

# **ELEMENTOS ESSENCIAIS PARA UM ENTENDIMENTO DOS LIMITES E VANTAGENS DA ESTATÍSTICA INFERENCIAL NA PESQUISA FONÉTICA**

**Plínio Almeida Barbosa<sup>1</sup>**

pabarbosa.unicampbr@gmail.com

**RESUMO:** Este trabalho tem a finalidade de apresentar e discutir os elementos essenciais para um entendimento dos limites e vantagens da estatística inferencial na pesquisa fonética, especialmente no que tange à relação entre amostra e população estatísticas. Para tanto discute a relação entre amostra e população com dois exemplos hipotéticos e um exemplo real de pesquisa em prosódia experimental. Questões relacionadas a testes de hipóteses, rejeição e aceitação de hipóteses, variáveis independentes, dependentes e a controlar, nível de significância, representatividade de informantes dentro de um perfil sociolinguístico são levantadas e discutidas concretamente a partir dos exemplos apresentados.

**PALAVRAS-CHAVE:** Estatística Inferencial; População e amostra; Fonética Experimental.

## **INTRODUÇÃO**

A observação de fatos de interesse científico não é ingênua, sendo assim não pode ser feita por um leigo. O neófito precisa receber instrução formal para ser capaz de ler adequadamente os traçados dos valores de uma variável relevante para um determinado arcabouço teórico ou ser capaz de montar um corpus cujos dados sejam susceptíveis, uma vez analisados, de responder a indagações levantadas por uma teoria, resposta que deve ser dada a partir da obediência a um conjunto pré-estabelecido, e de certa forma “codificado”, de procedimentos ditados pela disciplina científica. Para uma leitura sobre o papel da observação na Ciência ver Fleck (1992), Beveridge (1957: 102-105) e Bunge (1998: 181-189).

Como ilustração, tomemos o caso de um traçado de valores de frequência fundamental ao longo de um enunciado, frequência que é o correlato acústico da taxa de vibração das pregas vocais. Os picos locais dessa curva, que assinalam valores locais de maior taxa de vibração das pregas vocais, têm diferentes alturas, mas somente alguns desses picos são

---

<sup>1</sup> Grupo de Estudos de Prosódia da Fala, Departamento de Linguística, Instituto de Estudos da Linguagem, Universidade Estadual de Campinas – UNICAMP.

relevantes do ponto de vista perceptivo e um número ainda menor é relevante do ponto de vista estritamente linguístico. Essa relevância é aprendida a partir de instrução formal. Sendo assim, a questão “o que é um pico de frequência fundamental linguisticamente significativo?” não pode ser respondida por um leigo, mas por alguém que passou por instrução formal nas áreas de Fonética Linguística e Perceptiva. A observação é a primeira etapa de uma pesquisa científica, tanto experimental quanto não experimental, portanto, é a primeira etapa da pesquisa em Fonética.

Como em toda empresa científica, a descrição é a segunda etapa da pesquisa, aplicando-se também à área de Fonética. A descrição se serve dos recursos formais da estatística descritiva ou de métodos correlacionais que permitem ao pesquisador obter, no caso da estatística descritiva, medidas de centralidade como média e mediana, medidas de variação como desvio-padrão, amplitude, variância e semi-amplitude entre quartis e medidas de assimetria, inter alia, respectivamente chamados de descritores estatísticos de primeira, segunda e terceira ordens. No caso de métodos correlacionais, o pesquisador obterá medidas de correlação entre duas ou mais variáveis de interesse.

Entre os descritores estatísticos, a média assinala uma tendência central enquanto o desvio-padrão assinala um afastamento médio dos valores com relação à média. Esses descritores quantificam aspectos graficamente observáveis através do histograma da variável em análise. Essas duas etapas da pesquisa científica estabelecem os fundamentos para a terceira etapa, a da experimentação. Convém ressaltar que, por ser o alicerce sobre o qual se construirá a terceira etapa da investigação experimental em Fonética, a pesquisa descritiva, constituída por observação e descrição de fatos fônicos, não pode ser considerada menos importante ou não científica. Se de um lado é verdade que a experimentação alicerçada numa análise estatística inferencial avalia probabilisticamente a reprodutibilidade dos dados, critério fundamental de cientificidade, também é verdade que esse “poder” da análise inferencial pode ser mal interpretado quando não se compreende a necessária relação entre amostra e população estatísticas de dados, como veremos adiante.

Quanto à necessária relação entre experimentação e estatística inferencial, “one of the first things which the beginner must grasp is that statistics need to be taken into account when the experiment is being planned, or else the results may not be worth treating statistically.” (BEVERIDGE, 1957: 19). Como em toda ciência experimental, a experimentação em Fonética parte de um arcabouço teórico da área de pesquisa em produção e percepção da fala que primeiramente guia a observação de fatos fônicos. A teoria aliada aos fatos observados produz um conjunto de hipóteses explícitas e verificáveis que vão ser colocadas à prova a

partir da descrição que orienta a montagem de um desenho experimental. As hipóteses precisam ser testáveis e fazer previsões de dados novos. O modelo teórico é assim testado a partir de medidas de variáveis ligadas diretamente às hipóteses levantadas, medidas que são avaliadas por métodos estatísticos inferenciais que estabelecem uma ponte entre amostra e população estatísticas. Esse último passo permite, ou a confirmação, normalmente acrescida de uma quantificação das variáveis em escrutínio, ou a refutação, ou ainda o refinamento das hipóteses levantadas.

