

**Efeito de *Wordlikeness* no processamento de não-palavras por falantes do português brasileiro**

***Effect of Wordlikeness in Processing of Nonwords by Brazilian Portuguese Speakers***

Christina Abreu Gomes

Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Rio de Janeiro, Brasil.  
christina-gomes@uol.com.br

Suzana do Couto Mendes

Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Rio de Janeiro, Brasil.  
suzaninha\_15@yahoo.com.br

Marcela Branco Silva

Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro – Rio de Janeiro / Brasil.  
tchelarj@yahoo.com.br

Clara Oliveira Esteves

Universidade Estadual da Bahia (UNEB), Salvador, Bahia, Brasil.  
clarinhaesteves@yahoo.com.br

Gastão Coelho Gomes

Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Rio de Janeiro, Brasil.  
gastao@im.ufrj.br

**Resumo:** O presente estudo aborda a questão da emergência de uma competência fonológica abstraída das formas das palavras representadas no léxico, conforme estabelecido nos Modelos baseados no Uso, considerando resultados de um teste de aceitabilidade de não-palavras.

Efeito de wordlikeness, o grau de semelhança entre uma não-palavra e uma palavra real, tem sido observado em diversos estudos. Não-palavras formadas por sílabas de alta e baixa probabilidade e de diferentes tamanhos foram especialmente elaboradas para testar o efeito de seu grau de semelhança com as palavras reais no julgamento do falante. A frequência das sílabas foi obtida de uma amostra *online* de textos do português brasileiro. Os estímulos foram julgados por 22 adultos universitários. Os resultados mostraram que os julgamentos se basearam nas propriedades probabilísticas dos estímulos. Os resultados constituem evidência adicional da relação entre conhecimento fonológico e inferência estatística no léxico.

**Palavras-chave:** Wordlikeness; Não-palavra; Léxico; Gramática emergente; Modelos baseados no Uso.

**Abstract:** The present study addresses the issue of the emergent phonological competence abstracted from the wordforms represented in the lexicon, as stated in the Usage-based Models, considering the results of an acceptability test of nonwords. Effect of wordlikeness, the degree of similarity between a nonword and a real word, has been observed in several studies. Nonwords formed by syllables of high and low probability and of different length were especially elaborated in order to check the effect of their wordlikeness in the speaker judgment. The frequency of the syllables were obtained from an online database of texts of Brazilian Portuguese. The stimuli were judged by 22 university speakers. The results showed that judgments were based on the probabilistic properties of the stimuli. The results constitute further evidence for the relationship between phonological knowledge and statistical inference in the lexicon.

**Keywords:** Wordlikeness; Nonword; Lexicon; Emergent grammar; Usage-based Models.

Recebido em 23 de maio de 2014.

Aprovado em 13 de outubro de 2014.

## Introdução

Neste artigo, são apresentados os resultados relativos a um teste de julgamento de *wordlikeness* de não-palavras, grau de semelhança entre a não-palavra e palavras existentes no léxico, aplicado a falantes adultos do português brasileiro.<sup>1</sup> Há poucos estudos sobre reconhecimento de palavra falada tendo como língua de estudo o português brasileiro; portanto, há diversas lacunas a serem preenchidas, e o estudo proposto pode contribuir com evidências para um debate importante relativo ao conhecimento linguístico abstraído pelo falante no que diz respeito à relação entre gramática fonológica e léxico, gerenciamento probabilístico do conhecimento linguístico e seu efeito na produção, processamento e aquisição.

Diversos trabalhos que focalizam a maneira como os falantes reconhecem as palavras faladas têm mostrado que há vários fatores que afetam a percepção e o processamento dos estímulos linguísticos, tais como o grau de semelhança sonora entre os itens lexicais, frequência de uso do item e relação entre os segmentos que compõem a palavra. Também tem sido observado o efeito da densidade de vizinhança, relação entre palavras no léxico que compartilham segmentos, como também o efeito de semelhanças que envolvem relações morfológicas entre os itens lexicais (LUCE; PISONI, 1998; VITEVITCH; LUCE, 1999; VITEVITCH; STAMER, 2006).

