

Mascaramento de fala no Frequency-Following Response obscurece respostas espectro-temporais evocadas auditivas

Brainstem correlates of concurrent speech identification in adverse listening conditions

El enmascaramiento del habla en la respuesta de seguimiento de frecuencia oscurece las respuestas espectral-temporales evocadas auditivas

Leonardo Gleygson Angelo Venâncio*
Lílian Ferreira Muniz*

YELLAMSETTY, Anusha; BIDELMAN, Gavin M. Brainstem correlates of concurrent speech identification in adverse listening conditions. *Brain research*, v. 1714, p. 182-192, 2019.

Diversos estudos psicofísicos e neurofisiológicos têm sido realizados com intuito de investigar mecanismos neurais subjacentes à percepção de fala no ruído^{1,2,3,4}, situação de escuta considerada complexa que demanda que o sistema auditivo distinga a mensagem alvo mascarada comumente por pistas acústicas que se sucedem temporalmente. Entretanto, poucos trabalhos objetivam analisar a codificação subcortical auditiva para identificação da fala com mensagem competitiva em estágios pré-atencionais e pós-perceptivos com

estímulos vocálicos duplos que variaram em *pitch* (F0) e nível de ruído (SNR).

O artigo original intitulado “*Brainstem correlates of concurrent speech identification in adverse listening conditions*”, desenvolvido por Anusha Yellamsetty e Gavin M. Bidelman nos Estados Unidos, traz contribuições importantes para a atuação Fonoaudiológica em Audiologia, uma vez que o foco centra-se em mensurar, do ponto de vista neurofisiológico, as respostas evocadas auditivas subcorticais correlacionando-as ao comportamento

*Universidade Federal de Pernambuco, Recife, PE, Brasil

E-mail para correspondência: Leonardo Gleygson Angelo Venâncio e-mail: leonardoserrita@hotmail.com

Recebido: 26/06/2019

Aprovado: 02/12/2019

auditivo humano, além de demonstrar a aplicabilidade do exame Frequency-Following Response (FFR) para investigar a codificação neural do tronco encefálico em condições de escuta adversas.

Para esse fim, o estudo original avaliou a relação entre a atividade neural evocada auditiva do tronco encefálico de forma passiva ou pré-atencional e a identificação comportamental de vogais, ao testar a hipótese de que a identificação de vogais duplas, concorrentes, podem ser registradas objetivamente antes de se projetar corticalmente, ou seja, subcorticalmente, com alta fidelidade espectro-temporal por meio do FFR com estímulos vocálicos que variaram em seu *pitch* de voz, isto é, separação de F0, além de na relação entre o sinal e o ruído (SNR).

Do ponto de vista de caracterização amostral, a população foi homogênea, constituída por dezesseis adultos jovens, de ambos os sexos, com idade média de 24 anos, com limiares audiométricos tonais normais, nível de instrução formal semelhante e sem distúrbios neuropsiquiátricos. Esses foram submetidos, de maneira geral, às tarefas auditivas comportamentais e às eletrofisiológicas, respectivamente.

Para tarefa comportamental de identificação de vogal dupla, foram combinados *tokens* vocálicos sintéticos em estado estacionário (/ a /, / ε / e / u /), isto é, sem variações no tempo, de duração de 200 ms com onset/offset de 10 ms e F0 de 150 ou 190 Hz e frequências de formantes (F1, F2) que variavam entre 766 Hz, 1299 Hz; 542 Hz, 1780 Hz e 329 Hz, 810 Hz. Tais estímulos foram desenvolvidos por meio do sintetizador Klatt e implementado no MATLAB® 2014 (The MathWorks, Inc.) com sobreposição de vogais simples individuais e pares de vogais com F0s idênticas (0ST) ou diferentes (4ST).

O processamento das vogais duplas foi avaliado em dois blocos separados: na condição em silêncio e com ruído contínuo de balbucio multilocal (+5 dB SNR) dentro do intervalo inter-estímulo aleatório de 800 e 1000 ms com ordem de pares vocálicos randomizada entre os participantes. Para avaliar como as pistas espectrotemporais são afetadas pelo ruído, os participantes ouviram 50 repetições de cada combinação de vogais e foram solicitados a identificar as vogais pressionando duas teclas em um teclado, durante cada bloco. A resposta comportamental seguiu julgamento sequencial sem fornecimento de feedback, embora

tenha sido mencionado que cada tentativa conteria duas vogais únicas. Chama-se atenção para o fato de que antes do experimento, foi realizado um treinamento com identificação das vogais isoladas para familiarização com os estímulos.

