

SOBRE O POTENCIAL DISCRIMINANTE DAS PROPRIEDADES DE VOZ E FALA NA TAREFA DE COMPARAÇÃO DE LOCUTORES: UM ESTUDO DE CASO

ON THE DISCRIMINATING POWER OF VOICE/SPEECH PROPERTIES IN SPEAKER COMPARISON TASK: A CASE STUDY

*Cláudia Regina Brescancini*¹
*Cíntia Schivinski Gonçalves*²
*Dênis Fernandes*³
*June Campos*⁴
*Ana Paula Correa da Silva Biasibetti*⁵
*Felipe Bilharva da Silva*⁶
*Márcio Oppliger Pinto*⁷

RESUMO: Este estudo tem por objetivo replicar a tarefa pericial de comparação de locutores ao confrontar uma amostra de fala teste (amostra questionada), produzida por um indivíduo adulto do sexo masculino, com três amostras de fala de referência (amostras padrões), pertencentes ao irmão gêmeo do locutor questionado (LP1), ao próprio locutor questionado (LP2) e a outro locutor masculino com idade semelhante a do locutor questionado (LP3). Busca-se confrontar as propriedades de voz/fala provenientes das amostras para identificação da autoria a partir das características (socio)linguísticas e de voz que revelem poder discriminante. A análise é conduzida com base na verificação perceptivo-auditiva e acústica das amostras, o que compreende o chamado *método combinado*, tendo por guia o protocolo proposto por Gonçalves e Petry (2015). Os elementos técnico-comparativos encontrados na examinação evidenciam que mesmo irmãos gêmeos podem ser diferenciados em razão de suas

¹ Doutora em Letras, Professora Adjunta da Faculdade de Letras da Pontifícia Universidade Católica do Rio Grande do Sul, membro do Laboratório de Áudio e Fonética Acústica (LAFA)/PUCRS.

² Doutora em Letras, Fonoaudióloga, Perita Criminal do Instituto Geral de Perícias do Rio Grande do Sul, membro do Laboratório de Áudio e Fonética Acústica (LAFA)/PUCRS.

³ Doutor em Engenharia Elétrica, Professor Adjunto da Faculdade de Engenharia da Pontifícia Universidade Católica do Rio Grande do Sul, coordenador do Laboratório de Áudio e Fonética Acústica (LAFA)/PUCRS.

⁴ Doutora em Letras, Professora da Universidade Federal do Rio Grande do Sul, membro do Laboratório de Áudio e Fonética Acústica (LAFA)/PUCRS.

⁵ Doutoranda em Letras, Programa de Pós-Graduação em Letras da Pontifícia Universidade Católica do Rio Grande do Sul, membro do Laboratório de Áudio e Fonética Acústica (LAFA)/PUCRS.

⁶ Doutorando em Letras, Pontifícia Universidade Católica do Rio Grande do Sul, membro do Laboratório de Áudio e Fonética Acústica (LAFA)/PUCRS.

⁷ Doutorando em Letras, Pontifícia Universidade Católica do Rio Grande do Sul; Fonoaudiólogo; Perito Criminal do Instituto-Geral de Perícias do Rio Grande do Sul, membro do Laboratório de Áudio e Fonética Acústica (LAFA)/PUCRS.

propriedades de voz/fala e que, no caso em questão, o exame de aspectos referentes aos perfis de voz, à taxa de articulação, à frequência fundamental e à produção de róticos em coda são suficientes e eficientes para sustentar fortemente a hipótese de mesma origem no caso do confronto entre a amostra referente ao locutor questionado e a amostra referente ao LP2.

PALAVRAS-CHAVE: Fonética Forense. Comparação de Locutores. Sociolinguística. Fonética Acústica. Fonética Impressionística.

ABSTRACT: This study aims to reply the police task of speaker comparison by confronting a speech sample (unknown sample (US)) produced by a male adult and three other reference speech samples (known samples) belonging to the twin brother of the unknown speaker, to the unknown speaker himself and to other male adult with the same age as the unknown speaker. The goal is to compare the voice/speech properties of the samples (from now on US, Sp1, Sp2 and Sp3) in order to identify the (socio)linguistic characteristics and the voice characteristics that reveal discriminant power. The analysis is conducted through the perceptual and acoustic verification of the samples (the so-called *mixed method*) based on the protocol proposed by Gonçalves e Petry (2015). The technical and comparative elements found in the present exam evidence that even twin brothers can be distinguished by their voice/speech properties and that the exam of the voice profiles, the articulation rate, the fundamental frequency and the rhotic production in coda position are sufficient and effective to strongly sustain the hypothesis for the similarity between the unknown speaker sample and the Sp2 sample.

KEYWORDS: Forensic Phonetics. Speaker Comparison. Sociolinguistics. Acoustic Phonetics. Impressionistic Phonetics.

SOBRE O POTENCIAL DISCRIMINANTE DAS PROPRIEDADES DE VOZ E FALA NA TAREFA DE COMPARAÇÃO DE LOCUTORES: UME STUDO DE CASO

1 INTRODUÇÃO

O reconhecimento de um indivíduo a partir de sua voz e fala - o *reconhecimento de locutor* (*speaker recognition*) (HOLLIEN, 2002; NOLAN, 1983; ROSE, 2002) - pode envolver a tarefa de identificação ou de verificação. Na *identificação de locutor* procede-se à comparação da amostra de fala de um indivíduo desconhecido com um grupo de amostras de fala pertencentes a locutores de identidade sabida, reunidas para fins de confronto ou pertencentes a um determinado banco de dados de produções orais, ainda inexistente de forma consistente no Brasil.

