

Fernanda Leitão de Castro Nunes de Lima¹ 

Larissa Cristina Berti¹ 

Viviane Cristina de Castro Marino¹ 

Eliana Maria Gradim Fabbron¹ 

Julgamento perceptivo-auditivo e perceptivo-visual na identificação de produções gradientes em fricativas

Perceptive-auditive and perceptive-visual judgment in the identification of gradient productions in fricatives

Descritores

Percepção Auditiva

Percepção Visual

Ultrassonografia

Fricativas

Língua Portuguesa

Keywords

Auditory Perception

Visual Perception

Ultrasonography

Fricatives

Portuguese Language

RESUMO

Objetivo: Analisar qual método de julgamento, perceptivo-auditivo (JPA) de áudios ou julgamento perceptivo-visual (JPV) (imagens ultrassonográficas), é mais sensível para detectar produções gradientes entre a classe das fricativas coronais surdas e verificar se há correlação entre essas formas de julgamento. **Método:** Arquivos de áudio e vídeo de ultrassonografia de língua (USL) relativos às produções das palavras “sapo” e “chave”, de 11 crianças, entre 6 a 12 anos de idade, com produção de fala atípica, foram selecionados de um banco de dados e editados para os julgamentos. Após instrução e treino prévio, 20 juízes deveriam escolher, imediatamente à apresentação do estímulo (auditivo ou visual), uma dentre três opções dispostas na tela do computador. No JPA as opções eram: produção correta, incorreta ou gradiente, enquanto no JPV as opções eram imagens correspondentes à produção de [s], de [ʃ] ou indiferenciada. O tempo de apresentação dos estímulos e o tempo de reação foram controlados automaticamente pelo *software* PERCEVAL. **Resultados:** O JPV propiciou uma maior porcentagem de identificação dos estímulos gradientes e um menor tempo de reação na realização da tarefa comparativamente ao JPA, ambos estatisticamente significante. O teste de correlação de Spearman não mostrou significância estatística entre as respostas do JPA e JPV, nem para o tempo de reação. **Conclusão:** O JPV com o uso de imagens US mostrou-se o método mais sensível para a detecção da produção gradiente na produção de fricativas [s] e [ʃ], podendo ser utilizado como método complementar ao JPA na análise de fala.

ABSTRACT

Purpose: To analyze which method of judgment, auditory- perceptual (PAJ) of audios or perceptual-visual judgment (PVJ) (ultrasound images), is more sensitive to detect gradual productions between the class of deaf coronal fricatives and check if there is a correlation between these forms of judgment. **Method:** Audio and video files of language ultrasound (LUS) related to the production of the words “frog” and “key”, of 11 children, between 6 and 12 years old, with atypical speech production, were selected from a bank data and edited for judgments. After instruction and prior training, 20 judges should choose, immediately upon presentation of the stimulus (auditory or visual), one of three options arranged on the computer screen. In PAJ the options were: correct, incorrect or gradient production, while in PVJ the options were images corresponding to the production of [s], [ʃ] or undifferentiated. The presentation time of the stimuli and the reaction time were automatically controlled by the PERCEVAL software. **Results:** PVJ provided a higher percentage of identification of gradient stimuli and a shorter reaction time in performing the task compared to PAJ, both statistically significant. Spearman’s correlation test did not show statistical significance between PAJ and PVJ responses, nor for reaction time. **Conclusion:** PVJ using US images proved to be the most sensitive method for detecting gradient production in the production of fricatives [s] and [ʃ], and can be used as a complementary method to PAJ in speech analysis.

Endereço para correspondência:

Fernanda Leitão de Castro Nunes de Lima

Faculdade de Filosofia e Ciências,
Universidade Estadual Paulista –
UNESP

Rua Dante Buosi, 100, Casa 26, Jardim
Maracanã, São José do Rio Preto (SP),
Brasil, CEP: 15092-205.

E-mail:fernandaleitao@live.com

Recebido em: Junho 27, 2020

Aceito em: Outubro 02, 2020

Trabalho realizado na Faculdade de Filosofia e Ciências, Universidade Estadual Paulista Júlio de Mesquita Filho – UNESP - - Marília (SP), Brasil.

¹ Faculdade de Filosofia e Ciências, Universidade Estadual Paulista – UNESP - Marília (SP), Brasil.

Fonte de financiamento: nada a declarar.

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

Produções gradientes são aquelas consideradas intermediárias entre uma categoria fônica e outra⁽¹⁻⁴⁾. A presença de produções gradientes envolvendo a classe das fricativas tem sido detectada tanto pelo julgamento perceptivo-auditivo, usando-se escalas em seu julgamento^(5,6), quanto pelo uso de ferramentas instrumentais, seja pelo uso de análises acústicas⁽⁷⁻⁹⁾, seja pelo uso de análises articulatórias, tal como a análise ultrassonográfica do movimento da língua^(10,11).

Num estudo⁽⁷⁾, os autores analisaram acusticamente as produções de /s/ e /ʃ/ na fala de crianças falantes do Português Brasileiro (PB) com desenvolvimento típico e com Distúrbios dos Sons da Fala (DSF), que apresentavam o processo fonológico de anteriorização. Os resultados deste estudo indicaram a presença maciça de produções gradientes nas crianças com DSF, entre as duas fricativas investigadas. As autoras sugeriram que crianças com DSF não realizaram a substituição do fonema /ʃ/ pelo fonema /s/ categoricamente.

Outro estudo⁽⁸⁾, por sua vez, comparou as características acústicas das fricativas /s/ e /ʃ/ em adultos falantes da língua inglesa e da japonesa, bem como a aquisição de contrastes envolvendo esses sons, em crianças de dois e três anos de idade, de ambas as línguas. Os resultados da análise acústica das produções adultas revelaram diferenças interlinguísticas entre as duas fricativas, particularmente sobre os parâmetros acústicos utilizados para diferenciar as produções. A análise acústica dos dados infantis revelou a presença de produções gradientes, denominadas pelos autores de contrastes encobertos, tanto nas produções de crianças inglesas quanto nas produções de crianças japonesas.

