

LAUDO DE EXAME PERICIAL

ESPÉCIE: Análise Acústico-perceptual.

TIPO DE EXAME: Analítico-comparativo de Autenticidade e
Identificação de Locutor

Belo Horizonte
Outubro – 2020

1. APRESENTAÇÃO DO PERITO

João Sival Patrício da Silva, graduado em Psicologia pela Faculdade de Estudos Administrativos de Minas Gerais, CRP-MG 40202; Técnico de Informática, Perito do Instituto Mineiro de Perícia em áudio e vídeo; Trânsito; Degravação; Comparação de Voz e Psicologia Forense, Perito judicial do Fórum Lafaiete de Belo Horizonte/MG e de outras comarcas de MG, localizado na Av. Do Contorno 5491/4º andar, Belo Horizonte/MG, telefones para contato 31 – 99104-0501.

2. APARELHAGEM UTILIZADA

Pela complexidade das análises, o perito utilizou dois microcomputadores da marca Apple, modelos Imac 21.5 polegadas com processador Intel Core 2 Duo de 3.06 GHZ 1 GB de memória Ram, placa de vídeo NVIDIA GeForce 9400 e MacBook Pro 17 polegadas com processador Intel Core i5 de 2.53 GHZ 4 GB de memória Ram, placa de vídeo NVIDIA GeForce GT330M, placa de som marca MOTU, modelo TRAVELER, Fone de ouvido marca ROLAND, modelo RH-200 e os softwares Adobe Audition, PRAAT e MATLAB .

3. MATERIAL ANALISADO

3.1 Peça Motivo

1 (um) arquivo de áudio com aproximadamente 00:02:00 (dois minutos), aparentemente de propaganda política/partidária. Foi avaliado tão somente a voz atribuída ao Sr. Leonardo Ângelo Costa Ribeiro (entre 01 minuto e 35 segundos e 01 minuto e 46 segundos, aproximadamente).

3.2 Peça Padrão

Foi realizada a coleta da peça padrão na data de 22 de outubro de 2020, a partir das 14:00h, por este signatário. A peça é composta por um áudio, contendo algumas frases específicas.

4. OBJETIVO PERICIAL

Comparar as vozes nos arquivos enviados a exame (arquivo de áudio denominado por estes signatários como **“PEÇA PADRÃO”** e **“PEÇA MOTIVO”**), identificando se é do mesmo locutor, indicado pelo solicitante deste laudo como sendo a voz de **“Leonardo Ângelo Costa Ribeiro”**.

