

## A Fonotaxe dos ataques ramificados CCV em Português Brasileiro: frequência de uso, produtividade e aceitabilidade

Andressa Toni (University of São Paulo)

Esta pesquisa insere-se nos estudos sobre a fonotaxe do Português Brasileiro (PB), descrevendo em medidas empíricas (experimentais e de *corpora*) o conhecimento implícito do falante sobre as propriedades fonológicas, combinatórias e distribucionais das sílabas de ataque ramificado CCV (Consoante<sub>1</sub>+Consoante<sub>2</sub>+Vogal). Investigando a frequência de uso deste tipo silábico no *corpus* de médio-grande porte de Benevides & Guide (2016), observou-se que palavras contendo sílabas CCV são pouco frequentes no Léxico oral do PB (Tabela 1), totalizando apenas 7,3% de 1,9 milhões de *tokens*. Desse percentual, 74% concentra-se nas combinações segmentais /tr, pr, br/, sendo /tr/ a combinação mais frequente (28%), enquanto um total de 4,4% representa sequências pouco frequentes /dr, kl, gl/, sendo /gl/ a menos frequente (0,4%) – um número não tão distante dos 0,02% somados pelas sequências marginais /tl, dl, vl/. Considerando essas características quantitativas, o presente estudo questiona i) se CCV é realmente produtivo e gramatical em PB, ou se este seria um tipo silábico em desuso, lexicalizado/cristalizado (Cristófaros-Silva, 2003); e ii) se todas as combinações segmentais CCV são igualmente produtivas e bem formadas na língua, apesar de suas frequências de uso discrepantes. Visamos, com isso, contribuir à discussão sobre os efeitos que uma distribuição estatística esparsa e desbalanceada pode desencadear na Fonologia: como o conhecimento fonotático do falante é moldado/influenciado pelo input? (Coleman & Pierrehumbert, 1997; Gorman, 2011). A esse respeito, três comparações são oportunas: i) /tr/ vs. /dr/ (mesma classe natural, distribuição percentual diferente); ii) /gl/ vs. /tl, dl, vl/ (classes naturais diferentes, distribuição percentual semelhante); e iii) /tl, dl/ vs. /vl/ (uso marginal por motivos distintos: /tl, dl/ → dificuldade perceptual; /vl/ → acidente histórico (Chagas, 2020)). Para examinar a produtividade e a aceitabilidade fonotática de CCV e de suas combinações segmentais frequentes, pouco frequentes e marginais, conduzimos 2 experimentos com 25 universitários paulistanos: **Experimento 1 - Produtividade:** Os participantes foram solicitados a inventar pseudopalavras utilizando conjuntos pré-selecionados de 5 segmentos (2 obstruintes e 1 líquida estabelecidas conforme sua frequência em relação a CCV, ALTA, BAIXA, MARGINAL, além de 2 vogais), sendo levados a criar ataques ramificados ou codas. A tarefa examinou se a posição alocada às líquidas /l, r/ (CCV ou CVC) variaria dependendo da qualidade da obstruinte. O objetivo foi verificar se o molde CCV e todas as suas combinações segmentais estão disponíveis à criatividade fonotática do falante apesar de sua frequência de uso. Os resultados apontam que 28,45% das pseudopalavras formadas continham CCV, indicando que esta sílaba é produtiva na língua. Constatou-se significativa tendência à formação CVC nas condições BAIXA e MARGINAL, embora dentre os CCVs formados não haja diferença significativa entre a proporção de sequências de ALTA ou BAIXA frequência, especialmente em comparação ao estudo de *corpora* (Gráfico 2). Observa-se também diferença significativa entre a proporção de formação de /tl, dl, vl/ vs /tr, pr, br/ e /dr, kl, gl/, mas não entre /tl, dl/ vs. /vl/, indicando que causas perceptuais ou históricas não afetam diferentemente a produtividade fonotática. **Experimento 2 - Julgamento de aceitabilidade:** Os mesmos participantes foram solicitados a julgar em escala de 1 a 7 a aceitabilidade de pseudopalavras contendo CCVs de frequência ALTA, BAIXA, MARGINAL e de CCVs não atestados na língua, de sonoridade não-marcada (SSP: /ʃr, xl, bn/) ou marcada (\*SSP: /ft, lb, řt/). CCVs não-atestados apresentaram média entre 2,5 e 1,6 pontos (/bn, ʃr/ > /ft/ > /řt/ > /xl/ > /lb/); sequências de alta e baixa frequência apresentaram aceitabilidade entre 5,5 e 4,5 pontos (/pr, br/ > /tr/ > /gl/ > /dr/ > /kl/); e sequências marginais obtiveram nota 3 (/vl/ > /dl/ > /tl/) (Gráfico 3). Em conjunto, os Experimentos 1 e 2 indicam que CCV e suas combinações consonantais são produtivos e bem-formados na língua, não se caracterizando como uma estrutura cristalizada apesar de sua baixa frequência de uso – embora uma comparação à formação CVC possa sugerir um processo incipiente de desuso (Cristófaros-Silva, 2003). Nota-se também que a distribuição estatística não reflete diretamente a gramática