Em estudo posterior⁽⁹⁾, os autores afirmaram que os ouvintes não treinados de língua inglesa tendem a rotular as produções gradientes das fricativas das crianças entre /s/ e /ʃ/ como sendo /s/, já os ouvintes não treinados da língua japonesa, o oposto. Isso significa dizer que, a depender da língua, as produções gradientes entre as fricativas coronais tenderiam a ser rotuladas mais frequentemente como /s/ ou /ʃ/.

Em outro estudo⁽¹⁰⁾, os autores analisaram qualitativamente, a partir da análise ultrassonográfica, o contorno da superfície da língua para as produções de /s/ e /ʃ/ na fala de crianças falantes do PB, duas com desenvolvimento típico e quatro com DSF: duas delas apresentavam os processos fonológicos variados, incluindo o de anteriorização (as duas crianças produziam o /ʃ/ como /s/); e as outras duas apresentavam processos fonológicos que não envolviam a palatal (/ʃ/). O padrão ultrassonográfico para crianças com desenvolvimento típico na produção de /s/ foi descrito como contorno de língua mais achatado; e para produção de /ʃ/ foi descrito como ponta de língua abaixada em direção ao assoalho da boca e o dorso elevado. Para as duas crianças que apresentavam o processo fonológico que não envolvesse /ʃ/, uma criança com DSF produziu /s/ igualmente às crianças com desenvolvimento típico (contorno de língua mais achatado), enquanto a outra criança com DSF produziu /s/ com dorso de língua mais elevado. A produção de /ʃ/ das duas crianças com DSF com processo fonológico que não envolvesse a palatal foi semelhante ao das crianças com desenvolvimento típico (ponta

de língua abaixada em direção ao assoalho da boca e o dorso elevado). Contudo, não foram observadas diferenças quanto aos contornos de língua na produção de /s/ e /ʃ/ das crianças com DSF que apresentavam o processo fonológico de anteriorização. Assim, a descrição para as crianças que produziram o /ʃ/ como /s/ foi a de língua achatada com uma pequena elevação do dorso de língua. As autoras concluíram que a análise das imagens ultrassonográficas realizadas a partir da fala de crianças com DSF confirmou o julgamento perceptivo-auditivo feito pelos fonoaudiólogos, mostrando que as duas avaliações são complementares.

Em outra pesquisa⁽¹¹⁾, através de uma análise quantitativa, os autores descreveram os critérios que podem ser utilizados no ultrassom para medir as diferenças entre os contornos da língua na produção de /s/ e /ʃ/ na fala de adultos, crianças com desenvolvimento típico e crianças com DSF, com a presença do processo fonológico de anteriorização. O resultado da análise constatou que os 11 pontos analisados na USL contribuem na diferenciação entre os contornos de língua das fricativas estudadas entre os adultos e crianças com desenvolvimento típico. Contudo, os valores das análises dos contornos de língua para a maioria das crianças com DSF apontaram que não houve diferença no posicionamento de língua na produção de /s/ e /ʃ/.

Dentre o conjunto de ferramentas instrumentais de análise articulatória, o uso da ultrassonografia do movimento da língua (USL) tem se destacado em virtude do melhor custo/benefício para detectar produções gradientes⁽¹²⁻¹⁸⁾. No entanto, não há estudos que busquem verificar se o julgamento visual das imagens ultrassonográficas poderia ser utilizado na detecção de produções gradientes.

Como se pode observar nos estudos citados, a produção gradiente só foi detectada com o uso da análise instrumental. Mais recentemente, porém, autores têm questionado se essas produções gradientes também poderiam ser detectadas em uma avaliação perceptivo-auditiva^(5,6).

Considerando o valor clínico e linguístico da presença de produções gradientes, pesquisadores têm também buscado, cada vez mais, utilizar metodologias instrumentais que permitam a sua identificação na produção de fala. No entanto, não há estudos que investiguem qual o método mais sensível para a detecção de produções gradientes.

Neste contexto, o objetivo deste estudo foi analisar em qual método (isoladamente) o julgamento perceptivo-auditivo (JPA) de áudios ou o julgamento perceptivo-visual (JPV) (imagens ultrassonográficas) se mostra como o mais sensível para detectar produções de fala gradientes na classe das fricativas coronais surdas, bem como correlacionar as duas formas de julgamentos.

As hipóteses do estudo assim se definiram:

- Espera-se, no julgamento de imagens ultrassonográficas, maior porcentagem de respostas dos juízes e menor tempo de reação;
- Espera-se uma correlação positiva entre o JPA de áudios e o JPV de imagens ultrassonográficas, no que se refere à porcentagem de respostas dos juízes e ao tempo de reação da tarefa de identificação.

MÉTODO

O presente estudo foi aprovado pelo Comitê de Ética em Pesquisa (CEP) de uma unidade universitária, sob nº 1.268.673/2015. Todos os indivíduos incluídos na pesquisa tiveram ciência e assinaram o termo de consentimento livre e esclarecido (TCLE) que lhes foi apresentado.

Participantes

Participaram deste estudo 20 juízes recrutados no Curso de Graduação em Fonoaudiologia da Unesp de Marília. Para a seleção dos juízes, adotou-se como critério de inclusão o fato de eles apresentarem conhecimento prévio, pelo cumprimento e aprovação nas duas disciplinas de Fonética e Fonologia, sobre o processo de produção de fala e da classificação e descrição fonética dos fonemas do Português Brasileiro. Os juízes relataram não ter queixas auditivas e visuais.

