

Plurais em *-ão* do português brasileiro: efeitos de frequência

Ana Paula Huback (Columbia University)*

Resumo: Este artigo investiga palavras do português terminadas em *-ão* no singular, como *avião*, *capitão* e *cidadão*. A forma singular dessas palavras é igual, mas, no plural, cada palavra apresenta uma terminação diferente: *avião* – *aviões*, *capitão* – *capitães*, *cidadão* – *cidadãos*. Etimologicamente, cada palavra deveria ter somente uma forma de plural. No entanto, em corpora do português, é possível encontrar formas como “*capitões*” e “*cidadões*”, em vez de “*capitães*” e “*cidadãos*”.

Neste artigo, argumenta-se que a frequência de uso das formas linguísticas pode justificar variações neste grupo de palavras. O foco teórico adotado é o Modelo de Redes, proposto por Bybee (1985, 2001).

Palavras-chave: Modelo de Redes, estocagem lexical, pluralização.

Introdução

Neste artigo, investigamos o grupo de palavras do português brasileiro (PB) terminado em *-ão* no singular e com pluralização em *-ões*, *-ãos* ou *-ães*. O artigo descreve um teste de reação no qual falantes nativos do PB foram questionados sobre o plural desses itens. Os objetivos dessa pesquisa podem ser sumarizados da seguinte forma:

- Verificar se a frequência de uso da palavra interfere no tipo de plural que os falantes adotam;
- Encontrar indícios que possam elucidar se a unidade de estocagem no léxico mental dos falantes é a palavra inteira ou morfemas individuais;
- Analisar se o uso das palavras no cotidiano tem alguma interferência na forma como os falantes organizam informações linguísticas no léxico mental;
- Observar se os falantes preservam informações detalhadas sobre a estrutura das palavras analisadas.

Nos parágrafos subsequentes, vamos apresentar mais informações sobre o tema analisado nesta pesquisa.

* Department of Spanish and Portuguese.

Em nossa vida cotidiana, a frequência com que realizamos algumas atividades determina o grau de automatização delas. Rotinas como dirigir um carro ou amarrar o cadarço de um sapato são tão comuns no nosso dia-a-dia que podemos realizá-las sem raciocinar como elas devem ser feitas. No entanto, se lembrarmos como aprendemos a dirigir, vamos perceber que, no início, era necessário pensar ativamente em cada movimento e era difícil coordenar os movimentos simultâneos de pés e mãos. Aos poucos, com a prática e a repetição, dirigir um carro torna-se uma atividade totalmente automática. Enquanto dirigimos, não precisamos pensar ativamente no que estamos fazendo e podemos, inclusive, dirigir e realizar outras atividades ao mesmo tempo, como conversar, falar ao telefone ou cantar.

Em termos linguísticos, a repetição de itens lexicais também tem suas consequências. Em linhas gerais, podemos afirmar que palavras usadas com frequência são facilmente memoráveis. Por outro lado, palavras pouco usadas são recuperadas com mais dificuldade no nosso “dicionário mental”. É bastante comum que pessoas tenham um “lapso de memória” quando tentam se lembrar de palavras que usam pouco. Pode-se afirmar, grosso modo, que a frequência de uso da palavra é um dos fatores determinantes para que um item léxico seja lembrado ou esquecido.

A discussão sobre como a frequência de uso de palavras e/ou expressões afeta processos de variação e mudança é bastante explorada na literatura linguística. Provavelmente, a primeira referência a efeitos de frequência foi feita por Schuchardt (1885, p. 58). Esse autor afirma que palavras pouco usadas são mais resistentes à mudança, ao passo que palavras muito frequentes avançam mais rapidamente em processos de mudança. No entanto, o autor menciona também que exceções podem ser encontradas nos dois grupos de palavras (frequentes e não frequentes).

Leslau (1969), estudando as línguas etíopes, analisa mudanças fonéticas que afetam as palavras mais frequentes primeiro e as menos frequentes depois. Em sua pesquisa, o autor observa que, nas línguas investigadas, houve várias mudanças fonéticas, como aspiração, elisão, assimilação, etc., e todas elas afetaram as palavras mais frequentes dentro daquela realidade cultural: “comer”, “Deus”, “dizer”, “fazer”, “igreja”, “poder”, “saber”, “vestir”.

Fidelholtz (1975) aponta o fator frequência de ocorrência como determinante da suscetibilidade à mudança sonora. Analisando a redução vocálica no inglês, o autor observa uma correlação entre familiaridade da palavra e estocagem lexical.

Uma palavra familiar é muito usada, o que a tornará mais acessível no léxico mental e, conseqüentemente, fará com que seja mais suscetível a mudanças sonoras. Segundo o autor, as palavras frequentes e as infrequentes são estocadas em áreas distintas do cérebro, por isso exibem diferente propensão à mudança.

Phillips (1984) postula que a mudança sonora afeta alguns itens lexicais antes de outros e, além disso, a mudança não opera somente com base no condicionamento fonético. Para Phillips (1984), as mudanças que atingem as palavras mais frequentes primeiro incluem redução vocálica, apagamento e assimilação, ou seja, são mudanças baseadas na fisiologia da fala. A autora ressalva que, com essa afirmação, não está pressupondo que somente fatores fonéticos geram a mudança, mas salienta que a influência da superfície fonética é imprescindível para tais mudanças. Posteriormente, com relação às mudanças que afetam as palavras menos frequentes primeiro, a autora afirma que estas são mudanças não-fisiologicamente motivadas. Então, a hipótese de Phillips (1984) para efeitos de frequência na estrutura linguística pode ser resumida da seguinte forma: mudanças fisiologicamente motivadas afetam as palavras mais frequentes primeiro, enquanto que mudanças não-fisiologicamente motivadas afetam as palavras menos frequentes primeiro.

