

FREQUÊNCIA DE USO: TOKENS E TYPES NA HARMONIA VOCÁLICA

TOKENS AND TYPES IN VOWEL HARMONY

Eneida de Goes Leal¹

Leda Bisol²

eneidaleal@yahoo.com

lebisol@uol.com.br

RESUMO: Este texto trata da relação entre harmonia vocálica e a frequência de palavras, com base na fonologia de uso (Bybee 1985, 2001, 2002, 2010; Phillips 1984; Lindblom 1990, entre outros). Phillips (1984, p. 322) explica que palavras mais frequentes são mais sensíveis à variação do que as menos frequentes, principalmente no que se refere à redução vocálica e à assimilação, pois mudanças sonoras fisiologicamente motivadas (como a harmonia) afetam primeiro palavras mais frequentes. Dessa forma, investigamos neste artigo se a harmonia é mais aplicada em palavras mais frequentes e, de modo inverso, se há uma menor aplicação com palavras menos frequentes, com base em dados do VARSUL (Schwindt 1995), relacionando sua frequência ao Corpus Brasileiro (Berber Sardinha 2004).

PALAVRAS-CHAVE: harmonia vocálica; frequência de tokens; frequência de types; pretônica

ABSTRACT: This paper deals with the relationship between vowel harmony and word frequency, based on usage phonology (Bybee 1985, 2001, 2002, 2010; Phillips 1984; Lindblom 1990). Phillips (1984, p. 322) shows that frequent words are more susceptible to variation than less frequent words, mainly concerning to vowel reduction and assimilation. Physiologically motivated sound changes (as vowel harmony) firstly affect more frequent words. Therefore, we investigate whether vowel harmony occurs more often in most frequent words, and conversely whether it is less applied in least frequent words. VARSUL data (Schwindt 1995) was used in to investigate the phonological process and Corpus Brasileiro (Berber Sardinha 2004) was used as a reference corpus.

KEYWORDS: vowel harmony; token frequency; types frequency; pretonic vowel.

¹ Doutora; pós-doutorado na Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS).

² Doutora; Universidade Federal do Rio Grande do Sul (UFRGS).

INTRODUÇÃO

O tema em foco neste trabalho é o papel da frequência na Harmonia Vocálica (HV), um processo de assimilação regressiva, que tem por alvo a vogal média e por gatilho a vogal alta, /i/ ou /u/, cuja aplicação é variável. Algumas propriedades a notificam: a vogal /i/ é mais ativa do que /u/ e o contexto ideal é a sequência imediata de alvo e gatilho; havendo mais de uma vogal na pretônica, a regra tende a estender-se sem fazer "pulos". Como regra variável, está sujeita à influência de fatores linguísticos e extralinguísticos, mas indubitavelmente atende a todos os requisitos de um processo de assimilação regressiva.

Incorporada ao sistema fonológico do português brasileiro desde Camara Jr. (1970), HV é uma regra variável constatada de norte ao sul do país e amplamente estudada via modelo laboviano.

Assim concebida, passemos a observá-la sob o prisma da fonologia de uso (Bybee 1985, 2001, 2002, 2010; Phillips 1984; Lindblom 1990, entre outros). De acordo com Phillips (1984, p. 322), palavras de uso mais frequente são mais sensíveis à variação do que palavras menos frequentes, principalmente no que se refere à redução vocálica e à assimilação.

A análise de HV e efeitos de frequência foi considerada sob dois tópicos:

- Quanto à frequência de tokens (ocorrências) (cf. seções 1 e 2), Bybee (2010, p. 10-11) assim explica:

A FREQUÊNCIA DE TOKEN é a frequência de ocorrência de uma unidade, geralmente uma palavra, num texto corrido – ou seja, estabelece-se com que frequência uma determinada palavra aparece no texto.³

Assim, na frequência de tokens, são contadas quantas vezes uma determinada palavra apareceu no corpus e, com este tipo de análise, o domínio que está sendo observado para a aplicação de HV é a palavra.

- Quanto à frequência de tokens (tipos de um dado padrão) (cf. seção 2), é explicada por Bybee (2010, p. 10) do seguinte modo:

³ Do original: "TOKEN FREQUENCY is the frequency of occurrence of a unit, usually a word, in running text – how often a particular word comes up."

A FREQUÊNCIA DE TYPE diz respeito à frequência de dicionário em um determinado padrão.⁴

Em frequência de types, as palavras são analisadas como em um dicionário; por exemplo, os verbos são apresentados no infinitivo, os substantivos e adjetivos aparecem em sua forma singular e no gênero masculino. Bybee (2010, p. 10) apresenta como exemplo de types os verbos no passado do inglês, como **came**, **spoke**, **understood**, **liked**, **used**.⁵ Os verbos nesses exemplos podem ser separados quanto ao passado regular ou irregular: com o paradigma de verbos irregulares, representado por **came**, **spoke** e **understood**, e o paradigma dos verbos regulares, terminados em **-ed**, como em **liked** e **used**, o **type** de maior frequência nessa língua.

Para exemplificar a diferença entre types e tokens, observe os exemplos a seguir:

(1) **carregar carregam carrego carregavam carregar carregado**

Os itens em (1) estão relacionados ao significado "levar, conduzir, sustentar" e podem ser chamados de types de "**carreg**" ou "**carregar**". Quanto ao número de tokens, há 6, ou seja, seis formas diversas de uma palavra.

Adicionalmente, esses mesmos itens verbais apresentados em (1) podem ser categorizados do seguinte modo:

(2) singular:	(3) plural:
carregarØ	carregam
carregoØ	carregavam
recarregarØ	
carregadoØ	

Como apresentados nos exemplos (2) e (3), podemos classificar essas palavras em dois types: verbos no singular com os tokens **carregar**, **carrego**, **recarregar** e **carregado**, e um type de verbos no plural, com **carregam** e **carregavam**.

Para fins de análise de HV, contamos com os dados do VARSUL organizados por Schwindt (1995), referentes a Porto Alegre, Florianópolis e Curitiba, dados que

⁴ Do original: "TYPE FREQUENCY refers to the dictionary frequency of a particular pattern."

⁵ Came *vir*, 3ª pessoa, singular, passado simples; spoke *falar*, 3ª pessoa, singular, passado simples; understood *entender*, 3ª pessoa, singular, passado simples; liked *gostar*, 3ª pessoa, singular, passado simples; used *usar*, 3ª pessoa, singular, passado simples.

neste texto estão reorganizados com vistas à proposta da teoria de uso. Assim, uma vez que o contexto ideal de HV é uma sequência imediata entre o alvo e o gatilho, foram retirados do corpus do VARSUL todos os dados que não tivessem esses elementos adjacentes.

