

A INTERFERÊNCIA DA FREQUÊNCIA EM FENÔMENOS LINGUÍSTICOS (The interface between frequency and linguistic processes)

Ana Paula HUBACK (Columbia University)

Resumo: *Este artigo discute efeitos de frequência em alguns casos de variações fonológicas e morfológicas do português brasileiro. O Modelo de Redes (Bybee 1985, 2001, 2010) foi adotado como foco teórico. Quatro fenômenos linguísticos do português brasileiro foram comparados na tentativa de entender que tipos de processos linguísticos afetam as palavras mais frequentes ou as menos frequentes primeiro. Conclui-se que, nos casos analisados, variações de fundo analógico parecem afetar as palavras menos frequentes primeiro, ao passo que reduções fonéticas parecem afetar as palavras mais frequentes primeiro.*

Palavras-chave: *Efeitos de frequência, Modelo de Redes, variação e mudança linguística*

Abstract: *This article discusses frequency effects in some Brazilian Portuguese phonological and morphological variations. The theoretical background adopted is the Network Model (Bybee 1985, 2001, 2010). Different linguistic phenomena are compared to analyze which linguistic processes affect the most or the least frequent words first. As a conclusion, it is observed that variations based on analogy seem to affect the least frequent words first, whereas phonetic reductions seem to affect the most frequent words first.*

Key words: *Frequency effects, Network Model, language variation and change*

INTRODUÇÃO

Neste artigo, vamos discutir o papel da frequência em processos de variação e mudança linguística recorrentes na língua falada. O quadro teórico adotado é o Modelo de Redes (Bybee 1985, 2001, 2010). Segundo Bybee (2010), a repetição de estruturas linguísticas (sejam elas palavras ou expressões inteiras) tem efeitos cognitivos em nosso “dicionário mental”. Em linhas gerais, palavras que usamos muito são lembradas com mais facilidade, ao passo que palavras menos usadas são recuperadas com mais dificuldade no léxico mental e podem, inclusive, ser esquecidas.

Neste artigo, vamos analisar como efeitos de frequência atuam em processos linguísticos que ocorrem na fala. Basicamente, nossos objetivos são:

- Apresentar uma resenha sobre a literatura linguística que descreve o papel da frequência nos fenômenos linguísticos;
- Discutir que tipos de fenômenos linguísticos podem afetar palavras mais frequentes ou menos frequentes primeiro;
- Apresentar exemplos do português brasileiro (doravante PB) para discutir efeitos de frequência nas variações linguísticas;
- Debater como o uso da língua tem impacto nos processos cognitivos de estocagem e acesso a informações linguísticas.

Definidos os objetivos deste artigo, vamos passar para a subseção seguinte, que discute a literatura sobre efeitos de frequência nos fenômenos linguísticos.

A LITERATURA SOBRE EFEITOS DE FREQUÊNCIA

A primeira menção sobre o papel da frequência em processos linguísticos é feita por Schuchardt (1885:58): “Rarely-used words drag behind; very frequently used ones hurry ahead. Exceptions to the sound laws are formed in both groups.” Esse comentário é bastante importante, visto que, há mais de um século, o autor já percebia que palavras de diferentes faixas de frequência poderiam ser afetadas de forma distinta pela mudança linguística. No entanto, conforme veremos posteriormente, o papel da frequência em fenômenos linguísticos não pode ser simplista e afirmar que palavras raras resistem a variações, ao passo que as mais frequentes são as mais avançadas nesse processo. É necessário haver um refinamento maior sobre os efeitos de frequência.

Leslau (1969), em um estudo sobre as línguas da Etiópia, também observou efeitos de frequência. O autor analisou processos fonéticos como aspiração, elisão, assimilação, etc. e percebeu que esses processos afetaram as palavras mais frequentes dentro da comunidade de fala investigada.

Fidelholtz (1975) analisa a redução vocálica do inglês e observa que as palavras mais frequentes são mais suscetíveis a esse fenômeno. Outro ponto

que o autor menciona é que a familiaridade da palavra também determina se (ou como) ela será afetada por fenômenos linguísticos. Segundo ele, por exemplo, o topônimo “Manhattan” é mais reduzido por falantes da própria ilha. Finalizando sua abordagem, o autor menciona que, se uma palavra é usada mais frequentemente, será mais acessível no léxico mental e, conseqüentemente, mudará mais.

