

Fernanda Leitão de Castro Nunes de Lima¹ 
Larissa Cristina Berti¹ 
Viviane Cristina de Castro Marino¹ 
Eliana Maria Gradim Fabbron¹ 

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

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

Keywords

Auditory Perception
Visual Perception
Ultrasonography
Fricatives
Portuguese Language

Descritores

Percepção Auditiva
Percepção Visual
Ultrassonografia
Fricativas
Língua Portuguesa

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.

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.

Correspondence address:

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

Received: June 27, 2020

Accepted: October 02, 2020

Study conducted at 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.

Financial support: nothing to declare.

Conflict of interests: nothing to declare.



This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

Gradient productions are those classified as intermediate between two different phonic categories⁽¹⁻⁴⁾. The presence of gradient productions involving the fricative class has been detected both by the auditory-perceptual assessment, using scales^(5,6), and by the use of instrumental tools, either by the use of acoustic analysis⁽⁷⁻⁹⁾ or by the use of articulatory analysis, such as the ultrasound analysis of the movement of the tongue^(10,11).

The authors of a study⁽⁷⁾ performed an acoustic analysis of the productions of /s/ and /ʃ/ of Brazilian Portuguese-speaking (BP) children with typical development and with Speech Sounds Disorders (SSD), who had the phonological process of anteriorization. The results of this study reported a significant presence of gradient production in children with SSD, in the two fricatives being investigated. The authors suggested that children with SSD did not replace the phoneme /ʃ/ with the phoneme /s/ categorically.

In turn, another study⁽⁸⁾ compared the acoustic characteristics of the /s/ and /ʃ/ fricatives in English-speaking and Japanese-speaking adults, as well as the acquisition of contrasts involving these sounds, in children of two and three years of age, from both languages. The results of the acoustic analysis of the productions of adult participants showed interlinguistic differences between the two fricatives, particularly on the acoustic parameters used to differentiate the productions. The acoustic analysis of the data of child participants showed the presence of gradient productions, called by the authors as hidden contrasts, both in the productions of English- and Japanese-speaking children.

In a subsequent study⁽⁹⁾, the authors reported that untrained English-speaking listeners tend to report children's fricative gradient productions between /s/ and /ʃ/ as being /s/, while untrained Japanese-speaking listeners do the opposite. This means that, depending on the language, the gradient productions between the coronal fricatives would tend to be reported more frequently as /s/ or /ʃ/.

In another study⁽¹⁰⁾ using ultrasound analysis, the authors carried out a qualitative analysis of the contour of the surface of the tongue for the productions of /s/ and /ʃ/ in the speech of BP-speaking children, two of whom had typical development and four had SSD: two of them had varied phonological processes, including anteriorization (both children produced /ʃ/ as /s/); and the other two had phonological processes that not involving the palatal (/ʃ/). The ultrasound pattern for children with typical development in the production of /s/ showed a more flattened tongue contour; while for the production of /ʃ/, there was a pattern with the tip of the tongue lowered towards the floor of the mouth and the raised back. Two children had a phonological process that did not involve /ʃ/, one child had SSD and produced /s/ equally to children with typical development (more flattened tongue contour), while the other child with SSD produced the /s/ with higher back of tongue. The production of /ʃ/ of the two children with SSD with a phonological process that did not involve the palate was similar to the production of children with typical development (tip of the tongue lowered towards the floor of the mouth and the elevated back). However, no differences were observed regarding the language contours

in the production of /s/ and /ʃ/ of children with SSD who had anteriorization. Thus, the result of children who produced /ʃ/ as /s/ was reported as a flattened tongue with a slight elevation of the back of the tongue. Therefore, the authors concluded that the analysis of ultrasound images performed from the speech of children with SSD confirmed the auditory-perceptual assessment of speech-language pathologists, showing that the two assessments are complementary.

In another study⁽¹¹⁾ including a quantitative analysis, the authors reported the criteria that can be used in ultrasound to measure the differences between the contours of the tongue in the production of /s/ and /ʃ/ in the speech of adults, children with typical development and children with SSD, with the presence of anteriorization. The result of the analysis found that the 11 points analyzed in the TUS contribute to the differentiation between the contours of the tongue of the investigated fricatives between adults and children with typical development. However, the values of the analysis of the contours of the tongue for most children with SSD showed no difference in the positioning of the tongue in the production of /s/ and /ʃ/.