A análise de Phillips (1984) é importante porque a autora estabelece uma correlação entre frequência de uso e tipos de mudanças. Isso indica que, dependendo da natureza do fenômeno, palavras mais frequentes ou menos frequentes serão afetadas primeiro.

Bybee (1985, 1995, 2001) discute exaustivamente como a frequência de uso das palavras interfere na forma como elas são organizadas e acessadas no léxico mental dos falantes. Bybee (1985, 1995, 2001) propõe um modelo de léxico mental intitulado “Modelo de Redes” (“Network Model”). De acordo com esse modelo, a frequência de uso determina se a palavra será ou não guardada no dicionário mental do falante. Se a palavra for muito frequente, será armazenada; se não for, será recuperada através da associação com outros itens lexicais. Não é, portanto, a regularidade ou irregularidade da palavra que define se ela será armazenada no léxico mental. A frequência de uso é mais relevante nesse aspecto.

Bybee (2002) afirma que a frequência deve ser medida de duas maneiras distintas:

- Frequência de ocorrência (“token frequency”): indica quantas vezes uma determinada palavra é repetida em um certo corpus de uma língua. Se acessarmos o site www.google.com.br e digitarmos a palavra “gato”,

teremos um número superior a 30.900.000¹ resultados. Essa, então, será a frequência de ocorrência da palavra “gato” em toda a internet;

- Frequência de tipo (“type frequency”): indica quantas vezes um determinado padrão linguístico se repete no dicionário da língua. Consideremos, por exemplo, o sufixo *-vel* no português. Se consultarmos o Dicionário Eletrônico Houaiss da Língua Portuguesa (2001), veremos que 2.065 palavras do português apresentam esse sufixo. Esse número será, portanto, a frequência de tipo do sufixo *-vel* no português.

Essas duas abordagens de frequência apresentam as seguintes consequências:

a. Palavras com alta frequência de ocorrência têm representação lexical mais forte e, por isso, são acessadas mais rapidamente que palavras com baixa frequência de ocorrência;

b. Palavras com alta frequência de ocorrência podem ser acessadas de forma independente, porque a memória delas é mais ativa; por outro lado, palavras com baixa frequência de ocorrência dependem de conexões com outros itens lexicais para serem ativadas;

c. A manutenção de uma palavra com flexão irregular depende de sua frequência de ocorrência. Palavras irregulares e frequentes têm representação lexical mais forte, portanto costumam resistir a mudanças analógicas²; por outro lado, palavras irregulares e infrequentes não são suficientemente reforçadas no léxico mental para que sua irregularidade se mantenha, por isso é comum que adotem os paradigmas mais frequentes da língua;

d. 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 acontece 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 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 de ocorrência, que garante a produtividade de uma classe;

¹ Busca realizada em 25 de outubro de 2009, às 2h09min.

² Segundo Bybee (2001), palavras irregulares e de alta frequência de ocorrência resistem mais a mudanças analógicas (adoção de um paradigma diferente do de sua classe, a partir do efeito da analogia). No entanto, a repetição dos itens léxicos gera automatização e sobreposição dos gestos articulatórios, o que faz com que palavras altamente frequentes sejam mais propensas a reduções fonéticas em geral (Cf. Pagliuca; Mowrey, 1987, e Browman; Goldstein, 1992).

e. Como decorrência do item d acima, os padrões flexionais encontrados em palavras altamente frequentes não tendem a se espalhar para outros itens lexicais, porque o acesso a tais itens não reforça a rede; por outro lado, as flexões de itens com frequência baixa e média podem, mais facilmente, ser adotadas por novas palavras, já que o acesso a palavras pouco frequentes é crucial para a força lexical da rede: “[...] if a class contains both high-frequency and medium-frequency items, it is the medium-frequency items that contribute most to the productivity of the class, and to the establishment of the parameters that define the class.” (Bybee, 1985, p. 134).

Na subseção seguinte, vamos analisar as palavras que terminam em *-ão* no português.

O caso das palavras em *-ão* no PB

Neste artigo, vamos analisar como efeitos de frequência afetam um grupo específico de palavras do PB: os itens terminados em *-ão* no singular. Nesta subseção, vamos apresentar informações sobre essas palavras.

Em latim, as formas que evoluíram para palavras com singular *-ão* no português apresentavam terminações diferentes tanto para o singular quanto para o plural. O quadro abaixo, extraído de Sequeira (1943), ilustra essas formas:

Singular	Plural	Exemplos	
-anem	-anes	panem	panes
-anum	-anos	germanum	germanos
-anum	-anus	manum	manus
-onem	-ones	leonem	leones
-onum	-onos	patronum	patronos
-udinem	-udines	certitudinem	certitudines

Quadro 1: Formas de singular e plural existentes no latim

As formas acima continuaram sem alterações aproximadamente até o século XIV. Depois disso, uma série de mudanças fez com que todas as formas singulares se fundissem no que hoje é a forma *-ão* do português. As mudanças que desencadearam a fusão dessas formas singulares podem ser sumarizadas da seguinte maneira: a) As terminações relativas aos casos latinos se perderam por

causa do desenvolvimento gradual das preposições; b) Posteriormente, a nasalidade do fonema /n/ foi assimilada pela vogal precedente, fazendo com que o /n/ sofresse síncope; c) Mais tarde, as vogais contíguas se fundiram e acabaram se transformando em *-ão*.