Visto que em outro lugar (BARBOSA, 2012) nos detivemos amiúde sobre as propriedades de um estudo experimental bem como sobre a seleção de variáveis independentes e dependentes em prosódia experimental, procedimentos que também se aplicam à área mais ampla de Fonética Experimental, passaremos a discutir um aspecto mais básico, aquele da relação entre amostra e população estatísticas. Essa relação é discutida na próxima seção a partir da avaliação do caráter aleatório da seleção de informantes.

## **1. RELAÇÃO ENTRE AMOSTRA E POPULAÇÃO ESTATÍSTICA PARA ANÁLISE INFERENCIAL**

A análise estatística inferencial por testes de hipóteses, independentemente da seleção do teste estatístico apropriado, sempre envolve uma escolha probabilística entre duas hipóteses de trabalho: a hipótese nula ( $H_0$ ) e a hipótese alternativa ( $H_a$ ). A primeira é uma igualdade que impõe uma relação entre valores ou valor de um descritor estatístico do tipo: a média  $\mu$  da população é 200 ( $\mu = 200$ ) ou do tipo: as variâncias dos dois grupos ou níveis em análise são iguais na população ( $\sigma^2_1 = \sigma^2_2$ ). A hipótese alternativa é a negação da hipótese nula na forma de uma desigualdade (o valor é diferente, maior ou menor do que o da população (e.g.,  $\mu > 200$  ou  $\sigma^2_1 \neq \sigma^2_2$ ).

A decisão pela aceitação de uma dessas hipóteses é feita a partir da comparação (1) do resultado do cálculo de probabilidade de ocorrência do valor ou diferença de valores de um descritor estatístico da variável dependente assumindo que os dados pertencem à distribuição que representa a hipótese nula com (2) um limiar escolhido a priori chamado de nível de significância ( $\alpha$ ). Se a probabilidade de ocorrência envolvendo descritores estatísticos fundamentados na hipótese nula for menor que o nível de significância, essa eventualidade é considerada rara e rejeita-se a hipótese nula; por outro lado, caso a probabilidade de ocorrência envolvendo descritores estatísticos fundamentados na hipótese nula for maior ou igual ao nível de significância, se aceita a hipótese nula. É importante ressaltar que a aceitação não implica a veracidade da hipótese nula. Assim, a decisão por uma hipótese, em detrimento

da outra, é probabilística, a hipótese escolhida é sempre uma hipótese, não será nunca uma tese. A razão para isso envolve a relação entre amostra e população estatísticas e o procedimento de aleatorização associado.

Quando falamos em população estatística referimo-nos a um conjunto finito ou, as mais das vezes no caso da Fonética, um conjunto infinito de variáveis aleatórias, isto é, de valores de dados obtidos por um procedimento experimental. Esses valores podem ser medidos, como já apresentamos em outro lugar (BARBOSA, 2012), ao nível categórico, ordinal ou intervalar, em grau crescente de número de operações realizáveis com os valores medidos bem como em grau crescente de conteúdo informacional.

Como exemplo de variável categórica tomemos o tom numa descrição fonológico-entoacional de um enunciado, variável que pode assumir os valores discretos H (alto) ou baixo (L), sendo possível dizer apenas se dois tons são iguais ou diferentes. No caso de uma variável intervalar, como a duração silábica, essa pode assumir um valor real positivo qualquer. O pesquisador pode verificar por esse nível de mensuração se duas sílabas têm a mesma duração, se têm duração diferente, se uma é maior do que a outra, bem como pode ainda somar suas durações.

Os valores medidos num determinado nível de mensuração para variáveis dependentes especificando valores oriundos de enunciados podem pertencer a um único informante ou a um conjunto finito de informantes. Nos dois casos a amostra de dados é sempre finita e a população de dados que contém a amostra é infinita ou quase infinita, pois enunciados produzidos pelo mesmo informante podem ser obtidos enquanto ele viver e assim, os valores das variáveis dependentes podem ser medidos, constituindo um conjunto tão amplo que não poderia ser medido por um único pesquisador no seu tempo de vida. Se a amostra é formada por dados oriundos de enunciados de vários informantes representativos de uma comunidade linguística, por exemplo, o conjunto é potencialmente infinito, enquanto aquela comunidade linguística se definir como tal. A população é assim o supraconjunto de dados dos quais se obtém a amostra, o conjunto que subjaz a amostra colhida para análise, seja de valores de parâmetros fônicos oriundos de enunciados de um informante ou de mais de um informante.

Para que um subconjunto da população seja uma amostra estatística dessa mesma população não basta ser um mero subconjunto: é preciso que seja constituído aleatoriamente. E essa é uma condição *sine qua non* para que a análise estatística inferencial possa inferir uma propriedade da população a partir da amostra estatística. Esclareçamos com um exemplo.

Um mini-dicionário é, normalmente, um subconjunto do dicionário mais extenso da mesma equipe de lexicógrafos. É um subconjunto, mas não é uma amostra estatística, pois

não passa pelo critério de aleatoriedade. A razão disso é a de que as palavras que constituem os verbetes do mini-dicionário não são escolhidas ao acaso a partir do dicionário que o contém, mas a partir de determinados critérios envolvendo frequência de uso das palavras, faixa etária e escolaridade do potencial usuário, além de outras possibilidades. Sendo assim, não se pode, com base em análise descritiva realizada num mini-dicionário, inferir ou estimar pela via da estatística inferencial, alguma medida do dicionário maior. Por exemplo, o cálculo da extensão média em sílabas de palavras num mini-dicionário não permite dizer qual a extensão média das palavras do dicionário que o contém. Porque o mini-dicionário não é uma amostra estatística do dicionário maior. Caso, hipoteticamente, um pesquisador montasse um dicionário com mesmo número de verbetes do mini-dicionário regular a partir de uma seleção aleatória das palavras do dicionário que contém o mini-dicionário regular, o cálculo de descritores estatísticos do mini-dicionário aleatoriamente constituído permite estimar apropriadamente os descritores estatísticos do dicionário maior. O que garante a legitimação dessa estimativa é o procedimento de aleatorização, condição para a inferência estatística.