The ultrasound of the movement of the tongue (TUS) has stood out among the set of instrumental tools of articulatory analysis, due to the best cost-effectiveness to detect gradient productions⁽¹²⁻¹⁸⁾. However, there is a lack of studies investigating whether visual assessment of ultrasound images could be used to detect gradient productions.

As observed in the aforementioned studies, the gradient production was only detected with the use of instrumental analysis. However, authors have recently questioned whether these gradient productions could also be detected in an auditory-perceptual assessment.^(5,6)

Given the clinical and linguistic value of the presence of gradient productions, researchers have also increasingly used instrumental methodologies that allow their identification in speech production. However, there is also a lack of studies investigating the most sensitive method in order to detect gradient productions.

In this context, this study aimed to investigate in which method (in isolation) the auditory-perceptual assessment (APA) of audios or the visual-perceptual assessment (VPA) (ultrasound images) is shown to be the most sensitive to detect gradient speech productions in the class of voiceless coronal fricatives, as well as to correlate the two assessments.

Two hypotheses were elaborated for the study:

- When evaluating ultrasound images, a higher percentage of responses from the evaluators and a shorter reaction time are expected;
- A positive correlation is expected between the APA of audios and the VPA of ultrasound images, with regard to the percentage of responses from the evaluators and the reaction time of the identification task.

METHODS

This study was approved by the Research Ethics Committee of a University, under the no. 1.268.673/2015. All individuals

enrolled in the study were informed and signed the informed consent form (ICF).

Participants

This study included 20 evaluators from the Undergraduate Course in Speech-Language Pathology at Unesp (Marília). The inclusion criterion for the selection of evaluators included their prior knowledge, through compliance and approval in the two subjects of Phonetics and Phonology, about the speech production process and the phonetic classification and description of Brazilian Portuguese phonemes. The evaluators reported no auditory and visual complaints.

Initially, the study conducted a session to clarify to the evaluators the procedures involved in the task and a calibration of the participants to standardize the expected ultrasound image for the production of voiceless coronal fricatives, with their respective examples, presented in a systematic way in a Power Point presentation.

Procedures

Stimuli

Audio and video files (ultrasound images) related to the production of the words “frog” (*sapo*) and “key” (*chave*) were selected from a database of 11 Brazilian Portuguese-speaking children, aged 6 to 12 years old, being 9 boys and 2 girls, with atypical speech production.

Through Sound Forge Studio 6.0 software, the frames corresponding to the maximum point of constriction of the tongue in the production of the phonemes /s/ and /ʃ/, respectively, were selected in the production of each child reaching a total of 22 frames: 11 frames corresponding to the production of /s/ and 11 frames corresponding to the production of /ʃ/.

The data in this database were collected using a DP 6600 portable ultrasound, including a transducer coupled to a computer, unidirectional microphone and head stabilizer. The acoustic and image signals were recorded simultaneously using the Articulate Assistant Advanced (AAA) software, together with a synchronizer that allows the synchronization between the images and the acoustic signal. Ultrasound images (USI) were obtained with a 6.5 MHz frequency, 120° image field and 29.97 Hz sampling rate; while the acoustic signals were obtained using a unidirectional microphone positioned at 20 cm from the participant’s mouth.

The methodological procedure involved two distinct and independent steps: APA and VPA (ultrasound images), automatically performed using the PERCEVAL⁽¹⁹⁾ software.

Audio evaluation

The APA of the audios was prepared to be executed automatically by the PERCEVAL software, so that the evaluators could hear one stimulus at a time, among 22 randomized auditory stimuli. Then, the evaluators were asked to choose the category corresponding to the stimulus presented (audio involving /s/ or /ʃ/).

The experiment involved three stages: presentation of instructions, training and test. Therefore, the evaluators were individually arranged in a comfortable way in front of the computer screen, with KOSS headphones, in a quiet room.