Inicialmente, foi realizado um esclarecimento aos juízes sobre os procedimentos envolvidos na tarefa de julgamento e uma calibração dos participantes para padronização da imagem ultrassonográfica esperada para a produção das fricativas coronais surdas, com seus respectivos exemplos, apresentados de forma sistematizada em um arquivo de *PowerPoint*.

Procedimentos

Estímulos

Foram selecionados, de um banco de dados, arquivos de áudio e vídeo (imagens ultrassonográficas) relativos à produção de palavras “sapo” e “chave”, de 11 crianças falantes do Português Brasileiro, na faixa etária de 6 a 12 anos de idade, sendo 9 meninos e 2 meninas, com produção de fala atípica.

Para a produção de cada criança, com o uso do *software Sound Forge Studio 6.0*, foram selecionados os *frames* correspondentes ao ponto máximo de constrição da língua na produção dos fonemas /s/ e /ʃ/, respectivamente, computando um total de 22 *frames*: 11 *frames* correspondentes à produção de /s/ e 11 *frames* correspondentes à produção de /ʃ/.

Os dados que constituem esse banco foram coletados com o uso de um ultrassom portátil, modelo DP 6600, contendo transdutor acoplado a um computador, microfone unidirecional e estabilizador de cabeça. Os sinais acústicos e de imagem foram registrados simultaneamente com o uso do *software AAA (Articulate Assistant Advanced)* aliado a um sincronizador que permite a sincronização entre as imagens e o sinal acústico. As imagens de ultrassom (IUS) foram adquiridas com uma frequência de 6.5 MHz, 120° de campo de imagem e 29.97 Hz de taxa de amostragem; enquanto os sinais acústicos foram adquiridos com um microfone unidirecional posicionado a 20 cm da boca do participante.

O procedimento metodológico envolveu duas etapas distintas e independentes: o JPA e o JPV (imagens ultrassonográficas), executados automaticamente pelo *software PERCEVAL*⁽¹⁹⁾.

Julgamento de áudio

O JPA dos áudios foi preparado para ser executado automaticamente pelo *software PERCEVAL*, de modo que os juízes pudessem ouvir um estímulo por vez, dentre 22 estímulos auditivos aleatorizados. Foi solicitado aos juízes escolher a categoria correspondente ao estímulo apresentado (áudio envolvendo /s/ ou /ʃ/).

O experimento envolveu três etapas: apresentação das instruções, fase treino e fase teste. Para tanto, os juízes foram dispostos individualmente de maneira confortável diante da tela do computador, com fones KOSS acoplados aos ouvidos, em uma sala silenciosa.

A etapa de apresentação das instruções, escrita na tela do computador, explicava que o juiz ouviria consecutivamente duas vezes o estímulo auditivo correspondente às palavras que envolviam a produção de /s/ e /ʃ/ produzidas por crianças. Após a apresentação do estímulo auditivo, o juiz deveria pressionar a tecla relativa à categoria correspondente ao estímulo apresentado, dentre três possibilidades: 1) produção-alvo ou acurada; 2) produção incorreta ou com substituição; 3) ou produção gradiente (distorcida), pressionando as teclas 1 ou 2 ou 3.

Na sequência, a fase de treino simulava o julgamento, com o intuito de garantir a compreensão da tarefa pelos participantes. Essa fase consistia na realização da tarefa de identificação de palavras que envolviam produções de /s/ e /ʃ/. Foram aleatorizados estímulos do experimento e selecionadas apenas cinco apresentações para o treino. Tais estímulos envolviam produções categóricas e gradientes. Os resultados obtidos nesta fase não foram computados pelo *software* e, conseqüentemente, não foram considerados na análise. Os juízes poderiam discutir com a pesquisadora eventuais dúvidas sobre a execução da tarefa. Logo após, iniciava-se a fase teste propriamente dita.

Na fase teste, os juízes, individualmente, escutavam (com apresentação binaural numa intensidade de 50 dB - NPS) um estímulo auditivo correspondente à produção de uma palavra e, em seguida, decidiam e pressionavam a tecla relativa à categoria correspondente ao estímulo apresentado, dentre três possibilidades dispostas na tela do computador: 1) produção-alvo; 2) produção incorreta; 3) ou produção distorcida (que corresponde às produções gradientes), tecendo as teclas 1, 2 e 3, respectivamente.

O tempo de apresentação dos estímulos e o tempo de resposta (ou tempo de reação) foram controlados e mensurados automaticamente pelo *software PERCEVAL*. Cada apresentação dos estímulos auditivos teve a duração de três segundos, enquanto a resposta deveria ser dada em até cinco segundos (tempo estipulado na elaboração do experimento). Caso o juiz não respondesse nesse intervalo de tempo, o *software PERCEVAL* registrava o tempo de reação como não resposta (n.a) e não voltava para refazer.

Julgamento de imagens

Analogamente ao julgamento dos áudios, o JPV (imagens ultrassonográficas) envolveu uma tarefa de identificação ou tarefa de escolha forçada executada pelo *software PERCEVAL*. O julgamento de imagens foi feito após o julgamento de áudios,

sempre com um intervalo de cinco minutos entre eles, para todos os juizes.

O julgamento das imagens foi elaborado para que os juizes analisassem uma única imagem por vez, de um total de 22 estímulos visuais aleatorizados, relacionando-a a uma das categorias que lhes foi apresentada previamente ao experimento (teste) e que apresentavam exemplos representando cada uma das possibilidades de resposta para /s/ e /ʃ/. O experimento também envolveu três etapas: apresentação das instruções, fase treino e fase teste.

