

Caroline Nunes Rocha-Muniz^{1,2} 
Eliane Schochat¹ 

Investigação da discriminação neural das características acústicas dos sons de fala em normo-ouvintes por meio do Frequency Following Response (FFR)

Investigation of the neural discrimination of acoustic characteristics of speech sounds in normal-hearing individuals through Frequency-following Response (FFR)

Descritores

Audiologia
Eletrofisiologia
Vias Auditivas
Percepção Auditiva
Percepção de Fala

Keywords

Audiology
Electrophysiology
Auditory Pathways
Auditory Perception
Speech Perception

Endereço para correspondência:

Caroline Nunes Rocha-Muniz
Universidade de São Paulo – USP
Rua Cipotânea, 51, Cidade
Universitária, São Paulo (SP), Brasil,
CEP: 05360-160.
E-mail: carolrocha@usp.br

Recebido em: Maio 10, 2019

Aceito em: Março 12, 2020

RESUMO

Objetivo: Avaliar como as vias auditivas codificam e diferenciam as sílabas plosivas [ga],[da] e [ba], por meio do potencial evocado auditivo *Frequency Following Response* (FFR), nas crianças em desenvolvimento típico. **Método:** Vinte crianças (6-12 anos) foram avaliadas por meio do FFR para estímulos [ga],[da] e [ba]. Os estímulos foram compostos por seis formantes, sendo diferenciados na transição F₂ e F₃ (porção transiente). Os demais formantes foram idênticos nas três sílabas (porção sustentada). Foram analisadas latências de 16 ondas que compõe a porção transiente do estímulo (<70ms) e latências de 21 ondas da porção sustentada (90-160ms) nas respostas neurais obtidas para cada uma das sílabas. **Resultados:** As respostas eletrofisiológicas registradas por meio do FFR demonstraram que as latências da porção transiente da resposta neural foram diferentes nas três sílabas evocadas. Além disso, os valores de latência das ondas da porção transiente foram aumentando progressivamente, sendo [ga]<[da]<[ba]. Já na porção sustentada da resposta, não houve diferenças significantes nas latências das ondas que compõe essa porção. **Conclusão:** O FFR mostrou-se uma ferramenta eficiente na investigação da discriminação subcortical de diferenças acústicas dos sons de fala, uma vez que demonstrou diferentes resposta eletrofisiológica para três sílabas evocadas. Na porção transiente (consoantes) foram observadas mudanças de latência e na porção sustentada (vogal) não houve diferenças entre as latências para os três estímulos. Esses resultados demonstram a capacidade neural de distinção entre características acústicas dos estímulos [ga],[da],[ba].

ABSTRACT

Purpose: To evaluate how the auditory pathways encode and discriminate the plosive syllables [ga], [da] and [ba] using the auditory evoked Frequency-following Response (FFR) in children with typical development. **Methods:** Twenty children aged 6-12 years were evaluated using the FFR for the [ga], [da] and [ba] stimuli. The stimuli were composed of six formants and were differentiated in the F2 to F3 transition (transient portion). The other formants were identical in the three syllables (sustained portion). The latencies of the 16 waves of the transient portion (<70ms) and of the 21 waves of the sustained portion (90-160ms) of the stimuli were analyzed in the neural responses obtained for each of the syllables. **Results:** The transient portion latencies were different in the three syllables, indicating a distinction in the acoustic characteristics of these syllables through their neural representations. In addition, the transient portion latencies progressively increased in the following order: [ga] < [da] < [ba], whereas no significant differences were observed in the sustained portion. **Conclusion:** The FFR proved to be an efficient tool to investigate the subcortical acoustic differences in speech sounds, since it demonstrated different electrophysiological responses for the three evoked syllables. Changes in latency were observed in the transient portion (consonants) but not in the sustained portion (vowels) for the three stimuli. These results indicate the neural ability to distinguish between acoustic characteristics of the [ga], [da] and [ba] stimuli.

Trabalho realizado no Departamento de Fisioterapia, Fonoaudiologia e Terapia Ocupacional, Faculdade de Medicina, Universidade de São Paulo – USP - São Paulo (SP), Brasil.

¹ Departamento de Fisioterapia, Fonoaudiologia e Terapia Ocupacional, Faculdade de Medicina, Universidade de São Paulo – USP - São Paulo (SP), Brasil.

² Faculdade de Ciências Médicas da Santa Casa de São Paulo - São Paulo (SP), Brasil.

Fonte de financiamento: Fapesp – 2011/23131-8.

Conflito de interesses: nada a declarar.



Este é um artigo publicado em acesso aberto (Open Access) sob a licença Creative Commons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

INTRODUÇÃO

Na fala, a percepção auditiva de vogais pode ser determinada por um pequeno número de frequências dos primeiros formantes, os quais refletem as propriedades de ressonância do trato vocal⁽¹⁾. Consoantes plosivas são produzidas por uma obstrução temporária do fluxo de ar através do trato vocal com três fases distintas: obstrução total da cavidade oral, acúmulo de pressão enquanto a cavidade oral permanece obstruída e finalmente, a liberação repentina da corrente de ar, que ocasiona um ruído, também chamado de *burst*. O registro acústico correspondente à liberação do fluxo de ar refere-se à fonte de ruído transiente^(2,3).

Uma consoante plosiva é rica em pistas acústicas que subjaz a identificação do ponto de articulação e vozeamento, tais como: transição de formantes, espectro do *burst*, presença ou ausência de aspiração e a duração do *Voice Onset Time* (VOT) que corresponde ao tempo de início ou ataque de vozeamento⁽⁴⁾.

Estudos com modelos animais tem mostrado que a percepção destas informações acústicas é codificada através de muitos níveis do sistema auditivo e com eventos neurais distintos. Tanto as estruturas periféricas quanto centrais, como o nervo auditivo e fibras dos núcleos cocleares, mostram a capacidade de sincronia de fase (atividade de *phase-locking*) para os harmônicos (frequências múltiplas inteiras da frequência fundamental) de um estímulo de fala^(5,6). Além disso, essas estruturas, bem como a parte rostral do colículo inferior, também mostram um aumento na atividade (taxa de disparos) para o VOT⁽⁷⁾.