Na *verificação de locutor* são confrontadas propriedades de duas amostras, uma delas a amostra relativa ao locutor cuja identidade se questiona, obtida frequentemente através de interceptação telefônica, e a outra, a amostra relativa ao locutor de identidade conhecida, obtida pelos próprios peritos em procedimento de coleta técnica de padrão vocal. Tal contraponto situacional compreende o que atualmente é denominado *comparação de locutores* (doravante CL)⁸, cujo objetivo é definir a autoria de falas armazenadas em uma determinada mídia, avaliando-se se essas, de fato, foram produzidas (ou não) pelo aparelho fonador de um determinado indivíduo. Os perfis de voz e de linguagem verbal (na modalidade oral) do locutor questionado e do locutor de identidade sabida (padrão) são cotejados, identificando-se quais parâmetros, dos elencados como relevantes, indicam convergência ou divergência entre as amostras. O objetivo do exame é definir similaridade e tipicidade, ou seja, quão similar é o perfil de voz/fala do locutor da amostra questionada em relação ao do

⁸Para acesso à discussão mais aprofundada acerca da denominação envolvendo a referida verificação sugere-se consulta à Valente (2012).

locutor da amostra padrão e quão típico é o corpo probatório⁹, ou seja, qual é a probabilidade da evidência considerando-se o observado na população.

O método de examinação prevalentemente empregado no desenvolvimento da CL na perícia oficial brasileira está em consonância com o apregoado pela comunidade científica internacional. Segundo panorama apresentado por Gold e French (2011), baseado no depoimento de 36 peritos de 13 países distintos, há predomínio da utilização do *método combinado*, assim denominado por ser constituído das análises perceptivo-auditiva e acústica, o qual também é referido por Byrne e Foulkes (2004), Kuwabara e Sagisaka (1995), McDougall (2005), Nolan (2001), Rodman *et al.* (2002), Romito e Galatà (2004), Rose (2002) e Watt (2010). Complementarmente, a perícia oficial de alguns estados brasileiros considera ainda resultados provenientes de sistemas de reconhecimento automático de locutor.

Na análise perceptivo-auditiva são investigados elementos caracterizadores do comportamento de voz/fala dos locutores nas amostras. São observados fatores indicativos do sexo, da fase do ciclo de vida (se aparente infância, adolescência, fase adulta ou velhice), do estado de saúde dos órgãos fonoarticulatórios e da provável condição sociocultural e intelectual do falante. Atenta-se para vestígios, entre outros, relativos à qualidade vocal, aos ajustes musculares utilizados na fonação, à presença de hábitos vocais típicos (pigarro, estalos, cliques, etc), à forma de articulação, à existência de desvios fonéticos (distorções) e/ou fonológicos, a alterações temporais, de ritmo ou de fluência da fala, ao padrão entonacional empregado e ao nível de coordenação pneumofonoarticulatória. São consideradas, também, as variações de cunho diatópico e diastrático nos níveis fonético-fonológico, morfossintático, sintático, léxico-semântico e pragmático.

Já a análise acústica é empregada de duas formas, a saber, (a) na verificação de vestígios indicativos do efetivamente produzido, os quais são na maior parte das vezes detectados primeiramente de oitiva e que denotam o comportamento linguístico, paralinguístico e/ou extralinguístico dos falantes¹⁰ e (b) na extração de medidas, que contrastarão as amostras a partir de parâmetros físicos associados às configurações do aparelho fonador, tanto relativos à fonte quanto ao processo de filtragem. As informações são obtidas com a aplicação de recursos disponibilizados em *softwares* de edição e análise de áudio, sendo comumente utilizados como ferramentas oscilogramas (formas de onda); espectrogramas em banda larga e estreita de frequência; gráficos que ilustram a condição da f_0 , dos formantes e da intensidade em curto e longo termo em uma dada ocorrência ou em múltiplas ocorrências (curvas, histogramas, dispersões, etc), além de espectros FFT (*Fast Fourier Transform*), LPC (*Linear Predictive Coding*) e LTAS (*Long-term Average Spectrum*).

No cotejo entre as amostras podem ser consideradas sentenças, palavras, sílabas ou segmentos (fones), atentando-se para que os segmentos confrontados sejam pares quanto ao acento (ao menos lexical) e imersos em ambiente fonético antecedente e seguinte maximamente análogos.

Os resultados obtidos a partir das análises realizadas são apresentados tanto qualitativamente, especialmente no que se refere ao comportamento linguístico manifesto, quanto quantitativamente, por meio de estatística descritiva, restrita normalmente à exposição das medidas extraídas durante o exame acústico e à caracterização da diferença percentual existente entre as amostras questionada e padrão, no que se refere a um determinado parâmetro quantitativo (por exemplo, a frequência dos formantes). De uso mais recente, a

⁹Conjunto de elementos técnico-comparativos destacados na análise contrastiva das amostras.

