

Original articles

Ultrasonographic analysis of lateral liquids and coronal fricatives: judgment of experienced and non-experienced judges

Fernanda Leitão de Castro Nunes de Lima⁽¹⁾Cássio Eduardo Esperandino da Silva⁽¹⁾Lídia Maurício da Silva⁽¹⁾Aline Mara de Oliveira Vassoler⁽¹⁾Eliana Maria Gradim Fabbron⁽¹⁾Larissa Cristina Berti⁽¹⁾

⁽¹⁾ Faculdade de Filosofia e Ciências da Universidade Estadual Paulista - UNESP, Marília, São Paulo, Brasil.

Conflict of interests: Nonexistent



ABSTRACT

Purpose: to verify if the accuracy in the judgment of the ultrasound (US) images varies according to the experience of the judges and the sound class – alveolar liquids and coronal fricatives.

Methods: ultrasound images related to the production of liquids and fricatives in the intervocalic context of the vowel /a/, produced by 20 typical adults, were judged by 15 experienced judges and 15 non-experienced judges. A previous analysis of the images was performed to establish the typical ultrasound pattern of the liquids and fricatives. The accuracy (% of correctness and errors) of the judgments performed by the judges was considered in the analysis. Factorial ANOVA statistical test was used, considering as factors the sound class and the experience of the judges.

Results: the Factorial ANOVA showed a significant effect for the accuracy of the judgment only for the sound class, with no significance for the judges' experience or for the interaction between the experience of judges and the sound class. The liquids had a lower accuracy in the judgment, as compared to the fricatives, confirming their articulatory complexity, since they involve the production of two simultaneous gestures.

Conclusion: the accuracy in the judgment of the US images did not vary according to the experience of judges, but it was dependent on the sound class. It should be noted, however, that all judges presented previous knowledge about the processes of speech production which may have favored image interpretation.

Keywords: Speech; Phonetic; Ultrasonography

Received on: August 21, 2017

Accepted on: June 7, 2018

Corresponding author:

Fernanda Leitão de Castro Nunes de Lima
Rua José Bonifácio, 860, Apto 51 -
Bairro Vila Mendonça
CEP: 16015-050, Araçatuba, São Paulo,
Brasil
E-mail: fernandaleitao@live.com

INTRODUCTION

Ultrasonographic analysis, normally used to evaluate images of internal organs, has also been highly recommended for obtaining information about the real-time image of tongue movement during speech production¹⁻³.

In order to obtain the ultrasound image of tongue movement (hereinafter TUS), it is necessary to place a transducer in the submandibular region of the speaker. This transducer will emit high frequency waves, which will be propagated to the interior of the oral cavity. Due to the density difference between tongue/air and air/bone structures, the ultrasonic waves will be reflected and captured by the transducer. After the detection of these waves, the reflection point will be calculated resulting in the creation of an image⁴.

Some authors⁵⁻⁸ argue in favor of the use of ultrasonography of the tongue movement in clinical practice due to its affordable cost, when compared to the other equipment used in an articulatory analysis; portability of the equipment, not restricting the collection in laboratory situation; relative comfort provided to the individuals at the time of collection, since it is not an invasive instrument and nor requires previous preparation of them; beyond the visualization of the tongue surface- from its tip to the root - during the production of speech sounds.

However, it is known that the ultrasound image of the tongue surface (either in the coronal and/or sagittal plane) is not always clear, because the formation of the image depends on the difference in density between the structures of the vocal tract, varying from individual to individual. Additionally, in the ultrasound image of the tongue, the passive articulators are not visualized; many times, part of the tip-of-tongue information is lost; in addition to the difficulty in reducing the head movement during data collection, which presupposes the need for previous training of clinicians and/or researchers to interpret the images.

Another challenge for the TUS applicability would be the need for clinicians to present previous knowledge about the ultrasonographic pattern of tongue movement for the production of different phonemes. In the literature, there are authors that described different speech sounds⁹, while other authors have proposed to study classes of isolated sounds, such as occlusives¹⁰, fricatives¹¹, liquids^{5,12-14} and vowels^{15,16}.

Despite the difficulties inherent in the use of the TUS, the studies do not explicitly recommend the need of a previous training for researchers and/or clinicians

for this type of analysis^{2,7} nor they highlight whether there was a sound class that implies a greater difficulty of interpretation of the ultrasound image.