A análise dos dados comportamentais considerou a precisão de identificação das vogais por meio da porcentagem de acertos e o “benefício F0” que diz respeito à melhoria na precisão de identificação com aumento da F0 entre as vogais que foi calculado pelo desempenho sem ruído e nos diferentes níveis de ruído entre os diferentes semitons, 0ST e 4ST.

Para identificar os padrões de respostas através do FFR de vogais isoladas e duplas quando os sons da fala variaram em pitch (F0), e determinar o efeito de diferentes níveis do ruído mascarador sobre codificação neural, os participantes foram submetidos ao paradigma de escuta passiva das vogais / a / e / ε / em cabina eletroacusticamente blindada em condição confortável, enquanto assistiam a um filme legendado. Os estímulos foram apresentados binauralmente, conduzidos via fones de ouvido do tipo de inserção com introdução de ruído contínuo de balbucio por abordagem eletroencefalográfica seguindo condições da ordem de cada tipo de vogal, nível de ruído e *pitch*.

A análise do FFR foi realizada com o uso de transformadas rápidas de Fourier de representações das amplitudes de frequência F0, harmônicas e formantes para quantificar a codificação do “*pitch*” em cada condição nas faixas de frequências, bem como tempo de reação ou velocidade de resposta para cada tipo de vogal e nível de ruído. Em linhas gerais, a análise estatística utilizada foi robusta em ambas as abordagens, baseando-se em medidas controladas por testes ANOVA mista, comparações múltiplas de Tukey-Kramer e regressão com nível de significância estabelecido em $\alpha = 0,05$.

Os achados comportamentais mostraram melhor desempenho na identificação de vogais individuais comparados às duplas, piores respostas para vogais duplas quando apresentaram mesma F0, indicando que as diferenças de semitons (F0) entre as vogais ajudam os ouvintes a identificar a fala, porém, esse benefício perceptual da F0 é degradado com o ruído.

De maneira complementar, evidenciou-se que amplitudes F0 das vogais individuais foram mais robustas que em vogais duplas, mais fortes para as vogais duplas sem sinais de *pitch* e reduções estive-

ram relacionadas ao ruído, sugerindo não linearidade da resposta. Em outras palavras, o efeito do ruído foi maior quando cada par de vogais teve F0s idêntica de modo com que os efeitos de mascaramento da fala no FFR para estímulos de fala concorrentes foram sistematicamente diferentes de contrapartes da vogal individual, que variaram em função do componente de frequência e presença de ruído.

As respostas eletrofisiológicas eliciadas pelo FFR estiveram associadas ao desempenho comportamental, uma vez que análises de regressão lineares entre as amplitudes da F0 e a precisão comportamental revelaram influência da velocidade de resposta na condição de ruído. Ao fazer comparações entre codificação subcortical e identificação comportamental de vogais concorrentes para verificar a análise de dominância de vogal, refutou-se a hipótese que a força relativa de cada vogal individual em relação à resposta de vogal dupla determina qual vogal é perceptivamente dominante, corroborando com o pressuposto sugerido pelos autores de que a codificação subcortical prevê a velocidade perceptual, mas não a precisão na identificação perceptiva comportamental de vogais duplas na condição de ruído.

Os autores concluem, portanto, que as respostas evocadas auditivas de tronco encefálicas eliciadas por meio do FFR refletem a codificação atencional de estímulos concorrentes podendo prever fracamente o sucesso da identificação de fala simultânea, apesar de fornecer informações importantes acerca do processamento neural de sinais complexos, como a fala diante do mascaramento auditivo.

Pode-se questionar, na leitura do artigo, alguns aspectos importantes, como o fato de ser um estudo com delineamento amostral restrito à conveniência e caracterizado sem descrição detalhada do processo de recrutamento, além do rigor de outros fatores aditivos que podem interferir substancialmente nas respostas do FFR, como nível experiência auditiva, fatores cognitivos, individuais e dentre outros, como relata a literatura^{5,6,7}.

Além disso, por se tratar de um estudo experimental com abordagem diferencial com utilização de misturas de vogais duplas complexas e protocolo de duração de aproximadamente 2,5 horas, esperava-se a inclusão da realização de um teste piloto que não foi mencionado, para prever respostas normativas, aumentar a validade do procedimento empregado, possibilitar a comparabilidade com

medidas estabelecidas na literatura que se utiliza de estímulos de fala isolados e melhor distribuição do tempo de realização do exame prezando o conforto do participante. Essas questões podem ter colaborado com o número limitado de associações entre respostas eletrofisiológicas e comportamentais, associados ainda ao fato de que o FFR nem sempre varia com medidas perceptuais^{1,5}, como mencionado pelos autores.