¹⁰ Conforme Laver (1994, p.21) tratam-se, respectivamente, do comportamento comunicativo verbal; do comportamento comunicativo utilizado para veicular o estado afetivo, atitudinal e emocional do indivíduo (ou para regular o tempo compartilhado na conversação) e do comportamento não comunicativo, mas informativo, relacionado a atributos sociais, psicológicos e físicos, entre esses, a qualidade vocal.

razão de verossimilhança (*Likelihood Ratio*) constitui um novo paradigma na confecção e exposição dos resultados da perícia de CL (MORRISON, 2009; ROSE, 2002 e 2006).

Quanto à interpretação dos resultados, faz-se uso atualmente de escalas verbais, as quais são utilizadas tanto para indicação direta da conclusão, a exemplo da escala de nove pontos apresentada por Eriksson (2012), considerada neste estudo, a qual varia entre os extremos, a saber, os resultados suportam a hipótese com quase certeza e os resultados contrariam a hipótese com quase certeza, quanto para o enquadramento *a posteriori* de resultados quantitativos, como no caso da razão de verossimilhança, cuja classificação verbal tem associada uma escala de escores baseada no grau de suporte à hipótese de mesma origem.

Considerando a carência de estudos sobre o potencial discriminante de indivíduo das propriedades de voz/fala usualmente empregadas na perícia de CL em amostras de fala espontânea do português brasileiro, objetivou-se neste estudo verificar a suficiência de tais propriedades na diferenciação de indivíduos, ou seja, na determinação da autoria da manifestação oral. Com o propósito de testar a sensibilidade dos elementos técnico-comparativos dessa ordem, estabeleceu-se que entre os locutores a serem confrontados com o locutor da amostra linguística questionada estaria o irmão gêmeo do conhecido autor, conforme será detalhado a seguir, na seção 2, onde as informações referentes ao método adotado serão apresentadas. A condução do confronto entre as amostras, a partir do exame de aspectos referentes à qualidade de voz, à motricidade orofacial e aos aspectos suprasegmentais e sociolinguísticos da fala, são relatados na seção 3. Seguem-se conclusão e referências

2 METODOLOGIA

A metodologia empregada na pesquisa contemplou, ainda que parcialmente, o contexto de CL com fins forenses normalmente encontrado na perícia oficial dos estados¹¹. Quanto às amostras de fala utilizadas, buscaram-se gêmeos univitelinos do sexo masculino em uma comunidade acadêmica de Porto Alegre/RS. O estudo restringiu-se a investigar informantes do sexo masculino devido ao fato desse ser o sexo que até então predomina no contexto forense.

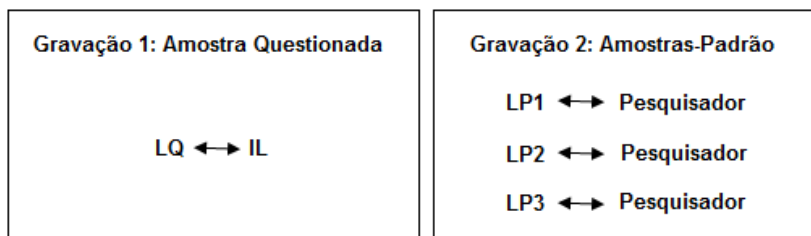
Os irmãos gêmeos selecionados para a pesquisa estavam com 24 anos e frequentavam o mesmo curso universitário à época da coleta de dados. Além disso, ambos viviam na região metropolitana de Porto Alegre-RS, na mesma comunidade em que nasceram e em que passaram a infância e a adolescência. Os dois outros participantes, um utilizado como interlocutor no áudio questionado e outro como um terceiro locutor-padrão, por vezes referido como padrão distrator, compartilhavam com os gêmeos características como sexo, idade, origem e escolaridade.

As gravações realizadas na coleta de dados geraram arquivos digitais de formato *Wave*, com taxa de amostragem de 44,1 kHz e resolução de 16 bits por amostra. Os equipamentos utilizados na captura ambiental dos áudios foram um gravador Marantz PMD661 e dois microfones Shure SM-58, acoplados a tripés de mesa posicionados a cerca de 20 cm de distância da boca dos locutores. A coleta compreendeu, conforme esquema apresentado na Figura 1, duas etapas distintas, denominadas Gravação 1 e Gravação 2.

¹¹ No contexto real de confronto forense de voz/ fala parte-se, via de regra, de uma amostra gravada sem a ciência do locutor. No presente estudo, todavia, todos os locutores tiveram ciência e autorizaram a gravação antes que essa iniciasse. A ressalva deve-se em razão de entender-se, amparados nas proposições de Labov (1972), que a ciência da gravação possibilita aos falantes a produção de uma fala mais cuidada, ou seja, menos vernacular.

Ressalta-se que as gravações somente foram iniciadas após a leitura e a assinatura, por parte dos participantes, do Termo de Consentimento Livre e Esclarecido¹².