For Brazilian Portuguese (hereinafter BP), in particular, there is already a qualitative TUS description for the typical production in adults of the coronal liquids⁵ and the coronal fricatives¹⁷. These sounds were privileged by the authors because they presented late acquisition and were commonly involved in substitution processes in atypical speech.

The ultrasound contour of the tongue of /j/ was characterized by the raised tongue with presence of curvature of the dorsum to generate palatal turbulence; while for /s/ four distinct patterns were described: 1) absence of sharp curvature of the dorsum or root of tongue; 2) presence of sharp curvature of root and dorsum in descending direction; 3) presence of a sharp curvature of root and dorsum in an upward direction, and 4) the presence of sharp curvature and more anterior dorsum¹⁷.

Differently, the ultrasound pattern of the liquids /l/ and /r/ is characterized by the presence of double tongue gestures: one related to the tip of the tongue and the other to the dorsum of the tongue, making more complex sounds from an articulatory point of view. However, the degree of constriction of the tip gesture of the tongue is greater in the /r/ when compared to /l/, while the gesture of dorsum of the tongue towards the pharynx is more evident in /l/⁵.

Therefore, it can be deduced from the ultrasound descriptions made above that the articulatory complexity varies according to the sound class and it can be an important factor in the image analysis.

Considering the inherent difficulties in the use of TUS in the clinical context allied to the different articulatory complexity of the sound classes, the hypotheses of the study are defined as follows:

(1) Experienced judges in TUS analysis would present better accuracy in the image judgement than non-experienced judges.

(2) TUS images related to coronal liquids (/l/ and /r/) would produce greater difficulty in the judgment than images related to coronal fricatives.

The purpose of the present study is to verify if the accuracy in the judgment of the ultrasound images varies according to the experience of the judges and the sound class - alveolar liquids and coronal fricatives.

METHODS

The present study was approved by the Research Ethics Committee of the Faculdade de Filosofia e Ciências da Universidade Estadual Paulista - UNESP of Marília (Faculty of Philosophy and Sciences of the São Paulo State University - UNESP of Marília), nº1.268.673/2015. All individuals included in the research were informed and signed the informed consent form (ICF) presented to them.

Participants

Thirty judges, divided into two groups, participated in the present study: 15 judges who were experienced in ultrasonographic analysis of tongue movement (G1) and 15 judges who had no previous experience with this tool of speech analysis (G2).

The judges were recruited in the Undergraduate and Post-Graduate courses of Speech-Language Pathology of the home institution. For the composition of the G1 (experienced judges) it was required previous knowledge on ultrasound analysis of tongue movement in the speech production, as a short duration course or technical training in ultrasound analysis of tongue movement. Because there were not many clinicians with previous experience in TUS analysis, the time of experience with the use of this technique was not controlled. For participants of G2, the inclusion criterion was the previous knowledge about the speech production process, besides the phonetic classification and description of the different phonemes of Brazilian Portuguese. All judges should necessarily have completed a course in Phonetic and Phonology of Brazilian Portuguese with an approved performance. The exclusion criterion was the lack of the necessary knowledge for the analyzes that would be performed. Figure 1 presents the characterization of the judges (G1 and G2).

Procedure

Stimulus

Ultrasound images related to the production of logatome involving the coronal liquids (/l/ - [a'la] and /r/ - [a'ra]) and coronal fricatives (/s/ -- [a'sa] e /f/ - [a'fa]), in the intervocalic context of the vowel /a/, of 20 adults, Brazilian Portuguese speakers, monolingual, with typical speech production in the 20-30 age (10 men and 10 women). For each production, using Sound Forge Studio 6.0 software, it was selected frames corresponding to the maximum point of tongue constriction

of the phoneme production, static ultrasound images, computing a total of 80 frames: 20 frames relative to the production of /l/, 20 relative to the production of /r/, 20 frames corresponding to the production of /s/ and 20 frames corresponding to the production of /f/.

The data that constitute this database were collected using a portable ultrasound, model DP 6600, containing transducer coupled to a computer, unidirectional microphone and head stabilizer. The acoustic signals and image were recorded simultaneously with the use of the AAA (Articulate Assistant Advanced) software combined with a synchronizer that allows synchronization between the images and the acoustic signal. The US images were acquired with a frequency of 6.5 MHz, 120 ° of image field and 29.97 Hz of sampling rate.

