



PARECER TÉCNICO IBP18072

REVISTA FÓRUM

PROJETO COMPROVA

CONSULENTES



SÃO PAULO, 21 DE SETEMBRO DE 2018



SUMÁRIO

1. OBJETIVO	3
2. METODOLOGIA	3
3. PEÇAS SUBMETIDAS A EXAMES.....	3
4. ANÁLISE PERCEPTIVO-AUDITIVA	5
5. ANÁLISE ACÚSTICA.....	6
5.1 Frequência fundamental.....	6
5.2 Formantes	8
5.3 Long Term Average Spectrum	10
5.4 VOT.....	11
5.5 Fonema [j].....	13
6. CONCLUSÕES	16





1. OBJETIVO

Atendendo à solicitação da Revista Fórum e do Projeto Comprova, analisar segundo metodologia pericial forense do áudio contido em arquivo de vídeo hospedado no “YouTube”, visando determinar se a voz presente no áudio pertence ou não ao candidato à presidência Sr. Jair Bolsonaro.

2. METODOLOGIA

As melhores práticas periciais recomendam que análises dessa natureza tenham como base os métodos adotados pelas Ciências Forenses que são próprios da área da Fonética e Fonoaudiologia.

Na análise de comparação de falantes são confrontadas amostras de voz e fala, com o objetivo de determinar se estas foram ou não produzidas pela mesma pessoa. Para realizar a comparação da voz presente no áudio questionado serão utilizados as análises percetivo-auditiva e acústica da fala.

3. PEÇAS SUBMETIDAS A EXAMES

Os consulentes solicitaram que fosse analisado o vídeo hospedado no site “YouTube” denominado “#BOLSONARO SURTA NO HOSPITAL #ELEIÇÕES 2018 A CASA CAIU!!!!” e disponível na URL “<https://www.youtube.com/watch?v=rxyuUoj9X-Y>”.

Como padrão de confronto (paradigma) foi utilizado o vídeo hospedado no site “YouTube” denominado “Datena entrevista Jair Bolsonaro” e disponível na URL “<https://www.youtube.com/watch?v=e1gSZl7LJ-k>”.



Os vídeos de interesse foram coletados do “YouTube”, de forma que características e o código *hash*¹ dos arquivos gerados estão sintetizadas no quadro a seguir, visando o registro da cadeia de custódia:

Nome do arquivo	Hash SHA-1	Tamanho	Duração
#BOLSONARO SURTA NO HOSPITAL #ELEIÇÕES 2018 A CASA CAIU!!!!.mp4	EF00AA74A5ECF6670CA4EA451FDD69B26B26A589	1,60 MB	00:01:03
Datena entrevista Jair Bolsonaro.mp4	B0BCB8A93FF2E82E2E558E23996668AB781E4A35	108 MB	00:14:10

Após a identificação das peças a serem examinadas, foi realizada análise segundo metodologia pericial forense do conteúdo dos arquivos visando averiguar se a voz presente no áudio questionado pertence ou não à mesma pessoa.

Para fins de análise, foram adotados os seguintes padrões:

- Vídeo “#BOLSONARO SURTA NO HOSPITAL #ELEIÇÕES 2018 A CASA CAIU!!!!”: **Amostra Questionada.**
- Vídeo “Datena entrevista Jair Bolsonaro”: **Amostra Padrão.**

¹ O código hash é gerado por um algoritmo de dispersão que transforma o conteúdo de um determinado conjunto de dados, não importando seu tamanho, em uma sequência hexadecimal de tamanho fixo, de forma que qualquer alteração no arquivo implica também na alteração do seu código hash, possibilitando verificar a qualquer tempo se a integridade daquele arquivo ou disco está preservada.



4. ANÁLISE PERCEPTIVO-AUDITIVA

A análise perceptivo-auditiva constitui-se de uma análise qualitativa dos parâmetros (marcadores de voz, fala e linguagem) que caracterizam o perfil de comunicação dos falantes pelas peculiaridades como são realizados. Posteriormente, é realizada a comparação desses parâmetros em cada amostra de voz e fala questionada.

Neste exame foram observados os seguintes aspectos:

- **Segmentais:** padrões das vogais e consoantes, suas qualidades, aplicações e articulações (presença ou ausência de desvio fonético e/ou fonológico).
- **Suprasegmentais:** ritmo de fala e padrões de entonação.
- **Fatores paralinguísticos:** qualidade de voz, velocidade de fala e hábitos apresentados pelo falante.
- **Dialetais:** aplicação de palavras nas sentenças, uso de regras gramaticais e prosódia.