Pesquisas posteriores às de Schuchardt (1885), Leslau (1969) e Fiedelholz (1975), apresentam maiores detalhamentos sobre a análise da frequência.

Phillips (1984) introduz a hipótese de que mudanças fisiologicamente motivadas afetam as palavras mais frequentes primeiro, enquanto que mudanças não-fisiologicamente motivadas afetam as palavras menos frequentes primeiro. A terminologia “fisiologicamente motivadas” refere-se a fenômenos como redução vocálica, apagamento e assimilação, ou seja, mudanças baseadas na fisiologia dos gestos articulatórios. A autora ressalva que, com essa hipótese, não pressupõe que somente fatores fonéticos geram a mudança, mas salienta que a influência da superfície fonética é imprescindível para tais mudanças. A análise de Phillips é importante porque menciona que os efeitos de frequência não são uniformes (conforme se imaginava anteriormente). Ao contrário disso, dependendo do tipo de variação linguística, as pioneiras no processo serão palavras mais ou menos frequentes.

Pagliuca e Mowrey (1987) e Browman e Goldstein (1992) apresentam evidências para o fato de que palavras muito usadas são mais suscetíveis a variações que implicam na redução de segmentos. Segundo os autores, os gestos articulatórios envolvidos na fala tornam-se mais automatizados por causa da repetição. Sendo assim, existe redução e sobreposição de gestos articulatórios, o que faz com que palavras mais frequentes sofram mais redução fonética.

Lindblom (1990) apresenta duas tendências competitivas que interferem na produção de sons pelo falante: a) a tendência de utilizar pouco esforço articulatório; b) a necessidade de produzir um estímulo auditivo que seja suficientemente claro para que o interlocutor possa compreender as sentenças pronunciadas. O equilíbrio entre essas duas tendências faria com que o falante reduzisse os segmentos de palavras ou expressões nos casos em que ele (o falante) soubesse que não gerariam dúvidas de compreensão

para o interlocutor. Nesse sentido, palavras muito frequentes já estão mais disponíveis no léxico mental do falante, por isso a economia de gestos articulatórios atuaria mais sobre elas. A consequência disso é que palavras mais frequentes sofreriam mais redução fonética. Considerando que ocorrências individuais dos itens léxicos (sejam eles palavras ou expressões inteiras) são estocadas na memória, a redução fonética vai ser estocada juntamente com a palavra e categorizada como um exemplar. Conforme a redução for sendo repetida, o exemplar reduzido pode se tornar mais forte que o não reduzido e, futuramente, pode haver uma mudança linguística.

Bybee (2002) propõe que o conceito de frequência seja dividido em duas medidas diferentes:

- **Frequência de ocorrência** (“token frequency”) indica quantas vezes uma determinada palavra ocorre em um corpus linguístico. Procuramos, por exemplo, quantas vezes a palavra “livros” é registrada na Amostra do Banco de Português versão 1.0¹ e encontramos 129 casos. Essa é, portanto, a frequência de ocorrência do item “livros” nesse corpus;
- **Frequência de conjunto de formas** (“type frequency”, ou frequência de tipo) indica a quantidade de itens da língua que contém uma determinada estrutura. No PB, por exemplo, o Dicionário Eletrônico Houaiss da Língua Portuguesa² registra 13.012 verbos terminados em [-ar]. Essa é, então, a frequência do conjunto dos verbos da primeira conjugação do PB.

Essas duas medidas de frequência interagem e geram consequências cognitivas e linguísticas:

- Palavras com alta frequência de ocorrência têm memória mais ativa e podem ser acessadas de forma independente no léxico mental; palavras com baixa frequência de ocorrência dependem de conexões com outros itens lexicais para ser ativadas;

1. A Amostra do Banco de Português versão 1.0, disponível em <<http://lael.pucsp.br/corpora/index.htm>>, contém 1.182.994 dados oriundos de fala e escrita.

