

Efeito da língua, estilo de elocução e sexo do falante sobre medidas globais da frequência fundamental¹

Effect of language, speaking style and speaker sex on global measures of fundamental frequency

Pablo Arantes

Maria Érica do Nascimento Linhares

Universidade Federal de São Carlos – São Carlos – São Paulo – Brasil



Resumo: Analisamos o efeito do sexo, do estilo de elocução e da língua (alemão, estoniano, francês, inglês, italiano, português e sueco) sobre estimadores de tendência central e dispersão da frequência fundamental da fala. Os estilos são entrevista, leitura de frases e de palavras. As línguas diferem entre si em termos do valor típico de F0 usado por seus falantes. O estilo leitura de frases apresenta valores maiores do que a entrevista, mas o efeito não é uniforme entre as línguas. Diferenças significativas entre o estilo leitura de palavras e entrevista são mais raros. Todos os estimadores de tendência central estudados sofreram efeitos dos três fatores testados. Os homens apresentaram variabilidade de F0 ligeiramente maior do que as mulheres de forma significativa em três de cinco estimadores. O estilo entrevista tem maior variabilidade em dois. As línguas não diferem em termos de variabilidade de F0.

Palavras-chave: Frequência fundamental; Fonética acústica; Estilos de fala; Sociofonética

Abstract: We analyze the effect of speaker sex, speaking style and language (English, Estonian, French, German, Italian, Portuguese, Swedish) on global statistical measures of the voice fundamental frequency. The styles studied are interview, sentence reading and word list reading. Typical F0 values used by speakers among languages differ both in male and female speakers. Sentence reading has slightly higher values than interview, but the effect is not uniform among languages. In few cases word reading and interview values differ significantly. All central tendency estimators are affected by the variables tested. Three out of five estimators of dispersion show slightly higher values for male than for female speakers. Among styles, the interview has higher standard deviation and coefficient of variation. Languages do not differ among themselves in terms of F0 variability.

Keywords: Fundamental frequency; Acoustic phonetics; Speaking styles; Sociophonetics

Introdução

Existem diversos trabalhos de natureza descritiva a respeito dos efeitos de variáveis de natureza sociolinguística, estilística, paralinguística e extralinguística sobre a caracterização da variação de longo termo da frequência fundamental (doravante F0) da voz. O tema interessa a muitas áreas de investigação, como a sociofonética, fonoestilística, síntese de fala e fonética forense.

Para a sociofonética (FOULKES, P.; SCOBIE; WATT, 2010; FOULKES, PAUL; DOCHERTY, 2006), estudos dessa natureza interessam na medida em que valores típicos de F0 e sua dispersão são dimensões que falantes podem explorar de maneira sistemática para produzir distinções no interior de comunidades linguísticas e entre línguas diferentes.

No campo da síntese de fala, o estudo da variabilidade acústica encontrada na fala natural, como a que é motivada por diferentes estilos de elocução, fornece dados importantes para o aperfeiçoamento de sistemas de síntese de fala (LLISTERRI, 1992), no sentido de torná-la mais flexível.

Ainda pensando em aplicações de natureza prática, dados dessa natureza, especialmente o estudo

¹ Os autores agradecem ao professor Anders Eriksson, da Universidade de Estocolmo, pela cessão do *corpus* analisado no trabalho. A segunda autora agradece à FAPESP pela Bolsa de Iniciação Científica (processo 2014/21161-5).



da influência dos estilos de elocução sobre a F0, são muito importantes para pesquisadores e praticantes do campo da fonética forense (JESSEN, 2008). A F0 é uma informação muito usada em tarefas de comparação de locutores (GOLD; FRENCH, 2011). Um cenário muito comum na prática forense é aquele em que a amostra de fala questionada, aquela cuja autoria se quer atribuir a um entre um conjunto de falantes suspeitos, foi produzida em um estilo e a amostra de que o perito forense dispõe para análise foi produzida em outro estilo. Esse cenário, que é chamado de incongruência estilística, coloca o seguinte problema: origem das diferenças observadas na F0 podem ser atribuídas à mudança no estilo ou à possibilidade de que outra pessoa tenha produzido a amostra questionada (JESSEN, 2009).

Apesar de numerosos, os estudos de natureza descritiva já publicados deixam lacunas importantes no campo. Muito frequentemente, fornecem poucos detalhes técnicos a respeito das condições testadas, analisam amostras pequenas (poucos estudos têm amostras de mais de 50 falantes), nem sempre fazem uso de estatística inferencial, e abrangem pouca diversidade linguística (predomina um conjunto pequeno de línguas germânicas, sobretudo inglês e alemão). Procuramos, com o presente estudo, preencher algumas dessas lacunas.

Os dois estimadores da variabilidade mais presentes nos estudos a respeito da caracterização estatística de amostras de F0 são a média aritmética, usada como indicador do valor de tendência central da F0 de um falante e o desvio padrão como indicador de dispersão. Esses dois estimadores serão analisados no presente trabalho para que os resultados obtidos possam ser comparados com os reportados na literatura existente. Apesar da popularidade, esses dois estimadores são conhecidos pela sua sensibilidade à presença de valores extremos ou de assimetrias nas amostras a que são aplicados, que pode conduzir a resultados enviesados e eventualmente não representativos da amostra que se quer caracterizar. Uma vez que as amostras de F0 muito frequentemente são assimétricas em algum grau (JASSEM, 1975) e em determinadas situações podem apresentar um número não desprezível de valores extremos (quando um falante produz qualidades de voz não modais, como voz laringalizada ou registro de falsete, por exemplo) a falta de robustez pode ser inconveniente. No presente estudo estudamos três estimadores de tendência central e cinco de dispersão, alguns deles elaborados especificamente tendo em vista amostras de F0.

No que toca a análise dos resultados, todas as comparações passaram por testes estatísticos. Tomamos o cuidado de reportar, sempre que possível, além dos valores de *p* dos testes, o tamanho do efeito. Essa é uma informação frequentemente negligenciada, embora seja

relevante para um entendimento mais completo dos testes estatísticos e útil, por exemplo, para o planejamento de tamanho de amostras de estudos futuros.

Em termos de diversidade linguística, contemplamos sete línguas, seis da família indo-europeia (três do ramo românico e três do ramo germânico) e uma da família urálica (estoniano).

Apresentamos, por fim, uma meta-análise na forma de uma tabela, que sistematiza resultados de mais de 30 estudos sobre o tema do efeito de variáveis como estilo de elocução (principalmente “fala espontânea” e “fala lida”) e outras de natureza paralinguística (fala em ambiente ruidoso) sobre estatísticas globais da F0.

O trabalho reportado aqui tem um caráter eminentemente descritivo e seu principal objetivo é fornecer estimativas de valor e de variabilidade para uma série de estimadores estatísticos disponíveis para o analista da língua e também para o perito envolvido em tarefas de comparação de vozes. Não buscamos testar nenhuma hipótese específica derivada de um quadro analítico ou teórico, mas sim detectar a presença ou não de efeitos nos padrões de F0 causados pelas variáveis linguísticas controladas. Os resultados descritos e analisados aqui devem ser encarados a partir dessa perspectiva.

1 Metodologia

1.1 Material de fala

Usamos gravações feitas no âmbito do projeto internacional “A typology for word stress and speech rhythm based on acoustic and perceptual considerations”, coordenado por Anders Eriksson, da Universidade de Estocolmo. O *corpus* de gravações foi desenhado para permitir a comparação, em condições semelhantes, de diferentes línguas, três estilos de elocução e o sexo dos falantes. O *corpus* abrangia sete línguas até o momento em que foi cedido aos autores do presente artigo: alemão, estoniano, francês, inglês britânico, italiano, português brasileiro e sueco. As amostras das línguas individuais foram coletadas por pesquisadores integrantes do projeto em países em que cada uma das línguas é falada.

Os estilos de elocução coletados no âmbito do *corpus* são: entrevista, leitura de frases e leitura de palavras. No estilo entrevista, um entrevistador (em geral um membro da equipe do projeto) faz perguntas ao participante sobre assuntos diversos, como trabalho, estudos e outros interesses do entrevistado, visando obter uma narrativa não planejada e de extensão variável. Para o estilo leitura de frases, um membro da equipe do projeto selecionou frases ditas pelo participante na entrevista, transcreveu-as ortograficamente e pediu que o participante as lesse em voz alta. As frases selecionadas são declarativas

e interrogativas, embora predominem as primeiras², lidas com foco amplo, isto é, com escopo na frase inteira. No estilo leitura de palavras, o procedimento consistiu na escolha de uma palavra de cada frase presente da etapa anterior e na sua apresentação ao participante na forma de uma lista a ser lida. Uma amostragem aleatória dos dados desse estilo mostra que nas sete línguas do *corpus* a entoação predominante na leitura das palavras foi declarativa, isto é, um contorno descendente, embora existam alguns falantes que produzem contornos com característica continuativa ou suspensiva (GONÇALVES, 1997), que são característicos do assim chamado “efeito de lista”.