As análises foram descritas no quadro a seguir comparando a amostra questionada e a amostra padrão:

Parâmetros	Amostra Questionada	Amostra Padrão
Caracterização provável do indivíduo	Adulto, sexo masculino	Adulto, sexo masculino
Qualidade de voz	Modal	Modal
<i>Pitch</i>	Adequado para o sexo	Adequado para o sexo
Ritmo de fala e entonação	Adequada	Adequada



Ressonância	Baixa	Adequada
Articulação	Alterada	Adequada
Velocidade de fala	Aumentada	Adequada
Respiração	Coordenação pneumo-fonoarticulatória normal	Coordenação pneumo-fonoarticulatória normal
Dialeto	Presença do fonema /r/ retroflexo, típico do interior	-
Desvio fonético e/ou fonológico	Distorção acentuada na pronúncia dos fonemas [ʀ] e [s]	Distorção leve na pronúncia do fonema [s]

5. ANÁLISE ACÚSTICA

A análise acústica constitui-se de uma análise quantitativa dos parâmetros acústicos da voz (frequência fundamental, formantes, tempo de ataque de vozeamento, entre outros) que possibilitam maior compreensão do sinal vocal, associações à análise perceptivo-auditiva, documentação visual dos parâmetros da voz e da fala, e assimilação e identificação de semelhanças e diferenças de cada som tanto inter quanto intra-falantes. Posteriormente, é realizada a comparação desses parâmetros em cada amostra de voz e fala questionada.

5.1 FREQUÊNCIA FUNDAMENTAL

A frequência fundamental (f_0) é determinada pelo número de ciclos de vibração que as pregas vocais realizam por segundo, ou seja, é o correlato acústico da frequência de vibração das pregas vocais na produção de voz. Os



valores de f_0 dependem de características fisiológicas, como tamanho, massa, densidade e tensão das pregas vocais, assim, os valores da frequência fundamental mudam de pessoa para pessoa. Esse valor é percebido pelo indivíduo como o pitch (altura tonal) da voz.

Após análises da f_0 , pode-se observar diferenças significativas entre as amostras. A Figura 1 demonstra o contorno do pitch em determinado período de tempo do áudio, sendo que a linha em preto é referente a amostra questionada e a linha em vermelho referente a amostra padrão. A Figura 2 mostra a estatística descritiva da f_0 .

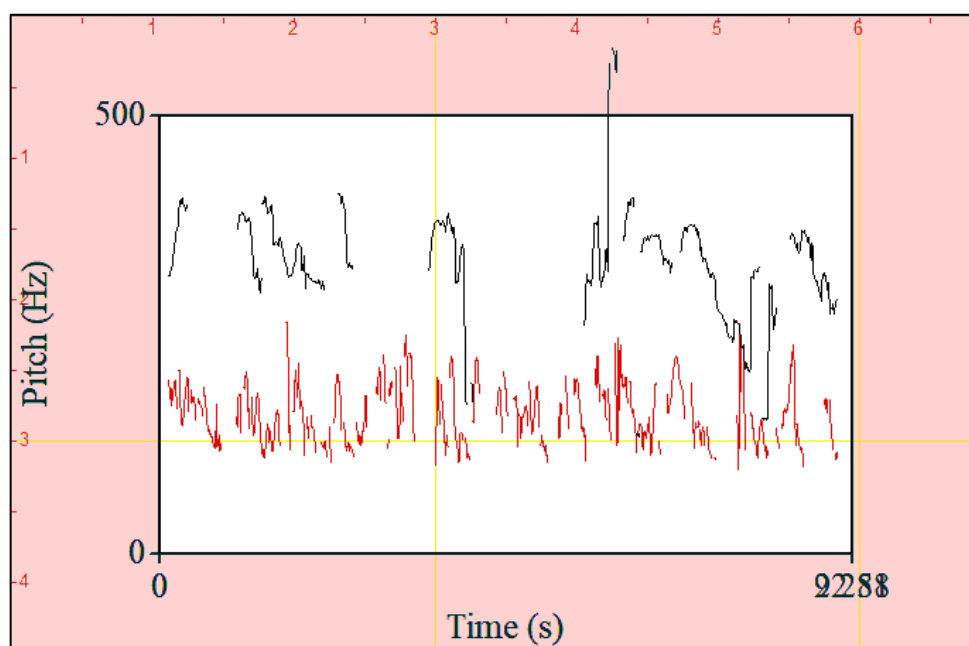


Figura 1: Análise do pitch para cada amostra.

