

Considerações sobre o ensino da Fonética Forense para estudantes de Letras e Linguística

DOI: <http://dx.doi.org/10.21165/el.v50i1.3109>

Renata Regina Passetti¹

Pablo Arantes²

Resumo

Este artigo é um relato do curso introdutório “Iniciação à Fonética Forense” ministrado pelos autores na I Escola de Estudos Linguísticos do GEL. No artigo, apresentamos os conteúdos trabalhados no curso, as adaptações feitas para adequá-lo à modalidade remota e as atividades propostas aos participantes. Considerações sobre a formação típica dos participantes, feitas a partir das respostas dadas às atividades do curso, são associadas aos conhecimentos necessários para atuar na área de Fonética Forense. A formação em Letras e/ou Linguística torna estudantes e pesquisadores dessas áreas sensíveis para fenômenos linguísticos e os equipa com vocabulário técnico fundamental para a pesquisa e o trabalho nessa área. Como forma de melhor preparar egressos desses cursos para atuação em Fonética Forense, consideramos essencial ampliar o espaço dado a disciplinas de Fonética e Fonologia, bem como valorizar a temática da variação linguística na grade curricular desses cursos.

Palavras-chave: ensino de Fonética Forense; formação em Letras e Linguística; relato de experiência.

1 Pontifícia Universidade Católica (PUC), São Paulo, São Paulo, Brasil; re.passetti@gmail.com; <https://orcid.org/0000-0002-1547-2831>

2 Universidade Federal de São Carlos (UFSCar), São Carlos, São Paulo, Brasil; pabloarantes@ufscar.br; <http://orcid.org/0000-0001-9707-8493>

Some remarks on the teaching of forensic phonetics to linguistics students

Abstract

This paper reports on the authors' experience in teaching an introductory course on forensic phonetics to a group of linguistics students and young researchers. We present the course syllabus and the assignments participants had to complete. We discuss participants' answers to the assignments and how they relate to the typical academic background students acquire in a linguistics major. We find that linguistics students and researchers are an apt audience for a course on subjects relating to forensic phonetics given their training in identifying and analyzing linguistic phenomena in an objective way, especially those relating to language and speech contextual variability. We suggest some areas where the curricula of linguistics majors in the Brazilian context could be improved to best prepare students to face the challenges of doing research or practical work in forensic phonetics.

Keywords: forensic phonetics teaching; excessive resolution; empirical segmentation.

Introdução

A Fonética Forense tem recebido destaque entre as áreas das Ciências da Linguagem no Brasil. Essa visibilidade deve-se, em parte, ao aumento da frequência com que os meios de comunicação noticiam investigações criminais, principalmente relacionadas ao cenário político brasileiro, envolvendo amostras de fala como elementos de prova, como a reportagem do jornal *O Globo*³ que revelou a gravação de uma conversa entre o então presidente Michel Temer e o empresário Joesley Batista, em 2017, a qual foi objeto de delação premiada do empresário. Soma-se a isso a curiosidade que essa área desperta, principalmente, entre alunos dos cursos de Letras e Linguística que veem a Fonética Forense como uma possibilidade de atuação profissional.

No entanto, ao mesmo tempo que a Fonética Forense desperta interesse no público estudantil jovem e/ou leigo, é uma área de estudos ainda em consolidação no ambiente acadêmico brasileiro. Esse cenário deve-se ao pouco espaço que, por muito tempo (e por vezes ainda hoje), foi (ou é) dado à discussão de assuntos relacionados a essa temática nas disciplinas de Fonética em cursos de graduação e pós-graduação de universidades do país e à quase ausência de disciplinas específicas dessa subárea nas instituições de ensino superior. Assim, a divulgação da área entre estudantes e pesquisadores das Ciências da Linguagem fica concentrada, em sua maioria, em grupos de estudos voltados

3 Disponível em: <https://oglobo.globo.com/brasil/dono-da-jbs-grava-temer-dando-aval-para-compra-de-silencio-de-cunha-21353935>. Acesso em: 01 abr. 2021.

especificamente para a Fonética Forense ou que têm a área como tópico adicional entre seus temas de discussão e pesquisa. Pode-se citar como exemplos de grupos de pesquisa que têm se dedicado à divulgação científica dessa área na última década o Grupo de Estudos em Fonética Forense (GEFF) da UNICAMP, o Laboratório de Fonética da UFSCar, o Laboratório Integrado de Análise Acústica e Cognição da PUCSP e o Laboratório de Áudio e Fonética Acústica (LAFA) da Escola Politécnica da PUCRS.

Uma iniciativa que busca preencher a lacuna quanto ao ensino e à divulgação da Fonética Forense para o público de Letras e Linguística no Brasil é a promoção de cursos de curta duração em eventos científicos. Além de introduzir alunos nessa área de estudos, essa formação tem como objetivo desmistificar estereótipos criados pela cultura popular, veiculados em grande parte pela mídia e produções ficcionais, em torno da Fonética Forense, bem como dar visibilidade para as possibilidades de atuação profissional de graduados em Letras e Linguística em órgãos do sistema judicial brasileiro, uma vez que o diploma desses cursos ainda não permite o ingresso por meio de concurso público em muitos desses órgãos nas carreiras de perito e/ou assistente técnico na área de Fonética Forense.

Tendo em vista esse contexto de crescente visibilidade e interesse pela área, conjugado à pouca oferta de cursos na área, fomos convidados para dar um curso introdutório de Fonética Forense na I Escola de Estudos Linguísticos do GEL. O curso, intitulado "Iniciação à Fonética Forense", foi oferecido de forma remota entre os dias 06 e 19 de setembro de 2020. Após sua realização, fomos convidados pela organização do evento a publicar um artigo sobre o tema do nosso curso na revista *Estudos Linguísticos* do GEL. Optamos por elaborar um relato sobre a nossa experiência, a fim de tecermos considerações sobre o ensino de Fonética Forense para estudantes de Letras e Linguística.

Neste relato, nos propomos a cumprir uma série de objetivos. Primeiramente, fazer uma discussão a respeito dos desafios de selecionar os tópicos relevantes na área, de condensar e adaptar esse conteúdo ao formato de um curso de curta duração e adaptado a um público com pouca ou nenhuma familiaridade com a área. Em seguida, descrevemos uma série de atividades práticas propostas aos participantes do curso e discutimos seus objetivos pedagógicos e as habilidades que elas pretendiam exercitar. Na terceira parte, apresentamos uma análise das respostas dos participantes às atividades propostas e discutimos de que forma essas respostas refletem achados de pesquisas da área de Fonética Forense. Ao final discutimos, com base nas observações a respeito do desempenho dos participantes, por que defendemos a inserção de linguistas no quadro de profissionais atuantes nesta área.

Descrição do curso e dos inscritos

Descrição do curso

O formato proposto inicialmente pela organização para o curso previa sua realização durante o recesso escolar de julho de 2020 em dois encontros na modalidade presencial, com duração total de 6 horas. Projetando um público de participantes com pouca experiência no tema da Fonética Forense e o objetivo de fazer uma introdução relativamente ampla da área, selecionamos os seguintes tópicos para discussão no curso: apresentação da Linguística e Fonética Forenses, breve histórico do campo, principais temas e métodos do campo, *linguistic profiling* e memória para vozes e filas de reconhecimento (1º encontro); e qualidade acústica do material de fala, comparação de locutor e vozes, efeitos da transmissão telefônica no sinal da fala e disfarce de voz (2º encontro). Devido à evolução da pandemia de COVID-19 no Brasil ao longo do primeiro semestre de 2020, a organização do evento suspendeu a realização de atividades presenciais e solicitou que docentes adaptassem seus cursos para que eles pudessem ser oferecidos na modalidade remota em setembro do mesmo ano. A proposta dos organizadores consistiu em promover dois encontros síncronos entre os instrutores e os participantes, cada um com duas horas de duração e separados por uma semana. Sugeriu-se, ainda, a disponibilização de materiais diversos na plataforma Google Classroom, como textos, vídeos e *podcasts*, que poderiam ser consultados pelos participantes em paralelo aos encontros síncronos e a partir dos quais atividades poderiam ser propostas.

Em função da adaptação do curso à proposta da organização, decidimos fazer algumas modificações no conteúdo do curso para adequá-lo melhor ao novo formato. Essas mudanças levaram em conta as respostas a um questionário enviado a nosso pedido pela organização aos participantes inscritos, que teve por objetivo obter informações a respeito do público, em especial seu conhecimento prévio sobre a área. De modo geral, essas respostas, que serão descritas em detalhes na seção 2.3, indicaram que o público que se interessou pelo curso era composto majoritariamente por pessoas que tinham conhecimento básico ou mediano sobre Fonética e Fonologia, pouco ou nenhum conhecimento a respeito de Fonética Acústica e do *software* de análise acústica Praat (BOERSMA, 2001), uma das principais ferramentas da área da Fonética Acústica, e que a grande maioria nunca teve contato com a área da Fonética Forense. Esse quadro foi importante para que ajustássemos o conteúdo a ser abordado no curso, tanto em termos da abrangência quanto da profundidade de tratamento dos tópicos.