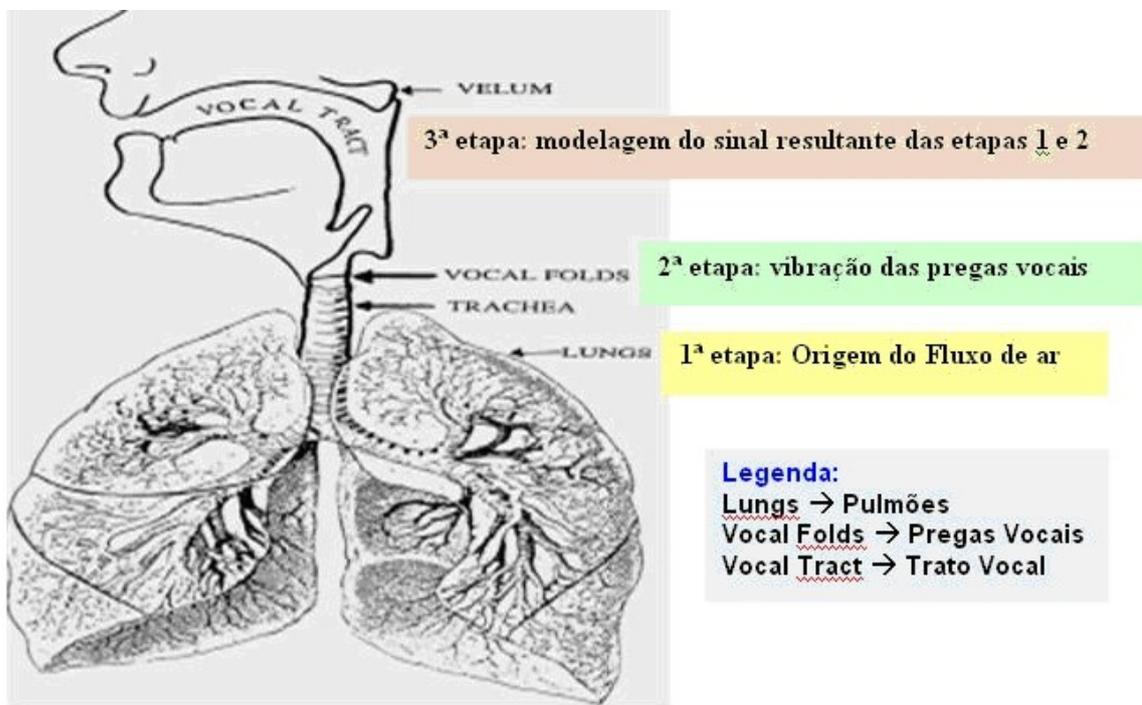
5. DOS EXAMES

A verificação de locutor é o exame pericial que atribui univocamente à autoria de determinadas falas armazenadas numa mídia de gravação a alguém ou desvincula a sua participação no crime, constatando que as falas examinadas não partiram do aparelho fonador do suspeito e fornecedor do material sonoro padrão. Para melhor entendimento acerca deste tipo de exame é necessária uma breve explanação sobre o processo de produção da fala, porquanto diversas técnicas de processamento digital de sinais de voz são baseadas em modelos matemáticos que procuram descrever o funcionamento do sistema humano de produção de voz.

O som consiste na sensação proporcionada pela vibração de delicadas membranas do ouvido, causada por uma onda de pressão que se propaga geralmente no ar. É importante observar que tal onda se origina em geral da vibração de algum material e, especificamente em tratando da voz humana, sabemos que ela é gerada pela vibração das pregas vocais.

O processo mecânico de produção da voz pode ser dividido em três etapas: 1ª) a criação pelos pulmões de um fluxo de ar; 2ª) a passagem desse fluxo de ar pela laringe, passando pelas pregas vocais; e 3ª) atuação do trato vocal. Na produção da voz, a laringe assume o papel de mecanismo fonatório, transformando o fluxo de ar contínuo proveniente dos pulmões em uma onda de pulsos de ar. Já o trato vocal

também pode atuar de várias maneiras a fim de produzir um fonema vocálico ou consonantal.



As vogais representam os sons mais puros porque são geradas a partir da vibração das pregas vocais e sem a ocorrência de constrições ou obstruções no trato vocal. Já nas consoantes existe alguma resistência à passagem de ar e pode ou não haver vibração das pregas vocais. Tanto os fonemas vocálicos quanto os consonantais possuem classificações (zona de articulação, modo de articulação, etc), mas que fogem ao escopo do objetivo desta breve consideração.

Nestes exames foram utilizadas as seguintes análises:

A - Análise qualitativa do sinal

O som pode ser definido como uma forma de energia ondulatória que se propaga em meios elásticos, o ar é um bom exemplo de meio elástico. Ao se interpretar as dimensões ou atributos da onda sonora nota-se que a ela estão relacionadas às qualidades de altura, intensidade e timbre.

A altura é a qualidade relacionada à frequência da onda sonora que, por sua vez, nos permite classificá-la em uma escala que varia de grave a aguda. Quanto mais alta a frequência mais aguda é o som, portanto quanto mais baixa a frequência mais grave é o som.

A intensidade é uma qualidade relacionada tanto à amplitude da onda sonora quanto à sua pressão efetiva e sua energia transportada, permitindo-se classificá-la dentro de uma escala que varia de fraco a forte. Dessa maneira, quanto maior for a amplitude, a pressão efetiva e a energia transportada pela onda sonora mais forte é o som.

Timbre não é uma qualidade do som, mas sim da fonte sonora. Por meio dele podemos diferenciar, por exemplo, vozes diferentes faladas ao telefone.

Em uma análise qualitativa do som procura-se definir e/ou identificar cada produção sonora presente no suporte tendo por base as qualidades sonoras descritas acima, contudo a análise pode ser prejudicada pela presença de ruídos presentes quando da gravação.