Os resultados reportados por Arantes e Eriksson (2014) mostram que um intervalo de tempo entre 10 e 20 segundos de amostra de fala é suficiente para que a estimativa de estimadores de tendência central como a média aritmética, mediana e valor de base se torne estável em torno do seu valor de longo termo. A adição de mais material após esse intervalo tende a não provocar mudanças significativas no valor e na variabilidade dos estimadores. Uma vez que as gravações utilizadas neste estudo têm duração superior a esse limiar, assumimos que os valores dos estimadores, obtidos levando em conta toda sua extensão, refletem predominantemente o efeito do comportamento de longo termo dos falantes e as influências específicas dos estilos de fala e são afetados minimamente por fatores que impactam os contornos de F0 em escala local, como a composição segmental das palavras, a modalidade dos enunciados e o assim “efeito de lista” na leitura de palavras.

Os falantes gravados para o *corpus* têm idades que variam entre 20 e 30 anos, aproximadamente. Para cada língua foram selecionados falantes que os entrevistadores julgaram ter o mesmo dialeto regional. No caso dos dados do português brasileiro, por exemplo, os falantes são todos do interior de São Paulo, de cidades próximas a Campinas.

Analisamos os dados de dez falantes de cada língua, cinco do sexo masculino e cinco do feminino, uma amostra de cada estilo, totalizando 210 amostras de fala (= 7 línguas × 10 falantes × 3 estilos).

1.2 Análise acústica

A extração dos contornos de F0 e das medidas estatísticas de interesse foi feita usando o programa de análise acústica Praat (BOERSMA, 2001) com o auxílio de *scripts* elaborados pelo primeiro autor do presente trabalho. Para a extração inicial dos contornos de F0 das

amostras de fala utilizamos o *script* *better_f0*, que otimiza a escolha dos parâmetros *floor* e *ceiling* da função *To Pitch* do programa Praat, baseada na técnica de autocorrelação. O procedimento de otimização, sugerido por Hirst (2011) tem como objetivo diminuir os erros de estimação de F0 mais comuns produzidos pela função *To Pitch*, como erros de oitava (quando a função sugere um valor que é o dobro ou metade do valor correto) e falsos vozeamentos e desvozeamentos.

Após a extração do contorno por meio do *script* *better_f0*, os arquivos são processados por um segundo *script*, chamado *f0_outliers*, que analisa candidatos vozeados sucessivos no contorno e identifica valores suspeitos, ajudando na detecção de erros não eliminados pelo procedimento aplicado na primeira etapa. O procedimento implementado pelo *script* para identificar um candidato cujo valor pode ser considerado suspeito é analisar a razão entre os valores de F0 de um candidato e seu sucessor no tempo. Se a razão entre o maior e o menor dos dois valores ultrapassa um limiar estipulado pelo usuário e os candidatos estiverem separados por um intervalo de tempo menor do que um limiar de tempo também definido pelo usuário, então o *script* marcará o segundo elemento do par como suspeito a ser examinado posteriormente pelo usuário. Para o presente estudo, adotamos os limiares 1,5 e 80 ms. A adoção do primeiro limiar implica que se o segundo valor do par analisado é 50% maior ou menor do que o primeiro e os dois estiverem separados no tempo por menos do que 80 ms, então o segundo candidato será marcado para análise posterior pelo usuário.

Os pontos no tempo onde se localizam os valores de F0 suspeitos indicados pelo algoritmo são armazenados em um arquivo no formato *TextGrid* do Praat. Nos casos em que o *script* identificou suspeitos, o oscilograma do arquivo de som correspondente ao contorno de F0 em questão foi examinado visualmente para que se decidisse se os valores estimados pelo programa naquele trecho correspondiam à frequência identificada visualmente.

1.3 Estimadores estatísticos coletados

Apresentamos a seguir a lista de estimadores cujo valor foi calculado a partir dos valores de F0 das amostras de cada uma das 210 amostras de fala do *corpus*. Definições e fórmulas são apresentadas para aquelas menos comumente usadas e justifica-se o uso de algumas delas. O cálculo do valor dos estimadores foi feito por meio de um *script* do programa Praat, elaborado pelo primeiro autor especificamente para esse propósito.

Os estimadores de tendência central ou valor típico são:

² Em função dos seus objetivos, os organizadores do *corpus* não tiveram a preocupação de fazer uma seleção balanceada em termos de modalidade.

- Média aritmética (M)
- Mediana (med): 50º quantil da amostra
- Valor de base (F_b): 7,4º quantil da amostra

Os estimadores de dispersão coletados são:

- Desvio padrão (DP).
- Coeficiente de variação (CV): $CV = DP/M$.
- Desvio mediano absoluto (MAD³): $\text{med}(|x_i - \text{med}(x)|)$.⁴
- Desvio mediano absoluto relativo ao valor de base (MADb): $\text{med}(|x_i - F_b(x)|)$.⁵
- Gama tonal: $\log_2(f_{\max}/f_{\min})$.⁶

Os valores da média aritmética e do desvio padrão de cada contorno de F0 foram calculados em Hertz e também em semitons re 1 Hz⁷. É natural que as mulheres apresentem valores de F0 maiores do que os homens em função da dependência do valor de F0 em relação a fatores anatômicos e fisiológicos, tais como o comprimento e a massa das pregas vocais (TITZE, 1994). Essa diferença não é anulada se, na passagem da escala física (Hertz) para a de semitons, se usa o mesmo valor de referência (1 Hz, no caso) para a conversão dos contornos produzidos por mulheres e homens. Por isso, a existência da diferença estatisticamente significativa é bastante previsível nas comparações envolvendo a média aritmética e o fator sexo quer na escala Hertz que na de semitons da maneira como fizemos a conversão. Para que seja possível comparar conjuntamente dados de média da F0 de homens e mulheres observamos, para cada falante, não os valores de cada estilo, mas as diferenças entre as médias dos diferentes estilos na escala st re 1Hz.

No caso do desvio padrão, comparações diretas entre falantes de ambos os sexos são possíveis após a transformação dos valores da amostra da escala Hertz para a de semitons porque o valor de desvio padrão em semitons será o mesmo qualquer que seja o valor escolhido como referência para essa última, seja 1 ou 100 Hz, por exemplo. Isso se dá em função da natureza do cálculo desse estimador, que toma os desvios de cada observação da amostra relativamente a um valor de tendência central. Em função disso, os valores de desvio padrão da F0 de homens e mulheres podem ser comparados diretamente sem recurso a nenhum artifício.

³ Da expressão em inglês *median absolute deviation*.

⁴ Na fórmula, $\text{med}()$ é a função mediana, x representa uma amostra qualquer e x_i é cada uma das observações que compõem a amostra x . O MAD é a mediana do conjunto de valores, tomados em seu valor absoluto, formado pelos desvios de cada valor de uma amostra em relação ao valor mediano dessa mesma amostra.

⁵ O MADb é semelhante ao MAD, mas os desvios são calculados relativamente ao valor de base da amostra e não à sua mediana.

⁶ Na fórmula, f_{\max} e f_{\min} são respectivamente os valores máximo e mínimo de F0 de um contorno, medidos em Hertz. O resultado é expresso em oitavas.

⁷ Use-se a expressão “semitons re 1 Hz” para indicar de maneira abreviada que os valores em semitons são calculados relativamente à frequência 1 Hz.

O valor de base (*base value* ou *base line* em inglês) é um estimador estatístico de localização, proposto por Traunmüller e Eriksson ([S.d.]) que leva em conta as especificidades típicas das amostras de F0, como a variabilidade causada por fatores paralinguísticos – atitude ou estado emocional do falante, por exemplo. O valor de base (F_b) de uma determinada amostra de F0 é obtido pela aplicação da fórmula $F_b = F_{\text{média}} - k\sigma$, em que $F_{\text{média}}$ e σ são, respectivamente, o valor da média aritmética e do desvio-padrão da F0 da amostra em questão e k é uma constante determinada empiricamente. Os autores sugeriram a princípio o valor de 1,5 para a constante, mas, em estudo posterior, Lindh e Eriksson (2007) revisam o valor para 1,43 e sugerem uma formulação alternativa para o cálculo do valor de base. Nessa nova formulação, chamada de *alternative base value*, o ponto $1,43\sigma$ abaixo da média corresponde, aproximadamente, ao 7,4º percentil da distribuição, assumindo-se uma distribuição normal para os dados de F0. Lindh e Eriksson (2007) mostraram que a formulação alternativa é mais robusta do que a original a fatores como a distorção causada pelo canal de gravação e mudanças na distância comunicativa entre interlocutores. Por esse motivo a formulação alternativa foi usada para o cálculo do valor de base neste estudo.