Encoding of images

A previous analysis of the ultrasound images was made by two experienced examiners, in order to verify not only the correspondence with the ultrasound pattern described as typical of the coronal liquid and fricatives, but also the need for adjustments in this description.

After the inspection of the images and inspiring on the existing description in the literature^{5,17}, the categories for the assessment of the TUS images described in figure 2 were defined.

Once the categories were defined, the two experienced examiners independently judged all ultrasound images to obtain an accuracy of 100% and 87.5%, respectively. Kappa was calculated by obtaining a degree of agreement of 1.00 ($p < 0.05$).

Judgment of the images

The judgment task was elaborated so that the judges could analyze the pair of static ultrasound images of the same speaker and not only a single image, in order to minimize differences in the structures of the vocal tract between the individuals, as well as the positioning of the transducer in the capture of the images.

Therefore, the ultrasound images related to the maximum point of tongue constriction in the production of liquids /l/ and /r/, on the one hand, and the fricatives /s/ and /f/, on the other hand, were paired and organized in two different PowerPoint files for the judgment, as shown in figures 3 and 4.

Before the judgment of the images itself, a description of the expected ultrasound pattern for the coronal liquids and fricatives was provided in a

Judges	Background	Gender	Experience in TUS analysis
J1	PhD	Female	Yes
J2	PhD	Female	Yes
J3	Undergraduate student	Female	Yes
J4	PhD	Female	Yes
J5	Undergraduate student	Male	Yes
J6	Undergraduate student	Female	Yes
J7	Master's student	Female	Yes
J8	Post-Doctor	Female	Yes
J9	Post-Doctor	Male	Yes
J10	Undergraduate student	Female	Yes
J11	Master's	Female	Yes
J12	PhD student	Female	Yes
J13	Master's	Female	Yes
J14	Master's	Female	Yes
J15	PhD	Female	Yes
J16	Undergraduate student	Female	No
J17	Master's	Female	No
J18	Undergraduate student	Female	No
J19	Master's	Female	No
J20	Undergraduate student	Female	No
J21	Undergraduate student	Female	No
J22	Undergraduate student	Female	No
J23	Undergraduate student	Female	No
J24	Undergraduate student	Female	No
J25	Undergraduate student	Female	No
J26	Undergraduate student	Female	No
J27	Undergraduate student	Female	No
J28	Undergraduate student	Female	No
J29	Undergraduate student	Female	No
J30	Undergraduate student	Female	No