Com o intuito de examinar como se dá essa relação entre HV e a frequência de types e tokens, os dados foram analisados no software R, através do Rbrul, um programa que roda efeitos mistos, em que as variáveis independentes podem ser fixas e aleatórias. Verificadas as ocorrências no VARSUL, relacionou-se a frequência de palavra no VARSUL com a frequência do **Corpus Brasileiro** (CBBS, Berber Sardinha 2004), corpus referencial do português brasileiro.

O VARSUL pode ser classificado como corpus de estudo, que tem um conteúdo dialetal, proveniente da variedade do sul do país, com características sociolinguísticas específicas, em que sexo/gênero, faixa etária, escolaridade e cidade delimitam os informantes que compõem o corpus; na composição desses tipos de corpora, a finalidade é um estudo de descrição de um determinado dialeto. O CBBS, por sua vez, é um corpus de referência, cuja finalidade é servir de base a um *corpus* de estudo, já que sua composição abarca frequências de palavras da língua em uso (ou seja, o português brasileiro). Neste caso, quanto maior e mais variada essa amostra, mais será representativa de uma língua. CBBS é um corpus formado de palavras classificadas quanto ao uso tanto na escrita quanto na fala sem delimitações regionais. É um corpus de referência, enquanto o VARSUL agrega um único tipo de texto – o falado. A relação entre os corpora pôde ser feita porque o CBBS conta com um dispositivo que seleciona especificamente dados da fala, utilizado neste trabalho, o que possibilitou a correspondência entre eles. Foi considerada a seguinte escala para as frequências do CBBS:

(4) Escala de frequência adotada para o CBBS:

Mais de 4 Algarismos: frequência alta

Com 3 Algarismos: frequência média

Menos de 2 Algarismos: frequência baixa

Observa-se em (4) que para frequência alta, foram considerados tokens com mais de 4 algarismos no CBBS; para frequência média, tokens com 3 algarismos; e, para frequência baixa, aqueles com menos de 2 algarismos.

Consideramos HV a partir da fonologia de uso (Bybee 1985, 2001, 2002, 2010; Phillips 1984; Lindblom 1990, entre outros) e, segundo' Phillips (1984, p. 322), palavras de uso mais, palavras de uso mais frequente são mais sensíveis à variação do que palavras menos frequentes, principalmente no que se refere à redução vocálica e à assimilação. Assim, por ser uma processo fonológico de assimilação, a expectativa é que seja mais aplicada em tokens e types de frequência alta e, de modo inverso, que haja menor aplicação do processo em frequências mais baixas – de acordo com Phillips (1984). Há ainda duas questões a serem verificadas:

- a) um efeito de frequência de tokens; e/ou
- b) um efeito de frequência de types.

Na seção 1 a seguir, apresentamos a análise feita com base em tokens.

1. AS ANÁLISES DE FREQUÊNCIAS DE TOKENS

Nas subseções a seguir, apresentamos as análises das vogais-alvo /e/ (cf. 1.1) e /o/ (ver 1.2), unindo os tokens quanto a nomes e verbos para, em seguida, separar essas categorias morfológicas, a fim de verificar possíveis efeitos diferentes entre elas, computando-se também em separado palavras mais frequentes e menos frequentes. Assim, foram feitas as seguintes análises:

- 15 contextos de HV com nomes e verbos, com tokens mais frequentes e menos frequentes;
- 15 contextos de HV somente com nomes, com tokens mais frequentes e menos frequentes; e
- 15 contextos de HV somente com verbos, com tokens mais frequentes e menos frequentes

São 6 análises para cada pretônica, abrangendo nomes e verbos e itens mais frequentes e menos frequentes.

1.1. OS TOKENS COM A PRETÔNICA /e/

Para a análise de /e/, relacionando HV dessa pretônica com a frequência de uso de palavra, o número total de tokens foi N=812/2124, com uma proporção de 38,2% de aplicação de HV.

Na Tabela ⁶ a seguir, apresentamos os resultados das 15 palavras mais frequentes com contexto de HV:

	categ.	p.r.	VARSQL	CBBS	freq.
segurar	verbo	1	4	5564	alta
queriam	verbo	1	20	1230	alta
segunda	nome	1	19	5944	alta
pedir	verbo	1	14	4941	alta
pedindo	verbo	1	14	2494	alta
perigo	nome	1	6	1260	alta
menina	nome	1	30	669	média
meninas	nome	1	19	913	média
menino	nome	1	16	915	média
senti	verbo	1	15	431	média
sentia	verbo	1	12	322	média
perigoso	nome	1	12	736	média
pedi	verbo	1	8	790	média
pedia	verbo	1	7	254	média
meninos	nome	1	7	939	média

Tabela 1: Os 15 Nomes e Verbos mais frequentes com /e/

Como se observa na Tabela 1:, todas as 15 palavras mais frequentes no VARSQL têm p.r. = 1, o que significa que a aplicação de HV é favorecida. Nota-se também que, com relação à categoria lexical, 8 palavras são nomes e 7 são verbos, o que parece indicar que há um equilíbrio entre essas classes gramaticais com relação à frequência de palavra na HV.

⁶ Nas Tabelas apresentadas, foram utilizados: **categ.** para indicar categoria morfológica da palavra; **p.r.** para peso relativo; **VARSQL** para representar o total de número de dados encontrados em nossos dados; **CBBS** para indicar o número de frequência dos dados do Corpus Brasileiro; e **freq.** para representar as frequências alta, média e baixa das palavras, de acordo com o Corpus Brasileiro.

Ao comparar a frequência de dados do VARSUL com a frequência do CBBS, observa-se que houve 6 palavras de frequência alta e 9 de frequência média; nota-se também que não aparecem palavras de baixa frequência na Tabela 1:. Assim, pode-se dizer que os contextos de HV se dão com os nomes e verbos de alta e média frequência – e não há casos de contextos com baixa frequência.

