

ENSAIO TEÓRICO

FONÉTICA, QUE BICHO É ESSE?

Plínio Almeida BARBOSA  

Instituto de Estudos da Linguagem – Universidade Estadual de
Campinas (UNICAMP)

Sandra MADUREIRA  

Programa de Pós-Graduação em Linguística Aplicada e Estudos da
Linguagem – Pontifícia Universidade Católica de São Paulo (PUC-SP)

Renata Regina PASSETTI  

Programa de Pós-Graduação em Linguística Aplicada e Estudos da
Linguagem – Pontifícia Universidade Católica de São Paulo (PUC-SP)

Cláudia Regina BRESCANCINI  

Escola de Humanidades – Pontifícia Universidade Católica do Rio
Grande do Sul (PUC-RS)



OPEN ACCESS

EDITORES

- Miguel Oliveira, Jr. (UFAL)
- René Almeida (UFS)

AVALIADORAS

- Aline Ponciano (UFRJ)
- Flaviane Svartman (USP)

SOBRE OS AUTORES

- Plínio Almeida Barbosa
Conceptualização, Escrita –
Rascunho Original, Escrita – Análise
e Edição.
- Sandra Madureira
Conceptualização, Escrita –
Rascunho Original, Escrita – Análise
e Edição.
- Renata Regina Passetti
Conceptualização, Escrita –
Rascunho Original, Escrita – Análise
e Edição.
- Cláudia Regina Brescancini
Conceptualização, Escrita –
Rascunho Original, Escrita – Análise
e Edição.

DATAS/ES

- Recebido: 27/10/2020
- Aceito: 27/11/2020
- Publicado: 17/03/2021

COMO CITAR

BARBOSA, P.A.; MADUREIRA, S.;
PASSETTI, R.R.; BRESCANCINI, C.R.
(2021). Fonética, que bicho é esse?
Cadernos de Linguística, v. 2, n. 1, p. 01-
19.

RESUMO

Este ensaio aprofunda temas discutidos na Mesa Redonda “Fonética, que bicho é esse?” conduzida por estes autores no evento ABRALIN Ao Vivo – Linguists Online, no dia 22 de maio de 2020. Assim, são apresentados o objeto de estudo da Fonética, suas três áreas de estudos, a Fonética Articulatória, a Fonética Acústica e a Fonética Auditiva, bem como técnicas de análises que permitem o estudo experimental de cada uma delas; além de suas aplicações para outros domínios comunicativos, como aqueles que emergem da veiculação de emoções na fala, do estudo de aspectos sociais e do contexto forense. Nosso objetivo é fornecer ao leitor e, principalmente, ao estudante da área de Ciências da Linguagem, uma primeira visão geral da área de atuação da Fonética.

ABSTRACT

This essay develops, in greater detail, the topics discussed in the Round Table “Phonetics, what kind of stuff is it?”, which was run by the authors in the Abralin Ao Vivo – Linguists Online event, on May 22, 2020. Three areas of Phonetic Sciences and their associated analytical techniques used in experimental research works are introduced: Articulatory Phonetics, Acoustic Phonetics and Auditory Phonetics. Their application to domains such as the ones which focus on the expression of psychological and sociobiological features and forensic contexts are also considered. Our objective is to provide the readers, especially the Language Sciences students, a more precise picture of the field of Phonetics.

PALAVRAS-CHAVE

Fonética e Seus Ramos; Técnicas de Análise Fonética;
Expressividade da Fala; Fonética Forense; Sociofonética.

KEYWORDS

Phonetics and Its Branches; Techniques In Phonetic Analysis;
Speech Expressivity; Forensic Phonetics; Sociophonetics.

INTRODUÇÃO

É inegável que, em todos os tempos, os sons da fala permeiam a paisagem sonora de nosso dia a dia. Nós os produzimos e os percebemos com variadas formas de expressividade, aliando nossas formas individuais do dizer com as formas do falar em distintas comunidades e ambientes comunicativos em seus variados extratos sociais.

A Fonética é a disciplina que estuda os sons da fala considerando os aspectos mencionados acima, além de muito outros, mas sofre de uma compreensão parcial de quais seriam, concretamente, os temas e métodos de pesquisa que lhe são associados. Muitos ainda a restringem a uma atividade meramente de transcrição da fala pelo uso de símbolos especiais, aliando percepção sonora e uma consciência articulatória do som. Mas, há muitas décadas, as atividades do foneticista são bem mais diversificadas.

O processo de transcrição fonética serve ao menos três principais funções: (1) a de permitir que outros foneticistas possam saber os sons que foram ditos num trecho, uma vez que, mesmo que se tenha acesso ao áudio correspondente, há sons que nem todos podem identificar, por não fazer parte de sua língua ou variedade linguística; (2) a de permitir a montagem de bases de dados com a transcrição sonora, o que possibilita consultas e análises como frequência de um som, frequência de sequências sonoras, entre outros; (3) a de servir de guia para o aprendizado da pronúncia de uma língua estrangeira.

O símbolo, no entanto, não basta para compreender os aspectos da produção ou da percepção de um som que nos é desconhecido, é realmente apenas um guia. Quantos de nós, ao consultar os símbolos fonéticos num dicionário de uma língua estrangeira e uma comparação com os sons de nossa língua, num dicionário bilíngue, ainda ficaram sem saber como realizá-los ou como soam! Por isso, desde cedo os foneticistas se interessaram pelos movimentos articulatórios que são necessários para produzir determinado som, objeto da Fonética Articulatória.

1. SOBRE AS SUBDISCIPLINAS DA FONÉTICA

Atualmente a Fonética Articulatória se serve de instrumentos que permitem a observação de aspectos diferenciados da dinâmica dos articuladores da fala. As técnicas mais comuns são: (1) a eletropalatografia, que permite saber os pontos de contato da língua com o palato duro e assim lançar luzes sobre os pontos de articulação dos sons; (2) a eletromiografia, que permite conhecer a atividade mioelétrica especialmente dos músculos da língua e descobrir seu posicionamento ao longo do tempo; (3) a eletroglotografia, que traça indiretamente a atividade vibratória das pregas vocais; (4) a ultrassonografia, que permite a observação 2D da superfície da língua de modo mais direto do que na eletromiografia; (5)

a articulografia eletromagnética, que permite observar em 3D, via eletrodos em posições previamente definidas do trato e da cabeça, o movimento de articuladores como língua e lábios; (6) a ressonância magnética, que permite a observação de todos os articuladores da fala num plano definido, frequentemente o sagital, na ressonância bidimensional, embora tenha restrições para movimentos muito rápidos, por conta da taxa de imagens por segundo que é capaz de produzir e (7) a pletismografia respiratória de indutância, que serve para o estudo da coordenação entre a fala e a respiração durante a fala a partir da medida na área da seção transversal tanto da caixa torácica quanto do abdômen por meio de duas cintas, uma na altura das axilas e outra na altura do umbigo.