Os ruídos podem ser ocasionados por uma série de variáveis como:

- Ambiente de gravação; ruidoso como uma oficina mecânica;
- Produção de ruídos pelo próprio aparelho gravador;
- Ecos e reverberações; etc...

B - Análise de conteúdo do sinal (análise do discurso)

Nesse momento serão avaliadas:

- A sequência lógica dos discursos, no contínuo de tempo;
- A sequência de falantes;
- Os ruídos de fundo.

C - Análise do sinal

O arquivo de áudio examinado apresentou, à época dos exames a gravação de um diálogo por pessoas do sexo masculino, não identificados na gravação.

A análise de sinal irá ater-se exclusivamente às medidas obtidas pela modelagem matemática e pela interpretação física do sinal enviado à exames, e será subdividida em:

C.1 - Análise de formantes

Para um melhor entendimento é válida uma breve análise do trato vocal e de suas particularidades:

Na produção da fala, o ar expelido pelos pulmões é conduzido pela traqueia, passa pelas pregas vocais, atinge as cavidades oral e/ou nasal, e o som (produzido de alguma forma dentro do trato vocal) é irradiado na boca e/ou narinas.

Os movimentos da língua, maxilar, e lábios dão forma à cavidade oral, dependendo do som desejado. As cavidades do trato vocal (filtros) são excitadas pela combinação de pelo menos três tipos de estímulos (fontes): uma sequência de pulsos de ar produzidos pela vibração quase-periódica das pregas vocais durante a geração de sons sonoros (vogais e consoantes como /l/ e /m/); fluxo turbulento de ar (ruído) causado por constrições em alguma parte do trato vocal, como fricativos “surdos” (não sonoros) /f/ e /s/; um único pulso produzido pela liberação súbita do ar comprimido em uma cavidade formada dentro do trato vocal, como nos oclusivos (surdos) /p/ e /t/.

Os fenômenos associados com a vibração das pregas vocais têm sido mais estudados, principalmente por estarem diretamente relacionados com a produção de vogais. Na maioria das aplicações, a qualidade da síntese e a precisão da análise de vogais são, em geral, mais importantes que a análise ou síntese dos outros fonemas.

Vogais são definidas basicamente pela amplificação inserida na energia glótica. As faixas de frequência amplificadas, ou seja, os picos de energia variam de acordo com

a vogal emitida e, no caso específico das vogais, representam grupos de harmônicos e recebem o nome de formantes do som.

Assim, formantes das vogais são “faixas de frequência que caracteristicamente contêm centroides de energia” (Black,1937), ou seja, uma concentração de energia acústica numa faixa de frequências.

Os formantes são em número variável e dependem da vogal emitida, das características anátomo-funcionais do indivíduo e do treino de voz que este possui. Os valores dos formantes representam as frequências naturais de ressonância do trato vocal na posição articulatória específica da vogal emitida e são expressos em Hertz (Hz) ou ciclos por segundo (c/s), através de um valor numérico que é geralmente a média das frequências que este contém. No presente trabalho foram analisadas tanto vogais como consoantes, através de exames de espectrografia acústica e análise de formantes (vogais) em tempo real.

C.2 - Análise de frequências

A análise das frequências que constituem o sinal de áudio enviado a exames foi realizada tanto em sua forma de onda como em gráficos espectrais.

5.1 Transcrição do Áudio

Cabe esclarecer que foi utilizada codificação específica para esta análise, que é dada a seguir:

- Números entre parêntesis, designam a quantidade de palavras ininteligíveis no período. Ex: (2/3) existem entre duas e três palavras ininteligíveis no período;
- Comentários são inseridos entre barras. Ex: /risos/;
- O discurso será transcrito da forma com que foi realizado, ou seja, não serão corrigidos erros gramaticais ou de pronuncia, uma vez que o exame pretende reconstituir um registro oral.

Arquivo: “PEÇA MOTIVO”

//INÍCIO DAS TRANSCRIÇÕES//

VOZ 1: Ontem chegaram lá na cooperativa já avisando pra todo mundo que se não fosse tava cortado da cooperativa, que tinha que ir nessa porra dessa carreata, pra receber 30 reais de gasolina. O quê que eu vou fazer com 30 reais de gasolina?

Arquivo: “PEÇA PADRÃO”

Nome Completo: Leonardo Ângelo Costa Ribeiro.