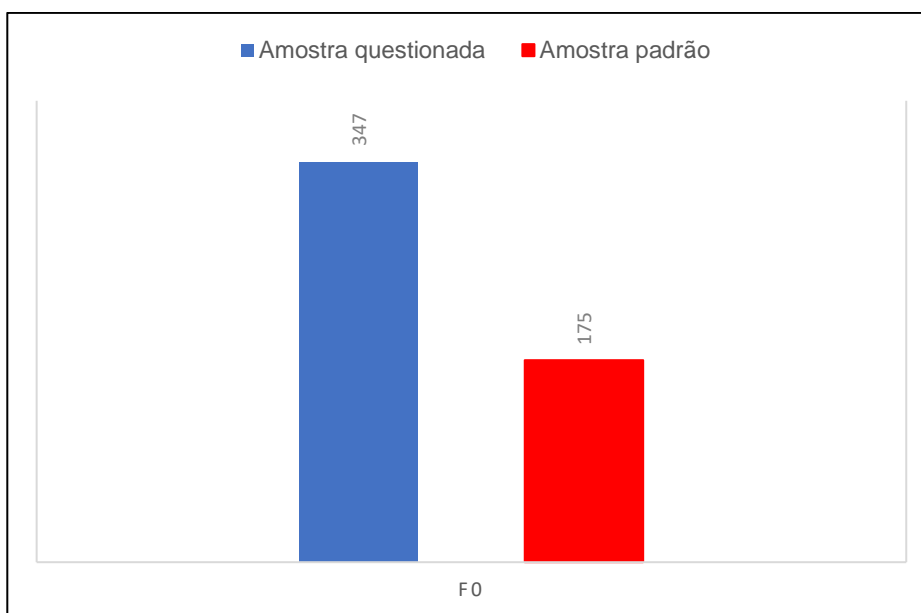


Figura 2: Estatística descritiva de f0 de cada amostra

Um dos fatores que podem influenciar na variação desse parâmetro, como observado nas Figura 1 Figura 2, é a tentativa de imitação da voz, pois quando à a tentativa de imitar a voz de outra pessoa, à a tendência de aumentar ou diminuir a frequência fundamental para, assim, tentar alcançar o pitch da outra pessoa.

5.2 FORMANTES

Os formantes apresentam informações importantes sobre a ressonância do trato vocal, revelando a fisiologia e o modo como o falante articula o som. Ele é determinado pelo comprimento e forma do trato vocal.

Embora existam infinitos formantes, os dois primeiros são os mais importantes, pois o F1 corresponde ao movimento de abertura da mandíbula e o F2 corresponde ao movimento ântero-posterior do dorso da



língua. Dessa forma, este parâmetro pode fornecer dados importantes sobre os falantes, tanto em termos numéricos quanto visuais.

A partir da combinação das frequências de F1 em ordenada e de F2 em abcissa foi construída a carta de formantes para cada amostra, medida em Hz. Para a análise dos formantes foram selecionadas as vogais “a”, “e” e “o”.

A Figura 3 mostra a carta de formantes para cada amostra, sendo as vogais em verde corresponde a amostra questionada e as vogais em azul corresponde a amostra padrão.

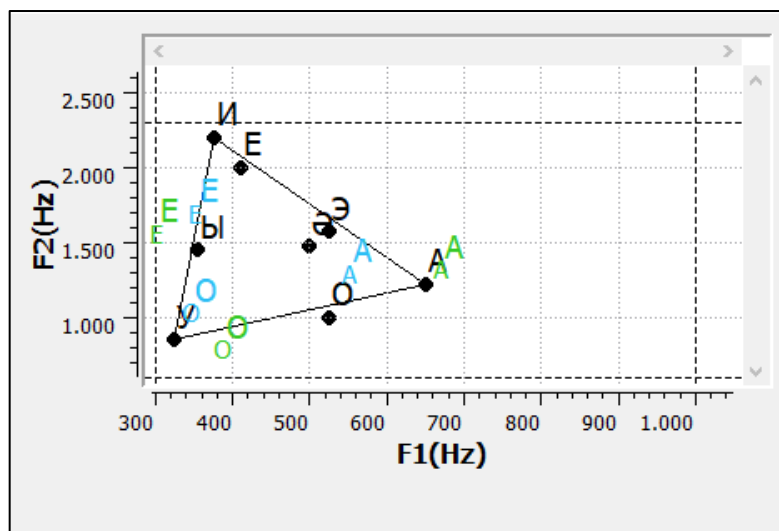


Figura 3: Carta de formantes para cada amostra

A dispersão dos valores das vogais está relacionada com as características individuais de cada falante. Assim, podemos verificar a partir da Figura 3 que os formantes das vogais de cada amostra apresentam significativa variação.

