

Artigos originais

Comparação de medidas de frequência fundamental e frequências dos formantes em duas tarefas de fala

*Comparison of fundamental frequency and formants frequency measurements in two speech tasks*Flávia Viegas^{1,3}<https://orcid.org/0000-0003-1993-8288>Danieli Viegas²<https://orcid.org/0000-0002-9522-1388>Glaucio Serra Guimarães³<https://orcid.org/0000-0003-4573-1053>Margareth Maria Gomes de Souza⁴<https://orcid.org/0000-0003-3937-415X>Ronir Raggio Luiz⁴<https://orcid.org/0000-0002-7784-9905>Marcia Simões-Zenari¹<https://orcid.org/0000-0003-0505-6655>Katia Nemr¹<https://orcid.org/0000-0002-8662-2702>

¹ Universidade de São Paulo - USP, São Paulo, São Paulo, Brasil.

² Prefeitura da Cidade do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro, Brasil.

³ Universidade Federal Fluminense - UFF, Rio de Janeiro, Rio de Janeiro, Brasil.

⁴ Universidade Federal do Rio de Janeiro - UFRJ, Rio de Janeiro, Rio de Janeiro, Brasil.

Fonte de auxílio à pesquisa: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001

Conflito de interesses: Inexistente



Recebido em: 05/08/2019

Aceito em: 30/10/2019

Endereço para correspondência:

Flávia Viegas

Instituto de Saúde de Nova Friburgo - Universidade Federal Fluminense

Rua Dr. Silvio Henrique Braune, 22 – Centro

CEP: 28625-650 - Nova Friburgo,

Rio de Janeiro, Brasil

E-mail: fviegas@id.uff.br

RESUMO

Objetivo: comparar as medidas de frequência fundamental (F0) e frequência dos dois primeiros formantes (F1 e F2) das sete vogais orais do português brasileiro em duas tarefas de fala em adultos sem distúrbios de voz e fala.

Métodos: oitenta participantes entre 18-40 anos pareados por gênero foram selecionados após avaliações orofacial, ortodôntica e perceptivo-auditiva da voz e fala. Os sinais de fala foram obtidos de sentenças-veículo e vogais sustentadas e foram estimados os valores de F0 e frequências de F1 e F2. As diferenças foram verificadas por meio do teste t e foi calculado o *Effect Size*.

Resultados: foram encontradas diferenças nas medidas de f0 entre as duas tarefas de fala em duas vogais no gênero masculino e em cinco vogais no gênero feminino. Nas frequências de F1 foram notadas diferenças em seis vogais nos homens e em duas nas mulheres. Nas frequências de F2 houve diferença em quatro vogais nos homens e em três nas mulheres.

Conclusão: a partir das diferenças encontradas, conclui-se que a tarefa de fala para avaliação de frequência fundamental e frequências dos formantes no português brasileiro pode demonstrar resultados distintos tanto em medidas glóticas como supraglóticas na produção das diferentes vogais orais deste idioma. Desta forma, sugere-se que os clínicos e pesquisadores considerem ambas formas de emissão para interpretação mais apurada das implicações destes dados na avaliação da comunicação oral e no direcionamento de condutas terapêuticas.

Descritores: Voz; Medida da Produção da Fala; Acústica da Fala; Fonética

ABSTRACT

Purpose: to compare the measurements of fundamental frequency (F0) and frequency of the first two formants (F1 and F2) of the seven oral vowels of the Brazilian Portuguese in two speech tasks, in adults without voice and speech disorders.

Methods: eighty participants in the age range 18 and 40 years, paired by gender, were selected after orofacial, orthodontic and auditory-perceptual assessments of voice and speech. The speech signals were obtained from carrier phrases and sustained vowels and the values of the F0 and frequencies of F1 and F2 were estimated. The differences were verified through the *t* Test, and the effect size was calculated.

Results: differences were found in the F0 measurements between the two speech tasks, in two vowels in males, and in five vowels, in females. In the F1 frequencies, differences were noted in six vowels, in men, and in two, in women. In the F2 frequencies, there was a difference in four vowels, in men, and three, in women.

Conclusion: based on the differences found, it is concluded that the speech task for evaluation of fundamental frequency and formants' frequencies, in the Brazilian Portuguese, can show distinct results in both glottal and supraglottal measures in the production of different oral vowels of this language. Thus, it is suggested that clinicians and researchers consider both forms of emission for a more accurate interpretation of the implications of these data in the evaluation of oral communication and therapeutic conducts.

Keywords: Voice; Speech Production Measurement; Speech Acoustics; Phonetics

INTRODUÇÃO

Avanços tecnológicos contribuem para a ampliação dos estudos das ciências da fala. Dentre as diversas formas de avaliação, a análise acústica da fala e da voz destaca-se por ser não invasiva e de relativo baixo custo¹ perceptual and instrumental assessment of disordered voice, including overall voice quality, should ideally sample both sustained vowels and continuous speech. This investigation assessed the utility of combining both voice contexts for the purpose of auditory-perceptual ratings as well as acoustic measurement of overall voice quality. Sustained vowel and continuous speech samples from 251 subjects with ($n = 229$, o que contribui para que frequentemente seja utilizada nas pesquisas realizadas por diferentes profissionais, dentre eles, o fonoaudiólogo²⁻⁴.

É possível observar na literatura diferentes metodologias para análise dos mesmos fenômenos. Na fonoaudiologia, os parâmetros acústicos frequentemente investigados são a frequência fundamental e as frequências dos formantes de vogais³⁻⁶.