No português atual, o grupo de palavras terminadas em *-ão* no singular não é homogeneamente distribuído nos plurais *-ãos*, *-ães* e *-ões*. Fazendo uma consulta ao Dicionário Eletrônico Houaiss da Língua Portuguesa (2001)³, encontramos um total de 7.416 palavras terminadas em *-ão* no singular. Considerando os plurais etimológicos dessas palavras, temos a seguinte distribuição:

Tipos	Número	% no grupo <i>-ão</i> singular	% no dicionário
<i>-ÕES</i>	7.260	97,8	3,17
<i>-ÃOS</i>	108	1,5	0,047
<i>-ÃES</i>	48	0,7	0,021
TOTAL	7.416	100	3,24

Tabela 1: Frequência de tipo de *-ões*, *-ãos* e *-ães* no Dicionário Eletrônico Houaiss

Na tabela acima, observamos que a quantidade de itens pluralizados em *-ões* é bastante superior à de itens pluralizados em *-ãos* e *-ães*. Os plurais em *-ões* representam 3,17% de todo o Dicionário Eletrônico Houaiss, ou 97,8% do grupo terminado em *-ão* no singular. Por outro lado, os plurais em *-ãos* e *-ães* apresentam um número muito menor de itens (1,5% do grupo é flexionado com *-ãos* e 0,7% é flexionado com *-ães*). Percebe-se, portanto, que a distribuição das terminações de plural nesse grupo é bastante desigual. Uma pergunta a se fazer neste ponto é por que o português apresenta preferência pelo plural em *-ões*, em vez de *-ãos* e *-ães*. A resposta pode ser que, na passagem do latim ao português, nossa língua já tenha recebido uma quantidade maior de itens em etimologia *-one*. Said Ali (1964, p. 60) ressalta esse ponto:

Os termos em *-ane* e *-anu*, donde se originaram os plurais em *-ães* (português antigo *-ães*) e *-ãos* (português antigo *-ãos*), recebidos do latim, foram mui poucos em comparação da onda de nomes em *-one* com que se enriqueceu o idioma português.

³ O Dicionário Eletrônico Houaiss da Língua Portuguesa (2001) apresenta aproximadamente 228.500 verbetes.

Parreira e Pinto (1985, p. 160-161) também apontam o mesmo fato:

Preferimos, no entanto, as formas *aldeões*, *corrimões*, *ermitões*, *guardiões*, *verões* e *vilões*, por tenderem a ser as mais usadas. Repare-se que a maioria dos nomes terminados em *-ão* que vieram do latim para o português formam o plural em *-ões*, daí que as palavras novas que entram na língua terminadas em *-ão* tendam a seguir a maioria, formando o plural em *-ões*.

Said Ali (1964, p. 59-60) também menciona a adoção da forma *-ões* para pluralizar quaisquer palavras novas que surjam na língua:

A regra geral do plural em *-ões* vigora para a linguagem moderna, sendo aplicada naturalmente a quaisquer termos novos: *civilizações*, *vagões*, *salões*, *montões*, *felicitações*, *estremeções*, etc. Dêste oceano de substantivos em *-ão*, mal tiramos umas duas dúzias de vocábulos com plural diferente...

Investigando as palavras terminadas em *-ão* no singular a partir do Modelo de Redes (Bybee, 1985, 1995, 2001), propomos que uma interação entre as frequências de tipo e de ocorrência pode justificar variações de plural que vêm acontecendo nesse grupo de palavras. Essa hipótese pode ser sumarizada a partir das seguintes afirmações:

- A terminação *-ões* apresenta maior frequência de tipo no grupo de palavras terminadas em *-ão* no singular. Essa frequência de tipo mais alta pode “atrair” membros que originalmente eram pluralizados em *-ãos* e *-ães*;
- Com relação à frequência de ocorrência, supomos que itens altamente frequentes, mesmo que pertençam às classes de *-ãos* ou *-ães*, deverão manter seus plurais etimológicos. Isso se justifica porque, quando um item tem alta frequência de ocorrência, adquire autonomia semântica e não depende de sua classe para ser recuperado no léxico mental. Defendemos, portanto, a preservação de itens de alta frequência de ocorrência, independentemente da classe de plural a que pertençam.

Essas hipóteses ratificam as afirmações de Phillips (1984), de que mudanças analógicas afetam as palavras menos frequentes primeiro.

A fim de confirmarmos ou refutarmos essas hipóteses, um trabalho de campo foi realizado. A seguir descrevemos os procedimentos metodológicos adotados nessa investigação.

Metodologia para a coleta e classificação dos dados

A coleta de dados sobre as palavras terminadas em *-ão* foi realizada entre janeiro e março de 2006. Ao todo, foram entrevistados 36 informantes, divididos segundo os critérios classe social, nível de escolaridade e faixa etária. A realização de entrevistas sociolinguísticas seria ideal para nossa pesquisa; no entanto, como as palavras terminadas em *-ão* representam um grupo pequeno de itens no léxico do português, optamos por realizar experimentos psicolinguísticos com palavras pré-determinadas. Dessa maneira, garantiríamos a ocorrência das palavras terminadas em *-ão* e a homogeneidade dos dados coletados. As palavras adotadas nos experimentos foram selecionadas a partir da frequência de ocorrência, já que nossa hipótese se fundamentava nesse critério. A frequência foi medida no corpus NILC/São Carlos (disponível em <www.linguateca.pt/ACDC/>), que é parte do Corpus CETEM/Público (Corpus de Extratos de Textos Eletrônicos MCT/Público). Esse corpus apresenta 41.372.943 dados, oriundos de textos didáticos, jornalísticos, revistas, enciclopédias, etc.

Na coleta de dados, foram realizados três experimentos psicolinguísticos: conjunto de figuras, leitura de frases e teste de reação. Abaixo explicamos cada um desses experimentos.