No entanto, para evitar a interferência dessas variáveis, há estudos que utilizam estímulos de não-palavras formados por sequências fonotáticas tanto existentes quanto não existentes na língua para avaliar o conhecimento fonológico dos falantes. Esses estudos abordaram a aquisição fonológica em crianças com desenvolvimento típico e população clínica (GATHERCOLE, 1995; ARCHIBALD; GATHERCOLE, 2006; BECKMAN; MUNSON; EDWARDS, 2007), como também o conhecimento fonológico de adultos (FRISCH; LARGE; PISONI, 2000; SAITO *et al.*, 2003) e sua relação com aspectos do acesso lexical (VITEVITCH; LUCE, 2005). Nos estudos linguísticos, não-palavras formadas por sequências possíveis na língua são também denominadas pseudopalavras (BROWN; HILDUM, 1956, *apud* FRISCH; LARGE; PISONI, 2000). Os estímulos baseados em pseudopalavras

---

<sup>1</sup> Essa pesquisa foi financiada pelo CNPq (Processo no. 304400/2010-6) e pela FAPERJ (Processo no. 26/102.405/2009).

variam o grau de similaridade em relação às palavras de uma língua. Esse grau de semelhança é tratado sob o rótulo de *wordlikeness* nos diversos estudos que envolvem reconhecimento de não-palavras.

Diversos trabalhos têm mostrado que o grau de *wordlikeness* tem uma forte influência na maneira como as palavras são processadas em tarefas de repetição, julgamento de aceitabilidade, reconhecimento e memória. Segundo Frisch, Large e Pisoni (2000), os estudos sobre o efeito de *wordlikeness* são de dois tipos: ou focam a semelhança entre palavras do léxico e estímulos de pseudopalavras ou examinam a influência da estrutura fonotática da pseudopalavra. Por exemplo, Greenberg e Jenkins (1964, *apud* FRISCH; LARGE; PISONI, 2000), em uma das primeiras investigações sobre a influência da similaridade de pseudopalavras em relação a palavras do léxico, mediram a distância entre pseudopalavras monossilábicas do tipo CCVC e palavras do inglês, usando o método de substituição de fonema. O resultado obtido para os julgamentos que os falantes fizeram dos estímulos mostrou estar relacionado ao escore de distância obtido entre não-palavras e palavras do inglês. Assim, Greenberg e Jenkins propuseram que há um espaço fonológico no qual as palavras são organizadas no léxico de acordo com a sua estrutura sonora, medido em função do número e da frequência de outras palavras que ocupam a similaridade de vizinhança nesse espaço. Assim, combinações de segmentos que são mais comuns normalmente ocorrem em vizinhança densa no léxico, ao passo que combinações não comuns se situam em vizinhanças com densidade esparsa.

Frisch, Large e Pisoni (2000) conduziram três experimentos para avaliar o efeito da estrutura das não-palavras na percepção de adultos. Os estímulos dos experimentos correspondiam a pseudopalavras de duas, três e quatro sílabas com padrão CV nas sílabas iniciais e mediais e CVC na sílaba final (CV.CVC; CV.CV.CVC; CV.CV.CV.CVC). O padrão acentual foi o mais comum do inglês para cada tamanho de item. Além disso, cada pseudopalavra era composta inteiramente ou por difones (dois sons adjacentes na mesma palavra independentemente de ocuparem a mesma sílaba) com relação fonotática de alta frequência ou de baixa frequência na língua. Os resultados indicaram que os estímulos com sequências de alta probabilidade foram julgados mais semelhantes a itens do léxico do que os com baixa probabilidade nos 3 tamanhos de estímulos, havendo um decréscimo na aceitabilidade em função do tamanho, e, ainda, que pseudopalavras com constituintes de alta

probabilidade foram reconhecidos mais acuradamente do que padrões de pseudopalavras com baixa probabilidade.