Em humanos, a sincronia neural em resposta a características acústicas da fala tem sido medida utilizando o *Frequency Following Response* (FFR, um potencial evocado auditivo que é também conhecido como potencial evocado auditivo de tronco encefálico com estímulos complexos ou de fala, PEATEc ou PEATEf). Essa terminologia tem sido mudada desde meados de 2015⁽⁸⁾, a fim de não limitar os conceitos envolvidos por este potencial, entre eles, a natureza integrada (*top-down* e *bottom-up*) e relacionada a experiências enriquecedoras e estimulações⁽⁹⁾.

O FFR reflete uma resposta neural composta por vários tipos diferentes de células, principalmente de células neurais da porção rostral do tronco encefálico.

O tronco encefálico responde com um alto grau de sincronia neural e é excepcionalmente bem sintonizado para as características espectrais e temporais do som, incluindo sons de fala. Contudo, os mecanismos envolvidos na precisão da codificação de muitas pistas acústicas na fala ainda permanecem especulativos.

Um grande número de pesquisas tem estudado como as respostas auditivas em tronco encefálico respondem para o som de fala [da]⁽¹⁰⁾. Para essa pesquisa, uma estrutura tem sido proposta que sugere que distintos mecanismos neurais são responsáveis por codificar diferentes aspectos acústicos dos sons de fala⁽¹¹⁾. Sons de fala consistem de três componentes fundamentais: pitch (frequência fundamental); formantes (harmônicos) e aspectos temporais. Todos esses aspectos são importantes para a percepção de fala e embora estejam simultaneamente presentes no sinal de fala e em suas respostas, componentes específicos do tronco encefálico respondem separadamente para cada um desses componentes⁽⁴⁾.

Em um sistema auditivo maduro, a região basal da cóclea é mais responsiva para frequências altas, enquanto a região

mais apical da cóclea responde para frequências baixas. Essa organização tonotópica é preservada no decorrer de toda a via auditiva até o córtex e acredita-se que essa organização auxilie a preservar a relação espectral no padrão de atividade neural^(12,13).

Estudos têm demonstrado que a percepção das diferenças entre fonemas por meio dos potenciais evocados auditivos corticais ([da], [ga] [ba], por exemplo) está relacionada as frequências contidas nos formantes dos estímulos utilizados⁽¹⁴⁻¹⁸⁾.

A transição de formantes é uma das pistas essenciais que subjazem a identificação de consoantes plosivas⁽¹⁹⁾. Assim, uma forma interessante de estudar como a codificação dessa transição ocorre no sistema auditivo seria estudar estímulos que se diferenciem apenas nas características de “filtro”(ou harmônicos), como no caso das sílabas [da], [ga] e [ba]. Uma das diferenças primárias entre essas sílabas é a transição de frequências entre o segundo e terceiro formante (F_2 e F_3).

Uma vez que F_2 e F_3 estão além da capacidade de *phase-locking* do tronco encefálico, a investigação a cerca das diferenças dessas pistas espectrais são observadas pelas latências das respostas neurais.

Baseado na organização tonotópica do sistema auditivo, sons de baixa frequência, localizadas na parte apical da cóclea, geram respostas com milésimos de segundos a mais em comparação às respostas geradas por sons de alta frequência, codificadas na parte basal da cóclea. Ou seja, o tempo de resposta de estímulos com frequências altas poderiam ter resposta de latência menores que respostas à estímulos com frequências mais baixas. Essa progressão de latência em função da frequência tem sido demonstrada nas respostas de tronco encefálico para tons puros⁽²⁰⁾.

Desta forma, a investigação da codificação neural para os traços distintivos das sílabas [da], [ga] e [ba] (os quais ocorrem na transição do formante F_2), por meio do FFR, pode nos auxiliar na investigação da codificação neural para os formantes e portanto, entender os processos que subjazem a distinção neural de contrastes acústicos de diferentes estímulos de fala, como consoantes plosivas.

A fim de expandir nosso conhecimento sobre a discriminação neural de diferentes características acústicas, o objetivo deste estudo consiste em avaliar como as vias auditivas codificam e diferenciam as sílabas plosivas consoante-vogal [ga], [da] e [ba], apresentadas sob a forma de estímulos de fala, por meio do FFR, em crianças em desenvolvimento típico (DT).

Nós hipotetizamos que:

- (1) Por causa da tonotopia do sistema auditivo - a qual promove a codificação mais rápida de frequências altas - as diferenças entre as frequências na transição dos formantes F_2 e F_3 dos estímulos apresentados devem se manifestar com a mudança de latência. Assim, esta mudança deve ocorrer por meio do aumento progressivo na latência das respostas para os estímulos [ga], [da] e [ba] (ou seja, [ga]<[da]<[ba]) devido a sincronia neural;
- (2) Diferença entre a latência deve diminuir ao longo do curso da resposta até desaparecer no momento que as três sílabas chegarem ao seu estado estável;
- (3) Não deve haver diferenças entre as latências das respostas eletrofisiológicas para os três estímulos na a porção sustentada.

MÉTODO

O presente estudo foi aprovado pela Comissão de Ética da Faculdade de Medicina da Universidade de São Paulo, sob Protocolo nº109/12. Os pais ou responsáveis pelas crianças receberam orientação sobre todos os procedimentos da pesquisa e assinaram do termo de consentimento livre e esclarecido.

Casuística

Participaram da pesquisa 20 crianças com desenvolvimento típico (de acordo com informações obtidas através de entrevista com os responsáveis pela criança e professores), ausência de transtornos neurológicos, cognitivos e psiquiátricos, ausência de queixas escolares, problemas de fala e linguagem.

Todos os indivíduos avaliados apresentaram limiares dentro do limite de normalidade (≤ 15 dB NA) para as frequências avaliadas (500 Hz a 4000 Hz), reconhecimento de fala com escore $> 90\%$, medidas timpanométricas normais, PEATE com estímulo clique dentro dos padrões de normalidade. Além disso, essas crianças tiveram desempenho normal para a avaliação do processamento auditivo. Os testes utilizados para descartar alteração no processamento auditivo seguiu os critérios preconizada pela AAA⁽²¹⁾ e ASHA⁽²²⁾, com a utilização de um teste monótico, um teste dicótico e dois testes temporais. Caso fossem encontradas alterações referentes aos aspectos auditivos, neurológicos, cognitivos ou psiquiátricos, os indivíduos seriam excluídos do estudo e encaminhado para serviço especializado. Todos os responsáveis pelas crianças assinaram o termo de consentimento livre e esclarecido.