Essa variação tem como uma das explicações, o fato de que os formantes apresentam uma relação direta com as características fisiológicas do trato



vocal, pois cada indivíduo opta por um ajustamento articulatorio diferente enquanto fala. Dessa forma, podemos verificar o indicio de falantes distintos entre as amostras.

5.3 LONG TERM AVERAGE SPECTRUM

A Long Term Average Spectrum (LTAS) é um desenho bidimensional da intensidade em função da frequência espectral da amostra. Ela irá refletir a contribuição da fonte glótica e do trato vocal na qualidade de uma voz.

A Figura 4 mostra a LTAS de cada amostra, em que os picos no espectro correspondem à média de cada formante de todos os sons amostrados. Podemos observar que há uma diferença significativa entre a amostra questionada (linha preta) e amostra padrão (linha vermelha):

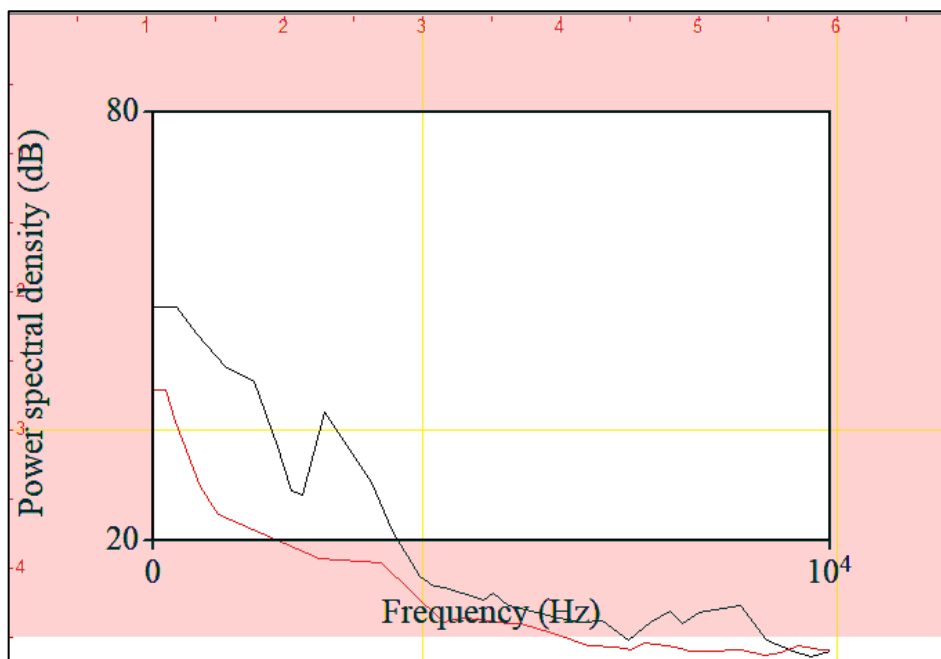


Figura 4: LTAS de cada amostra



5.4 VOT

O VOT (Tempo de Ataque de Vozeamento) é um parâmetro utilizado para identificar as consoantes oclusivas surdas e sonoras, através do intervalo de tempo entre a liberação articulatória da oclusão e o início de vozeamento a ele associado.

Esse parâmetro é eficaz na percepção e produção de oclusivas em início de palavras, característicos para cada falante. Dessa forma, esse parâmetro é uma ferramenta muito eficaz.

O VOT foi medido a partir da duração absoluta. O padrão seguido para a extração das medidas foi baseado no parâmetro seguido por Prestes (2013) em relação às oclusivas surdas, coletando a duração em milissegundos entre o burst ²e o primeiro pulso regular da vogal seguinte. A oclusiva utilizada para cálculo do VOT foi o fonema [t].

As análises evidenciaram importante diferença entre as VOTs de uma amostra com relação a outra. A Figura 5 mostra o valor em milissegundos do VOT correspondente a amostra questionada, qual seja de 0.018 milissegundos:

² “Explosão de energia” caracterizada por uma barra vertical no espectrograma e por uma forma de onda aperiódica, imediatamente após um sinal sem energia acústica.