Assim, tendo em vista o novo formato e o perfil dos participantes do curso, a primeira mudança envolveu a seleção dos materiais a serem disponibilizados. Uma das exigências da organização era que o material disponibilizado aos participantes deveria estar em formato digital e ser de livre acesso, de forma a não violar direitos autorais. Essa foi uma restrição importante, pois reduziu a gama potencial de materiais. Em uma oferta presencial, normalmente usamos materiais em formato que não o digital, como fotocópias

de textos, livros de biblioteca etc. Dessa forma, selecionamos livros e artigos científicos publicados em plataformas e periódicos, respectivamente, que seguiam a política de livre acesso, além de materiais audiovisuais, como vídeos da plataforma YouTube e *podcasts*. Os materiais utilizados para as atividades propostas, como arquivos de áudio, também foram retirados de plataformas de livre acesso ou de nossos arquivos pessoais.

Além disso, tivemos que reestruturar os tópicos selecionados para cada encontro. Decidimos que o primeiro encontro síncrono seria destinado ao embasamento teórico-científico da área de Fonética Forense, enquanto o segundo trabalharia aspectos práticos dos tópicos abordados anteriormente, além de outros pertinentes para a temática do curso. Assim, no primeiro encontro apresentamos a área de Fonética Forense, por meio de uma introdução ao histórico de formação desta área, à apresentação de suas principais tarefas e à discussão dos métodos de análise de dados e de tomada de decisão. Já no segundo encontro, discutimos a avaliação da qualidade acústica de áudios, por meio da metodologia proposta pelo Protocolo Geral para Exame de Comparação de Locutor (BARBOSA *et al.*, 2018), desenvolvido pelo Grupo de Estudos em Fonética Forense (GEFF) da UNICAMP, e também sobre a possibilidade de redução de ruído em gravações de áudio, por meio de um breve tutorial usando o Praat.

Descrição das atividades propostas

No formato final que o curso tomou, cada encontro síncrono foi acompanhado de uma atividade anterior, que designamos pré-síncrona, e outra posterior a ele, que designamos pós-síncrona. Essas atividades tinham os objetivos de estimular e fixar o conteúdo apresentado nos encontros e presente nos materiais disponibilizados anteriormente. Como atividade pré-síncrona do primeiro encontro, pedimos aos alunos que respondessem a três questões elaboradas com base em materiais introdutórios sobre Fonética Forense e sobre princípios gerais da ciência forense.

A atividade pós-síncrona do primeiro encontro, por sua vez, tinha o objetivo de criar o perfil linguístico de cinco falantes, dos quais três eram falantes de português brasileiro (doravante PB) e dois eram falantes de inglês britânico. As amostras de falantes do PB foram retiradas de *podcasts* de entrevistas. Uma delas do *podcast Fronteiras da Ciência*, da UFRGS, episódio “Crises Econômicas I”⁴, e as outras duas do *podcast Mamilos*⁵, episódios “Rio Doce – da lama ao caos” e “Era uma vez... uma família feliz”. As duas amostras de falantes do inglês britânico eram provenientes de material produzido pelo foneticista John Laver para treinamento no uso do protocolo *Voice Profile Analysis* (doravante VPA)

4 Disponível em: http://multimidia.ufrgs.br/conteudo/frontdaciencia/Fronteiras_da_Ciencia-T07E34-Crises.Economicas.1-24.10.2016.mp3. Acesso em: 01 abr. 2021.

5 Disponível em: <https://www.b9.com.br/shows/mamilos/mamilos-159-rio-doce-da-lama-ao-caos/?highlight=Rio%20Doce>. Acesso em: 01 abr. 2021.

(LAVIER, 1980). Cada amostra de fala que compôs essa atividade tinha a duração de 20 segundos, aproximadamente.

A elaboração dos perfis linguísticos desses falantes consistia na resposta a seis perguntas para as amostras de falantes do PB e quatro para as amostras de falantes do inglês britânico. As perguntas para cada grupo de falantes estão apresentadas no quadro 1.

Quadro 1. Perguntas a serem respondidas para traçar o perfil linguístico das amostras de fala

Perfil linguístico de falantes do PB	Perfil linguístico de falantes do inglês britânico
<ul style="list-style-type: none">• Sexo• Idade estimada• Com base no sotaque, de qual região geográfica do país o falante é?• O falante apresenta alguma idiossincrasia no nível articulatório? Se sim, especifique qual(is)• O falante apresenta alguma idiossincrasia de natureza linguística? Se sim, especifique qual(is).• Características de qualidade de voz	<ul style="list-style-type: none">• Sexo• Idade estimada• O falante apresenta alguma idiossincrasia no nível articulatório? Se sim, especifique qual(is)• Características de qualidade de voz

Fonte: Elaboração própria

Não incluímos perguntas sobre proveniência regional do falante no caso das amostras de língua inglesa porque essa informação não é dada no material de onde retiramos a amostra. Além disso, como não tínhamos informação sobre o nível de conhecimento dos participantes a respeito da língua inglesa, julgamos que não faria sentido incluir a pergunta a respeito de idiossincrasias linguísticas nessas amostras.

Em relação ao segundo encontro, a atividade pré-síncrona tinha como objetivo avaliar o efeito da adição de ruídos a uma gravação com boa qualidade acústica sobre a transcrição do seu conteúdo linguístico. Para isso, duas variáveis foram testadas: o tipo e a intensidade do ruído sobreposto. A base para o material do exercício foram duas amostras de fala de cerca de 20 segundos cada, produzidas por uma falante da região de Campinas/SP em um contexto de entrevista semiespontânea, gravadas em ambiente silencioso com microfone de boa qualidade.

A essas duas amostras foram adicionados, separadamente, ruídos de dois tipos: tráfego de automóveis (com perfil espectrográfico próximo ao ruído branco) e um conjunto de áudios de fala em cinco línguas diferentes. Em cada série, o nível de ruído foi variado em

três níveis, com intensidade crescente. Ao final, cada série continha quatro amostras: o áudio original, sem ruído adicionado (nível 0) e outras três, com a sobreposição de ruído de intensidade crescente (níveis 1 a 3). A relação sinal-ruído em cada amostra é mostrada na tabela 1. A tarefa dos alunos consistia em transcrever ortograficamente o conteúdo linguístico de cada gravação⁶, começando por aquela com mais ruído até chegar à gravação original, sem adição de ruído. As transcrições deveriam estar acompanhadas de um relatório, no qual os alunos tinham que: (1) comparar de maneira objetiva qual dos dois tipos de ruído era mais efetivo para encobrir a fala, isto é, aquele que mais dificultava a transcrição; e (2) fazer considerações sobre a natureza linguística dos erros de transcrição, investigando, por exemplo, se havia alguma relação entre a palavra correta e a identificada erroneamente. Essas relações poderiam envolver, ainda, semelhanças entre tonicidade e/ou fones ou o contexto linguístico de produção. Para avaliar a dificuldade de transcrição (tarefa 1), foi disponibilizado um *script* para o Praat, que tomava como base o total de palavras transcritas na gravação original (sem ruídos) para calcular a porcentagem de palavras corretamente transcritas em cada nível de ruído subsequente (ARANTES, 2020).

Tabela 1. Relação sinal-ruído nas amostras do exercício de transcrição em função do tipo de ruído e do nível de ruído

Tipo de ruído	Nível do ruído	Relação sinal-ruído (dB)
Fala sobreposta	0	36,5
	1	18,5
	2	12,6
	3	7,4
Tráfego	0	44,2
	1	5,1
	2	2,9
	3	1,3

Fonte: Elaboração própria

⁶ Idealmente os alunos deveriam realizar a transcrição no *software* de análise acústica Praat, uma vez que este viabilizaria a utilização do *script* para contagem da porcentagem de palavras. No entanto, aos alunos que não tinham familiaridade com o programa foi sugerido que fizessem a transcrição em um arquivo de texto e disponibilizada uma planilha para o cômputo das porcentagens de acertos.

Na atividade pós-síncrona deste encontro, os alunos deveriam avaliar a qualidade acústica do conjunto de áudios utilizados na atividade anterior, baseando-se na proposta do protocolo fonético-forense apresentado durante o encontro síncrono. Assim, ao final da análise, eles deveriam atribuir uma etiqueta a cada um dos seis arquivos de áudio analisados e elaborar um breve relatório comentando o processo de análise e os resultados obtidos.

Mais informações sobre as atividades, bem como a discussão dos resultados obtidos, serão apresentadas na seção “Discussão das atividades”.