A frequência fundamental (F0) produzida pela vibração das pregas vocais e seus harmônicos serão modificados nas cavidades supraglóticas, que atuam como um filtro atenuando algumas frequências e amplificando outras. As faixas de frequências amplificadas são conhecidas como frequências dos formantes e as mais estudadas são os dois primeiros (F1 e F2) por fornecerem identidade fonética às vogais. A frequência do primeiro formante (F1) apresenta relação com a posição vertical da língua e com o grau de abertura da mandíbula, sendo seu valor inversamente proporcional à posição do complexo linguomandibular. A frequência do segundo formante (F2) é influenciada pelo deslocamento anteroposterior da língua, sendo que quanto mais anterior for a constricção da língua, maior será o valor de F2 e quanto mais posterior, menor será esta medida⁷⁻¹².

Tanto os valores de frequência fundamental quanto os das frequências dos formantes apresentam correlações com o idioma. No português, de acordo com a posição da língua no eixo vertical as vogais podem ser divididas em: baixa [a], médias baixas [ɛ] e [ɔ], médias altas [e] e [o] e altas [i] e [u]. Desta forma, formam quatro regiões de F1 sendo que com exceção da vogal [a], cada vogal anterior tem frequência de F1 similar à sua correspondente vogal posterior. No eixo anteroposterior, as vogais são classificadas de acordo com as seguintes regiões da cavidade oral: média [a], anteriores [ɛ], [e], [i] e posteriores [ɔ], [o], [u].

Como estas regiões se relacionam com medidas de F2, as vogais anteriores apresentam maiores valores de F2 e as posteriores medidas mais baixas deste parâmetro^{7,13}. (Figura 1). A altura da língua também apresenta reflexos nos valores de F0, sendo que vogais altas possuem valores mais agudos do que as mais baixas. Desta forma, no português, as vogais com maiores valores de F0 são [u] e [i]. A posição da vogal no eixo anteroposterior também influencia este parâmetro, uma vez que vogais posteriores geralmente apresentam F0 mais agudas que suas correspondentes anteriores¹³.

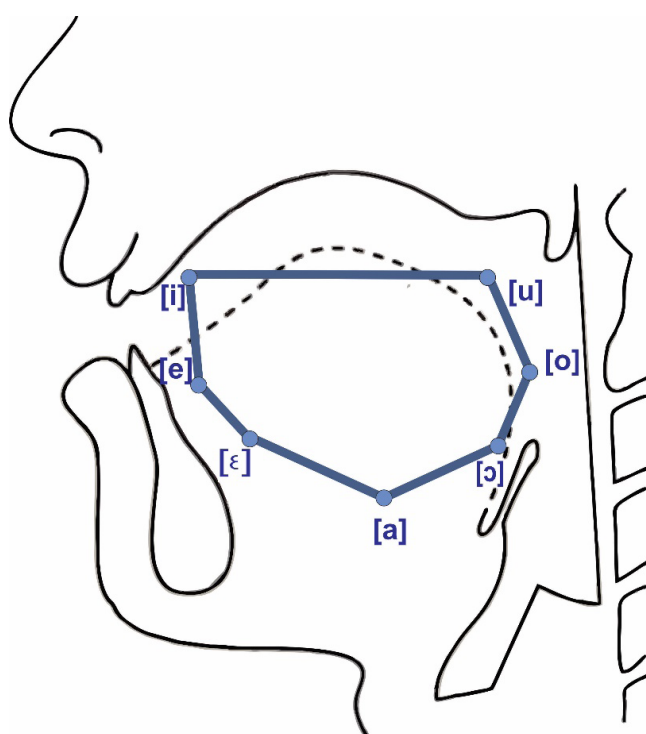


Figura 1. Representação esquemática da posição das vogais orais na cavidade oral no português brasileiro

Nestas medidas também são observadas distinções entre os gêneros, devido principalmente às diferenças anatômicas. De forma geral, como no gênero feminino, as pregas vocais e o trato vocal são mais curtos, são esperados valores mais altos de F0 e das frequências dos formantes em relação aos homens, que possuem trato vocal e pregas vocais mais longos e, portanto, frequências mais baixas¹⁴.

A frequência fundamental é o parâmetro mais robusto para os estudos da voz e as frequências dos formantes são essenciais para a identificação das

vogais e permitem interpretações articulatórias de dados acústicos⁷.

A proposição de estudar diferenças entre parâmetros acústicos de duas formas de emissão (frases-veículo e vogais sustentadas) foi baseada no fato de que estas são as tarefas de fala mais utilizadas em pesquisas e na clínica fonoaudiológica, como referido em uma recente revisão sistemática sobre formantes e produção da voz e fala⁶.

Estudos que investigaram as diferenças entre fala encadeada e vogais sustentadas concentraram-se na percepção do grau de disфонia^{1,15-17}, sendo, portanto, um tema pouco abordado nos indivíduos com vozes sem alterações. Apenas uma das pesquisas encontradas analisou parâmetros glóticos de pessoas saudáveis em duas tarefas de fala no português brasileiro¹⁸. No entanto, não foram encontrados dados de frequências dos formantes com este mesmo delineamento.

Desta forma, a comparação das medidas de F0 e dos dois primeiros formantes nestas tarefas de fala em pessoas sem distúrbios articulatórios e vocais se mostra relevante, sobretudo no português brasileiro devido à carência que existe na literatura. É importante que todas as vogais sejam analisadas devido às distinções na posição dos articuladores para a produção delas. A caracterização de tais aspectos irá colaborar com o conhecimento mais refinado das variantes de produção de fala e poderá auxiliar o trabalho do fonoaudiólogo tanto na área clínica como no aperfeiçoamento da comunicação oral uma vez que diferentes resultados podem ser encontrados de acordo com a forma de emissão escolhida para avaliação dos clientes.