Essa panóplia de instrumentos permite o avanço do conhecimento de como são feitos determinados sons. Por exemplo, permite saber que um “r” caipira pode ser feito com a retroflexão da lâmina da língua ou pelo recuo de seu corpo; que a posição da língua já na altura do “p” de “prata” vs. “preta” é aquela da vogal tônica. Esse conhecimento pode auxiliar, por exemplo, no ensino de pronúncia em língua estrangeira, orientando como os sons são feitos, o que acaba contribuindo também para a percepção desses sons.

Se não se dispõe de instrumentos como esses, que normalmente são caros ou podem ser de difícil acesso para o leigo, a observação das características acústicas dos sons pode ser feita dispondo de uma simples gravação.

A Fonética Acústica investiga as características materiais do próprio som. Um dos primeiros instrumentos para se examinar a composição sonora foi o espectrógrafo, construído durante a Segunda Grande Guerra. De lá para cá, de analógico passou a ser um simples algoritmo de um programa de software. Grandes avanços foram conseguidos com esse instrumento nas décadas e 1940 a 1960, com os primeiros experimentos de percepção usando fala sintética. A Figura 1 mostra um espectrograma de um trecho da palavra “chaga” em que se exhibe a sílaba tônica.

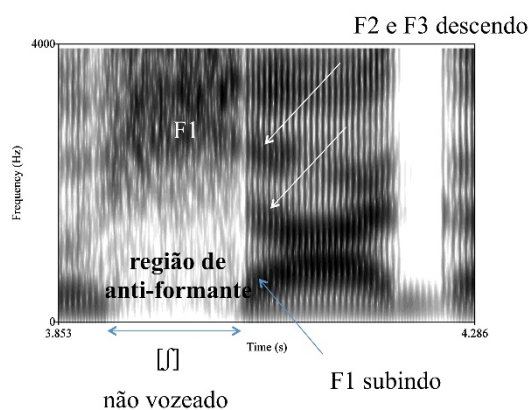


Figura 1. Espectrograma da sílaba tônica de “chaga” de locutor mineiro.

Observe que o espectrograma é um gráfico com o tempo no eixo x, com a frequência no eixo y e com o nível de energia mostrado pelo grau de escurecimento numa determinada faixa de frequência. As regiões assinaladas por F1, F2 e F3 são zonas de mais energia que ocorrem tanto na consoante [ʃ] (o “ch”) quanto na vogal [a]: são as ressonâncias desses sons. As estrias verticais durante a vogal [a], onde foram assinalados F1, F2 e F3, indicam a vibração das pregas vocais, padrão que não se encontra na consoante, pois não há vibração das pregas durante sua produção. Abaixo da região de F1 da consoante, há uma zona clara que assinala a presença de menos energia, aspecto que também a caracteriza. Todos esses pormenores nos permitem conhecer aspectos finos da produção dos sons com apenas a gravação do trecho que se deseja analisar.

A Fonética Acústica permite conhecer detalhes da produção sonora sem a necessidade de registro direto da articulação, reservando à Fonética Articulatória a tarefa de dirimir alguma dúvida sobre a forma como os articuladores se coordenam para a realização de um som. Ao espectrograma se somam outras técnicas, como o espectro de Fourier, os algoritmos que calculam as frequências dos formantes e descrevem aspectos da atividade laríngea, incluindo o algoritmo que extrai a taxa de vibração das pregas vocais a cada ciclo. No entanto, mesmo com todas essas técnicas não é possível avaliar a forma como percebemos esses sons, que cabe à Fonética Auditiva.

A Fonética Auditiva se serve do conhecimento de como identificamos e diferenciamos os sons entre si, através do estudo do processamento do som pelo sentido da audição, para apontar o conjunto de sons passíveis de serem usados na nossa comunicação. Por exemplo, se sabe que não é qualquer diferença em taxa de vibração das pregas vocais que percebemos como distintas. Na década de 1980, t'Hart (1981) faz experimentos que indicam que, na fala, é preciso uma diferença de cerca de 2 a 4 semitons, unidade musical, para percebermos que dois tons que usamos para dar ênfase em dois momentos são de fato diferentes, sendo que um é mais agudo do que o outro. Por conta do funcionamento de nossa audição, as frequências das ressonâncias já mencionadas são melhor discriminadas quando menores do que 3000 Hz do que acima desse valor. O que se passa em nosso sistema auditivo é um reflexo de como funciona nosso cérebro durante a percepção dos sons, do ritmo e da entoação da fala.

Todas essas subdisciplinas da Fonética ajudam a construir um conhecimento do universo sonoro que encontra muitas aplicações e que se abre para outros domínios comunicativos. Aqui nos referiremos aos estudos sobre a expressividade na fala para entender a variedade e riqueza de sentidos geradas pelo uso da criatividade na produção sonora, à Fonética Forense, que aplica o conhecimento fonético-acústico e sociolinguístico para entender como indivíduos diferem em termos sonoros e à Sociofonética para entender os aspectos sociais ou a vida social dos sons da fala.

2. SOBRE A EXPRESSIVIDADE DA FALA

Na análise fonética da expressividade da fala, percorremos o caminho do som ao sentido, pois, central a qualquer consideração sobre a expressividade da fala, está o papel das propriedades da materialidade fônica na veiculação de efeitos de sentido. A materialidade fônica compreende as propriedades físicas dos sons e os efeitos de sentido se referem às associações semânticas que se estabelecem a partir da percepção das propriedades físicas do som.

Evidências robustas a favor da existência de vínculos motivados entre som e sentido são apontadas por Blasi *et al.* (2016) em seu estudo com cem itens lexicais do vocabulário básico em dois terços das línguas do mundo e interpretadas pelos autores como reflexo da preferência cultural, histórica e geográfica pela comunidade humana.

O foco da análise da expressividade da fala recai sobre o detalhe fonético e sua valoração sonora simbólica, indexical ou sinestésica. Consideramos aqui o simbólico para se referir às associações motivadas de natureza icônica ou imitativa, por exemplo, na atribuição do sentido de vastidão a uma emissão de som longa; o indexical para se referir às associações entre as características sonoras e pessoais, sejam elas de natureza psicológica, social, orgânica ou estilística, por exemplo, os julgamentos de tipo de personalidade, nível social, condições de saúde, ou característica de estilo de fala; e as sinestésicas que remetem a associações multimodais e multisensoriais, visto que abrangem as variadas linguagens de expressão e os cinco órgãos dos sentidos humanos, como quando associamos na poesia o som do [u] a algo lúgubre ou à cor azul ou o som do [i] a objetos pontudos.