2. O Dicionário Eletrônico Houaiss da Língua Portuguesa (2001) apresenta aproximadamente 228.500 verbetes.

- Palavras irregulares e frequentes têm memória mais forte no léxico mental, portanto costumam resistir a mudanças analógicas; por outro lado, palavras irregulares e infrequentes não são suficientemente reforçadas no léxico mental para manter sua irregularidade, então podem adotar os paradigmas mais frequentes da língua;
- Palavras derivadas e de frequência de ocorrência baixa ou média formam conexões mais fortes do que palavras derivadas e de alta frequência de ocorrência. Isso ocorre porque os itens infrequentes dependem de sua classe para ser lembrados, portanto, a cada vez que são acessados, reforçam a coesão de sua rede; por outro lado, itens derivados altamente frequentes ganham autonomia lexical e sua ativação no léxico mental não é feita necessariamente através de sua classe. O acesso a itens derivados e frequentes não reforça a conexão da classe como um todo. Portanto, é a frequência de tipo, não a de ocorrência, que garante a produtividade de uma classe.

A frequência, portanto, parece exercer papel importante nos fenômenos linguísticos. Mesmo assim, ainda existe muito a ser investigado e discutido sobre esse tema. Uma das questões controversas nesse ponto é a definição de limites precisos para faixas de frequência. Bybee (2006) afirma que não há, até o presente momento, medidas precisas para definir os limites entre frequência baixa, média e alta. Segundo a autora, essa é uma questão empírica e valores categóricos de frequência para cada fenômeno linguístico serão definidos conforme as análises linguísticas evoluírem.

Uma outra questão a ser discutida é que medida de frequência pode, também, ser relativa a cada informante. Um grupo de falantes, ainda que estes pertençam à mesma comunidade de fala, tem experiências linguísticas diferentes, por isso não se pode afirmar que uma certa medida de frequência seja válida para todos os indivíduos.

A questão da familiaridade da palavra também parece ser uma referência bastante individual. Considerando a rede de contatos e as diferentes atividades que cada indivíduo desempenha, cada pessoa terá sua lista de itens mais ou menos familiares ou frequentes de acordo com sua realidade social. Outra questão a considerar é que toda medida de frequência de ocorrência é sempre relativa ao corpus que se está adotando. Existe uma dicotomia básica entre língua falada e normalmente se considera que a fala é mais formal e a escrita, mais informal. No entanto, essas duas caracterizações

(fala/escrita = formal/informal) refletem, na verdade, dois polos de um mesmo continuum. Existem discursos orais que são mais formais do que os escritos e vice-versa. No uso da língua em situações reais, a dicotomia fala/escrita não pode ser categoricamente considerada.

É importante distinguir, também, os conceitos de familiaridade e frequência. A familiaridade de um item nem sempre significa que ele será mais frequente na língua ou em um corpus específico. Objetos domésticos, como mesa, cadeira, cama, etc., são muito familiares para todos nós. No entanto, a frequência desses itens é provavelmente baixa na maioria dos corpora do PB (mesmo de língua falada). Isso significa que esses dois conceitos (frequência e familiaridade) interagem, mas não são sinônimos.

A fim de investigar como a frequência afeta fenômenos do PB, vamos analisar algumas variações linguísticas e verificar quais delas afetam as palavras mais frequentes ou as menos frequentes primeiro.

FENÔMENOS QUE AFETAM PALAVRAS MAIS FREQUENTES PRIMEIRO

Tomando-se como base as pesquisas de Phillips (1984), Pagliuca e Mowrey (1987) e Browman e Goldstein (1992), as variações que afetam as palavras mais frequentes primeiro podem ser entendidas como motivadas pela fisiologia da fala e podem ocorrer em função da redução e sobreposição de gestos articulatórios. Vamos verificar se essa hipótese se confirma em casos concretos do PB.

Fenômeno 1 – Apagamento do (R) final em formas nominais

O cancelamento do (R) final em verbos é um fenômeno bastante estudado na literatura linguística do PB (cf.: Oliveira 1983, 1997). Oliveira (1997), a partir de entrevistas com 50 falantes de Belo Horizonte, observa que, nos infinitivos verbais, a taxa de cancelamento do (R) final é praticamente categórica: 95,4%. Com essa porcentagem, podemos afirmar que o apagamento do (R) final em verbos é muito mais frequente do que sua realização, que só se dá em 4,6% dos casos.

O cancelamento do (R) final em categorias nominais não foi tão estudado como o (R) em verbos. A fim de determinar que fatores linguísticos

afetavam esse fenômeno, foram feitas, em 2001, entrevistas com 30 falantes de Belo Horizonte³. Nessa pesquisa, constatou-se que o (R) final em nominais foi apagado em 22% dos casos. A tabela abaixo lista as palavras líderes em cancelamento do (R) em nominais e sua frequência no Corpus NILC/São Carlos⁴.