Por analogia, o exemplo dos dicionários possibilita entender o que está em jogo na relação entre as amostras de dados fônicos e a população que as subjaz. Mas então se apresenta um problema prático: se a amostra precisa ser aleatoriamente constituída, é também preciso assegurar que possa ser obtida em termos economicamente viáveis. Para entender isso, ilustremos com um caso hipotético de enunciados produzidos por um informante a partir de sentenças isoladas selecionadas para o estudo dos correlatos acústicos do contraste de vozeamento de oclusivas do português brasileiro (doravante PB).

Suponhamos que um pesquisador queira caracterizar acusticamente o contraste fonológico de vozeamento em oclusivas no PB, isto é, avaliar quais são as diferenças e características acústicas das realizações de /p t k/ vs /b d g/. Antes de começar a medir os parâmetros acústicos desses sons, é preciso conceber o desenho experimental a partir: (1) das teorias fônicas que atestam o contraste, (2) do elencamento de hipóteses a partir dessas teorias e (3) da seleção de variáveis experimentais que guiam a montagem de um corpus a partir do qual as medidas acústicas serão feitas. Depois da mensuração, a análise estatística inferencial intimamente associada à questão teórica permitirá obter resultados que vão cotejar as hipóteses apontando os elementos passíveis de distinção fônica do traço de vozeamento. Antes de examinarmos essa questão consideremos a escolha da amostra estatística e sua relação com a população que a subjaz, considerando um único informante.

Se o pesquisador deseja que os dados do informante que vai selecionar reflitam ou revelem um padrão fônico do dialeto que estuda, ele precisa escolher um informante

representativo desse dialeto e de vários outros aspectos da população cujas características acústicas deseja inferir. Um informante representativo do dialeto é aquele que não apenas se reconhece identitariamente como pertencente a uma determinada comunidade linguística, como também aquele que leigos e especialistas identificam como pertencente àquela comunidade em estudo. Além disso, supondo no exemplo apresentado que sejam características não patológicas que estejam em estudo, por constituírem a situação da maioria dos falantes, é preciso que o informante representativo não tenha problemas fonoarticulatórios. Observe que a seleção ou busca de um indivíduo com essas características, se de um lado não configura o procedimento como puramente aleatório, por outro lado torna o procedimento de seleção economicamente viável (cf. Woods et al., 1986). Ao relatar os resultados dessa investigação, o pesquisador deverá informar as características do informante quanto à escolaridade, faixa etária, gênero, classe social, dialeto, inter alia, para que se tenha em mente que a população de dados acústicos que subjazem a amostra diz respeito a dados desse indivíduo *ad infinitum*, dados que fazem parte da população de dados de indivíduos com o mesmo perfil sociolinguístico. Assim, se as variáveis em estudo concernem variáveis linguísticas, como no caso do vozeamento ilustrado aqui, é muito provável que mais de um indivíduo da mesma comunidade linguística e com o mesmo perfil sociolinguístico apresente um comportamento fônico similar, isto é, no caso em tela, realize contraste de vozeamento se servindo dos mesmos parâmetros acústicos encontrados para o informante representativo, salvaguardadas as diferenças quantitativas inter-individuais para os valores dos parâmetros contrastados.

O pesquisador não pode esquecer que variáveis estranhas ao seu desenho experimental, denominadas de variáveis a controlar, podem interferir inadvertidamente no comportamento do informante. Para citar duas, consideremos primeiramente o intervalo de tempo entre a ingestão de alimento e a gravação, porque uma ingestão próxima ao momento da gravação tem consequências para a atividade e estado atencional do informante. Outra variável a controlar é o cansaço. Por conta do cansaço causado ao informante pela atividade e enunciação, a gravação não pode se estender de tal forma que o canse e altere sua voz. Há variáveis a controlar de natureza linguística, referindo-se aos fatores dessa natureza que interferem na medida e precisam ser mantidos constantes ao contrastar níveis distintos das variáveis independentes, como consideraremos adiante. Assegurados esses cuidados básicos, consideremos o tipo de enunciado a ser obtido para mostrarmos outro aspecto importante da relação entre amostra estatística e população.