Passamos a examinar as palavras menos frequentes, como apresentamos na Tabela abaixo:

	categ.	p.r.	VARSUL	CBBS	freq.
crescimento	nome	0.018	5	20611	alta
terminar	verbo	0.017	11	1367	alta
comercial	nome	0.017	12	6193	alta
atendimento	nome	0.017	7	10969	alta
região	nome	0.017	5	38223	alta
interior	nome	0.017	11	8128	alta
coletivo	nome	0.017	8	2238	alta
teria	verbo	0.016	10	6906	alta
resultado	nome	0.015	12	10999	alta
conhecimento	nome	0.011	33	13958	alta
revistas	nome	0.017	5	912	média
terminal	nome	0.013	15	528	média
verniz	nome	0.018	4	23	baixa
aprendia	verbo	0.018	11	38	baixa
serpentina	nome	0.015	8	8	baixa

Tabela 2: Os 15 Nomes e Verbos menos frequentes com /e/

Os pesos relativos apresentados na Tabela 2: indicam o desfavorecimento da aplicação de HV nos dados do VARSUL – os tokens menos frequentes. A expectativa é que, dos 15 nomes e verbos menos frequentes nos dados do VARSUL, haveria palavras também com baixa frequência no CBBS; no entanto, vemos que há 10 palavras de frequência alta, 2 médias e 3 baixas.

Uma vez verificada a relação entre palavras com contexto de HV, passamos a examinar os nomes separadamente, que seguem na Tabela abaixo:

	p.r.	VARSQL	CBBS	freq.
segundo	1	55	30120	alta
segunda	1	19	5944	alta
perigo	1	6	1260	alta
político	1	5	14763	alta
polícia	1	33	4127	alta
possível	1	5	20485	alta
precisando	1	4	587	média
menina	1	30	669	média
meninas	1	19	913	média
menino	1	16	915	média
perigoso	1	12	736	média
meninos	1	7	939	média
piquenique	1	6	13	baixa
manequim	1	2	2	baixa
segundas-feiras	1	2	62	baixa

Tabela 3: Os 15 Nomes mais frequentes com /e/

Nos 15 nomes mais frequentes, houve 6 palavras de frequência alta, 6 de média e 3 de baixa. Esses valores mostram que não há uma sistematização quanto à frequência, e parece não haver uma preferência de HV quanto a nomes mais frequentes, uma vez que o processo ocorre com palavras de alta, baixa e média frequências, em um contexto em que o esperado é tokens de alta frequência.

	p.r.	VARSQL	CBBS	freq.
conhecimento	0.017	33	13958	alta
comercial	0.022	12	6193	alta
resultado	0.021	12	10999	alta
interior	0.023	11	8128	alta
coletivo	0.022	8	2238	alta
atendimento	0.022	7	10969	alta
abertura	0.023	6	5812	alta
crescimento	0.023	5	20611	alta
região	0.023	5	38223	alta
terminal	0.019	15	528	média
vestibular	0.023	12	761	média
revistas	0.023	5	912	média
maternidade	0.023	5	406	média
verdura	0.023	10	19	baixa
serpentina	0.021	8	8	baixa

Tabela 4: Os 15 Nomes menos frequentes com /e/

Para nomes menos frequentes, houve 9 itens lexicais de frequência alta, 4 de média e 2 de baixa, resultado em que a maioria de palavras de frequência alta no CBBS corresponde a tokens menos frequentes no VARSQL.

Com relação aos verbos, apresentamos a seguir os tokens mais frequentes para essa categoria lexical:

	p.r.	VARFUL	CBBS	freq.
pedir	1	14	4941	alta
queriam	1	20	1230	alta
pedindo	1	14	2494	alta
senti	1	15	431	média
sentia	1	12	322	média
pedi	1	8	790	média
pedia	1	7	254	média
sentindo	1	5	511	média
pedimos	1	3	989	média
vestir	1	4	169	média
precisando	1	4	587	média
mentir	1	3	266	média
acontecia	1	2	482	média
vestia	1	4	19	baixa
pedisse	1	3	60	baixa

Tabela 5: Os 15 Verbos mais frequentes com /e/

Como se observa nos pesos relativos da Tabela 5:, há uma tendência de aplicação de HV nos verbos mais frequentes do VARSUL. Comparando-se ao CBBS, esses verbos correspondem a 3 palavras de frequência alta, 10 de média e 2 de baixa.

	p.r.	VARFUL	CBBS	freq.
recebi	0	6/12	1930	alta
repito	0	4	4042	alta
regular	0	2	1302	alta
prefiro	0.001	3/4	289	média
recuar	0	1	219	média
recupere	0	1	197	média
recebiam	0	2	456	média
refiro	0	1	649	média
regularizar	0	1	320	média
beneficiam	0	2	456	média
percebi	0	2	328	média
recebia	0	12	680	média
recuperaram	0	1	53	baixa
percebia	0	2	65	baixa
recuam	0	1	10	baixa

Tabela 6: Os 15 Verbos menos frequentes com /e/

Os pesos relativos de verbos menos frequentes no VARSUL indicam um desfavorecimento de HV, como se observa na Tabela 6.: Ao comparar a frequência dos 15 verbos menos frequentes no VARSUL com os dados do CBBS, nota-se que houve 3 palavras de frequência alta, 9 de frequência média e 3 de baixa.

Nesta seção, pudemos observar que não houve um padrão nos resultados para a pretônica /e/ (cf. discussão na seção 3). Depois da descrição dos tokens apresentada nesta seção, passamos à análise da frequência de tokens com a pretônica /o/.

1.2. OS TOKENS COM A PRETÔNICA /o/

Da análise entre a frequência dos tokens da pretônica /o/ e HV, o total de tokens foi N= 588/1379, com 42.6% de aplicação de HV.

	categ.	p.r.	VARFUL	CBBS	freq.
polícia	nome	1	33	4127	alta
políticos	nome	1	8	11001	alta
possível	nome	1	5	20485	alta
podia	verbo	1	72	2750	alta
policial	nome	1	3	3991	alta
bonito	nome	1	35	776	média
bonita	nome	1	25	814	média
podiam	verbo	1	11	664	média
costura	nome	1	7	146	média
bonitas	nome	1	3	272	média
domingos	nome	1	5	319	média
sobrinha	nome	1	17	66	baixa
chovia	verbo	1	6	24	baixa
acostumei	verbo	1	5	16	baixa
cochiladas	nome	1	2	1	baixa

Tabela 7: Os 15 Nomes e Verbos mais frequentes com /o/

Para nomes e verbos mais frequentes, observamos que os valores dos pesos relativos são todos 1, o que indica um favorecimento de aplicação da regra no VARFUL. Quanto à relação com o CBBS, há 5 palavras de frequência alta, 6 de frequência média e 4 de baixa frequência e, posto que são 15 tokens mais frequentes com contexto de harmonia, o esperado era que não houvesse frequências baixas, o que não ocorreu.