Figura 1 - Etapas da coleta de dados



Legenda: LQ = Locutor Questionado; IL = Interlocutor; LP1 = Locutor Padrão 1 (LQ); LP2 = Locutor Padrão 2 (irmão gêmeo); LP3 (padrão distrator) = Locutor Padrão 3

Conforme esquematizado na Figura 1, a *amostra questionada* consiste na coleta de dados a partir da interação entre o locutor questionado (o irmão gêmeo escolhido como *alvo*) e um interlocutor conhecido seu. Já as *amostras-padrão* contam com dados de três participantes, a saber, do já referido LQ, de seu irmão gêmeo e de um locutor padrão utilizado como um distrator.

A interação que culminou no material questionado surgiu da resolução do problema intitulado *Decidindo para sobreviver: sobrevivência na lua*¹³, a partir do qual os participantes tiveram que realizar três tarefas, a saber: (a) classificar individualmente os itens segundo seu grau de importância; (b) consensualizar sobre qual classificação seria a mais adequada e (c) discutir e problematizar a classificação.

A escolha pela realização de uma atividade de natureza argumentativa entre indivíduos que apresentam algum grau de intimidade decorreu da necessidade de analisar dados de fala espontânea em contexto similar àquele verificado em diálogos em interceptações telefônicas, registro esse que normalmente é o objeto da CL de âmbito oficial. Já a eliciação da fala padrão considerou um roteiro de entrevista de experiência pessoal, com duração de aproximadamente 30 minutos, e a repetição de sentenças previamente selecionadas¹⁴ do conteúdo da fala de LQ.

As gravações foram conduzidas por um mesmo pesquisador, que não participou do processo de análise dos áudios, mantendo, portanto, total sigilo com relação à identidade do locutor questionado. As coletas ocorreram através de interação entre o pesquisador e cada um dos três participantes (doravante LP1, LP2 e LP3, conforme especificado na Figura 1) individualmente. Na entrevista de experiência pessoal foram feitas perguntas relacionadas à vida dos participantes (infância, escola, família, trabalho, etc), de forma a propiciar o envolvimento emocional com os fatos narrados, conforme orientação da entrevista sociolinguística (LABOV, 1972).

¹² Esta pesquisa foi aprovada pelo Comitê de Ética em Pesquisa da PUCRS em 24/04/2014 (CAAE 26712014.4.0000.5336).

¹³ Disponível em <http://www.facom.ufu.br/~ronaldoliveira/ASI-2012-1/Sobreviv%EAcia%20na%20Lua.pdf>.

¹⁴ Foram selecionadas dezoito sentenças da fala de LQ. Sentenças que apresentaram elementos não linguísticos, a exemplo, risadas e suspiros, bem como disfluências e alongamentos não foram consideradas.

A análise contrastiva entre a amostra de LQ e de cada um dos padrões valeu-se do *método combinado*, conforme já mencionado, sendo os resultados qualitativos e quantitativos referentes aos três confrontos apresentados e discutidos a seguir.

A conclusão foi expressa através do uso da escala, proposta por Eriksson (2012) e apresentada na Figura 2, destacando-se que a definição do grau de suporte/contradição à hipótese de mesma origem considera o corpo probatório e não qualquer dos indícios isoladamente.

Figura 2 – Escala utilizada na conclusão das três CLs do estudo (ERIKSSON, 2012)

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

Assim, neste estudo, a conclusão do confronto, considerando-se as evidências disponíveis, será uma afirmação correspondente ao nível mais representativo da escala exibida na Figura 2, composta de expressões verbais que denotam da inconclusão (nível 0) a quatro níveis de suporte (níveis +1 a +4) ou contradição (níveis -1 a -4) à hipótese de que as amostras de fala cotejadas foram proferidas por um mesmo locutor.

3 APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS

Considerando os elementos relacionados à qualidade vocal, ao emprego de suprassegmentos, à motricidade orofacial e às características do estado estável de desenvolvimento de fala, observa-se que os três locutores-padrão analisados (LP1, LP2 e LP3) igualmente se equiparam à LQ no que se refere aos seguintes aspectos: voz atribuível a falante do sexo masculino; adulto jovem, sem os indicativos básicos de patologia vocal, ao menos em nível significativo, como rugosidade, sopro, astenia e tensão; modo de fonação modal e ressonância equilibrada. Dessa forma, logra-se que tais parâmetros não apontam dentre os três locutores-padrão qual seria o autor da *amostra questionada*. Por outro lado, foram encontradas convergências e divergências, apontadas na Figura 3, a partir de parâmetros articulatórios estabelecidos com base em Laver (1994), que subsidiam a formulação de três hipóteses de trabalho, a saber: (1) há divergência entre o perfil de voz de

LQ e LP3; (2) há divergência entre o perfil de voz de LQ e de LP1, embora haja convergências e (3) LP2 é o provável autor da *amostra questionada*, ou seja, LP2 e LQ são o mesmo indivíduo.