Conjunto de figuras: Foram selecionadas fotografias ou gravuras de palavras terminadas em *-ão* no singular. As gravuras ou fotografias sempre apresentavam esses itens em quantidade plural. Uma figura era mostrada por vez a cada informante e pedia-se que ele(a) fizesse um comentário sobre o que via na fotografia ou gravura.

Leitura de frases: Foram montadas frases com palavras terminadas com *-ão* inseridas em forma de figuras. O informante teria que ler essas frases, substituindo os desenhos pelas palavras na sua forma plural. Abaixo apresentamos dois exemplos dessas frases:

“Os [*figura de dois vulcões*] da Califórnia sempre entram em erupção.”

“As [*figura de duas mãos*] dele estavam muito ressecadas pelo frio.”

Teste de reação: Nesta etapa dos experimentos, perguntamos aos informantes o plural de determinadas palavras⁴. Neste caso, estávamos interessados em analisar duas informações: 1) O plural utilizado pelo informante e 2) A reação que os falantes

¹ Em cada etapa dos experimentos foram inseridas “palavras distratoras”, ou seja, itens que não seriam contabilizados nos dados finais, mas cuja função era fazer com que os informantes não percebessem que estávamos analisando somente as palavras terminadas em *-ão*.

teriam antes de responder o plural da palavra solicitada. Em alguns casos, os falantes não demonstraram nenhum problema para prover o plural; em outros, no entanto, houve hesitação e os informantes precisaram pensar um pouco antes de pronunciar o plural da palavra.

Conforme afirmamos anteriormente, o objetivo de nossa pesquisa era investigar se palavras cujo plural etimológico em *-ãos* ou *-ães* estão migrando para o grupo em *-ões*. No entanto, esse objetivo parte de uma premissa não investigada: de que palavras em *-ões* não estão migrando para as classes em *-ãos* ou *-ães*. Concluímos que os resultados de nossa pesquisa não seriam válidos se não verificássemos, primeiramente, que palavras em *-ões* de fato não estão adotando plurais em *-ãos* ou *-ães*. Em função disso, optamos por incluir em nossos experimentos palavras dos três grupos etimológicos: *-ões*, *-ãos* e *-ães*. Portanto, nossa primeira análise estatística tinha como objetivo verificar se nossa investigação partia do pressuposto correto: de que palavras em *-ãos* e *-ães* estavam adotando plurais em *-ões*, não o oposto.

Abaixo apresentamos as palavras usadas nos experimentos.

Palavras	Ocorrências	Palavras	Ocorrências	Palavras	Ocorrências
pião	1	balão	110	opinião	571
pavão	05	pão	118	caminhão	742
limão	13	televisão	122	avião	874
escrivão	14	leão	125	cartão	941
escorpião	23	anão	142	alemão	1.025
bênção	26	pulmão	164	irmão	1.125
guardião	26	religião	191	órgão	1.578
união	34	cristão	235	mão	2.949
vulcão	62	cão	350		
		grão	471		
		exposição	494		

Tabela 2: Palavras selecionadas para os experimentos psicolinguísticos (frequência plural)

Com relação à codificação dos dados, utilizamos fatores linguísticos e extralinguísticos a fim de verificar possíveis condicionamentos linguísticos ou sociais que justificassem o fenômeno analisado. A seguir discutimos cada um dos fatores codificados.

Fatores extralinguísticos:

- **Faixa etária:** indivíduos entre 15 e 30 anos, entre 31 e 45 anos, entre 46 e 60 anos. Analisando este fator, nosso objetivo era verificar se haveria mudança em progresso perceptível através da análise de tempo aparente (Cf. Labov 1972, 2001; Chambers, 1995);
- **Escolaridade:** informantes com ensino fundamental ou superior. Estávamos interessados em saber se o nível de escolaridade interferiria na escolha de um ou outro tipo de plural;
- **Gênero:** falantes de sexo feminino ou masculino. Queríamos analisar se o gênero poderia interferir nas variações observadas (Cf. Chambers, 1995);
- **Informante:** a fim de verificar se o comportamento do indivíduo interferiria no processo, acrescentamos um código para cada indivíduo separadamente.

Fatores linguísticos:

- **Número de sílabas:** foram selecionadas palavras monossílabas (“mão”, “pão”) ou polissílabas (“órgão”, “sacristão”);
- **Estrutura morfológica:** palavras com sufixo (“guardião”, “sacristão”) ou palavras sem sufixo (“pião”, “vulcão”);
- **Plural etimológico:** etimologia em *-ões*, em *-ãos* ou em *-ães*⁵;
- **Segmento precedente:** vogal (“avião”, “leão”) ou consoante (“cristão”, “grão”);
- **Frequência de ocorrência:** frequência baixa (entre 0 e 99 ocorrências), média (entre 100 e 500), alta (mais de 500 ocorrências);
- **Palavra:** foi acrescentado um código a cada item lexical, a fim de verificar se haveria palavras específicas liderando o processo analisado.

Os dados coletados foram submetidos ao programa Goldvarb 2005. Foi feita uma análise binária cujas características serão descritas na subseção seguinte.

⁵ A etimologia das palavras foi definida através de consultas ao Dicionário Eletrônico Houaiss da Língua Portuguesa (2001).