Descrição do perfil dos inscritos

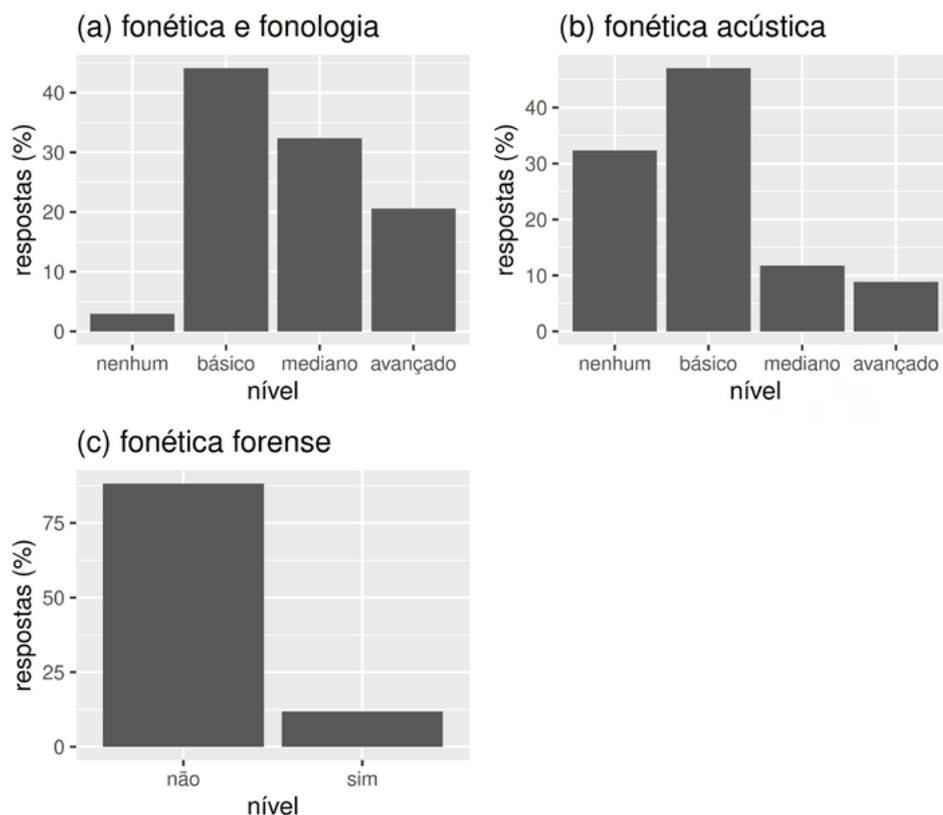
O curso teve um total de 45 alunos oficialmente matriculados. Para podermos fazer o planejamento do curso de maneira que ele se adequasse ao nível de conhecimento dos participantes, pedimos à organização que enviasse a eles um questionário. O questionário foi enviado cerca de um mês antes do início das atividades, continha três perguntas e as respostas eram abertas: 1) Qual é o seu grau de conhecimento na área de Fonética e Fonologia? 2) Qual é o seu grau de conhecimento da área de Fonética acústica? Já usou o programa de análise acústica Praat? e 3) Você tem alguma experiência prévia com Fonética Forense? O questionário foi aplicado com o objetivo de conhecer melhor a experiência dos participantes com o tema central do curso, a Fonética Forense, e com conhecimentos prévios mobilizados, Fonética, Fonética Acústica e Fonologia.

Do total de inscritos, 34 responderam ao questionário. As respostas para as duas primeiras perguntas foram classificadas em quatro categorias: nenhum, básico (se a pessoa mencionou haver cursado uma disciplina introdutória na graduação), mediano (mais do que um curso na área em nível de graduação ou pós-graduação), avançado (participante faz ou fez pesquisa na área em nível de graduação ou pós-graduação). As respostas à terceira questão foram categorizadas simplesmente em “não” ou “sim”. As respostas “sim”, em geral, mencionaram contato prévio com a Fonética Forense através de cursos de curta duração, ou de conteúdos da área discutidos como parte de disciplinas de graduação ou por meio da leitura de algum texto de apresentação da Fonética Forense.

A figura 1 mostra os percentuais de resposta para as três perguntas. Para todos, o total de respostas é 34 e as porcentagens foram calculadas tendo como referência esse valor. Podemos observar que quase a metade dos participantes tinha pelo menos conhecimentos básicos de Fonética e Fonologia e a outra metade conhecimentos mediano ou avançado. Apenas um participante reportou não ter nenhum conhecimento. A situação foi diferente para a segunda pergunta. Quase um terço dos participantes declarou não ter nenhum conhecimento de Fonética acústica e experiência com o programa Praat e quase 50% reportou conhecimento de nível básico. Houve em torno de 10% de respostas que indicaram conhecimento mediano e outros 10% de nível avançado.

Sobre Fonética Forense, 88% dos participantes declaram não ter tido contato prévio com conteúdos a respeito do assunto.

Figura 1. Respostas ao questionário aplicado aos participantes antes do início do curso sobre grau de familiaridade com áreas do conhecimento relacionadas ao tema do curso



Fonte: Elaboração própria

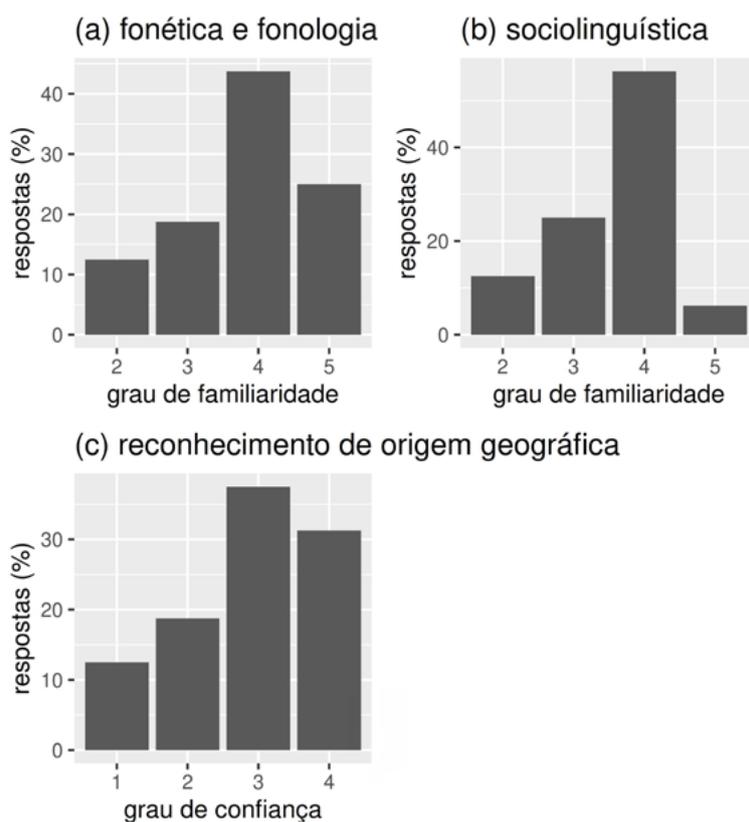
Depois de terminado o curso e com o convite feito pela organização do evento para a escrita de um texto a respeito deste, enviamos um segundo questionário aos participantes em que foi pedido seu consentimento para uso das respostas dadas aos exercícios no presente texto, além de algumas perguntas que nos permitiram fazer um quadro demográfico mais completo dos participantes. Essas informações foram coletadas para serem correlacionadas posteriormente com as respostas dados ao exercício de perfil linguístico. Responderam a esse questionário 16 dos participantes.

A idade mediana é 25 anos (mín. 20 e máx. 43 anos), 75% deram como resposta “feminino” e 25% “masculino” para o gênero. Os participantes vêm de 16 cidades diferentes em 5 estados brasileiros: São Paulo (N = 9), Minas Gerais (N = 3), Paraná (N = 2), Rio de Janeiro (N = 1) e Distrito Federal (N = 1). Metade dos respondentes já residiu por mais do que dois anos em outra cidade no mesmo estado da federação que não aquela em que nasceu (N

= 8); entre os demais, 25% (N = 4) residiram por mais do que dois anos em cidade de outro estado e outros 25% (N = 4) nunca residiram fora da cidade em que nasceram.

Em termos de formação acadêmica, os que responderam dividem-se entre graduação em andamento (50%, N = 8), pós-graduação completa (31%, N = 5) e pós-graduação em andamento (19%, N = 3). Entre os que especificaram o curso de graduação, predomina o curso de Letras (N = 11), seguido de Linguística (N = 2) e a dupla habilitação Letras-Linguística (N = 1). Entre os participantes, 75% (N = 12) desenvolveram alguma atividade de pesquisa durante a graduação, em áreas como Fonética e Fonologia, sociolinguística, sintaxe, linguística de *corpus*, semiótica ou análise do discurso. Entre os três respondentes que no momento frequentam a pós-graduação, predominam aqueles que desenvolvem pesquisa em temas ligados à área de Fonética e Fonologia ou em áreas de interseção com essa (como aquisição de L2, variação linguística e linguística histórica, por exemplo).

Figura 2. Respostas ao questionário aplicado após o curso sobre grau de familiaridade com áreas do conhecimento relacionadas ao tema do curso (1: baixa, 5: alta)



Fonte: Elaboração própria

Nesse questionário, também pedimos aos participantes que estimassem o grau de familiaridade com conceitos da área de Fonética e Fonologia e sociolinguística e que

estimassem sua habilidade para reconhecer de que região do Brasil uma pessoa vem pela maneira como ela fala. Nos três casos, a resposta foi dada em uma escala de 1 a 5, 1 sendo pouca familiaridade/habilidade e 5 muita familiaridade/habilidade. As porcentagens de respostas em cada grau para as três perguntas são mostradas na figura 2.

No caso da pergunta sobre familiaridade com conceitos na área de Fonética e Fonologia, a mediana das respostas é o grau 4 e 70% das respostas correspondem a respostas 4 e 5. Isso pode parecer contraditório com o que é mostrado na figura 1a, que indica que metade dos participantes tem conhecimento básico na área. Duas interpretações para esses dois resultados são: a) os participantes julgaram dominar bem os conceitos da área, mesmo quando seus conhecimentos não são muito avançados ou b) responderam ao questionário pós-curso um subconjunto de participantes com maior familiaridade em relação à área.