1.4 Análise estatística

A manipulação dos dados produzidos na etapa anterior e sua análise estatística reportadas nas seções a seguir foram feitas no ambiente de computação estatística R (TEAM, 2016).

Empregamos o teste estatístico paramétrico análise de variância (ANOVA) para comparar o efeito das variáveis independentes controladas no estudo sobre as variáveis dependentes, descritas na seção anterior, quando a variável independente cujo efeito se queria estabelecer tinha mais do que dois níveis. O teste- t de amostras independentes foi usado para comparar as médias de variáveis independentes com dois níveis. O teste- t de amostra única foi usado quando foi preciso comparar a média de uma variável a um valor específico dado pela situação de análise. Antes da aplicação dos testes de comparação entre médias testamos sempre o pressuposto da homogeneidade de variância por meio do teste Fligner-Killeen. Recorremos a testes não-paramétricos nos casos em que esse pressuposto não era satisfeito: testes de Wilcoxon e Kruskal-Wallis em lugar respectivamente do teste- t e da ANOVA.

Adotamos um nível de rejeição da hipótese nula de 5%, abaixo do qual declaramos significativos os efeitos apontados por todos os testes, paramétricos ou não.

Reportamos, sempre que possível, medidas de tamanho de efeito em conjunto com os valores de p de cada

teste. Os indicadores calculados foram o g de Hedges e o δ de Cliff. Usamos uma função da biblioteca de funções `compute.es` do R para obter os valores. O primeiro é uma versão não enviesada do d de Cohen (COHEN, 1988), e pode ser entendido como a distância entre duas médias, escalonada em termos de seus respectivos desvios-padrão. O δ de Cliff é uma medida alternativa ao g de Hedges e é robusta à presença de heterogeneidade de variância e não normalidade nas amostras analisadas. Denota a probabilidade de um valor qualquer retirado de uma das duas amostras comparadas ser maior ou menor do que um valor qualquer retirado da outra amostra. Varia entre -1 e 1 e quanto mais próximo dos extremos, maior o tamanho do efeito. Outra maneira de entender seu significado é imaginar que ele mede o grau de sobreposição entre as duas amostras: 0 indica sobreposição completa e ± 1 nenhuma sobreposição. Cohen (1988) sugere que o tamanho do efeito medido por d (e, por extensão, g) seja interpretado como pequeno caso o índice tenha valor até 0,2, médio para valores em torno de 0,5 e grande a partir de 0,8.

1.5 Estratégias de análise para os estimadores

Esperamos um efeito do sexo dos falantes sobre o valor da F0 média significativo e de tamanho grande em função do conhecimento a respeito da influência de fatores anatômicos e fisiológicos sobre a produção da F0. Por conta disso, testamos separadamente na amostra de mulheres e homens o possível efeito combinado do sexo dos falantes e da língua sobre a média da F0. No caso do estudo do efeito do estilo de elocução e de sua interação com o sexo dos falantes, a variável dependente a ser analisada é a diferença, na escala dos semitons, entre as médias da F0 dos três estilos, tomados em dois pares: leitura de frases-entrevista e leitura de palavras-entrevista. As amostras de falantes de ambos os sexos poderão ser analisadas conjuntamente neste caso porque as diferenças serão calculadas entre os valores na escala de semitons.

O valor do desvio padrão foi calculado a partir da amostra de F0 expressa na escala de semitons re 1 Hz. Por conta disso, nas análises desse estimador os dados dos falantes dos dois sexos serão analisados conjuntamente.

Os valores da gama tonal também são expressos em oitavas relativas ao valor mínimo de F0 de cada contorno, uma escala não linear assim como os semitons. Os dados de mulheres e homens foram analisados conjuntamente.

2 Resultados

A seção 2.1 começa descrevendo o efeito do sexo dos falantes sobre os três estimadores de tendência central. As

seções 2.1.1 a 2.1.3 explora a interação entre o sexo dos falantes e as línguas nos três estimadores de tendência central.

A seção 2.2 trata do efeito dos estilos de elocução sobre os três estimadores de tendência central e de suas interações com o sexo dos falantes e as línguas.

A seção 2.3 trata dos efeitos das três variáveis sobre os estimadores de dispersão.

2.1 Efeito do sexo dos falantes sobre os estimadores de tendência central

Tabela 1. Valores médios dos estimadores de tendência central, em Hertz, separados pelo sexo dos falantes e pelas línguas. Na última linha indica-se entre parênteses os valores dos estimadores em semitons.

| Língua | Média | | Mediana | | Valor de base | |
|--------------|----------------------|----------------------|----------------------|----------------------|--------------------|---------------------|
| | Fem. | Masc. | Fem. | Masc. | Fem. | Masc. |
| Alemão | 206 | 118 | 200 | 109 | 165 | 92 |
| Estoniano | 215 | 113 | 212 | 110 | 185 | 97 |
| Francês | 207 | 113 | 202 | 109 | 177 | 93 |
| Inglês | 202 | 112 | 202 | 115 | 177 | 97 |
| Italiano | 201 | 127 | 198 | 135 | 166 | 107 |
| Português | 212 | 142 | 208 | 137 | 180 | 110 |
| Sueco | 191 | 116 | 188 | 114 | 161 | 98 |
| Total | 205 (91,9) | 120 (82,5) | 201 (91,7) | 117 (82,2) | 173 (89) | 99 (79,4) |

O efeito do sexo dos falantes sobre o valor médio dos três estimadores é altamente significativo, como se pode ver a seguir.

- Média:
 - Hertz [$X^2(1) = 153,7 p < 0,001$]; semitons [$X^2(1) = 154,6 p < 0,001$]
 - Tamanho do efeito: 85 Hz; 9,4 st [$g = 4,2 \delta = 1$ nas duas escalas].
- Mediana:
 - Hertz [$F(1, 208) = 960,1 p < 0,001$]; semitons [$X^2(1) = 154,7 p < 0,001$]
 - Tamanho do efeito: 84 Hz [$g = 4,26 \delta = 1$]; 9,5 st [$g = 4,22 \delta = 1$].
- Valor de base:
 - Hertz [$X^2(1) = 140,1 p < 0,001$]; semitons [$F(1, 208) = 616,5 p < 0,001$]
 - Tamanho do efeito: 74 Hz [$g = 3,37 \delta = 0,98$]; 9,6 st [$g = 3,41 \delta = 0,98$].

A magnitude do efeito do sexo dos falantes sobre a média dos três estimadores de tendência central pode ser considerada extremamente alta de acordo com os critérios apresentados na seção 2.4.

2.1.1 Média

Mulheres: [Hz: $X^2(6)=13,7 p<0,05$; st: $X^2(6)=13,4 p<0,05$]. O português é a língua com valor médio mais alto. A diferença mais expressiva é em relação ao sueco: 21 Hz ou 1,8 st a favor do português [Hz: $g=1,55 \delta=0,74$; st: $g=1,52 \delta=0,73$]. Esta é a única diferença significativa do ponto de vista estatístico segundo as comparações múltiplas ($p<0,01$ em ambas as escalas).

Homens: [Hz: $X^2(6)=28,5 p<0,001$; st: $X^2(6)=29 p<0,001$]. O valor médio da F0 do português é aproximadamente 30 Hz ou 4 st maior do que o valor médio do estoniano, do francês, do inglês e do sueco, 24 Hz ou 3,3 st maior do que o valor do alemão e 15 Hz ou 1,6 st maior do que o valor do italiano. Comparações pareadas indicam que o português e o italiano são as línguas que mais diferem das demais, conforme a listagem abaixo (apenas diferenças significativas ou marginalmente significativas são mencionadas):

- Português: estoniano, inglês ($p<0,01$), francês ($p<0,05$), alemão, sueco ($p<0,10$).
- Italiano: inglês, estoniano ($p<0,05$), francês ($p<0,10$).

Com base nos resultados das comparações pareadas é possível dividir as línguas em dois grupos homogêneos: de um lado português, alemão e italiano, com médias mais altas, e o restante das línguas, que apresentam médias mais baixas. A diferença de média entre os dois grupos é de 15,5 Hz em favor do primeiro grupo [Hz: $g=0,88 \delta=0,47$; st: $g=0,84 \delta=0,45$].