	categ.	p.r.	VARISUL	CBBS	freq.
modificar	verbo	0.023	7	1234	alta
convivência	nome	0.021	6	2371	alta
domínio	nome	0.021	3	2371	alta
volume	nome	0.021	3	3717	alta
condução	nome	0.02	11	2106	alta
português	nome	0.018	21	1056	alta
motoristas	nome	0.017	8	1154	alta
desenvolvimento	nome	0.015	18	43176	alta
voluntário	nome	0.024	2	500	média
modificando	verbo	0.024	6	239	média
sorriso	nome	0.022	4	300	média
continuasse	verbo	0.019	7	190	média
motorista	nome	0.011	26	832	média
mordomia	nome	0.023	4	48	baixa
mordomias	nome	0.023	4	74	baixa

Tabela 8: Os 15 Nomes e Verbos menos frequentes com /o/

Na Tabela 8., podemos observar que os pesos relativos são baixos, entre 0,011 e 0,024, o que indica desfavorecimento na aplicação de HV nos dados de nomes e verbos menos frequentes com a pretônica /o/. Comparando-se esses dados com os do CBBS, observamos que houve 8 palavras com alta frequência, 5 com média e 2 com baixa frequência.

	p.r.	VARSQL	CBBS	freq.
polícia	1	33	4127	alta
possível	1	5	20485	alta
policial	1	3	3991	alta
políticos	1	8	11001	alta
político	1	5	14763	alta
bonito	1	35	776	média
bonita	1	25	814	média
bonitas	1	3	272	média
bonitos	1	3	151	média
costumes	1	2	911	média
sobrinha	1	17	66	baixa
assobio	1	3	1	baixa
cochiladas	1	2	1	baixa
frigorífico	1	3	2	baixa
cozinheira	1	2	34	baixa

Tabela 9: Os 15 Nomes mais frequentes com /o/

Para os nomes mais frequentes, nota-se que houve um equilíbrio, comparando-se os dados do VARSQL apresentados na Tabela 9: com os dados do CBBS: houve 5 palavras com frequência alta, média e baixa em cada uma delas.

	p.r.	VARSQL	CBBS	freq.
volume	0.024	3	3717	alta
autorização	0.024	2	2346	alta
continuidade	0.024	2	4955	alta
domínio	0.021	3	2371	alta
português	0.02	21	1056	alta
condução	0.019	11	2106	alta
motoristas	0.017	8	1154	alta
desenvolvimento	0.015	18	43176	alta
sorriso	0.024	4	300	média
construtora	0.024	2	200	média
noturno	0.024	2	220	média
voluntário	0.024	2	500	média
motorista	0.013	26	832	média
solúvel	0.024	2	87	baixa
modista	0.021	3	2	baixa

Tabela 10: Os 15 Nomes menos frequentes com /o/

Como se observa na Tabela 10:, os nomes menos frequentes no VARSQL correspondem a 8 palavras de alta frequência, 5 de média e 2 duas de baixa frequência no CBBS, em um contexto em que a previsão era de palavras com baixa frequência.

	p.r.	VARSQL	CBBS	freq.
podia	0.99	72/74	2750	alta
comunicar	0.974	2	1565	alta
comia	0.985	14/17	128	média
podiam	0.984	11	664	média
costuma	0.978	2	474	média
costumo	0.937	3	147	média
procuro	0.927	2	248	média
chovia	0.992	6	24	baixa
convidava	0.977	1	39	baixa
acostuma	0.973	5	16	baixa
complica	0.97	1	71	baixa
costurar	0.965	4	85	baixa
acostumando	0.962	1	32	baixa
acostumei	0.938	5	16	baixa
costurando	0.918	3	15	baixa

Tabela 11: Os 15 Verbos mais frequentes com /o/

Com relação a verbos, os pesos relativos apresentados na Tabela 11: são próximos de 1 (variando entre 0,992 e 0,918), o que indica um favorecimento de HV nas palavras mais frequentes do VARSUL. Confrontando esses dados com o CBBS, observamos que há 2 palavras de frequência alta, 5 de frequência média e 8 frequência baixa.

	p.r.	VARSQL	CBBS	freq.
negociar	0.104	2	2196	alta
continuam	0.038	5	2989	alta
modificar	0.036	7	1234	alta
continuar	0.032	1/18	9676	alta
correspondia	0.089	2	190	média
continuamos	0.083	1	541	média
solidão	0.083	1	144	média
continuava	0.062	5	249	média
consultar	0.052	1	544	média
modificando	0.032	6	239	média
continuasse	0.016	7	190	média
depositava	0.089	2	19	baixa
valorizasse	0.083	1	11	baixa
soluciona	0.081	2	32	baixa
fortificar	0.065	1	14	baixa

Tabela 12: Os 15 Verbos menos frequentes com /o/

Na Tabela 12:, expõem-se os 15 verbos menos frequentes no VARSQL distribuídos em 4 palavras de frequência alta, 7 de média e 4 de baixa no CBBS.

Assim como para a pretônica /e/, não encontramos um padrão para a pretônica /o/ (ver discussão na seção 3).

Uma vez descritas as frequências de tokens, relacionando os dados do VARSQL ao CBBS, passamos, na próxima seção, à análise de HV no que diz respeito aos types.

2. A ANÁLISE DE FREQUÊNCIAS DE TYPES

Nesta segunda análise, o objetivo é verificar se paradigmas de palavras afetam a aplicação de HV, em que foram computados os types. Para separá-los, observamos palavras que têm a mesma base morfológica, como exemplificado em (1) (cf. Introdução).

Já que o critério nos dados do VARSQL foi unir tokens com uma mesma base para formar os types, o mesmo foi feito no CBBS: foram somadas as ocorrências nesse corpus para formar um type. Tomando-se um exemplo deste trabalho, com o type

queria, houve 95 aplicações de HV de um total de 101 dados, o que corresponde a 53187 no CBBS – é, portanto, um type de alta frequência em ambos os corpora. Esses totais de 101 no VARSUL e de 53187 no CBBS correspondem a types, isto é, a um conjunto de dados. Olhando para dentro do type **queria**, temos:

	VARSUL	CBBS
queria	80	40233
queriam	20	11049
queríamos	1	1905
total	101	53187

Tabela 13: Exemplo de cômputo dos dados: o type **queria** no VARSUL e no CBBS

Na Tabela 13: está o type **queria** no VARSUL e no CBBS, com a seguinte correspondência: houve 80 e 40233 tokens de **queria**, respectivamente; 20 e 11049 de **queriam**; e 1 e 1905 de **queríamos**, totalizando 101 no VARSUL e 53187 no CBBS.