Estímulos e parâmetros de captação de resposta

O FFR foi obtido por meio da apresentação de estímulos acústicos de fala - consoantes plosivas [da], [ga] e [ba]. Os estímulos de fala foram sintetizados⁽²³⁾ com 20 KHz, 16-bit, e duração de 170 ms. Os estímulos foram compostos por seis formantes, sendo diferenciados nas frequências “onset” (porção inicial do estímulo), na transição do segundo para o terceiro formante (Tabela 1). Esses estímulos foram os mesmos utilizados por Johnson et al.⁽⁴⁾.

Procedimentos

Os estímulos foram apresentados pelo eletroneuromiógrafo, modelo SmartEP equipado com o módulo cABR (Intelligent Hearing Systems, Miami, FL, USA). Os estímulos foram

apresentados à uma velocidade de 4,35 estímulos por segundo e intensidade de 80dBnNA, apenas na orelha direita.

As respostas eletrofisiológicas geradas pelos estímulos [da], [ga] e [ba] foram processadas com um filtro de 50-3000Hz (filtro off-line de 70-2000Hz). O critério de rejeição de artefatos foi de $\pm 35 \mu V$.

O FFR foi captado através dos eletrodos de superfície nas posições Cz, M2 (mastóide direita) e o Fpz como terra com uma janela de análise de 230 ms (45ms correspondendo ao período pré-estímulo e 185 ms correspondendo ao período pós-estímulo).

Foram realizadas duas varreduras de 2000 estímulos para cada sílaba apresentada com polaridade alternada. As duas ondas geradas pelas varreduras foram somadas de forma ponderada e a onda resultante final, com 4000 estímulos, foi analisada.

Análise das respostas

Período de transição dos formantes

O período de transição dos formantes foi definido como a porção da resposta correspondente ao “onset” e o período de transição dos formantes no estímulo (0-50ms). Baseado em nossa primeira hipótese, são esperadas diferenças entre os estímulos para as latências nesta porção da resposta.

Nesta porção, foram marcados um total de 16 picos, sendo seis picos positivos e 10 picos negativos (Figura 1) nos 70 ms iniciais da resposta eletrofisiológica.

Período da resposta sustentada

A resposta sustentada foi definida como a porção correspondente a parte estável do estímulo (51-170 ms).

Nessa porção, a resposta eletrofisiológica correspondente encontra-se entre 90ms à 160ms e foram visualizados 21 picos, sendo sete picos positivos e 14 picos negativos.

Análise dos dados

Análise qualitativa

Para essa análise, utilizamos a técnica “Cross-phaseogram”⁽²⁴⁾. Essa técnica calcula as diferenças de fase de onda entre as duas respostas eletrofisiológicas como uma função tempo e frequência e permite ilustrar as diferenças na porção transiente nas comparações [ga] vs. [ba], [da] vs. [ba], [ga] vs. [da].

Tabela 1. Valores em Hz da frequência fundamental e os seis formantes de cada estímulo

	Frequências dos Formantes dos Estímulos										
	F ₀		F ₁		F ₂		F ₃		F ₄	F ₅	F ₆
	Plano	Onset	Estado Estável	Onset	Estado Estável	Onset	Estado Estável	Plano	Plano	Plano	
[ga]	100	400	720	3000	1240	3100	2500	3300	3750	4900	
[da]	100	400	720	1700	1240	2580	2500	3300	3750	4900	
[ba]	100	400	720	900	1240	2400	2500	3300	3750	4900	

Adaptado de Johnson et al.⁽⁴⁾

Quando a resposta para o estímulo [ga] está à frente da fase relativa para a resposta [ba], a representação gráfica consiste em tonalidades de amarelo, laranja e vermelho, e uma maior diferença representada pela cor vermelho escuro. Se ocorre o oposto, a representação é feita pela tonalidade azul e quando não há diferença entre as fases, pela tonalidade verde.

Análise quantitativa

Com o objetivo de normalizar as latências de todos os 16 picos (porção transiente) e 21 picos (porção sustentada) para que pudessem ser descritos na mesma escala, uma grande média geral das latências (MG) para os picos obtidos nos três estímulos foi computada. Em seguida, essa grande média foi subtraída para cada pico de latência individual ($\text{Latência}_{\text{Individual}} - \text{Latência}_{\text{MG}}$). Assim, os picos com latências mais precoces são números negativos, picos com latências mais tardias são números positivos e picos com latências próximas a grande média geral são próximos de zero⁽⁴⁾.

Para comparar as médias dos testes nos três estímulos estudados foram aplicadas as técnicas de análise multivariada de variância com medidas repetitivas (MANOVA repeated measure), respectivamente⁽²⁵⁾. Na MANOVA, o p-valor e a razão F (que é utilizada para testar a diferença global entre grupos) foram analisadas por meio do teste Lambda de Wilks (λ de Wilks).

Para complementar a análise descritiva, foi utilizado o intervalo de confiança (IC) para avaliar a extensão para qual a média poderia variar em um certo nível de confiança. O IC estabelecido para nossa análise de dados foi 95% com nível de significância de 0,05 (5%).

RESULTADOS

Os resultados dos picos de latência resultantes da estimulação das sílabas [da], [ga] e [ba] foram analisadas de acordo com os períodos: transiente (16 picos) e sustentado (21 picos).

Foi hipotetizado que durante o período de transição do formante (porção transiente), a diferença nas frequências dos formantes F_2 e F_3 seriam manifestas nas latências de suas respostas eletrofisiológicas, com as respostas para os estímulos $[\text{ga}] < [\text{da}] < [\text{ba}]$ elevando progressivamente os valores de latência devido a progressão entre as diferenças de frequências dos sons. Já para a porção sustentada do estímulo (90-170 ms), foi hipotetizado que não haveria diferenças significantes nas respostas eletrofisiológicas entre os três estímulos utilizados.

Na análise qualitativa realizada por meio da técnica do *cross-phaseogram*⁽²⁴⁾, verificamos que a maior discriminação ocorreu entre [ga] e [ba] e depois para [da] e [ba] (Figura 2). Essa discriminação é representada na Figura 2 por meio das tonalidades amarelo, laranja e vermelho. Uma menor discriminação ocorreu para os pares com menor diferença, ou seja, [ga] e [da], representada na Figura 2 por meio da tonalidade - mais predominante - verde. Nota-se que as diferenças ocorreram apenas na porção transiente, na qual ocorre a diferença entre F_2 e F_3 para as três sílabas (10-50 ms). Já na porção sustentada (semelhante nas três sílabas), verificamos que não houve diferença entre as fases das respostas, representada na Figura 2 por meio da tonalidade verde.