Conhecimentos sobre a articulação e suas consequências acústicas e perceptivas são necessários para se examinar o detalhe fonético. Na produção dos segmentos fônicos, dois tipos de fontes estão presentes: o vozeamento e o ruído. A fonte de voz caracteriza todos os sons ressoantes (laterais, vibrantes, tenses, nasais, aproximantes e vogais) e os sons obstruintes vozeados, que englobam os oclusivos, fricativos e africados vozeados. Os ruídos transientes, isto é, de curta duração, estão presentes nas consonantais oclusivos, os contínuos nos fricativos e ambos nos africados. Do ponto vista expressivo, estabelece-se, dessa maneira, a expressividade sonora, o confronto entre sonoridades melodiosas e ruidosas e entre ruídos explosivos e contínuos, que vão ser explorados na prosa, na poesia e no discurso oral para a criação de efeitos de sentido.

Ressalta-se que cada segmento fônico ao ser comparado com outro pode ter algumas propriedades compartilhadas e outras não. Se compararmos, por exemplo, as ressoantes nasais e laterais com as obstruintes fricativas e as oclusivas, verificamos que continuidade é compartilhada entre as fricativas, as laterais e as nasais, mas não a propriedade do ruído, que é compartilhada entre fricativas e oclusivas. A observação sobre a combinatória de

características fonéticas é referida por Tsur e Gafni (2019) como *double-edged* pelo fato de que os sons da fala têm múltiplas características, cada uma exibindo um potencial expressivo, isto é, expressão múltiplos polos do sentido.

A expressividade sonora também reflete a sonoridade das características prosódicas, o modo de falar (BARBOSA, 2019), que confere proeminências e padrões rítmicos e entoacionais e pausas à fala. Os correlatos físicos da prosódia são a frequência fundamental, a duração, a intensidade, o silêncio e a qualidade de voz. (Barbosa, 2019). Variações acústicas de frequência fundamental, de duração e de intensidade são determinadas, respectivamente, por vibrações das pregas vocais, faseamento de gestos articulatórios e força expiratória e provocam, também respectivamente, sensações auditivas de *pitch* (eixo grave/agudo), quantidade e volume (*loudness*). O nível prosódico interage com o segmental contribuindo para a caracterização de determinadas expressividades sonoras, como a fala carismática, por exemplo. Um nível de *pitch* elevado, o *timing* e a configuração dos acentos de *pitch* (*pitch accents*) combinados com segmentos vocálicos não reduzidos e com menos ocorrências de redução vocálica e de assimilação consonantal de ponto de articulação são apontados por Niebuhr, Thumm e Michalsky (2018) como aspectos relevantes para a caracterização da fala carismática.

Vozeamento, ruídos transientes e contínuos e correlatos físicos da prosódia presentes na fala conferem à materialidade fônica uma textura sonora que suscita no ouvinte a atribuição de efeitos de sentido. Revela-se, dessa maneira, o simbolismo sonoro, a base da expressividade sonora, pois os efeitos de sentido atribuídos pelos ouvintes às emissões de fala são suscitados pelas características acústicas dos sons que derivam das manobras articulatórias e fonatórias, e impressionam os ouvidos (MADUREIRA, 2019).

Dois conceitos relevantes para se compreender a expressividade da fala são os códigos sonoros e as metáforas sonoras. Quatro tipos de códigos sonoros são mencionados na literatura fonética: o código de frequência (OHALA, 1997) e o código sirênico (GUSSENHOVEN, 2016), o de esforço e o de respiração (Gussenhoven, 2004). As metáforas sonoras, termo cunhado por Fónagy (1983), por sua vez, são definidas como gestos articulatórios expressivos que emergem de sinestésias entre os cinco sentidos do homem (FÓNAGY, 2001).

As emissões de uma narradora ao contar a lenda “A festa no céu” ilustram bem o conceito do código de frequência: valores altos e extensão aumentada de f_0 , ao fazer a voz do sapo jovem e animado e valores baixos de f_0 ao fazer a voz do sábio sapo velho. Na Figura 2, a seguir, os traçados de frequência fundamental de três emissões da narradora podem ser confrontados.

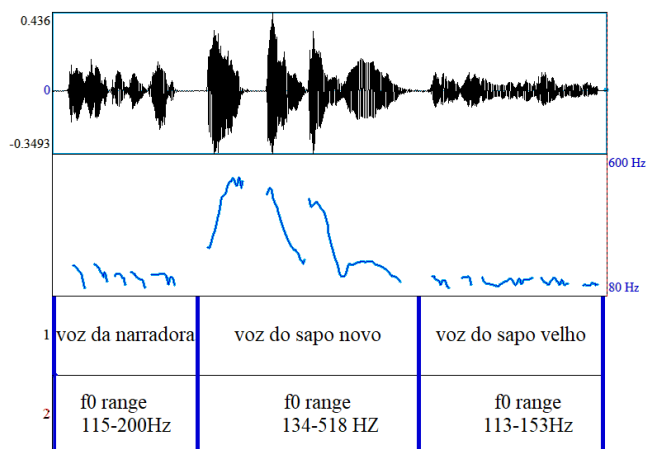


Figura 2. Oscilogramas e traçados de frequência fundamental das emissões das frases “Tinha sido contemplado” (voz da narradora), “Eu também vou” (voz do sapo novo) e “É melhor para a sua saúde não sair do chão” (voz do sapo velho).

Do mesmo modo, se confrontarmos as emissões com ajustes de qualidade de voz relaxado e tenso, verificaremos contrastes em termos de efeitos de sentido de tranquilidade e nervosismo que podem ser explicados por meio dos códigos de esforço e de respiração. O poder de atratividade da voz com escape de ar, como explicitado por Rosenberg e Hirshberg (2020), por sua vez, demonstra concordância com os pressupostos do código sirênico, e a falta de atratividade de vozes tensas (*harsh*), que se caracteriza por apresentar irregularidade, pode ser entendida a partir da constatação de presença de não atratividade de vozes com aperiodicidade (BELIN, 2020).

Associações sinestésicas entre som, cores, odores, sabores, formas geométricas e sensações táteis remetem às metáforas sonoras, conceituação que permite explicar relações em que se estabelecem associações, como a associação entre a vogal posterior alta e a cor preta (NEWMAN, 1933; TSUR, 1992), entre a vogal anterior alta e o tamanho diminuto (SAPIR, 1929; NEWMAN, 1933; CHUENWATTANAPRANITHI *et al.*, 2008), vogais com *pitch* agudo em palavras com sentido proximal em oposição ao distal (WOODWORTH, 1991), sons faríngeos e sensação de náusea (FÓNAGY, 2001), associações de formas arredondadas com sons ressoantes e pontiagudas com sons plosivos (KOHLE, 1929) e associações entre qualidades de voz *harsh* e laringalizada (*creaky*) com texturas ásperas (MOOS *et al.* 2013). Uma inclusiva apreciação de estudos das associações sinestésicas entre som e experiências ligadas aos diversos órgãos dos sentidos é encontrada em Nobili (2019), trabalho que também engloba evidências a favor do simbolismo sonoro a partir de achados de pesquisa com ressonância magnética funcional.

