

LINGÜÍSTICA

Na escuta contra o crime



Especialistas em fonética forense analisam vozes e modos de falar para auxiliar a investigação de delitos

MÔNICA MANIR — ilustrações JULIA JABUR

Ramo da linguística dedicado ao estudo dos sons da fala humana, a fonética tem uma subárea que vem ganhando espaço não apenas na academia como também entre peritos criminais: a fonética forense. Trata-se da aplicação de técnicas de análise de fala para caracterizar quem está falando em gravações obtidas pela polícia e, a partir disso, utilizar essas informações como evidência em investigações. A possibilidade de empregar conhecimento científico em uma prática concreta atrai não apenas especialistas do próprio campo da linguística, como também de outras áreas, como engenharia. “É importante para nós, linguistas, ter esse olhar voltado à aplicação dos nossos estudos na solução de demandas sociais”, afirma Cláudia Regina Brescancini, do Programa de Pós-graduação em Letras da Pontifícia Universidade Católica do Rio Grande do Sul (PUC-RS).

A “comparação de locutor” é a aplicação mais tradicional da fonética forense. O método consiste em confrontar uma amostra de autoria desconhecida – chamada de questionada, como um áudio de escuta telefônica – com a voz de um ou mais suspeitos por um crime, denominada amostra de referência ou padrão. Com base em estudos de anatomia do trato vocal, da análise acústica e de informações sobre o contexto em que vivem os indivíduos, é possível avaliar a probabilidade de que as duas amostras pertençam, ou não, à mesma pessoa.

Como sociolinguista, Brescancini se dedica ao estudo das variações sonoras, ou seja, variações de pronúncia que não afetam o significado das palavras ou frases e que são motivadas por fatores tanto linguísticos quanto sociais. Uma dessas questões é a geográfica, relacionada às diferentes colonizações de uma região. Nesse cenário, o português falado no Brasil representa um desafio para a fonética forense. “Nosso país é um baú de sotaques”, constata a pesquisadora. Ela destaca o “s” chiado, característico da influência portuguesa e açoriana em cidades como Rio de Janeiro, Recife e Florianópolis, e o chamado “r” caipira, típico do interior paulista, do sul de Minas Gerais e de Goiás.

Porém as variações sonoras não são determinadas apenas pelo local onde se vive. Pessoas com maior nível de escolaridade, por exemplo, tendem a apresentar padrões linguísticos distintos daquelas com menor escolaridade. Há ainda diferenças associadas ao gênero. Em geral, a voz masculina apresenta mais nasalidade e crepitação, enquanto a feminina tende à soprosonidade, à suavidade. Além disso, muitas pessoas circulam pelo país incorporando sotaques. “A migração traz um tempero a mais nessa variação”, afirma Brescancini. No caso de suspeitos por crimes, o histórico de confinamento por longos períodos em penitenciárias acrescenta outra complexidade: o contato de diferentes dialetos em razão do convívio entre presos de origens diversas.

Além da similaridade, os peritos precisam considerar a tipicidade: o grau de raridade de um determinado traço linguístico. Mesmo quando há forte semelhança entre os áudios comparados, se as características observadas no material questionado forem comuns na população à qual se supõe que o suspeito pertença, o valor probatório da evidência diminui. Brescancini é coautora de um artigo sobre tipicidade publicado em 2024 no periódico *Cadernos de Estudos Linguísticos*, da Universidade Estadual de Campinas (Unicamp). “Na comparação de locutor, é preciso avaliar em que medida similaridade e tipicidade sustentam a hipótese de que duas amostras provenham do mesmo aparelho fonador, ou seja, da mesma pessoa”, diz. “A resposta será sempre em termos de probabilidade.”

Diante desse cenário, Brescancini e a perita criminal Cíntia Schivinski Gonçalves, doutora em linguística pela PUC-RS, vêm desenvolvendo um método voltado a auxiliar peritos oficiais no tratamento das variações sonoras. A primeira versão da proposta virou um dos capítulos do livro *Análise fonético-forense: Em tarefa de comparação de locutor* (Editora Millennium, 2020), coorganizado pelo linguista Plínio Barbosa, do Instituto de Estudos da Linguagem da Unicamp, onde coordena o Grupo de Estudos de Fonética Forense (Geff).

Fruto da parceria entre o Geff e a Escola Superior do Ministério Público de São Paulo, a obra é resultado de projeto realizado por Barbosa entre 2016 e 2018 na Unicamp com apoio da FAPESP. Primeiro, o livro apresenta os procedimentos para o ensino sistemático e a aplicação da análise de comparação de locutor no contexto forense. Na sequência, pesquisadores detalham técnicas fonético-acústicas ou sociolinguísticas relacionadas a esse tipo de comparação.