Figura 3 – Principais Convergência e Divergências de Parâmetros entre as Amostras

Vestígios	Locutores			
	LQ	LP1	LP2	LP3
Tensão muscular do trato vocal	Não observada	Não observada	Não observada	Hiperfunção, com contenção da movimentação vertical de laringe
Ocorrências vocais de curto-termo	Laringalizações em final de sentenças	Não observadas	Laringalização variável em final de palavra	Não observadas
Respiração	Evidente, com soltura de ar após o final da fala	Não evidente	Evidente, com expiração ruidosa	Não evidente
Continuidade	Sem elementos destacáveis	Ritmo de fala bloqueado, com recorrência de eventos de disfluência (entre eles bloqueios)	Sem elementos destacáveis	Uso recorrente de alongamentos em pausa preenchida sugerindo planejamento da fala ou hesitação
Taxa de articulação (TA)	Tendente à normal com ocorrências de TAs baixas	Tendente à normal com ocorrências de TAs altas	Tendente à normal com ocorrências de TAs baixas	Tendente rápida
Extensão de (movimentação) de articuladores	Extensão vertical diminuída de ponta/ lâmina de língua	Sem elementos destacáveis	Extensão vertical diminuída de ponta/ lâmina de língua	Sem elementos destacáveis
Estado particular dos articuladores	Ajuste de lábio diferenciado (estirados) e ponta da língua avançada	Sem elementos destacáveis	Ajuste de lábio diferenciado (estirados) e ponta da língua avançada	Sem elementos destacáveis
Desvios de fala	Ceceo anterior/	Não	Ceceo	Não

	aparente não aproximação da região anterior de língua em direção ao palato	observados	anterior/ aparente não aproximação da região anterior de língua em direção ao palato	observados
--	--	------------	--	------------

Com relação à primeira hipótese, a de divergência no perfil de voz de LQ e LP3, tem-se, como observado na Figura 3, o fato de LP3, diferentemente de LQ, apresentar durante a emissão de fala: a) dinâmica vocal associada à contenção da movimentação vertical de laringe; b) recorrentes alongamentos segmentais em contextos que sugerem planejamento da fala ou hesitação; c) taxa de articulação tendente à rápida (situação adiante devidamente documentada e, especialmente, d) inevidência da condição miofuncional diferenciada de língua durante a realização do gesto de segmentos alveolares, especificamente / l, r, s, z /.

Fundamentando a segunda hipótese, a de divergência entre o perfil de voz de LQ e de LP1, ressalvada a presença de elementos concordantes entre as amostras, mas que, devido ao baixo poder discriminante, apenas atenuam o grau de convencimento acerca da disparidade, tem-se, como observado na Figura 3: a) a não observância de laringalizações pareáveis com as manifestadas por LQ e de ruído expiratório após a conclusão das emissões e, principalmente; b) a frequência e a espécie de eventos de disfluência, que configuram um ritmo bloqueado de fala, constituído de falsos começos/reparos, repetições, bloqueios e alongamentos não previsíveis, e a inevidência da condição miofuncional diferenciada caracterizada; c) tendência à taxa de articulação de normal à alta e d) a inevidência da condição miofuncional diferenciada caracterizada.

Relativamente à terceira hipótese, a qual indica LP2 como provável autor da *amostra questionada*, observam-se as seguintes características como deladoras da unicidade de origem, conforme exposto na Figura 3: a) laringalizações em finais de emissão; b) soltura de ar após a finalização da emissão, evidente à oitiva e cujo vestígio consta no sinal acústico, e, especialmente; c) ocorrência de valores baixos de taxa de articulação e, especialmente, d) condição miofuncional de língua que leva a alterações de fala (distorções), especificamente ao ceceo anterior com comprometimento das formas de veiculação de /s z/, à dorsalização do tepe alveolar e à imprecisão de ponto e modo articulatorios da lateral alveolar. Destaca-se que a característica descrita em “d” é compatível com a dificuldade para elevação/posicionamento da ponta da língua nos alvéolos por parte de LP2 à ocasião da coleta de padrão, dificuldade relatada pelo pesquisador responsável pela coleta dos dados.

Com relação ao uso do tempo de fala por parte dos locutores tem-se na Figura 4 os valores de taxa de articulação, obtidos em dez intervalos interpausais¹⁵ da fala espontânea de cada um dos quatro locutores envolvidos nos confrontos. Foram admitidos para fins de cálculo exclusivamente intervalos sem pausas preenchidas e/ ou eventos de disfluência (falsos começos/reparos, repetições, bloqueios e alongamentos não associáveis à ênfase). As taxas apresentadas na figura referem-se ao resultado da divisão do número de sílabas fonéticas, identificadas manualmente através da inspeção do oscilograma e espectrograma

¹⁵ Trecho de fala compreendido entre o silêncio do início de emissão e uma pausa silenciosa (com duração igual ou maior à 130ms) ou entre duas pausas silenciosas.

correspondente, pela duração do respectivo intervalo de fala. Em seguida, realizou-se a média aritmética das TA locais para confecção da TA local média dos locutores.