Sobre os conceitos da área de sociolinguística, a resposta mediana também é o grau 4 e mais de 80% das respostas correspondem aos graus 3 e 4. Nesse grupo, o domínio percebido dos conceitos da área de sociolinguística tende a ser um pouco menor do que o domínio dos conceitos da área de Fonética e Fonologia, embora sejam números relativamente altos. O fato de muitos dos participantes já terem realizado pesquisa que envolvia de algum modo a área de Fonética e Fonologia pode explicar a percepção por parte dos participantes de maior domínio dessa área em comparação com a sociolinguística.

Na pergunta a respeito da habilidade de identificar proveniência geográfica a partir da fala de alguém, o grau de confiança mediano dos participantes é 3, o ponto médio da escala, com cerca de 37% das respostas. Note-se a ausência de qualquer resposta de grau 5 e a presença de cerca de 13% de respostas de grau 1, que não aparece nas outras duas perguntas. Mais de 30% de todas as respostas correspondem aos graus 1 e 2. É evidente que os participantes têm menor grau de confiança nessa habilidade, apesar de julgarem ter um bom domínio dos conhecimentos de Fonética, Fonologia e Sociolinguística, que em princípio poderiam ajudar na tarefa.

A proveniência geográfica dos participantes concentrou-se em estados da região Sudeste do país (São Paulo, Minas Gerais e Rio de Janeiro perfazem mais de 80% dos participantes). Os outros eram naturais de PR e DF. Entre os participantes, 75% sempre residiram em sua cidade natal ou mudaram-se para cidades dentro do próprio estado. Portanto, uma parcela grande dos participantes teve, em tese, menos oportunidades de ter experiências cotidianas com falantes de variedades linguísticas diferentes⁷. Apenas 25% mudaram de estado ao longo da vida – três de quatro participantes nessa situação mudaram-se para estados em outra região geográfica do país.

⁷ Excluímos a possibilidade de que algum participante possa ter relações próximas com falantes de uma outra variedade linguística. Nosso questionário não levantou essa informação.

Com relação às variedades regionais brasileiras com as quais o participante julga ter familiaridade, variedades da região onde o participante mora ou já morou foram nomeadas por 62,5% (N = 10) e de regiões vizinhas por 37,5%. Quase 44% dos participantes disseram ter familiaridade com variedades da região Nordeste, independentemente de terem residido lá (variedades nomeadas de forma específica foram “baiano”, “cearense” e “pernambucano”, além do rótulo genérico “nordeste”). O fato de ser uma região populosa, historicamente importante e que é frequentemente representada em produtos culturais (romances, teatro, programas de TV etc.) pode ser uma das razões para essa resposta dos participantes.

A referência às variedades regionais nas respostas ao questionário é feita majoritariamente em três níveis geográficos: as grandes regiões geográficas do país (Sul, Norte, Centro-Oeste etc.), estados (“mineiro”, “paulista”, “gaúcho”) e cidades individuais (“carioca”, “Porto Alegre”). Essa escolha pode ter derivado da maneira como a pergunta foi formulada no questionário, que fazia referência a “variedades regionais” ou de uma assunção por parte dos participantes de que estados e cidades são as unidades dialetais mais relevantes. A exceção às denominações com base na divisão político-administrativa é a referência ao “dialeto caipira”, que em princípio tem uma base histórico-cultural e não coincide necessariamente com fronteiras estaduais ou regionais. O uso que os participantes fazem da expressão, no entanto, parece igualar o “caipira” à variedade falada no interior de São Paulo em contraposição à variedade paulistana.

Discussão das atividades

Bases da ciência forense

Essa atividade pré-síncrona do primeiro encontro consistiu na resposta a uma pergunta baseada na leitura de Eriksson (2014) e duas baseadas na leitura de Robertson, Vignaux e Berger (2016).

A primeira pergunta dizia: *“O conhecimento leigo sobre Fonética Forense, muitas vezes influenciado pelo ‘maravilhoso mundo dos filmes e seriados policiais’, pode nos levar a pensar que identificar um indivíduo por sua voz e fala trata-se de uma tarefa simples e totalmente automática, realizada de maneira rápida e eficaz, por exemplo, pela comparação visual de espectrogramas – o que não é verdade. Mas, por que não podemos assumir que o padrão representado por um espectrograma se equipara à fala de um indivíduo da mesma forma que a impressão digital corresponde ao padrão da superfície de um dedo?”*.

Além da resposta objetiva à questão, alguns alunos reportaram que, antes do contato com as leituras introdutórias, acreditavam que o reconhecimento de falantes era uma tarefa intuitiva e totalmente automatizada. Conforme comentado na questão, esse relato ilustra a difusão de um estereótipo sobre a condução de práticas forenses

sustentado, principalmente, pela maneira como essas práticas são representadas em produções ficcionais. Entretanto, após a introdução dos alunos à bibliografia do curso, todas as respostas foram categóricas em apontar a principal razão que faz a tarefa de reconhecimento de falantes ser um grande desafio: a variabilidade da fala.

As fontes de variabilidade apontadas nas respostas puderam ser agrupadas em duas categorias: linguísticas e extralinguísticas. Dentre os fatores linguísticos, foram destacados: o contexto linguístico de produção dos sons da fala, o dialeto, disfarces vocais envolvendo modificações entoacionais e/ou articulatórias e a qualidade de voz. Como fatores extralinguísticos, foram mencionados: a qualidade acústica das gravações (fator relacionado à presença de ruídos ambientais, técnicos e de sobreposição de fala), a não contemporaneidade e/ou a duração das amostras de fala comparadas, a (não) familiaridade entre os falantes, o estado de saúde e/ou emocional, a idade, as diferenças fisiológicas, as alterações hormonais e o uso de substâncias químicas (como entorpecentes, álcool e anabolizantes).

A segunda e a terceira perguntas apresentavam a descrição de três princípios fundamentais de práticas da ciência forense, a saber: o Princípio de Locard, o Princípio da individualidade e o Princípio da individualização (ROBERTSON; VIGNAUX; BERGER, 2016) e diziam: *“Em que medida material linguístico de forma geral e de natureza fonético-fonológica em particular se adequa aos princípios apresentados acima? Quais são as possibilidades e os limites da atribuição de material linguístico a um falante/usuário em particular?”*.

Quanto à primeira questão, relacionada ao Princípio de Locard, os alunos mencionaram como traços deixados por um falante na cena do crime: material linguístico proveniente de gravações ou escrita e de natureza fonético-fonológica, principalmente, mas também sintática, morfológica, semântica, lexical e discursiva, e que pode indicar marcas dialetais, como a palatalização de /t d s/ e o abaixamento de vogais médias átonas; características fisiológicas provenientes de aparelho fonador, cuja dimensão e configuração diferem entre homens e mulheres, por exemplo; e a memória de testemunhas auriculares em crimes envolvendo-as. Como traços que o falante pode levar da cena do crime foram apontados ruídos ou sons de fundo presentes em gravações questionadas, os quais podem dar pistas sobre o local onde estas ocorreram.

Quanto ao Princípio da individualidade, as respostas mencionaram a variabilidade presente na fala, a qual garante que duas amostras de fala produzidas por um mesmo falante nunca sejam exatamente iguais. Além disso, mencionaram, ainda, que os vestígios deixados pelo falante (cf. Princípio de Locard) podem revelar características idiossincráticas, as quais, descartada a possibilidade de coincidência casual, podem indicar a mesma fonte, conforme assegurado pelo Princípio da individualização.

As respostas indicaram como possibilidades de atribuição de material linguístico a um mesmo falante o elevado número de similaridades entre os materiais comparados, o que permite um maior número de análises comparativas e o emprego de diferentes técnicas de análise. Como limites a essa atribuição foram citadas as diferenças de qualidade acústica entre os materiais comparados (gravação interceptada *versus* controlada), a imprecisão no uso das técnicas, a deterioração da memória acústica de testemunhas auriculares e o tempo decorrido entre a produção da prova e a obtenção do material de comparação.

A análise geral das respostas indicou um elevado grau de consciência por parte dos participantes acerca de fenômenos linguísticos inerentes à fala, em especial, o efeito de uma série de fatores linguísticos e extralinguísticos na variabilidade da fala. Mesmo antes do encontro síncrono e da discussão de conteúdos do curso, os alunos demonstraram em suas respostas conhecimentos que estão relacionados ao próprio funcionamento das línguas naturais. Destaca-se, também, a presença e o uso apropriado feito por parte dos participantes, de termos técnicos da Linguística e da Fonética Forense, como “palatalização” e “abaixamento”, referentes a processos fonológicos, e “idiossincrasias” e “amostras (não) contemporâneas”, recuperados das leituras indicadas para o primeiro encontro síncrono.

Perfil linguístico

De forma geral, embora as respostas dadas pelos participantes às perguntas sobre a presença de idiossincrasias (tanto no nível articulatório quanto em outros níveis linguísticos) e características de qualidade de voz tenham demonstrado algumas limitações, que comentaremos a seguir, elas também sinalizaram aspectos interessantes a serem considerados nas práticas fonético-forenses.