Para Barbosa, o principal desafio da fonética forense enquanto ciência é a busca por parâmetros de comparação mais robustos. Ele ressalta que houve avanço na área com a popularização de áudios obtidos por aplicativos de mensagem, como o WhatsApp. “Antigamente, a maioria das amostras vinha da telefonia fixa ou do celular, cujas bandas alcançam no máximo 3.400 Hertz (Hz) de frequência”, conta. Esse limite dificultava a análise de segmentos que podem ser diferenciadores de indivíduos, como a pronúncia do “s”, que ultrapassa essa faixa. A frequência do WhatsApp, de até 8 mil Hz, ampliou o espectro da investigação.

O efeito da transmissão telefônica na fala foi objeto de estudo da linguista Renata Passetti tanto no mestrado (2015) quanto no doutorado (2018) realizados na Unicamp, ambos com apoio da FAPESP. “Quando nos comunicamos pelo telefone, adotamos inconscientemente algumas modificações acústicas e prosódicas que caracterizam esse tipo de interação”, afirma a pesquisadora. Prosódia é o componente da linguagem oral relacionado à sua musicalidade, que abrange aspectos como ritmo e acentuação, ajudando a transmitir emoções e intenções.

Passetti também investigou bancos de dados de língua portuguesa no Brasil, destacando seu potencial como apoio para a comparação de locutor na avaliação da tipicidade de traços linguísticos. Esses acervos reúnem tanto registros brutos de fala do português brasileiro quanto atlas linguísticos, nos quais os dados já foram sistematizados e analisados. O levantamento desse material é realizado principalmente por linguistas, dialetólogos e geolinguistas vinculados a centros de pesquisa.

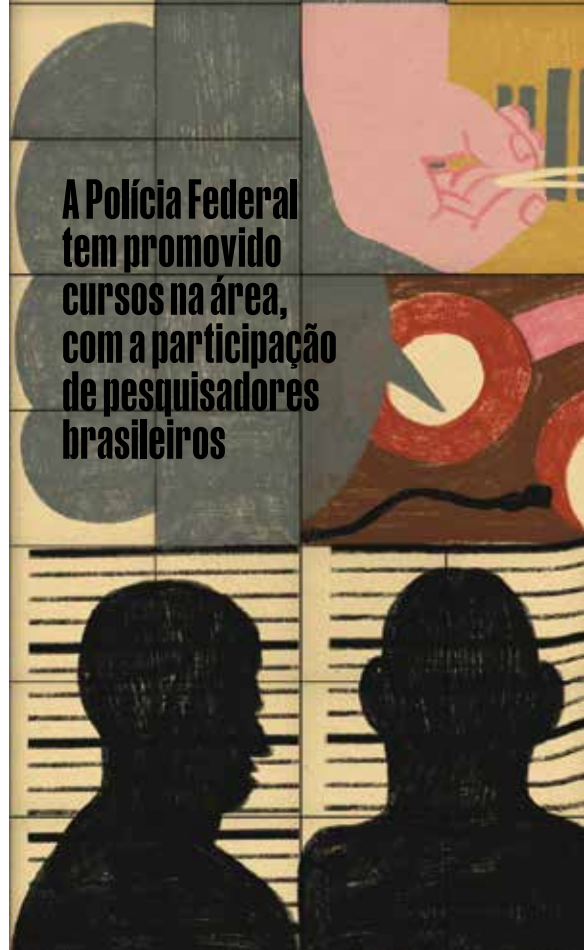
A pesquisadora assina com os linguistas Pablo Arantes e Daniel Fonseca Vieira, ambos da Universidade Federal de São Carlos (UFSCar), um artigo publicado em 2025 no qual avaliam acervos que registram variações em diferentes níveis da língua: o lexical (diferenças no vocabulário usado para expressar o mesmo conceito, sem alteração de sentido), o sintático (variações na construção, organização e ordem das palavras em frases e

orações) e o conversacional (a adaptação da língua ao contexto, ao falante e ao grupo social). “Dos 45 acervos analisados, apenas 13 possuem áudios, material essencial para a fonética forense, e 46% dessa parcela estão concentrados no Sudeste”, informa Arantes, coordenador do Laboratório de Fonética da UFSCar. “Não identificamos acervos com amostras de áudio provenientes de falantes das regiões Centro-Oeste e Norte.”

Nos Países Baixos, um dos centros de referência mundial em fonética forense, há um conhecimento acústico avançado dos diferentes modos de falar, resultado sobretudo do trabalho do Instituto Forense dos Países Baixos (NFI, na sigla em inglês), em Haia, vinculado ao Ministério da Segurança e Justiça. A Suécia também abriga um centro de excelência: o Instituto Real de Tecnologia (KTH), em Estocolmo, que vem avançando no reconhecimento de falantes com o uso de inteligência artificial (IA).