A apresentação das instruções foi feita em um arquivo do programa *PowerPoint*, contendo não apenas informações importantes para interpretação das imagens (por exemplo, onde se localizam as diferentes partes da língua), como também exemplos do padrão ultrassonográfico de cada uma das possibilidades de análise, em dois padrões típicos do /s/ (ponta da língua elevada e ponta da língua abaixada), um padrão típico de /ʃ/ (Figura 1) e padrão ultrassonográfico indiferenciado, isto é, correspondente às produções gradientes (Figura 2). Na fase treino, foi realizada a simulação do julgamento, mostrando as imagens e as opções de julgamento relativas às imagens apresentadas, a fim de garantir a compreensão da tarefa pelos juizes. Aleatoriamente, selecionaram-se pelo *software* apenas cinco imagens do experimento para o julgamento, mas as respostas dos juizes e os respectivos tempos de reação não foram computados nem analisados pelo *software* PERCEVAL.

Na fase teste, cada imagem era selecionada aleatoriamente pelo *software* e apresentada na tela do computador. Em seguida, eram dispostas na tela as três possibilidades de julgamento (1) padrão de /s/ (considerando as duas possibilidades de

posicionamento da ponta da língua); 2) padrão de /ʃ/; 3) e imagem indiferenciada, para que os juizes decidissem e pressionassem a tecla correspondente à categoria escolhida.

O tempo de apresentação dos estímulos e o tempo de resposta (ou tempo de reação) foram controlados e mensurados automaticamente pelo *software* PERCEVAL. Cada apresentação dos estímulos visuais teve a duração de três segundos. A resposta deveria ser dada em um intervalo de tempo de até cinco segundos; caso o juiz não respondesse neste intervalo de tempo, o *software* mensuraria o tempo de reação como não resposta (n.a). A duração total do experimento foi de, aproximadamente, dez minutos por participante.

Finalmente, fez-se uma análise estatística descritiva dos dados, considerando a porcentagem de resposta dos juizes e o tempo de reação nos julgamentos perceptivo-auditivo dos áudios e perceptivo-visual das imagens ultrassonográficas, a partir das categorias prévias de cada experimento. Utilizaram-se os seguintes testes: o teste de Mann-Whitney, para comparar a porcentagem de respostas e o tempo de reação no julgamento dos áudios e das imagens; o teste Wilcoxon, para comparar a porcentagem de respostas e o tempo de reação dos estímulos categóricos e gradientes; e o teste de correlação Spearman, para correlacionar as duas formas de julgamentos. O valor de p estabelecido foi o de < 0,05.

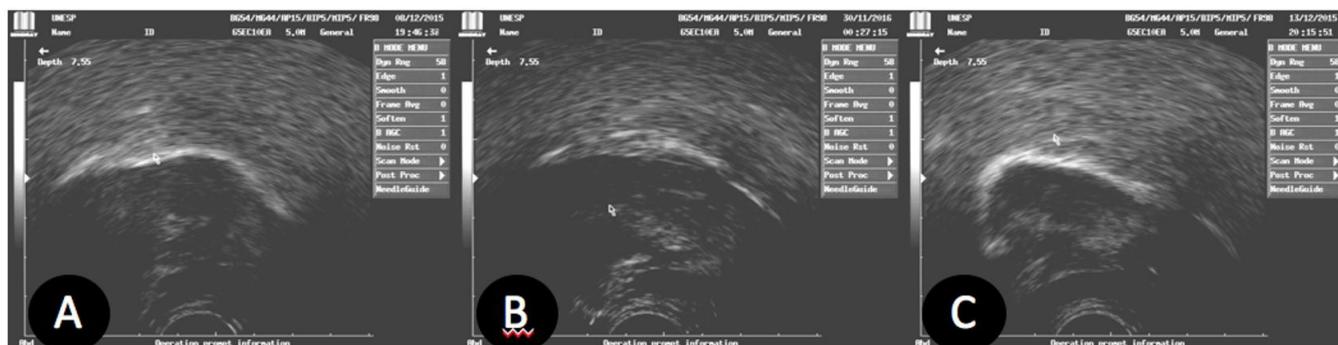
RESULTADOS

A Tabela 1 apresenta os resultados das análises do JPA, a porcentagem de respostas dos juizes e a média do tempo de reação em função das categorias adotadas. Na categoria



Na imagem (A), a seta indica /s/ com sulco; em (B), a seta indica /s/ sem sulco; e na imagem (C), referente à produção de /ʃ/, a seta indica lâmina de língua elevada em forma côncava com ponta de língua para baixo. Em todas as imagens ultrassonográficas, a ponta de língua encontra-se para esquerda.

Figura 1. Padrão ultrassonográfico das produções alvo de /s/ e /ʃ/



A imagem (A) e (B) correspondem ao padrão ultrassonográfico indiferenciado de /s/; em (C), corresponde ao padrão ultrassonográfico indiferenciado de /ʃ/. Em todas as imagens ultrassonográficas, a ponta de língua encontra-se para esquerda

Figura 2. Padrão ultrassonográfico das produções consideradas atípicas de /s/ e /ʃ/

Tabela 1. Porcentagem de respostas dos juizes por categorias e o tempo médio de reação no julgamento perceptivo-auditivo (áudios de amostras de fala)

Estímulos auditivos	Categorias do julgamento	Categorias de respostas	Respostas dos juizes	Tempo de reação (ms)
		(Áudio)	N (%)	Média (DP)
<i>Alvo /s/</i>	Categórico	Típica	157 (71,36)	2505,48 (±724,77)
		Incorreta	22 (10)	2741,40 (±714,85)
	Gradiente	Gradiente	38 (17,27)	3132,75 (±835,70)
	Não resposta	Não resposta	3 (1,36)	0
<i>Alvo /l/</i>	Categórico	Típica	32 (14,54)	2963,71 (±1046,89)
		Incorreta	124 (56,36)	2690,52 (±862)
	Gradiente	Gradiente	42 (19,09)	3119,78 (±768,94)
	Não resposta	Não resposta	22 (10)	0
<i>Total</i>	Categórico (alvos /s/ e /l/)	Produção típica e incorreta	335 (76,13)	2725,27 (±188,52)
	Gradiente (alvos /s/ e /l/)	Produção gradiente	80 (18,18)	3126,26 (±9,17)
	Não resposta	Não resposta	25 (5,45)	0

Teste Wilcoxon

Legenda: ms = milissegundos

Tabela 2. Percentual de respostas dos juizes, por categorias, e tempo médio de reação no julgamento perceptivo-visual (imagens de ultrassom).