As análises dos types contam apenas com itens mais frequentes: não foi possível fazer uma análise com types menos frequentes, pois há 158 deles que apareceram uma única vez no VARSUL.

Como foi dito na Introdução deste artigo, o VARSUL é um corpus de estudo, cujo propósito é descrever o dialeto falado no sul do Brasil. Em meio às entrevistas, pode ser que haja types específicos a informantes. Supondo-se que um informante com a profissão de serralheiro tenha dito, na entrevista sociolinguística, 20 vezes a palavra “esmeril”, que pode ser considerado um índice de frequência alta no Varsul, essa frequência não corresponderia à do português falado no sul nem à do corpus referencial – é um item específico de um informante. Ademais, ainda que se analisassem esses 158 types menos frequentes, os resultados tenderiam a ser aleatórios: por exemplo, houve uma ocorrência da palavra "destino" nos dados do VARSUL (de Porto Alegre, Florianópolis e Curitiba), ou seja, essa palavra é de frequência baixa nesse corpus, enquanto que, no CBBS, há 3320 ocorrências – uma vez que há mais de 4 Algarismos na ocorrência, é um dado de alta frequência no CBBS.

Os types apresentados nas tabelas desta seção estão separados de acordo com as categorias morfológicas: (1) nomes e verbos, (2) nomes e (3) verbos.

Na Tabela a seguir, há 15 types (nomes e verbos) mais frequentes com contextos para HV no VARSUL e a relação com o CBBS:

	categ.	VARSQL	%	CBBS	freq.
quer/ia	verbo	101	94	5515	alta
consequ/ir	verbo	99	74	14418	alta
pod/ia	verbo	98	94	21992	alta
precis/ar	verbo	97	86	114857	alta
serv/iço	nome	77	6	15695	alta
condi/ção	nome	70	13	39826	alta
segund/o	nome	56	98	30120	alta
acredit/ar	verbo	54	43	8821	alta
doming/o	nome	45	84	2839	alta
conhec/i	verbo	94	10	1831	média
modific/ar	verbo	75	0	1478	média
menin/o	nome	73	100	3487	média
bonit/o	nome	68	100	2019	média
perd/i	verbo	56	0	446	média
cozinh/ar	verbo	53	30	106	média

Tabela 14: Os 15 Nomes e Verbos mais frequentes com contexto de harmonia

Como se observa na Tabela 14:, dos 15 types mais frequentes no VARSQL, há 9 deles que correspondem a frequências altas no CBBS e outras 6 que correspondem a types de frequência média nesse corpus; observa-se também que não há casos de types com frequência baixa.

Podemos notar que, nesses contextos de nomes e verbos mais frequentes no VARSQL, a aplicação de HV é bastante heterogênea: nos casos de **menin/o** e **bonit/o**, houve 100% de aplicação; por outro lado, em **modific/ar** e **perd/i**, não houve aplicação de HV, indicado com 0% na Tabela 14:.

Assim, optamos por analisar, além de contextos de HV, também a aplicação do processo, tanto para nomes e verbos, quanto para essas categorias morfológicas analisadas separadamente. Assim, as análises de types compreendem:

- 15 types mais frequentes de nomes e verbos, com **contextos** de HV;
- 15 types mais frequentes de nomes e verbos, com **aplicação** de HV;
- 15 types mais frequentes com nomes, com **contextos** de HV;
- 15 types mais frequentes com nomes, com **aplicação** de HV;

- 15 types mais frequentes com verbos, com **contextos** de HV; e
- 15 types mais frequentes com verbos, com **aplicação** de HV.

Na Tabela 15: a seguir, estão os 15 types mais frequentes de nomes e verbos, com contextos de HV:

	categ.	VARSQL	%	CBBS	freq.
quer/ia	verbo	101	94	5515	alta
pod/ia	verbo	98	94	21992	alta
precis/ar	verbo	97	86	114857	alta
consequ/ir	verbo	99	74	14418	alta
segund/o	nome	56	98	30120	alta
ped/ir	verbo	53	100	4527	alta
sent/ir	verbo	48	98	2818	alta
dev/ia	verbo	46	85	16089	alta
doming/o	nome	45	84	2839	alta
polícia	nome	37	100	14439	alta
polític/a	nome	38	92	91149	alta
menin/o	nome	73	100	3487	média
dorm/ir	verbo	43	56	850	média
bonit/o	nome	68	100	2019	média
sobrinh/o	nome	32	100	262	média

Tabela 15: Os 15 Nomes e Verbos mais frequentes com aplicação de harmonia

Como se vê na Tabela 15:, de 15 palavras (entre nomes e verbos) mais frequentes no VARSUL, houve 11 palavras de frequência alta e 4 de frequência média no CBBS.

Na Tabela 16: a seguir, apresentamos a frequência de nomes em que HV foi aplicada:

	VARISUL	%	CBBS	freq.
serv/iço	77	6	15695	alta
condi/ção	70	13	39826	alta
segund/o	56	98	30120	alta
doming/o	45	84	2839	alta
necessi/dade	42	38	23861	alta
religi/oso	39	3	4290	alta
polícia	37	100	14439	alta
polític/a	38	92	91149	alta
motorist/a	34	0	1986	alta
português	28	4	1913	alta
comid/a	24	92	1500	alta
avenid/a	19	47	280	alta
menin/o	73	100	3487	média
bonit/o	68	100	2019	média
sobrinh/o	32	100	262	média

Tabela 16: Os 15 Nomes mais frequentes com contexto de harmonia

Como se observa na Tabela 16:, há 12 nomes mais frequentes no VARISUL com contexto de HV que têm frequência alta no CBBS e 3 com frequência média. Novamente, não houve nenhum caso com frequência baixa.

Nas duas Tabelas a seguir, estão 15 types, mas, desta vez, a frequência diz respeito à aplicação de HV (sem contar os casos em que HV não ocorreu), e nomes e verbos estão separados. O interesse é relacionar frequência de types com aplicação de HV.

	VARISUL	%	CBBS	freq.
segund/o	56	98	30120	alta
doming/o	45	84	2839	alta
polícia	37	100	14439	alta
polític/a	38	92	91149	alta
comid/a	24	92	1553	alta
perig/o	19	100	2261	alta
necessidade	42	38	23861	alta
seguranç/a	22	64	20244	alta
avenid/a	19	47	280	alta
condição	70	13	39826	alta
precis/ando	1	4	587	média
menin/o	73	100	3487	média
bonit/o	68	100	2019	média
sobrinh/o	32	100	262	média
vestid/o	12	92	317	média

Tabela 17: Os 15 Nomes mais frequentes com aplicação de harmonia

Na Tabela 17: acima, observamos que os 15 types mais frequentes com aplicação de HV correspondem a 11 types de alta frequência e 4 de média no CBBS; adicionalmente, constata-se que não houve nenhum type de frequência baixa.