No estudo considerado aqui, uma escolha apropriada para constituir um corpus que permita o entendimento dos aspectos essenciais do contraste de vozeamento se dá pela gravação de enunciados isolados em que a palavra-chave esteja inserida numa frase-veículo como “Digo *palavra-chave* baixinho.” A estrutura /CV'CVCV/ para a palavra-chave possibilita trabalhar com a estrutura silábica mais frequente do PB, bem como possibilitaria avaliar as oclusivas em três condições de tonicidade: pré-tônica, tônica e pós-tônica. Observe que a pronúncia do enunciado com a palavra-chave na porção central tem a vantagem de colocar a palavra numa posição mais frequente nos enunciados produzidos, muito embora numa posição potencialmente de proeminência, isto é, em que a palavra-chave será produzida com maior destaque em relação às demais. Assim, os resultados obtidos dirão respeito às características acústicas do contraste de vozeamento em palavras proeminentes, uma vez que a população estatística que subjaz os dados dizem primariamente respeito a dados produzidos *ad infinitum* nas mesmas condições. A inferência, da amostra para a população de dados, não pode revelar assim as características acústicas do contraste de vozeamento nas palavras não proeminentes, a não ser que a oposição entre palavras nas duas categorias faça parte das variáveis independentes (variáveis manipuladas pelo experimentador a partir da hipótese de trabalho) com a consequente gravação de palavras-chave com essa característica acústica no mesmo contexto fônico. Isso pode ser obtido pela gravação adicional de enunciados a partir da sentença “Digo *palavra-chave* BAIXINHO.”, que permite ao informante realizar foco estreito em “baixinho”, forçando a pronúncia não saliente da palavra-chave e portanto resultando num efeito de não proeminência no campo da percepção. Essa decisão aumenta a generalidade dos resultados, mas dobra o número de dados a serem medidos e, portanto, aumenta o tempo para a realização da análise, tempo cujas consequências práticas deverão ser consideradas pelo pesquisador.

Para que se tenha uma ideia do número de dados que um desenho experimental simples como o apresentado aqui produz, suponhamos que, a partir dos enunciados com a palavra-chave proeminente, o pesquisador desejasse avaliar a influência das vogais /i a u/ seguintes às seis oclusivas em tônicas e pré-tônicas mantendo a sílaba pós-tônica final fixa em /ta/, por exemplo. Assim, teríamos um número de palavras-chave obtido pela combinação de 6 consoantes pré-tônicas x 3 vogais pré-tônicas x 6 consoantes tônicas x 3 vogais tônicas = 324 enunciados. Para garantir a consistência dos dados e tornar a análise estatística viável, o pesquisador deve gravar certo número de repetições dos enunciados, por exemplo, de 5 a 10 repetições de cada enunciado, tendo então que gravar de 1620 a 3240 enunciados no total.

Não é possível determinar a priori o número de repetições necessário para permitir uma decisão estatística, porque depende da variabilidade dos parâmetros acústicos considerados no informante escolhido. Por conta disso, um estudo piloto com poucas repetições é desejável para permitir estimar o número de repetições ainda necessário para uma decisão estatística.

Apesar do número elevado, a amostra ou corpus que constitui o conjunto de enunciados em exame permite inferir aspectos acústicos de contraste de vozeamento em enunciados semelhantes, isto é, de três palavras paroxítonas, na sequência, dissílabo, trissílabo, dissílabo e com trissílabo central proeminente. O grau de generalidade obtido depende do que se deseja inferir. Aspectos quantitativos das medidas são específicos a um informante. Por conta disso, não se pode inferir, a partir de um informante, faixas de valores dos parâmetros acústicos em estudo para contextos linguísticos distintos do considerado na ilustração. A análise estatística inferencial permitirá obter, por outro lado, para os dados do informante estudado, intervalos de variação dos parâmetros, denominados de intervalos de confiança que são estimativas de valores para enunciados similares. Assim, um número finito de enunciados permite inferir faixas de valores de enunciados similares produzidos *ad infinitum*.

Se o que se deseja inferir são aspectos qualitativos, a análise estatística da ilustração em tela permite determinar a extensão do efeito (*effect size* em inglês), isto é, a porcentagem da influência de um fator sobre a variável dependente, o que permite consequentemente hierarquizar os fatores para explicar a variação de determinado parâmetro acústico quando do contraste e, reciprocamente, por Análise Discriminante Linear (LDA em inglês), determinar que parâmetro acústico discrimina, separa, melhor as oclusivas em contraste de vozeamento. O padrão acústico contrastado no contexto das vogais /i a u/ permitirá formular novas hipóteses de comportamento fônico das oclusivas no contexto das demais vogais do PB, vogais que se encontram em situação intermediária em termos de altura e ponto de articulação com relação às vogais /i a u/. As novas hipóteses exigirão a gravação de um novo corpus e realização assim de um novo experimento.

Resumindo, se o informante é representativo da comunidade linguística em estudo, a análise das características fônicas de sua fala permite hipotetizar aspectos fônicos da fala de informantes da mesma comunidade e do mesmo perfil sociolinguístico<sup>2</sup>.

---

<sup>2</sup> Supomos que não haja patologia de fala envolvida, a menos que o comportamento fônico de determinada patologia seja o tema da pesquisa. Nesses casos a inferência estatística do indivíduo para a comunidade de informantes com aquela patologia é muito delicada, uma vez que os aspectos anátomo-fisiológicos de dois indivíduos com a mesma patologia normalmente diferem em mais de um aspecto, não constituindo amostras da mesma população estatística.

Assim, dito de outra forma, se uma análise acústica, ao nível segmental ou prosódico, revelar um determinado padrão a partir de poucos ou de apenas um único indivíduo de uma comunidade linguística, os aspectos qualitativos das medidas apontados acima não podem ser descartados como aspectos que não estariam presentes na comunidade linguística como um todo. Isso é possível porque a amostra em questão é uma amostra aleatória da população de dados acústicos produzidos pelos informantes da comunidade linguística em questão. Por isso a preocupação em escolher um informante que seja representativo da mesma. Por outro lado, somente a continuidade da pesquisa e a inclusão de novos informantes podem refinar ou reorientar a inferência inicial. Mas essa inferência inicial não pode ser descartada como irrelevante por dizer respeito a um único ou a poucos indivíduos.