2.1.2 Mediana

Mulheres: [Hz: $X^2(6)=12,8 p<0,05$; st: $F(6, 98)=2,2 p<0,05$]. Comparações múltiplas mostram que as únicas diferenças significativas são aquelas entre o sueco e português ($p<0,01$) na escala Hertz e sueco e estoniano na escala de semitons ($p<0,05$). A magnitude das diferenças é de 20,8 Hz/1,8 st em favor do português [$g=1,61 \delta=0,76$] e 24,6 Hz/2 st em favor do estoniano [$g=1,24 \delta=0,63$], ambas consideradas altas de acordo com a interpretação das métricas.

Homens: [Hz: $X^2(6)=29,7 p<0,001$; st: $X^2(6)=29,8 p<0,001$]. Comparações pareadas indicam que o português tem média estatisticamente diferente do estoniano, inglês ($p<0,01$) e francês ($p<0,05$) e marginalmente diferente do alemão e sueco ($p<0,10$) e o italiano tem média diferente do inglês ($p<0,01$), estoniano e francês ($p<0,05$). Os resultados das comparações pareadas permitem dividir as línguas em dois grupos sem sobreposição: de um lado português, e italiano, com medianas mais altas, e estoniano, francês e inglês, que apresentam as medianas mais baixas. A diferença no valor médio da mediana entre

os dois grupos é de 21,7 Hz/3 st em favor do primeiro grupo. As métricas de tamanho de efeito indicam que essa diferença pode ser considerada de magnitude alta [Hz: $g=1,4 \delta=0,68$; st: $g=1,47 \delta=0,71$].

2.1.3 Valor de base

Mulheres: [Hz: $X^2(6)=12,8 p<0,05$; st: $F(6, 98)=2,2 p<0,05$]. Comparações múltiplas mostram que as únicas diferenças significativas são aquelas entre o sueco e português ($p<0,01$) na escala Hertz e sueco e estoniano na escala de semitons ($p<0,05$). A magnitude das diferenças é de 20,8 Hz/1,8 st em favor do português [$g=1,61 \delta=0,76$] e 24,6 Hz/2 st em favor do estoniano [$g=1,24 \delta=0,63$], ambas consideradas altas.

Homens: [Hz: $X^2(6)=29,7 p<0,001$; st: $X^2(6)=29,8 p<0,001$]. Comparações pareadas indicam que o português tem média estatisticamente diferente do estoniano, inglês ($p<0,01$) e francês ($p<0,05$) e marginalmente diferente do alemão e sueco ($p<0,10$) e o italiano tem média diferente do inglês ($p<0,01$), estoniano e francês ($p<0,05$). Os resultados das comparações pareadas permitem dividir as línguas em dois grupos sem sobreposição: de um lado português, e italiano, com valores de base mais altos, e estoniano, francês e inglês, que apresentam os valores de base mais baixos. A diferença no valor médio entre os dois grupos é de 21,7 Hz/3 st em favor do primeiro grupo. As métricas de tamanho de efeito indicam que essa diferença pode ser considerada de magnitude alta [Hz: $g=1,4 \delta=0,68$; st: $g=1,47 \delta=0,71$].

2.2 Efeito dos estilos de elocução sobre os estimadores de tendência central

2.2.1 Média

Tabela 2. Diferença entre a média dos três estilos (st). O símbolo que segue o número na coluna "Todos" indica se aquele valor pode ser considerado significativamente diferente de zero (** indica $p<0,01$, * $p<0,05$ e • $p<0,1$).

| Língua | Estilos de elocução | | | | | |
|--------------|---------------------|-------------|---------------|---------------------|-------------|----------------|
| | frases-entrevista | | | palavras-entrevista | | |
| | Fem. | Masc. | Todos | Fem. | Masc. | Todos |
| Alemão | 0,01 | 1,17 | 0,59* | -1 | -0,16 | -0,58 |
| Estoniano | 0,52 | 1,52 | 1,02* | 0,04 | 1,03 | 0,54 |
| Francês | 0,28 | 1,95 | 1,12* | 0,22 | 1,31 | 0,77 |
| Inglês | 1,44 | 0,87 | 1,16 | 0,14 | -0,03 | 0,05 |
| Italiano | 2,03 | 1,24 | 1,63* | 1,17 | 0,82 | 0,99 |
| Português | -1 | -0,15 | -0,57* | -1,48 | -1,34 | -1,41** |
| Sueco | 0,42 | 1,28 | 0,85* | 0,18 | 1,31 | 0,74 |
| Geral | 0,53 | 1,31 | 0,83* | -0,1 | 0,42 | 0,16 |

A ANOVA indica um efeito significativo do fator estilo: a diferença frases-entrevista (0,83 st) e palavras-

entrevista (0,16 st) é significativa [$F(1, 138)=5,8 p<0,05$], uma diferença de 0,67 st em favor da leitura de frases em relação à de palavras [$g=0,4 \delta=0,23$].

As ANOVAs aplicadas separadamente aos dois níveis da variável diferença entre estilos de elocução para testar o efeito da variável língua mostram os resultados a seguir:

- frases-entrevista: [$F(6, 63)=2,1 p<0,1$], diferença média de 0,76 st entre as seis línguas.
- palavras-entrevista: [$F(6, 63)=3 p<0,05$], diferença significativa entre o português e o italiano ($p<0,05$). Tamanho do efeito de 2,4 st em favor do italiano [$g=1,45 \delta=0,72$].

Procedimento semelhante foi adotado para verificar a influência do sexo dos falantes sobre a diferença entre os estilos. Os resultados das ANOVAs são reportados a seguir:

- frases-entrevista: [$F(1, 68)=2,6 ns$].
- palavras-entrevista: [$F(1, 68)=1,6 ns$].

Para verificar se a diferença média global de 0,83 st em favor do estilo leitura de frase em relação ao estilo entrevista é distinta de zero, aplicamos um teste-*t* de amostra única, que indicou uma diferença significativa [$t(69)=4,4 p<0,001$]. A diferença global de 0,16 st em favor do estilo leitura de palavras não é significativamente diferente de zero [$t(69)=0,8 ns$]. Testes semelhantes foram realizados separadamente por língua (os resultados estatísticos são mostrados na Tab. 2). A diferença frases-entrevista é estatisticamente diferente de zero em três das sete línguas da amostra: estoniano, francês e italiano. A diferença é positiva e maior do que 1 st em todos os casos. A diferença palavras-entrevista só é significativamente diferente de zero no caso do português (a leitura de palavras tem média 1,41 st abaixo da média da entrevista).

2.2.2 Mediana

Tabela 3. Diferença entre a mediana dos três estilos (st). O símbolo que segue o número na coluna "Todos" indica se aquele valor pode ser considerado significativamente diferente de zero (** indica $p<0,01$, * $p<0,05$ e • $p<0,1$).

| Língua | Estilos de elocução | | | | | |
|--------------|---------------------|-------------|---------------|---------------------|-------------|----------------|
| | frases-entrevista | | | palavras-entrevista | | |
| | Fem. | Masc. | Todos | Fem. | Masc. | Todos |
| Alemão | 0,29 | 1,46 | 0,87* | -1,12 | -0,21 | -0,66 |
| Estoniano | 0,74 | 1,62 | 1,18* | 0,42 | 0,95 | 0,68• |
| Francês | 0,53 | 2,12 | 1,33* | 0,52 | 1,66 | 1,09 |
| Inglês | 1,11 | 1,35 | 1,23* | 0,36 | 0,78 | 0,57 |
| Italiano | 1,56 | 1,63 | 1,60** | 0,64 | 1,17 | 0,90* |
| Português | -0,69 | 0,20 | -0,25 | -1,13 | -1,22 | -1,17** |
| Sueco | 0,25 | 1,35 | 0,80• | 0,04 | 1,39 | 0,71 |
| Geral | 0,54 | 1,39 | 0,97 | -0,04 | 0,65 | 0,30 |

A ANOVA indica um efeito significativo do fator estilo: a diferença frases-entrevista (0,97 st) e palavras-entrevista (0,3 st) é significativa [$F(1, 138)=7,5 p<0,01$], uma diferença de 0,67 st em favor da leitura de frases em relação à de palavras [$g=0,47 \delta=0,26$].

As ANOVAs aplicadas separadamente aos dois níveis da variável diferença entre estilos de elocução para testar o efeito da variável língua mostram os resultados a seguir:

- frases-entrevista: [$F(6, 63)=2,5 p<0,05$], comparações múltiplas indicam diferença significativa entre o português, que apresenta a maior diferença negativa entre os estilos, e o italiano, que tem a maior diferença positiva ($p<0,05$). Tamanho do efeito de 1,8 st em favor do italiano [$g=1,81 \delta=0,82$].
- palavras-entrevista: [$F(6, 63)=3,8 p<0,01$], comparações múltiplas indicam diferenças significativas entre o português, que apresenta a maior diferença negativa entre os estilos, e o italiano e o francês, que têm as maiores diferenças positivas ($p<0,05$). Tamanho do efeito de 2,2 st em favor do italiano [$g=1,42 \delta=0,7$].