Na Tabela 18: a seguir, estão os verbos mais frequentes encontrados no corpus.

	VARSQL	%	CBBS	freq.
quer/ia	101	94	5515	alta
consequ/ir	99	74	14418	alta
pod/ia	98	94	21992	alta
precis/ar	97	86	114857	alta
acredit/ar	54	43	8821	alta
ped/ir	53	100	4527	alta
sent/ir	48	98	2818	alta
dev/ia	46	85	16089	alta
progred/ir	40	8	275	alta
receb/i	35	17	3066	alta
conhec/i	94	10	1831	média
modific/ar	75	0	1478	média
perd/i	56	0	446	média
cozinh/ar	53	30	106	média
dorm/ir	43	56	850	média

Tabela 18: Os 15 Verbos mais frequentes com contexto de harmonia

Como se pode verificar na Tabela 18:, houve 10 Verbos com contexto de HV que correspondem à frequência alta no CBBS e 5 de frequência média.

	VARISUL	%	CBBS	freq.
quer/ia	101	94	5515	alta
pod/ia	98	94	21992	alta
precis/ar	97	86	114857	alta
consequ/ir	99	74	14418	alta
ped/ir	53	100	4527	alta
sent/ir	48	98	2818	alta
dev/ia	46	85	16089	alta
acredit/ar	54	43	8821	alta
segur/ar	17	100	5564	alta
dorm/ir	43	56	850	média
acostum/ar	18	100	168	média
cozinh/ar	53	30	106	média
com/i	18	83	153	média
costur/ar	12	100	141	média
conhec/i	94	10	1831	média

Tabela 19: Os 15 Verbos mais frequentes com aplicação de harmonia

Para a Tabela 19:, há 9 types com verbos cuja frequência é alta e 6 com frequência média. Percebe-se que a aplicação de HV se dá em types verbais de frequência alta – em sua maioria e, novamente, não há paradigmas com baixa frequência.

Podemos observar que, nesta seção, não houve nenhum type com frequência baixa, imperando types com frequência alta.

Na seção a seguir, há uma discussão dos resultados encontrados nas subseções 1.1 e 1.2 e na seção 2.

3. FREQUÊNCIA DE TOKENS E DE TYPES NA HARMONIA VOCÁLICA

Nas seções anteriores, descrevemos os resultados de frequência de uso de palavras na harmonia vocálica, tanto para os tokens com a pretônica /e/ (cf. seção 1) e com a pretônica /o/ (ver seção 1.2), quanto para os types (cf. 2) com a pretônica. Nesta seção, fazemos a relação entre HV e a frequência de types e de tokens.

Para facilitar a discussão a seguir, apresentamos de forma resumida quantas

vezes ocorreram frequências altas, médias e baixas de tokens/types.

Os resultados de tokens **mais** frequentes com /e/ e /o/ (expostos nas subseções 1.1 e 1.2):

Mais frequente	Ambiente	alta	média	baixa
/e/	Contextos de HV: nomes e verbos	6	9	0
	Contextos de HV: nomes	6	6	3
	Contextos de HV: verbos	3	10	2
/o/	Contextos de HV: nomes e verbos	5	6	4
	Contextos de HV: nomes	5	5	5
	Contextos de HV: verbos	2	5	8

Tabela 20: Resumo dos resultados de **tokens mais frequentes** com /e/ e /o/ na HV

Para as palavras mais frequentes no VARSUL, o esperado era que fossem palavras de alta frequência no CBBS. No entanto, como se observa na Tabela 20:, os resultados não mostram isso, pois na maioria dos casos, palavras de frequência média superaram os casos de frequência alta, tanto para /e/, nos contextos de nomes e verbos e contextos de verbos, quanto para /o/, em contextos de nomes e verbos. Para contextos de HV com nomes em que a pretônica é /e/, o total de palavras de alta e média frequências foram iguais, superando os tokens de frequência baixa; para a pretônica /o/, os resultados com nomes são iguais. Para contextos de /o/ com verbos, o resultado é o contrário do esperado, já que palavras de frequência baixa superaram aquelas de frequência média e alta.

Assim, podemos interpretar que a suposição levantada foi parcialmente confirmada: tokens mais frequentes no VARSUL com contexto de HV não são aqueles que correspondem a palavras de alta frequência; de qualquer forma, na maior parte dos casos, tanto as palavras de frequência alta quanto as de frequência média sujeitas à HV superam as palavras de baixa frequência.

Na Tabela 21: a seguir, estão os casos de palavras **menos** frequentes no VARSUL:

Menos frequente	Ambiente	alta	média	baixa
/e/	Contextos de HV: nomes e verbos	10	2	3
	Contextos de HV: nomes	9	4	2
	Contextos de HV: verbos	3	9	3
/o/	Contextos de HV: nomes e verbos	8	5	2
	Contextos de HV: nomes	8	5	2
	Contextos de HV: verbos	4	7	4

Tabela 21: Resumo dos resultados de **tokens menos frequentes** com /e/ e /o/ na HV

A suposição para as tokens menos frequentes foi que essas palavras seriam de baixa frequência no CBBS. Como se observa no contexto de HV com nomes e verbos para a pretônica /e/, há mais palavras de alta frequência do que as de baixa frequência e essas, por sua vez, aparecem mais do que os tokens de média frequência. Os resultados de /e/ e /o/ para contextos de nomes e verbos e de /o/ para contextos de nomes mostram que há mais palavras de alta frequência, seguidas por tokens de frequência média e de baixa. Nos contextos de verbos com a pretônica /e/ e com a pretônica /o/, há mais tokens de frequência média, e aqueles de alta e baixa são iguais, ou seja, palavras de frequência média superam as de média e baixa frequência nesses casos.

Desses resultados, podemos constatar que a frequência de tokens parece não exercer um papel na HV, já que a relação entre esse processo e a frequência não resultou em um padrão. Passamos então aos paradigmas, ou seja, aos tokens.