## **2. VARIÁVEIS INDEPENDENTES E DEPENDENTES EM PESQUISA FONÉTICA**

As variáveis independentes são aquelas manipuladas pelo experimentador porque as hipóteses que levantou apontaram que são variáveis importantes para a investigação em questão. Assim, no estudo do vozeamento, o experimentador levantou a hipótese de que, além da especificação fonológica de vozeamento, vogais altas podem alterar o grau de vozeamento da oclusiva, bem como a tonicidade da sílaba, por isso também selecionou essas duas variáveis independentes ou fatores. Fator é uma palavra técnica que designa a variável categórica que, ao ser variada entre seus níveis, pode afetar a variável dependente (e.g. a mudança do nível “surdo” para “sonoro” no fator VOZEAMENTO provoca mudanças na duração total da oclusiva, no vozeamento fonético, *inter alia*). Assim, na investigação ilustrada acima, o pesquisador escolheu os seguintes fatores: VOZEAMENTO (com os níveis “surda” e “sonora”), TONICIDADE (com os níveis “pré-tônica” e “tônica”) e VOGALSEGUINTE (com três níveis correspondentes às vogais /i a u/).

Quanto às variáveis dependentes que podem ser medidas para cada realização de oclusiva, podemos citar a duração total, o VOT (intervalo entre a soltura da plosão e o vozeamento) e o espectro do ruído de plosão da oclusiva. Além disso, por conta de seu papel na percepção do traço de vozeamento, o valor do início de F1 na vogal seguinte é outra medida relevante. Teríamos assim quatro variáveis dependentes relacionadas ao traço de vozeamento.

Para cada variável dependente o experimentador se interessa em avaliar diferenças no fator VOZEAMENTO para cada par de oclusiva homorgânica independentemente e em interação com os outros dois fatores citados acima. Nesse caso, admitindo que os valores das

variáveis dependentes se comportem de maneira tal que permitam o uso de uma análise paramétrica<sup>3</sup>, o teste estatístico indicado seria a Análise de Variância (ANOVA) fatorial que inclua a interação entre os fatores em seu modelo (consulte o livro de Crawley, 2007 para aprender a usar esse teste estatístico bem como as condições para seu uso num programa de software estatístico gratuito, o R).

Aplicada para prever cada uma das quatro variáveis dependentes acima, uma ANOVA para cada um dos três pares de oclusivas homorgânicas, avalia se há diferenças significativas nos valores médios entre os níveis de cada fator, bem como se cada fator interage ou não com os demais. Adiante veremos o que se entende por ‘diferença significativa’, que ocorre quando da rejeição da hipótese nula já apontada na seção anterior.

Deve-se ressaltar que uma análise fatorial exige que haja valores das variáveis medidas para todas as combinações entre os níveis de todos os fatores, o que, na prática e em nosso exemplo, impossibilita o uso de palavras reais em todos os contrastes. Se o interesse for apenas o emprego geral dos parâmetros acústicos para assinalar vozeamento sem considerar, por exemplo, o efeito da frequência de uso de uma palavra sobre o contraste de vozeamento, então não há problema em se trabalhar com logatomas, vocábulos sem sentido, desde que se instrua o informante para não se importar com isso e pronunciar como se fosse uma palavra da língua. Se o experimentador cuida para que as estruturas silábicas da palavra sejam preferencialmente CV, como em nosso exemplo, essa tarefa é realizada com facilidade pelo informante. O imprescindível nesses casos é respeitar a condição *ceteris paribus*, isto é, se gravou a palavra “batata”, por exemplo, para testar o [b] pré-tônico diante de [a], deve gravar o logatoma “bitata” para testar o [b] pré-tônico diante de [i], e assim por diante.

Cada ANOVA de três fatores permitirá avaliar, inferencialmente, a significância da diferença entre médias do respectivo parâmetro acústico considerando os contrastes entre os níveis de cada fator. Assim, o fator VOZEAMENTO em cada ANOVA de três fatores realizada permitirá dizer, para cada ponto de articulação da oclusiva, se há diferença média significativa para cada parâmetro acústico entre as oclusivas sonora e surda, independentemente e/ou em interação com os níveis dos dois demais fatores, TONICIDADE e VOGALSEGUINTE. Diferenças que consideram interação entre fatores funcionam da seguinte maneira: um contraste no fator VOZEAMENTO num parâmetro acústico qualquer

---

<sup>3</sup> Na análise paramétrica são os próprios valores dos dados que são usados para se calcular os descritores. Na análise não paramétrica é uma variável ordinal obtida a partir da variável original que é usada no lugar, perdendo-se informação. No entanto, se os pressupostos da análise paramétrica, entre eles a normalidade da distribuição de erro entre valor medido e obtido a partir do modelo de ANOVA não forem obedecidos, uma análise paramétrica deve ser usada. A análise não paramétrica correspondente a uma ANOVA de um fator é a Kruskal-Wallis.

pode se dar na tônica, mas não na pré-tônica, considerando o fator TONICIDADE. Se isso ocorrer se diz que há interação entre o fator VOZEAMENTO e o fator TONICIDADE. Se, por outro lado, um contraste num determinado parâmetro acústico se dá no fator VOZEAMENTO independentemente dos níveis dos outros dois fatores, diz-se que não há interação entre o fator VOZEAMENTO e os outros dois fatores.