Neste sentido, o objetivo deste estudo foi comparar a frequência fundamental e frequências dos dois primeiros formantes (F1 e F2) de todas as sete vogais orais do português brasileiro (PB) entre as emissões em vogais sustentadas e em fala encadeada com a utilização de frases-veículo em pessoas sem disфонia e sem distúrbios de fala.

MÉTODOS

Trata-se de estudo do tipo transversal observacional descritivo, no qual os participantes foram divididos em dois grupos de acordo com os gêneros. O projeto foi aprovado pelos Comitês de Ética em Pesquisa das instituições envolvidas: Faculdade de Medicina da Universidade de São Paulo (Parecer no. 1.540.289/2016) e do Hospital Universitário Antônio

Pedro da Universidade Federal Fluminense (Parecer no. 1.585.551/2016) e todos os participantes assinaram o Termo de Consentimento Livre e Esclarecido.

Seleção dos Participantes

Para compor a amostra deste trabalho foram incluídas 80 pessoas pareadas por gênero entre 18 e 40 anos (homens: = 23,3 anos, DP=2,71; mulheres: = 22,2 anos, DP= 2,66).

Os participantes foram entrevistados pela primeira autora da pesquisa e responderam a um questionário com dados pessoais e questões relacionadas às condições de saúde. Em seguida foram avaliados pelos ortodontistas coautores deste estudo e passaram por avaliação fonoaudiológica das estruturas orofaciais e perceptivo-auditiva da voz e fala.

Os critérios de inclusão foram: não possuir histórico de distúrbios respiratórios, auditivos, vocais ou de fala, não ser tabagista, ser falante nativo do português brasileiro da cidade do Rio de Janeiro, possuir normoclusão ou Classe I de Angle com relação maxilo-mandibular equilibrada nas 3 dimensões do espaço, perfil harmônico e pequenas alterações de posicionamento dentário. A inclusão de participantes com Classe I de Angle foi adotada uma vez que pacientes com oclusão normal, portanto, sem alterações dento-esqueléticas são raros. Os participantes deveriam apresentar *scores* correspondentes a nota 4 (normal) na avaliação das estruturas orofaciais avaliadas com uso do Protocolo AMIOFE-E¹⁹, pontuação zero no grau geral de disфонia (G) de acordo com a escala de avaliação perceptivo-auditiva GRBAS²⁰, ressonância equilibrada e não apresentar distúrbios de fala. Os fatores de exclusão foram: presença de mordida aberta, mordida cruzada anterior ou posterior, ausências dentárias ou presença de dentes supranumerários. Os participantes que relataram a presença de resfriado ou processos alérgicos no dia da coleta das amostras de fala, ou que por algum motivo não conseguiram realizar as emissões de forma adequada, foram excluídos da amostra.

Gravação dos sinais de fala e processamento dos sinais

A gravação dos sinais de fala seguiu metodologia testada em pesquisas anteriores^{3,5,21}.

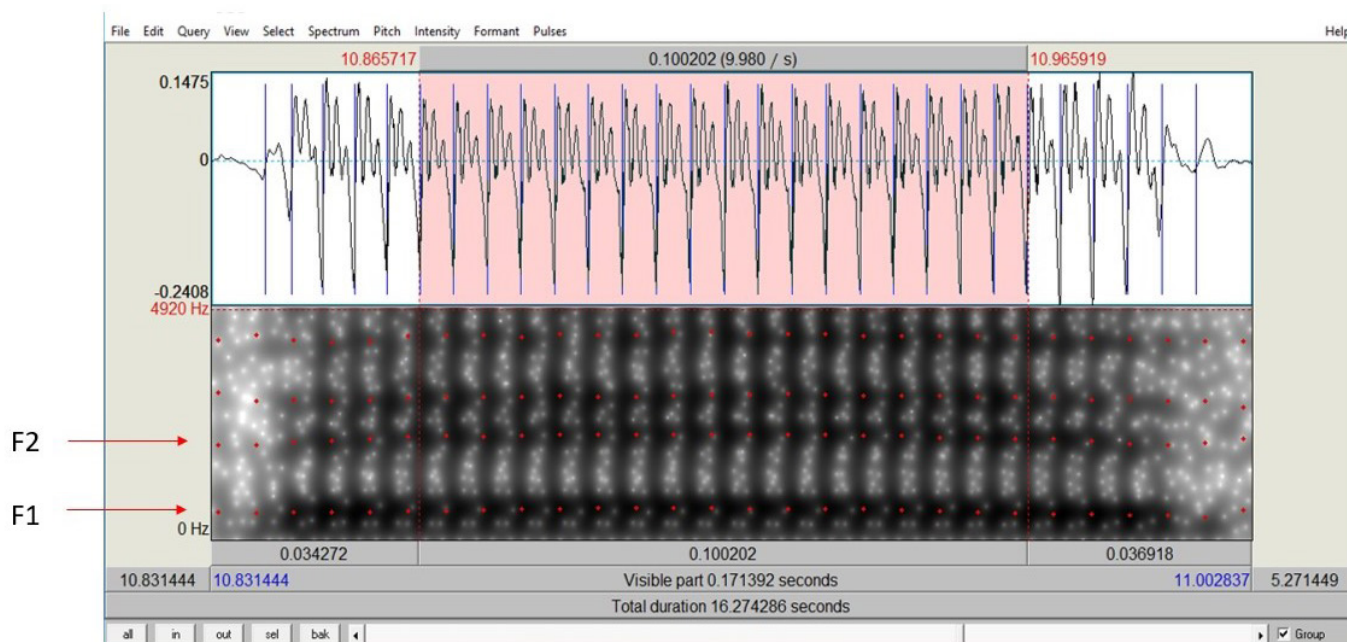
Para estimação da frequência fundamental e das frequências dos formantes os sinais de fala foram obtidos a partir de: a) gravações de frases-veículo:

“Fale ___ para mim”, preenchidas com os vocábulos “pápa”, “pépe”, “pêpe”, “pípi”, “pópo”, “pôpo” e “púpu” e, b) em emissão prolongada das sete vogais orais do português brasileiro (PB) por três segundos. Os participantes leram as instruções e emitiram tarefas com *pitch* e *loudness* confortáveis. Cada tarefa de fala foi repetida quatro vezes e foram selecionadas as duas emissões com melhor definição do traçado dos formantes.

As gravações ocorreram em uma sala silenciosa, com utilização do *software* Praat, versão 6.0.16 (P. Boesma e D. Weenink, Universidade de Amsterdã, Holanda, gratuito, disponível em <http://www.fon.hum.uva.nl/praat/>), em mono canal, com uma taxa de

amostragem de 22.050Hz e em formato “wav”. Foram utilizados um notebook marca HP (Hewlett-Packard, EUA), com sistema operacional Windows 10, e um microfone da marca Shure, modelo SM 58 (Shure, EUA), posicionado a distância de 10 cm dos lábios do indivíduo.

Foram identificados os trechos com melhor definição no traçado dos formantes de duas emissões de cada tarefa a partir dos traçados LPC sobrepostos em espectrogramas de banda larga. Foram recortados manualmente os dez milissegundos da porção intermediária de cada vogal para estimação dos valores (Figura 2).



Legenda: F1= Frequência do primeiro formante; F2= Frequência do segundo formante

Figura 2. Exemplo de seleção dos dez milissegundos da porção intermediária da vogal [E] a partir de espectrograma de banda larga com auxílio do software Praat

Após o recorte, cada segmento foi salvo em arquivo com extensão .wav. Para o processamento digital dos sinais foi empregado um *script* criado com o *software* Praat e testado em pesquisas anteriores^{3,5,22}. As medidas foram obtidas de duas amostras de frases-veículos (FV) e vogais sustentadas (VS) de cada vogal para todos os participantes. Desta forma, foram coletados 3.360 valores paramétricos, compostos por três parâmetros (F0, F1 e F2) X sete vogais X duas amostras X 80 indivíduos. Todos os

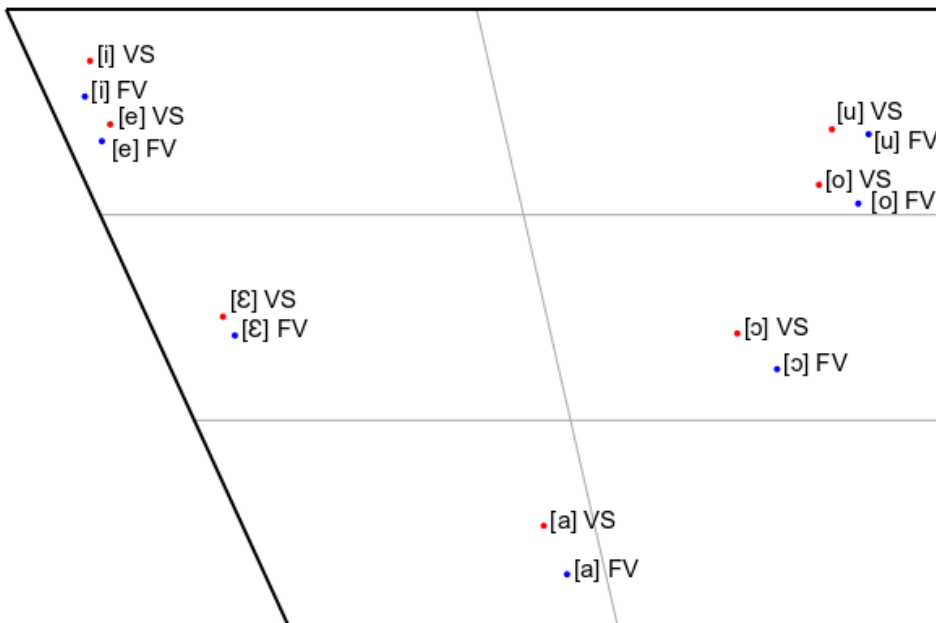
recortes das vogais foram realizados pela mesma investigadora.

Os valores obtidos com o *script* foram revistos em três momentos diferentes para garantir que as mensurações estavam corretas. Assim, a primeira pesquisadora fez uma conferência manual das medidas e outra autora verificou as frequências por meio do *script* e conferiu os valores manualmente. Nos casos em que houve divergência entre as estimações automática e manual foram consideradas as medidas obtidas de forma manual. Estes procedimentos foram adotados

para evitar erros de estimação, principalmente nas vogais posteriores, que, por apresentarem proximidades das frequências dos dois primeiros formantes¹², são passíveis de ocorrência de erros de estimação automática pelos programas.