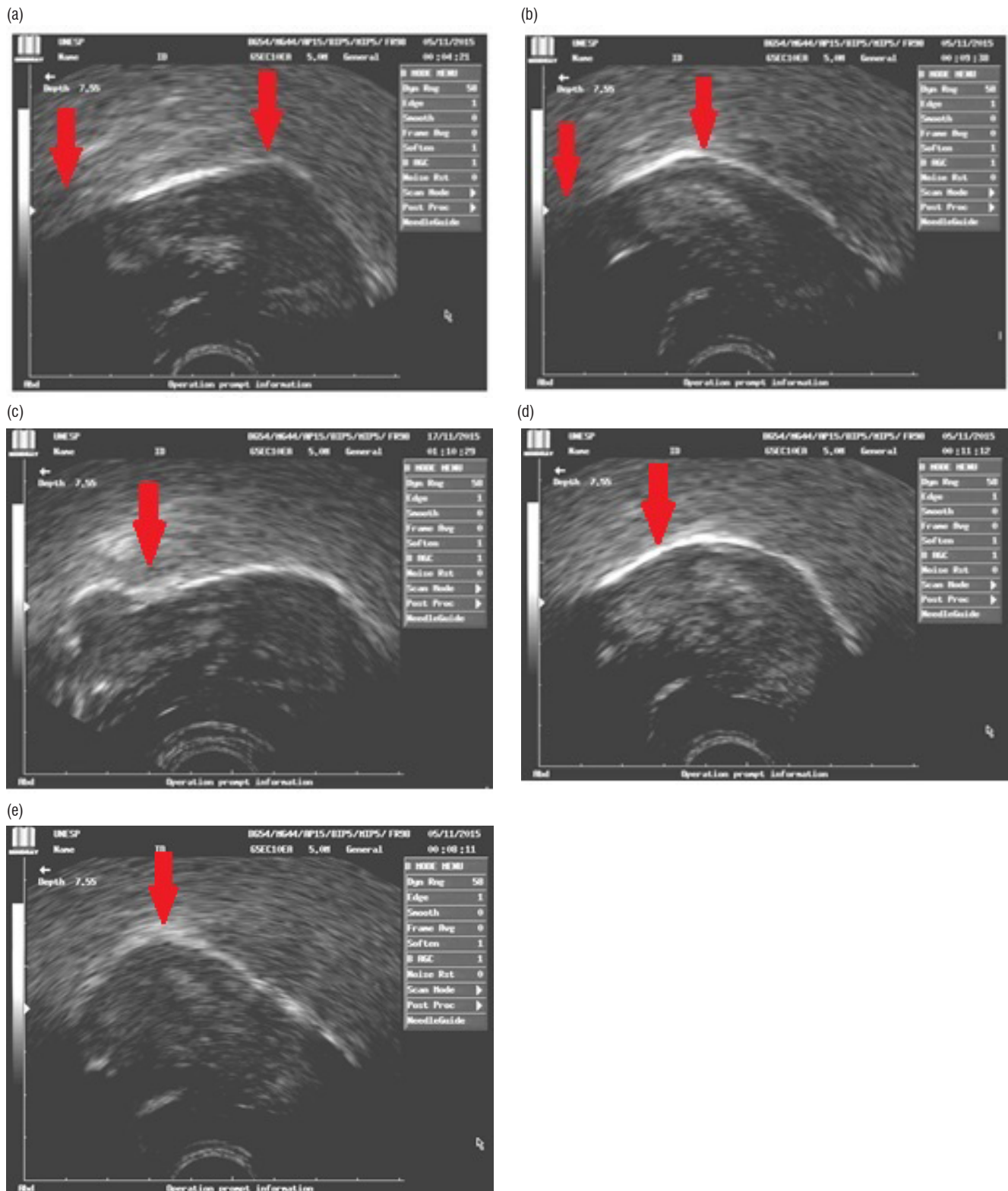
TUS = Tongue Ultrasound Imaging

Figure 1. Characterization of experienced and non-experienced judges

PowerPoint program file, since not only important information was presented for the interpretation of the images (for example, where different parts of the tongue are localized), but also examples of the ultrasound pattern of each one of the possibilities of analysis.

The judges could return to the description and initial example as many times as they thought necessary. In both judgments, that is, of the liquids and fricatives separately, the task did not exceed 10 minutes.

Finally, a statistical and inferential analysis of the data was made considering the accuracy of the judges in the judgment in terms of percentage of correctness. A Factorial ANOVA was used in the analysis to verify if the accuracy in the judgment of the images varied according to the previous experience of the judges and the sound class.



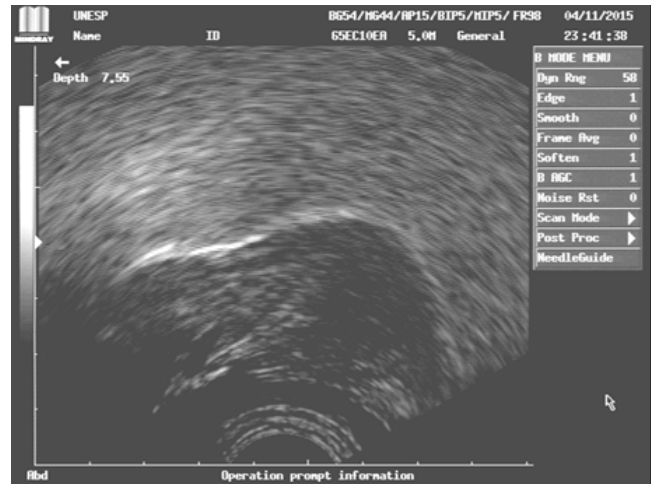
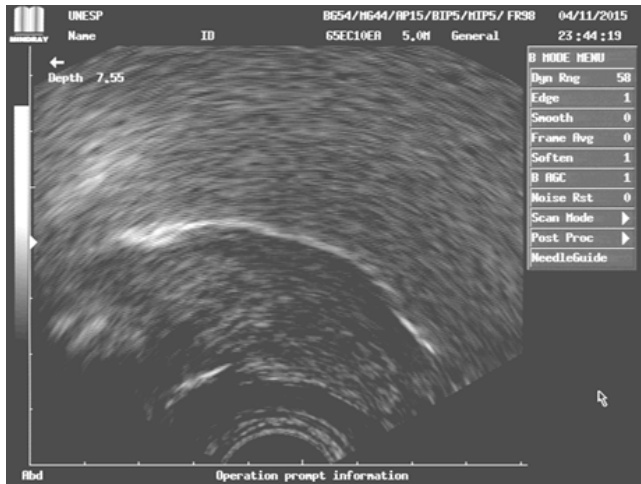
The image (a) represents the production of /l/ and the arrows indicate the presence of two simultaneous gestures, the tip and the dorsum of the tongue (with dorsum gesture of the posteriorized tongue, towards the pharynx); the image (b) illustrates the production of /ɺ/ and the arrows show the presence of two simultaneous gestures of the tip and the dorsum of the tongue (in this figure, the dorsum gesture of the tongue is found less posteriorized); the image (c) refers to the production of /s/ and shows the tip and blade of the anteriorized tongue (the arrow shows the presence of groove); the image (d) refers to the production of /s/ and shows the tip and blade of the anteriorized tongue (the arrow shows the absence of groove); Finally, the image (e) shows the production of /ʃ/ and the arrow identifies the raised tongue in concave shape.