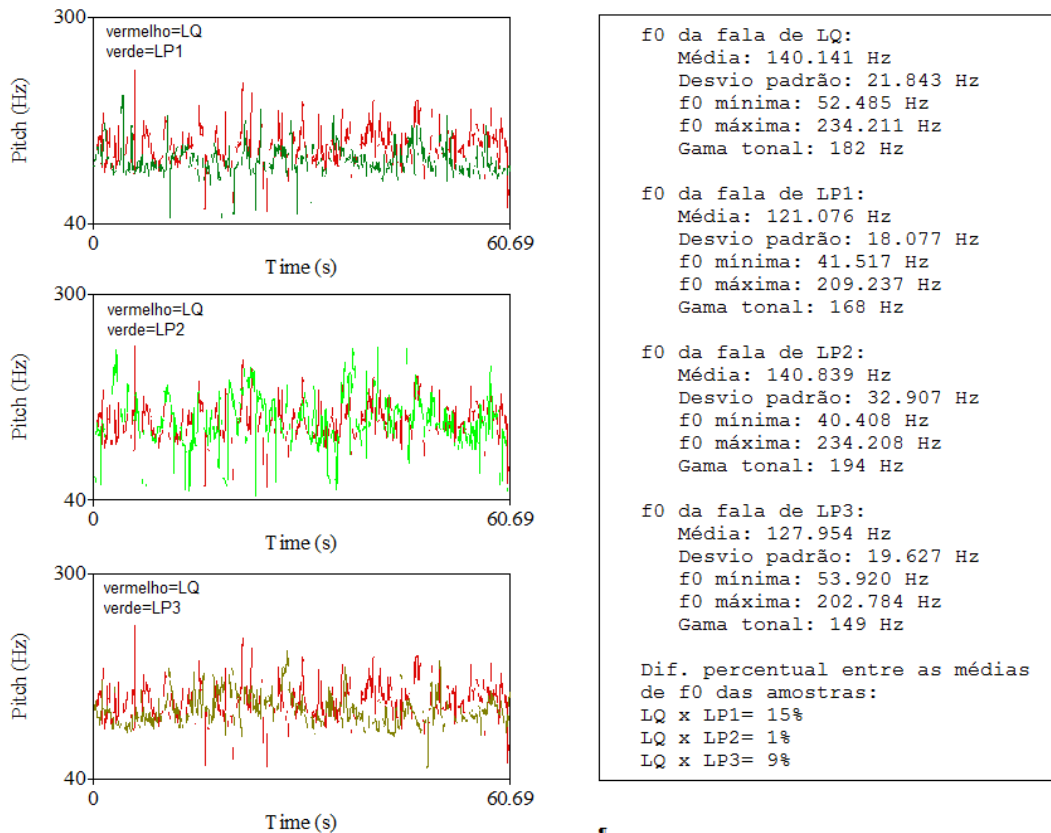
Figura 4 – Taxa de articulação (TA) local média de LQ, LP1, LP2 e LP3

	LQ	LP2	LP1	LP3
TAs locais (em sílabas/ segundo)	4.77	4.53	5.59	6.17
	4.87	4.76	5.85	6.2
	5.97	5.07	6.53	6.28
	6.12	5.82	6.76	6.3
	6.68	6	7.17	8
	6.81	6.56	7.31	8.39
	7.18	6.63	7.88	8.92
	7.52	7.34	8.05	9.06
	7.55	7.72	8.42	9.43
	7.97	8.99	9.11	10.37
TA média	6.544	6.342	7.267	7.912

A análise dos dados expostos na Figura 4 revela que apenas LP2 tem, assim como LQ, TAs inferiores à 5 sílabas/segundo. Esses locutores convergem, ainda, quanto à prevalência de taxas entre 6-7 sílabas/segundo e à TA local média próxima à 6.5 sílabas/segundo. Já contrastando os tempos de fala de LQ com os dos outros dois locutores percebe-se que: a) nos intervalos de LP3 analisados há predomínio do emprego de TAs mais altas e inobservância de valores inferiores à 6 sílabas/segundo e b) nos intervalos de LP1 analisados prevalecem as TAs médias e altas.

Quanto à variação de f_0 em intervalo com duração aproximada de 60 segundos, composto por justaposição de trechos de fala exclusiva, verificou-se o valor médio, o desvio padrão (indicativo da variabilidade intrassujeito na amostra), a alocação dos valores mínimo e máximo e, ainda, a gama tonal (intervalo entre o valor mínimo e máximo de f_0) nas amostras dos quatro locutores. Na Figura 5 a seguir, são exibidos, à esquerda, os gráficos com as sobreposições das curvas de variação de f_0 nos trechos de fala questionado (traçado em vermelho) e padrão, respectivamente de cima para baixo, de LP1, LP2 e LP3. Nos gráficos a escala vertical representa os valores de frequência e a escala horizontal os temporais. No quadro maior, à direita, são exibidos os valores das medidas referidas para os quatro locutores e a diferença percentual entre as médias de LQ e cada um dos padrões. Destaca-se que o resultado do confronto do comportamento do parâmetro nas amostras indica o já observado qualitativamente, qual seja, a existência de divergência entre LQ e LP3, o locutor-padrão que mostra uma gama tonal restrita se comparado às largas excursões produzidas por LP1 e especialmente por LP2, e de convergência entre LQ e LP2, confronto cuja diferença percentual entre as médias não ultrapassa 1%.