Estímulos visuais	Categorias do Julgamento	Categorias de resposta	Resposta dos juizes	Tempo de reação (ms)
		(Imagem)	N (%)	Média (DP)
<i>Alvo /s/</i>	Categórico	Imagem /s/	78 (35,45)	978,54(±524,48)
		Imagem /l/	55(25)	1052,38(±609,51)
	Gradiente	Indiferenciada	85(38,63)	1011,39(±591,82)
	Não resposta	Não resposta	2(0,90)	0
<i>Alvo /l/</i>	Categórico	Imagem de /s/	83(37,72)	1006,84(±626,02)
		Imagem de /l/	84(38,18)	1021,31(±622,72)
	Gradiente	Imagem indiferenciada	52(23,63)	1134,86(±615,07)
	Não resposta	Não resposta	1(0,45)	0
<i>Total</i>	Categórico	Imagem de /s/ Imagem de /l/	300 (68,18)	1014,76(±30,72)
	Gradiente	Imagem indiferenciada	137 (31,13)	1073,12(±87,30)
	Não resposta	Não resposta	3 (0,68)	0

Teste Wilcoxon ms = milissegundos

de julgamento “categórico”, consideraram-se as respostas de produção típica e de produção incorreta/substituída; e para categoria gradiente, as respostas de produção gradiente (distorcida, auditivamente).

Na última linha da Tabela 1, observa-se que o número de não respostas foi muito pequeno (5,45%), por esse motivo, essa categoria não foi considerada na análise estatística.

A Tabela 2 apresenta os resultados das análises do JPV por imagens ultrassonográficas, a média da porcentagem de respostas dos juizes e do tempo de reação em função das categorias adotadas. Na categoria de julgamento “categórico”, foram consideradas as respostas de imagem de /s/ e imagem de /l/; e para a categoria gradiente, as respostas de imagem indiferenciada.

Da mesma forma que o JPA, o número de observações de não resposta foi muito pequeno, conforme disposto na última linha da Tabela 2, por isso não foram considerados na análise estatística realizada.

A Tabela 3 apresenta comparativamente a porcentagem de respostas dos juizes e o tempo médio de reação nos julgamentos de áudio e de imagens.

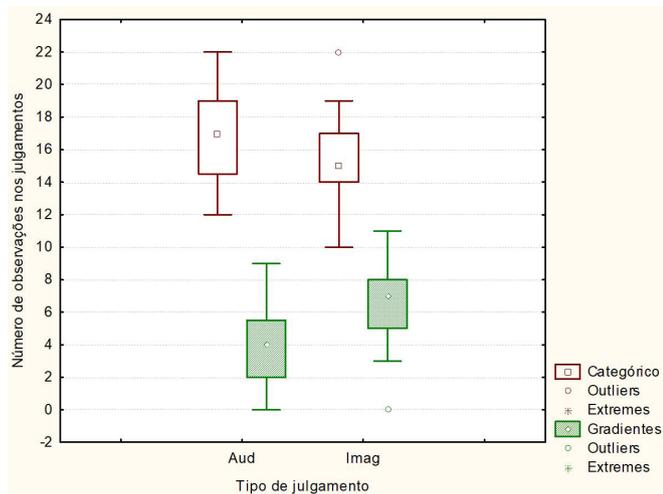
Na comparação da porcentagem de respostas dos juizes e do tempo de reação entre o julgamento dos áudios e das imagens, observou-se, de acordo com o teste Mann-Whitney, que todas

Tabela 3. Comparação da porcentagem de respostas dos juizes e o tempo médio de reação nos julgamentos de áudio e imagem.

Julgamentos	Categorias do Julgamento	Porcentagem de Respostas dos juizes N (%)	Tempo de reação (ms)
			Média (DP)
<i>Julgamento de áudio</i>	Categórico	335 (76,13%)	2725,27 (±188,52)
	Gradiente	80 (18,18%)	3126,26 (±9,17)
<i>Julgamento de imagens</i>	Categórico	300 (68,18%)	1014,76 (±30,72)
	Gradiente	137 (31,13%)	1073,12 (±87,30)

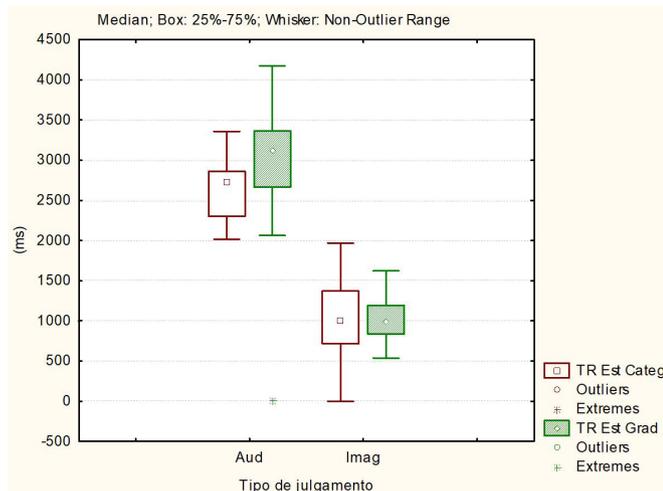
Teste Mann-Whitney ms = milissegundos

as categorias comparadas apresentaram diferenças significativas (todos $p < 0,05$). Em relação aos estímulos julgados como categóricos, os resultados da comparação entre os julgamentos de áudios e de imagens mostraram que o JPA apresentou maior número de ocorrências comparativamente ao JPV ($Z(20)=2,00$, $p=0,04$). Por outro lado, os estímulos julgados como gradientes foram identificados com maior ocorrência pelo JPV, julgamento perceptivo-visual de imagens ($Z(20)=2,00$, $p=0,04$), ao passo