Endereço: Rua Marlene Alevato Ferrari, número 96, Bairro Osvaldo Barbosa Pena – é, Nova Lima – Minas Gerais.

Nome dos pais: Paulo Jorge Ribeiro e Maria Ângela da Costa Ribeiro.

Data de nascimento: 13/08/1981

Segundo o último levantamento semanal disponibilizado pela Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP), entre os dias 16 e 22 de agosto, o preço médio da gasolina comum no país era, na semana pesquisada, de R\$ 4,268. O diesel S-500, de R\$ 3,374. O etanol, de R\$ 2,782.

A agenda de campanha dos quatro candidatos posicionados entre o primeiro e quarto lugares na última pesquisa Ibope de intenção de voto teve carreatas, encontro com empresários, médicos e representantes de cooperativas.

"Tivemos aumento de quase 30% no IPTU e nos impostos municipais nos últimos dois anos. A população não aguenta mais aumento de imposto. Então nós tivemos o cuidado de assumir o compromisso e registrar em cartório que nos próximos quatro anos não haverá aumento de impostos no município", ressalta Alysson Lima.

Ontem chegaram lá na cooperativa já avisando pra todo mundo que se não fosse tava cortado da cooperativa, que tinha que ir nessa porra dessa carreatã, pra receber 30 reais de gasolina. O quê que eu vou fazer com 30 reais de gasolina?

Ontem chegaram lá na cooperativa já avisando pra todo mundo que se não fosse tava cortado da cooperativa, que tinha que ir nessa porra dessa carreatã, pra receber 30 reais de gasolina. O quê que eu vou fazer com 30 reais de gasolina?

//FIM DA TRANSCRIÇÃO//

6 - DOS RESULTADOS DOS EXAMES

6.1 Análise qualitativa do sinal

O sinal analisado apresentou boa qualidade sonora, não necessitando os peritos a realizarem processos de filtragem.

6.2 Processos realizados

Primeiramente foi realizada a comparação auditiva, no qual observa os parâmetros da fala: de características culturais, como sotaque e vocabulário, as pessoais, como respiração e entonação, incluindo traços como rouquidão.

Já a comparação visual utiliza o chamado espectrograma, espécie de raio X das ondas sonoras, antes produzido pelo espectrógrafo de som.

7 – CONCLUSÃO

Ao final dos exames técnicos constatou-se que o suporte magnético apresentou quantidade e qualidade suficiente para que se possa confirmar que a voz masculina identificada no arquivo de áudio “PEÇA PADRÃO” **NÃO** está presente no arquivo de áudio da “PEÇA MOTIVO”.

Com 100% de probabilidade, este signatário conclui que a voz encontrada na peça motivo, transcrita e denominada como “VOZ 1” da peça motivo, NÃO pertence ao padrão fonético do Sr. Leonardo Ângelo.

Tal conclusão pode ser alcançada baseando-se nos parâmetros apresentados a seguir:

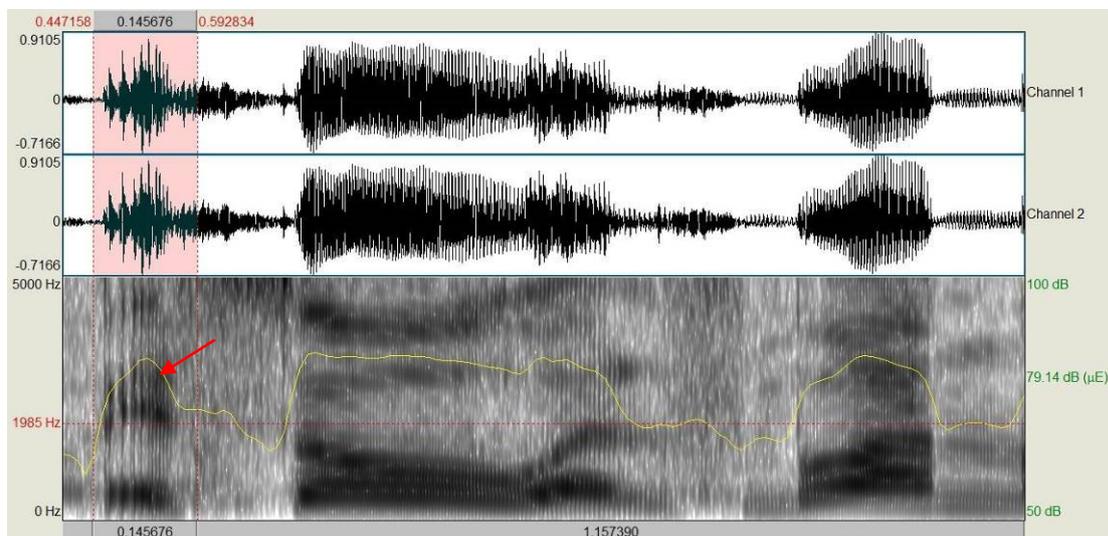
A emitente da voz na **peça motivo, voz 1**, emite linguagem informal, com mau sincronismo da fala e uso de vários vícios de linguagem.