Uma vez que não identificamos diferenças entre os estímulos na porção sustentada, a análise estatística foi realizada apenas na porção transiente.

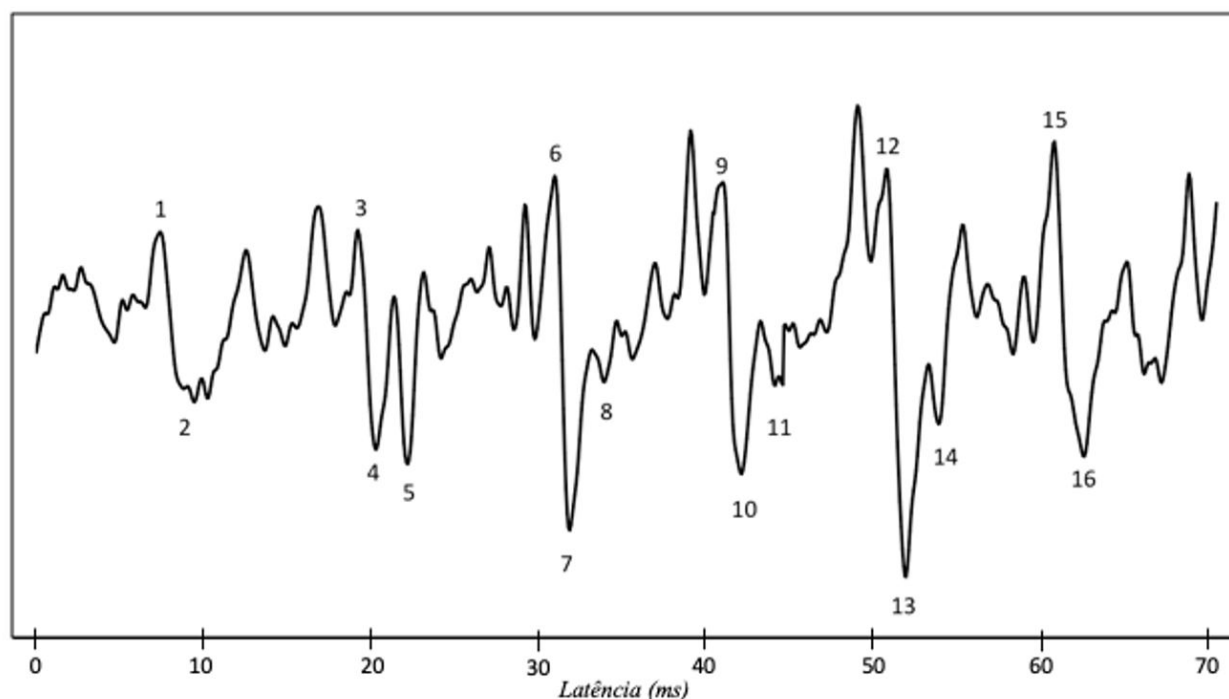


Figura 1. Representação da marcação das 16 ondas na porção transiente da resposta eletrofisiológica gerada pelo estímulo /da/

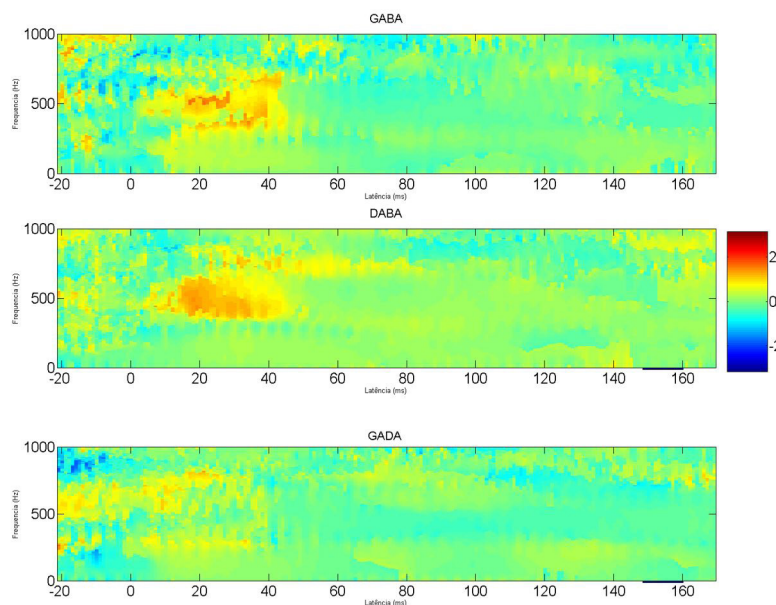


Figura 2. Diferença nas respostas médias para as sílabas [ga] e [ba]; [da] e [ba]; [ga] e [da] nas 20 crianças avaliadas por meio da técnica de análise “Cross-phaseogram”

Tabela 2. Valores descritivos das latências absolutas de cada pico da resposta do FFR para todos os estímulos

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
		<i>Latência dos picos (ms)</i>															
[ga]	Média	8,10	9,00	21,27	22,48	24,26	31,53	32,70	34,50	41,79	43,11	44,58	51,75	53,19	54,82	62,10	63,39
	DP	0,43	0,78	0,49	0,66	0,79	0,45	0,56	0,69	0,50	0,71	0,84	0,66	1,21	0,90	0,82	0,97
	N	20	19	20	20	18	20	20	14	20	20	15	20	20	15	20	20
[da]	Média	8,03	9,25	21,78	22,81	24,57	31,78	32,96	34,90	41,72	42,91	44,78	51,83	53,12	55,02	61,94	63,23
	DP	0,38	0,59	0,82	0,82	1,19	0,51	0,59	0,87	0,51	0,55	0,93	0,84	0,97	0,91	0,97	1,10
	N	20	20	20	20	17	20	20	16	20	20	15	20	20	18	20	20
[ba]	Média	8,24	9,57	22,04	23,19	25,07	32,21	33,43	35,23	42,00	43,31	45,24	52,00	53,24	55,28	62,01	63,39
	DP	0,50	0,49	0,99	0,86	0,91	0,47	0,47	0,85	0,48	0,56	0,81	1,00	1,06	1,10	0,98	1,09
	N	20	20	18	19	15	20	20	13	20	20	19	20	20	17	20	20

DP = Desvio Padrão; N = número de indivíduos

Foram analisadas as latências das 16 ondas que compõe a parte transiente do estímulo (latência de 0-70 ms) em cada uma das estimulações. A análise descritiva das medidas de latência das 16 ondas de todas as crianças em DT está na Tabela 2. Uma vez que alguns sujeitos não apresentaram todas as 16 ondas, foi inserido a sigla ‘N’ para especificar a quantidade de sujeitos que apresentaram aquela onda e conseqüentemente, a quantidade de sujeitos utilizadas para as demais análises.