Procedimento semelhante foi adotado para verificar a influência do sexo dos falantes sobre a diferença entre os estilos. Os resultados das ANOVAs são reportados a seguir:

- frases-entrevista: [$F(1, 68)=8,5 p<0,01$], tamanho do efeito de 0,85 st em favor dos falantes masculinos [$g=0,7 \delta=0,38$].
- palavras-entrevista: [$F(1, 68)=3,5 p<0,01$], tamanho do efeito de 0,69 st em favor dos falantes masculinos [$g=0,44 \delta=0,25$].

Para verificar se a diferença média global, isto é, aquela que engloba as sete línguas, de 0,97 st em favor do estilo leitura de frase em relação ao estilo entrevista é distinta de zero, aplicamos um teste-*t* de amostra única, que indicou uma diferença significativa [$t(69)=6,3 p<0,001$]. A diferença global de 0,3 st em favor do estilo leitura de palavras não é significativamente diferente de zero [$t(69)=0,8 ns$]. Testes semelhantes foram realizados separadamente por língua (os resultados estatísticos são mostrados na Tab. 3). A diferença frases-entrevista só não é estatisticamente diferente de zero em duas das sete línguas da amostra: português e sueco. Nos casos de significância a diferença é positiva e maior do que 0,8 st. A diferença palavras-entrevista só é significativamente diferente de zero no caso do português (a leitura de palavras tem média 1,17 st abaixo da média da entrevista) e do italiano (leitura de palavras tem média 0,9 st acima da entrevista).

2.2.3 Valor de base

Tabela 4. Diferença entre o valor de base dos três estilos (st). O símbolo que segue o número na coluna "Todos" indica se aquele valor pode ser considerado significativamente diferente de zero (** indica $p < 0,01$, * $p < 0,05$ e $\cdot p < 0,1$).

| Língua | Estilos de elocução | | | | | |
|--------------|---------------------|-------------|---------------|---------------------|-------------|------------------------------|
| | frases-entrevista | | | palavras-entrevista | | |
| | Fem. | Masc. | Todos | Fem. | Masc. | Todos |
| Alemão | -0,67 | 0,51 | -0,08 | -0,91 | -0,34 | -0,63 |
| Estoniano | -0,13 | 1,20 | 0,54 | -1,65 | 1,17 | -0,24 |
| Francês | 0,69 | 1,46 | 1,07** | 0,89 | 1,32 | 1,1\cdot |
| Inglês | 2,42 | 0,84 | 1,63 | 0,47 | -0,17 | 0,15 |
| Italiano | 2,53 | 0,77 | 1,65 | 2,52 | 0,71 | 1,61 |
| Português | -0,17 | 0,37 | 0,10 | -1,89 | -1,27 | -1,58* |
| Sueco | 2,24 | 0,99 | 1,61 | 2,31 | 0,96 | 1,64 |
| Geral | 0,99 | 0,88 | 0,93 | 0,25 | 0,34 | 0,29 |

A ANOVA não apontou a existência de um efeito significativo do fator estilo: a diferença frases-entrevista (0,93 st) e palavras-entrevista (0,29 st) não é significativa [$F(1, 138)=2 ns$]. A diferença é de 0,64 st em favor da leitura de frases em relação à de palavras [$g=0,24 \delta=0,13$].

As ANOVAs aplicadas separadamente aos dois níveis da variável diferença entre estilos de elocução para testar o efeito da variável língua mostram os resultados a seguir:

- frases-entrevista: [$F(6, 63)=0,8 ns$], diferença média de 0,89 st entre as seis línguas.
- palavras-entrevista: [$F(6, 63)=2 p < 0,1$], diferença média de 1,5 st entre as seis línguas.

Procedimento semelhante foi adotado para verificar a influência do sexo dos falantes sobre a diferença entre os estilos. Os resultados das ANOVAs são reportados a seguir:

- frases-entrevista: [$F(1, 68)=0,03 ns$], tamanho do efeito de 0,11 st em favor dos falantes masculinos [$g=0,04 \delta=0,02$].
- palavras-entrevista: [$F(1, 68)=0,02 ns$], tamanho do efeito de 0,09 st em favor dos falantes masculinos [$g=0,03 \delta=0,02$].

Para verificar se a diferença média global, isto é, aquela que engloba as sete línguas, de 0,93 st em favor do estilo leitura de frase em relação ao estilo entrevista é distinta de zero, aplicamos um teste- t de amostra única, que indicou uma diferença significativa [$t(69)=3 p < 0,01$]. A diferença global de 0,3 st em favor do estilo leitura de palavras não é significativamente diferente de zero [$t(69)=0,9 ns$]. Testes semelhantes foram realizados separadamente por língua (os resultados estatísticos são mostrados na Tabela 4). A diferença frases-entrevista só é estatisticamente diferente de zero no caso do francês (valor de base 1,07 st maior no estilo leitura de frases).

A diferença palavras-entrevista só é significativamente diferente de zero no caso do português (a leitura de palavras tem média 1,58 st abaixo da média da entrevista).

2.3 Estimadores de dispersão

2.3.1 Desvio padrão

Tabela 5. Desvio padrão médio (st) agrupado pelos estilos de elocução, sexo dos falantes e pela língua. A última linha mostra a média global do DP.

| Língua | Sexo | | Estilos de elocução | | |
|--------------|------------|-------------|---------------------|-------------|-------------|
| | Fem. | Masc. | entrevista | frases | palavras |
| Alemão | 2,09 | 2,52 | 2,26 | 2,42 | 2,24 |
| Estoniano | 1,9 | 2,09 | 1,97 | 2,1 | 1,91 |
| Francês | 2,05 | 2,52 | 2,58 | 2,28 | 1,99 |
| Inglês | 2,68 | 2,52 | 3,19 | 2,29 | 2,32 |
| Italiano | 2,44 | 2,11 | 2,64 | 2,24 | 1,95 |
| Português | 2,08 | 3,06 | 2,92 | 2,1 | 2,7 |
| Sueco | 2,15 | 2,27 | 2,56 | 2,09 | 1,99 |
| Geral | 2,2 | 2,44 | 2,59 | 2,22 | 2,16 |

Análise de variância de 3 fatores (sexo, estilo e língua), tendo como variável dependente o valor do desvio padrão, medido em semitons.

- Sexo: [$F(1, 200)=4,4 p < 0,05$], falantes masculinos apresentam valores médios de desvio-padrão significativamente maiores do ponto de vista estatístico do que os falantes do sexo feminino. Tamanho do efeito é 0,24 st em favor dos falantes masculinos [$g=0,28 \delta=0,16$].
- Estilo: [$F(2, 200)=5,5 p < 0,01$]. Testes de comparação pareada mostram que o estilo entrevista tem variabilidade maior do que os outros dois estilos [frases-entrevista: $p < 0,05 g=0,43 \delta=0,24$; palavras-entrevista: $p < 0,01 g=0,45 \delta=0,25$], frases-palavras: $ns [g=0,08 \delta=0,05]$.
- Língua: [$F(6, 200)=1,9 p < 0,1$], diferença média de 0,24 st entre as línguas.

2.3.2 Coeficiente de variação

Análise de variância de 3 fatores (sexo, estilo e língua), tendo como variável dependente o valor do coeficiente de variação, expresso como porcentagem do desvio padrão da F0 relativo à média da F0.

- Sexo: [$F(1, 200)=19 p < 0,001$], mulheres 2,4% homens 2,96%. Tamanho do efeito é 0,56 ponto percentual em favor dos falantes masculinos [$g=0,57 \delta=0,31$].
- Estilo: [$F(2, 200)=5,6 p < 0,01$]. entrevista 2,99%; leitura de frases 2,56% st; leitura de palavras 2,5% st. Testes de comparação pareada mostram

que o estilo entrevista tem variabilidade maior do que os outros dois estilos [frases-entrevista: $p < 0,05$ $g = 0,43$ $\delta = 0,24$; palavras-entrevista: $p < 0,05$ $g = 0,44$ $\delta = 0,25$], frases-palavras: *ns* [$g = 0,06$ $\delta = 0,04$].

- Língua: [F(6, 200) = 1,8 *ns*], diferença média de 0,27 ponto percentual entre as línguas.

2.3.3 MAD

Para testar o efeito do sexo do falante, estilo de elocução e língua sobre o valor médio do MAD, rodamos uma ANOVA de três fatores. Apenas a variável sexo mostrou efeito significativo sobre a média da variável dependente.