Em termos de associações entre som e sentido que apresentam valoração indexical, abrem-se perspectivas para trabalhos de natureza sociofonética e forense. Ouvintes, leigos ou não, atribuem, a partir de características da fala dos sujeitos, julgamentos sobre: nacionalidade, regionalidade, escolaridade, características corpóreas, condições físicas,

atitudes, estados emotivos e atividades de fala espontânea, leitura, narração ou comentário. A comparação entre as características de fala de locutores também é viabilizada por meio de análise de natureza fonética experimental, envolvendo julgamentos de análise perceptiva e acústica.

Considerações sobre a metodologia de pesquisa de natureza fonética experimental em expressividade de fala são encontradas em Barbosa (2009), Barbosa e Madureira (2016), Silva e Barbosa (2017) Madureira, Fontes e Camargo (2019), Madureira e Fontes (2019) e Barbosa *et al.* (2020). Características auditivas, acústicas, semânticas, respiratórias e visuais são contempladas por meio de variadas técnicas instrumentais que permitem a visualização e a inferência de mecanismos de produção, como as mencionadas na seção 2 deste trabalho.

3. SOBRE FONÉTICA FORENSE

Enquanto falantes (e ouvintes) de uma língua, temos a habilidade de reconhecer indivíduos unicamente por características de sua voz e fala. Essa habilidade, frequente em nossas interações diárias, é desempenhada com exatidão quando associada ao reconhecimento de vozes familiares. Experienciamos isso, por exemplo, quando atendemos ao telefone ou ligamos o rádio e, de imediato, reconhecemos quem nos ligou ou o cantor da música que está tocando, mesmo que a letra nos seja desconhecida.

Ao longo da história, o reconhecimento de locutores também se fez presente no campo jurídico, como peça pericial em investigações criminais. No início do século XX, as práticas forenses envolvendo amostras de fala eram essencialmente de base auditiva-perceptiva e apoiavam-se, sobretudo, em depoimentos de testemunhas auriculares. Entretanto, o uso de uma tarefa auditiva em contextos forenses, principalmente por ouvintes não treinados, e o peso dado a ela em julgamentos, levantaram questionamentos sobre sua acuidade, uma vez que a eficácia de um reconhecimento de locutor também é influenciada por fatores como a atenção dirigida pelo ouvinte à voz do locutor-alvo, a quantidade de tempo em que o ouvinte ficou exposto à voz e, ainda, as circunstâncias em que essa exposição ocorreu.

Assim, era preciso que essas análises auditivas fossem aprimoradas por técnicas que lhe conferissem exatidão. Essas melhorias foram possibilitadas pelo avanço da tecnologia e dos estudos fonéticos, que trouxe para a área forense ferramentas que permitiam a análise acústica dos sons da fala, como o espectrógrafo. Elegeu-se, então, a Fonética como a área de estudos linguísticos cujos conhecimentos teóricos e métodos de análise supriam as lacunas existentes na área de reconhecimento de locutor forense. É nesse contexto que surge a Fonética Forense.

Desde 1991, a Fonética Forense possui reconhecimento oficial conferido pela *International Association for Forensic Phonetics and Acoustics* (IAFPA), que promove a área por meio de conferências anuais e publicações científicas no *International Journal of Speech, Language and Law* e tem sede em York, na Inglaterra.

Atualmente, o exame de comparação de locutor é a principal tarefa da Fonética Forense. Por meio dele, comparam-se os padrões acústicos e sociolinguísticos da fala de um locutor associado a um crime (amostra de fala questionada) com aqueles da fala de um locutor suspeito (amostra de referência ou padrão) a fim de determinar a probabilidade de as duas amostras serem provenientes de um mesmo indivíduo.

Porém, comparar duas amostras de fala está longe de ser uma tarefa simples. Isso porque o potencial biométrico do trato vocal e, por conseguinte, da fala e da voz é afetado pelo fenômeno da variabilidade. Dentre as principais fontes de variabilidade no sinal da fala relatadas em análises fonético-forenses podemos destacar aquelas que se manifestam em função de diferenças de qualidade de voz, provenientes de modificações no trato vocal nos níveis fonatório, articulatório e de tensão muscular (cf. LAYER, 1980); dos estados emocionais, distintos nas amostras de fala questionada e de referência, pois aquela frequentemente envolve situações de tensão e estresse emocional, e físico do falante, podendo este ser afetado por resfriados e consumo de álcool, por exemplo; e de diferenças fisiológicas, como aquelas associadas a tratos e pregas vocais de diferentes dimensões e massa ou a alterações hormonais.

Soma-se a esses fatores de variabilidade o impacto da não contemporaneidade das amostras de fala comparadas, pois alterações, principalmente na voz, podem ocorrer a depender do intervalo de tempo entre a obtenção das amostras de fala questionada e de referência e, conseqüentemente, afetar a precisão da análise de parâmetros associados à qualidade vocal, por exemplo (cf. ERIKSSON, 2005).

Outros desafios que se colocam ao exame de comparação de locutor estão relacionados aos efeitos da transmissão telefônica e de disfarces de voz. Grande parte das amostras de fala questionadas é proveniente de interceptações telefônicas e, por conta das características da transmissão e do filtro telefônico, essas amostras possuem qualidade acústica inferior à das gravações das amostras de fala de referência, para as quais geralmente se utiliza equipamentos com alta precisão acústica.

Os aparelhos celulares possuem um filtro passa-faixas que filtra a energia espectral do sinal sonoro numa faixa de frequências entre 300 e 3400 Hz, aproximadamente. Assim, toda informação espectral que excede esses limiares é perdida e valores de frequências próximos a eles sofrem distorções, responsáveis por aumentar aqueles que estão próximos ao filtro passa-baixa (300 Hz) e diminuir os próximos ao filtro passa-alta (3400 Hz) (cf. ROSE, 2003). A Figura 3 mostra as formas de onda e os espectrogramas da palavra “ciclo”

proferida por um mesmo falante que foi gravado simultaneamente de forma direta (à esquerda) e via celular (à direita).