The stage of presentation of the instructions, shown on the computer screen, explained that the evaluator would hear consecutively the auditory stimulus corresponding to the words that involved the production of /s/ and /ʃ/ produced by children. After the presentation of the auditory stimulus, the evaluator was asked to press the key related to the category corresponding to the presented stimulus, among three possibilities: 1) target or accurate production; 2) incorrect or substituted production; 3) or gradient production (distorted), pressing keys 1 or 2 or 3.

Then, the training stage simulated the assessment in order to guarantee the participants’ understanding of the task. This stage included performing the task of identifying words involving the production of /s/ and /ʃ/. Experiment stimuli were randomized and only five presentations were selected for training. These stimuli involved categorical productions and gradients. The results obtained in this stage were not computed by the software and, consequently, were not included in the analysis. The evaluators could clarify any doubts about the execution of the task with the researcher. Then, the test phase was started.

In this test phase, the evaluators individually listened (with binaural production at an intensity of 50 dB - SPL) to an auditory stimulus corresponding to the production of a word, and then decided and pressed the key related to the category corresponding to the stimulus presented, among three possibilities provided on the computer screen: 1) target production; 2) incorrect production; 3) or distorted production (corresponding to gradient productions), pressing keys 1, 2 and 3, respectively.

The presentation time of the stimuli and the response time (or reaction time) were monitored and measured automatically through the PERCEVAL software. Each presentation of the auditory stimuli lasted for three seconds, while the answer should be provided in up to five seconds (as defined in the design of the experiment). If the evaluator did not provide a response within that time, the PERCEVAL software recorded the reaction time as “no answer” (n.a) and did not offer the option to redo.

Image evaluation

Analogously to the audio evaluation, the VPA (ultrasound images) involved an identification task or a forced choice task performed by the PERCEVAL software. The evaluation of images was performed after the evaluation of audios, always providing a 5-minute break between them, for all the evaluators.

The image evaluation was designed so that the evaluators could analyze a single image at a time, out of a total of 22 randomized visual stimuli, relating it to one of the categories that was presented to them prior to the experiment (test) and that had examples representing each of the answer possibilities for /s/ and /ʃ/. The experiment also involved three stages: presentation of instructions, training and test.

The instructions were presented in a PowerPoint presentation, including not only important information for interpreting the images (such as the location of different parts of the tongue), but

also examples of the ultrasound pattern of each of the possibilities of analysis, in two typical patterns of /s/ (tip of the tongue raised and tip of the tongue lowered), a typical pattern of /j/ (Figure 1) and undifferentiated ultrasound pattern, corresponding to the gradient productions (Figure 2). The training stage included a simulation of the evaluation, showing the images and the evaluation options related to the images presented, in order to ensure that the evaluators understood the task. Only five images of the experiment were selected at random by the software for the evaluation, but the responses of the evaluators and their respective reaction times were not computed or analyzed by the PERCEVAL software.

In the test stage, each image was selected randomly by the software and presented on the computer screen. Then, the screen showed the three choices: (1) pattern of /s/ (considering the two possibilities of positioning the tip of the tongue); (2) pattern of /j/; (3) and undifferentiated image, so that the evaluators could decide and press the key corresponding to the chosen category.

The presentation time of the stimuli and the response time (or reaction time) were monitored and measured automatically through the PERCEVAL software. Each visual stimulus was presented for three seconds. The answer should be provided within five seconds; if the evaluator did not respond within this time interval, the software would classify the reaction time as “no answer” (n.a). The experiment had lasted approximately ten minutes per participant.

Finally, the study conducted a descriptive statistical analysis of the data, considering the percentage of response of the evaluators and the reaction time in the auditory-perceptual and auditory-visual evaluations of the ultrasound images, from the previous categories of each experiment. The following tests were applied: the Mann-Whitney U test, in order to compare the percentage of responses and the reaction time in the evaluation of audios and images; the Wilcoxon signed-rank test, in order to compare the percentage of responses and the reaction time of categorical stimuli and gradients; and the Spearman’s rank correlation coefficient, in order to correlate the two different assessments. A p -value $< 0,05$ was adopted to conduct the tests.