Aud – Auditivo Imag – Imagens Teste Mann-Whitney

Figura 3. Comparação da porcentagem de respostas dos juízes em geral entre os tipos de julgamentos



Aud – Auditivo Imag – Imagens Teste Mann-Whitney

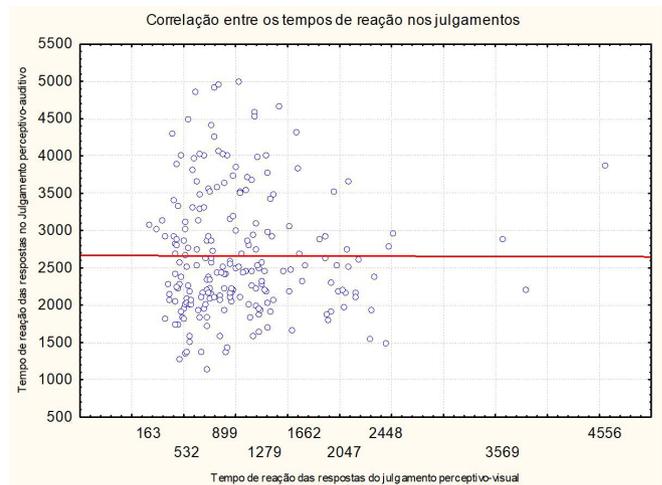
Figura 4. Comparação do tempo de reação para cada tipo de julgamento

que o JPV propiciou maior identificação dos estímulos julgados como gradientes ($Z(20)=-3,24, p<0,00$).

A Figura 3 ilustra a diferença de porcentagem de respostas dos juízes entre os tipos de julgamentos.

Analogamente, na comparação da média do tempo de reação despendido durante o julgamento pelos juízes, verificou-se, a partir do teste Mann-Whitney, que o tempo de reação no julgamento auditivo foi significativamente superior ao tempo de reação no julgamento de imagens, tanto para o julgamento dos estímulos julgados como categóricos ($Z(2)=5,41, p<0,00$) quanto para o julgamento dos estímulos julgados como gradientes ($Z(20)=4,32, p<0,00$) (Figura 4).

Finalmente, ao correlacionar a porcentagem de respostas dos juízes e o tempo de reação em ambas as tarefas de julgamento, o teste de correlação de Spearman não mostrou significância estatística para porcentagem de respostas ($t(N-2)=-1,03, p=0,29$) nem para o tempo de reação ($t(N-2)=0,36, p=0,71$), conforme ilustra a Figura 4, respectivamente.



Teste de correlação de Spearman

Figura 5. Correlação entre o tempo de reação nos julgamentos perceptivo-auditivo e visual-perceptivo

A Figura 5 apresenta correlação entre o tempo de reação em ambos os julgamentos.

DISCUSSÃO

Considerando a existência de produções categóricas e gradientes nos erros de produção de fala⁽¹⁻⁴⁾, o objetivo deste estudo foi analisar qual método (JPA ou JPV) é mais sensível para detectar produções gradientes na classe das fricativas coronais surdas, bem como correlacionar as duas formas de julgamentos.

Embora fosse possível detectar produções gradientes tanto no JPA quanto no JPV, na primeira hipótese, no julgamento de imagens ultrassonográficas, esperava-se maior porcentagem de respostas e menor tempo de reação, e os resultados obtidos confirmaram integralmente esta hipótese. Inicialmente, observou-se que, em ambos os julgamentos, as produções gradientes foram detectadas, sugerindo que estes dois métodos de julgamento perceptivos são válidos para detectar se há presença de produções gradientes, embora a ocorrência das respostas dos juízes varie entre os métodos.

Estudos prévios que utilizaram exclusivamente o JPA também reportaram a presença de produções gradientes, entretanto constataram que a inclusão de categorias intermediárias no julgamento é condição imprescindível para que as produções gradientes sejam identificadas^(5-7,20,21). No presente estudo, a inclusão da categoria gradiente (intermediária) favoreceu a identificação auditiva de produções gradientes nas fricativas coronais surdas, pelos juízes.

Especificamente, em outro estudo⁽⁵⁾, os autores ressaltaram que a informação auditiva como única forma de análise pode não ser totalmente eficiente na detecção de um som gradiente. A percepção dos ouvintes pode ser alterada quando lhes é apresentada uma categoria intermediária juntamente a uma produção categórica. Adicionalmente, os autores sugerem o uso de informações instrumentais que facilitem a detecção da produção gradiente durante o JPA.

Da mesma forma, um estudo⁽²¹⁾ que utilizou a imagem ultrassonográfica para identificar produções gradientes na

produção de líquidas do inglês, evidenciou a pertinência desse método e alertou sobre a importância da apresentação da categoria intermediária aos juízes, uma vez que, em geral, as tarefas de julgamento de amostras de fala envolvem apenas respostas relativas às categorias fonêmicas, por exemplo, ou o estímulo é julgado como /s/ ou como /ʃ/. Neste mesmo estudo, a imagem ultrassonográfica para identificar produções gradientes nas fricativas surdas também foi utilizada com apresentação da categoria intermediária aos juízes, o que pode ter favorecido a identificação desta categoria nas imagens apresentadas.