Destaca-se, contudo, a relevante contribuição que os resultados da pesquisa trazem no sentido da exploração de mudanças no FFR relacionadas ao mascaramento da fala em situações adversas, do ponto de vista comportamental e eletrofisiológico, que revela que o mascaramento de fala ocasionado pelo ruído no *Frequency-Following Response* modifica e obscurece as respostas espectrotemporais evocadas auditivas, aspectos cruciais para o seu entendimento, pois envolvem o domínio de tempo e frequência^{8,9}.

O FFR, dessa forma, pode ser considerado como indicador complementar de comprometimento na precisão temporal auditiva e na tarefa de explorar sinais degradados acusticamente⁴, haja vista que os ouvintes do estudo tenderam a ser mais lentos na identificação dos estímulos vocálicos duplos diante do ruído. Assim, apesar de terem poucos estudos com usuários de implante coclear, e outras populações, seria interessante ampliar essa abordagem, pois pode auxiliar na tomada de decisões para ajustes do aparelho em ambientes de competição sonora e no monitoramento da evolução terapêutica a fim de obterem sucesso na adaptação do IC, por exemplo.

Igualmente, registros simultâneos com o FFR e potenciais cortical em diferentes frequências, como sugerido pelos autores, poderiam explorar a correlação neural existente entre os centros de integração da informação sonora em diferentes contextos e populações. Em suma, a avaliação da codificação neural auditiva evocada pelo FFR é uma medida objetiva promissora que pode ser utilizada para verificação da integridade do sistema auditivo para identificação da fala em situações de escuta desafiadoras. A diversidade de protocolos de realização e paradigmas utilizados conforme amostra de interesse devem ser testados com cautela para diminuir vieses internos e externos dos estudos. Cabe aos pesquisadores maior aprofundamento e apropriação do tema para que possam ser evidenciados o real impacto do mascaramento auditivo

temporal no processamento auditivo central e saúde da comunicação humana nas habilidades sociais do sujeito.

Referências bibliográficas

1. Smith SB, Krizman J, Liu C, White-Schwoch T, Nicol T, Kraus N. Investigating peripheral sources of speech-in-noise variability in listeners with normal audiograms. *Hear Res.* 2019; 371(2019): 66–74.
2. Anderson S, Parbery-Clark A, White-Schwoch T, Kraus N. Auditory Brainstem Response to Complex Sounds Predicts Self-Reported Speech-in-Noise Performance. *J Speech Lang Hear Res.* 2013; 56(1): 31–43.
3. Morris DJ, Tondering J, Lindgren M. Electrophysiological and behavioral measures of some speech contrasts in varied attention and noise. *Hear Res.* 2019; 373(2019): 1–9.
4. Coffey EBJ, Chepesiuk AMP, Herholz SC, Baillet S, Zatorre RJ. Neural Correlates of Early Sound Encoding and their Relationship to Speech-in-Noise Perception. *Front Neurosci.* 2017; 11(479): 1-14.
5. Kraus N, Anderson S, White-Schwoch T, Fay RR, Popper AN. *The Frequency-Following Response: A Window into Human Communication.* 1.ed. United States: Springer International Publishing; 2016.
6. Coffey EBJ, Herholz SC, Chepesiuk AMP, Baillet S, Zatorre RJ. Cortical contributions to the auditory frequency-following response revealed by MEG. *Nat Commun [Internet]. Nature Publishing Group;* 2016 Apr 24 [cited 2019 May 18];7(1):11070. Available from: <http://www.nature.com/articles/ncomms11070>
7. Yeend I, Beach EF, Sharma M, Dillon H. The effects of noise exposure and musical training on suprathreshold auditory processing and speech perception in noise. *Hear Res.* 2017; 353(2017): 224–36.
8. Samelli AG, Schochat E. Processamento auditivo, resolução temporal e teste de detecção de gap: revisão da literatura. *Rev CEFAC.* 2008; 10(3): 369–77.
9. Anderson S, Kraus N. The Potential Role of the cABR in Assessment and Management of Hearing Impairment. *Int J Otolaryngol.* 2013; 2013(604729): 1-10.