajuste mais radical no que diz respeito, principalmente, à fonética (WEINRICH; LABOV; HERZOG, 2006). É no intuito de melhor

---

\* Doutoranda do Programa de Pós-Graduação em Letras da Universidade Federal do Rio Grande do Sul (UFRGS), na área de Teoria e Análise Linguística.

entender esse ajuste radical fonético, principalmente no que tange a palavras de maior frequência de uso, que tencionamos desenvolver este estudo.

Em termos mais amplos, através de diversos estudos já publicados em 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, apud VIGÁRIO, 2012) 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.

Este artigo tem por intuito realizar uma revisão teórica dos estudos que tomam a frequência lexical como fator importante na variação fonológica, principalmente a condicionada morfologicamente. Para tanto, na seção 1, apresentamos a justificativa do nosso estudo, respaldando-nos nas constatações de Guy (1991), Coetzee (2008) e Schwindt (2012). Nas seções seguintes, traçamos alguns dos principais expoentes teóricos acerca do estudo da frequência lexical, incluindo os modelos neogramático e difusionista, a Fonologia do Uso e a Teoria da Otimidade, incluindo a discussão acerca dos modelos baseados em regras comparativamente a dos baseados no uso. Dando continuidade, relatamos alguns estudos realizados no Português Brasileiro até então acerca dos efeitos de frequência, os quais são relativos a três pesquisadoras: Cristófarosilva (2001), Haupt (2011) e Huback (2013). Por fim, apresentamos algumas ressalvas no que diz respeito à metodologia de trabalho com frequência, com exemplos de Walker (2012) e Myers & Guy (1997).

## **1 Justificativa: variação fonológica, frequência lexical e Morfologia**

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).

Com o passar do tempo, muitos exemplos de variação foram descobertos, os quais dependiam, crucialmente, da morfologia. Essa constatação, pois, nos faz repensar as propostas teóricas que limitam a variação fonológica aos últimos estágios de implementação fonética.

Os estudos variacionistas percebem a heterogeneidade como parte do sistema, o que também fundamenta a inclusão de variáveis morfológicas nas suas análises. Essas variáveis, então, podem 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, ainda, 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 (ANTTILA, 1997; BOERSMA, 1998; BOERSMA & HAYES, 2001; COETZEE, 2006; LABOV, 1969, apud COETZEE, 2008). Apesar de esses modelos lidarem bem com a influência da gramática na aplicação de fenômenos variáveis, há, segundo Coetzee, 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 (BYBEE, 2002; COETZEE, 2008).

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, p. 2-3)

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 ambos esses fatores influenciam na variação e 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; BOPP DA SILVA, 2009) e vocalização da lateral pós-vocálica (COLLISCHONN; QUEDNAU, 2009).

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 parece acessar a distinção entre nomes e verbos, o que sustentaria 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 nas suas ocorrências, 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* parece ter sido 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 a constatação de Coetzee (2008) relativamente ao caráter imprescindível da projeção de modelos gramaticais que abarquem também as influências da frequência de uso lexical e tendo em mente as considerações de Bybee (2002) e Schwindt (2012), percebe-se que temos uma necessidade premente de ampliação e aprofundamento dos estudos acerca da frequência lexical e de 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. O trabalho aqui desenvolvido tem, pois, 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 no aprofundamento de nosso conhecimento acerca do papel da frequência lexical e da sua relação com a variação fonológica condicionada morfológicamente.

## **2 Efeitos de frequência: dos modelos neogramático e difusionista à Fonologia do Uso e à Teoria da Otimidade**

A primeira menção à questão dos efeitos de frequência em processos linguísticos foi feita por Schuchardt (1885, apud PHILLIPS, 1984, p. 321): “Rarely-used words drag behind; very frequently used ones hurry ahead. Exceptions to the sound laws are formed in both groups.” A importância dessa colocação reside no fato de que Schuchardt já havia atentado, há mais de cem anos, para o fato de que palavras que apresentam diferentes

faixas de frequência tendem a ser afetadas diferentemente no que tange à mudança linguística.