Análise de dados

Conforme afirmamos anteriormente, a primeira análise estatística que realizamos foi para verificar empiricamente a direção das migrações de plural: de *-ãos* e *-ães* para *-ões* ou o inverso. Finalizada a coleta de dados, contamos com um total de 1.277 palavras. A tabela abaixo representa o percentual de palavras que mudou de classe:

	Respostas em <i>-ões</i>		Respostas em <i>-ãos</i>		Respostas em <i>-ães</i>	
Plurais Etím.	N	%	N	%	N	%
Etím. <i>-ões</i>	626/669	93,5	41/669	6,1	02/669	0,2
Etím. <i>-ãos</i>	140/361	38,7	206/361	57	15/361	4,1
Etím. <i>-ães</i>	37/247	14,9	13/247	5,2	197/247	79,7

Tabela 3: Resultados gerais para migrações de plurais

Na tabela acima, destacamos em negrito os casos em que houve maiores ocorrências de migração de plurais. Observamos que os itens etimológicos em *-ãos* foram pluralizados em *-ões* em 38,7% dos casos. Esse é o maior índice de migrações de plural observado no grupo. Em segundo lugar, observamos que as palavras de plural etimológico em *-ães* também migraram para *-ões*, mas em uma porcentagem menor (14,9%). Por outro lado, quando analisamos os plurais etimológicos em *-ões*, percebemos que eles mantiveram seus plurais em 93,5% dos casos. Apenas 6,3% desses itens foi pluralizado em *-ãos* ou *-ães*.

Os dados apresentados nessa tabela confirmam nossa hipótese de que palavras etimológicas em *-ãos* e *-ães* estão sendo pluralizadas como *-ões*, mas o oposto (itens de etimologia *-ões* serem pluralizados como *-ãos* ou *-ães*) ocorre com muito menos frequência. Após corroborar a direção em que as migrações de plural ocorrem, optamos por retirar de nossa análise todos os itens cujo plural etimológico era *-ões*. Agindo dessa maneira, buscamos analisar os grupos em que a variação linguística ocorre de forma mais proeminente (nos itens etimológicos em *-ãos* e *-ães*). Essa recodificação dos dados nos permitiu fazer uma regressão binária no programa Goldvarb. As variáveis dependentes adotadas foram o uso de plural em *-ões* versus plurais em *-ãos* e *-ães* (codificados em conjunto). Excluindo os itens de etimologia em *-ões*, ficamos com 608 dados em *-ãos* e *-ães* que serão analisados a seguir.

Partindo para a análise binária, o programa Godvarb selecionou os seguintes fatores como determinantes na adoção de plural em *-ões*: **faixa etária**, **plural etimológico** e **frequência de ocorrência**. Comentaremos esses resultados em

separado em cada uma das tabelas a seguir. Antes, porém, de fazermos essa análise, queremos comentar a interferência do fator número de sílabas. Esse fator não foi selecionado pelo Goldvarb como relevante para a adoção do plural em *-ões*. No entanto, essa não seleção tem um motivo: todas as palavras monossílabas foram pluralizadas em *-ãos* ou *-ães*. Esse resultado gerou um “knockout” no programa Goldvarb, por isso foi necessário remover os códigos para número de sílabas. Esse foi o motivo pelo qual o número de sílabas não foi selecionado na análise binária. Como houve adoção categórica de *-ãos* ou *-ães* quando as palavras eram monossílabas, achamos relevante comentar esse fator. Vejamos os resultados na tabela abaixo:

Número de Sílabas	Dados Obtidos	%
Monossílabos	0/260	0
Polissílabos	177/348	50,8

Tabela 4: Efeito do número de sílabas

Essa tabela confirma a informação de que não houve itens monossilábicos pluralizados em *-ões*. É importante lembrar que, no PB, não existe nenhuma palavra monossílabo cujo plural seja em *-ões*. Itens como *cão*, *grão*, *mão*, *pão* fazem plural sempre em *-ãos* ou *-ães*. Em nossos experimentos, os falantes demonstraram ter conhecimento dessa informação, por isso não utilizaram plurais em *-ões* para palavras monossílabas. Como o Modelo de Redes propõe, palavras inteiras são estocadas no léxico mental a partir de similaridades fonéticas e semânticas. Conexões morfológicas emergem das semelhanças entre os itens léxicos estocados. Provavelmente, as palavras monossílabas terminadas em *-ão* no singular formam um esquema local (Cf. Bybee; Slobin, 1982) que as previne contra mudanças analógicas. Embora o tipo *-ões* seja mais frequente para itens singulares em *-ão*, as palavras monossílabas estão imunes a essa variação porque representam uma classe separada no léxico mental. A peculiaridade de possuir apenas uma sílaba previne essas palavras de sofrerem generalizações analógicas. Preservando o plural etimológico das palavras monossílabas, o falante demonstra ter em seu léxico mental a informação de que itens assim não poderiam ser pluralizados como “*cões*”, “*grões*”, “*mões*”, “*pões*”.

A seguir, vamos comentar os fatores que o programa Goldvarb selecionou como relevantes para a migração de plurais em *-ãos* e *-ães* em direção a *-ões*. Em primeiro lugar, vamos comentar o único fator sociolinguístico selecionado, faixa etária.

Faixa Etária	Dados Obtidos	%	Peso Relativo
15-30	74/207	35.7	0.532
31-45	61/206	29.6	0.505
46-60	41/195	21	0.460

Tabela 5: Efeito da faixa etária

Os dados relativos à faixa etária nos permitem propor que as migrações de *-ãos* e *-ães* para *-ões* são um caso de mudança em progresso (Cf. Labov 1972, 2001; Chambers, 1995). Observa-se que indivíduos mais idosos tendem a conservar os plurais etimológicos (peso relativo de 0.460) e falantes em idade mediana favorecem ligeiramente o fenômeno (0.505). Por outro lado, falantes mais jovens são os líderes no processo (0.532), por isso pode-se afirmar que se trata de uma mudança em progresso. Seria oportuno realizar outros experimentos linguísticos com os mesmos falantes depois de um intervalo de mais ou menos 10 anos. Assim poder-se-ia verificar se a mudança em progresso estaria mais avançada.

Em seguida vamos analisar os fatores linguísticos, começando por origem etimológica.