Uma das limitações encontradas para a análise das amostras em PB diz respeito ao uso de vocabulário não técnico, baseado em impressões vocais sobre as amostras analisadas. Na descrição de idiossincrasias no nível articulatório, por exemplo, muitas descrições não foram feitas em termos dos ajustes acústico-articulatórios objetivos responsáveis pelas idiossincrasias, mas sim recorrendo a impressões emocionais, discursivas e/ou pragmáticas evocadas pelas falas analisadas. Exemplos disso são descrições como:

1. “Me pareceu uma voz animada, empolgante (pensei em um *coach*).”
2. “[...] só achei uma voz tranquila, de alguém experiente.”

Além disso, muitas descrições de características de qualidade de voz fizeram uso de rótulos mais ou menos correntes no uso comum da língua, geralmente associados à área da música, a estados fisiológicos cotidianos, como gripes e resfriados, ou a disfonias funcionais, causadas por uso inadequado da voz. Exemplos de rótulos dessa natureza que apareceram nas respostas são: voz nasalizada, rouca, áspera, grave e aguda⁸.

No entanto, o uso de rótulos não técnicos não é de todo mal, uma vez que, em situações envolvendo testemunhas auriculares, provavelmente esses rótulos estarão presentes e caberá ao foneticista forense relacioná-los a características acústicas da voz-alvo. Além disso, é importante considerar a intuição linguística de ouvintes não treinados em tarefas fonético-forenses como essas, uma vez que ouvintes leigos são capazes de atribuir de forma consistente rótulos descritivos a amostras de voz (WATT; BURNS, 2012). Por isso, estudos que correlacionem rótulos não técnicos, dados por ouvintes leigos, a descrições técnicas de qualidade de voz e/ou articulatórias são necessários para aprimorar e otimizar o trabalho dos peritos nessas situações.

Por outro lado, nas amostras em inglês britânico, idioma distinto da língua nativa dos participantes, a descrição das características linguísticas das vozes-alvo foi aprimorada em relação ao uso de termos técnicos. Algumas das descrições apresentaram termos e/ou símbolos fonéticos para se referir a características dos sotaques dos falantes, por exemplo, a menção ao desvozeamento de sons róticos em posição de coda silábica e à produção da oclusiva alveolodental vozeada ([d]) ao invés da fricativa dental vozeada ([ð]). Esse resultado demonstra que, para esse conjunto de dados, os alunos, no geral, deram mais importância à "forma" linguística, isto é, a características acústico-articulatórias das vozes analisadas. Ao comparar esse resultado com aquele das amostras de fala do PB, observa-se que o conteúdo da mensagem pode ter um peso na tarefa perceptiva desempenhada por ouvintes leigos. Por isso, também é importante que o foneticista forense leve esse fato em consideração ao fazer uso dessas descrições.

Outro fator percebido neste exercício diz respeito ao impacto da qualidade acústica das gravações na análise das amostras de fala em inglês britânico. Essas amostras foram obtidas na década de 1980, quando da elaboração do VPA, e disponibilizadas por meio de gravações em fitas cassetes (LAVIER, 1980). Devido a isso, possuem qualidade acústica inferior às demais gravações do conjunto de áudios, que foram obtidas por meio de equipamentos e tecnologias de gravação mais desenvolvidos. A dificuldade de análise em função da qualidade acústica dessas gravações foi demonstrada por meio de comentários como:

⁸ Também apareceram termos mais técnicos ou pouco usuais na linguagem ordinária, como voz gutural e crepitante. Porém, nesses casos, não sabemos com que grau de precisão esses termos foram usados, pois um termo como "gutural" pode se referir a alterações vocais produzidas por diferentes estratégias articulatórias e/ou fonatórias.

3. "Parece uma gravação antiga, televisionada, talvez da década de 50."
4. "Áudio abafado, com muito ruído."
5. "Qualidade antiga, demorei um pouco para poder identificar."

Esses comentários reforçam observações feitas por Barbosa *et al.* (2018) no Protocolo Geral para Exame de Comparação de Locutor de que a baixa qualidade acústica de gravações pode afetar tarefas perceptivas (e, conseqüentemente, acústicas) conduzidas em materiais de fala.

Quanto às respostas dos participantes referentes à identificação de idiossincrasias de natureza linguística e seu potencial de discriminar falantes, ocorreram relatos como:

6. "não monotongou em 'justificou'".

Embora a observação seja correta do ponto de vista observacional, uma vez que o processo fonológico mencionado no exemplo (monotongação) de fato foi aplicado na amostra em questão, o peso que esse fenômeno teria enquanto evidência sociofonética numa tarefa de Comparação de Locutor (doravante CL) é provavelmente baixo, uma vez que é amplamente difundido no território brasileiro e usado por diferentes grupos sociais. Por outro lado, observações de fenômenos de caráter morfossintático, como o aludido no exemplo 7, podem ter um peso discriminante maior, dados os condicionantes diastrático e diatópico a que podem estar associados.

7. "(uso) 'de' minha mãe ao invés 'da' (minha mãe)."

Devido ao fato de tarefas de CL buscarem idealmente por características linguísticas com alta variabilidade interfalante e baixa variabilidade intrafalante (NOLAN, 1993), essas respostas ensejaram uma discussão no encontro síncrono sobre o potencial de uso de evidências sociofonéticas em práticas fonético-forenses, em especial o fato de que não basta identificar os fenômenos linguísticos presentes no material de fala, é preciso também (e principalmente) avaliar qual o potencial discriminante que possuem (BRESCANCINI; GONÇALVES, 2020).

Comentamos agora as variações encontradas nas idades atribuídas aos falantes de cada uma das cinco amostras analisadas. Vale ressaltar que esses comentários não levarão em conta as taxas de acerto e erro das estimativas dadas pelos participantes,

pois informações sobre as idades dos falantes não foram fornecidas nos materiais de onde retiramos as amostras. Os resultados mostraram que as estimativas para a maioria dos falantes ficaram abaixo dos 50 anos e se concentraram em faixas com variações de até 14 anos. A idade estimada para o primeiro falante brasileiro concentrou-se na faixa de 31 a 45 anos e para a segunda falante brasileira na faixa de 26 a 35 anos. Em ambos os casos, essas faixas representaram em torno de 58% das escolhas dos ouvintes. Já as idades estimadas tanto para a terceira falante brasileira quanto para o quinto falante estrangeiro concentraram-se na faixa de 36 a 50 anos, as quais foram responsáveis por aproximadamente 69% e 89% das respostas, respectivamente. A situação foi diferente apenas para o quarto falante estrangeiro, cuja estimativa de idade distribuiu-se em diferentes faixas etárias. Embora não tenhamos como calcular a taxa de acerto dessas respostas, esses resultados são interessantes porque, de modo geral, apontam para uma consistência entre as estimativas dos participantes.

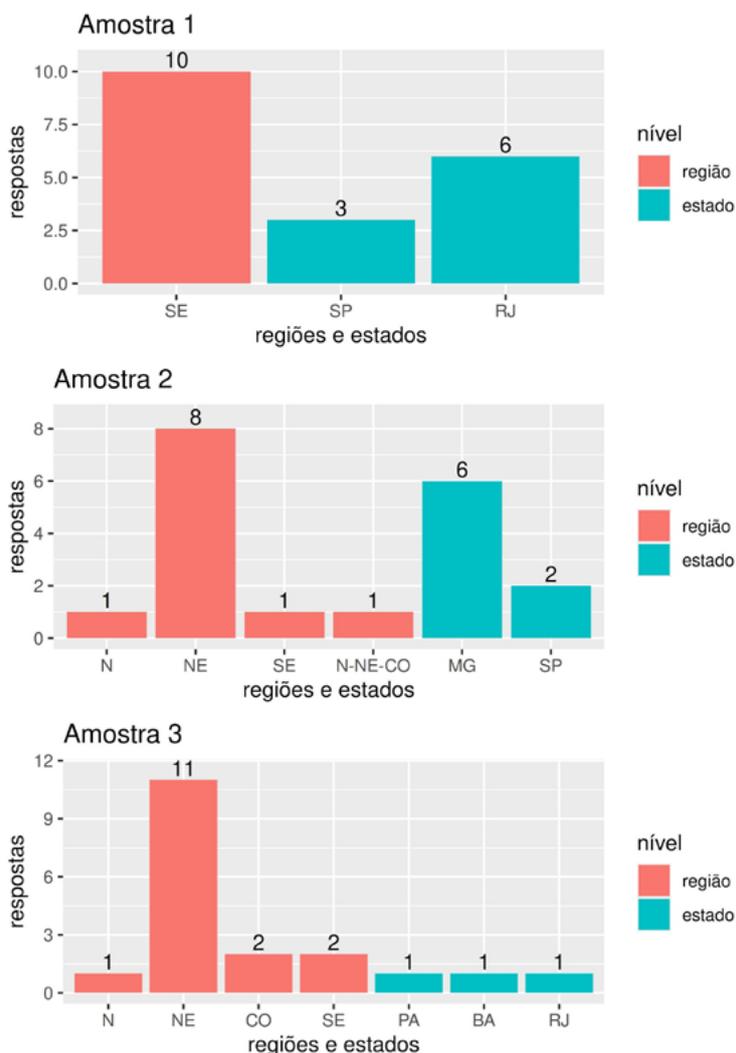
Por fim, cabe-nos comentar sobre as respostas dos participantes em relação à proveniência geográfica dos falantes das amostras 1 a 3. O falante da amostra 1 é natural do Rio de Janeiro. Em torno de 84% das respostas indicaram corretamente a origem do falante, embora em 53% dos casos a região Sudeste tenha sido mencionada e apenas em 31% o estado do Rio de Janeiro tenha sido especificamente apontado. O restante das respostas indicou erroneamente o estado de São Paulo como a origem. Nenhum outro estado ou região é citado. A identificação foi mais precisa, provavelmente pelo fato de a maior parte dos participantes também ser da região SE. Além disso, muitos declararam ter familiaridade com a variedade “carioca”.