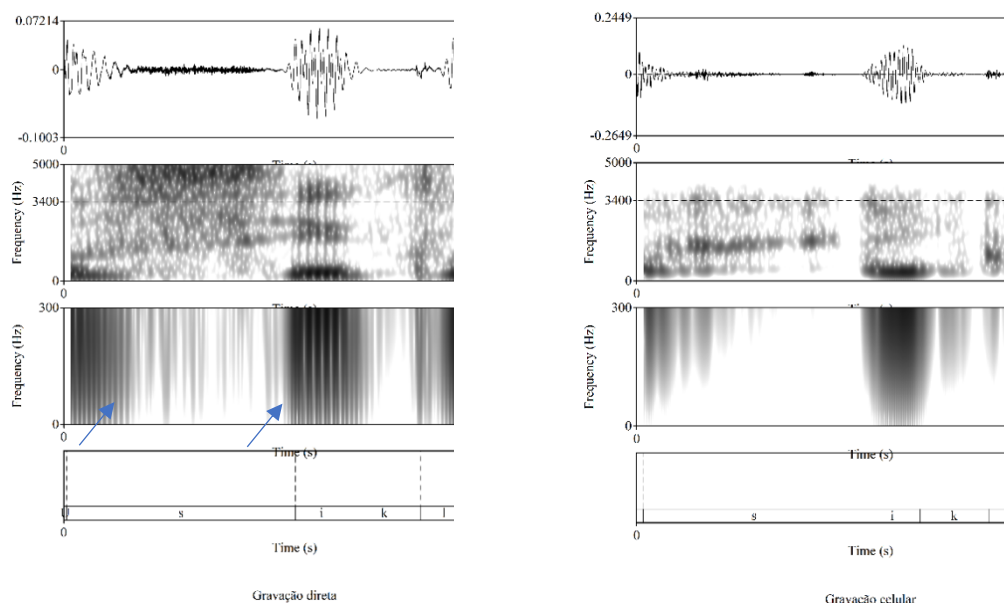


Figura 3. Formas de onda (acima) e espectrogramas até 5000 Hz (intermediário) e até 300 Hz (abaixo) da palavra “ciclo” de um falante paulista (Fonte: *corpus* “Sem Fio” (PASSETTI, 2015)).

Observe que nos espectrogramas da gravação telefônica há perda da informação espectral a partir dos 3400 Hz e abaixo dos 300 Hz, o que não acontece na gravação direta, cuja taxa de amostragem é de 22050 Hz. Embora haja ausência de informação espectral nos trechos apontados pelas setas no espectrograma de 0 a 300 Hz da gravação direta, ela se deve a regiões anti-formânticas do [s] e à fase de oclusão de [k], respectivamente, e não a efeitos de filtragem. Comparando esse espectrograma com o da gravação telefônica, ambos de banda larga, nota-se a impossibilidade de distinguir as estrias verticais dos trechos vocálicos, em razão da ação do filtro passa-baixa (300 Hz) do telefone.

A comparação entre os espectrogramas intermediários (faixa frequencial de 0 a 5000 Hz) das gravações direta e telefônica revelam, ainda, o efeito do ruído e da distorção espectral causado pela transmissão telefônica. Na gravação telefônica, a inspeção visual das regiões formânticas das vogais [i] e [v] em regiões frequenciais mais elevadas e próximas ao filtro passa-alta (3400 Hz) é altamente prejudicada. Além disso, por serem vogais altas, os valores de frequência de F1 dessas vogais situam-se próximos ao filtro passa-baixa, o que faz com que o efeito de distorção espectral os aumente.

Os disfarces de voz, por sua vez, apesar de não serem frequentes no cenário forense, podem causar sérios danos às tarefas fonético-forenses (cf. ERIKSSON, 2010). Dentre os tipos

mais comuns relatados na literatura, estão aqueles que envolvem modificações entoacionais pelo aumento ou diminuição de *pitch*. Outros tipos de disfarces já reportados estão relacionados ao uso de objetos na boca, à obstrução das vias nasais ou oral, a alterações na qualidade de voz e à imitação de indivíduos ou sotaques (KÜNZEL, 2000).

A Figura 4 mostra os cepstros da fricativa alveolar surda ([s]) na sílaba tônica da palavra “licença” produzida por uma falante com sua voz habitual (linha preta) e usando um lápis atrás dos seus dentes incisivos como disfarce (linha vermelha).

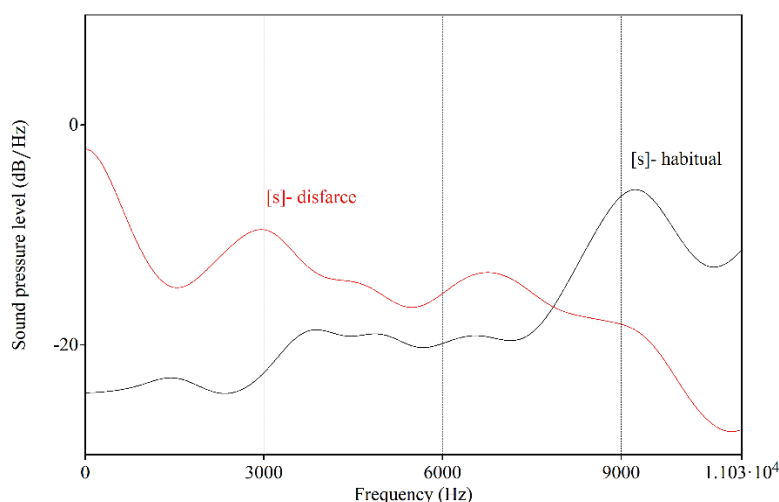


Figura 4: Cepstros da fricativa [s] na sílaba tônica da palavra “licença” de uma falante paulista nas situações disfarce (vermelho) e voz habitual (preto) (Fonte: corpus de fala de Passetti (2012)).

Observe que a concentração de energia do primeiro formante do [s] na amostra com a voz habitual da falante está por volta dos 9000 Hz, enquanto que, na situação de disfarce, a concentração de energia da mesma fricativa aproxima-se dos 3000 Hz. Este valor está mais próximo da faixa frequencial esperada para a fricativa pós-alveolar surda ([ʃ]), articulada em posição posterior do trato oral. A diferença nos valores relaciona-se à presença do lápis atrás dos dentes incisivos, o que dificulta a realização do alvo articulatorio do [s] na região dos alvéolos. Assim, possivelmente, a falante produziu essa consoante de forma mais retraída, próxima à região pós-alveolar, o que, inclusive, a fez soar como um [ʃ] (algo como “linchença”).

Os efeitos das fontes de variabilidade na fala apresentados aqui demonstram a importância de uma sólida formação em Fonética Acústica para a adequada condução de tarefas fonético-forenses. É essencial que essa formação esteja acompanhada de conhecimentos nas áreas de Sociolinguística e de Voz, uma vez que, por meio daquela, conhece-se a distribuição de traços fonéticos na população e os efeitos da variação condicionada por elementos internos e externos ao sistema linguístico (cf. FOULKES; SCOBIE; WATT, 2010). O conhecimento na área de Voz, por sua vez, permite

compreender aspectos fonatórios e de tensão relacionados ao subsistema laríngeo com potencial discriminatório.

Ainda que incipiente no Brasil, a área de Fonética Forense e a atuação de linguistas vêm se consolidando no país, com destaque para o crescente número de publicações científicas com dados do português brasileiro e para a atuação de grupos de pesquisas de universidades brasileiras na formação de pesquisadores e peritos na área.