A pesquisadora Betty Phillips, em seu artigo intitulado *Word Frequency and the Actuation of Sound Change* (1984) defende 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 defende que fatores fonéticos não precisam ser necessariamente os únicos responsáveis pela mudança, mas a influência deles, certamente, precisa ser levada em conta 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)

Hoje, temos clara a noção da importante influência causada pelo fator frequência lexical nas mudanças sonoras motivadas fisiologicamente. No entanto, nem sempre foi assim, tendo em vista, principalmente, as concepções de mudança advogadas no modelo neogramático, o qual defende que o principal lócus da mudança não é a palavra, mas antes o som. Uma mudança sonora, segundo esse modelo, afeta, pois, todas as palavras que satisfaçam as condições estruturais que regem a implementação da mudança, ou seja, todas as palavras que apresentam ambiente ou contexto que condiciona a mudança sonora em questão. Disso, segue-se que as mudanças sonoras são foneticamente graduais e lexicalmente abruptas: todos os itens lexicais que estiverem sujeitos à aplicação de determinada regra serão modificados concomitantemente e da mesma forma.

Essas mudanças sonoras são condicionadas por determinadas regras fonológicas e são, portanto, regulares. Mudanças irregulares, quando houver, são explicadas apenas através de processos de empréstimo ou analogia. A proposta neogramática foi formulada com o propósito de ir de encontro às ideias dos dialetologistas, os quais

defendiam que cada palavra tem sua própria história. Com esse contraponto, vemos que a proposta dos neogramáticos apresenta um forte princípio metodológico, regendo a maioria dos trabalhos desenvolvidos em fonologia diacrônica (CRISTÓFARO-SILVA, 2001).

Já o modelo da Difusão Lexical, como proposto e difundido por Wang (1969, 1977, apud HUBACK, 2006) alega, diferentemente, que a palavra é a principal unidade da mudança e, como consequência, teremos mudanças sonoras lexicalmente graduais e foneticamente abruptas, já que determinados itens lexicais é que vão aderir gradativamente às mudanças. Com isso em vista, o modelo propõe que as mudanças não são mais condicionadas por regras fonológicas, mas por características lexicais, assim como a frequência e a previsibilidade de ocorrência e a familiaridade do item léxico em questão. Seriam, pois, esses aspectos extragramaticais os responsáveis pela maior ou menor vulnerabilidade da palavra à mudança sonora. Nessa perspectiva, uma mudança sonora ocorre inicialmente em alguma palavra e propaga-se para outras que apresentam estrutura fonológica semelhante. Pode haver casos em que a difusão lexical não atue em determinadas palavras, assim como há também a possibilidade de a alteração atingir todas as palavras potenciais. Irregularidades, nesse panorama, são justificadamente esperadas, já que a mudança sonora não precisa ser generalizada.

Mais recentemente, outra teoria que tem merecido destaque no estudo da frequência lexical é a Fonologia de Uso, preconizada por Joan Bybee, a qual advoga pela existência de um modelo de estocagem das palavras na mente dos falantes, apresentando 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 do modelo de Bybee está na

determinação das unidades de categorização e estocagem mental, quais sejam palavras e/ou frases inteiras. Não é através dos fonemas, portanto, que os indivíduos farão, por exemplo, generalizações morfológicas, mas através da associação fonética e semântica realizada entre os itens estocados no léxico mental.

A Fonologia do Uso visa a abarcar todos os subsistemas (fonologia, morfologia, sintaxe, semântica), não permanecendo apenas no 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 na substância da mesma. Com isso, a representação de um item lexical, constituída basicamente pela fonética e pela semântica, 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 artifícios 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 a palavra 'você' é registrada em determinado banco de dados; o número de casos encontrados será a frequência de ocorrência dessa palavra. Já a frequência de tipo corresponde à frequência de um determinado padrão ou estrutura linguística. Podemos procurar em certo corpus o número de palavras que apresenta o sufixo /-agem/, por exemplo; esse número será a frequência do conjunto desses nomes. 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, os mesmos 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)



Algumas consequências da interação das medidas de frequência apontadas também por Huback (2013) são as seguintes:

- Palavras com alta frequência de ocorrência estão mais ativas no léxico mental e, por isso, são acessadas mais rapidamente; já as 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. Diferentemente, palavras irregulares pouco frequentes não apresentam força suficiente para manter sua irregularidade; com isso em vista, essas palavras 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 forem 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. É a frequência de tipo, pois, a responsável por garantir a produtividade de seu grupo.