Na Figura 3 podemos verificar a média das respostas eletrofisiológicas (*Grand-Average*) obtidas pelos 20 indivíduos para as três estimulações ([ga], [da] e [ba]).

O resultado da GM subtraída de cada um dos 16 pico da porção transiente ($Latência_{Individual} - Latência_{MG}$) está exposto na Figura 4. Os picos com latências mais precoces são números negativos [ga], picos com latências mais tardias são números positivos [ba] e picos com latências próximas a grande média geral são próximos de zero [da].

A Figura 5 apresente os IC para as 16 ondas da porção transiente.

O modelo multivariado de análise utilizando medidas repetidas (Repeated Measures MANOVA) foi conduzido para

determinar se houve diferenças entre as três sílabas estudadas. O resultado das análises foram divididos em quatro partes: a) Latência dos picos “Onset” (1,2); b) Latência dos picos maiores (3,4,6,7,9,10,12,13,15,16); c) Latência dos picos menores (5,8,11,14) e Latência dos picos end-point (15,16).

Análise das latências dos picos “onset”

A MANOVA com medidas repetidas não mostrou diferença multivariada nas medidas de latências entre as respostas eletrofisiológicas dos estímulos [da],[ga] e [ba] $[F(16,4) = 1,90, p=0,16]$.

Análise das latências dos picos maiores

A MANOVA com medidas repetidas mostrou diferença multivariada nas medidas de latências entre as respostas eletrofisiológicas dos estímulos [da], [ga] e [ba] $[F(16,4) = 92,05, p<0,001, \eta^2_{parcial}=0,99, Wilks \lambda=0,99]$.

Para avaliar as relativas contribuições de cada medida de latência das ondas analisadas na diferença encontrada, análise univariada foi realizada. As análises indicaram que houve

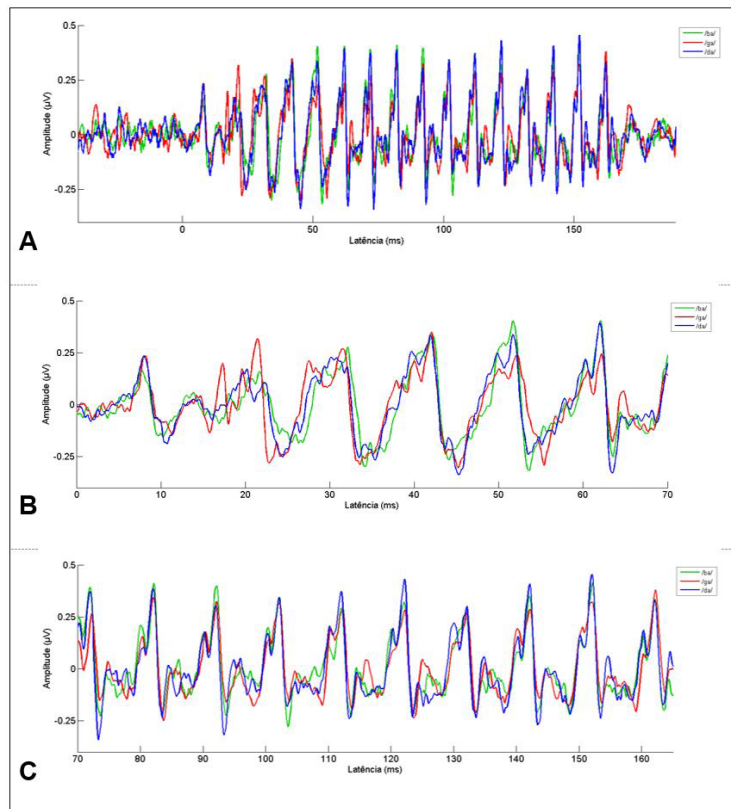


Figura 3. (A) Média das respostas eletrofisiológicas obtidas pela FFR com estímulo /ga/ (verde), /da/ (vermelho) e /ba/ (azul) nos 20 indivíduos; (B) Porção transitória da resposta eletrofisiológica; (C) Porção sustentada da resposta eletrofisiológica

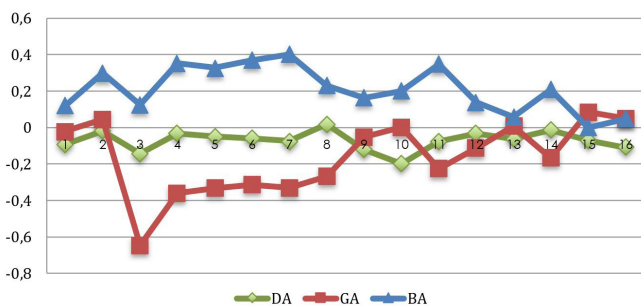


Figura 4. GM subtraída de cada um dos 16 picos da porção transitória ($Latência_{Individual} - Latência_{MG}$) nos três estímulos estudados

diferença estatisticamente significativa entre os três estímulos estudados para as ondas: 3 ($p < 0,001$), 4 ($p < 0,001$), 6 ($p < 0,001$), 7 ($p < 0,001$), 9 ($p < 0,001$), 10 ($p < 0,001$).

Na comparação pareada, foi aplicado o teste-t pareado para verificar as diferenças entre os estímulos (Tabela 3).

Análise das Latências dos picos menores

Na análise dos picos menores não foi possível aplicar a análise MANOVA com medidas repetidas uma vez que os números de ausências encontradas nessas ondas (Tabela 2) prejudicou a aplicação da análise multivariada. Esse achado demonstra maior inconstância dessas ondas quando comparadas aos picos das ondas maiores, ondas “onset” e end-point.

Desta forma, aplicamos apenas o teste-t pareado para verificar as diferenças entre os estímulos (Tabela 4).

Análise das latências dos picos end-point

A MANOVA com medidas repetidas mostrou diferença multivariada nas medidas de latências entre as respostas eletrofisiológicas dos estímulos [da], [ga] e [ba] [$F(16,4) = 3,37$, $p = 0,035$, η^2 parcial = 0.45, Wilks $\lambda = 0,54$]

Na comparação pareada, foi aplicado o teste-t pareado para verificar as diferenças entre os estímulos (Tabela 5).