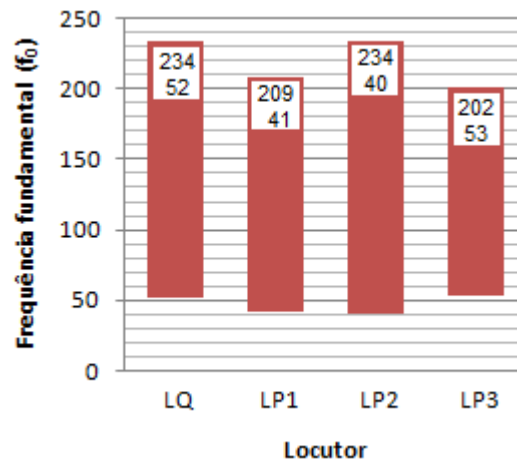
Figura 5 – Análise da f_0 em Longo-Termo



Legenda: Vermelho = LQ, Tons de verde = LP1, 2 e 3 (sendo o verde-limão referente ao LP2)

A percepção da paridade entre LQ e LP2, no que se refere ao uso da f_0 na fala, pode ser percebida, ainda, ao serem observados os valores relativos à dimensão e alocação da gama tonal no eixo da frequência, ilustrados da Figura 6.

Figura 6 – Dimensão e Alocação da Gama Tonal



Nota-se que os valores de f_0 máxima, exibidos no topo de cada coluna, distanciam-se do de LQ no caso de LP1 (209 Hz) e de LP2 (202 Hz), mas coincide com o de LQ no caso de LP2 (234 Hz). Apesar de apresentar f_0 mínima mais baixa, e por consequência uma dimensão de gama tonal um pouco maior se comparada com a de LQ, LP2 é o locutor-padrão com valor de gama tonal mais próximo do de LQ (LP2= 12Hz, LP1= 14 Hz e LP3=33 Hz).

Considerando os elementos relacionados ao comportamento linguístico manifesto pelos locutores das amostras em trechos de áudio com duração de 00:10:42 (LQ), 00:11:47 (LP1), 00:11:34 (LP2) e 00:12:24 (LP3), especificamente quanto à aplicação de processos variáveis, verificou-se que LQ1, LQ2 e LQ3 não se diferenciam em relação ao obtido na análise da *amostra questionada*. Os processos variáveis identificados são produzidos nos trechos de áudio considerados com taxas típicas da variedade porto-alegrense do português brasileiro, atestadas em Bisol e Brescancini (2002), Bisol e Collischonn (2010) e Bisol e Battisti (2014), um resultado previsível diante do fato de os sujeitos da pesquisa compartilharem as mesmas características sócio-culturais.

Atestaram-se, desse modo, variação moderada entre vogais altas e médias-altas em posição pré-tônica em casos de harmonização vocálica (ex.: fós[f]o~ro ~ fós[f]u~ro; t[e]cido ~ t[i]cido); pequena variação entre vogais altas e médias-altas em posição pré-tônica em casos de alçamento sem motivação aparente¹⁶(ex.: c[o]légio ~ c[u]légio, mas c[o]lega); produção praticamente categórica de vogais altas em posição pós-tônica final (ex.: semestr[i]; cab[u]); ausência de palatalização da fricativa sibilante em coda (ex.: e[s]telar; de[z]de); vocalização praticamente categórica da lateral em coda (ex.: portáti[w]); redução variável do ditongo decrescente diante de tepe (ex.: terc[e]ro para *terceiro*); ditongação variável diante de fricativa em coda (ex.: m[a]js para *mais*) e redução variável do ditongo nasal (ex.: falar[u] para *falaram*).

A variação envolvendo a produção de róticos, por outro lado, especialmente a verificada em posição de coda, mostrou-se pertinente na tarefa de pareamento das amostras, conforme apresenta a Figura 7 a seguir, onde são computadas as ocorrências para a posição de onset em início de palavra, coda medial e coda final em não-verbos.

Figura 7¹⁷ – Produção de Róticos: número de ocorrências por posição

LQ		☼	☼8	□	↖ / ↗	☒ □ † □ ≍ □ ♪	✕ □ ☼	≍ □ ♪	∅
	N								
Onset inicial	15					15			
Coda medial	30	13		4	9	1			3
Coda final	20	8		6	6				
LP3		☼	☼8	□	↖ □ □ □ ↗	☒ □ † □ ≍ □ ♪	✕ □ ☼	≍ □ ♪	∅
	N								
Onset inicial	11					10	1		
Coda	64	64							

¹⁶ Sobre o processo de alçamento sem motivação aparente, ver Bisol (2010).

¹⁷ Não foram computadas as ocorrências em contextos em que o rótico não apresentou variação nas amostras, a saber, posição intervocálica (como em *arame* e *carro*); final de palavra seguida por vogal (como em *compra*[☼ *o*] *carro*); final de verbo no infinitivo (como em *compra*∅ *o cabo*) e onset complexo (como em *p*[☼ *ato*]).

medial									
Coda final	2	2							
LP2		☸	☸8	□	↶□□□↷	☒□†□☞□♯	✱□☸	☞□♯	∅
	N								
Onset inicial	14					14			
Coda medial	54	26			22				6
Coda final	10	4			3			2	
LP1		☸	☸8	□	↶□□□↷	☒□†□☞□♯	✱□☸	☞□♯	∅
	N								
Onset inicial	13					13			
Coda medial	56	37		9					10
Coda final	3	2		1					

De acordo com o exposto na Figura 7 anterior, em onset inicial a produção posterior (fricativa velar ([☒□†]) ou glotal ([☞□♯]) em *receptor*, por exemplo) é predominante em todas as amostras, impossibilitando assim qualquer conclusão a respeito da identificação de amostras equivalentes. Em coda medial (ex.: *arma*), embora haja predominância de [☸] em todas as amostras, a aproximante alveolar ([□]) foi verificada também em LQ (4 ocorrências) e em LP1 (9 ocorrências) e a variante retroflexa ([↶□□□↷]) em LQ (9 ocorrências) e LP2 (22 ocorrências). O apagamento de coda medial foi verificado em LQ, LP1 e LP2, corroborando a hipótese de maior semelhança entre essas amostras.