Vícios de Linguagem são alterações defeituosas das normas da língua padrão, provocadas por ignorância, descuido ou descaso por parte do falante.

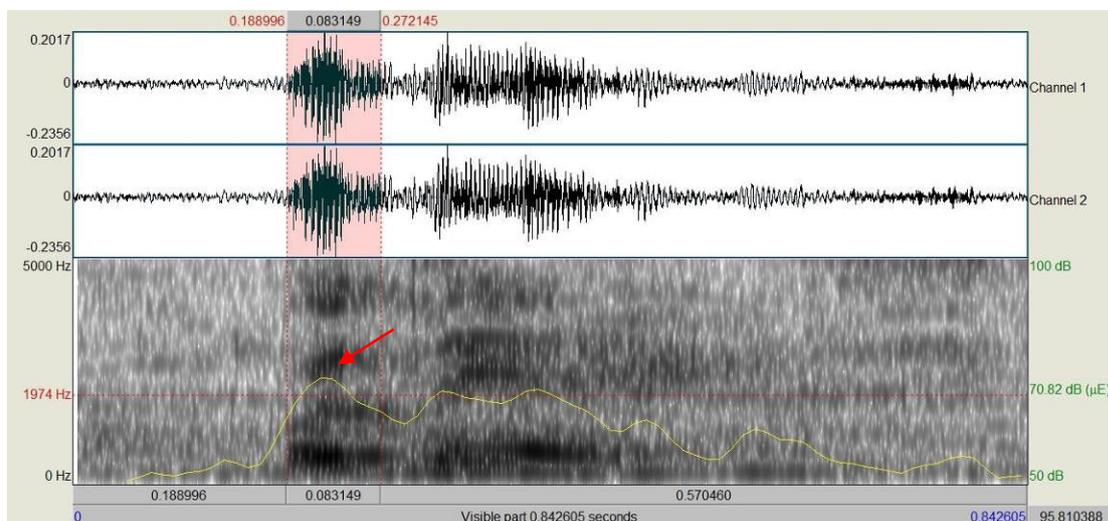
O tom de voz apresentado no áudio contestado possui dicção distinta à do Sr. Leonardo, fala dessincronizada (período entre uma palavra e outra), e em nenhum momento percebe-se tom de voz semelhante (alterações como exaltação, morosidade, nuança).

Já a linguagem emitida pelo Sr. Leonardo é sincronizada, com boa dicção, sem o uso de vícios de linguagem, reproduzindo integralmente o som de cada palavra, não demonstrando vício cultural explícito.

1 - Distinção entre os padrões melódicos da peça motivo e peça padrão - Dessincronização da fala.



Peça motivo: padrão acelerado – (palavra “coope” (coopertativa)).