DISCUSSÃO

Devido a importância do processamento neural da transição de elementos acústicos temporalmente para a integridade do processamento de fala, há um grande interesse em compreender como o sistema auditivo central codifica estas informações em um sistema nervoso auditivo normal para então, compreender o que acontece quando essa codificação é rompida ou ainda está em desenvolvimento.

O objetivo desse estudo foi entender como as vias auditivas localizadas no tronco encefálico refletem diferenças acústicas sutis que existem entre as sílabas plosivas consoante-vogal [ga], [da] e [ba], as quais diferem apenas na transição das frequências F_2 e F_3 .

Os resultados encontrados por esse estudo confirmaram nossa primeira hipótese, ou seja, as diferenças entre as

Tabela 3. Teste-T pareado para cada contraste entre os estímulos nos picos maiores

ondas	p-valor					
	/ga/ vs. /ba/		/da/ vs. /ba/		/ga/ vs. /da/	
	t	p	t	p	t	p
3	-4,54	<0,001*	-1,49	0,154	4,35	<0,001*
4	-4,71	<0,001*	-2,40	0,002*	3,17	0,005*
6	-8,852	<0,001*	-6,981	<0,001*	0,432	0,007*
7	-8,713	<0,001*	-3,708	0,001*	2,358	0,03*
9	-3,203	0,005*	-5,653	<0,001*	-1,172	0,26*
10	-1,478	0,16	-4,693	<0,001*	-1,803	0,09 [#]
12	-1,355	0,19	-2,492	0,02*	0,481	0,64
13	-0,223	0,83	-1,39	0,18	-0,302	0,77

*Estatisticamente significativa; [#]Marginalmente significativa; t = t-valor; p = p-valor

Tabela 4. Teste-T pareado para cada contraste entre os estímulos nos picos menores

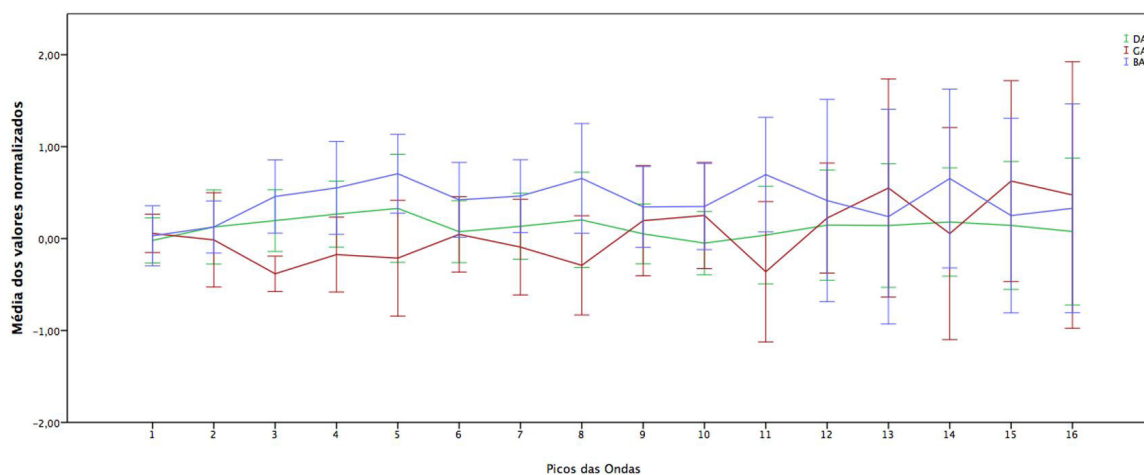
ondas	p-valor					
	/ga/ vs. /ba/		/da/ vs. /ba/		/ga/ vs. /da/	
	t	p	t	p	t	p
5	-4,71	<0,001*	-2,40	0,027*	3,17	0,005*
8	-5,01	0,001*	-1,42	0,19	1,74	0,1
11	-4,45	0,001*	-2,51	0,026*	0,88	0,40
14	-2,73	0,019*	-3,34	0,005*	1,91	0,08

*Estatisticamente significativa; t = t-valor; p = p-valor

Tabela 5. Teste-T pareado para cada contraste entre os estímulos nos picos finais (end-point)

ondas	p-valor					
	/ga/ vs. /ba/		/da/ vs. /ba/		/ga/ vs. /da/	
	t	p	t	p	t	p
15	1,68	0,11	-2,38	0,03*	-2,811	0,01*
16	0,03	0,97	-2,33	0,03*	-2,58	0,02*

*Estatisticamente significativa; t = t-valor; p = p-valor

**Figura 5.** Intervalo de Confiança (IC 95%) dos valores normalizados para cada um dos 16 picos nos três estímulos

frequências F_2 e F_3 foram manifestadas no processamento neural das características acústicas dos estímulos estudados. Em outras palavras, foi demonstrado mudanças no tempo de latência da resposta eletrofisiológica com aumento progressivo na latência das respostas para os estímulos [ga], [da] e [ba] (ou seja, [ga]<[da]<[ba]).

Essa diferença entre período de latência e estímulos foi observada de forma evidente, principalmente para as latências dos picos maiores e menores. Entretanto, notamos que os picos

maiores possuem morfologia mais nítida e estável, estando presente em todos os sujeitos, diferente do observado nos picos menores (Tabela 2).

Segundo a teoria de Johnson et al.⁽⁴⁾, essa distinção nas respostas entre os picos maiores e menores suportam a ideia de que as informações fornecidas pelos picos maiores e menores são codificadas por mecanismos neurais separados. Os picos maiores representariam a frequência fundamental (F_0) e corresponderiam ao pulso glótico no estímulo, transmitindo assim informações

sobre o *pitch*. Já as latências dos picos menores refletem os formantes de transição do estímulo, os quais variam entre as sílabas [ga], [da] e [ba] e são expressos no domínio de tempo na resposta eletrofisiológica, uma vez que a variação dessas frequências está acima da capacidade de phase-locking do sistema auditivo.

Visto que os picos maiores refletem o F_0 do estímulo, seria esperado que esses picos fossem idênticos entre as respostas neurais obtidas entre todas as sílabas utilizadas nesse estudo. Contudo, diferenças entre as latências foram observadas também para os picos maiores.