4. SOBRE SOCIOFONÉTICA

Como campo de estudo de intersecção entre a Fonética e a Sociolinguística, a Sociofonética é caracterizada como fruto da integração dos princípios, técnicas e pressupostos desses dois ramos da Linguística. Estudos que se identificam como sociofonéticos, como Preston e Niedzielski (2010), Thomas (2011), Di Paolo e Yaeger-Dror (2011) e Celata e Calamai (2014), entre outros, revelam a Acústica como o ramo da Fonética que compõe um dos conjuntos em intersecção. Quanto ao outro ramo da Linguística, o da Sociolinguística, é o modelo da Teoria da Variação (LABOV, 1972) que mais claramente revela sua vocação sociofonética, como é possível constatar desde a publicação de *A quantitative study of sound change in progress*, de Labov, Yaeger e Steiner, em 1972, sobre variação e mudança de vogais do inglês americano.

Estudos subsequentes comprovaram que o exame do detalhe fonético fino pode espelhar, dentre outros aspectos sociolinguísticos, mudanças na estrutura social de uma comunidade ou perda de traços culturais de um grupo. Harrington, Palethorpe e Watson (2000) constataram o primeiro caso de detalhe fino a partir da análise das vogais do inglês britânico produzidas pela rainha Elizabeth II durante sua mensagem anual de Natal entre a década de 50 e de 80.

Do ponto de vista social, o estudo confrontou dois momentos distintos da Inglaterra: os anos 50, com estratificação social marcada, e os anos 80, com demarcação social comparativamente bem menos rígida. Do ponto de vista linguístico, foram comparadas duas variedades do Inglês Britânico: a da rainha, o RP (*Received Pronunciation*), o sotaque das classes mais altas, e a variedade da classe média, o SSB (*Standard Southern British*) dos anos 80, falado pelos mais jovens e pela classe média. Assim, além da amostra composta pelas mensagens de Natal, com mesmo estilo de fala, duração aproximada de gravações e boa qualidade de áudio, os autores examinaram também as vogais da classe média, obtidas a partir de uma amostra de fala de mulheres que, nos anos 80, atuavam como apresentadoras de programas da BBC.

Foram consideradas 11 vogais, [i ɪ ε ʌ a ɒ æ ɔ ɜ ʊ u], em uma amostra analisada acusticamente através das medidas de F1 e F2. Os resultados indicaram que, de fato,

houve mudança na posição das vogais produzidas pela rainha: as vogais produzidas na década de 80 localizaram-se, no espaço acústico, entre as vogais produzidas na década de 50 e as vogais representativas da classe média. Adicionalmente, as mudanças sonoras constatadas na variedade da classe média nos últimos 50 anos, especificamente o abaixamento de /æ/ e a anteriorização de /u/, mostraram-se como uma tendência na fala da rainha nos anos 80, indicando assim que, de fato, a mudança gradual verificada nas vogais produzidas pela rainha parece espelhar as mudanças na sociedade inglesa dos anos 50 aos 80.

O segundo caso foi constatado por Mileski (2017) através da análise acústica das vogais tônicas do português produzidas por 16 indivíduos, entre 30 e 70 anos, nascidos e residentes na Serra Gaúcha, em localidades de colonização polonesa. Todos eram bilíngues (português/polonês), embora com graus variados de bilinguismo.

A variação vocálica em posição tônica não é esperada no Português Brasileiro, diferentemente da variação vocálica em posição átona. No sul do Brasil, onde o estudo foi realizado, as vogais átonas grafadas com as letras “e” são percebidas, de oitiva, sem grande dificuldade, ora como [e] e ora como [i], até mesmo na fala de um mesmo indivíduo. As grafadas “o”, ora como [o] e ora como [u]. São exemplos “leit[**e**] ~ leit[**i**]” e “carr[**o**] ~ carr[**u**]”, em posição pós-tônica final; “núm[**e**]ro ~ núm[**i**]ro” e “abób[**o**]ra ~ abób[**u**]ra”, em posição pós-tônica não final e “m[**e**]nino ~ m[**i**]nino” e “c[**o**]ruja ~ c[**u**]ruja”, em posição pré-tônica. No entanto, não se percebe, de oitiva, variação das vogais grafadas “e” e “o” quando estão na posição tônica, como em “gota” e “dedo”.

A observação da fala dos descendentes de poloneses na Serra Gaúcha revelou tanto um leve abaixamento das tônicas “e” e “o”, como em “c[**e**]do” e “perig[**o**]so”, difícil de se perceber de oitiva, como também um abaixamento mais nítido, como o verificado em “s[**ɛ**]xta” e “d[**ɔ**]ze”, o que se explica pelo fato de que o sistema vocálico do polonês apresenta seis vogais, /i ɪ ɛ a ɔ u/, diferenciando-se do sistema do português pela ausência das médias-altas /e/ e /o/.

A análise acústica das ocorrências referentes tanto à vogal /e/ quanto à vogal /o/, no entanto, apontou a correlação positiva entre os valores de F1 e a idade dos falantes, indicando que as vogais tônicas tendem a ser produzidas como mais baixas pelos indivíduos com mais idade. Tal resultado revela que o abaixamento das vogais se dá em graus variados na comunidade, tanto para “e” quanto para “o”, e que são os indivíduos mais velhos, que utilizam frequentemente o polonês em sua vida diária, os que produzem essas vogais tônicas mais abaixadas. Como os mais jovens produziram menos o abaixamento dessas vogais, é possível prever o desaparecimento desse traço da língua dos descendentes dos imigrantes poloneses no futuro.

Harrington, Palethorpe e Watson (2000) e Mileski (2017), estudos aparentemente bem diferentes, têm em comum o fato de considerarem o detalhe fonético fino como

um elemento de expressão da construção social dos indivíduos. Embora os casos de variação citados envolvam propriedades segmentais, entende-se, na perspectiva sociofonética, que as propriedades prosódicas, especialmente temporais, assim como as segmentais, são inerentes à produção fonética, apresentam realidade psicoacústica e portam significado sociolinguístico.

Um dos principais desafios para a investigação sociofonética do detalhe fonético fino é, certamente, a definição dos procedimentos metodológicos a serem adotados nos estudos, tarefa que exige a conjugação da orientação do campo da Fonética Acústica e do da Sociolinguística. O Quadro 1 a seguir ilustra três possíveis diferenças referentes a técnicas de amostragem entre esses dois campos.

Sociolinguística Variacionista	Fonética Acústica
Amostra coletada por meio de entrevistas de experiência pessoal	Amostra coletada por meio de instrumentos
Amostra estratificada por características sociais e demográficas	Amostra geralmente composta por indivíduos de mesmo sexo e idade
Grande quantidade de participantes e de dados	Poucos participantes e grande quantidade de dados

Quadro 1. Orientações Metodológicas: Sociolinguística Variacionista e Fonética Acústica.