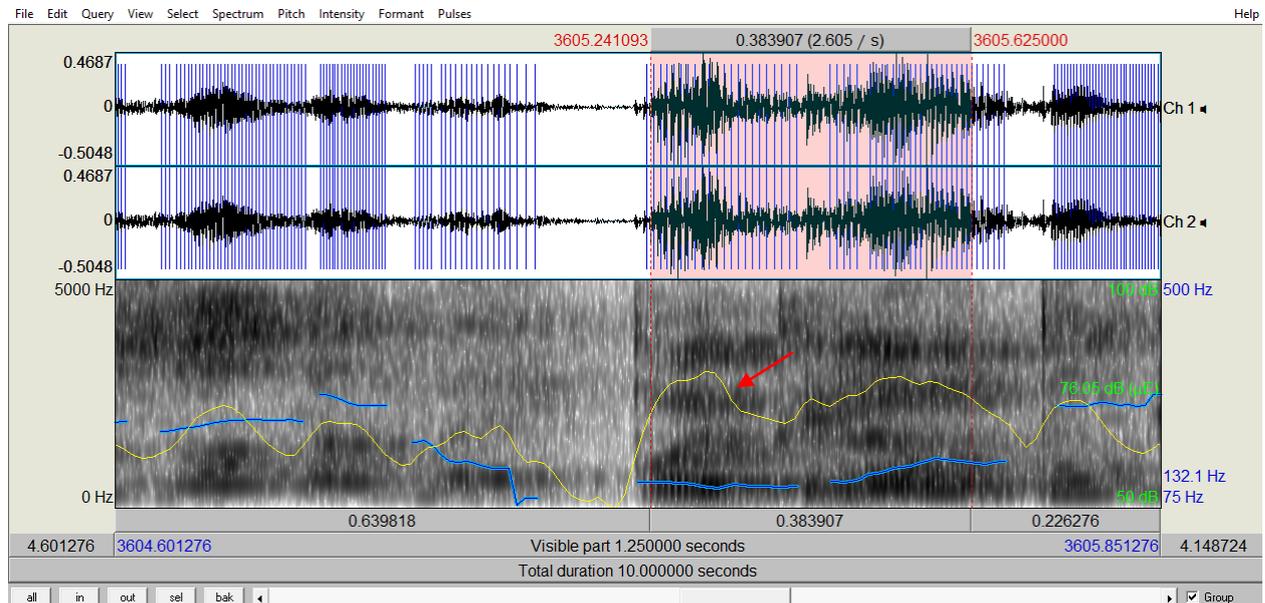


Peça Padrão: padrão melódico - (palavra “coope” (coopertativa)).

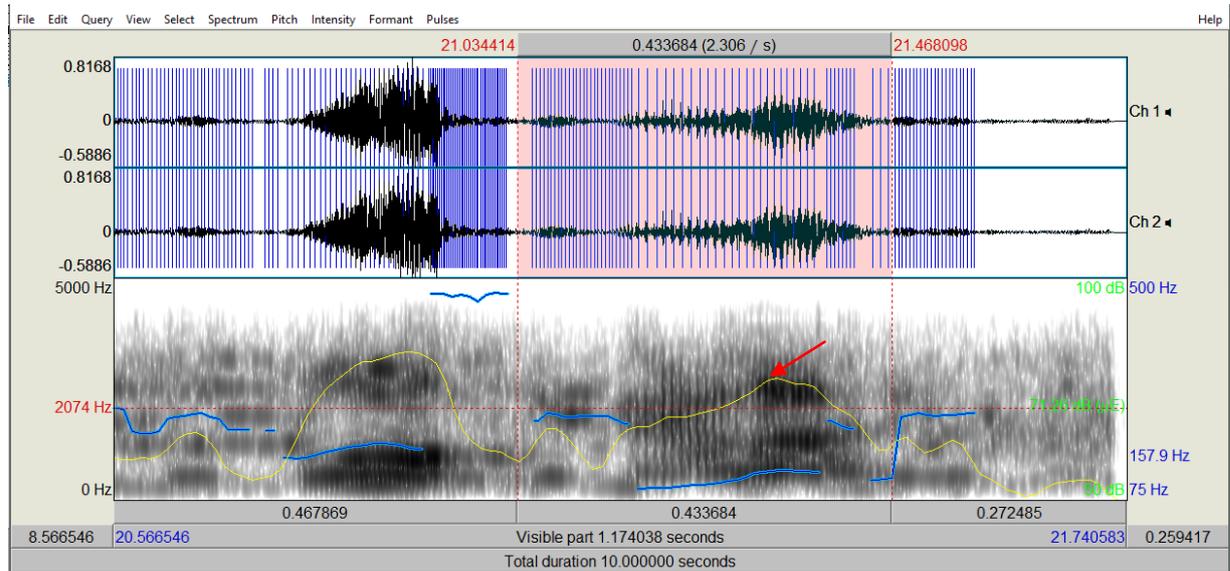
Percebe-se que a voz identificada no arquivo questionado possui “língua presa” (condição que restringe a faixa de movimento da língua) ao expressar o o início da palavra coopertativa, tornando em sua expressão menos limpa, ou seja, sua dicção é falha.