A falante da amostra 2 é da cidade de Mariana, estado de Minas Gerais (MG). Apenas 31% das respostas apontaram corretamente o estado de MG como a origem. É notável que em torno de 42% das respostas apontem a região Nordeste como a origem. Uma hipótese para a predominância dessa resposta errônea é o fato de a falante produzir na amostra abertura das vogais pretônicas /e o/ nas palavras *morava*, *morar* e *apegada*, além de alongar de forma perceptível algumas das sílabas tônicas, traços comumente associados a variedades faladas na região NE. O fato de uma porcentagem muito alta dos participantes ter respondido no questionário pós-curso ter familiaridade com variedades dessa região pode ser também um fator que tenha colaborado para a identificação errônea da origem da falante. A indicação de origem dessa amostra foi a menos precisa. Entre as respostas coletadas aparecem a região Norte e uma que englobava num rótulo só as regiões Norte, Nordeste e Centro-Oeste, nenhuma das quais a região a que pertence o estado de Minas Gerais.

A falante da amostra 3 é natural da cidade de Recife, estado de Pernambuco. Do total de respostas, 58% identificaram corretamente a região Nordeste como origem da falante. Outras regiões que não a NE foram mencionadas em 26% das respostas. Das respostas que mencionaram estados específicos, 16% do total, nenhuma indicou Pernambuco

e duas das respostas são de estados fora do NE. Esse resultado é interessante, na medida em que corrobora o julgamento dos participantes a respeito de sua familiaridade percebida com variedades daquela região – a região apareceu em 44% das respostas ao questionário aplicado após a realização do curso. Por outro lado, dadas as grandes diferenças linguísticas dentro de cada uma das grandes regiões brasileiras, um perfil que indica apenas a região reduz o espectro de possibilidades, mas ainda tem um grau de incerteza razoável. É interessante analisar a maneira como a região NE aparece nas respostas tanto a respeito da amostra 2 quanto da 3.

Figura 3. Atividade de perfil linguístico: respostas dos participantes a respeito da origem geográfica dos falantes nas três amostras do português brasileiro. As respostas especificaram região geográfica ou um estado da federação específico

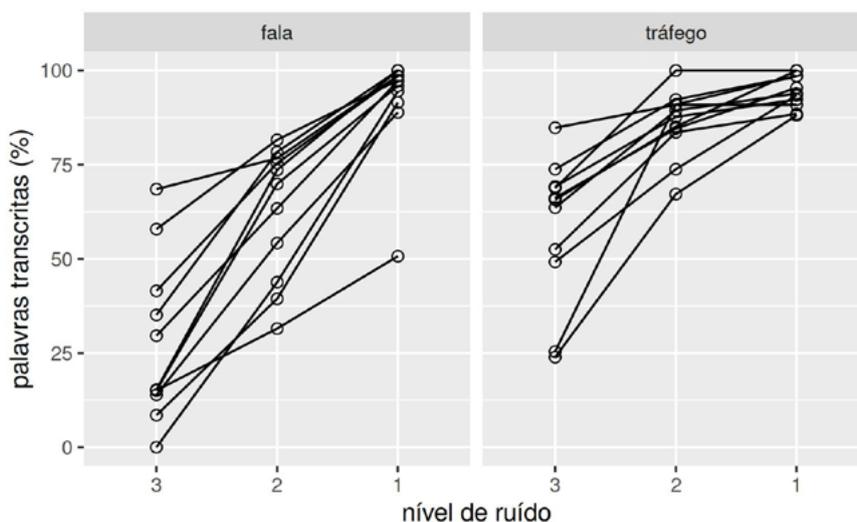


Fonte: Elaboração própria

Transcrição de fala em ruído

Como descrevemos anteriormente, nesta tarefa os participantes deveriam fazer a transcrição ortográfica de duas amostras de fala, cada uma encoberta por ruídos de diferentes intensidades e de diferente tipo. Para cada tipo de ruído havia quatro níveis de ruído, desde o nível "0", no qual não havia ruído adicionado até o nível "3", no qual o ruído tinha a maior intensidade e a relação sinal ruído era a pior. Assumimos que no nível de ruído "0" cada participante seria capaz de transcrever a totalidade das palavras presentes na gravação e por isso omitimos esse nível no gráfico. Em cada caixa da figura, o nível de ruído diminui da esquerda para a direita no eixo horizontal. A observação do gráfico mostra que a porcentagem de palavras transcritas aumenta conforme diminui o ruído, muito embora a proporção no aumento seja diferente para os dois tipos de ruído. Cada linha liga três círculos, que são os percentuais transcritos por um mesmo participante nos três níveis. Esse recurso permite ver como os diferentes participantes variam entre si. Alguns conseguem transcrever mais palavras dos que os outros, independentemente do nível de ruído.

Figura 4. Porcentagem de palavras transcritas nos níveis de ruído 3, 2 e 1 relação ao total de palavras no nível de ruído 0 (sem ruído adicionado)



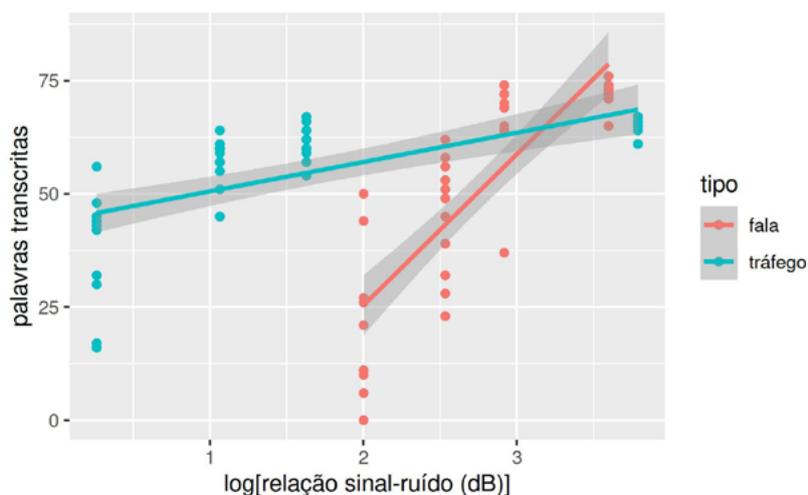
Fonte: Elaboração própria

Em cada caixa da figura, o nível de ruído diminui da esquerda para a direita no eixo horizontal. A observação do gráfico mostra que a porcentagem de palavras transcritas aumenta conforme diminui o ruído, muito embora a proporção no aumento seja diferente para os dois tipos de ruído. Cada linha liga três círculos, que são os percentuais transcritos por um mesmo participante nos três níveis. Esse recurso permite ver como os diferentes participantes variam entre si. Alguns conseguem transcrever mais palavras do que os outros, independentemente do nível de ruído.

Comparando a média dos dois tipos de ruído nível a nível por meio de testes-*t* pareados⁹, vemos que no nível “3” a condição “fala” fica atrás da condição “tráfego” em média 31 pontos percentuais na medida de palavras transcritas [$t(10) = -4,94; p < 0,001$], uma diferença que é estatisticamente significativa; no nível “2”, a diferença entre as condições de ruído diminui para 22,2 pontos percentuais, diferença que também é estatisticamente significativa [$t(10) = 4,48; p = 0,001$]; finalmente, no nível “1”, a condição “fala” fica atrás em média -2,26% da condição “tráfego” [$t(10) = -0,57; p = 0,58$], uma diferença que não é significativa – neste caso, ambos os tipos apresentam percentuais de transcrição muito próximos da totalidade das palavras das amostras (92,2% contra 94,4%).

A observação da dispersão vertical dos dados permite fazer a hipótese de que o desempenho na transcrição não é uniforme, tanto no que diz respeito ao tipo de ruído quanto aos níveis. Um teste-*F* para comparação das variâncias dos valores dos dois tipos de ruído, agrupando os três níveis, indica que a diferença é significativa, com a do grupo “fala” sendo maior do que a do grupo “tráfego” [$F(32, 32) = 2,67; p = 0,007$]. Fazendo a comparação entre os tipos de ruído nível a nível, os testes mostram que no nível “3” a diferença não é significativa (SD: fala 21; tráfego 18,5) [$F(10, 10) = 1,28; p = 0,7$]; no nível “2” a diferença é significativa, a variância do tipo “fala” apresentando o maior dos valores (SD: fala 17,6; tráfego 9) [$F(10, 10) = 3,83; p = 0,045$]; no nível “1” a diferença também é significativa e na mesma direção (SD: fala 14,2; tráfego 4,5) [$F(10, 10) = 10,1; p = 0,001$]. De modo geral, os resultados mostram que a variabilidade no desempenho da transcrição tende a diminuir conforme diminui o nível de ruído que encobre a fala-alvo, embora o ruído de tipo “fala” provoque um nível maior de variabilidade entre os transcritores.