RESULTS

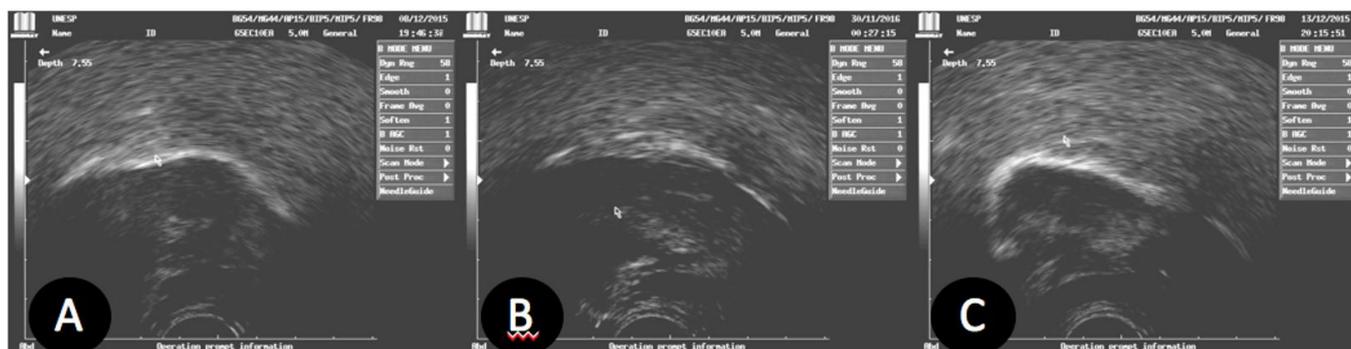
Table 1 shows the results of the APA analyzes, the percentage of responses by the evaluators and the average reaction time according to the categories adopted. In the “categorical” evaluation category, the responses of “typical production” and “incorrect/substituted production” were considered; while the “gradient” category included the “gradient” (distorted, auditory) production responses.

The last line of Table 1 shows that there was a very small number (5.45%) of “no answers” and, for this reason, this category was not considered in the statistical analysis.



The arrow in the image (A) indicates /s/ with groove; while the arrow in (B) indicates /s/ without groove; and the arrow in the image (C), which refers to the production of /j/, indicates a concave-shaped raised tongue blade with the tip of the tongue down. In all ultrasound images, the tip of the tongue is to the left

Figure 1. Ultrasonographic pattern of the target productions of /s/ and /j/



The images (A) and (B) correspond to the undifferentiated ultrasound pattern of /s/; while (C) corresponds to the undifferentiated ultrasound pattern of /j/. In all ultrasound images, the tip of the tongue is to the left.

Figure 2. Ultrasonographic pattern of the atypical productions of /s/ and /j/

Table 1. Percentage of responses by evaluators by category and average reaction time in the auditory-perceptual evaluation (audios of speech samples)

Auditory stimuli	Evaluation categories	Response categories	Responses from the evaluators	Reaction time (ms)
		(Audio)	N (%)	Mean (SD)
<i>Target /s/</i>	Categorical	Typical	157 (71.36)	2505.48 (±724.77)
		Incorrect	22 (10)	2741.40 (±714.85)
	Gradient	Gradient	38 (17.27)	3132.75 (±835.70)
	No answer	No answer	3 (1.36)	0
<i>Target /j/</i>	Categorical	Typical	32 (14.54)	2963.71 (±1046.89)
		Incorrect	124 (56.36)	2690.52 (±862)
	Gradient	Gradient	42 (19.09)	3119.78 (±768.94)
	No answer	No answer	22 (10)	0
<i>Total</i>	Categorical (Target /s/ and /j/)	Typical and incorrect production	335 (76.13)	2725.27 (±188.52)
	Gradient (Target /s/ and /j/)	Gradient production	80 (18.18)	3126.26 (±9.17)
	No answer	No answer	25 (5.45)	0

Wilcoxon signed-rank test

Caption: ms = milliseconds

Table 2. Percentage of responses by evaluators by category and the average reaction time in the visual-perceptual evaluation (ultrasound images)