O presente estudo traz informações inéditas ao explicitamente comparar resultados de ambos os métodos (perceptivo-auditivo e perceptivo-visual) para a detecção de produções gradientes, particularmente na classe das fricativas coronais surdas. Embora não haja pesquisas prévias que comparem ambos os métodos de detecção de produções gradientes (independentemente da classe fônica), os estudos apresentados na literatura, até o momento, mostraram que o julgamento de imagens ultrassonográficas facilita essa tarefa⁽¹²⁻¹⁸⁾. Ao longo do desenvolvimento da linguagem, os falantes aprendem a perceber a fala categoricamente, ou seja, atrelados às categorias pré-existentes da língua⁽²²⁾, o que levaria os ouvintes a julgarem os estímulos auditivos de forma categórica. A imagem, diferentemente, propicia a visualização direta do movimento articulatorio que gerou o estímulo, podendo ser um facilitador (uma vez que independe de uma inferência a partir do estímulo auditivo) para a detecção de produção gradiente⁽²³⁻²⁵⁾, fato que justificaria maior porcentagem de respostas e menor tempo de reação de produções gradientes, a partir das imagens ultrassonográficas (JPV) analisadas.

A segunda hipótese não foi confirmada nos resultados obtidos, pois se esperava uma correlação positiva entre o JPA e o JPV no que se refere à porcentagem de respostas dos juízes e ao tempo de reação.

Esses resultados não corroboraram os achados de um estudo^(26,27), segundo os quais as medidas qualitativas do movimento da língua na imagem ultrassonográfica tiveram uma correlação com as medidas perceptivas e acústicas dos sons /r/ em crianças com DSF e de fala típica, porém confirmaram os resultados descritos noutro estudo⁽²⁸⁾. Neste último estudo citado, os autores verificaram uma variabilidade de produção da retroflexa no inglês britânico, em que os falantes asiáticos típicos produzem /l/ e /r/ com uma constrição anterior, com a frente ou ponta da língua, enquanto que os falantes anglos normalmente produzem /l/ e /r/ com uma constrição mais posterior, com o dorso da língua retraído. Assim, os autores⁽²⁸⁾ ainda verificaram que, apesar da correlação entre a articulação, a acústica e a audição, os padrões articulatorios individuais do inglês britânico nem sempre se correlacionam com as variações auditivas e acústicas das formas esperadas. Os autores concluíram que a produção de fala recebe influência fonética da língua materna, e cada indivíduo adquire uma variabilidade de produção de acordo com o seu contato com a língua.

A depender da língua materna, a produção de fala pode sofrer variabilidades e os padrões articulatorios, acústicos e auditivos não se correlacionam, como foi demonstrado neste estudo.

Ainda no presente estudo, a ultrassonografia de língua (USL) mostrou ser uma técnica eficaz e facilitadora na detecção da

produção gradiente das fricativas coronais surdas: ela pode acrescentar critérios valiosos para avaliação, diagnóstico e processos terapêuticos fonoaudiológicos, bem como ser uma ferramenta importante para monitorar resultados após intervenção terapêutica. A detecção da produção gradiente, por meio da USL, evidencia o conhecimento fonológico da criança, podendo tornar o processo terapêutico mais rápido, ou seja, se a criança apresentar uma produção gradiente, ela já estará a meio do caminho para atingir o alvo, o que significa menor tempo de terapia. Além disso, vale ressaltar o maior custo-benefício da avaliação da USL se comparada com as demais análises articulatórias, o que favorece seu uso para fins clínicos e de pesquisa.

Embora o presente estudo tenha trazido contribuições relevantes sobre o uso da USL como ferramenta para análise gradiente da produção de fala, é importante realizar, em estudos posteriores, uma investigação perceptivo-auditiva e perceptivo-visual através de imagens ultrassonográficas, de outros contrastes fônicos, a fim de verificar se as produções gradientes poderão ser resgatadas por essas medidas em outras classes de sons da língua.

CONCLUSÃO

O estudo constatou a capacidade dos juízes em detectar produções gradientes dentre a classe das fricativas coronais surdas, tanto no JPA quanto no JPV. Foi possível verificar maior porcentagem de respostas e menor tempo de reação para estímulos gradientes no JPV, evidenciando que o uso de imagens ultrassonográficas é o método mais sensível e facilitador para a detecção da produção gradiente na produção de fricativas coronais surdas.

REFERÊNCIAS

1. Albano E. O Gesto e suas Bordas: esboço de Fonologia Acústico-Articulatória do Português Brasileiro. Campinas: Mercado de Letras; São Paulo: FAPESP; 2001.
2. Pouplier M, Goldstein L. Asymmetries in the perception of speech production errors. *J Phonetics*. 2005;33(1):47-75. <http://dx.doi.org/10.1016/j.wocn.2004.04.001>.
3. Goldstein L, Pouplier M, Chen L, Saltzman E, Byrd D. Dynamic action units slip in speech production errors. *Cognition*. 2007;103(3):386-412. <http://dx.doi.org/10.1016/j.cognition.2006.05.010>. PMID:16822494.
4. Rodrigues L, Freitas M, Berti L, Albano E. Acertos Gradientes nos chamados erros de pronúncia. *Letras*. 2008;36:86-112. <https://doi.org/10.5902/2176148511968>.
5. Munson B, Edwards J, Schellinger SK, Beckman ME, Meyer MK. Deconstructing phonetic transcription: covert contrast, perceptual bias, and an extraterrestrial view of Vox Humana. *Clin Linguist Phon*. 2010;24(4-5):245-60. <http://dx.doi.org/10.3109/02699200903532524>. PMID:20345255.
6. Munson B, Schellinger SK, Carlson KU. Measuring speech-sound learning using visual analog scaling. *Perspect Lang Learn Educ*. 2012;19(1):19-30. <http://dx.doi.org/10.1044/lle19.1.19>.
7. Berti L, Marino V. Marcas linguísticas constitutivas do processo de aquisição do contraste fônico. *Revista do GEL*. 2008;5(2):103-21.
8. Li F, Edwards J, Beckman ME. Contrast and covert contrast: the phonetic development of voiceless sibilant fricatives in English and Japanese toddlers. *J Phon*. 2009;37(1):11-24. <http://dx.doi.org/10.1016/j.wocn.2008.10.001>. PMID:19672472.
9. Li F, Munson B, Edwards J, Yoneyama K, Hall K. Language specificity in the perception of voiceless sibilant fricatives in Japanese and English: implications for cross-language differences in speech-sound development. *J*