Uma diferença média entre parâmetros acústicos é significativa com relação à raridade (baixa probabilidade com relação ao nível de significância) da possibilidade de igualdade entre os valores médios desses parâmetros, considerando a inferência populacional. A hipótese nula para o teste de ANOVA considera que a diferença média numérica observada entre os níveis dos dados da amostra é devida ao caráter aleatório da amostragem de uma população homogênea de médias de estímulos. Por conta disso, os valores médios não diferem na população que contém a amostra e portanto avalia-se a probabilidade da diferença observada considerando a premissa de inexistência de diferença. Ora, se a probabilidade da diferença observada for menor que o limiar de raridade escolhido previamente pelo pesquisador (o nível de significância  $\alpha$ ), rejeita-se a hipótese nula e considera-se que a diferença é significativa para aquele nível de significância.

Como apontamos anteriormente, a decisão por rejeitar ou aceitar a hipótese nula é probabilística. Sendo assim, a probabilidade de cometer um erro com a decisão acima (isto é, rejeitar uma hipótese nula verdadeira, conhecido como erro do tipo I em Estatística Inferencial) é dada pela probabilidade “p” (*p-value* em inglês) de ocorrência da diferença entre as médias, considerando a distribuição de diferença nula. O valor dessa probabilidade “p” nunca é nulo, embora possa ser desprezível (e.g.  $p = 10^{-7}$  ou 0,00001 %). Assim, há sempre uma margem de erro que, de certa forma, nos coloca no caminho certo da experimentação, pois fatores não considerados podem afetar um comportamento fônico e é preciso estar alerta para os mesmos.

### **3. UM EXEMPLO REAL ACOMPANHADO DA INTERPRETAÇÃO ESTATÍSTICO-INFERENCIAL DOS DADOS**

Ilustremos os aspectos delineados acima com um exemplo retirado de uma pesquisa na área de prosódia experimental. Num estudo em andamento, Constantini e Barbosa (2013) avaliaram oito parâmetros prosódico-acústicos a partir da análise de dados de informantes de sete regiões dialetais brasileiras não homogeneamente delimitadas por conta do número desigual de informantes representativos de suas regiões, como será assinalado adiante. As

regiões escolhidas foram as correspondentes aos Estados brasileiros de São Paulo, Rio de Janeiro, Minas Gerais, Paraná e Amazonas além do Distrito Federal e de informantes de cinco Estados da região Nordeste (Ceará, Rio Grande do Norte, Pernambuco, Paraíba e Sergipe) considerados como um grupo a ser separado a posteriori, se necessário. O número de informantes do Amazonas é de dois, do Distrito Federal é de cinco, do Paraná de quatro, e das demais regiões de seis cada uma. O estudo considera, de forma inédita no País, possíveis marcas dialetais prosódico-acústicas, e não segmentais. Como se trata de um estudo acústico de natureza prosódica, a classificação e discriminação entre os dialetos na pesquisa anterior com base em análises de oitiva e, mais raramente, com base em estudos acústico-segmentais, não implica a obtenção das mesmas classes do ponto de vista tipológico, pois é possível que duas regiões previamente classificadas numa mesma região dialetal possam pertencer a outras com base em aspectos prosódicos. Somente a investigação científica da prosódia da fala pode dar elementos de resposta a essa questão, contribuindo para uma melhor caracterização fonética da fala dos dialetos brasileiros.

O corpus usado por Constantini e Barbosa faz parte de um banco de dados coletado pela Polícia Federal de Brasília. Todas as gravações são de oficiais da Polícia Federal, do sexo masculino, com média de idade de 35,4 anos. Para serem incluídos nessa pesquisa os informantes precisavam ter nascido e morado até a idade adulta na mesma região. O grau de escolaridade variou de ensino superior completo até pós-graduação completa. Todos os demais dados sociolinguísticos foram colhidos, incluindo o perfil dialetal dos pais dos informantes. As gravações são de pequenos relatos de vida, entre cinco e cinco minutos e meio cada um. Os trechos gravados foram segmentados em sílabas fonéticas delimitadas por dois *onsets* de vogal consecutivos. A seguir foram etiquetados por uma transcrição fonética larga que distingue segmentos tônicos e átonos. Um programa desenvolvido para o software Praat (BOERSMA e WEENINK, 2013), o ProsodicDescriptorExtractor (BARBOSA, 2011), foi utilizado para extração de oito medidas acústicas ao longo dos relatos, entre as quais destacamos para ilustração a mediana da frequência fundamental (expressa em Hz) e a ênfase espectral (expressa em dB). Cada uma foi calculada para cada sílaba fonética do relato de cada informante, num total de 35 informantes com um número médio de sílabas de 1281 e desvio-padrão 155 para cada informante. A primeira variável é uma medida de centralidade do valor da frequência fundamental de cada informante, enquanto a segunda é uma medida relativa da intensidade que revela a quantidade de energia acima da frequência fundamental, sendo uma medida que não muda com a posição do microfone, por isso seu interesse na

pesquisa fonética. Além disso, Traunmüller e Eriksson (1995) mostraram que essa medida é um correlato acústico do esforço vocal.

Nesse estudo, a variável independente ou fator é a região dialetal, que chamaremos de DIALETO. As variáveis dependentes são os oito parâmetros prosódico-acústicos medidos, entre os quais destacamos a mediana da frequência fundamental e a ênfase espectral por conta da significância dos resultados obtidos, bem como por ensejarem uma discussão sobre a relação entre amostra e população estatísticas.