Etimologia	Dados Obtidos	%	Peso Relativo
Em <i>-ãos</i>	140/361	38.8	0.860
Em <i>-ães</i>	36/247	14.6	0.066

Tabela 6: Efeito da origem etimológica

A tabela acima demonstra que os plurais etimológicos em *-ãos* favorecem muito as migrações de plurais em direção a *-ões* (peso relativo de 0.860). Plurais etimológicos em *-ães*, por outro lado, desfavorecem o fenômeno (0.066). Para explicar esse favorecimento, vamos recorrer à tabela em que contabilizamos a quantidade de itens em *-ãos*, *-ães* e *-ões* no léxico do PB. Naquela tabela, observamos que as palavras em *-ães* representam a menor porcentagem do grupo (0,7%). Justamente por ser um grupo pequeno dentro da classe de palavras em *-ão* no singular, o plural desse grupo acaba sendo marcado e preservado. Além disso, nesse grupo (em *-ães*) temos itens monossílabos (*cães*, *pães*) que não mudam sua pluralização. Talvez esses dois fatores em conjunto (pequeno número de itens léxicos e a estabilidade dos monossílabos) contribua para que a classe em *-ães* desfavoreça a migração em direção aos plurais em *-ões*.

Por fim, vamos comentar os efeitos de frequência de ocorrência.

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

Tabela 7: Efeito da frequência de ocorrência

A tabela acima mostra que palavras de baixa frequência são as líderes na migração de plurais (0.995), palavras de frequência média desfavorecem ligeiramente o fenômeno (0.419) e palavras de alta frequência desfavorecem muito a adoção de plurais em *-ões* (0.019). De acordo com o Modelo de Redes, esse resultado já era esperado. Conforme afirmou-se anteriormente, a frequência de uso das palavras determina se elas serão ou não estocadas no léxico mental. Itens altamente frequentes são armazenados individualmente, ao passo que itens pouco frequentes são relembrados através das conexões com sua classe de origem. Partindo desse princípio, palavras frequentes têm memória lexical mais forte e por isso são mais imunes a processos analógicos, como o caso da migração de plurais em direção a *-ões*. Por outro lado, palavras pouco frequentes não têm memória lexical suficientemente forte para que sejam relembradas individualmente. O acesso lexical a elas se faz primeiro através do acesso à sua classe. Nesse caso, a informação que emerge a partir de conexões fonéticas e semânticas é o fato de que, para palavras terminadas em *-ão* no singular, o plural em *-ões* é o tipo mais frequente. Por causa disso, toda vez que o falante não tem memória suficiente para lembrar o plural de um determinado item terminado em *-ão* no singular, ele adota o plural em *-ões*, que é o mais frequente para a classe. Dessa forma, itens pouco frequentes estão desencadeando a mudança em progresso no grupo de palavras em *-ão* no singular.

Além da análise de frequência de ocorrência como um todo, vamos observar, também, as palavras individualmente, a fim de verificar se essa análise também corrobora a informação de que palavras mais frequentes favorecem a adoção de plurais em *-ões*. Organizamos as palavras terminadas em *-ãos* e *-ões* no gráfico abaixo mostrando a quantidade de itens léxicos que sofreram mudança em direção a *-ões*:

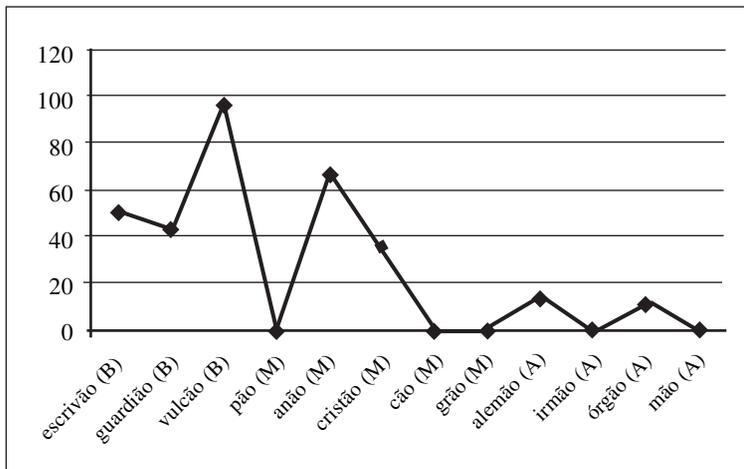


Gráfico 1: Efeito da palavra

No gráfico acima, as palavras estão organizadas em ordem crescente por frequência de ocorrência. Na frente de cada palavra colocamos uma letra identificando as faixas de frequência: Baixa, Média e Alta. O gráfico demonstra de forma visual o que observamos anteriormente com relação à frequência de ocorrência. Percebe-se nitidamente que a migração de plurais diminui conforme aumenta a frequência de ocorrência das palavras individualmente. Dentro do grupo de frequência média, observamos que os itens monossílabos (*pão*, *cão* e *grão*) não foram afetados pela migração de plurais. Conforme afirmou-se anteriormente, a característica de ser monossílabo é uma referência muito importante para a manutenção dos plurais em *-ãos* e *-ães*. Por outro lado, itens de frequência de ocorrência média e com mais de uma sílaba, como *anão* e *cristão*, já apresentam um certo grau de variação linguística. Entre as palavras de frequência alta, observa-se que a variação diminui consideravelmente.

Na subseção seguinte, vamos proceder a uma discussão geral dos resultados alcançados com nossos experimentos psicolinguísticos.