Em artigo de 2008, Coetzee aponta 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, apud COETZEE, 2008), 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 predições advêm da arquitetura básica do modelo de gramática por ele proposto e tem-se a oportunidade e a necessidade de, a partir de agora, testá-lo com outros fenômenos.

### 3 Modelos baseados em regras *versus* modelos baseados no uso

Seguindo a linha argumentativa de Coetzee (2008), em artigo publicado em 2014, Gregory Guy aponta a difícil conciliação entre 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) defende que a regularidade apontada pelos neogramáticos por meio de análise de registros históricos é bem instanciada, uma vez que a maioria das mudanças fonológicas não deixou resíduos históricos de segmentos que não sofreram o processo de mudança em palavras excepcionais. 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 o modelo ter uma boa adequação explanatória, o mesmo não tem a capacidade de dar conta de alguns tipos de fatos fonológicos, especialmente os que envolvem itens lexicais em específico. Para remediar esses defeitos, a teoria dos exemplares (PIERREHUMBERT, 2001; BYBEE, 2001) surge com o intuito de lidar com a variação, a gradiência, a difusão lexical na mudança sonora, os efeitos de frequência lexical, os contextos favorecedores e os efeitos probabilísticos, pois possui uma rica representação dos itens lexicais: os falantes retêm na memória os exemplares das palavras que eles já ouviram com alto detalhamento fonético. Essa pretensa nuvem de exemplares fornece ao falante 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.

Tendo em vista essas considerações baseadas no uso, Guy (2014), refletindo sobre as análises de dados por meio da teoria dos exemplares, levanta alguns possíveis impasses para o modelo teórico: 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?

Com isso em vista, o pesquisador sugere que uma análise alternativa, que incorpore tanto a produtividade gerativa quanto a precisão quantitativa, é o modelo de Regra Variável (doravante RV) da Sociolinguística, haja vista que o mesmo é capaz de preservar as vantagens dos modelos baseados em regras, tais como a abstração e a capacidade de representação dos processos categóricos. Ademais, o modelo de RV é capaz de resolver muitas das limitações desses modelos formais ao fazer uso da quantificação probabilística: qualquer processo ou restrição fonológica deve estar associado a uma probabilidade, o que permite o tratamento da variação e da gradiência.

Contudo, Guy (2014) assinala que um dos atributos fundamentais da teoria dos exemplares não é incorporado no modelo de RV, quais sejam os efeitos lexicais: frequência lexical, difusão lexical e contextos comuns. Para lidar, pois, com exceções lexicais de processos fonológicos variáveis, o autor defende a postulação de múltiplas representações subjacentes. *And*, por exemplo, é a palavra mais frequente do inglês com oclusiva coronal final, e a mesma apresenta o maior índice de apagamento desse traço: isso é facilmente explicado pela teoria dos exemplares. Analogamente, o modelo de RV também pode explicitar essa constatação por meio do pressuposto de que *and* apresenta múltiplas representações subjacentes: uma com e a outra sem o /d/ final.

Na conclusão de seu artigo, o pesquisador defende que a abstração fonológica não deve ser abandonada simplesmente porque a sua implementação tradicional não inclui quantificação: 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 (2014), é capaz de fornecer uma versão da fonologia mais equilibrada, a qual apresenta abstração e, ao mesmo tempo, também presta atenção nas quantidades.

Na próxima seção, relatamos alguns estudos realizados no Português Brasileiro até então acerca dos efeitos de frequência. Os mesmos são relativos, principalmente, a três pesquisadoras: Cristófaros-Silva (2001), Haupt (2011) e Huback (2013).

#### **4 Relato de alguns estudos já realizados no PB**

Em artigo de 2001, levando em conta a proposta da Difusão Lexical, Cristófaros-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 [tʃ] é preponderante. No caso em 3, a mudança ainda é incipiente. Cristófaros-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, tʃi] 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 [aI, eI, oI, uI] 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 e a palatalização de /S/ e apagamento de /tʃ/. 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.