Source: Research Data

Figure 2. Example of the ultrasound pattern of the target productions of the liquids /l/ and /ɺ/ and fricatives /s/ and /ʃ/

Anterior dorsum () or Posterior dorsum ()

Anterior dorsum () or Posterior dorsum ()

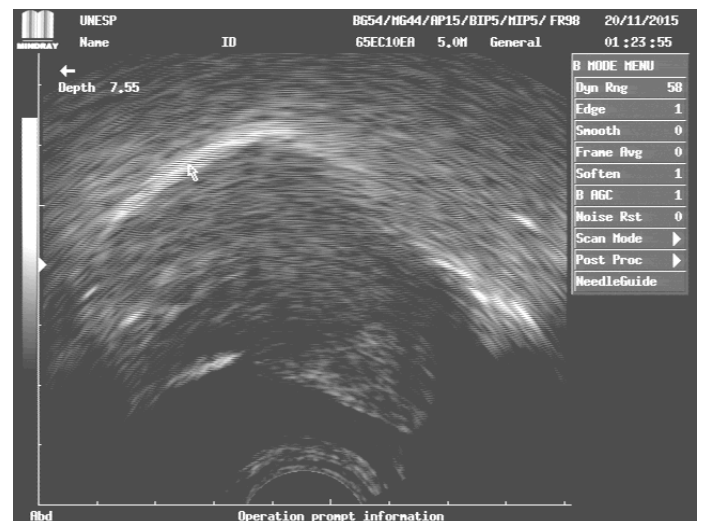
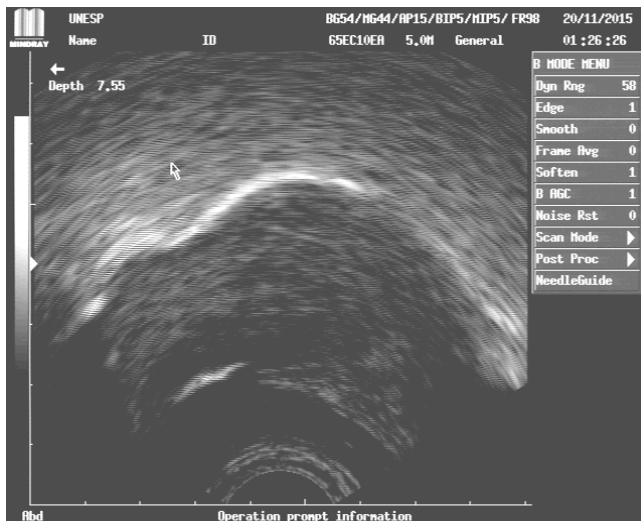


Source: Research Data

Figure 3. Example of image showing the maximum point of constriction of the tongue to the production of liquids /l/ and /r/ presented to the judges for the judgment task.

/s/ () or /ʃ/ ()

/s/ () or /ʃ/ ()



Source: Research Data

Figure 4. Example of image showing the maximum point of constriction of the tongue to the production of the fricatives /s/ and /ʃ/ presented to the judges for the judgment task.

RESULTS

The accuracy in the images judgment of the liquids and fricatives in function of the experience of the judges can be seen in the Table 1.