O resumo dado na Tabela 22: abaixo diz respeito aos types:

Ambiente	alta	média	baixa
Contextos de HV: nomes e verbos	9	6	0
Contextos de HV: nomes	12	3	0
Contextos de HV: verbos	10	5	0
Aplicação de HV: nomes e verbos	11	4	0
Aplicação de HV: nomes	10	5	0
Aplicação de HV: verbos	9	6	0

Tabela 22: Resumo dos resultados de **types** na HV

Como se observa na Tabela acima, em todos os resultados de types mais frequentes no VARSUL, tantos de contextos de HV (de nomes e verbos, de nomes e de verbos) quanto de aplicação (de nomes e verbos, de nomes e de verbos), há mais palavras de frequências alta do que de média, e não aparece nenhum type de baixa

frequência.

Dessa forma, podemos identificar um padrão na frequência de types de HV, já que são todos de frequência altas, que sobrepujam os de frequência média, e não há types de baixa frequência.

Com base nas análises de tokens feitas no presente trabalho, pode-se dizer que a ideia condutora desse estudo foi corroborada em parte: há uma preferência de HV ser aplicada em frequências altas e médias, mas não há um padrão em frequências baixas.

Com relação às questões apresentadas em a) e b) na Introdução deste artigo, que seguem reescritas a seguir, pudemos verificar o seguinte:

- a) o efeito de frequência de tokens: existe até certo ponto, visto que há palavras de frequências alta e média, que superam as de frequência baixa; e
- b) o efeito de frequência de types: existe e se dá especificamente em types de frequência alta.

Como vimos na seção 2, os tokens neste trabalho foram constituídos quanto à igualdade da base morfológica das palavras, o que possibilitou verificar que há um efeito de frequência alta de types na HV. Pode-se dizer que:

- HV "enxerga" a base morfológica da palavra; ou
- HV tende a aplicar-se em conjuntos de palavras que tenham a mesma base morfológica.

Todavia, o que os resultados da análise apresentada até esse ponto permitem afirmar é o seguinte: HV tende a aplicar-se em types em que se manifestam tokens de alta frequência.

4. STATUS DA HARMONIZAÇÃO VOCÁLICA

Como um adendo ao exposto, busca-se, nessa seção, indícios do status de HV em dados do sul do país. Para tanto, foram organizados paradigmas constituídos por um conjunto de palavras fonológicas derivadas e não derivadas, flexionadas e não flexionadas. A presença de um prefixo não interfere nos resultados. A morfologia não é o pressuposto em discussão. Mas a condição para a entrada em paradigma é ser palavra fonológica com o contexto de HV: sílaba com vogal média seguida de sílaba

com vogal alta. Por sua vez, o paradigma subdivide-se em dois subconjuntos: a) sem efeito de HV e b) com efeito de HV. Da soma de ocorrências de cada subconjunto, depreende-se o status de HV no paradigma e do conjunto dos resultados dos paradigmas depreende-se o status de HV no sistema.

As palavras de cada paradigma têm em comum a base fonética. Por vezes, a base fonética e a fonológica coincidem. A base fonológica é delimitadora da extensão da palavra. Qualquer palavra, independentemente do sentido, pode entrar no paradigma desde que satisfaça as duas bases, fonética e fonológica. A base fonética uniformiza os conjuntos, a base fonológica especifica o núcleo da palavra fonológica e o identificador (Ident), delimita a sua abrangência a partir do núcleo. O identificador pode ser constituído de um só segmento, como /o/ em **menino**, ou mais, como em **meninos** e outros mais, como em **seguiria**.

Dada a existência na amostra de palavras repetitivas como **divia** e **pidia** sem a contraparte com vogal média, estabeleceu-se que cada subconjunto deve ser constituído, minimamente, de duas palavras fonológicas e cinco ocorrências e quanto à classificação dos conjuntos, tomou-se por norma o seguinte: Se o subconjunto com efeito de HV tiver dez ocorrências a mais de HV em relação ao subconjunto sem efeito de HV, é rotulado de progressão, isto é, HV em expansão; se os valores de ocorrência dos dois subconjuntos forem idênticos ou diferentes, mínima ou moderadamente, o conjunto é rotulado de estabilidade, isto é, HV estável no sistema.

Paradigma 1 : base fonética [segi], base fonológica /segi +IDENT/

a) Sem efeito de HV

seguir 1 seguido 3, seguidos 2
seguisse 1, seguissem 1, segui 1
prosseguir 2 ,perseguidor 1, seguidor 1
prossegua 2, perseguir 1, persegua 1

b) Com efeito HV

siguir 4, sigui 3, seguida 5 seguido 2, seguidos 1
siguinte 18, consiguir 7, consiguindo 4, seguindo 14
consiguimos 2, consiguia2, consigui 18,
consiguirem 2, consiguisses 1, conseguisse 1
prossiguia 1, seguissem 1

In absentia: seguimento, seguidamente , perseguição, perseguido, etc.

Sumário

Palavras por subconjunto	12	17
Não ocorrência e ocorrência de HV	17	86
Ocorrência no paradigma e classificação	103	progressão

Constituído de 29 palavras fonológicas e 103 ocorrências das quais 86 com HV e 17 sem HV, esse paradigma se mostra um condutor da expansão de HV, classificando-se na linha de progressão.

Paradigma 2 : base fonética [kostum], base fonológica /kostum+IDENT/

a) Sem efeito de HV

costuma 3, acostumando 2
acostumado 1, acostumo 6
acostumar 1, costumam 3, costumou 1
costumes 3, costumavam 2, costuma 3,
acostumou 3, acostumava 2,
costume 2, costume 4, costumamos 2,
acostumam 4, costumava 2, acostumar 1

In absentia: costumaz, costumagem, costumança, etc.

b) Com efeito HV

custume 4, costumes 3, costumava 2 costume 4
acostumou 3, acostumado 3, acostumei 3
acostumada 7, acostumadas 2, acostumam 1,
acostuma 3, costumamos 2, acostumo 1
custumei 2, costumavam 1, costumava 4,
acostumei 3, acostumar 2, acostumando 1,
custumados 1, custuma 3

Sumário

Palavras por subconjunto	18	21
Não ocorrência e ocorrência de HV	45	55
Ocorrência no paradigma e classificação	100	Progressão

Esse paradigma reúne 39 palavras fonológicas com 100 ocorrências, entre as quais 55 com efeito de HV e 45 sem esse efeito, uma diferença de 11 pontos que, de acordo com os cânones estabelecidos, é um paradigma com HV em progressão.