Discussão geral dos resultados

Analisando o processo de variação de plurais nas palavras que terminam em *-ão* no singular, observamos, primeiramente, que, dentre as três possibilidades de plural para esse grupo, o morfema *-ões* apresenta frequência bastante superior aos demais morfemas (*-ões* pluraliza 97,8% dos itens, *-ãos* 1,5% e *-ães*, 0,7%). Ao analisar os dados coletados com nossos experimentos psicolinguísticos, buscamos,

primeiramente, observar em que direção as migrações de plurais aconteciam. Observou-se que itens etimológicos em *-ãos* e *-ães* estavam adotando plurais em *-ões*. Por outro lado, foram poucas (6,2%) as ocorrências de itens etimológicos em *-ões* adotarem plurais em *-ãos* ou *-ães*. Com essa análise prévia, definiu-se a direção das migrações de plurais: palavras de origem etimológica em *-ãos* e *-ães* estavam migrando para a classe de plurais em *-ões*. Nossa justificativa para isso tem como base um dos pressupostos do Modelo de Redes: palavras inteiras são estocadas no léxico mental e redes morfológicas emergem a partir das conexões estabelecidas entre esses itens. Como os itens terminados em *-ão* não apresentam distinção no singular, eles provavelmente são estocados em conjunto no léxico mental. A informação de que a pluralização em *-ões* é mais frequente para esse grupo emerge a partir de comparações entre os itens armazenados. Sendo assim, sempre que o falante precisa pluralizar uma palavra, mas não sabe que plural utilizar, acaba lançando mão do morfema mais frequente para esse grupo, que é *-ões*. Os demais morfemas (*-ãos* e *-ães*) são preservados apenas em itens específicos, como monossílabos ou palavras altamente frequentes e que, por isso, têm memória lexical independente. O que observamos nessas palavras é que, uma vez que não existe uma distinção no singular, em teoria qualquer uma das três desinências pode ser aplicada no plural. Sendo assim, a alta frequência de tipo do morfema *-ões* está abarcando itens que anteriormente eram pluralizados em *-ãos* e *-ães*. O morfema *-ões* é, portanto, o mais produtivo para o grupo de palavras que termina em *-ão* no singular.

Como vimos, esse fenômeno linguístico indica mudança em progresso, então pode ser que, em gerações futuras, os plurais em *-ãos* e *-ães* se apliquem a cada vez menos itens, até que fiquem restritos a pequenos grupos altamente marcados, como os monossílabos.

Nossos experimentos também nos fizeram perceber que, apesar da alta frequência de tipo do morfema *-ões*, ele não é aplicado indiscriminadamente a todas as palavras que terminam em *-ão* no singular. Observamos claramente que nenhum dos monossílabos foi pluralizado em *-ões*, apesar da massiva frequência de tipo desse morfema. Essa suposta isenção dos monossílabos à variação linguística nos faz concluir que os falantes possuem informações detalhadas sobre os itens estocados no léxico mental. Apesar de saberem que *-ões* é muito frequente, os falantes também sabem que há ambientes em que a aplicação de *-ões* não se dá, como no caso dos monossílabos. Já que todos os monossílabos terminados em *-ão* no singular são pluralizados em *-ãos* ou *-ães*, seria estranho se algum falante pluralizasse uma dessas palavras com –

ões. A informação sobre a pluralização dos itens monossílabos emerge a partir da comparação entre as palavras estocadas no léxico mental. Esses itens representam uma “subclasse” dentro do grupo em *-ão*. O fato de serem monossílabos os torna imunes à variação linguística.

Em nossa análise, percebemos também que a frequência de ocorrência foi determinante para que certos itens etimológicos em *-ãos* e *-ães* adotassem o plural em *-ões*. Observamos claramente que palavras menos frequentes foram as primeiras a mudar de plural. Conforme afirmou-se anteriormente, palavras altamente frequentes têm memória lexical mais forte, por isso sofrem menos mudanças analógicas. Por outro lado, palavras de baixa frequência de ocorrência não são lembradas com facilidade, por isso, quando usadas, os falantes tendem a flexioná-las de acordo com o paradigma mais frequente para aquela classe, no caso, o plural em *-ões*. Conforme Phillips (1984) afirma, mudanças não-fisiologicamente motivadas afetam as palavras menos frequentes primeiro, ao passo que mudanças fisiologicamente motivadas afetam as palavras mais frequentes primeiro. No caso dos plurais em *-ão* no singular, estamos lidando com uma variação não-fisiologicamente motivada. A motivação para essa mudança não reside na fisiologia da fala, mas sim na analogia, que faz com que classes linguísticas com menos itens sejam flexionadas a partir de um outro paradigma, desde que haja alguma similaridade entre elas. Este parece ser o caso exato das palavras sob análise: todos esses itens lexicais terminam em *-ão* no singular, então a força da analogia pode interferir, porque existe uma identidade fonética no singular que permite que, em princípio, qualquer um dos três plurais possa ser aplicado a qualquer palavra. Não existe nenhuma informação no singular da palavra que faça com que uma forma de plural prevaleça sobre as outras, então formou-se um ambiente propício para que a analogia, baseada na força do tipo *-ões*, atuasse sobre esse grupo.

Analisando as palavras individualmente, observamos que essa variação linguística vem afetando o léxico gradualmente, de acordo com os postulados da Difusão Lexical (Wang, 1969, 1977). A variação é foneticamente abrupta, porque não existe estágio intermediário entre uma forma e outra de plural. Por outro lado, os itens lexicais vão adotando o plural em *-ões* de forma gradual. Conforme a Difusão Lexical propõe, não vai haver, necessariamente, uma regularidade na mudança. Pode ser que alguns itens mudem seu plural definitivamente e outros não mudem nunca. Neste momento estamos assistindo a uma mudança em progresso que pode ou não alcançar resultado regular. Como há itens muito marcados, como os monossílabos, pode ser que essa classe seja sempre imune à mudança.