Esse design de experimento se baseia no fato de que há evidências de que padrões fonotáticos probabilísticos têm influência no processamento e na produção fonológica. Coleman e Pierrehumbert (1997) propuseram que a aceitabilidade dos julgamentos de pseudopalavras envolve uma avaliação completa da composição da pseudopalavra, calculada na probabilidade fonotática cumulativa, não sendo, portanto, o resultado de um julgamento categórico que recusa como palavra possível de uma língua um estímulo formado por combinações não possíveis de segmentos da língua. Eles observaram que estímulos formados por sequências não existentes no inglês, mas incluindo também sequências fonotáticas de alta frequência, não foram rejeitados pelos falantes; ao contrário, receberam um escore avaliativo semelhante ao dos estímulos formados somente com sequências fonotáticas de baixa frequência. Esses resultados indicam que os falantes têm uma percepção das relações fonotáticas em termos probabilísticos e não de forma categórica e que essa probabilidade tem um papel importante nos julgamentos de aceitabilidade.

Segundo Frisch *et al.* (2000), uma gramática estocástica, que gerencia probabilisticamente as relações entre os constituintes, pode prover uma maneira de unificar a influência da probabilidade fonotática e da vizinhança lexical, essa última definida como a quantidade de itens que compartilham semelhança de pelo menos um segmento com outras palavras no léxico. Uma vez que a probabilidade métrica esperada leva em consideração a composição total da pseudopalavra, a gramática estocástica reflete o número de palavras no léxico que são semelhantes à pseudopalavra em algum grau. Uma pseudopalavra que agregue diversos constituintes de baixa probabilidade estará localizada numa região esparsa no espaço de similaridade do léxico, ao passo que uma pseudopalavra formada por constituintes de alta frequência estará situada em uma região densa. A hipótese principal para os autores é a de que padrões abstraídos do léxico provêm a base de uma competência fonológica emergente, que permite o processamento de não-palavras.

Susan Gathercole (1995) também investigou em que medida a performance de crianças com desenvolvimento típico entre 4 e 5 anos em teste de repetição de não-palavra é condicionada pela memória de trabalho fonológica ou pelas representações lexicais de longo termo. A autora também observou que a acurácia de repetição foi melhor para

estímulos classificados com alta probabilidade de *wordlikeness*. O escore de *wordlikeness* foi obtido em estudo realizado em Gathercole, W.; Emslie e Baddelly (1991). A autora argumenta que a repetição de pseudopalavras com baixo grau de *wordlikeness* depende da memória fonológica, ao passo que as com alto grau de *wordlikeness* são mediadas pelas representações lexicais de longo termo e, portanto, menos afetadas por restrições da memória fonológica.

Este artigo se estrutura da seguinte maneira: na seção a seguir apresentamos as hipóteses e a metodologia utilizada na elaboração das não-palavras, a descrição dos sujeitos e a aplicação do teste. Em seguida, são apresentados e discutidos os resultados obtidos, seguidos das considerações finais.

## **1 Hipótese e metodologia**

A hipótese básica da pesquisa é a de que a gramática fonológica é probabilística e gradiente. As informações relativas à forma sonora das palavras, às relações fonotáticas e morfofonológicas estão organizadas em um conjunto de representações com diferentes graus de abstração, que correspondem a diferentes tipos de conhecimento fonológico: representação fonético-paramétrica das propriedades acústicas e articulatórias dos sons que compõem as formas das palavras, representação detalhada das palavras no léxico e níveis mais abstratos que envolvem as relações fonotáticas, a prosódia e as relações morfofonológicas (PIERREHUMBERT, 2003). Nessa perspectiva, a gramaticalidade, ou a percepção do falante em relação às formas de sua língua, não é uma questão de pertencer ou não a uma categoria de forma categórica (gramatical vs. agramatical), mas se dá de forma probabilística e também gradiente, baseada nas distribuições de sequências de segmentos abstraídas das representações das palavras no léxico. Em outras palavras, o conhecimento gramatical probabilístico é derivado do léxico, mas também é independente deste, no sentido de que diferentes tipos de abstração são estabelecidos, com base nas representações. Diversos trabalhos têm mostrado que essa gramática que contém inferências probabilísticas é capaz de acessar a probabilidade de uma sequência independentemente de sua existência nas palavras estocadas no léxico, conforme foi mencionado na seção anterior em referência aos trabalhos de Coleman e Pierrehumbert (1997) e Frisch, Large e Pisoni (2000).