Paradigma 3: base fonética [poli], base fonológica /polit+ IDENT /

a) Sem efeito de HV

policial 4
policiais 2

In absentia: polícia, policiamento, policiado, políticas, politicagem, etc.

b) Com efeito HV

pulítica 23, pulítico 4, puíticos 6,
pulicial 31, puliciais 1

Sumário

Palavras por conjunto	2	5
Não ocorrência e ocorrência de HV	6	65
Ocorrências no paradigma	74	Progressão

Com 7 palavras fonológicas e 74 ocorrências das quais 68 com efeito de HV e 6 sem esse efeito, esse paradigma tem HV em progressão. Embora morfológicamente seja discutível a relação estabelecida entre **política** e **policia**, ela se sustenta, contando-se com a regra de assimilação que converte /t/em [s] diante da terminação -ia, uma regra de aplicação categórica.

Paradigma 4 : base fonética [vesti]; base fonológica /vesti + IDENT]

a) Sem efeito de HV

vestiam 1, vestida 4

vestidas 2, vestidos 1

vestimenta 1

investissem1

In absentia : investimento, investigação, vestidinho, etc.

b) Com efeito HV

vistir 2, vestida 3

vistido 5, vistidos 2

vistimentas 1

vistia 2

Sumário

Palavras por subconjunto	6	6
Não ocorrência e ocorrência	10	15
Ocorrências no Paradigma	25	Estabilidade

No paradigma acima, há 12 palavras fonológicas com 25 ocorrências das quais 10 sem efeito de HV e 15 com efeito, o que indica HV com estabilidade.

Paradigma 5: base fonética [felis]; base fonológica /feliS+IDENT/

a) Sem efeito de HV

feliz 2, infeliz 1, infelizes 2,

infelizmente 1, infeliz 1

In absentia: felicitar, felicitação, felicidades, etc.

b) Com efeito HV

filiz 1, filicidade 3, infiliz 1,

filizmente 1, infelizes 2

Sumário

Total de palavras por subconjunto	5	5
Não ocorrência e ocorrência de HV	7	8
Ocorrências no paradigma	15	Estabilidade

Constituído de 10 palavras fonológicas, entre as quais 7 sem efeito de HV e 8 com efeito, uma diferença mínima entre os subconjuntos (7/8), este paradigma tem HC com estabilidade.

Paradigma 6 : Base fonética [segu]; base fonológica [segur+ IDENT]

a) Sem efeito de HV

segurança 7, seguiu 2

seguro 3, segure 3

insegurança 1, segurava 2

segura 3, seguir 1

In absentia : inseguro, assegurar, assegurado, etc.

b) Com efeito HV

sigurança 13, siguiu 1, sigura 3

sigurador 1, sigurava 2

sigurando 2, sigure 2 segura

siguro 1, sigurar 3, siguradora 1

Sumário

Palavras por subconjunto	8	10
Não ocorrência e ocorrência	22	28
Ocorrências no paradigma	50	Estabilidade

Constituído de 18 palavras fonológicas e 50 ocorrências, das quais 22 sem efeito de HV e 28 com efeito, distinguindo-se apenas por quatro pontos (22/28), este paradigma tem HV com estabilidade.

Paradigma 7 Base fonética: [kozi]; bas fonológica /kozi+Ident /

a) Sem efeito de HV

cozinha 2, cozinhar 6

cozido 1, cozinjava 2

cozinhas 3, cozinheira 2

cozinheiro 4, cozinho 2

cozinheiras 2

In absentia: cozimento, cozidura, cozido. recozido , etc.

b) Com efeito HV

cuzinha 18, cuzinhar 10

cuzinhava 4, cuzido 1

cuzinho 2, cuzinheiro 2

cuzinhou 1, bicuzinho 1

cuzinheira 1

Sumário

Palavras por subconjunto	9	9
Não ocorrência e ocorrência de HV	24	40
Ocorrências no paradigma	64	Progressão

Com 18 palavras fonológicas e 64 ocorrências, este paradigma expõe uma diferença de 16 pontos entre os dois subconjuntos, com o privilégio do subconjunto com efeito de HV, caracterizando-se como paradigma em progressão.

Assim finda-se esse estudo que reorganizou os dados da amostra em pauta em paradigmas de acordo com os critérios estabelecidos. Foram encontrados sete paradigmas dos quais quatro (1, 2, 3, 7) com HV em progressão e três (4, 5, 6) com HV em estabilidade. Não é uma grande diferença, mas é expressiva, pois permite afirmar que HV é um regra com estabilidade no sistema que dá sinais de expansão. Todavia, não é um indicativo de mudança, mas é um indicativo de alteração no status de uma regra variável do sistema que pode tanto permanecer nesse estado por tempo indeterminado, assim como progredir ou regredir. Ressalte-se, porém, que a análise contou somente com dados do sul do país. Fica a demanda de dar continuidade a esse estudo com dados de outras variedades do PB com vistas a uma visão mais abrangente.

CONCLUSÃO

Nas três primeiras seções deste artigo, constatou-se que, na amostra em estudo representativa do sul do país, palavras com HV de alta frequência correspondem à frequência alta e média no corpus referencial proporcionalmente, confirmando o pressuposto de que regras variáveis são sensíveis a palavras de uso frequente. E que types constituídos de tokens de frequência alta refletem melhor a frequência de uso de HV. Na última seção (4), em que se busca informação sobre o status de HV, chega-se à conclusão de que HV é uma regra variável com estabilidade no sistema fonológico do português brasileiro com alguns indícios de expansão.

REFERÊNCIAS BIBLIOGRÁFICAS

- BERBER SARDINHA, T. (2004) **Linguística de Corpus**. Barueri-SP: Editora Manole.
- BYBEE, J.L. (1985) **Morphology: A study of the relation between meaning and form**. Philadelphia: Benjamins.
- _____ (2001) **Phonology and language use**. Cambridge: Cambridge University Press.
- _____ (2002) Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. **Language Variation and Change**, 14, 261-290. Cambridge University Press, 2002.
- _____ (2010) **Language, Usage and Cognition**. Cambridge: Cambridge University Press.
- CAMARA Jr., J.M. (1970) **Estrutura da língua portuguesa**. Petrópolis: Vozes.
- LINDBLOM, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In **Speech production and speech modeling**, ed. by William J. Hardcastle & A. Marchal, Kluwer, 403-39.
- PHILLIPS, B.S. (1984) Word Frequency and the Actuation of Sound Change. **Language**, Vol. 60, No. 2, pp. 320-342.
- SCHWINDT, L. C. (1995) **A harmonia vocálica em dialetos do sul do país: uma análise variacionista**. Dissertação de Mestrado. Pontifícia Universidade. Católica do Rio Grande do Sul.