- Sexo: [F(1, 200) = 9,8 $p < 0,01$] falantes femininas 1,32 st; falantes masculinos 1,55 st [$g = 0,43$ $\delta = 0,24$].
- Estilo: [F(2, 200) = 0,6 *ns*] entrevista 1,45 st; leitura de frases 1,48 st; leitura de palavras 1,39 st.
- Língua: [F(6, 200) = 1,8 $p < 0,1$] alemão 1,5 st; estoniano 1,21 st; francês 1,4 st; inglês 1,56 st; italiano 1,44 st; português 1,59 st; sueco 1,36 st. Diferença média de 0,15 st entre as línguas.

2.3.4 MAD_b

Para testar o efeito do sexo do falante, estilo de elocução e língua sobre o valor médio do MAD_b, rodamos uma ANOVA de três fatores. Nenhuma das variáveis independentes tem efeito significativo sobre a média da variável dependente.

- Sexo: [F(1, 200) = 0,02 *ns*] falantes femininas 2,79 st; falantes masculinos 2,82 st [$g = 0,02$ $\delta = 0,01$].

- Estilo: [F(2, 200) = 0,004 *ns*] entrevista 2,8 st; leitura de frases 2,82 st; leitura de palavras 2,8 st.
- Língua: [F(6, 200) = 0,9 *ns*] alemão 2,68 st; estoniano 2,42 st; francês st; inglês 3,27 st; italiano 2,9 st; português 3,17 st; sueco 2,71 st. Diferença média de 0,4 st entre as línguas.

2.3.5 Gama tonal

Uma ANOVA de três fatores foi usada para estabelecer a influência das variáveis sexo dos falantes, estilo de elocução e língua sobre os valores médios de gama tonal.

- Sexo: [F(1, 200) = 2,1 *ns*], falantes femininas 1,36 8va; falantes masculinos 1,41 8va [$g = 0,15$ $\delta = 0,08$].
- Estilo de elocução: [F(2, 200) = 54 $p < 0,001$], entrevista (1,66 8va), frases (1,35 8va) e palavras (1,14 8va); frases-entrevista ($p < 0,001$ $g = 0,84$ $\delta = 0,45$); palavras-entrevista ($p < 0,001$ $g = 1,4$ $\delta = 0,68$); frases-palavras ($p < 0,001$ $g = 0,77$ $\delta = 0,42$).
- Língua: [F(6, 200) = 11 $p < 0,001$], análise de agrupamentos homogêneos das comparações pareadas indicam que o estoniano (1,09 8va) e o português (1,6 8va), as línguas com menor e maior gama média, formam grupos sem interseção [$g = 1,29$ $\delta = 0,65$].

Inspeção visual dos *boxplots* mostrados na Figura 1 sugere que as línguas não são homogêneas no que diz respeito às diferenças de extensão da gama tonal existentes entre os três estilos. É possível propor a divisão das línguas em dois grandes grupos. De um lado alemão, francês, inglês, e português, nas quais o estilo

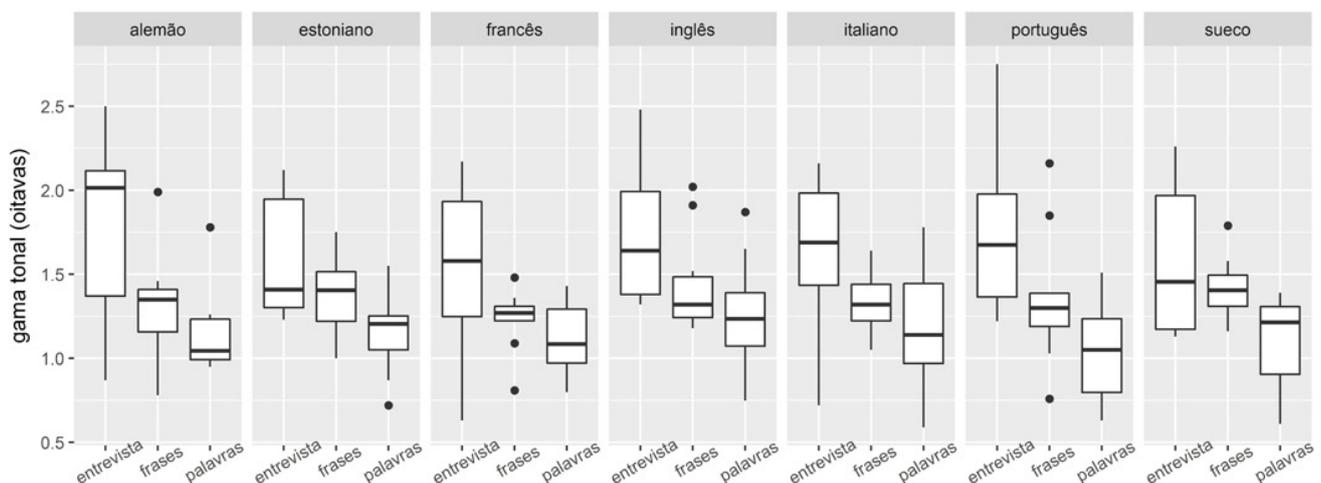


Figura 1. *Boxplots* mostrando distribuição dos dados de gama tonal (oitavas), das sete línguas do corpus nos três estilos de elocução.

entrevista apresenta valores de gama maiores do que os demais: em torno de duas oitavas contra valores entre 1 e 1,5 oitavas para os demais. De outro lado, estoniano, italiano e sueco, nas quais não há diferenças marcantes entre os três estilos ou o estilo leitura de palavras tem o menor valor comparado com os demais. Essa análise é corroborada pelos resultados das comparações pareadas, apresentadas a seguir.

- Alemão: diferenças entre entrevista e leitura de palavras ($p < 0,05$) e leitura de palavras e leitura de frases ($p < 0,01$).
- Estoniano: nenhuma diferença significativa entre os estilos.
- Francês: diferenças significativas entre todos os estilos ($p < 0,05$)
- Inglês: diferenças significativas entre todos os estilos ($p < 0,05$)
- Italiano: diferença apenas entre entrevista e leitura de palavras ($p < 0,05$)
- Português: estilo entrevista diferente dos demais ($p < 0,05$)
- Sueco: diferença apenas entre entrevista e leitura de palavras ($p < 0,05$).

3 Discussão

Os resultados que mostram um efeito significativo e de grande magnitude do sexo dos falantes sobre a média de F0, reportados na seção 3.1, são previsíveis em função do que se sabe sobre a influência de fatores anatômicos e fisiológicos na produção da F0, em especial o comprimento e a massa das pregas vocais (TITZE, 1994). A diferença entre os dois sexos a esse respeito é sabida e bastante descrita na literatura. Os valores da média obtidos no presente estudo, 205 Hz para as mulheres e 120 Hz para os homens, são muito semelhantes aos reportados no estudo meta-analítico de Traunmüller e Eriksson ([S.d.]): 210 Hz para mulheres e 120 Hz para homens. O mesmo princípio explica também as diferenças observadas nos estimadores mediana e valor de base, muito embora esses dois sejam muito menos reportados na literatura do que a média. Note-se que o tamanho do efeito é um pouco menor no caso do valor de base, o que pode ser explicado pelo fato do valor de base ser, entre os três estimadores, o menos afetado pelos valores extremos introduzidos nas amostras pelas excursões positivas da F0.

Os resultados relativos aos estimadores de dispersão mostram efeito significativo do sexo no caso do desvio padrão, coeficiente de variação e MAD. Em todos os casos, os valores dos falantes homens são maiores do que os das mulheres, embora o tamanho do efeito seja pequeno no caso do DP e médio nos demais. Esses resultados divergem

daquele relatado por Traunmüller e Eriksson, que não encontram diferenças no DP de homens e mulheres quando ele é expresso em semitons. A magnitude dos valores encontrados por eles é ligeiramente maior (2,7 para homens e 2,8 para mulheres), embora não seja significativa do ponto de vista estatístico. Gama tonal e MAD_b não mostram efeito do sexo dos falantes. O estudo dos autores suecos não traz dados sobre gama tonal assim como definimos essa medida no presente trabalho. Não temos notícia de estudos que tenham reportado valores de MAD_b para que se pudesse comparar os resultados. Os resultados que mostram haver diferença significativa entre os sexos nos estimadores de variabilidade são inesperados se tomarmos o trabalho dos autores suecos como referência. Essa diferença de variabilidade não pode ser facilmente atribuída a diferenças anatômicas ou fisiológicas que explicam diferenças na média. Resta como hipótese explicativa que ela seja produto de diferenças culturais ou sociais entre os dois sexos.