Conforme se postula nos Modelos baseados no Uso (BYBEE, 2001, 2010; PIERREHUMBERT, 2003, 2012), o léxico é concebido de forma dinâmica e redundante. Nele estão representadas todas as ocorrências dos itens lexicais que fazem parte da experiência do falante de produzir e ouvir. Estão também representadas tanto as palavras irregulares quanto as formações regulares, formas de singular e de plural, formas flexionadas verbais e *chunks*, unidades maiores que as palavras e que funcionam como uma unidade lexical. Além disso, o léxico é concebido não como uma lista não estruturada de itens, mas como uma rede de relações lexicais baseadas em similaridade sonora e semântica entre os itens representados. As relações morfofonológicas, relações fonotáticas e a gramática fonológica são emergentes das representações no léxico. Observa-se, segundo os Modelos baseados no Uso, que a frequência de ocorrência do item tem impacto na representação e que a produtividade em que um padrão é aplicado a novos itens incorporados no léxico e aos neologismos tem relação com a frequência de tipo das estruturas abstratas emergentes.

Assim, o esperado no processamento de não-palavras por falantes do português brasileiro através de uma tarefa de aceitabilidade é que o julgamento dos falantes reflita as distribuições probabilísticas observadas para os constituintes dos estímulos utilizados no teste. Essas hipóteses se inserem no quadro teórico dos Modelos baseados no Uso ou Modelos Multirrepresentacionais, que propõem, conforme foi apresentado anteriormente, uma organização probabilística do conhecimento linguístico para dar conta da representação mais detalhada da forma sonora dos itens lexicais. O conhecimento linguístico é entendido não como um conjunto mínimo de restrições ou regras categóricas, mas como um conjunto redundante de informações gradientes, que podem ser caracterizadas por uma distribuição estatística. Assim, as estruturas emergem do sistema de armazenagem e não são apriorísticas.

Um teste de não-palavras com sequências de sílabas possíveis do português foi elaborado por Esteves (2013), de acordo com a metodologia proposta por Frisch, Large e Pisoni (2000) com o objetivo de avaliar o conhecimento fonológico de crianças com desenvolvimento típico e atípico. Esses estímulos foram avaliados quanto à aceitabilidade para testar o grau de *wordlikeness*, isto é, se o grau de semelhança com palavras do léxico tem relação com o julgamento de falantes adultos.

Para elaboração do teste foi realizado um levantamento na Base Aspa/UFMG ([www.projetoaspa.org](http://www.projetoaspa.org)) da frequência de ocorrência de todas as combinações de consoantes e vogais do português em todas as posições prosódicas de dissílabas, trissílabas e palavras de quatro sílabas do português com estrutura silábica CV (consoante e vogal) e acentuação paroxítona. Optou-se por trabalhar com a frequência fonotática da sílaba e não de difones. Também optou-se, primeiramente, por estímulos com pseudopalavras estritamente formadas por sílabas CV, devido às restrições de sílabas com travamento silábico (que terminam em consoantes) ou formadas por grupos consonantais no português brasileiro. A frequência foi utilizada para calcular a probabilidade de cada sílaba, tomando o total de palavras que contêm aquela sílaba particular dividido pelo total de palavras que contêm aquele tipo de sílaba naquela posição prosódica (tônica ou átona) de acordo com a metodologia de Frisch *et al.* (2000). A métrica final (probabilidade da pseudopalavra) foi obtida através da multiplicação da probabilidade de cada sílaba que compõe o estímulo.