Com base na primeira linha do Quadro 1, pode-se afirmar que a busca pelo vernáculo, objeto da Sociolinguística Variacionista, orienta para a formação de amostras coletadas por meio de entrevistas de experiência pessoal, menos comuns nos estudos em Acústica, os quais, devido à preocupação com o controle dos contextos linguísticos, valem-se, muitas vezes, de instrumentos de coleta. Com relação à segunda linha, vê-se que o método sociolinguístico envolve a consideração de dimensões sociais, como gênero, etnia, classe social, faixa etária, anos de escolarização, localidade de origem e moradia, por exemplo, não exigidas nos estudos em Fonética, os quais podem se valer de amostras compostas por indivíduos de mesmo sexo e de idades semelhantes. Por último, no Quadro 1, é possível apontar uma semelhança, a quantidade de ocorrências para análise, que deve ser considerável para os dois campos. No entanto, nos estudos de variação fonético-fonológica são previstos muitos participantes e um número menor de ocorrências por participante, diferentemente dos estudos puramente fonéticos, que consideram poucos participantes, mas muitas ocorrências por participante.

A construção de uma metodologia híbrida não é, no entanto, o único desafio apresentado. Conforme apontam Foulkes, Scobbie e Watt (2010), a interface com outros campos relacionados, como a Psicolinguística, a Linguística Clínica, a Linguística Forense, a Aquisição de L1 e de L2, a Fonologia e a Linguística Computacional, também impõe o confronto de abordagens teóricas, além da implementação de recursos metodológicos para coleta e análise de dados. Estudos conduzidos nessa perspectiva revelam o potencial

de aplicação da Sociofonética para a solução de problemas da sociedade nas áreas da saúde, da segurança pública, do ensino e da tecnologia.

5. CONSIDERAÇÕES FINAIS

Procuramos mostrar algumas das áreas de estudo e de aplicação da Fonética para que o leitor tenha uma ideia mais precisa de seu campo de atuação deixando de considerá-la como um bicho de sete cabeças. As seções aqui apresentadas aprofundam os temas que discutimos quando de nossa apresentação para o ABRALIN ao Vivo – *Linguists Online*, disponível no YouTube na URL < <https://www.youtube.com/watch?v=O80ZamzKQ3U>>.

No que segue, indicamos uma lista de obras de referência para o estudante se iniciar nas áreas de pesquisa em Fonética ou mesmo para melhor se inteirar de suas aplicações.

6. OBRAS BÁSICAS

- BALDWIN, J. R.; FRENCH, P. *Forensic phonetics*. Londres: Pinter Publishers, 1990.
- BARBOSA, P.A.; CAZUMBÁ, L.A.F; CONSTANTINI, A.C.; MACHADO, A.P.; PASSETTI, R.R.; SANCHES, A.P. [Orgs.] *Análise Fonético-Foreense: em tarefa de Comparação de Locutor*. Campinas: Millennium Editora, 2020.
- BARBOSA, P. A.; MADUREIRA, S. *Manual de fonética acústica experimental: aplicações a dados do português*. São Paulo: Cortez, 2015.
- DI PAOLO, M.; YAEGER-DROR, M. (Eds.). *Sociophonetics: a student's guide*. New York: Routledge, 2011.
- HARDCASTLE, W. J.; LAVER, J. (Eds.). *The Handbook of Phonetic Sciences*. Oxford: Blackwell, 1997.
- HOLLIEN, H. *The acoustics of crime: The new science of forensic phonetics*. Springer Science & Business Media, 2013.
- HINTON, L.; NICHOLS, J.; OHALA, J. J. (Eds.). *Sound symbolism*. Cambridge: Cambridge University Press, 2006.
- JESSEN, M. Forensic phonetics. *Language and linguistics compass*, 2(4), 671-711, 2008.
- JOHNSON, K. *Acoustic and auditory phonetics*. Oxford: Blackwell, 2011.
- KENT, R. D.; READ, C. *The acoustic analysis of speech*. San Diego: Singular Publishing Group, 1992.
- LADEFOGED, P. *Elements of acoustic phonetics*. Chicago: University of Chicago Press, 1996.

- PISONI, D.; REMEZ, R. (Eds.). *The handbook of speech perception*. John Wiley & Sons, 2008.
- SHRIBERG, L. D.; KENT. *Clinical phonetics*. Boston, MA: Allyn and Bacon, 2003.

7. AGRADECIMENTOS

Os autores agradecem à ABRALIN pelo convite para participação no evento virtual ABRALIN ao Vivo – Linguists Online. O primeiro autor, a segunda e a quarta autoras agradecem ao CNPq pela Bolsa de Produtividade concedida. A terceira autora agradece à FAPESP pelo financiamento de duas pesquisas, cujos resultados foram apresentados aqui, referentes aos Processos nº 2011/23054-3 e nº 2013/12516-1, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). Agradecemos também às duas pareceristas.

REFERÊNCIAS

BARBOSA, P. A. Detecting changes in speech expressiveness in participants of a radio program. *Proceedings of the 10th Annual Conference of the International Speech Communication Association (Interspeech 2009)*, Brighton, 2155-2158, 2009.

BARBOSA, P. A.; MADUREIRA, S. Elicitation techniques for cross-linguistic research on professional and non-professional speaking styles. In: *Speech Prosody 2016*, 2016, Boston. Proc. Speech Prosody 2016, p. 503-507, 2016.

BARBOSA, P. A.; MADUREIRA, S.; FONTES, M. A. S.; MENEGON, P. Speech Breathing and Expressivity: An Experimental Study in Reading and Singing Styles. In: QUARESMA P.; VIEIRA R.; ALUÍSIO S.; MONIZ H.; BATISTA F.; GONÇALVES T. (Org.). *Lecture Notes in Computer Science*. Springer International Publishing, 2020, v. 12037, p. 393-398, 2020.

BARBOSA, P. A. *Prosódia*. São Paulo: Parábola Editorial, 2019.

BELIN, P. On Voice Averaging and Attractiveness. In: WEISS, B.; TROUVAIN, J.; BARKAT-DEFRADAS, M.; OHALA, J. J. (Eds.). *Voice Attractiveness Studies on Sexy, Likable, and Charismatic Speakers*, 2020. p.139-149.

BLASI, D.; WICHMANN, S.; HAMMARSTRÖM, H.; STADLER, P.; CHRISTIANSEN, M. Sound-meaning association biases evidenced across thousands of languages. *PNAS*, September 27, 113(39), 2016, 10818-10823. DOI: <https://doi.org/10.1073/pnas.1605782113>

CELATA, C.; CALAMAI, S. *Advances in Sociophonetics*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2014.

CHUENWATTANAPRANITHI S.; XU Y.; THIPAKORN B.; MANEEWONGVATANA, S. Encoding Emotions in Speech with the Size Code. *Phonetica*, 65:210-230, 2008. DOI: 10.1159/000192793.

DI PAOLO, M.; YAEGER-DROR, M. (Eds.). *Sociophonetics: a student's guide*. New York: Routledge, 2011.

ERIKSSON, A. The disguised voice: imitating accents or speech styles and impersonating individuals. In: C. Llamas and D. Watt (Eds.). *Language and Identities*. Edinburgh, Edinburgh University Press: 86-96, 2010.