Visual stimuli	Evaluation categories	Response categories	Responses from the evaluators	Reaction time (ms)
		(Image)	N (%)	Mean (SD)
<i>Target /s/</i>	Categorical	Image /s/	78 (35.45)	978.54(±524.48)
		Image /j/	55(25)	1052.38(±609.51)
	Gradient	Undifferentiated	85(38.63)	1011.39(±591.82)
	No answer	No answer	2(0.90)	0
<i>Target /j/</i>	Categorical	Image /s/	83(37.72)	1006.84(±626.02)
		Image /j/	84(38.18)	1021.31(±622.72)
	Gradient	Undifferentiated image	52(23.63)	1134.86(±615.07)
	No answer	No answer	1(0.45)	0
<i>Total</i>	Categorical	Image /s/ and Image /j/	300 (68.18)	1014.76(±30.72)
	Gradient	Undifferentiated image	137 (31.13)	1073.12(±87.30)
	No answer	No answer	3 (0.68)	0

Wilcoxon signed-rank test ms = milliseconds

In turn, Table 2 shows the results of the VPA analysis by ultrasound images, the average percentage of responses by the evaluators and the reaction time according to the categories adopted. The “categorical” evaluation category included the responses of image of /s/ and image of /j/; while the gradient category includes undifferentiated image responses.

Similarly to APA, the number of “no answers” was very small, as shown in the last line of Table 2, so they were not considered in the statistical analysis performed.

Table 3 shows comparatively the percentage of responses by the evaluators and the average reaction time in the audio and image evaluations.

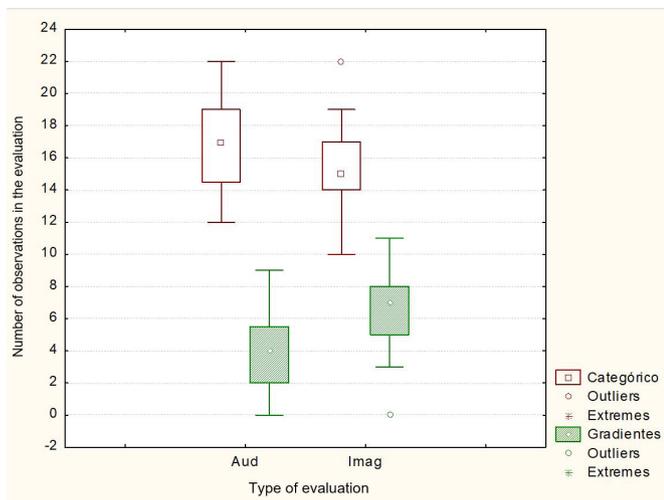
According to the Mann-Whitney U test, when comparing the percentage of responses by the evaluators and the reaction time between the evaluation of audios and images, there were significant differences (all $p < 0.05$) in all categories compared. In relation to the stimuli evaluated as categorical, the results of the comparison between the audio and image evaluations showed that the APA had a higher number of occurrences compared to

Table 3. Comparison of the percentage of responses from the evaluators and the average reaction time in the audio and image evaluations.

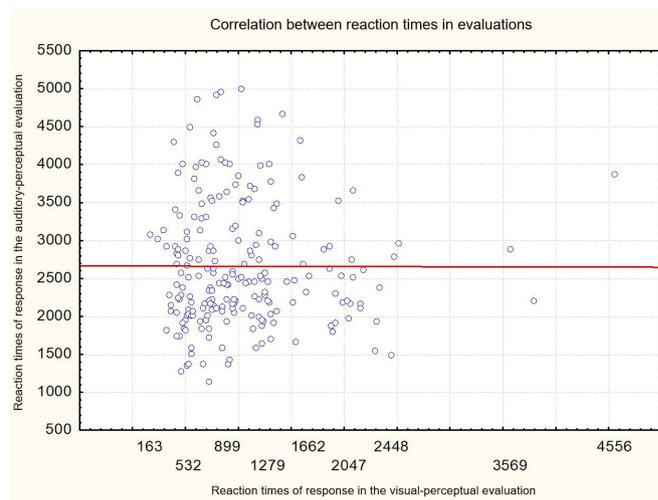
Evaluations	Evaluation categories	Percentage of judges' evaluations N (%)	Reaction time (ms)
			Mean (SD)
<i>Audio evaluation</i>	Categorical	335 (76.13%)	2725.27 (±188.52)
	Gradient	80 (18.18%)	3126.26 (±9.17)
<i>Image evaluation</i>	Categorical	300 (68.18%)	1014.76 (±30.72)
	Gradient	137 (31.13%)	1073.12 (±87.30)