Uma hipótese para essa diferença observada é que os picos maiores sofrem influência dos padrões observados nos picos menores. Outro fator a ser considerado seria que, em articulações naturais, perturbações de *pitch* causadas por movimentos articulatórios no trato vocal poderiam estar presentes. Neste caso, o padrão sistemático observado nos picos menores, poderia ser evidenciado também nos picos maiores.

Uma vez que as latências dos picos menores são representações neurais dos formantes desses estímulos, a menor diferença encontrada entre as respostas eletrofisiológicas [ga] e [da] sugere uma codificação neural similar entre os estímulos. A diferença entre os formantes dos estímulos acústicos [ga] e [da] é mais discreta em comparação à [ga] e [ba], que são os pares com maior distinção em relação aos seus formantes. Desta forma, a codificação neural por meio das medidas de latências dos 16 picos da porção transiente mostrou que a representação eletrofisiológica [da] e [ga] são mais similares em F_2 e F_3 e [ga] e [ba] possui maior diferença entre F_2 e F_3 . Esses contrastes na representação neural das respostas eletrofisiológicas também foram verificados por meio da análise de *crossphaseogram*, na Figura 2.

Portanto, os resultados apresentados corroboram a hipótese de Johnson et al.⁽⁴⁾, e Hornickel et al.⁽²⁶⁾, demonstrando que a codificação neural para os diferentes elementos acústicos se manifesta de forma distinta, independente e podem ser estudadas por meio do potencial evocado auditivo de tronco encefálico.

A segunda hipótese do estudo também foi confirmada por essa pesquisa. Por meio da Tabela 2 e Figura 3 é possível verificar que a diferença entre as latências médias obtidas entre os três estímulos diminuiu ao longo do curso da resposta, até o desaparecimento no momento que as três sílabas chegam ao seu estado estável (vogal) (Figura 2).

Entretanto, nossos resultados não observaram diferenças estatisticamente significantes entre as latências médias na parte inicial ou “onset” da resposta (ondas 1 e 2). Esses achados corroboram Johnson et al.⁽⁴⁾ os quais relacionam o “onset” da resposta neural com a explosão inicial do estímulo de sílaba plosivo-sonora, semelhante nos três estímulos estudados.

A terceira hipótese de nosso estudo também foi demonstrada por nossos resultados, uma vez que não houve diferenças entre as latências das respostas eletrofisiológicas para os três estímulos na porção sustentada (vogal). Esse resultado já era esperado uma vez que as propriedades acústicas nesta porção são idênticas entre os estímulos estudados.

Desta forma, a diferente representação eletrofisiológica das características acústicas das porções transientes e sustentadas

dos estímulos de fala no tronco encefálico, em crianças normais, mostra que diferentes mecanismos neurais, mediados pela sincronia neural ou phase-locking, codificaram separadamente essas pistas acústicas.

O presente estudo contribui para o entendimento dos mecanismos neurais subcorticais que subjazem a codificação de formantes em transição. Nossos resultados demonstraram que as respostas eletrofisiológicas nos primeiros 70 ms de respostas foram responsáveis pela diferenciação entre pistas espectrais que auxiliam na distinção entre as consoantes, sugerindo que diferentes neurônios possuem respostas específicas para aspectos acústicos diferentes. Ou seja, o estímulo que possui frequências mais altas tem respostas com latências mais precoces em comparação aos estímulos com frequências mais baixas. Essa progressão no tempo de latência em função da frequência já tem sido demonstrada na resposta em tronco encefálico para tons puros⁽²⁰⁾. Em um sistema auditivo maduro, a região basal da cóclea é mais responsiva para altas frequências e a região apical é mais responsiva para baixas frequências. Essa organização tonotópica é preservada ao longo das vias auditivas neurais, o que ajudaria a preservar a informação espectral na atividade de codificação neural^(12,13).

Embora o estudo descrito aqui tenha contribuído com novas informações com relação a representação de pistas acústicas transiente e sustentada nas vias auditivas subcorticais, ainda há muito a ser estudado. Com relação a codificação normal das características acústicas, nós esperamos que futuros estudos acrescentem um repertório mais amplo de sílabas, incluindo consoantes com diferentes pontos articulatórios.

Por fim, acreditamos que o FFR com estímulos de fala (ou outros estímulos complexos) – juntamente com outras medidas e avaliações clínicas - podem informar processos que subjazem a natureza biológica do processamento auditivo e alterações de fala e linguagem, auxiliar estratégias de terapias e promover um index objetivo de evolução terapêutica. Por exemplo, pode ser que algumas populações apresentem déficits na codificação neural para elementos específico de “onset” e/ou end-point, ou específicos para codificação da transição rápida de formantes. Por outro lado, outras populações podem exibir dificuldades somente na codificação de pistas acústicas sustentadas. Além disso, há a possibilidade de existir também algumas crianças que apresentem déficits na codificação neural, tanto das informações transientes quanto sustentadas.

Assim, os resultados demonstrados por esse estudo, nos permite vislumbrar que tais populações poderiam ser mais precisamente identificadas e programas e estratégias terapêuticas mais precisas poderiam ser desenvolvidos para adequar a área de dificuldade específica.

CONCLUSÃO

O FFR demonstrou-se uma ferramenta eficiente na investigação da discriminação subcortical de diferenças acústicas dos sons de fala, uma vez que os dados demonstram que as respostas eletrofisiológicas apresentam diferenças pertinentes a cada uma das três sílabas evocadas. Na porção transiente (consoantes) foram observadas mudanças de latência e na porção sustentada (vogal)

não houve diferenças entre as latências para os três estímulos. Em outras palavras, foi possível demonstrar representações neurais distintas para as diferentes características acústicas das sílabas [ga],[da],[ba].

Considerando o conhecimento já existente sobre a codificação de características acústicas de sons de fala, nossos dados auxiliam na compreensão de como o SNA codifica as diferenças perceptuais importantes da fala por meio do FFR. Além disso, acreditamos na significância deste estudo para ampliar os conhecimentos de como a codificação neural dessas diferenças acústicas ocorrem em populações clínicas.

AGRADECIMENTOS

À Fapesp, pelo apoio financeiro a pesquisa (n processo 11/23131-8).