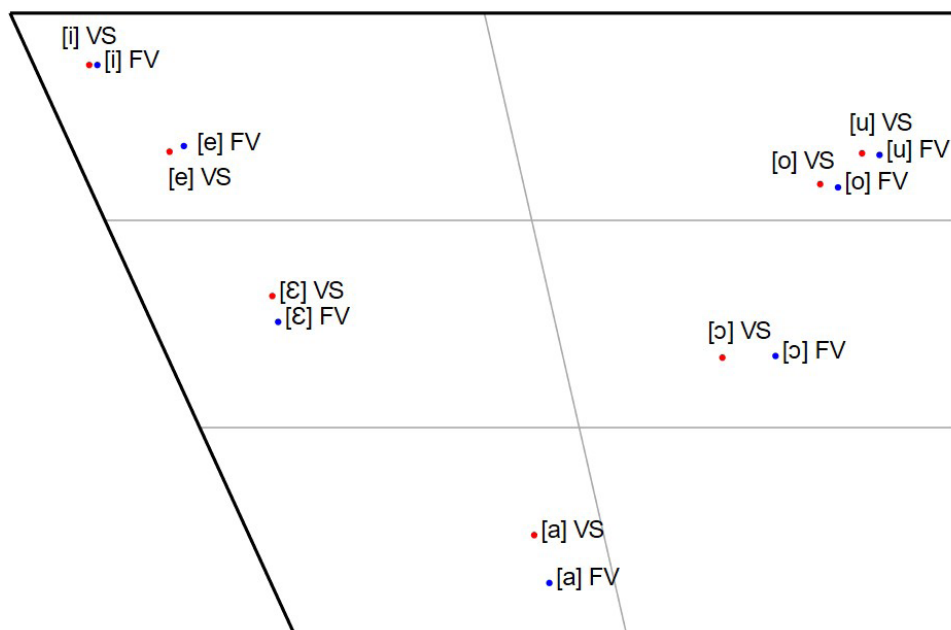
Após esta etapa foram calculadas as médias entre as duas emissões de VS e FV de cada participante,

portanto, o valor final de cada medida correspondeu à média das duas emissões em cada modalidade. As frequências dos formantes de cada gênero foram plotadas por meio de programa disponível em <http://www.adambaker.org/formant-chart/formant-chart.html> (Figuras 3 e 4).



Legenda: FV= Frase-veículo; VS= Vogal sustentada

Figura 3. Plotagem de F1 e F2 nas tarefas de vogal sustentada e frase-veículo no grupo masculino



Legenda: FV= Frase-veículo; VS= Vogal sustentada

Figura 4. Plotagem de F1 e F2 nas tarefas de vogal sustentada e frase-veículo no grupo feminino

Análise estatística

A análise estatística foi realizada com uso do Programa *Statistical Package for Social Sciences* para Windows (SPSS®, Inc. Chicago, Illinois, Chicago) e consideraram-se as medidas de tendência central média, mediana e desvio-padrão.

Para verificar a normalidade da distribuição dos dados foi utilizado o teste não paramétrico Kolmogorov-Smirnov e foram observadas evidências que as variáveis apresentavam distribuição normal.

Para comparação das medidas de F0, F1 e F2 entre as duas formas de emissão pesquisadas foi utilizado o teste t pareado. O nível de significância adotado para rejeição da hipótese nula (frequências nas duas formas de emissão eram iguais) foi igual ou menor que 0,05 (5%). A hipótese alternativa foi que existiriam diferenças entre as duas formas de emissão analisadas.

Calculou-se ainda o *Effect size* que consiste em um importante complemento do teste de significância estatística. O objetivo foi verificar o grau em que o fenômeno esteve presente na população

estudada, sendo que quanto maior o seu valor maior foi a presença do fenômeno. Os valores de ES são considerados pequenos ($0,20 \leq d < 0,50$), médios ($0,50 \leq d < 0,80$) ou grandes ($d \geq 0,80$)²².

RESULTADOS

Observou-se diferença entre as médias das duas formas de emissão (FV e VS) tanto da F0 quanto das frequências dos dois primeiros formantes em diversas vogais orais do português brasileiro.

No gênero masculino foram encontrados valores mais agudos de F0 em duas vogais e valores mais altos de F1 em seis vogais na emissão em FV. Os valores de F2 apresentaram-se mais baixos na emissão em FV em quatro vogais. Neste grupo o maior valor de *Effect size* foi encontrado em F1 da vogal [i] (Tabela 1).

No gênero feminino foram observados valores mais graves de F0 em cinco vogais, mais altos de F1 em duas vogais, além de valores mais baixos de F2 em três vogais na emissão em FV. (Tabela 2).

Tabela 1. Valores descritivos e tratamento estatístico de frequência fundamental e frequências dos dois primeiros formantes nas emissões de frase-veículo e vogal sustentada no gênero masculino