Para cada um dos três tamanhos considerados de pseudopalavras, há 5 pseudopalavras com combinações inteiramente de sílabas de alta frequência de tipo e 5 pseudopalavras com combinações de sílabas de baixa frequência, totalizando 30 estímulos (Anexo I). Estão previstas etapas futuras de ampliação dos estímulos com inclusão de todas as estruturas silábicas da língua. O total de estímulos será ampliado, mas não será possível atingir a mesma quantidade de estímulos por tipo (tamanho x frequência) do teste de Frisch *et al.* (2000) que foi constituído por estímulos com um molde lexical que inclui sílaba final do tipo CVC, o que amplia as possibilidades de combinação entre segmentos em três posições, considerando-se ainda o fato de que há mais possibilidades de ocorrência de consoantes na coda tendo como base a língua inglesa, se comparado com o português. Além disso, também concorre para uma menor quantidade de estímulos no presente teste o fato de o molde lexical para os três tamanhos ser de estímulos com acentuação na penúltima sílaba, sendo, portanto, a última sílaba átona, já que há restrição, no português brasileiro, do tipo de vogal que ocorre nesta posição, reduzindo as possibilidades de combinação de consoantes e vogais. Não houve o uso de distratores no teste devido ao tipo de estímulo usado. O teste já foi aplicado na avaliação de crianças com desenvolvimento atípico (ESTEVES, 2013), adultos com afasia (SENNA, 2013), crianças com desenvolvimento típico entre 5 e 9 anos



(MENDES, 2014) e em estudo longitudinal de crianças com aquisição tardia (SILVA, 2014).

Os limites de alta e baixa frequência foram estabelecidos através da distribuição dos valores probabilísticos em quartis para cada posição prosódica da sílaba em cada tamanho considerado. Uma métrica probabilística foi obtida através da multiplicação das probabilidades de cada sílaba da pseudopalavra, o que resulta em um contínuo de probabilidades que decresce em função do tamanho da palavra (COLEMAN e PIERREHUMBERT, 1997).

Além disso, conforme já foi mencionado anteriormente, as pseudopalavras são todas com acentuação paroxítona, molde lexical mais frequente no português para os 3 tamanhos considerados segundo a pesquisa realizada na Base Aspa. O levantamento indicou ser o padrão de acentuação paroxítona predominante nos 3 tamanhos de palavras considerados, conforme pode ser observado na Tabela 1, a seguir.

**Tabela 1** - Frequência de padrão acentual em função do número de sílabas da palavra

	2 SÍLABAS	3 SÍLABAS	4 SÍLABAS
TOTAL	13.929	35.427	39.986
PROPAROXÍTONA		2.278	3.881
PAROXÍTONA	<b>8.427</b>	<b>23.858</b>	<b>29.776</b>
OXÍTONA	5.502	9.293	6.270

Foram excluídas as pseudopalavras cujas combinações de sílabas se assemelhavam a morfemas do português para evitar sequências que podem contribuir para uma análise morfológica da palavra, sobrepondo-se à frequência de tipo das sílabas ou competindo com esta, como, por exemplo, [pade'nava] com a terminação –ava, que poderia ser analisada morfológicamente como um verbo da 1ª conjugação no pretérito imperfeito na terceira pessoa do singular, como *cantava*. Também foram descartadas as pseudopalavras que formam “pares mínimos” com palavras do léxico, como ['bada]/bata e ['kaba]/cada, e sequências com

segmentos repetidos, como [patu'tuva].<sup>2</sup>

Os estímulos foram gravados por uma mulher de 22 anos em um estúdio de gravação com equipamento digital. Na gravação, a forma fonética das pseudopalavras se acomodou às realizações fonéticas presentes na variedade dos sujeitos testados (p. ex. redução de vogal átona final, realização de africada pós-alveolar seguida de [i]).

O teste foi aplicado a 22 sujeitos falantes do português brasileiro, alunos de graduação da Universidade Federal do Rio de Janeiro (UFRJ). A aplicação individual do teste foi feita através do Programa PsyScope, que apresenta uma ordenação aleatória dos itens para cada julgador.

Os sujeitos foram expostos a 5 estímulos (também pseudopalavras) na fase de treino e, em seguida, aos estímulos do teste. Cada sujeito foi solicitado a aceitar ou rejeitar se o estímulo ouvido se assemelha a uma palavra existente na língua, respondendo sim ou não e apertando a tecla correspondente do teclado. Embora, como hipótese, tenhamos definido a aceitabilidade de forma gradiente, a aplicação do piloto do teste considerou somente respostas do tipo sim vs. não, considerando que a gradualidade poderia emergir das variáveis controle: frequência dos constituintes e tamanho dos estímulos.