Mann-Whitney U test ms = milliseconds

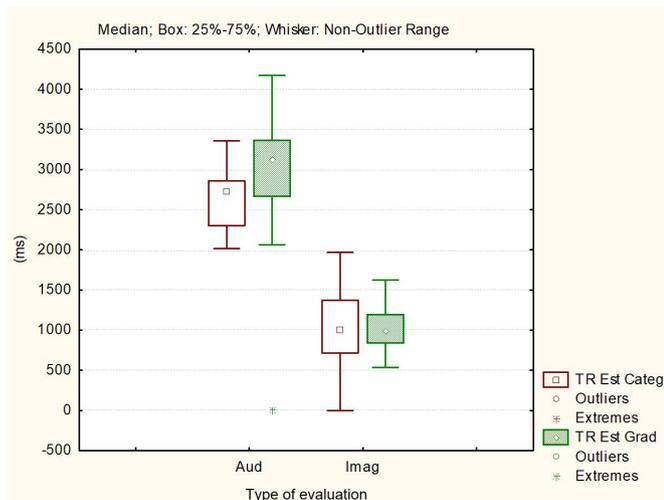
the VPA ($Z(20)=2.00$, $p=0.04$). On the other hand, the stimuli evaluated as gradients were identified more frequently by the VPA, visual-perceptual image evaluation ($Z(20)=2.00$, $p=0.04$),



Aud – Auditory Imag – Images Mann-Whitney U test
Figure 3. Comparison of the percentage of responses from the general judges between the types of judgments



Spearman's rank correlation coefficient
Figure 5. Correlation between the reaction time in the auditory-perceptual and visual-perceptual judgments



Aud – Auditory Imag – Images Mann-Whitney U test
Figure 4. Comparison of reaction time for each type of judgment

while the VPA provided greater identification of the stimuli evaluated as gradients ($Z(20)=-3.24, p<0.00$).

Figure 3 shows the difference in the percentage of responses by the evaluators between the types of evaluations.

Similarly, when comparing the average reaction time spent during the evaluation, the Mann-Whitney U test showed that the reaction time in the auditory evaluation was significantly higher than the reaction time in the image evaluation, both for the evaluation of the stimuli evaluated as categorical ($Z(2)=5.41, p<0.00$) and for the evaluation of the stimuli evaluated as gradients ($Z(20)=4.32, p<0.00$) (Figure 4).

Finally, when correlating the percentage of responses by the evaluators and the reaction time in both evaluation tasks, the Spearman's rank correlation coefficient did not show statistical significance for the percentage of responses ($t(N-2)=-1.03, p=0.29$) nor for the reaction time ($t(N-2)=0.36, p=0.71$), as shown in Figure 4, respectively.

Figure 5 shows a correlation between the reaction time in both evaluations.

DISCUSSION

Given the existence of categorical and gradients productions in speech production errors⁽¹⁻⁴⁾, this study aimed to analyze which method (APA or VPA) is more sensitive to detect gradual productions in the voiceless coronal fricative class, as well as to correlate the two forms of evaluations .

Although it was possible to detect gradient productions in both APA and VPA, in the first hypothesis, in the evaluation of ultrasound images, a higher percentage of responses and shorter reaction time were expected, and the results were fully in line with this hypothesis. Initially, the gradient productions were detected in the two evaluations, suggesting that these two perceptual evaluation methods are valid for detecting the presence of gradient productions, although the occurrence of the evaluators' responses varies between the methods.

Although previous studies that exclusively used APA also reported the presence of gradient productions, these studies found that the inclusion of intermediate categories in the evaluation is an essential condition to identify the gradient productions^(5-7,20,21). In this study, the inclusion of the gradient (intermediate) category favored the auditory identification of gradient productions in the voiceless coronal fricatives by the evaluators.

However, the authors of another study⁽⁵⁾ reported that auditory information as the only form of analysis may not be fully efficient in detecting a gradient sound. This means that the listeners' perception can be changed when an intermediate category is presented together with a categorical production. In addition, the authors suggest the use of instrumental information that facilitates the detection of gradient production during the APA.