Os resultados do presente estudo que dizem respeito ao efeito dos estilos de elocução sobre os estimadores estatísticos indicam que apenas no caso do estimador média a diferença entre os estilos leitura de frases e entrevista (LF-E) é significativamente diferente de zero (0,83 st). A diferença entre a leitura de palavras e a entrevista (LP-E) não é significativa. As duas amostras (LF-E e LP-E) são estatisticamente diferentes entre si. Considerando todas as línguas, observa-se que o estilo leitura de frases tende a apresentar valores mais altos de tendência central do que o estilo entrevista, embora essa diferença seja diferente de zero do ponto de vista estatístico apenas no caso da média. Tomando as línguas isoladamente, essa diferença é significativa em muitos casos. O português é uma exceção a esse padrão. Para essa língua, o estilo entrevista tende a apresentar valores de tendência central maiores do que o do estilo leitura de frases. Uma possível explicação para essa diferença pode estar no procedimento de coleta do material. Nas gravações do estilo entrevista feitas pela equipe brasileira do projeto o entrevistador e o entrevistado em geral se conheciam e tinham algum grau de intimidade. Pode-se imaginar que a familiaridade entre os interlocutores pode ter colaborado para um maior grau de coloquialidade e espontaneidade na gravação, que conduziu à produção de fala com mais vivacidade no estilo entrevista. No caso das gravações das outras línguas não havia familiaridade entre entrevistado e entrevistador.

A maioria dos estudos presentes na amostra reunida na meta-análise não aplicou testes de inferência para estabelecer o grau de significância das diferenças observadas. Em três deles, essa informação é explicitamente mencionada e apenas em um caso a

diferença atinge significância. Em alguns casos, o valor numérico da média de cada estilo é reportado (na maior parte dos casos em Hz) e, em outros, apenas qual dos estilos investigados obteve o maior valor. A F0 média do estilo leitura é maior do que a do estilo fala espontânea em 22 dos conjuntos de resultados. O contrário acontece em sete casos. Nos casos em que o estilo leitura teve valor maior do que o estilo fala espontânea, a média da diferença é de aproximadamente 7,9 Hz ou 0,8 st relativo ao valor da média da fala espontânea. Esse resultado é bastante próximo ao que foi encontrado no presente estudo para a diferença entre o estilo leitura de frases e entrevista.

Nenhuma diferença entre leitura de frases e entrevista e leitura de palavras e entrevista é significativa no caso dos estimadores mediana (LF-E: 0,97 st, LP-E: 0,3 st) e valor de base (LF-E: 0,93 st, LP-E: 0,29 st), muito embora as diferenças sejam, em termos absolutos, maiores do que as observadas no caso da média. No caso da mediana, as diferenças LF-E são significativamente maiores do que as LP-E, embora nenhuma delas seja estatisticamente diferente de zero quando tomadas isoladamente.

Os resultados do presente estudo corroboram em alguma medida os resultados encontrados na literatura prévia, que sugerem uma diferença positiva de baixa magnitude na média de F0 do estilo leitura em comparação ao que neste trabalho chamamos estilo entrevista, isto é, uma forma de produção verbal de tipo narrativo, produzido com baixo grau de automonitoramento por parte do falante.

O sexo dos falantes interagiu com os estilos de forma significativa apenas nos estimadores média e mediana. As diferenças entre estilos foram maiores nos homens do que nas mulheres. Essa diferença não foi explorada sistematicamente na literatura prévia.

A variável língua exerceu efeito significativo sobre os valores dos três estimadores de tendência central. As diferenças entre línguas não são uniformes entre mulheres e homens. As diferenças significativas no caso das mulheres, considerando os três estimadores, foram encontradas entre o português e o estoniano, com valores maiores, e sueco, com valores menores. No caso dos falantes masculinos, o grupo das línguas com valores maiores nos três estimadores é integrado pelo português e pelo italiano, enquanto sueco (média), estoniano, francês e inglês (mediana e valor de base) têm os valores mais baixos. A magnitude do efeito é muito elevada em todos os casos (em torno de 20 Hz, $g > 1$ no mínimo).

O estoniano apresenta uma peculiaridade: os valores dos estimadores das falantes femininas estão entre os mais altos e valores dos falantes masculinos entre os mais baixos no conjunto das sete línguas investigadas.

Esse padrão aconteceu nos três estimadores de tendência central.

Houve variabilidade interlinguística também na maneira como os estilos de elocução influenciaram os valores dos estimadores de tendência central. O português foi a única língua na qual a diferença entre os estilos leitura de palavras e entrevista é negativa e significativamente diferente de zero em todos os estimadores. Apresentou também diferenças negativas entre os estilos leitura de frases e entrevista no caso da média e da mediana, embora não significativamente diferentes de zero. Além desses casos envolvendo o português, houve ocorrência de diferença negativa envolvendo o alemão (todos os estimadores, diferença LP-E; falantes femininos, valor de base, diferença LF-E) e o estoniano (valor de base, falantes femininos, diferença LP-E).

Conclusões

De maneira geral, o que os resultados do presente estudo sugerem é que, em termos dos estimadores de tendência central, a mediana e valor de base são mais robustos à influência dos fatores investigados. O valor de base é um estimador notável na medida em que ele se mostrou robusto à influência dos três fatores. O uso desse estimador no âmbito de aplicações forenses tem a vantagem de ser menos suscetível ao problema da incongruência entre estilos, mencionado na introdução. Entre os estimadores de dispersão, aqueles que se mostraram mais robustos às influências das variáveis estudadas foram o MAD_b e o MAD . É interessante observar que a média e o desvio padrão, os dois parâmetros mais estudados na literatura e mais usados para a caracterização dos padrões de F0 em cenários forenses, sejam os mais afetados pelas variáveis estudadas no presente trabalho. Os resultados sugerem que as práticas no campo devem ser revistas, especialmente em situações em que é importante que os estimadores precisem ser robustos à influência de fatores como a mudança no estilo de elocução.

Os dados documentam a existência de grande variabilidade motivada pelas línguas, mostrando que não é possível assumir a universalidade dos padrões observados nos dados já coletados e analisados na literatura. Como sugerem Traunmüller e Eriksson, o valor típico de F0 (que a literatura reputa ser uma característica determinada fisiologicamente) e a sua variabilidade são dimensões que podem ser definidas culturalmente no interior das comunidades falantes. Por isso é importante coletar dados de um grande número de línguas para que seja possível ter uma ideia mais clara a respeito da diversidade efetivamente existente.

Referências

- ARANTES, Pablo; ERIKSSON, Anders. *Temporal stability of long-term measures of fundamental frequency*. 2014, Dublin: ISCA, 2014. p. 1149-1152.
- BOERSMA, Paul. Praat, a system for doing phonetics by computer. *Glott International*, v. 5, n. 9/10, p. 341-345, 2001.
- COHEN, Jacob. *Statistical Power Analysis for the Behavioral Sciences*. New Jersey: Lawrence Erlbaum Associates, 1988.
- ESKÉNAZI, Maxine. Trends in Speaking Styles Research. 1993, Berlin: ISCA, 1993. p. 501-509. Disponível em: <http://www.isca-speech.org/archive/eurospeech_1993>.
- FOULKES, Paul; SCOBIE, James M.; WATT, Dominic J.L. Sociophonetics. In: HARDCASTLE, William; LAVER, John; GIBBON, Fiona (Org.). *Handbook of Phonetic Sciences*. 2. ed. Oxford: Blackwell, 2010. p. 703-754.
- FOULKES, Paul; DOCHERTY, Gerard. The social life of phonetics and phonology. *Journal of Phonetics*, v. 34, p. 409-438, 2006.
- GOLD, Erica; FRENCH, Peter. International practices in forensic speaker comparison. *The International Journal of Speech, Language and the Law*, v. 18, n. 2, p. 293-307, 2011.
- GONÇALVES, Carlos A. *Focalização no português do Brasil*. Rio de Janeiro: UFRJ, 1997.
- GUTZEIT, Suska; ARANTES, Pablo. *Estabilidade de medidas estatísticas de frequência fundamental: comparação da variabilidade intra e interlinguística*. São Carlos: Sociedade Brasileira para o Progresso da Ciência, 2015.
- HIRST, Daniel J. The Analysis by Synthesis of Speech Melody: from Data to Models. *Journal of Speech Sciences*, v. 1, n. 1, p. 55-83, 2011.
- HOLLIEN, Harry; HOLLIEN, Patricia; JONG, Gea De. Effects of three parameters on speaking fundamental frequency. *Journal of the Acoustical Society of America*, v. 102, n. 5, p. 2984-2992, 1997.
- JASSEM, Wiktor. Normalisation of F0 curves. In: FANT, Gunnar; TATHAM, M. A. A. (Org.). *Auditory Analysis and Perception of Speech*. London: Academic Press, 1975. p. 523-530.
- JESSEN, Michael. Forensic Phonetics. *Language and Linguistics Compass*, v. 2, n. 4, p. 671-711, 2008.
- JESSEN, Michael. Forensic phonetics and the influence of speaking style on global measures of fundamental frequency. In: GREWENDORF, Günther; RATHERT, Monika (Org.). *Formal linguistics and law*. Berlin: Mouton de Gruyter, 2009. p. 115-139.
- KARLSSON, I. et al. *Within-speaker variability due to speaking manners*. 1998, Sydney: [s.n.], 1998. p. 2379-2382.
- KÜNZEL, Hermann. Some general phonetic and forensic aspects of speaking tempo. *Forensic Linguistics*, v. 4, n. 1, p. 48-83, 1997.
- LINDH, Jonas; ERIKSSON, Anders. *Robustness of Long Time Measures of Fundamental Frequency*. Antwerp, Belgium: [s.n.], 2007. p. 2025-2028.
- LLISTERRI, Joaquim. *Speaking styles in speech research*. Dublin, Ireland: [s.n.], 1992.
- TEAM, R Core. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2016. Disponível em: <<https://www.R-project.org/>>.
- TITZE, Ingo. *Principles of voice production*. Englewood Cliffs: Prentice Hall, 1994.
- TRAUNMÜLLER, Hartmut; ERIKSSON, Anders. *The frequency range of the voice fundamental in the speech of male and female adults*. [S.d.]. Disponível em: <http://www2.ling.su.se/staff/hartmut/f0_m&f.pdf>.