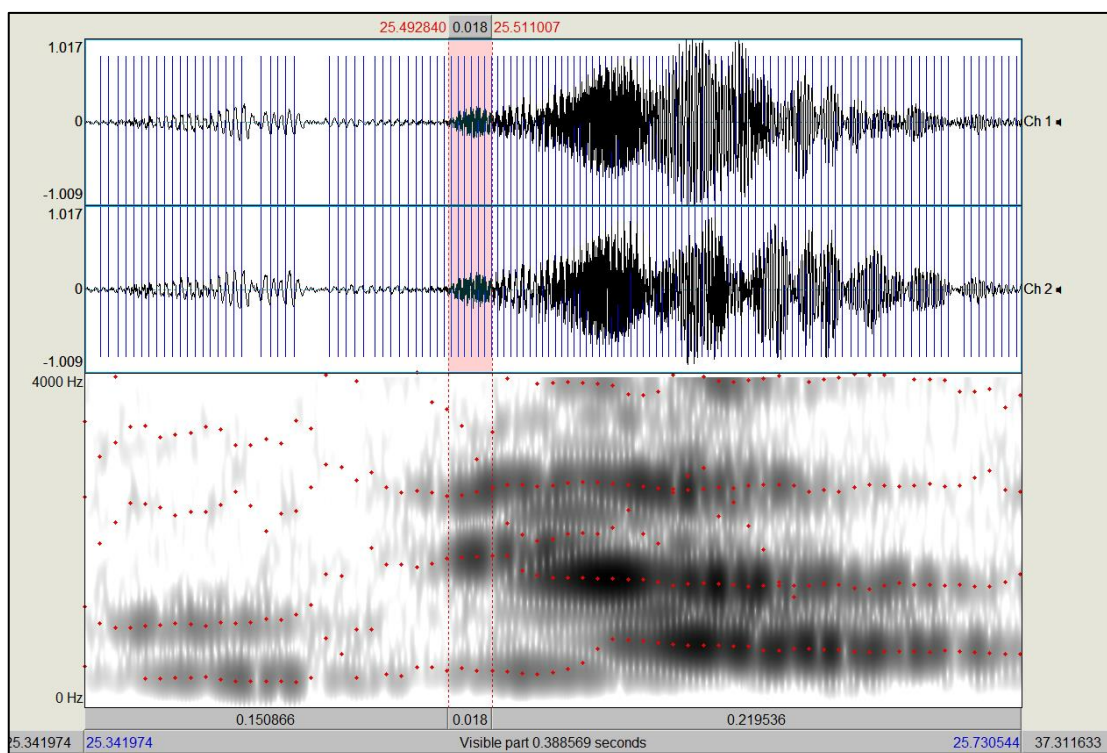


Figura 5: VOT correspondente a amostra questionada

A Figura 6 mostra o valor em milissegundos do VOT correspondente a amostra padrão, qual seja de 0.010 milissegundos:

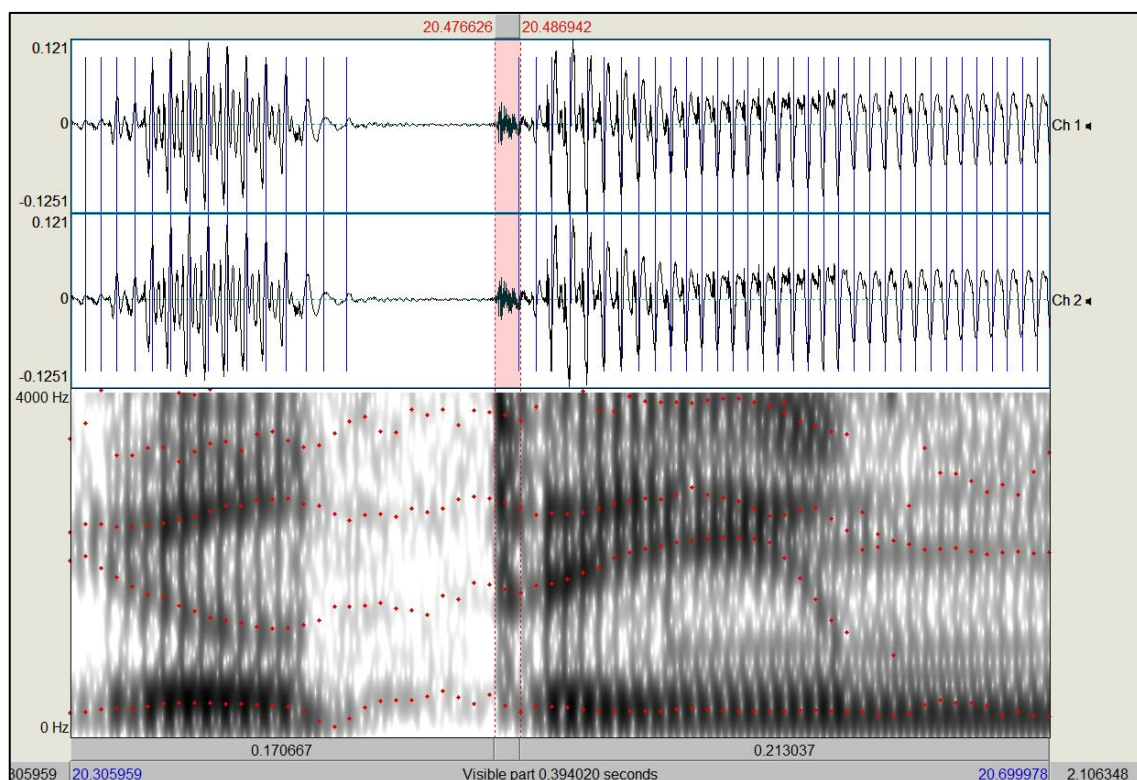


Figura 6: VOT correspondente a amostra padrão

5.5 FONEMA [j]

De acordo com a análise perceptivo-auditiva, na amostra questionada à a presença de distorção no fonema [j], o que não é observado na amostra padrão. O fonema [j] é classificado como uma consoante fricativa, para que ocorra sua produção a corrente de ar passa por obstruções parciais formadas num ponto do trato vocal, assim, com a passagem forçada da corrente de ar, irá gerar turbulência. Auditivamente essa turbulência irá gerar a sensação auditiva de ruído.

Em uma análise acústica, como no exemplo abaixo (Figura 7), essa turbulência pode ser vista pela forma de onda aperiódica, que



se contrasta drasticamente com a forma de onda à direita e à esquerda, como pode ser vista pela seta em vermelho. Já no espectrograma, a turbulência pode ser vista como um “chuveirco”, como demonstrado pela seta azul.

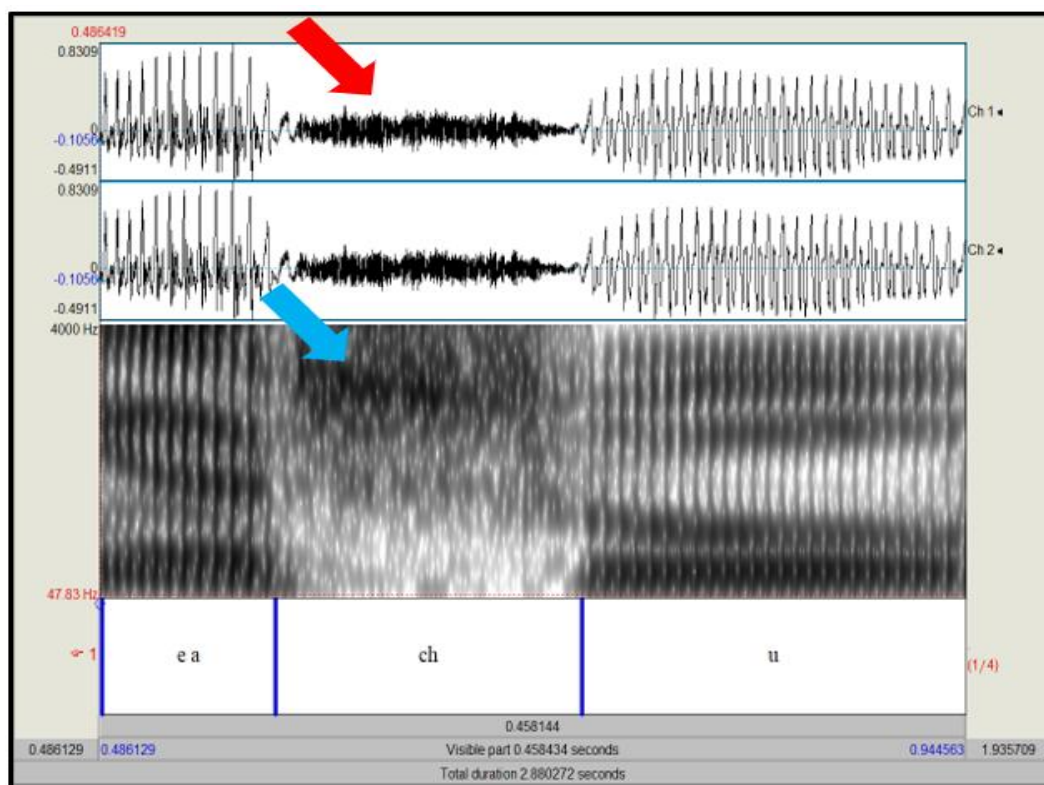


Figura 7: Exemplo de onda sonora e espectrograma do fonema [j]

Na amostra questionada não observamos esse padrão, tanto na forma de onda quando no espectrograma, demonstrando assim sua distorção. Porém na amostra padrão podemos observar esse padrão perfeitamente, indicando que não há dificuldade em sua produção. Dessa forma, podemos constatar que o falante correspondente ao áudio padrão não apresenta alteração articulatória na pronúncia do fonema [j].