Tabela 1: Líderes no cancelamento do (r) e a sua frequência no Corpus NILC/São Carlos

| Itens | Cancelamento % (número) | Frequência NILC/S. Carlos por milhão de palavras |
|---------------|----------------------------|---|
| 1. a partir | 91% (11/12) | 265.735 |
| 2. qualquer | 70% (28/40) | 331.738 |
| 3. particular | 48% (13/27) | 58.2 |
| 4. computador | 45% (16/35) | 70.7 |
| 5. apesar | 41% (10/24) | 90.518 |
| 6. lugar | 37% (34/89) | 246.537 |
| 7. mulher | 31% (23/74) | 192.758 |
| 8. melhor | 25% (23/93) | 351.195 |

Na tabela 1 *não* há uma correlação categórica entre frequência de ocorrência e cancelamento do (R), no sentido de palavras mais frequentes serem as líderes no fenômeno de apagamento. No entanto, é importante mencionar que os líderes no apagamento do (R) são palavras altamente frequentes no Corpus NILC/São Carlos. Além disso, alguns outros dados dessa tabela são dignos de nota: a) Primeiramente, três dos oito itens mais propensos ao cancelamento do (R) são conectivos (“a partir”, “qualquer”, “apesar”). Por serem itens gramaticais, não têm referência lexical fixa e, portanto, podem sofrer redução fonética sem que o sentido do enunciado se perca; b) Esses mesmos itens são muito frequentes em construções pré-estabelecidas na língua, como “a partir de”, “qualquer um”, “apesar de”, etc. O fato de serem muito presentes nesses “chunks” também pode causar automatização dos gestos articulatórios e conseqüente apagamento do (R) final; 3) Por fim, devemos comentar os itens “particular” e “computador”.

3. Para maiores informações sobre a metodologia adotada para o trabalho de campo, consulte-se Huback (2003).

4. O Corpus NILC/São Carlos, disponível em <<http://www.linguateca.pt/CETEMPublico/>>, contém um total de 41.372.943 dados de língua escrita.

Apesar de sua frequência de ocorrência relativamente mais baixa, também foram líderes no cancelamento do (R). A presença de sufixos ({-ar} e {-dor}) é uma similaridade entre esses itens. Tais sufixos são bastante frequentes nas classe de nominais em (R), por isso a frequência do sufixo em si já é alta e talvez isso faça com que esses itens sejam mais propensos ao cancelamento do (R).

Em linhas gerais, considerados os comentários acima, percebemos que o cancelamento do (R) final em nominais parece afetar as palavras mais frequentes primeiro. Um último ponto a mencionar sobre esse fenômeno é o fato de que o apagamento do (R) final é mais frequente em verbos que em formas nominais (95,4% de apagamento em verbos e 22% de apagamento em nomes). Essa informação já nos fornece 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.

A seguir vamos analisar outro fenômeno linguístico que afeta as palavras mais frequentes primeiro.

Fenômeno 2 – Palatalização de /s/ e apagamento de {tʃ}

Cristófaros-Silva e Oliveira (2004) fizeram uma análise sobre o apagamento de {tʃ} na cidade de Belo Horizonte. Na pronúncia padrão de Belo Horizonte, 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” – [deʃtʃinʊ], “castigo” – [kaʃtʃigʊ], “ginástica” – [ʒiˈnaʃtikɐ].

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 “destino” [deʃinʊ], “castigo” – [kaʃigʊ], “ginástica” – [ʒiˈnaʃikɐ]. As autoras realizaram um experimento para analisar melhor esse fenômeno. Quinze palavras com a sequência *-sti-* foram selecionadas, e divididas de

acordo com a frequência de ocorrência⁵. Nove falantes do PB foram entrevistados. Os resultados de apagamento da sequência *-sti-* são apresentados na tabela abaixo:

Tabela 2: Efeito da frequência de ocorrência no apagamento da africada [tʃ]

| Faixas de frequência | Dados Obtidos | % | Peso relativo |
|----------------------|---------------|----|---------------|
| alta | 25/44 | 57 | 0.73 |
| média | 15/36 | 42 | 0.48 |
| baixa | 20/45 | 44 | 0.29 |

Apesar do baixo número de dados coletados, esta tabela sugere que as palavras mais frequentes 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.