2 – Dessincronização da fala.

Peça motivo: padrão acelerado e entonação de língua presa – (palavra “trinta”).



Peça Padrão: padrão melódico - (palavra “trinta”).



O locutor da peça motivo, devido ao que, aparentemente ser “linguam presa” (condição que restringe a faixa de movimento da língua), não expressa uma dicção clara ao falar a palavra trinta (referente a 30 litros de gasolina). É possível perceber tal falha na dicção nas duas vezes em que repete o gesto, contrário ao Sr. Leonardo, que expressa a palavra de forma correta, com devido espaçamento entre uma sílaba e outra.

8. ENCERRAMENTO

Este signatário baseou-se tal conclusão não somente nos arquivos coletados e análises demonstradas, mas também nas observações realizadas antes do momento das gravações, ouvindo o Sr. Leonardo, identificando as particularidades de seu padrão fonético, a fim de identificar possível tentativa de simulação ou mudança em sua dicção. Também foram analisados arquivos encontrado na internet (https://youtu.be/kvM2L0907_Y - <https://youtu.be/OHfR4ldjQGU>), a fim de corroborar as informações acima prestada.

Durante as conversas pré-coleta e pós-coleta, foi observado se, fora do momento de gravação, o Sr. Leonardo emitia padrões parecidos com a voz do arquivo questionado (analisado previamente por este signatário), sendo eles: sotaque, vícios de linguagem, dicção falha; que não foram percebidos.

Uma vez que consta no parecer contestado que há **certo grau de probabilidade** de acerto para autenticidade, entende-se que as conclusões não são finais, contrário ao provado neste laudo pericial, onde este signatário afirma com 100% de certeza de que o Sr. Leonardo Ângelo Costa Ribeiro **NÃO É O LOCUTOR DO ARQUIVO QUESTIONADO.**

Entende-se ainda que o arquivo foi alterado em seu formato original e a possibilidade de apuração e identificação se torna questionável. O perito que modificou o arquivo, deverá apresentar a metodologia utilizada, bem como os parâmetros empregados para conseguir proporcionar a equalização do áudio a ponto de coincidir com uma voz aleatória e ainda, após esse feito, demonstrar, cientificamente, que é compatível com a voz do senhor Leonardo. Caso o perito tenha recebido o áudio de um terceiro,

já com a retirada da distorção, ou seja, o arquivo foi trabalhado sem o seu certame técnico, a conclusão da perícia se torna nula, pois seu resultado é dúbio e a sua fonte referencial é torna duvidosa. Induzindo-o ao erro.

Portanto, baseia sua conclusão nestes parâmetros que julga de relevância e ainda esclarece que TODOS convergem para a conclusão apresentada.

Belo Horizonte, 23 de outubro de 2020.



João Sinval Patrício da Silva
PERITO FONOGRÁFICO

BIBLIOGRAFIA UTILIZADA

Araújo, G.M. & Regazzi, R.D. (1999). *Perícias e avaliação de ruído e calor*. 1º ED. Rio de Janeiro.

Baldwin, J. & French, P. (1990). *Forensic Phonetics*. Billing & Sons. Bristol.

Kent, R. D. & Read, C. (1992). *The Acoustic Analysis of Speech*. McNaughton & Gunn. San Diego.

Black JW. The quality of a spoken vowel. Arch. Speech 1937

Russo, I.C.P. (1999). *Acústica e Psicoacústica aplicadas à Fonoaudiologia*. 2ª Ed. Editora Lovise. São Paulo

Russo, I & Behlau, M. (1993). *Percepção da fala: Análise Acústica*. Editora Lovise. São Paulo.

Silva, T.C. (1999). *Fonética e Fonologia do Português*. Editora Contexto. São Paulo.

Vieira, M.N. (1997). *Acústica Princípios da produção e análise de voz*. Departamento de Física/Icex/UFMG. Belo Horizonte.

Vieira, M.N. (1993). *Processamento digital de sinais*. Departamento de Física/Icex/UFMG. Belo Horizonte.