A pesquisa de Huback acerca do cancelamento do (R) final em categorias nominais contou com entrevistas de 30 falantes de Belo Horizonte no ano de 2001. Nesse estudo, constatou-se que o /R/ final em nominais foi apagado em 22% dos casos e foi percebido que esse cancelamento final em nominais parece afetar as palavras mais frequentes primeiro. Pertinentemente, Huback aponta que o fato de o apagamento do (R) final ser mais frequente em verbos que em formas nominais (95,4% de apagamento em verbos (OLIVEIRA, 1997 apud HUBACK, 2013) e 22% de apagamento em nomes) já nos forneceria indícios sobre efeitos de frequência:

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. (HUBACK, 2013, p. 86).

Acerca do segundo fenômeno que afeta palavras mais frequentes, Cristóforo-Silva e Oliveira (2004 apud HUBACK, 2013) 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” – [deʃˈtʃinʊ], “castigo” – [kaʃˈtʃiɡʊ], 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 [deʃˈinʊ] e [kaʃˈiɡʊ]. 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 há uma tendência das palavras mais frequentes estarem mais propensas a fenômenos articulatórios 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.

## 5 Uma ressalva

Em artigo intitulado *Form, function and frequency in phonological variation*, o pesquisador James Walker (2012), seguindo o debate presente em Myers & Guy (1997), 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 primeiramente e morfológicas em segundo lugar, 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.

Outro caso interessante consta em Myers & Guy (1997), os quais analisaram o papel da frequência em verbos monomorfêmicos e regulares no passado, concluindo que o mesmo só influenciava a classe dos primeiros. De forma diversa, Bybee (1996 apud MYERS & GUY, 1997) apresenta resultados conflitantes ao tratar desses mesmos verbos, uma vez que suas análises mostraram que a frequência opera também nos verbos

regulares. Os pesquisadores, então, explicam a discrepância nos resultados: o dialeto examinado por Bybee apresenta taxas de apagamento muito mais altas que o examinado por eles; ademais, a pesquisadora restringiu sua análise apenas para os *tokens* regulares em ambientes não pré-vocálicos, haja vista que esse ambiente tende a favorecer o apagamento.

Por meio dessa constatação, juntamente com as ideias de Coetzee (2008) e Guy (2014), parece-nos premente a consideração prévia de vários fatores gramaticais que podem estar em jogo antes de nos atermos a questões de frequência, tendo em vista a possibilidade de obtermos resultados conflitantes acerca de um mesmo fenômeno, muitas vezes até em um mesmo *corpus*. Portanto, em pesquisas que obtiverem o fator frequência como determinante no que diz respeito a certos processos fonológicos, tais como apagamento ou assimilação, torna-se plausível e até necessária a realização de algumas reanálises de dados para averiguar se algum fator gramatical não figurou como negligenciado (como nas análises iniciais de Walker (2012)) ou tratado de forma diferenciada (como no caso de Bybee (1996)).

### **Palavras finais e desafios futuros**

Este artigo teve por objetivo fazer uma revisão teórica dos principais pressupostos e análises feitas com base na consideração do papel da frequência lexical. Para tanto, apresentamos os objetivos do nosso estudo, o tracejamento de alguns dos principais expoentes teóricos da frequência lexical e o relato de relevantes análises realizadas no PB acerca dos efeitos de frequência.

Alguns problemas enfrentados pelos pesquisadores que trabalham com esse tema residem, a nosso ver, em dois pontos principais: a) a falta de uma definição mais acurada acerca dos limites precisos para faixas de frequência e b) a ideia de que a medida de frequência tende a ser relativa a cada informante em particular. Por ora, medidas para definir limites de frequência baixa, média e alta devem ser elaboradas empiricamente, com base no fenômeno linguístico e no corpus em questão.