## 2 Resultados

Os resultados obtidos se referem ao efeito das duas variáveis controladas, tamanho da pseudopalavra e da frequência das sílabas constituintes da pseudopalavra, e as variáveis dependentes são o tipo de resposta (sim vs. não) e o tempo de resposta.<sup>3</sup> Com relação ao tempo de resposta não houve diferença significativa entre os tempos de resposta e as variáveis analisadas. O teste ANOVA não se revelou significativo para tempo de resposta e tamanho ( $p(>f)$  0.1613), tempo de resposta e frequência da sílaba ( $p(>f)$  0.3813) e para a correlação entre ambas as variáveis ( $p(>f)$  0.5921). Esse resultado pode se dever ao fato de que não está em jogo uma tarefa de decisão que envolve palavras do léxico e pseudopalavras, mas o grau de semelhança entre estímulos e palavras reais. No teste em questão todos os estímulos são da mesma natureza, ou seja, inexistentes na língua.

<sup>2</sup> Não foi possível, no entanto, evitar a situação de par mínimo para algumas das pseudopalavras de duas sílabas. É o caso de [tʃibə], que forma par com tina.

<sup>3</sup> Frisch *et al.* (2000) não consideraram tempo de resposta como variável dependente.

Por outro lado, os resultados do teste de  $X^2$  (chiquadrado) indicaram que as variáveis tamanho dos estímulos (2, 3 ou 4 sílabas) e a frequência de ocorrências das sílabas CV (alta ou baixa) foram significativas em relação à resposta apresentada ( $X$ -squared = 32.25,  $df = 1$ ,  $p$ -value = 1.356e-08). Conforme pode ser observado na Tabela 2 a seguir, pseudopalavras formadas com sílabas de alta frequência foram mais aceitas do que aquelas com constituintes de baixa frequência.

**Tabela 2** - Tipo de resposta em função do número de sílabas

FREQUÊNCIA DAS SÍLABAS	SIM		NÃO		TOTAL
BAIXA	102	31%	228	69%	330
ALTA	175	53%	155	47%	330
TOTAL	272	41%	382	59%	660

Com relação ao tamanho das pseudopalavras, o teste do  $X^2$  revelou ser significativo o efeito do tamanho da pseudopalavra na avaliação dos estímulos ( $X$ -squared = 57.4952,  $df = 2$ ,  $p$ -value = 3.274e-13). Os resultados obtidos, e apresentados na Tabela 3 abaixo, indicaram que estímulos com 2 sílabas foram mais aceitos que os de três sílabas, que, por sua vez, foram mais aceitos que os de 4, ou seja, palavras menores foram mais aceitas que palavras maiores.

**Tabela 3** -Tipo de resposta em função do tamanho

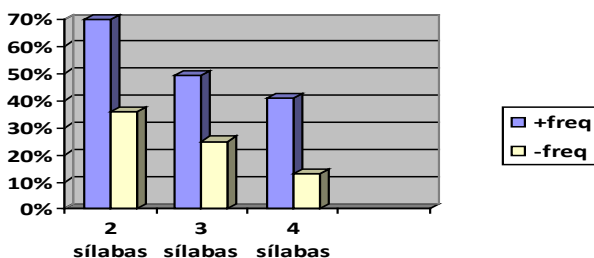
TAMANHO	SIM		NÃO		TOTAL
2 SÍLABAS	136	62%	84	38%	220
3 SÍLABAS	81	37%	139	63%	220
4 SÍLABAS	60	28%	160	72%	220
TOTAL	277	42%	383	58%	660

Os resultados obtidos fornecem evidências sobre o gerenciamento probabilístico de relações fonotáticas na gramática fonológica. A hipótese principal é a de que padrões abstraídos do léxico provêm a base de uma

competência fonológica emergente, que permite o processamento de pseudopalavras. Ou seja, padrões recorrentes no léxico, que constituem a memória de longo termo, fornecem evidência para o processamento das pseudopalavras. Esses resultados se assemelham aos encontrados por Frisch, Large e Pisoni (2000) para o inglês.