Similarly, a study that used ultrasound imaging to identify gradient productions in the production of English liquids, showed the relevance of this method and warned about the importance of presenting the intermediate category to the evaluators, since, in general, the tasks of assessing speech samples involve only responses related to phonemic categories, for example, or the stimulus is evaluated as /s/ or as /ʃ/. In this same study,

the ultrasound image to identify gradient productions in the voiceless fricatives was also used with the presentation of the intermediate category to the evaluators, which may have favored the identification of this category in the images presented.

This study provides unprecedented information by explicitly comparing the results of both methods (auditory-perceptual and visual-perceptual) for the detection of gradient productions, particularly in the class of voiceless coronal fricatives. Although there is no previous research comparing both methods of detecting gradient productions (regardless of the phonic class), the studies available in the literature so far have shown that the evaluation of ultrasound images facilitates this task⁽¹²⁻¹⁸⁾. Throughout language development, speakers learn to perceive speech categorically, that is, linked to pre-existing language categories⁽²²⁾, which would lead listeners to categorize auditory stimuli categorically. On the other hand, the image provides a direct visualization of the articulatory movement that generated the stimulus, and can be a facilitator (since it does not depend on an inference from the auditory stimulus) for the detection of gradient production⁽²³⁻²⁵⁾, which would explain the higher percentage of responses and less reaction time of gradient productions from the analyzed ultrasound images (VPA).

In turn, the second hypothesis was not confirmed in the results obtained, as a positive correlation was expected between the APA and the VPA with regard to the percentage of responses by the evaluators and the reaction time.

These results did not corroborate the findings of a study^(26,27), according to which the qualitative measures of the movement of the tongue in the ultrasound image had a correlation with the perceptual and acoustic measures of the sounds /r/ in children with SSD and typical speech, but are in line with the results described in another study⁽²⁸⁾. In this latest study, the authors noticed a variability in retroflex production in British English, in which typical Asian speakers produce /l/ and /r/ with an anterior constriction, with the front or tip of the tongue, while Anglo speakers usually produce /l/ and /r/ with a more posterior constriction, with the back of the tongue retracted. Thus, the authors⁽²⁸⁾ also found that, despite the correlation between articulation, acoustics and hearing, individual articulatory patterns in British English are not always correlated with auditory and acoustic variations in the expected forms. Therefore, the authors concluded that speech production has a phonetic influence on the mother tongue, and each individual acquires a production variability according to their contact with the language.

Depending on the mother tongue, speech production may have variability and the articulation, acoustic and auditory patterns do not correlate, as shown in this study.

In addition, the tongue ultrasound (TUS) of this study proved to be an effective and facilitating technique in detecting the gradient production of voiceless coronal fricatives: it can add valuable criteria for assessment, diagnosis and speech-language pathology processes, as well as being an important tool for monitor results after the therapeutic intervention. The detection of gradient production by through the TUS shows the child's phonological knowledge, which can accelerate the therapeutic process, that is, if the child has a gradient production, they will already be halfway to reach the target, which means shorter

therapy time. In addition, it should be noted that the evaluation of TUS has best cost-effectiveness compared to other articulatory analyzes, which favors its use for clinical and research purposes.

Although this study has provided relevant contributions on the use of TUS as a tool for gradient analysis of speech production, further studies are required to conduct an auditory-perceptual and visual-perceptual investigation through ultrasound images, other phonic contrasts, in order to investigate if the gradient productions can be monitored by these measures in other classes of sounds of the language.

CONCLUSION

The study showed the ability of the evaluators to detect gradient productions among the class of voiceless coronal fricatives, both in APA and VPA. It was possible to notice a higher percentage of responses and a shorter reaction time for gradient stimuli in the VPA, thus showing that the use of ultrasound images is the most sensitive and facilitating method in order to detect gradient production in the production of voiceless coronal fricatives.

REFERENCES

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):111-24. <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 /ʃ/ 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 glossectomized. *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 Intern 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 /r/ 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, Fabbion 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.

Authors contributions

FLCNL: project design, sample collection and selection, data analysis and tabulation, and article writing; LCB: study design; data analysis; discussion; co-supervision and article writing; VCCM: data analysis and article writing; EMGF: project and study design, discussion and general guidance, development and article writing.