The Factorial ANOVA (2 factors) was used in the analysis, considering the experience of the judges (experienced vs. non-experienced) and the sound class (liquids and fricatives) as factors that could influence the accuracy of the judgment (Table 2).

Table 1. Mean and standard deviation of the percentage of correctness in the judgment of the ultrasound images

Type of Judges	% of correctness	
	Liquids	Fricatives
Experienced	89.05 (± 7.58)	96.13 (± 6.02)
Non-experienced	86.08 (± 10.12)	92.97 (± 10.32)

Table 2. Factorial ANOVA values regarding the main effects of the sound class and experience of the judges, as well as the interaction between sound class and experience of the judges

Main effects and interactions between effects		Results
Sound class	F	F(1,56)=9.66
	P	0.00
Experience of the judges	F	F(1,56)=1.82
	P	0.17
Interaction between experience of the judges and sound class	F	F(1,56)=0.002
	P	0.96

P = Probability value

F = Ratio between the model and its error in the analysis of variance

A significant main effect was verified only for the accuracy of the judgment in function of the sound class (F (1,56)=9,66, $p > 0.00$), with no significance for the experience of the judges (F (1, 56)=1.82, $p > 0.17$) nor for the interaction between experience of the judges and sound class (F (1,56)=0.002, $p > 0.96$). It is observed that although, numerically, the experienced judges present a greater accuracy than the non-experienced

ones, this difference was not statistically significant. Differently, the ultrasound images corresponding to the fricative class had a greater percentage of correctness in the judgments, compared to images related to liquids. Figure 5 illustrates the results found.

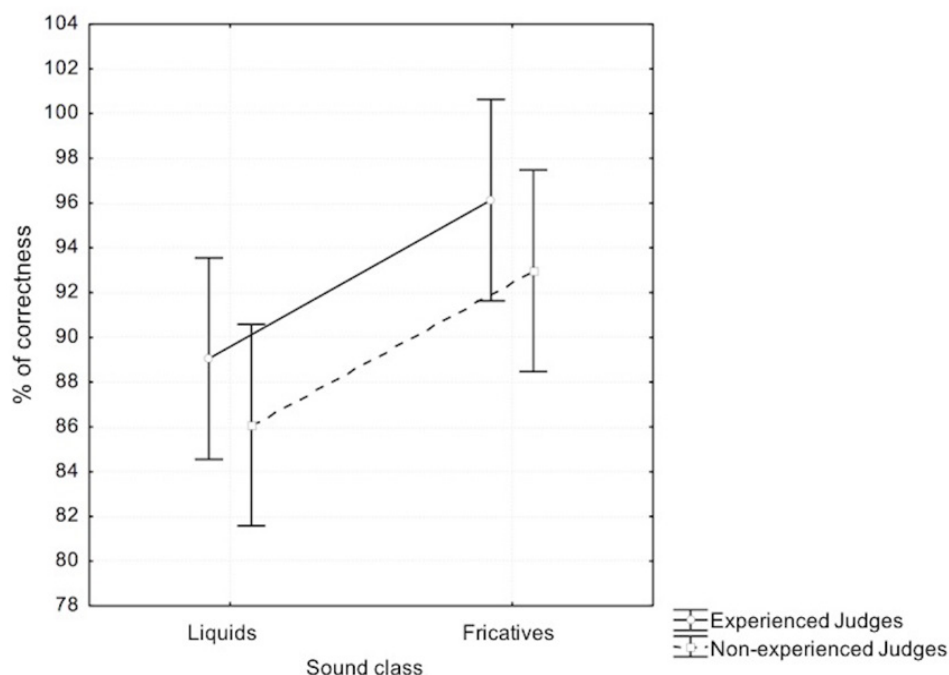


Figure 5. Accuracy of the judgment based on the experience of the judges and sound class

DISCUSSION

The present study aimed to verify if the accuracy in the judgment of the ultrasound images varies in function of the experience of the judges and the sound class - coronal liquids and fricatives.

Considering the difficulties inherent in the use of TUS in the clinical context combined with the different articulatory complexity of the sound classes, it was expected that (1) experienced judges in TUS analysis would present better accuracy in the image judgment than non-experienced judges and (2) the TUS images relative to the coronal liquids (/s/ e /r/) would imply greater difficulty in judgment than images relative to the coronal fricatives (/s/ e /f/).