Como há mais de dois níveis no fator DIALETO, Constantini e Barbosa usaram um teste ANOVA de um fator (One-Way ANOVA), adotando o nível de significância de 5 %. Como o teste apontou uma diferença significativa entre os níveis do fator DIALETO, o teste post-hoc de Tukey Honestly Significant Difference foi usado para apontar as diferenças entre níveis do fator, para o mesmo nível de significância. Os testes post-hoc testam níveis dois a dois e mantêm o nível de significância adotado inalterado, uma vez que múltiplas comparações com um teste t não garantem que a probabilidade final de ocorrência do descritor em estudo seja menor do que  $\alpha$  (para uma explicação detalhada sobre isso ver o livro de Dowdy e Wearden, 1991).

A ênfase espectral e a mediana da frequência fundamental foram os parâmetros que diferenciaram algumas regiões dialetais com valor de probabilidade p menor que 5 %, isto é, a probabilidade de que haja igualdade entre as médias de cada parâmetro foi baixa para o nível de significância adotado e, por conta dessa raridade, rejeitou-se a hipótese nula. Para a ênfase espectral, os informantes do Distrito Federal e da Amazônia apresentaram valores médios significativamente mais elevados do que os dos outros dialetos, sem se diferenciarem entre si quanto ao valor médio desse parâmetro prosódico-acústico.

Pelos resultados, podemos concluir, ao menos inicialmente, que o Distrito Federal e a Amazônia utilizam maior esforço vocal que os falantes das demais regiões dialetais em análise. Os dados dos informantes da Região Nordeste apresentaram os menores valores médios de ênfase espectral e a menor variação desse valor, apontando para certa homogeneidade quanto aos dados de fala obtidos dos informantes dos cinco Estados gravados. Pode-se concluir que, nesse corpus, os informantes da região compreendida pelos cinco Estados analisados fazem menor esforço vocal, diferenciando-os dos da Amazônia e do Distrito Federal. O que se pode concluir, a partir desse resultado, quanto à relação entre a amostra gravada e a população de dados acústicos inferida a partir dessa amostra?

A população que subjaz a produção dos dois informantes do Amazonas e dos cinco informantes do Distrito Federal é a repetição *ad infinitum* do comportamento fonético que os

informantes exibiram durante a gravação. Assim, essa população concerne a situação de entrevista informal, bem como traz as características dialetais que esses informantes representativos trazem consigo, marcadas pelo seu perfil sociolinguístico. Embora seja a fala de indivíduos cultos de formação universitária de classe média e que exercem uma profissão específica, tais indivíduos nasceram e moraram até o momento da gravação em suas regiões e são considerados falantes representativos dessas regiões. São indivíduos escolhidos aleatoriamente no perfil sociolinguístico mencionado no sentido de que foram conduzidos até a gravação a partir de sua história pessoal, em relação à qual o pesquisador não tem nenhum controle. Sendo assim, o resultado significativo quanto ao maior esforço vocal dos dois informantes do Amazonas e dos cinco informantes do Distrito Federal aponta para essa característica prosódica da população de falantes dessas duas unidades da Federação. Isso não significa dizer, de nenhuma forma, que essas duas regiões se comportam prosodicamente da mesma maneira, pois se trata de apenas um parâmetro prosódico-acústico, a ênfase espectral. Outros parâmetros precisam ser analisados para revelar o que os separa e o que os aproxima. É evidente que esse resultado, por conta do que foi dito sobre a população de estímulos acústicos que subjaz as amostras das duas regiões, não pode dizer necessariamente algo sobre informantes com outro perfil sociolinguístico, simplesmente porque não foram gravados, não fazem parte da amostra analisada. No entanto, a hipótese de que informantes com outro perfil sociolinguístico das mesmas regiões respectivas exibam características similares ou similares em parte não pode ser descartada, embora requeira uma investigação ulterior. É assim que caminha a ciência: resultados de análises guiadas por hipóteses verificáveis conduzindo a novas hipóteses e novas investigações.

No mesmo estudo, a mediana da frequência fundamental também permitiu separar o Distrito Federal e a Amazônia das demais regiões estudadas ( $p < 1\%$ ), sem que as duas regiões apresentassem diferenças significativas entre si. Os dados de ambas as regiões apresentam valores médios mais elevados de mediana de frequência fundamental do que os das outras regiões estudadas: 144 Hz para o Distrito Federal e 149 Hz para a Amazônia contra menos de 130 Hz para as demais regiões. Os valores médios das duas regiões apontadas estão em faixa superior aos dos valores masculinos dessa faixa etária (BEHLAU, 2001). Além disso, os valores das duas regiões dialetais são os que apresentam maior variabilidade nos resultados, com desvios-padrão em torno de 30 Hz e valores mínimos próximos de 120 Hz. As mesmas observações sobre a relação entre amostra e população que fizemos acima para os valores de ênfase espectral valem para os valores de mediana de frequência fundamental. Tendo em vista que o maior esforço vocal tende a gerar uma taxa mais elevada de vibração

das pregas vocais, produzindo maiores valores de frequência fundamental, os dois resultados juntos são concordantes, apontando para uma maior tensão na região laríngea nos informantes do Distrito Federal e da Amazônia. Esses resultados apontam para um comportamento prosódico-acústico possivelmente distinto nessas duas regiões em relação às demais avaliadas aqui. Essa diferença não pode, em princípio, ser atribuída à profissão ou fala culta, uma vez que as demais regiões apresentam valores inferiores para esses dois parâmetros acústicos, mesmo sendo da mesma profissão e faixa etária. A menos, e isso não pode ser descartado, que as condições de trabalho nessas duas regiões sejam tais que condicionam uma maior tensão vocal do que nas demais (e.g., maior volume de trabalho acarretando maior uso da voz). É evidente que não se pode dizer que esse resultado seja necessariamente válido para todos os falantes das duas regiões concernidas, uma vez que a amostra não selecionou aleatória ou quase-aleatoriamente indivíduos de todas as faixas etárias, de todas as classes sociais, de todo tipo de escolaridade, de todo perfil sociolinguístico. Mas não se pode descartar que esse resultado aponta para uma característica de indivíduos representativos de seus dialetos e que, portanto, os resultados não podem ser atribuídos a uma característica idioletal, uma vez que há mais de um indivíduo por região aleatória ou quase-aleatoriamente selecionados.