A Figura 8, correspondente ao áudio questionado, mostra o fonema [j], indicado pela seta em vermelho, em que podemos observar que a forma de onda não se contrasta drasticamente com a forma de onda à direita e à esquerda, e no espectrograma não a o “chuvisco” como demonstrado no exemplo. Esse padrão indica a distorção na pronúncia do fonema.

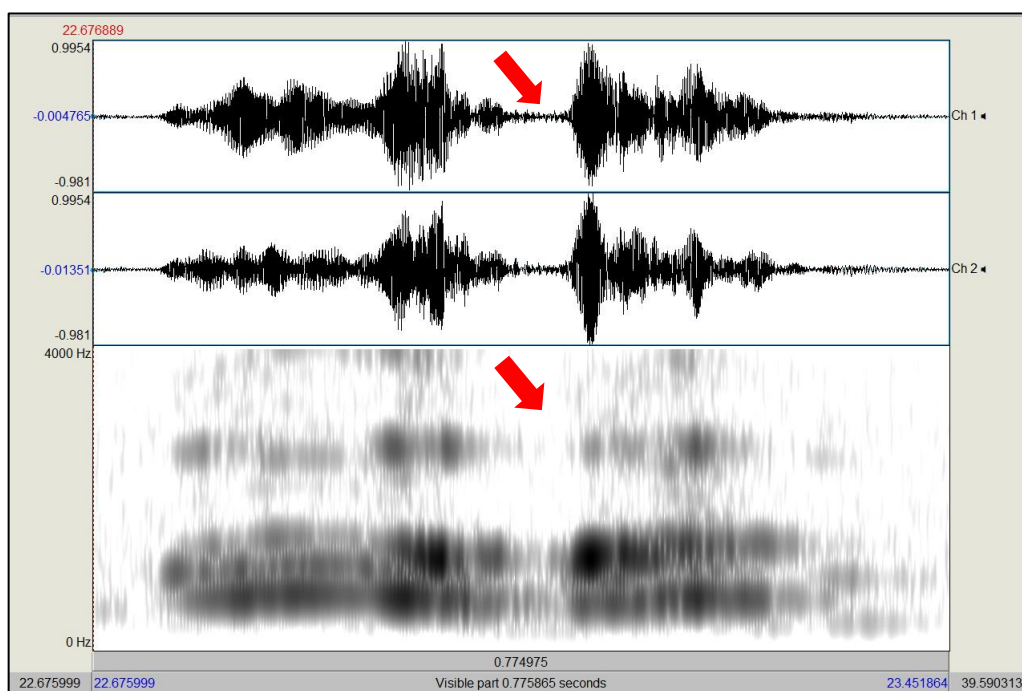


Figura 8: Forma de onda e espectrograma do fonema [j] referente a amostra questionada

A correspondente ao áudio padrão, mostra o fonema [j], indicado pela seta em vermelho, em que podemos observar que a forma de onda é aperiódica e se contrasta drasticamente com a forma de onda à direita e à esquerda, e no espectrograma à a presença do “chuvisco”, como demonstrado no exemplo. Esse padrão indica que o falante não apresenta dificuldade em sua pronúncia.