The first hypothesis related to the experience of the judges was not confirmed as far as there was no significant statistical difference in the percentage of correctness in the judgments of both groups of judges: experienced and non-experienced on the ultrasound analysis. It was verified that in both groups of judges the accuracy in the judgment was greater than 85%, suggesting an excellent performance by the judges. All judges (experienced and non-experienced) came from the same institution. Although they did not have specific training in the visualization of the US images, the non-experienced judges had a proven phonetic background and quite detailed regarding the production of speech sounds. In addition, they

have received previous explanations regarding the US pattern. This training of the non-experienced judges may have been a differential that resulted in the results found. The widespread use of US must be carefully estimated, especially from more research and with different levels of judges.

It is interesting to note that there is no explicit description in the literature of the need for previous training by researchers and/or clinicians to enable the interpretation of ultrasound images of the tongue. In the two classic studies of the area^{1,2}, the authors highlight the TUS as an effective instrument in the analysis of speech production, compared to other articulation techniques (magnetic resonance (MRI), X-rays, EMA, micros X-rays), since these, in addition to high cost of the equipment, the handling and portability are also limiting aspects for clinical use. In addition, the TUS allows to capture the contour of the tongue in a dynamic way and real time, allowing the study of the movement of dorsum to the tip of the tongue during the production of lingual consonants and vowels, being an allied tool in the diagnostic process of speech disorders, as well as in the therapeutic process (use of biofeedback in the therapeutic context).

However, it should be emphasized that both groups of judges in the present study had previous knowledge not only about the speech production process, but also about the articulatory characteristics of Brazilian

Portuguese phonemes, a fact that may justify similar performance between them. In this way, what allows the correct interpretation by the part of the judges is the previous training on aspects of speech production and not the specific one, on the technique of speech production analysis, in our case, the ultrasonography of the tongue movement. It is assumed, however, that training in the Phonetics area, such as expected in the curricular guidelines of the Speech-Language Pathology and Audiology courses (National Council of Education, 2002), provided the judges with not only a familiarity in identifying the main active articulator of the speech production - the tongue - in the sagittal plane in the TUS images, but also recognition of the different contours of the tongue surface corresponding to the production of the coronal liquids and fricatives.

In relation to the judgment of the two sound classes, it was hypothesized that the liquids class would impose a greater difficulty in judgment than the fricatives one. This hypothesis was corroborated to the extent that there was a significant difference of accuracy as a function of the sound class. The liquids presented a lower accuracy of judgment when compared to the fricatives.

An explanatory possibility for this accuracy difference lies in the very articulatory complexity present in the liquids ones.

Previous studies^{18,19} that have used TUS images to describe coronal liquids /l/ and /r/ characterized these sounds as being the most complex as far as they involve the production of two simultaneous articulatory gestures: one relative to the tip of the tongue and another relative to the dorsum of the tongue. The difference of images between the liquids was detected by overlapping the curves of /l/ and /r/. The result of this overlap showed that in /l/, the dorsum gesture of the tongue towards the pharynx is more evident in /l/ than in /r/, while the tip gesture of the tongue is greater in /r/ when compared to /l/⁵.

Another argument in favor of liquids complexity is that, even though there is variability in the production of /s/, as described in a previous study¹⁷ this factor did not contribute to the decrease of the accuracy of this class in comparison to liquids ones.

In addition, because the tip of the tongue is not always clear in TUS images, the judges were requested to observe mainly the direction of the dorsal gesture in the judgment of the liquids. Perhaps, this fact may have contributed to explain the lower accuracy of the liquids in relation to the fricatives, since there is usually

no articulatory description emphasizing the movement of the posterior part of the tongue. It is observed in the phonetic description of /l/ and /r/ of BP that the difference in production between these sounds lies in the duration parameter and in the ballistic movement of the tip of the tongue necessary to the production of /r/²⁰⁻²².

Based on the limitations of the present study, regarding the non-standardization of the control of the experience time of the first group of judges, besides the non-control of the response time during the judgment task, the feasibility of the use of TUS in clinical practice as a complementary tool for analysis of speech production was confirmed.

CONCLUSION

The study showed that experienced and non-experienced judges did not present difference in terms of accuracy in the judgment of ultrasound images. However, both groups had proven training in phonetics, received previous instructions for the judgement and were originated from the same institution.