REFERÊNCIAS

- Hillenbrand J, Gayvert RT. Vowel classification based on fundamental frequency and formant frequencies. *J Speech Hear Res.* 1993;36(4):694-700. <http://dx.doi.org/10.1044/jshr.3604.694>. PMID:8377482.
- Ladefoged P, Maddieson I. *The sounds of the world's languages.* Oxford: Blackwell. 1996.
- Johnson K. *Acoustic and auditory phonetics.* Malden, MA: Blackwell; 2003.
- Johnson KL, Nicol T, Zecker SG, Bradlow AR, Skoe E, Kraus N. Brainstem encoding of voiced consonant-vowel stop syllables. *Clin Neurophysiol.* 2008;119(11):2623-35. <http://dx.doi.org/10.1016/j.clinph.2008.07.277>. PMID:18818121.
- Sachs MB, Young ED. Encoding of steady-state vowels in the auditory nerve: representation in terms of discharge rate. *J Acoust Soc Am.* 1979;66(2):470-9. <http://dx.doi.org/10.1121/1.383098>. PMID:512208.
- Young ED, Sachs MB. Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers. *J Acoust Soc Am.* 1979;66(5):1381-403. <http://dx.doi.org/10.1121/1.383532>. PMID:500976.
- Chen GD, Nuding SC, Narayn SS, Sinex DG. Responses of single neurons in the chinchilla inferior colliculus to consonant-vowel syllables differing in voice-onset time. *Aud Neurosci.* 1996;3:179-98.
- White-Schwoch T, Woodruff Carr K, Thompson EC, Anderson S, Nicol T, Bradlow AR, et al. Auditory processing in noise: a preschool biomarker for literacy. *PLoS Biol.* 2015;13(7):e1002196. <http://dx.doi.org/10.1371/journal.pbio.1002196>. PMID:26172057.
- Kraus N, White-Schwoch T. Unraveling the biology of auditory learning: a cognitive-sensorimotor-reward framework. *Trends Cogn Sci.* 2015;19(11):642-54. <http://dx.doi.org/10.1016/j.tics.2015.08.017>. PMID:26454481.
- Banai K, Hornickel J, Skoe E, Nicol T, Zecker SG, Kraus N. Reading and subcortical auditory function. *Cereb Cortex.* 2009;19(11):2699-707. <http://dx.doi.org/10.1093/cercor/bhp024>. PMID:19293398.
- Johnson KL, Nicol T, Kraus N. Brain stem response to speech: a biological marker of auditory processing. *Ear Hear.* 2005;26(5):424-34. <http://dx.doi.org/10.1097/01.aud.0000179687.71662.6e>. PMID:16230893.
- Langner G. Neural processing and representation of periodicity pitch. *Acta Otolaryngol Suppl.* 1997;532(sup532):68-76. <http://dx.doi.org/10.3109/00016489709126147>. PMID:9442847.
- Merzenich MM, Reid MD. Representation of the cochlea within the inferior colliculus of the cat. *Brain Res.* 1974;77(3):397-415. [http://dx.doi.org/10.1016/0006-8993\(74\)90630-1](http://dx.doi.org/10.1016/0006-8993(74)90630-1). PMID:4854119.
- McGee T, Kraus N, King C, Nicol T, Carrell TD. Acoustic elements of speech like stimuli are reflected in surface recorded responses over the guinea pig temporal lobe. *J Acoust Soc Am.* 1996;99(6):3606-14. <http://dx.doi.org/10.1121/1.414958>. PMID:8655792.
- Sharma A, Dorman M. Cortical Auditory evoked potential correlates of categorical perception of voice-onset time. *J Acoust Soc Am.* 1999;106(2):1078-83. <http://dx.doi.org/10.1121/1.428048>. PMID:10462812.
- Tremblay K, Piskosz M, Souza P. Effects of age and age related hearing loss on the neural representation of speech cues. *Clin Neurophysiol.* 2003;114(7):1332-43. [http://dx.doi.org/10.1016/S1388-2457\(03\)00114-7](http://dx.doi.org/10.1016/S1388-2457(03)00114-7). PMID:12842732.
- Korczak P, Stapells DR. Effects of various articulatory features of speech on cortical event-related potentials and behavioral measures of speech-sound processing. *Ear Hear.* 2010;31(4):491-504. <http://dx.doi.org/10.1097/AUD.0b013e3181d8683d>. PMID:20453651.
- Elangovan S, Stuart A. A cross-linguistic examination of cortical auditory evoked potentials for categorical voicing contrast. *Neurosci Lett.* 2011;490(2):140-4. <http://dx.doi.org/10.1016/j.neulet.2010.12.044>. PMID:21193015.
- Blumstein SE, Isaacs E, Mertus J. The role of the gross spectral shape as a perceptual cue to place articulation in initial stop consonants. *J Acoust Soc Am.* 1982;72(1):43-50. <http://dx.doi.org/10.1121/1.388023>. PMID:7108042.
- Gorga M, Abbas P, Worthington D. Stimulus calibration in ABR measurements. In Jacobsen J, editor. *The auditory brainstem response.* San Diego: College-Hill Press; 1985. p. 49-62.
- AAA: American Academy of Audiology. Diagnosis, treatment, and management of children and adults with central auditory processing disorder [Internet]. Reston: AAA; 2010 [citado em 2019 Maio 10]. Disponível em: <https://www.audiology.org/publications-resources/document-library/central-auditory-processing-disorder>
- ASHA: American Speech and Hearing Association. (Central) auditory processing disorders. Technical report [Internet]. Washington: ASHA; 2005 [citado em 2019 Maio 10]. Disponível em: <https://www.asha.org/policy/TR2005-00043/>
- Klatt DH. Software for a cascade/parallel formant synthesizer. *J Acoust Soc Am.* 1980;67(3):971-95. <http://dx.doi.org/10.1121/1.383940>.
- Skoe E, Nicol T, Kraus N. Cross-phaseogram: objective neural index of speech sound differentiation. *J Neurosci Methods.* 2011;196(2):308-17. <http://dx.doi.org/10.1016/j.jneumeth.2011.01.020>. PMID:21277896.
- Dancey CP, Reidy J. *Estatística sem matemática para psicologia.* Porto Alegre: Artmed; 2006.
- Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N. Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. *Proc Natl Acad Sci USA.* 2009;106(31):13022-7. <http://dx.doi.org/10.1073/pnas.0901123106>. PMID:19617560.

Contribuição dos autores

CNRM: proposição da idéia original, coleta de dados, coleta e análise de dados, revisão da literatura, redação e revisão do manuscrito; ES: proposição da ideia original, coordenação do estudo, revisão do manuscrito.