ANEXO

Meta-análise

Neste anexo apresentamos os resultados de uma meta-análise de parte da literatura a respeito da variação na média e no desvio padrão da F0 motivada por diferentes estilos de fala. Pesquisas sobre o tema em mecanismos de busca de periódicos científicos permitiram encontrar seis artigos que compilam e comentam um número grande de resultados previamente reportados na literatura. Os artigos são: ESKÉNAZI, 1993; HOLLIEN; HOLLIEN; JONG, 1997; JESSEN, 2009; KARLSSON et al., 1998; KÜNZEL, 1997; LLISTERRI, 1992.

Organizamos e apresentamos os dados presentes nesses trabalhos no QUADRO a seguir. Informações relevantes sobre os estudos foram coletadas sempre que estavam disponíveis ou podiam ser inferidas a partir do texto. São elas: autor da pesquisa, língua investigada, número, sexo e idade de falantes observados, estilos de fala investigados, valores dos estimadores medidos, significância estatísticas das diferenças observadas. Células em branco indicam que a informação não estava disponível.

Abreviações usadas na tabela: leitura (L), fala espontânea (FE), fala semiespontânea (FS), contagem de números (CN), comparações estatísticas indicam diferença significativa (*), comparações estatísticas não indicam diferença significativa.

| Autor | Idioma | N | Sexo | Idade | Estilo | Tendência central | Dispersão |
|--------------------------------|----------|-----|-----------|-------|--|--|---|
| Snidecor (1943) | | 6 | M | | L, FE | Média: L 132 Hz FE 120 Hz | DP: L > FE |
| Hirson et al. (1995) | Inglês | 20 | M | 21-37 | L, FE | Média: L 121,5 Hz, FE 106,9 Hz | |
| | | | | 21-37 | L e FE por telefone | Média: L 121.2 Hz, FE 108,9 Hz | |
| Hollien et al. (1997) | Inglês | 324 | M | 18-30 | L, FE | Média: L 129 Hz, FE 123 Hz | DP: L = FE |
| | | | M | 18-30 | L, FE | Média: L 122 Hz, FE 116 Hz | DP: L = FE |
| | | | M | 18-30 | L, FE | Média: L 114 Hz, FE 110 Hz | DP: FE levemente > L |
| Laan (1997) | Holandês | 2 | M | | L, FE | Média: L > FE (diferença significativa apenas para um falante) | Gama: L 90% > para o sujeito com diferença significativa na média, o outro sujeito não apresentou diferença significativa |
| Schultz-Coulon (1975) | Alemão | 11 | F, M | | L, FE, CN | Média: L 117 Hz (apenas homens), CN 110 Hz, FE 107 Hz | |
| Batliner et al. (1995) | | 4 | 3F, 1M | | L, FE | Média: FE > L | DP: L > FE para três sujeitos. FE > L para uma das falantes |
| Künzel (1997) | Alemão | 10 | 5M, 5F | | L, FE, FS | L obteve maior média * | |
| | | | | | L, FE, FS por telefone | L obteve maior média * | |
| Tjaden (2000) | Inglês | 9 | M | 50-80 | L, FE | | Gama: L 64 Hz, FE 55Hz |
| Köster (2002) | Alemão | | | | 3 estilos: dentre eles FE e L | Sem ruído: FE levemente > L ** Com ruído: L levemente > FE ** | Interquartis muito parecidos para L e FE, |
| Mixdorff and Pfitzinger (2005) | Alemão | 4 | 3F, 1M | | L, FE | Média: F: FE > L para duas falantes; para a terceira praticamente não houve diferença. H: L > FE | DP: F: FE > L. H: FE = L |
| Jessen (2009) | Alemão | 100 | M | 21-63 | FE e L com ruído | Média: FE > L | CV: FE > L |
| | | | | | FE e L com ruído neutro | Média: FE > L** | CV: FE > L |
| | | | | | FE e L por telefone | Média: FE > L** | CV: FE > L |
| Koopmans-van Beinum (1991) | Holandês | | | | L, FE | Média: L > FE. Mediana obteve pequeno aumento nas palavras em foco | |
| Blaauw (1991) | Holandês | | | | L, FE | Média: L > FE | Gama: L > FE |
| Bhatt e Léon (1991) | Francês | | | | Noticiários, introduções a concertos, comentários esportivos | | Gama: noticiários obtiveram o menor valor |
| Picheny et al. (1985) | | | | | Fala clara, conversação | | F0 máxima: Fala clara > FE. Gama: Fala clara > FE |
| Koopmans-van Beinum (1991) | | | | | L, FE | Mediana: L obteve maior valor | |
| Koopmans-van Beinum (1992) | | | | | L, FE | Mediana: L obteve maior valor | |
| Koopmans-van Beinum (1991) | | | | | FE, L | Mediana: L > FE | |
| Clark et al. (1987) | | | | | | Mediana: sentenças com ruído obtiveram maior valor | |
| Eskénazi (1992) | | | | | L, Fala clara, FE | | Gama: L obteve maior valor (para um subconjunto de falantes) |
| Ayers (1991) | | | | | L, FE | | Gama: L = FE. Movimento de F0: FE apresentou maior variedade de tipos de contornos que L |
| Hollien et al. (1973) | Inglês | 157 | M | 18-26 | L, FE | Média: L 129HZ, FE 123Hz. Em semitons: L 35.7, FE 34.8 | DP em tons: L 1.6, FE 1.6 |
| (dados não declarados) | | 142 | M | 18-26 | L, FE | Média: L 122Hz, FE 116Hz. Em semitons: L 34.6, FE 33.8 | DP em tons: L 1.6, FE 1.7 |

(continua)

(conclusão)

| Autor | Idioma | N | Sexo | Idade | Estilo | Tendência central | Dispersão |
|------------------------|--------|----|--------|-------|--|--|---|
| Hollien et al. (1997) | Inglês | 25 | M | 21-30 | L, FE | <i>Média:</i> L 114Hz, FE 110Hz. Em semitons: L 33.6, FE 33 | <i>DP em tons:</i> L 1.3, FE 1.3 |
| | | 18 | F | 21-30 | L, FE | <i>Média:</i> L 198Hz, FE 190Hz. Em semitons: L 43.2, FE 42.4 | <i>DP em semitons:</i> L 1.2, FE 1.1 |
| Karlsson et al. (1998) | Sueco | 50 | M | 22-78 | Fala neutra (dados memorizados), fala com ruído alto, fala com estresse (raciocínio lógico e tarefa de reconhecimento auditivo com ruído ao fundo) | <i>Média:</i> sentenças com ruído 5% @ > neutras. Sentenças com ênfase diferiram ligeiramente entre os dois grupos | |
| Künzel (1997) | | 10 | 5F, 5H | 20-26 | FE, FS, L | <i>Média:</i> F: FE 199.1Hz, FS 208.2Hz, L 213Hz. M: FE 126.5Hz, FS 126.8Hz, L 132.3Hz. | <i>DP:</i> F: FE 17.3, FS 19.1, L 18.9. M: FE 24.8, FS 29.2, L 24.8 |

Recebido: 30 de setembro de 2016
Aprovado: 07 de março de 2017
Contato: pabloarantes@gmail.com