The liquids class presented a lower accuracy, in their judgment, as compared to the fricatives one.

TUS can be carefully disseminated in the clinical practice of the speech-language therapist as a therapeutic and diagnostic support resource, however, one should have theoretical knowledge of the phonemes production and knowledge about the TUS pattern in function of sound class.

REFERENCES

1. Gick B. The use of ultrasound for linguistic phonetic fieldwork. *J Intern Phon Assoc.* 2002;32(2):113-22.
2. Stone M. A guide to analyzing tongue motion from ultrasound Images. *Clin Linguist Phon.* 2005;19(6/7):455-502.
3. Scobbie J. Ultrasound-based tongue root imaging and measurement. Oral paper at the Workshop on Pharyngeals and Pharyngealisations, Newcastle Upon Tyne, March 2009.
4. Gonçalves GF, Paula MRB. *Dinâmica dos movimentos articulatorios: sons, gestos e imagens.* Pelotas: editora UFPel, 2013.
5. Barbarena LS, Soares MK, Berti LC. Description of the articulatory gestures concerned in the production of the sounds /r/ and /l/. *Audiol Commun Res.* 2014;19(4):338-44.

6. Preston JL, McCabe P, Rivera-Campos A, Whittle JL, Landry E, Maas E. Ultrasound visual feedback treatment and practice variability for residual speech sound errors. *J Speech Lang Hear Res.* 2014;57(6):2102-15.
7. Bressmann T, Flowers H, Wong W, Irish JC. Coronal view ultrasound imaging of movement in different segments of the tongue during paced recital: Findings from four normal speakers and a speaker with partial glossectomy. *Clin Linguist Phon.* 2010;24(8):589-601.
8. Zharkova N. A normative-speaker validation study of two indices developed to quantify tongue dorsum activity from midsagittal tongue shapes. *Clin Linguist Phon.* 2013;27(6-7):484-96.
9. Stolar S, Gick B. An index for quantifying tongue curvature. *Can Acoust.* 2013;41(1):11-6.
10. Houde RA. A study of tongue body motion during selected speech sounds. *Speech Commun Res Lab, No. 2 Santa Barbara, 1968.*
11. Wertzner HF, Francisco DT, Neves LOP. Tongue contour for /s/ and /f/ in children with speech sound disorder. *CoDAS.* 2014;26(3):248-51.
12. Campbell F, Gick B, Wilson I, Batesons EV. Spatial and temporal properties of gestures in north american english /r/. *Lang Speech.* 2010;53(1):49-69.
13. Lawson E, Scobbie JM, Smith JS. The social stratification of tongue shape for postvocalic /r/ in Scottish English. *J Sociolinguistics.* 2011;15(2):256-68.
14. Vassoler AMO, Berti LC. Coordenação gestual na produção de encontros consonantais em crianças com desenvolvimento de linguagem típico e atípico [Tese]. São José do Rio Preto (SP): Universidade Estadual Paulista; 2015.
15. Stone M, Goldstein MH, Zhang Y. Principal component analysis of cross sections of tongue shapes in vowel production. *Speech Commun.* 1997;22(2-3):173-84.
16. Menard L, Aubin J, Thibeault M, Richard G. Measuring tongue shapes and positions with ultrasound imaging: a validation experiment using an articulatory model. *Folia Phoniatr Logop.* 2012;64(2):64-72.
17. 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.
18. Gick B, Bacsfalvi P, Bernhardt BM, Oh S, Stolar S, Wilson I. A motor differentiation model for liquid substitutions: english /r/ variants in normal and disordered acquisition. *Proc Meet Acoust.* 2007;1(1):060003.
19. Proctor M. Towards a gestural characterization of liquids: evidence from Spanish and Russian. *Lab Phonology.* 2011;2(2):451-85.
20. Cagliari LC. Elementos de fonética do português brasileiro [Tese]. Campinas (SP): Universidade Estadual de Campinas; 1981.
21. Adler-bock M, Bernhardt BM, Gick B, Bacsfalvi P. The use of ultrasound in remediation of North American English /r/ in 2 adolescents. *Am J Speech Lang Pathol.* 2007;16:128-39.
22. Pagan LO, Wertzner HF. Análise acústica das consoantes líquidas do Português Brasileiro em crianças com e sem transtorno fonológico. *Rev Soc Bras Fonoaudiol.* 2007;12(2):106-13.