ERIKSSON, A. Tutorial on Forensic Speech Science. Paper presented at tutorial session on Forensic Speech Science, *Interspeech* (9th European Conference on Speech Communication and Technology), Lisbon, Portugal, 2005.

- FÓNAGY, I. *Languages within Language: an Evolutive Approach*. Amsterdam: John Benjamins, 2001.
- FÓNAGY, I. *La vive voix : essais de psycho-phonétique*. Paris: Payot, 1983. ISBN 9782228131308.
- FOULKES, P.; SCOBBIIE, J. M.; WATTS, D. Sociophonetics. In: W. HARDCASTLE, W.; LAVER, J.; GIBBON, F. (Eds.). *Handbook of Phonetic Sciences* (2nd Edition), Oxford: Blackwell, 2010. p. 703-754.
- GUSSENHOVEN, C. Foundations of intonation meaning: anatomical and physiological factors. In: OLLER, D. K.; DALE, R.; GRIEBEL, U. (Eds.). *Topics in Cognitive Science*, 8(2), 425–434, 2016.
- GUSSENHOVEN, C. *The Phonology of Tone and Intonation*. Cambridge University Press, 2004.
- HARRINGTON, J.; PALETHORPE, S.; WATSON, C. Does the Queen speak the Queen's English? *Nature*, 408, 21/28 dez., 2000.
- KÖHLER, W. *Gestalt Psychology*. New York: Liveright, 1929.
- KÜNZEL, H. Effects of voice disguise on speaking fundamental frequency. *Forensic Linguistics*, 7, p.149-179, 2000.
- LABOV, W. *Sociolinguistic patterns*. Philadelphia: Pennsylvania University Press, 1972.
- LABOV, W.; YAEGER, M.; STEINER, R. *A quantitative study of sound change in progress*. Philadelphia: US Regional Survey, 1972.
- LAVER, J. *The phonetic description of voice quality*. Cambridge University Press, 1980.
- MADUREIRA, S. Portuguese rhotics in poem reciting: perceptual, acoustic, and meaning-related issues. In: GIBSON, M. and GIL, J. (orgs.) *Romance Phonetics and Phonology*. Oxford: Oxford University Press, 2019. p. 191-215.
- MADUREIRA, S.; SOUZA FONTES, M. A.; CAMARGO, Z. Sound symbolism, speech expressivity and crossmodality. *Significances (Signifying)*, 2019. DOI 10.18145/SIGNIFICANCES.V3I1.234. Disponível em: <https://revues.bu.uca.fr/index.php/Significances/article/view/234>. Acesso em: 27/10/2020.
- MILESKI, I. *Varição no português de contato com o polonês no Rio Grande do Sul: vogais médias tônicas e pretônicas*. 2017. Tese (Doutorado em Letras) – Programa de Pós-Graduação em Letras, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2017.
- MOOS, A.; SIMMONS, D. R.; SIMNER, J.; SMITH, R. C. Color and texture associations in voice-induced synaesthesia. *Frontiers in Psychology*, 4, 568. DOI:10.3389/fpsyg.2013.00568, 2013.
- NEWMAN, S. Further experiments on phonetic symbolism. *American Journal of Psychology*, 45, p. 53-75, 1933.
- NIEBUHR, O.; THUMM, J.; MICHALSKY, J. (2018). Shapes and timing in charismatic speech - Evidence from sounds and melodies, *Proc. 9th International Conference on Speech Prosody 2018*, 582-586, DOI: 10.21437/SpeechProsody, p. 2018-118, 2018.
- NOBILE, L. Introduction - Sound symbolism in the age of digital orality. A perspective on language beyond «nature» and «culture». *Significances (Signifying)*, 2019. DOI 10.18145/SIGNIFICANCES.V3I1.248. Disponível em: <https://revues.bu.uca.fr/index.php/Significances/article/view/248>. Acesso em: 27/10/2020.
- OHALA, J. J. Sound symbolism, *Proceedings of the 4th Seoul International Conference on Linguistics [SICOL]*, 11-15 Aug 1997, p.98-103, 1997.
- PASSETTI, R. R. Desviando da própria fala: implicações para verificação de locutor em situações de disfarces vocais. Fundação de Amparo à Pesquisa do Estado de São Paulo. *Relatório Científico de Bolsa de Iniciação Científica* (processo n° 2011/23054-3). Campinas, 2012.
- PASSETTI, R. R. *O efeito do telefone celular no sinal da fala: uma análise fonético-acústica com implicações para a verificação de locutor em português brasileiro*. Dissertação de Mestrado. Universidade Estadual de Campinas. Instituto de Estudos da Linguagem, 2015.
- PRESTON, D. R.; NIEDZIELSKI, N. (Eds.). *A reader in sociophonetics*. New York: De Gruyter Mouton, 2010.

ROSE, P. J. Effect of telephone transmission. In: SELBY, H.; FRECKELTON, I. (Eds.) *Expert Evidence*. Sydney: Thomson Lawbook Company, 2003.

ROSENBERG, A.; HIRSCHBERG, J. Prosodic Aspects of the Attractive Voice. In: WEISS, B.; TROUVAIN, J.; BARKAT-DEFRADAS, M.; OHALA, J. J. (eds) *Voice Attractiveness Studies on Sexy, Likable, and Charismatic Speakers*. New York: Springer, 2020. p.17-40.

SAPIR, E. A study in phonetic symbolism. *Journal of Experimental Psychology*, 12, p. 225-239, 1929.

SILVA, W.; BARBOSA, P. A. Perception of emotional prosody: investigating the relation between the discrete and dimensional approaches to emoticons. *Revista de Estudos da Linguagem*, [S.l.], v. 25, n.3, p. 1075-1103, June 2017, 2017, ISSN 2237-2083. Disponível em: <http://www.periodicos.letras.ufmg.br/index.php/relin/article/view/10972>. Acesso em: 24 out. 2020. DOI: <http://dx.doi.org/10.17851/2237-2083.25.3.1075-1103>

T HART, J. Differential sensitivity to pitch distance, particularly in speech. *The Journal of the Acoustical Society of America*, 69(3), p. 811-821, 1981.

THOMAS, E. *Sociophonetics: an introduction*. Basingstroke, Hampshire: Palgrave Macmillian, 2011.

TSUR, R.; GAFNI, C. *Phonetic Symbolism: double-edgedness and aspect-switching*. Literary Universals Project, 2019. Disponível em: <https://literary-universals.uconn.edu/2019/07/20/phonetic-symbolism-double-edgedness-and-aspect-switching/>. Acesso em: 27/10/2020.

TSUR, R. *What Makes Sound Patterns Expressive: the poetic mode of speech-perception*. Durham, NC: Duke University Press, 1992.

WOODWORTH, N. L. Sound symbolism in proximal and distal forms. *Linguistics*, 29, p. 273-299, 1991.