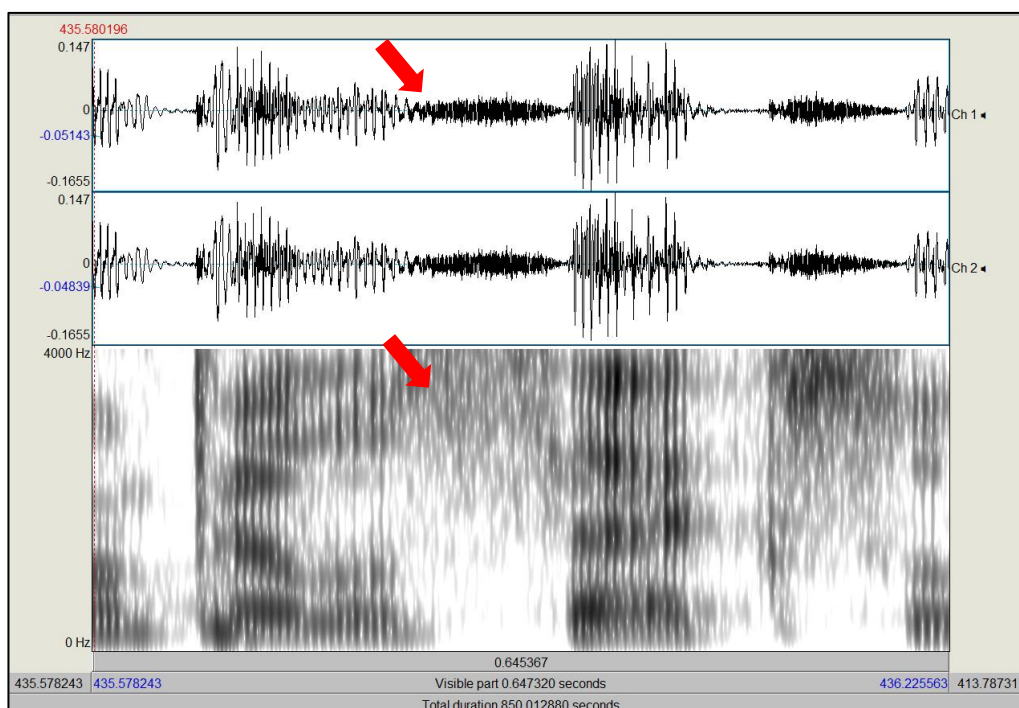


Figura 9: Forma de onda e espectrograma do fonema [j] referente a amostra padrão

6. CONCLUSÕES

A partir dos resultados das análises realizadas, conclui-se que:

- É possível identificar diferenças quanto aos parâmetros perceptivo-auditivos: ressonância, articulação, velocidade de fala, dialeto e desvio fonético entre as amostras de fala e voz analisadas;
- É possível identificar diferenças quanto aos parâmetros acústicos: f_0 , formantes, LTAS, VOT e na pronúncia do fonema [j], entre as amostras de fala e voz analisadas;

Dessa forma, a conclusão, considerando as análises perceptivo-auditiva e acústica da fala, será uma afirmação correspondente ao nível mais



representativo da escala exibida na figura abaixo (Figura 10), composta de expressões verbais que denota, da inconclusão (nível 0) a quatro níveis de suporte (níveis +1 a +4) ou contradição (níveis -1 a -4) à hipótese pretendida. O corpo probatório sustenta muito fortemente a hipótese de que as amostras de fala não foram proferidas pelo mesmo locutor, o que equivale ao nível -4 da escala apresentada na Figura 10.

GRAU	CORRELATO VERBAL
+ 4	o resultado suporta muito fortemente a hipótese
+ 3	o resultado suporta fortemente a hipótese
+ 2	o resultado suporta moderadamente a hipótese
+ 1	o resultado suporta levemente a hipótese
0	o resultado nem suporta nem contradiz a hipótese
- 1	o resultado contradiz levemente a hipótese
- 2	o resultado contradiz moderadamente a hipótese
- 3	o resultado contradiz fortemente a hipótese
- 4	o resultado contradiz muito fortemente a hipótese

Figura 10 – Escala utilizada na conclusão dos exames (Eriksson, 2012)

Portanto, as evidências indicam que as amostras de fala presentes no áudio questionado e no áudio paradigma não foram proferidas pelo mesmo locutor.



ENCERRAMENTO

Eventuais divergências ou contestações de qualquer natureza quanto à apropriação aos falantes deverão ser prontamente e formalmente submetidas pelos consulentes aos subscritores no endereço indicado abaixo, para avaliação técnica e, caso sejam consideradas procedentes, imediata republicação do parecer ou edição de adendo. Caso seja necessário, outras análises poderão ser eventualmente apresentas em Parecer Técnico Complementar. Os consulentes, ao receberem, utilizarem ou divulgarem o presente parecer ou as informações nele contidas, automaticamente concordam em assumir exclusivamente para si as responsabilidades pertinentes, comprometendo-se a isentar e proteger o IBP, seus diretores, peritos e funcionários de qualquer reivindicação eventualmente resultante. Encerra-se o presente parecer técnico que contém 18 folhas impressas em seu anverso, sendo esta última datada e assinada.

São Paulo, 21 de setembro de 2018.

Giuliano Giova
Diretor

Gustavo Batistuzzo
Gerente Técnico

Priscilla Donaire Brasil
Fonoaudióloga CRFa 2-19397