Parâmetros	MASCULINO (n = 40)							<i>Effect size</i>
	Frase-Veículo (FV)			Vogal sustentada (VS)			Teste t	
	Média (Hz)	Mediana (Hz)	Desvio-padrão	Média (Hz)	Mediana (Hz)	Desvio-padrão	p valor	
F0 [a]	116	114	14,12	118	115	15,01	0,128	- 0,139
F0 [ɛ]	117	115	16,49	118	114	15,45	0,790	- 0,063
F0 [e]	123	120	17,80	122	119	16,29	0,258	0,059
F0 [i]	132	131	19,17	126	122	18,46	0,013*	0,322
F0 [ɔ]	119	118	15,16	120	118	16,72	0,660	- 0,063
F0 [o]	126	124	19,37	124	122	18,70	0,384	0,106
F0 [u]	136	135	22,35	127	124	18,19	<0,001*	0,447
F1 [a]	795	791	72,80	746	753	72,09	<0,001*	0,684
F1 [ɛ]	554	558	45,12	535	540	41,14	0,015*	0,445
F1 [e]	358	358	31,83	341	346	30,37	0,006*	0,553
F1 [i]	313	320	27,55	277	276	25,27	<0,001*	1,379
F1 [ɔ]	588	584	41,89	552	544	54,51	<0,001*	0,750
F1 [o]	421	418	32,11	402	394	32,59	<0,001*	0,594
F1 [u]	351	347	34,90	346	343	34,62	0,510	0,145
F2 [a]	1304	1309	83,62	1349	1346	86,79	0,009 *	- 0,534
F2 [ɛ]	1946	1952	114,86	1969	1973	134,99	0,188	- 0,185
F2 [e]	2203	2189	175,07	2187	2189	128,36	0,479	0,105
F2 [i]	2236	2251	144,33	2226	2239	134,07	0,491	0,072
F2 [ɔ]	898	916	67,17	975	969	69,93	<0,001*	- 1,137
F2 [o]	741	724	72,23	817	813	82,68	<0,001*	- 0,801
F2 [u]	721	722	63,86	792	797	64,95	<0,001*	- 1,116

Legenda: FV= Frase-veículo; VS= Vogal sustentada

* p valor < 0,05; Teste t pareado

Tabela 2. Valores descritivos e tratamento estatístico de frequência fundamental e frequências dos dois primeiros formantes nas emissões de frase-veículo e vogal sustentada no gênero feminino

Parâmetros	FEMININO (n = 40)							
	Frase-Veículo (FV)			Vogal sustentada (VS)			Teste t	
	Média (Hz)	Mediana (Hz)	Desvio-padrão	Média (Hz)	Mediana (Hz)	Desvio-padrão	p valor	Effect size
F0 [a]	192	190	16,44	207	208	18,59	<0,001*	- 0,865
F0 [ɛ]	192	191	15,07	206	202	19,78	<0,001*	- 0,806
F0 [e]	199	199	15,65	210	209	19,06	<0,001*	- 0,638
F0 [i]	210	209	23,41	218	214	24,25	0,062	- 0,339
F0 [ɔ]	198	198	16,40	207	206	19,59	<0,001*	- 0,504
F0 [o]	205	207	16,14	212	209	21,07	0,004*	- 0,377
F0 [u]	222	226	19,27	222	218	27,02	0,980	0,000
F1 [a]	945	941	98,26	886	867	127,05	0,003*	0,526
F1 [ɛ]	623	634	54,71	591	607	67,65	0,010*	0,526
F1 [e]	406	403	28,70	413	409	39,09	0,109	- 0,206
F1 [i]	306	303	36,50	306	305	37,84	0,599	1,251
F1 [ɔ]	665	662	55,59	667	655	65,39	0,221	1,302
F1 [o]	457	455	33,05	453	441	41,49	0,566	- 0,033
F1 [u]	417	420	41,92	415	417	36,71	0,900	0,051
F2 [a]	1598	1588	176,50	1637	1645	201,03	0,219	-0,208
F2 [ɛ]	2297	2286	174,80	2312	2288	194,89	0,452	1,310
F2 [e]	2540	2555	231,79	2577	2590	139,52	0,324	- 0,082
F2 [i]	2763	2767	127,73	2784	2806	127,48	0,190	- 0,195
F2 [ɔ]	1015	1019	68,17	1152	1108	74,54	<0,001*	- 1,942
F2 [o]	854	847	65,25	900	891	61,09	<0,001*	- 0,737
F2 [u]	747	759	92,96	816	837	90,31	<0,001*	- 0,762

Legenda: FV= Frase-veículo; VS= Vogal sustentada

* p valor < 0,05; Teste t pareado

DISCUSSÃO

Neste estudo foram comparadas as médias da frequência fundamental e frequências dos dois primeiros formantes nas emissões em frase-veículo e vogal sustentada em pessoas sem disfonia e sem transtornos de fala. Após pesquisa bibliográfica foi observada carência de estudos que analisassem as diferenças entre estas duas tarefas de fala em diferentes vogais no português brasileiro em indivíduos vocalmente saudáveis, o que limita a comparação com os presentes resultados.

Frequência fundamental

Ao analisar os valores de frequência fundamental, observou-se que medidas mais agudas nas FV foram encontradas apenas nas vogais altas anterior [i] e posterior [u] no grupo masculino, com valores de *Effect Size* considerados pequenos. Assim, uma hipótese para tais achados foi que estes resultados podem ter

sido favorecidos por elevação do complexo hioideo laríngeo durante o processo de coarticulação presente nesta forma de emissão. A presença de diferenças apenas nas vogais altas pode apresentar correlação com a simetria entre altura da língua e medidas de F0 presentes no português¹³. Na literatura foram encontrados dois estudos que compararam medidas da vogal [a] entre vogais sustentadas e fala encadeada. Em um deles, foram observadas médias de F0 próximas nas vozes masculinas, tendência essa que também foi observada na presente investigação nesta vogal. No entanto, outro trabalho observou redução da F0 na emissão sustentada na vogal [a] em relação à emitida por meio de leitura de texto neste gênero²³.