O efeito do tamanho do estímulo no julgamento de aceitabilidade da pseudopalavra para o português brasileiro também segue o mesmo padrão observado para o inglês, em que palavras menores são mais aceitas que palavras maiores. No entanto, é importante ressaltar que o levantamento realizado na Base ASPA/UFMG, para estabelecer o padrão acentual mais recorrente para cada conjunto de palavras reais em função do número de sílabas, revelou que as palavras de 2 sílabas são menos frequentes no léxico (3.929) do que as de 3 (35.427) e 4 sílabas (39.986), que apresentam uma quantidade aproximada de itens. Embora o molde lexical com 2 sílabas seja menos frequente no léxico se comparado com o das palavras de 3 e 4 sílabas, a probabilidade de ocorrência de cada pseudopalavra obtida através da métrica probabilística (multiplicação das probabilidades de cada sílaba) indica que quanto maior a palavra menor sua probabilidade de ocorrência, uma vez que as palavras maiores multiplicam mais probabilidades. Assim, palavras maiores de alta probabilidade podem ter uma métrica probabilística menor que palavras menores de baixa probabilidade. De fato, conforme pode ser observado no Gráfico 1, a seguir, há um decréscimo na aceitabilidade do estímulo gradativamente em função do tamanho e da probabilidade de ocorrência. Esse gráfico sumariza os resultados apresentados nas Tabelas 2 e 3.

**Gráfico 1** - Aceitabilidade dos estímulos em função do tamanho e da frequência



No Gráfico 1 estão os percentuais de respostas “sim” para o tamanho e a frequência dos constituintes dos estímulos, indicando que pseudopalavras com constituintes mais frequentes foram mais aceitas que as com constituintes de baixa frequência e que estímulos menores foram mais aceitos que constituintes maiores. A distribuição observada no Gráfico 1 também indica que o efeito do tamanho do estímulo parece ter como base a probabilidade de ocorrência daquele estímulo em função das probabilidades de suas subpartes e não a frequência do molde lexical, definido em função da quantidade de palavras com o mesmo número de sílabas no léxico.

### **Considerações finais**

Estudos como os descritos neste artigo fornecem evidências acerca da natureza do léxico e da relação entre léxico e fonologia. Segundo Pierrehumbert (2003), estudos de percepção sustentam a proposição teórica de uma fonologia emergente em diversos níveis de abstração, a partir da experiência do falante com sua língua em ouvir e produzir e do léxico armazenado. De acordo com a autora, um léxico adulto é de tamanho suficiente para permitir a abstração de todas as estruturas relevantes para a produção e a percepção.

Esses resultados podem ser tomados como evidência de que o conhecimento fonotático emerge de padrões segmentais inferidos no léxico mental e tem relação com sua probabilidade de ocorrência no léxico (COLEMAN; PIERREHUMBERT, 1997).

Os resultados obtidos nessa etapa de trabalho subsidiarão a continuidade do projeto de constituição de um instrumento de avaliação do conhecimento fonológico, tanto de adultos quanto de crianças, com possibilidade de aplicação na clínica fonoaudiológica. A ampliação do teste em número de estímulos dos tipos utilizados no presente estudo e com outros tipos silábicos do português permitirá avaliar melhor a inferência probabilística de estruturas fonológicas abstraídas no léxico.

### **Referências**

ARCHIBALD, L. M. D.; GATHERCOLE, S. Nonword Repetition: a Comparison of Tests. *Journal of Speech, Language and Hearing Research*, v. 49, p. 970-983, 2006.

BYBEE, J. *Phonology and Language Use*. Cambridge: Cambridge University Press, 2001.

BYBEE, J. *Language, Use and Cognition*. Cambridge: Cambridge University Press, 2010.

BECKMAN, M. E.; MUNSON, B.; EDWARDS, J. The Influence of Vocabulary Growth on Developmental Changes in Types of Phonological Knowledge. In: COLE, J. & HUALDE, J. (Eds.). *Laboratory Phonology*, 9, p. 241-264. New York: Mouton de Gruyter, 2007.