Figura 5. Número absoluto de palavras transcritas em função do valor da relação sinal-ruído (valor em dB convertido para seu logaritmo). As duas condições de ruído são representadas por cores diferentes. As linhas contínuas são dadas pelas equações de regressão linear que melhor aproximam os dados de cada condição de ruído



Fonte: Elaboração própria

⁹ Assumimos para as análises estatísticas um limiar de 5% para o erro do tipo 1.

O gráfico da figura 5 mostra o número absoluto de palavras transcritas em função da relação sinal-ruído de todas as quatro amostras de cada condição de ruído, incluindo o nível "0", ou seja, sem nenhum ruído adicionado. As retas de regressão mostram que a variação na queda no número de palavras transcritas conforme a diminuição da relação sinal-ruído não é similar quando comparamos os dois tipos de ruído. A análise de regressão do número de palavras transcritas tomado em função do logaritmo do valor da relação sinal-ruído. Para a condição de ruído "fala", a reta de regressão encontrada foi significativa [$F(1, 42) = 92,4; p < 0,001$], com R^2 de 0,69, coeficiente de inclinação igual a 33,4 [$t(42) = 9,71; p < 0,001$] e intercepto igual a -41,4 palavras [$t(42) = -4,26; p < 0,001$]. Para a condição de ruído "tráfego", a reta de regressão encontrada foi significativa [$F(1, 42) = 34,9; p < 0,001$], com R^2 de 0,45, coeficiente de inclinação igual a 6,48 [$t(42) = 5,91; p < 0,001$] e intercepto igual a 44,1 palavras [$t(42) = 18,8; p < 0,001$]. Os resultados apontam que tanto os interceptos quanto o coeficiente de inclinação das duas retas são bastante diferentes: na condição "fala", o aumento na relação sinal ruído provoca um aumento maior no número de palavras transcritas do que na condição "tráfego"; a diferença no intercepto indica que na condição "fala" o número de palavras transcritas chegaria a zero com a queda na relação sinal-ruído muito antes do que na condição "tráfego".

É interessante notar que os valores da relação sinal-ruído da condição "tráfego" dos níveis "1" a "3" são todos menores do que os níveis "1" a "3" da condição "fala". Mesmo assim, o gradiente de queda no número de palavras transcritas na condição "tráfego" é significativamente menor em comparação ao da condição "fala". Esses resultados sugerem que o tipo de ruído que é sobreposto à fala-alvo tem um impacto maior do que a relação sinal ruído. Há boas hipóteses para explicar porque o ruído de fala exerce um efeito mais disruptivo na atividade de transcrição: o ruído de tráfego tem um perfil de variação espectral relativamente estável ao longo do tempo, enquanto a fala apresenta flutuações espectrais não previsíveis. Além disso, na situação de transcrição na condição de ruído "fala", a tarefa de concentrar a atenção na fala-alvo e inibir a percepção da fala sobreposta é mais custosa cognitivamente do que a de inibir um ruído uniforme, que, portanto, tem muito menos características em comum com a fala-alvo. Em todos os relatórios apresentados pelos participantes, houve comentários a respeito da maior dificuldade causada pela condição de ruído "fala".

O número total de palavras transcritas por um dos instrutores do curso nas amostras (e usado como referência) é 73 e 65 nas condições "fala" e "tráfego", respectivamente. Como é possível ver na figura 4, houve variação no número total de palavras transcritas pelos participantes, mesmo na amostra sem nenhum ruído adicionado (observar as duas séries de valores mais à direita na figura 4). Na condição "fala", a mediana foi de 73 palavras, com mínimo de 65 e máximo de 76 e na condição "tráfego", a mediana foi de 65 palavras, com mínimo de 61 e máximo de 67. As medianas coincidiram com o total usado como referência. Nos casos em que os participantes registraram um número maior do que o de referência, isso se deveu, em geral, a casos de marcadores conversacionais ou disfluências que foram registrados, embora as instruções recomendassem que casos

como esses fossem ignorados. Houve também casos em que a falante nas amostras a serem transcritas produziu contrações (como “pra”), que foram registradas por alguns dos participantes na forma expandida na transcrição (“para a”).

Ainda a respeito da análise quantitativa, as instruções para a atividade incluíram uma informação adicional, que era o cômputo, em cada nível de ruído, das palavras transcritas corretamente. A ideia era chamar a atenção para o fato de que nem todas as palavras transcritas correspondem ao que de fato foi dito (tomando por base a transcrição no nível 0 de ruído). Apenas uma participante relatou essa informação. Na condição de ruído “tráfego”, o número de erros na transcrição para os níveis 3, 2, e 1 foram, respectivamente 4, 3 e 0. Na condição de ruído “fala”, o número de erros nos níveis 3, 2 e 1 foram, respectivamente, 17, 13 e 1. Corroborando os resultados anteriores, este mostra que o ruído de tipo “fala” tem poder maior de dificultar a transcrição em relação ao ruído da condição “tráfego”. No caso específico, aumenta a chance não de impedir a audição da fala, mas de uma percepção errada do que foi ouvido.

Além da análise quantitativa, cujos resultados agregados descrevemos acima, as instruções para a atividade de transcrição pediram aos participantes que refletissem linguisticamente a respeito das dificuldades que tiveram para realizar a transcrição: o que foi mais difícil ou impossível de transcrever? O que foi transcrito de forma errada? Algum fator linguístico poderia explicar o erro de transcrição? Discutiremos a seguir os comentários feitos pelos participantes a esse respeito.

Os comentários de caráter mais geral foram na direção de dizer que a dificuldade de transcrição foi maior nos trechos produzidos com taxa de elocução mais alta e menos proeminência prosódica, em especial com menor intensidade. Apontou-se também que palavras gramaticais, aquelas usadas para dar coesão textual e preenchedores conversacionais, especialmente (embora não só) os monossílabos, foram mais difíceis de transcrever – os exemplos arrolados foram palavras como “não”, “que”, “daí”.

Alguns participantes listaram parte dos seus erros de transcrição e esboçaram análises que apontaram possíveis mecanismos linguísticos que poderiam explicar esses erros. Apresentamos a seguir uma seleção deles e das hipóteses levantadas pelos participantes. Trechos entre parênteses são expressões usadas pelos próprios participantes em seus relatórios. Na apresentação dos erros, o que vem antes da seta indica a transcrição correta e o que vem depois o que foi efetivamente transcrito pelo participante.

Uma classe de erros é explicada pela dificuldade de categorização fonética de certos sons em função do ruído:

8. *daí* → *aí*: “não foi possível ouvir a explosão (soltura da oclusiva) do /d/”
9. Dificuldade no trecho *que não sei o quê*, causada pela “predominância de consoantes desvozeadas”, cuja audição teria sido dificultada na condição de ruído de tráfego.
10. *ficando assim um tempo parada* → *ficando separada*: dificuldade de percepção da fricativa /s/ em *assim*.
11. *que não sobra* → *e não sobra*: ruído da soltura da oclusão do /k/ não foi escutado.
12. *não* → *um*: manteve-se a “percepção do traço da nasalidade” na palavra.

Alguns participantes apontaram o papel importante do processamento *top down* na compreensão da fala, que cria expectativas a partir do contexto dado pela narrativa e preenchem lacunas deixadas quando o processamento *bottom up* do sinal acústico sofre com a degradação causada pelo ruído.

13. 13. *ver se eu ainda tava* → *ver se eu tentava*: dificuldade de percepção da sonorização do /d/ em *ainda* levou ao que o participante chamou de “inferência semântica”.
14. *que, assim, você acaba* → *porque, assim, você acaba*: “tendo escutado a sílaba tônica /ke/ ao lado da palavra ‘assim’, concluí ser ‘porque, assim’. Provavelmente mantive essa escuta errônea porque o contexto semântico também reforçava a hipótese de ser uma oração explicativa”.
15. *você acaba doando sua vida* → *você acaba dando conta*: a troca foi atribuída por conta da “intuição linguística relacionada a *collocations* da língua”, isto é, *acaba* seguido de *dando conta* é uma combinação frequente na língua.
16. 16. *você acaba doando sua vida* → *acaba doando muito*: uma troca semelhante à anterior, mas atribuída à “inferência semântica”.

Considerações sobre os resultados

Estudantes e pesquisadores da área de Letras e Linguística constituem um público especialmente apto a desenvolver conhecimentos na área da Fonética Forense. Em primeiro lugar, o interesse e a curiosidade prévios que demonstram em relação ao assunto tornam esse grupo aberto aos novos conhecimentos. Em segundo lugar, porque a formação acadêmica que esse público recebe o equipa com uma sensibilidade para os fenômenos linguísticos, além de conceitos e um vocabulário técnico que são fundamentais para a pesquisa e o trabalho na área da Fonética Forense. Nos comentários às respostas da primeira atividade pré-síncrona, ressaltamos que os participantes compreendem bem como a natureza do fenômeno linguístico implica que ele se manifeste de maneira muito flexível e adaptada ao contexto, de forma que um mesmo falante pode apresentar padrões muito variáveis tanto do ponto de vista da fala propriamente dita quanto de padrões linguísticos de outros níveis (sintático, escolha lexical etc.). Essa compreensão é muito importante para entender o potencial e as limitações do trabalho na Fonética Forense. A capacidade de entender analiticamente essas dimensões da variabilidade da fala é um dos diferenciais relevantes da atuação de um profissional formado em Letras ou Linguística em equipes que trabalham em Fonética Forense.