No grupo feminino, medidas mais agudas foram encontradas nas emissões sustentadas em cinco vogais, com tamanho do efeito médio em quatro delas, sendo, portanto, o parâmetro que mais se diferenciou neste gênero. Assim como nos homens também foi observada uma simetria entre altura da língua e

medidas de F0 com diferenças estatísticas. Desta forma, as vogais média-baixa [ɛ] e média-alta [e] anteriores apresentaram tendência semelhante às suas correspondentes posteriores, vogais média-baixa [ɔ] e média-alta [o]. Igualmente, medida de F0 mais elevada na VS foi observada na vogal [a], que por ser central, não apresenta correspondência com outra vogal no eixo anteroposterior¹³.

Uma vez que essas diferenças foram notadas na maioria das vogais no grupo feminino na tarefa de fala mais próxima da situação de comunicação usual, foi hipotetizado que ao se basear em aspectos fisiológicos que indicam que valores mais altos de F0 podem demonstrar maior elevação do complexo hioideo laríngeo e maior velocidade de vibração das pregas vocais, estes achados podem auxiliar no entendimento de algumas implicações clínicas a partir da observação de qual tarefa de fala foi usada na avaliação. Assim, partindo do pressuposto que as vogais sustentadas são historicamente a forma de emissão mais investigada e mais utilizada na clínica fonoaudiológica¹, e que os presentes resultados demonstraram valores mais agudos de F0 em VS na maioria das vogais em mulheres sem disfonia, provavelmente esse ajuste muscular possa se refletir no aumento do grau de disfonia relatado em pesquisas¹⁵⁻¹⁷ tendo em vista que fisiologicamente, sons mais agudos requerem maior refinamento muscular para sua produção.

Uma hipótese para valores mais elevados de F0 nas VS na maioria das vogais produzidas pelas mulheres, seria que como esse tipo de emissão não faz parte de um contexto comunicativo usual¹⁸, há maior probabilidade de interferência do falante²³. Valores mais agudos de F0 na vogal sustentada [a] em relação à fala encadeada também foram observados em três grupos etários de mulheres em estudo que investigou vozes normais, porém as diferenças foram sutis¹⁸. A mesma tendência de aumento da F0 na vogal sustentada em relação à fala encadeada de vozes femininas foi observada em outro estudo²³.

Outra possível justificativa para simetria encontrada entre vogais anteriores e posteriores com distinções estatísticas pode ser baseada na correlação que existe entre posição da constrição da língua no eixo vertical e frequência fundamental¹³. Assim, mesmo diante da avaliação de diferentes tarefas de emissão, os valores de frequência fundamental seguiram tendências semelhantes de acordo com a altura da vogal em ambas formas de emissão.

As diferenças encontradas entre as formas de emissão também encontram respaldo em outros trabalhos^{24,25} que relacionaram a mudança de tom a aspectos articulatórios e destacaram a interação entre fonte glótica e filtro. Ao analisar a relação entre o controle de frequência e articulação das vogais, uma pesquisa²⁴ relatou que as mudanças encontradas na frequência fundamental, além de serem oriundas da musculatura intrínseca tensora da laringe puderam, em parte, também decorrer de movimentos dos músculos geni-hioideo, geni-glosso e osso hioide. Os músculos extrínsecos da língua e da laringe influenciam direta e indiretamente a posição do complexo hioideo-laríngeo e a configuração intra-laríngea.

Frequências dos formantes

No grupo masculino, as medidas que mais se diferenciaram entre as tarefas analisadas foram as frequências do primeiro formante. Os resultados demonstraram valores mais elevados nas FV em seis vogais, com tamanho do efeito médio em quatro delas (Figura 3). No grupo feminino, a mesma tendência foi observada, porém apenas nas vogais [a] e [ɛ] com tamanho do efeito médio (Figura 4). Desta forma, ao se realizar uma correspondência acústico-articulatória é possível inferir que a língua estava em posição mais baixa e a mandíbula em posição mais aberta, além de ter ocorrido maior estreitamento da faringe nas frases-veículos em relação às vogais sustentadas⁷. Uma hipótese para a diferenciação encontrada entre as duas tarefas de fala seria a interferência do fenômeno da coarticulação presente na fala encadeada, uma vez que um determinado segmento influencia segmentos adjacentes, ou seja, na vogal analisada estão contidas pistas acústicas da consoante que a precede^{7,12,26}. Nas mulheres, as diferenças encontradas apenas nas vogais [a] e [ɛ] podem ter ocorrido devido à posição da altura da língua inerente à produção destas, ou seja, língua baixa e com mandíbula aberta na vogal [a] e língua média baixa anterior na vogal [ɛ].

Os valores de F2 apresentaram medidas mais baixas nas FV em todas as vogais posteriores ([ɔ], [o], [u]) em ambos os gêneros, além da vogal [a] no grupo masculino (Figuras 3 e 4). A partir da observação destes dados, é possível inferir por meio de uma correspondência acústico-articulatória que a posição de constrição da língua foi mais posterior e a conformação da faringe foi mais estreita⁷ do que nas vogais sustentadas. Uma hipótese para uma redução nos valores de F2 seria uma maior interferência da

coarticulação presente na fala encadeada nestas vogais, uma vez que os movimentos dos articuladores para a produção de um som irão se modificar em função dos sons próximos^{7,12,26}. Os valores de F2 mais baixos nas FV em todas as vogais posteriores em ambos os gêneros podem ter sido favorecidos pela constrição posterior da língua inerente à produção destas. E, no gênero masculino, ao se analisar a vogal [a], foi possível observar abaixamento do complexo oromandibular pelo aumento de F1 nas frases-veículo, o que provavelmente pode ter contribuído para constrição mais posterior da língua, colaborando para redução da frequência de F2 nesta vogal.