A seguir vamos mencionar dois fenômenos linguísticos que afetam as palavras menos frequentes primeiro. Depois dessa análise, vamos elaborar comentários mais gerais sobre todos os fenômenos linguísticos analisados.

FENÔMENOS QUE AFETAM PALAVRAS MENOS FREQUENTES PRIMEIRO

Para ilustrar fenômenos linguísticos que afetam palavras de baixa frequência de ocorrência, vamos citar dois casos de plural do PB: o plural das palavras terminadas em [-ão] no singular (“leão”, “pão”) e o plural das palavras terminadas em ditongo em [-u] (“céu”, “museu”). A fim de analisar variações nesses dois grupos de plurais, experimentos psicolinguísticos foram feitos com 36 falantes do PB. Os dados foram coletados em 2006,

5. A frequência foi verificada no CRPC (Corpus de Referência do Português Contemporâneo), que contém 4.646.737 palavras.

no estado do Rio de Janeiro⁶. O programa SPSS foi utilizado para fazer a regressão logística e selecionar as variáveis significativas para a ocorrência dos dois fenômenos linguísticos analisados.

Fenômeno 3 – Plural das palavras terminadas em {-ão} no singular

Em linhas gerais, esse fenômeno refere-se ao fato de que, no PB, palavras que terminam em {-ão} no singular apresentam três formas distintas de pluralização: {-ões} (“balão” – “balões”), {-ãos} (“irmão” – “irmãos”), {-ães} (“capitão” – “capitães”). No latim, essas formas eram diferentes no singular e no plural, portanto não havia problema quanto à forma de plural que se aplicava a cada tipo de palavra. No entanto, no PB atual, como todas as formas de singular são 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”. Fazendo uma busca pela frequência de tipo dessas terminações no Dicionário Eletrônico Houaiss, encontramos os seguintes resultados:

Tabela 3: Frequência de tipo de {-ões}, {-ãos} e {-ães} no Dicionário Eletrônico Houaiss

| Típos | Número | % no grupo {-ão} singular | % no dicionário |
|--------|--------|---------------------------|-----------------|
| {-ões} | 7.260 | 97,8 | 3,17 |
| {-ãos} | 108 | 1,5 | 0,047 |
| {-ães} | 48 | 0,7 | 0,021 |
| TOTAL | 7.416 | 100 | 3,24 |

Com base no Dicionário Houaiss, o morfema {-ões} apresenta frequência de tipo muito mais alta que as demais terminações. Por causa desse diferencial, nossa hipótese é 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

6. Detalhes sobre a metodologia para a coleta de dados podem ser encontrados em Huback (2007).

{-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}). A tabela abaixo demonstra os resultados dos experimentos⁷:

Tabela 4: Efeito da frequência de ocorrência nos plurais em {-ão} no singular

| Faixas de frequência | Dados Obtidos | % | Peso Relativo |
|----------------------|---------------|------|---------------|
| Baixa | 131/171 | 76.6 | 0.995 |
| Média | 36/227 | 15.9 | 0.419 |
| Alta | 9/210 | 4.3 | 0.019 |

Total de dados = 608, $\chi^2 = 270$, $df = 2$, $p = 0.000$

Os dados demonstram a tendência de que palavras mais frequentes resistem mais à variação analógica que palavras pouco frequentes. Vamos analisar o segundo grupo de plurais para depois tecer comentários gerais sobre os fenômenos do PB que afetam as palavras menos frequentes primeiro.

Fenômeno 4 – Plural das palavras terminadas em ditongo em {-u} no singular

Neste grupo, estamos lidando primeiro com o fato de que o [-l] em posição de coda é vocalizado na maioria dos dialetos do PB. Sendo assim, não existe diferença fonética entre as letras L e U nos itens a seguir: “carnal” / “berimba”, “confortável” / “museu”. A partir desse fenômeno, as

7. A regressão logística do SPSS selecionou como relevantes os fatores plural etimológico, frequência de ocorrência e número de sílabas. Como este artigo discute apenas efeitos de frequência, as demais variáveis não serão comentadas.

classes de palavras terminadas em ditongo em [-u] e em [-l] começaram a mostrar variações linguísticas. Itens como “degrais” e “troféis” (em vez de “degrus” e “troféus”) já são encontrados em corpora do PB. Em termos de frequência de tipo, o Dicionário Houaiss apresenta os seguintes dados para as duas classes:

Tabela 5: Frequência de tipo dos itens terminados em [-l] e em ditongo em [-u]

| Grupos | Número | % no dicionário |
|------------|--------|-----------------|
| TOTAL [-U] | 741 | 0,32 |
| TOTAL [-L] | 6.225 | 2,72 |

A tabela 5 mostra 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] podem ser 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]. Com relação à frequência de ocorrência, também se espera que palavras mais frequentes resistam à variação, por sua representação autônoma no léxico mental. Por outro lado, as palavras de baixa frequência podem ser a primeiras a sofrer variações⁸. Vejamos os resultados na tabela a seguir:

8. Na variação de plural para ditongo em [-u], o SPSS considerou relevantes as variáveis número de sílabas, frequência de ocorrência e vogal precedente. No entanto, como este artigo discute apenas efeitos de frequência, os demais fatores não serão comentados.

Tabela 6: Efeito da frequência de ocorrência nos plurais terminados em ditongo em [-u]

| Frequência | Dados Obtidos | % | Peso Relativo |
|------------|---------------|------|---------------|
| Baixa | 112/339 | 33 | 0.609 |
| Média | 75/327 | 22,9 | 0.502 |
| Alta | 6/288 | 2,1 | 0.370 |

Total de dados = 954, $\chi^2 = 94.7$, $df = 2$, $p = 0.000$

A tabela 6 apresenta resultados análogos à tabela 4: palavras de alta frequência resistiram à variação linguística, ao passo que itens menos frequentes foram os primeiros a mudar.

Na subseção seguinte vamos traçar comentários gerais sobre os fenômenos analisados.

ANÁLISE GERAL DOS FENÔMENOS INVESTIGADOS

A partir da análise dos efeitos de frequência nas variações linguísticas discutidas, percebemos que determinados fenômenos afetam palavras mais frequentes primeiro, enquanto outros afetam itens menos frequentes primeiro.

Observamos que as palavras mais frequentes foram afetadas em dois fenômenos que implicam em redução fonética. A frequência de ocorrência faz com que alguns itens se tornem previsíveis no discurso. Por causa disso, quando são utilizados, seus segmentos não precisam ser pronunciados de forma plena para que sejam reconhecidos pelos interlocutores. Por sua vez, a frequência de ocorrência também faz com que os itens sejam pronunciados mais rapidamente. Uma consequência disso é a sobreposição e a consequente redução dos gestos articulatórios. A automatização, portanto, parece responder pelos processos que afetam as palavras mais frequentes primeiro. Considerando que cada ocorrência de uma palavra (mesmo que idêntica) é estocada na memória, exemplares com redução fonética acabam se tornando mais fortes do que exemplares com os segmentos pronunciados de forma plena. A mudança linguística pode ser a consequência futura desse processo cognitivo.

Com relação aos dois processos que afetaram as palavras menos frequentes primeiro, observamos que são fenômenos decorrentes da analogia

entre classes de plurais. Devido ao nivelamento das diferenças no singular (todas as formas analisadas terminam analisadas em [-ão] ou em [w]), formou-se um ambiente propício para que efeitos de frequência pudessem atuar. Sendo os grupos em [-ões] e [-l] mais frequentes que os demais, acabam atraindo membros das outras classes. A frequência de tipo é, portanto, imprescindível para que a analogia possa atuar. Observamos que as palavras mais frequentes são menos afetadas pelos fenômenos, justamente porque sua memória lexical é autônoma e elas não precisam do acesso à rede para que sejam utilizadas. Por outro lado, os itens de baixa frequência de ocorrência são os primeiros a mudar porque não têm memória lexical suficiente para que sejam lembrados de forma independente. Os itens de baixa frequência geralmente são acessados através de sua rede e, como essa é de baixa frequência de tipo, não garante que os itens sejam pluralizados na forma etimológica.

A seguir vamos apresentar as conclusões deste artigo.