COLEMAN, J.; PIERREHUMBERT, J. B. Stochastic Phonological Grammars and Acceptability. *Computational phonology. 3rd Meeting of the ACL Special Interest Group in Computational Phonology: Proceedings of the Workshop*, 12 July 1997. Association for Computational Linguistics, Somerset, NJ, p. 49-56, 1997.

ESTEVES, C. O. *O conhecimento fonológico de crianças com dislexia, desvio fonológico e distúrbio específico da linguagem: uma análise multirrepresentacional da linguagem*. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2013.

FRISCH, S. A.; LARGE, N. R.; PISONI, D. B. Perception of Wordlikeness: Effects of Segment Probability and Length on the Processing of Nonwords. *Journal of Memory and Language*, 42, p. 481-549, 2000.

GATHERCOLE, S. Is Nonword Repetition a test of Phonological Memory or Long-term Knowledge? It All Depends on Nonwords. *Memory & Cognition*, 23 (1), p. 83-94, 1995.

GATHERCOLE, W. C.; EMSLIE, H.; BADDELLY, A. D. The Influence of Number of Syllables and Word-likeness on Children's Repetition Nonword. *Applied Psycholinguistics*, 12, p. 349-367, 1991.

LUCE, P; PISONI, D. Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, n. 19, p. 1-36, 1998.

MENDES, S. do C. *Habilidades do processamento auditivo e conhecimento fonológico em crianças com desenvolvimento típico*. Dissertação (Mestrado em Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2014.

PIERREHUMBERT, J. Probabilistic Phonology. In: BOD, R.; HAY, J.; JANNEDY, S. (Eds.). *Probabilistic Linguistics*. Cambridge, MA: MIT Press, 2003. p. 177-228.

PIERREHUMBERT, J. B. The Dynamic Lexicon. In: COHN, A.; HUFFMAN, M.; FOUGERON, C. (Eds.). *Handbook of Laboratory Phonology*. Oxford: Oxford University Press, 2012. p. 173-183.

SAITO, A.; YOSHIMURA, T.; ITAKURA, T.; LAMBOM, R. Demonstrating a Wordlikeness Effect on Nonword Repetition Performance in a Conduction Aphasic Patient. *Brain and Language*, 85, p. 222-230, 2003.

SENNA, F. D. *Acesso e representação lexical na produção de afásicos sob a ótica da Fonologia de Uso*. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2013.

SILVA, M. B. *Aquisição fonológica em crianças falantes tardios: um estudo de caso*. Dissertação (Mestrado em Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2014.

VITEVITCH, M. S.; LUCE, P. A. Probabilistic Phonotactics and Neighborhood Activation in Spoken Word Recognition. *Journal of Memory and Language*, 40, p. 374-408, 1999.

VITEVITCH, M. S.; LUCE, P. A. Increases in Phonotactic Probability Facilitate Spoken Nonword Repetition. *Journal of Memory and Language*, 52, p. 193-204, 2005.

VITEVITCH, M.; STAMER, M. The Curious Case of Competition in Spanish Speech Production. *Language Cognitive Processes*, 21 (6), p. 760-770, 2006.

## ANEXO 1

Teste de repetição de pseudopalavras (ESTEVEES, 2013, p. 128)

### **Dissílabas Não-Frequentes**

- 1- gɔfʊ
- 2- ʒefə
- 3- feʎʊ
- 4- noʎɪ
- 5- zubi

### **Dissílabas Frequentes**

- 1- tʃibə
- 2- haɖʒɪ
- 3- sebə
- 4- masɪ
- 5- mibə

### **Trissílabas Não-Frequentes**

- 1- zigefə
- 2- mɔʒəpɪ
- 3- gufoʃʊ
- 4- vubohɪ
- 5- gəpɔnə

### **Trissílabas Frequentes**

- 1- kopizə
- 2- tarekʊ
- 3- mesivə
- 4- vabitʊ
- 5- detukə

### **4 sílabas Não-frequentes**

- 1- ketonisu
- 2- givɛpɔfʊ
- 3- vurehupə
- 4- zeʃɔfubɪ
- 5- ʒɔpɛgefə

### **4 sílabas Frequentes**

- 1- helanitu
- 2- kovitunə
- 3- dʒimerɔtə
- 4- mekulivə
- 5- moliratu