- Acoust Soc Am. 2011;129(2):999-1011. <http://dx.doi.org/10.1121/1.3518716>. PMID:21361456.
10. Wertzner HF, Francisco DT, Pagan-Neves LO. Tongue contour for /s/ and /ʔ/ in children with speech sound disorder. *CoDAS*. 2014;26(3):248-51. <http://dx.doi.org/10.1590/2317-1782/201420130022>. PMID:25118923.
 11. Francisco DT, Wertzner HF. Differences between the production of [s] and [ʃ] in the speech of adults, typically developing children, and children with speech sound disorders: an ultrasound study. *Clin Linguist Phon*. 2017;31(5):375-90. <http://dx.doi.org/10.1080/02699206.2016.1269204>. PMID:28085504.
 12. Bressmann T, Thind P, Uy C, Bollig C, Gilbert RW, Irish JC. Quantitative three-dimensional ultrasound analysis of tongue protrusion, grooving and symmetry: data from 12 normal speakers and a partial glossectomee. *Clin Linguist Phon*. 2005;19(6-7):573-88. <http://dx.doi.org/10.1080/02699200500113947>. PMID:16206485.
 13. Zharkova N, Hewlett N, Hardcastle WJ. An ultrasound study of lingual coarticulation in /sV/ syllables produced by adults and typically developing children. *J Inter Phon Assoc*. 2012;42(2):193-208.
 14. Berti L. Investigação ultrassonográfica dos erros de fala infantil à luz da Fonologia Gestual. In: Ferreira-Gonçalves G, Brum-de-Paula M, editores. *Dinâmica dos Movimentos Articulatorios: sons, gestos, imagens*. Pelotas: UFPel; 2013. p. 127-44.
 15. Barberena LS, Keske-Soares M, Berti LC. Descrição dos gestos articulatorios envolvidos na produção dos sons /t/ e /l/. *Audiol Commun Res*. 2014;19(4):338-44. <http://dx.doi.org/10.1590/S2317-6431201400040000135>.
 16. Berti L, Boer GD, Bressmann T. Tongue displacement and durational characteristics of normal and disordered Brazilian Portuguese liquids. *Clin Linguist Phon*. 2016;30(2):131-49. <http://dx.doi.org/10.3109/02699206.2015.1116607>. PMID:26853548.
 17. Lima FLCN, Silva CEE, Silva LM, Vassoler AMO, Fabbron EMG, Berti LC. Ultrasonographic analysis of lateral liquids and coronal fricatives: judgment of experienced and non-experienced judges. *Rev CEFAC*. 2018;20(4):422-31. <http://dx.doi.org/10.1590/1982-0216201820412317>.
 18. Vassoler A, Berti L. Padrões silábicos no desenvolvimento fonológico típico e atípico: análise ultrassonográfica. *CoDAS*. 2018;30(2). <http://dx.doi.org/10.1590/2317-1782/20182017067>.
 19. Andre C. Perceval: perception evaluation auditive e visuelle. Version 5.0. France: Aix-en-Provence; 2009.
 20. Urberg-Carlson K, Kaiser E, Munson B. Assessment of children's speech production 2: testing gradient measures of children's productions. Chicago: ASHA Convention; 2008.
 21. McAllister Byun T, Buchwald A, Mizoguchi A. Covert contrast in velar fronting: an acoustic and ultrasound study. *Clin Linguist Phon*. 2016;30(3-5):249-76. <http://dx.doi.org/10.3109/02699206.2015.1056884>. PMID:26325303.
 22. Lamprecht R. Aquisição fonológica do Português: perfil de desenvolvimento e subsídios para terapia. Porto Alegre: Artmed; 2004. 232 p.
 23. Gick B. The use of ultrasound for linguistic phonetic fieldwork. *J Intern Phon Assoc*. 2002;32(2):113-22.
 24. Stone M. A guide to analyzing tongue motion from ultrasound Images. *Clin Linguist Phon*. 2005;19(6-7):455-501. <http://dx.doi.org/10.1080/02699200500113558>. PMID:16206478.
 25. Scoobie J. Ultrasound-based tongue root imaging and measurement. In: Workshop on Pharyngeals and Pharyngealisations; 2009; Newcastle Upon Tyne Proceedings. Newcastle: Upon Tyne; 2009.
 26. Klein HB, McAllister Byun T, Davidson L, Grigos MI. A multidimensional investigation of children's /r/ productions: perceptual, ultrasound, and acoustic measures. *Am J Speech Lang Pathol*. 2013;22(3):540-53. [http://dx.doi.org/10.1044/1058-0360\(2013/12-0137\)](http://dx.doi.org/10.1044/1058-0360(2013/12-0137)). PMID:23813195.
 27. Preston JL, McCabe P, Tiede M, Whalen DH. Tongue shapes for rhotics in school-age children with and without residual speech errors. *Clin Linguist Phon*. 2019;33(4):334-48. <http://dx.doi.org/10.1080/02699206.2018.1517190>. PMID:30199271.
 28. Kirkham S, Wormald J. Acoustic and articulatory variation in British Asian English liquids. In: XVIII International Congress of Phonetic Sciences; 2015; Glasgow. Proceedings. Glasgow: Glasgow University; 2015. p. 1-5.

Contribuição dos autores

FLCNL foi responsável pela elaboração do projeto e coleta, seleção da amostra, tabulação e análise dos dados, e redação do manuscrito; LCB foi responsável pelo delineamento do estudo, análise dos dados, discussão dos achados, coorientação da pesquisa e escrita do artigo; VCCM foi responsável pela análise dos dados e escrita do artigo científico; EMGF foi responsável pelo projeto, delineamento do estudo, discussão dos achados e orientação geral das etapas de execução, elaboração e escrita do manuscrito.