CONCLUSÕES

Os fenômenos analisados neste artigo nos permitem delinear algumas conclusões. Primeiramente, a hipótese de Phillips (1984) parece ser corroborada pelos dados, uma vez que fenômenos que são relacionados à fisiologia da fala afetaram as palavras mais frequentes primeiro, ao passo que processos baseados em analogia afetaram as palavras menos frequentes primeiro. É importante notar, também, que os falantes, aparentemente, são capazes de armazenar detalhamento sobre padrões probabilísticos disponíveis no léxico. Isso pode ser demonstrado através do fato de que os falantes usaram o plural em [-ões] mais que os de [-ãos] e [-ães], provavelmente devido à alta frequência de tipo do primeiro. Se os falantes não tivessem conhecimento refinado de dados probabilísticos do léxico do português, provavelmente os experimentos não teriam mostrado variabilidade nos plurais em [-ão] (no singular). Finalmente, os processos cognitivos de categorização e acesso a informações linguísticas parece interferir na variação e mudança, fazendo com que algumas palavras (ou padrões linguísticos) sejam preservadas e outras se tornem obsoletas.

Recebido em fevereiro de 2011

Aprovado em outubro de 2012

E-mail: aph2129@columbia.edu

REFERÊNCIAS

- Amostra do Banco de Português versão 1.0. Disponível em <<http://lael.pucsp.br/corpora/index.htm>>).
- BROWMAN, Catherine & GOLDSTEIN, Louis. 1992. Articulatory phonology: An overview. *Phonetica*. 49 (3-4): 155-80.
- BYBEE, Joan. 1985. *Morphology: A Study of the Relation between Meaning and Form*. Philadelphia: John Benjamins.
- _____. 2001. *Phonology and Language Use*. Cambridge: Cambridge University Press.
- _____. 2002. Mechanisms of change in grammaticization: the role of frequency. In: Brian D. Joseph & Richard D. Janda. Eds.. *Handbook of Historical Linguistics*. Oxford: Blackwell.
- _____. 2006. From usage to grammar: the mind's response to repetition. *Language*, Washington, n. 82(4), p. 529-551.
- _____. 2010. *Language, Usage and Cognition*. New York: Cambridge University Press.
- Corpus de Referência do Português Contemporâneo. Disponível em <http://www.clul.ul.pt/sectores/linguistica_de_corpus/projecto_crpc.php>.
- Corpus NILC/São Carlos. Disponível em <www.linguateca.pt/ACDC/>.
- CRISTÓFARO-SILVA, Thaís & OLIVEIRA, Daniela Mara. 2004. Efeitos da frequência na produção de fricativas alveopalatais emergentes. In: 51º SEMINÁRIO DO GEL (Grupo de Estudos Linguísticos do Estado de São Paulo), Taubaté. 1-7.
- Dicionário Eletrônico Houaiss da Língua Portuguesa. 2001. Rio de Janeiro: Objetiva.
- FIDELHOLTZ, James L. 1975. Word frequency and vowel reduction in English. In: REGIONAL MEETING CHICAGO LINGUISTICS SOCIETY, 7, 1975, Chicago. Papers from... Chicago: Chicago Linguistic Society. p. 200-213.
- HUBACK, Ana Paula. 2003. Cancelamento do (R) final em nominais na cidade de Belo Horizonte: uma abordagem difusionista. (Mestrado em Linguística) – Universidade Federal de Minas Gerais.
- _____. 2007. Efeitos de frequência nas representações mentais. (Mestrado em Linguística) – Universidade Federal de Minas Gerais.

- LESLAU, Wolf. 1969. Frequency and change in the Ethiopian languages. *Word* 25: 180-189.
- LINDBLOM, Björn. 1990. Explaining phonetic variation: a sketch of the H&H theory. In: William Hardcastle & Alain Marchal. Eds.. *Speech Production and Speech Modeling*. Dordrecht: Kluwer.
- OLIVEIRA, Marco Antônio de. 1983. Phonological variation and change in Brazilian Portuguese: the case of the liquids. (Doutorado em Linguística) – University of Pennsylvania, Philadelphia.
- _____. 1997. Reanalizando o processo de cancelamento do (r) em final de sílaba. *Revista de Estudos da Linguagem* 6 (nº 2): 31-58.
- PAGLIUCA, William and Richard MOWREY. 1987. Articulatory evolution. In A. G. Ramat, O. Carruba and G. Bernini Eds.. *Papers from the 7th International Conference on Historical Linguistics*. Amsterdam: John Benjamins. p. 459-472.
- PHILLIPS, Betty. 1984. Word frequency and the actuation of sound change. *Language* 60 (n. 2): 320-42.
- SCHUCHARDT, Hugo. 1972 [1885]. On sound laws: against the neogrammarians. In: Schuchardt, the neogrammarians, and the transformational theory of phonological change. Frankfurt: Athenaeum.