Bybee (2006) aponta que, com o passar do tempo e com a evolução dos estudos linguísticos nessa área, valores de frequência poderão ser definidos para diferentes fenômenos linguísticos, tendo em vista uma maior abrangência de análises. No que

concerne à frequência relativa a cada indivíduo, temos claro que cada falante, inserido em seu meio sociocultural, apresenta um conjunto de itens léxicos que usa com maior ou menor frequência. Com isso em mente, há uma premência em se realizar também estudos qualitativos nessa área, em que haja a eleição de um determinado número de falantes pertencentes a diferentes níveis de escolaridade para a averiguação, através de conjuntos lexicais particulares, da influência da frequência lexical na variação fonológica de indivíduos em específico. Por fim, tudo isso deve ser feito sem prescindir de análises exaustivas prévias acerca dos vários fatores gramaticais passíveis de estar em jogo.

## Referências

- BYBEE, J. *Phonology and Language Use*. Cambridge: Cambridge University Press, 2001.
- \_\_\_\_\_. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, v. 14, p. 261-290, 2002.
- \_\_\_\_\_. From usage to grammar: the mind's response to repetition. *Language*, v. 82, p. 529-551, 2006.
- \_\_\_\_\_. *Frequency of use and the organization of language*. New York: Oxford University Press, 2007.
- COETZEE, A. W. Phonological Variation and lexical frequency. In: Anisa Schardl, Martin Walkow & Muhammad Abdurrahman, eds. NELS 38. v. 1. Amherst: GLSA, p. 189-202, 2008.
- COLLISCHONN, G.; QUEDNAU, L. R. As laterais variáveis da região sul. In: BISOL, L.; COLLISCHONN, G. (Org.). *Português do Brasil: variação fonológica*. Porto Alegre: EDIPUCRS, 2010, p. 152-173.
- CRISTÓFARO-SILVA, T. Difusão lexical: estudos de caso do português brasileiro. In: MENDES, E. A. M.; OLIVEIRA, O. M.; BENN-IBLER, V. (Org.) *O Novo Milênio: interfaces linguísticas e literárias*. Faculdade de Letras. Belo Horizonte: 2001, p. 209-218.
- GUY, G. Contextual conditioning in variable lexical phonology. *Language Variation and Change*, v. 3, p. 223-239, 1991.
- \_\_\_\_\_. Linking usage and grammar: Generative phonology, exemplar theory and variable rules. *Lingua*, v. 142, p. 57-65, 2014.
- \_\_\_\_\_.; MYERS, J. Frequency effects in Variable Lexical Phonology. *University of Pennsylvania Working Papers in Linguistics*, v. 4, p. 215-227, 1997.

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.

HYMES, Dell. *Foundations in Sociolinguistics – An Ethnographic Approach*. University of Pennsylvania Press, Philadelphia, 1974.

HUBACK, A. P. Cancelamento do /r/ final em nominais: uma abordagem difusionista. *Scripta (PUCMG)*, v. 9, p. 13-30, 2006.

\_\_\_\_\_. A interferência da frequência em fenômenos linguísticos. *Revista Delta*, v. 29, n. 1, p. 79-94, 2013.

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

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, 2001, p. 137-157.

SCHWINDT, L. C. *A harmonia vocálica em dialetos do sul do país: uma análise variacionista*. 1995. 78f. Dissertação (Mestrado em Letras) – Curso de Pós-Graduação em Letras, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre.

\_\_\_\_\_. Condicionamento morfológico em fenômenos fonológicos variáveis do português brasileiro. *Revista Letras & Letras, Uberlândia*. v. 28, n. 1, p. 115-127, 2012.

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

VIGARIO, M.; FROTA, S.; MARTINS, F.; CRUZ, M. Frequência na Fonologia do Português: recursos e aplicações. In: COSTA, A.; DUARTE, I. (eds.) *Nada na linguagem lhe é estranho. Estudos em homenagem a Isabel Hub Faria*. Porto: Edições Afrontamento, 2012, p. 613-631.

WALKER, J. A. Form, function and frequency in phonological variation. *Language Variation and Change*, v. 24, p. 397-415, 2012.

WEINREICH, U; LABOV, W.; HERZOG, M. *Fundamentos empíricos para uma mudança linguística*. Tradução: Marcos Bagno; revisão técnica Carlos Alberto Faraco. São Paulo: Parábola, 2006.

Recebido em junho de 2014.

Aceito em outubro de 2014.