#### **4. CONSIDERAÇÕES FINAIS**

A finalidade deste trabalho foi a de apresentar e discutir os elementos essenciais para um entendimento dos limites e vantagens da estatística inferencial na pesquisa fonética, especialmente no que tange à relação entre amostra e população estatísticas. Para uma leitura detalhada da questão, ver os livros de Woods et al. (1986: 48-58) e Bunschaft e Kellner (2001). Para uma leitura sobre vários aspectos discutidos aqui a respeito do desenho experimental, ver o livro de Kirk (1995).

Procuramos mostrar o caráter absolutamente essencial da estatística inferencial para apontar comportamentos fônicos para além da amostra, caracterizando a reprodutibilidade e o caráter de predição de toda empresa científica. O não uso de técnicas de análise estatística inferencial na Fonética experimental não é mais justificável, pois a comparação descritiva entre duas medidas não permite dizer nada sua consistência. Afinal, a repetição do experimento pode apontar que uma diferença numérica seja devido ao acaso da amostragem. Não se faz ciência sem revelar a consistência de um resultado. Por outro lado, o hábito de mostrar a significância de um resultado estatístico, por maior que seja, sem apresentar médias, diferenças numéricas e intervalos de confiança dos dados também não permite um diálogo

profícuo com a teoria que guiou a montagem do experimento e, portanto, a possibilidade de investigação detalhada do papel de variáveis independentes, dependentes e a controlar sobre o comportamento fônico.

## AGRADECIMENTOS

O trabalho é financiado com verba da Bolsa de Produtividade do CNPQ “Análise e modelamento dinâmicos da prosódia da fala espontânea”, proc. nº 301387/2011-7. Agradeço a Ana Carolina Constantini, André N. Xavier e Cristiane C. Silva pela leitura atenta e comentários, assumindo a responsabilidade pelo eventual mau uso das sugestões e pelos problemas que ainda ficaram.

## REFERÊNCIAS BIBLIOGRÁFICAS

1. BARBOSA, Plínio A. ProsodyDescriptorExtractor. [Programa de software]. Disponível com o autor.
2. BARBOSA, Plínio A. Panorama of Experimental Prosody Research. In: *Proceedings of the VII<sup>th</sup> GSCP International Conference - Speech and Corpora*. Florença: Firenze University Press, p. 33 – 42, 2012.
3. BEHLAU, Mara S. *Voz: o livro do especialista*. Vol. 1. Rio de Janeiro: Revinter, 2001.
4. BEVERIDGE, William I. B. *The Art of Scientific Investigation*. New Jersey, USA: The Blackburn Press, 1957.
5. BOERSMA, Paul; WEENINK, David. *Praat: doing phonetics by computer*. [Programa de software]. Acessível em <<http://www.praat.org/>>, 2013.
6. BUNGE, Mario. *Philosophy of Science. From Explanation to Justification*. New Brunswick, NJ: Transaction Publishers, 1998.
7. BUNSCHAFT, Guenia; KELLNER, Sheilah, R. O. *Estatística sem mistérios*. v. 1. Petrópolis: Vozes, 2001.
8. CONSTANTINI, Ana Carolina; BARBOSA, Plínio A. Características prosódicas de sete dialetos brasileiros. In: OLIVEIRA, Miguel; ALMEIDA, Ayane; ALMEIDA, René (org.) *Caderno de Resumos do 4º Colóquio Brasileiro de Prosódia da Fala*. Maceió, Alagoas, 16-24 de outubro. p. 60-65, 2013.
9. CRAWLEY, Michael J. *The R book*. Chichester: John Wiley and sons, 2007.

10. DOWDY, S.; WEARDEN, S. *Statistics for research*. New York: John Wiley & Sons. 2<sup>a</sup> ed., 1991.
11. FLECK, L Observation scientifique et perception en général. In : BRAUNSTEIN, J.-F. (Org.) *L'Histoire des sciences*. Paris: Librairie Philosophique J. Vrin, p. 245-272, 1992 [1935]).
12. KIRK, R. E. *Experimental Design: procedures for the Behavioral Sciences*. Brooks/Cole., 1995.
13. WOODS, A.; FLETCHER, P.; HUGHES, A. *Statistics in language studies*. Cambridge: Cambridge University Press, 1986.
14. TRAUNMÜLLER, Hermut; ERIKSSON, Anders. The frequency range of the voice fundamental in the speech of male and female adults. Acessível em <http://www.ling.su.se/staff/hartmut/aktupub.htm>, 1995.

**ABSTRACT:** This work aims at presenting and discussing crucial aspects for an understanding of the advantages and limits of inferential statistics in phonetic research, especially the relationship between sampling and population. This relationship is illustrated and discussed with two hypothetical and one real example from experimental prosody research. The main issues raised and discussed at length by these examples are: hypothesis testing, rejection and acceptance of hypotheses, independent, dependent and control variables, significance level, typicality of informants of a given sociolinguistic profile.

**KEYWORDS:** Inferential Statistics; Population and sample; Experimental Phonetics.