Na subseção seguinte partimos para as conclusões finais alcançadas com nosso trabalho.

Conclusões

As conclusões obtidas com nossos experimentos sobre as palavras terminadas em *-ão* no singular podem ser assim definidas:

- Palavras inteiras são estocadas no léxico mental. Se esse não fosse o caso, não haveria efeitos de frequência de ocorrência no grupo de palavras analisado;
- Falantes possuem um conhecimento bastante minucioso sobre as peculiaridades do léxico mental, por isso preservam a informação de que itens monossílabos e palavras etimológicas em *-ães* não favorecem a migração de plurais;
- A melhor análise para o fenômeno linguístico estudado reside em uma interação entre frequência de tipo e de ocorrência, da seguinte forma: o tipo *-ões* é o mais frequente para itens terminados em *-ão* no singular, por isso a maioria das palavras é flexionada com esse morfema. No entanto, a frequência de ocorrência faz com que os itens lexicais sejam mais ou menos imunes à mudança linguística: palavras muito frequentes não adotam plurais em *-ões* porque sua memória lexical é mais forte. Por outro lado, palavras de baixa frequência de ocorrência são as primeiras a serem afetadas pela mudança, porque sua memória lexical é mais fraca e, por isso, se flexionam de acordo com o padrão mais proeminente para sua classe, que é *-ões*;
- De uma maneira geral, observamos que o léxico mental é remodelado a partir do uso da língua em situações concretas de comunicação. De modo inverso, novos padrões decorrentes de associações no léxico mental também podem causar variações e/ou mudanças linguísticas.

Plural of *-ão* ending words in Brazilian Portuguese: frequency effects

Abstract: This article investigates the Portuguese words that end with the diphthong *-ão* in the singular form (*avião*, *capitão* and *cidadão* (*airplane*, *captain* and *citizen*)). When pluralized, these words have a different plural form: *avião* – *aviões*, *capitão* – *capitães*, *cidadão* – *cidadãos*. Although etymologically each word should have a different plural, it is noticeable that some of these words are pluralized using a morpheme that does not correspond to the etymological origin. The purpose of this article is to investigate why and how these variations happen to this group. The theoretical background adopted is the Network Model, as proposed by Bybee (1985, 2001).

Key-words: Network Model, lexical storage, pluralization.

Bibliografia

BROWMAN, Catherine P.; GOLDSTEIN, Louis M. Articulatory phonology: an overview. *Phonetica*, n. 49, p. 155-80, 1992.

BYBEE, Joan. *Morphology: A study of the relation between meaning and form*. Philadelphia: John Benjamins, 1985.

_____. Regular morphology and the lexicon. *Language and Cognitive Processes*, Cambridge, n. 10, p. 425-455, 1995.

_____. *Phonology and language use*. Cambridge: Cambridge University Press, 2001. (Cambridge Studies in Linguistics, 94).

_____. Mechanisms of change in grammaticization: the role of frequency. In: JANDA, Richard; JOSEPH, Brian (Eds.). *Handbook of historical linguistics*. Oxford: Blackwell, 2002. p. 602-623.

BYBEE, Joan; SLOBIN, Dan. 1982. Rules and schemas in the development and use of the English past tense. *Language*, Washington, n. 58, p. 265-289, 1982.

CHAMBERS, J. K. *Sociolinguistic theory: linguistic variation and its social significance*. Oxford: Blackwell, 1995. (Language in society).

Corpus NILC/São Carlos. Disponível em: <www.linguateca.pt/ACDC/>.

FIDELHOLTZ, James L. Word frequency and vowel reduction in English. In: REGIONAL MEETING CHICAGO LINGUISTICS SOCIETY, 7, 1975, Chicago. *Papers from...* Chicago: Chicago Linguistic Society, 1975. p. 200-213.

GOLDVARB (2005). Disponível em <http://individual.utoronto.ca/tagliamonte/goldvarb.htm>.

HOUAISS, Antônio. *Dicionário Eletrônico Houaiss da Língua Portuguesa*. Rio de Janeiro: Editora Objetiva, 2001.

LABOV, William. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press, 1972.

_____. *Principles of linguistic change: social factors*. Oxford: Blackwell, 2001. (v.2)

LESLAU, Wolf. Frequency and change in the Ethiopian languages. *Word*, New York, v. 25, p. 180-189, 1969.

PAGLIUCA, William; MOWREY, Richard. Articulatory evolution. In: RAMAT, Anna Giacalone; CARRUBA, Onofrio; BERNINI, Giuliano. *Papers from the 7th International Conference on Historical Linguistics*. Amsterdam: Benjamins, 1987. p. 459-72.

PARREIRA, Manuela; PINTO, J. Manuel de Castro. *Prontuário ortográfico moderno*. Lisboa: Edições Asa, 1985.

PHILLIPS, Betty S. Word frequency and the actuation of sound change. *Language*, Washington, v. 60, n. 2, p. 320-42, 1984.

SCHUCHARDT, Hugo. On sound laws: against the neogrammarians. In: *Schuchardt, the neogrammarians, and the transformational theory of phonological change*. Frankfurt: Athenaeum, 1972 [1885]. p. 39-72.

SAID ALI, Manuel. *Gramática histórica da língua portuguesa*. 3. ed. São Paulo: Edições Melhoramentos, 1964.

SEQUEIRA, Francisco Martins de. *Aspectos do português arcaico*. Lisboa: Livraria Popular, 1943.

WANG, William S-Y. Competing changes as a cause of residue. *Language*, Washington, v. 45, p. 9-25, 1969.

WANG, William S-Y. (Ed.). *The lexicon in phonological change*. The Hague: Mouton, 1977.