fonotática da língua: i) CCVs de frequências distintas são igualmente produtivos e aceitáveis (/tr/ vs. /dr/); e ii) CCVs de frequências semelhantes apresentam produtividade e aceitabilidade diferentes (/gl/ vs. /tl/) ou semelhantes (/dl/ vs. /vl/). Esses resultados indicam que diferenças na aceitabilidade e produtividade não motivam a distribuição estatística dos dados do PB.

Types	r	l	Total	Types	r	l	Total
p	1.554 (23,1%)	381 (5,7%)	1.935 (28,8%)	p	52.453 (36,9%)	5.444 (3,8%)	57.897 (40,7%)
b	685 (10,2%)	111 (1,7%)	796 (11,8%)	b	13.182 (9,3%)	2.527 (1,8%)	15.709 (11%)
t	1.796 (26,7%)	11 (0,2%)	1.807 (26,9%)	t	39.889 (28%)	24 (0,0%)	39.913 (28%)
d	218 (3,2%)	0 (0%)	218 (3,2%)	d	2.342 (1,6%)	0 (0%)	2.342 (1,6%)
k	527 (7,8%)	255 (3,8%)	782 (11,6%)	k	7.578 (5,3%)	3.379 (2,4%)	10.957 (7,7%)
g	598 (8,9%)	60 (0,9%)	658 (9,8%)	g	8.765 (6,2%)	584 (0,4%)	9.349 (6,6%)
f	346 (5,2%)	144 (2,1%)	490 (7,3%)	f	3.894 (2,7%)	820 (0,6%)	4.714 (3,3%)
v	31 (0,5%)	1 (0,0%)	32 (0,5%)	v	1.431 (1,0%)	3 (0,0%)	1.434 (1%)
Total	5.755	963	6.718 (100%)	Total	129.534	12.781	142.315 (100%)