A posição de coda final (ex.: *comprar*), por sua vez, preferencialmente produzida com tepe pelos sujeitos investigados, apresenta variação com a produção retroflexa apenas em LQ (6 ocorrências) e LP2 (3 ocorrências). Adicionalmente, o cômputo apresentado na Figura 7 revela a presença da variante aproximante em posição de coda (medial e final) apenas em LQ (10 ocorrências) e LP1 (10 ocorrências).

O exposto corrobora, desse modo, a hipótese aventada a partir dos resultados referentes às divergências entre os perfis de voz e aspectos suprasegmentais, segundo a qual LQ e LP2 são um mesmo falante e que, pela proximidade dos resultados obtidos, LP1 é o irmão gêmeo de LP2 e que LP3 é o locutor-distrator, já que apresenta apenas a produção de tepe em coda final.

Tais hipóteses recebem ainda o apoio nas amostras consideradas do comportamento de outro processo variável característico da variedade porto-alegrense do português brasileiro (KAMIANECKY, 2002), a palatalização de /t,d/ (ex.: sé[t♠]imo; [d♠]inheiro). Nas amostras referentes tanto a LP1 e LP2 quanto a LQ, a palatalização de /t, d/ é bastante frequente, diferentemente do que ocorre na amostra referente a LP3, em que a palatalização de /t/ é mais frequente do que a de /d/ devido à produção predominantemente alveolar da preposição *de* em sintagmas como *de repente*, *de esconde-esconde*, *de skate* e *de esporte*.

BYRNE, C.; FOULKES, P. The 'mobile phone effect' on vowel formants. **Speech, Language and the Law**, v.11, n.1, p.83-102, 2004.

ERIKSSON, A. Aural/acoustical vs. automatic methods in forensic phonetic case work. In: NEUSTEIN, A.; PATIL, H.A. **Forensic speaker recognition: law enforcement and counter-terrorism**. New York: Springer-Verlag, 2012. p.41-69.

GOLD, E.; FRENCH, P. International practices in forensic speaker comparison. **International Journal of Speech, Language and the Law**, v. 18, p. 293-307, 2011.

GONÇALVES, C.S.; PETRY, T. Comparação forense de locutores no âmbito da perícia oficial dos estados. In: REHDER, M.I.; CAZUMBÁ, L.A.F.; CAZUMBÁ, M.A. (Orgs.). **Identificação de falantes: uma introdução à Fonoaudiologia Forense**. Rio de Janeiro: Revinter, 2014. p.241-264.

HOLLIEN, H. **Forensic voice identification**. London: Academic Press, 2002.

KAMIANECKY, F.A. **A palatalização das oclusivas dentais /t/ e /d/ nas comunidades de Porto Alegre-RS e Florianópolis-SC: uma análise quantitativa**. 2002. Dissertação (Mestrado em Letras) – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre-RS.

KUWABARA, H.; SAGISAKA, Y. Acoustic characteristics of speaker individuality: control and conversion. **Speech Communication**, n.16, p.165-173, 1995.

LABOV, W. **Sociolinguistic patterns**. Philadelphia: University of Pennsylvania Press, 1972.

LAVER, J. **Principles of Phonetics**. Cambridge: Cambridge University Press, 1994.

McDOUGALL, K. **The role of formant dynamics in determining speaker identity**. 2005. Tese (Doctoral Dissertation) – University of Cambridge, Cambridge, UK.

MORRISON, G.S. Forensic voice comparison and the paradigm shift. **Science and Justice**, v.49, p.298-308, 2009.

NOLAN, F. **The phonetic bases of speaker recognition**. Cambridge: Cambridge University Press, 1983.

NOLAN, F. Speaker identification evidence: its forms, limitations, and roles. **Proceedings of the conference 'Law and language: prospect and retrospect'**, Levi (Finnish Lapland), p. 12-15, dez. 2001

ROMITO, L.; GALATÀ, V. Towards a protocol in speaker recognition analysis. **Forensic Science International**, v. 146S, p. S107-S111, 2004.

ROSE, P. **Forensic speaker identification**. London: Taylor & Francis, 2002.

ROSE, P. Technical forensic speaker recognition: evaluation, types and testing of evidence. **Computer Speech & Language**, v.20, n.2-3, p.159-191, 2006.

VALENTE, C.R. Perspectivas da fonética forense num cenário de quebra do dogma da unicidade. **Anais da Conferência Internacional de Ciências Forenses em Multimídia e Segurança Eletrônica**, Brasília/DF, Brasil, p.7-27, 2012.

WATT, D. The identification of the individual through speech. In: LLAMAS, C.; WATT, D. **Language and identities**. Edinburgh: Edinburgh University Press, 2010. p.76-85.

Recebido em: 24 de abril de 2016.
Aceito em: 28 de setembro de 2016.