Essa aptidão e adaptabilidade a novas tarefas evidencia-se também em outros exercícios, como o de transcrição da fala em situação de ruído. Os resultados mostraram que os participantes têm uma boa percepção das razões linguísticas pelas quais um tipo de ruído é mais prejudicial à tarefa de transcrição e, mais importante, são capazes de identificar mecanismos linguísticos que podem subjazer aos erros de transcrição que cometem. Essa habilidade, em particular, é importante na medida em que isso permite entender que a transcrição no contexto forense não exige apenas um “bom ouvido”, mas um entendimento de que a compreensão da fala implica não só a percepção do sinal de fala quanto sua integração com fatores *top-down*, como expectativas criadas pelo contexto semântico e pragmático. O entendimento dos processos linguísticos que subjazem a transcrição permite um maior realismo em relação à capacidade de um transcritor humano e o entendimento da necessidade de um treinamento técnico rigoroso para realizar a tarefa (FRASER, 2003).

No exercício que simulou a elaboração de perfis linguísticos, os participantes mostraram habilidade de avaliar as amostras de fala de maneira mais objetiva e analítica, sobretudo de aspectos segmentais, e o vocabulário técnico que apareceu foi usado de maneira apropriada. Em relação à qualidade de voz, notamos a predominância de avaliações de natureza impressionística ou o uso de um vocabulário mais específico, como voz rouca ou gutural, embora usados com a mesma carga de ambiguidade encontrada no emprego cotidiano desses termos. Notamos, também, a ausência de menção a características prosódicas, mesmo que feita de forma impressionística e não técnica. Entendemos que essas observações se explicam pela exposição relativamente limitada aos temas de Fonética e Fonologia que parte dos cursos de Letras e Linguística proporcionam aos

estudantes. A Fonética e a Fonologia são tematizadas, no geral, apenas em uma disciplina introdutória no início do curso, nem sempre ministrada por um docente especializado no assunto. Por conta disso, dificilmente assuntos como prosódia e qualidade de voz poderão ser abordados para além do nível mais elementar. Devemos também notar que, na comunidade acadêmica brasileira, o tema da qualidade de voz historicamente foi mais pesquisado pela área da fonoaudiologia e só mais recentemente passou a ser um tema de pesquisa mais frequente entre os linguistas.

Ainda em relação à atividade de elaboração de perfis linguísticos, notamos que a habilidade de identificar a proveniência geográfica dos falantes nas amostras é limitada à indicação das grandes regiões geográficas do país. Ressaltamos, sobretudo, o nível de confusão em relação à amostra 2, de uma falante de Minas Gerais, que recebeu mais indicações de origem incorretas do que corretas. Esse resultado parece incompatível com a autoavaliação dos participantes em relação ao grau de confiança em sua habilidade de identificar a origem de um falante por sua fala. Nessa pergunta, quase 70% das respostas indicaram graus 3 ou 4 para essa habilidade em uma escala entre 1 e 5. O resultado seria incompatível também com a autoavaliação dos participantes em relação ao seu grau de conhecimento de sociolinguística, que quase 60% dos participantes julgaram ser 4 em uma escala de 1 a 5. A interpretação que damos para esses resultados se divide em duas partes. Interpretamos o primeiro resultado, isto é, o fato de que os participantes julgam ter mais habilidade para reconhecer a origem geográfica de um falante do que de fato têm, como: (1) subestimação da diversidade real das variedades regionais ou (2) superestimação da abrangência do próprio repertório de diversidade linguística. Quanto à aparente contradição entre o grau de familiaridade alto com a área da sociolinguística e o desempenho mediano no reconhecimento da origem, nossa interpretação também é dupla: (1) o conhecimento da literatura, dos conceitos e das escolas da área de sociolinguística não é um bom preditor do conhecimento efetivo da variedade linguística do Brasil e (2) apesar da rica tradição de estudo da variação linguística que existe na comunidade acadêmica no Brasil (vide a existência de projetos como o NURC (CASTILHO, 1984), ALIB (AGUILERA; MILANI; MOTA, 2004) e VARSUL (VANDRESSEN, 1995), entre outros), a descrição dessa variabilidade parece não se traduzir, nos estudantes de Letras e Linguística, em um domínio mais consistente dessa variação.

A experiência do curso que descrevemos aqui, aliada às experiências anteriores que ambos os autores já tiveram ministrando cursos semelhantes, nos leva a propor algumas sugestões que podem ajudar a preparar ainda melhor egressos dos cursos de Letras e Linguística para atuar na área de Linguística ou Fonética Forenses. Em primeiro lugar, ampliar o espaço que a Fonética e Fonologia têm atualmente, para permitir a oferta de cursos básicos mais completos e também de disciplinas optativas, onde se possa discutir com mais abrangência e profundidade os temas da área. Em especial, há uma carência muito grande de cursos na área de Fonética Acústica, um requisito básico para atuar na área de Fonética Forense. Além disso, nos parece importante tematizar com mais frequência e mais profundidade a questão da variação linguística, não apenas no nível da

Fonética e da Fonologia, mas em todos os demais. Entendemos que o tema da variação deveria ser valorizado em todas as disciplinas e que mais atividades tematizassem o assunto, de maneira a aumentar significativamente a exposição dos estudantes à gama de variação linguística existente no Brasil e nos demais territórios da lusofonia. Entendemos que grande parte dos materiais didáticos de nível introdutório na área de Fonética e Fonologia voltados para o público universitário poderia explorar a variação nesse componente de maneira mais abrangente.

Esclarecemos que não queremos dizer com essas sugestões que os currículos de Letras e Linguística devam ser transformados em cursos preparatórios ou de especialização em Linguística ou Fonética Forenses. Aspectos mais especializados da área devem ser ensinados por profissionais especialistas com anos de experiência prática no contexto das instituições onde a profissão de perito é exercida. Acreditamos, no entanto, que aquilo que sugerimos representaria um incremento à formação dos estudantes para atuação em diversos campos, não apenas na área da Fonética Forense.

REFERÊNCIAS

AGUILERA, V. de A.; MILANI, G. A. L.; MOTA, J. A. (ed.). *Projeto Atlas Lingüístico do Brasil – Documentos I*. Salvador: ILUFBA–EDUFBA, 2004.

ARANTES, P. *Parantes/porcentagem_palavras: Initial release*. [s.l.] Zenodo, 2020. Disponível em: https://github.com/parantes/porcentagem_palavras. Acesso em: 05 mar. 2021.

BARBOSA, P. A.; CARRENHO, J. M.; CAZUMBA, L. A. F.; CONSTANTINI, A. C.; MACHADO, A. P.; PASSETTI, R. R.; SANCHES, A. P.; SILVA, B. A. *Protocolo Geral para Exame de Comparação de Locutor*. Campinas: Grupo de Estudos em Fonética Forense (GEFF), IEL/UNICAMP, 2018. Disponível em: <https://drive.google.com/drive/folders/1I-bcm5uVLgaiZDQpci8agkJMf2fdULXf>. Acesso em: 05 mar. 2021.

BOERSMA, P. Praat, a system for doing phonetics by computer. *Glott International*, v. 5, n. 9/10, p. 341-345, 2001.

BRESCANCINI, C. R.; GONÇALVES, C. S. O peso da evidência sociofonética na perícia de Comparação de Locutor. In: BARBOSA, P. A.; CAZUMBA, L. A. F.; CONSTANTINI, A. C.; MACHADO, A. P.; PASSETTI, R. R.; SANCHES, A. P. (org.). *Análise Fonético-Forense: em tarefa de Comparação de Locutor*. Campinas: Millenium Editora, 2020. p. 67-87.

CASTILHO, A. T. de. Informações sobre o Projeto de Estudo da Norma Urbana Lingüística Culta (Projeto NURC). *Cadernos de Estudos Lingüísticos*, v. 6, p. 187-190, 1984.

ERIKSSON, A. Tutorial sobre Fonética Forense. *ReVEL*, v. 12, n. 23, p. 297-322, 2014.

FRASER, H. Issues in transcription: factors affecting the reliability of transcripts as evidence in legal cases. *Forensic Linguistics*, v. 10, n. 2, p. 203-226, 2003.

LAVIER, J. *The phonetic description of voice quality*. New York: Cambridge University Press, 1980.

NOLAN, F. *The Phonetic Bases of Speaker Recognition*. Cambridge, UK: Cambridge University Press, 1993.

ROBERTSON, B.; VIGNAUX, G. A.; BERGER, C. E. H. *Interpreting Evidence: Evaluating Forensic Science in the Courtroom*. Oxford: John Wiley & Sons Inc, 2016.

VANDRESSEN, P. O Projeto Varsul: avaliação e perspectivas sobre pesquisas do português falado na Região Sul. Encontro Nacional sobre Língua Falada e Ensino. *Anais...* Maceió: EDUFAL, 1995.

WATT, D.; BURNS, J. Verbal descriptions of voice quality differences among untrained listeners. *York Papers in Linguistics*, v. 12a, p. 1-28, jun. 2012.