Tabela 1: Frequência de uso das sílabas CCV em PB – corpus ABG de língua oral

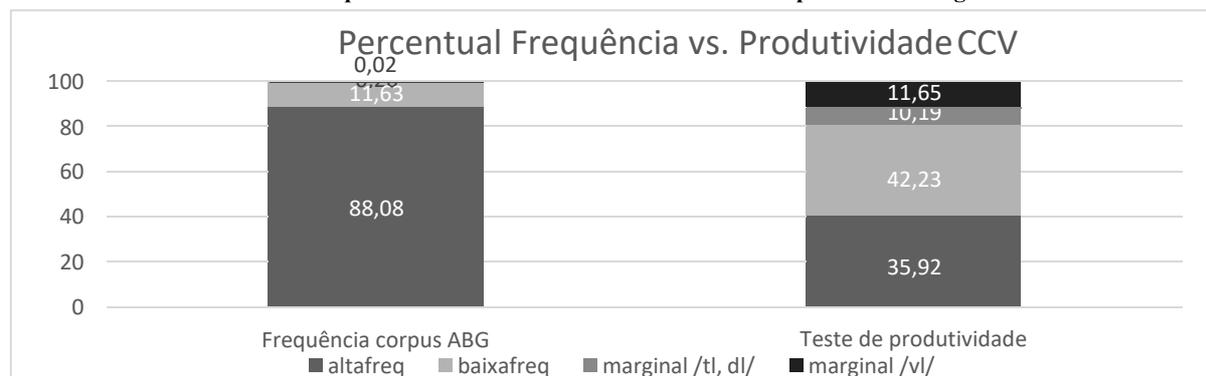


Gráfico 1: Frequência versus produtividade experimental de CCVs de frequência alta, baixa e marginal

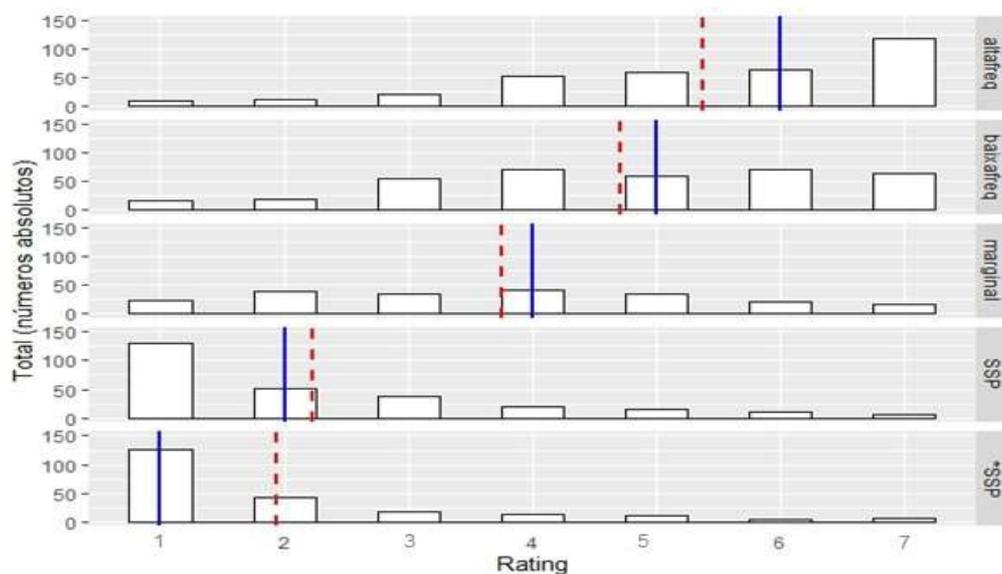


Gráfico 2: Julgamentos de aceitabilidade - frequência alta, baixa, marginal e ausente (não-marcado, marcado) (Vermelho: nota média; Azul: nota mediana; 1: palavra inaceitável; 7: palavra bem-formada)

Referências: ♦ Benevides, A.; Guide, B. Corpus ABG. Revista Texto Livre: Linguagem e Tecnologia, volume 10, número 1, 2017 ♦ Chagas, P. Ataques Complexos no Português. Pré-print, 2020 ♦ Coleman, J.; Pierrehumbert, J. Stochastic phonological grammars and acceptability. In: Third meeting of the ACL Special Interest Group in Computational Phonology, Somerset: Association for Computational Linguistics, p.49–56, 1997 ♦ Cristófaros-Silva, T. Sound change in tautosyllabic consonantal clusters in Brazilian Portuguese. In: Solé, Recasens & Romero, eds. Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS). Barcelona: Causal Productions. pp. 1675–1678, 2003. ♦ Gorman, K. A program for phonotactic theory. In: Proceedings of CLS, n. 47(1), p. 79–93, 2011.