É no KTH que Julio Cesar Cavalcanti realiza atualmente estágio de pós-doutorado, com apoio da FAPESP, como desdobramento de sua pesquisa na Pontifícia Universidade Católica de São Paulo (PUC-SP). Com doutorado em linguística pela Unicamp, defendido em 2021, ele tem analisado a transparência dos modelos de IA aplicados à identificação de autoria de voz. “Há um problema ético em permitir que uma máquina ou um modelo de IA incrimine um ser humano sem que se compreendam exatamente os critérios utilizados para chegar a essa conclusão”, alerta.

Em artigo publicado em 2024 no periódico *Frontiers in Artificial Intelligence*, Cavalcanti, Barbosa, o linguista sueco Anders Eriksson, da Universidade de Estocolmo, e o perito da Polícia Federal (PF) Ronaldo Rodrigues da Silva analisaram gravações de 20 pares de gêmeos idênticos do sexo masculino em diálogos e entrevistas espontâneas. Os resultados evidenciam o desafio significativo que gêmeos idênticos impõem aos sistemas automáticos de reconhecimento de falantes, uma tecnologia de IA voltada à comparação de vozes com base em características acústicas individuais. Segundo o estudo, vozes humanas carregam mar-



A Polícia Federal tem promovido cursos na área, com a participação de pesquisadores brasileiros



cas individualizantes importantes, mas o caso dos gêmeos idênticos mostra que essa individualidade não deve ser tratada como absoluta. “Há diferentes graus de semelhança vocal, e algumas vozes podem ser especialmente difíceis de distinguir até mesmo pela IA”, comenta Cavalcanti.

Por causa de desafios assim, desde 2007, a PF oferece cursos de capacitação em comparação de locutor, com participação de universidades brasileiras, entre elas Unicamp, PUC-SP, PUC-RS, UFSCar e Universidade de Brasília (UnB). Até o momento, foram realizados cinco ciclos de formação, com intervalos médios de dois anos entre eles, ministrados no Instituto Nacional de Criminalística (INC), em Brasília, com a participação do Serviço de Perícias em Audiovisual e Eletrônicos do próprio instituto.

Os cursos são voltados aos peritos da PF e também aos que atuam nos institutos de criminalística dos governos estaduais e do Distrito Federal. O mais recente ocorreu entre 2021 e 2023. Nessas formações, os profissionais aprendem a analisar registros de áudio por meio de softwares de análise acústica, como o Praat (“falar”, em holandês), e, em uma etapa posterior, recebem instruções de pesquisadores sobre linguística, fonética, sociolinguística e etiquetagem, processo que consiste em converter o sinal sonoro em dados anotados, permitindo comparações mais sistemáticas.

Além disso, os peritos analisam a qualidade vocal, tema central das pesquisas da linguista Sandra Madureira, do Programa de Pós-graduação em Linguística Aplicada e Estudos da Linguagem da PUC-SP. “A singularidade de cada voz não resulta apenas da vibração das pregas vocais, mas também da maneira recorrente como posicionamos os órgãos articulatórios da fala, como os lábios”, explica Madureira, uma das líderes do Grupo de Pesquisa em Estudos sobre a Fala, da mesma instituição. “A voz é influenciada ainda por fatores fisiológicos, como o prognatismo [projeção da mandíbula, por exemplo], e por hábitos, como o tabagismo.”

Madureira, Barbosa, Passetti, Arantes e Brescancini lecionaram em cursos da PF, mas, por força da legislação, não poderiam ingressar na carreira de perito criminal federal, caso desejassem. O Decreto nº 5.116, de 24 de junho de 2004, não inclui linguistas e fonoaudiólogos entre as formações habilitadas para o cargo.

Segundo o engenheiro elétrico André Luiz da Costa Morisson, chefe do Serviço de Perícias em Audiovisual e Eletrônicos da PF, seis peritos federais especializados em fonética forense, todos com formação em engenharia, compõem hoje a equipe do órgão. “Produzimos cerca de 50 laudos por ano”, contabiliza. O espectro de apurações abrange desde crimes eleitorais e tráfico de drogas a casos de abuso sexual infantil e difamação. “Nossas conclusões não se baseiam em resultados categóricos, mas em avaliações de natureza probabilística”, diz.

Morisson explica que todo resultado de perícia apresenta uma escala de probabilidade de nove pontos, que vai do -4 (quando é muitíssimo mais provável que a evidência seja obtida caso as amostras padrão e questionada pertençam a pessoas diferentes) ao +4 (quando é muitíssimo mais provável que a evidência seja obtida caso as amostras pertençam à mesma pessoa). O zero seria inconclusivo.

Nesse sentido, os pesquisadores ouvidos para esta reportagem ressaltam de forma unânime que, apesar dos avanços na precisão da técnica, os resultados da fonética forense devem ser interpretados como parte de um quadro mais amplo de evidências. “A comparação de locutor é um elemento adicional na investigação, não o único. É fundamental considerar outros vestígios, que, em conjunto com a voz, ajudam a compor o contexto do crime”, conclui Arantes. ●

Os projetos, os artigos científicos e o livro consultados para esta reportagem estão listados na versão on-line.