O fato de a produção de algumas vogais não apresentar diferenças estatísticas entre as tarefas de fala encontra respaldo na literatura, que destaca que embora diferentes formas de emissão possam usar ajustes musculares diferentes para a produção das mesmas vogais, adaptações nos articuladores podem ocorrer e desta forma, não produzir tantas diferenças em especial nas frequências dos formantes¹⁴.

Os valores das medidas acústicas apresentadas neste trabalho representam as médias da população estudada, de acordo com metodologia testada em outros estudos^{3,5,21}, não havendo a intenção de propor parâmetros de normalidade.

As tendências observadas de elevação da f0 em algumas vogais nos homens e agravamento em diversas vogais nas mulheres, assim como elevação dos valores de F1 e abaixamento das medidas de F2 nas vogais posteriores nas frases-veículos podem ser complementadas com outros estudos, por isso sugere-se que mais pesquisas com este escopo possam ser desenvolvidas para ampliar as informações sobre essas diferenças no português brasileiro.

Limitações

Embora o presente estudo tenha apresentado contribuições sobre diferenças que podem ser encontradas em parâmetros acústicos de acordo com a tarefa de fala avaliada em pessoas sem distúrbios de voz e fala, algumas limitações devem ser reconhecidas. Primeiro, examinou-se apenas as duas tarefas de fala mais utilizadas em pesquisas e na clínica fonoaudiológica, no entanto, não foram consideradas outras formas de emissão. Segundo, embora outros pesquisadores possam, de acordo com seus objetivos, se concentrar na avaliação de medidas extraídas de fala espontânea ou semi-espontânea com trechos mais complexos, optou-se por utilizar as frases-veículos

mais referidas na literatura, o que pode no futuro permitir a comparação de dados com outras pesquisas que utilizaram o mesmo corpus.

Contribuições

Os resultados desta pesquisa podem auxiliar o trabalho dos fonoaudiólogos que atuam tanto na clínica quanto nas pesquisas e no aperfeiçoamento de fala e voz ao reforçar os pressupostos de que diferentes tarefas de emissão podem produzir medidas acústicas distintas. Desta forma, ressalta-se a importância de que o fonoaudiólogo considere mais de uma forma de emissão na avaliação dos seus clientes e que isso possa auxiliar no direcionamento do seu trabalho de acordo com os objetivos terapêuticos para cada caso. Deve-se ressaltar ainda que nenhuma medida isolada na clínica fonoaudiológica é suficiente para definir conduta, contudo o conjunto de informações auxilia o clínico a tomar melhores decisões.

CONCLUSÃO

Houve diferença entre as medidas de frequência fundamental e frequências de F1 e F2 entre as duas tarefas de fala. No grupo masculino as vogais [i] e [u] apresentaram-se mais agudas e no grupo feminino as vogais [a], [ɛ], [e], [ɔ], [o] mais graves nas frases-veículos. As frequências de F1 se distinguiram nas [a], [ɛ], [e], [i], [ɔ] e [o] nos homens e [a] e [ɛ] nas mulheres, sendo mais altas na FV. E, as frequências de F2 demonstraram valores mais baixos em todas as vogais posteriores em ambos os gêneros, além da vogal [a] no grupo masculino. Portanto, conclui-se que a tarefa de fala para avaliação de frequência fundamental e frequências dos formantes no português brasileiro pode demonstrar resultados distintos tanto em medidas glóticas como supraglótica na produção das diferentes vogais orais deste idioma. Desta forma, sugere-se que os clínicos e pesquisadores considerem ambas formas de emissão para interpretação mais apurada das implicações destes dados na avaliação da comunicação oral e no direcionamento de condutas terapêuticas.

AGRADECIMENTOS

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) por ter financiado parcialmente a realização deste estudo (Código de Financiamento 001).

REFERÊNCIAS

- Maryn Y, Corthals P, Van Cauwenberge P, Roy N, De Bodt M. Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels. *J Voice*. 2010;24(5):540-55.
- Lima MFB, Camargo ZA, Ferreira LP, Madureira S. Qualidade vocal e formantes das vogais de falantes adultos da cidade de João Pessoa. *Rev. CEFAC*. 2007;9(1):99-109.
- Viegas F, Viegas D, Baeck HE. Frequency measurement of vowel formants produced by Brazilian children aged between 4 and 8 years. *J Voice*. 2015;29(3):292-8.
- Braga JN, Oliveira DSF, Sampaio TMM. Frequência fundamental da voz de crianças. *Rev. CEFAC*. 2009;11(1):119-26.
- Viegas F, Viegas D, Atherino CCT, Baeck HE. Frequência fundamental das 7 vogais orais do português em vozes de crianças. *Rev. CEFAC*. 2010;12(4):563-70.
- França FP, Evangelista DS, Lopes LW. Revisão sistemática sobre os formantes e a produção da voz e fala. *Rev Prolíngua*. 2017;12(1):2-16.
- Barbosa PA, Madureira S. Manual de fonética acústica experimental. São Paulo: Cortez Editora; 2015.
- Stevens KN, House AS. An acoustical theory of vowel production and some of its implications. *J Speech Hear Res*. 1961;4(4):303-20.
- Johnson K. Acoustic and auditory phonetics. 3rd ed. Oxford: Wiley-Blackwell; 2012.
- Fant G. Acoustic theory of speech production. The Hague: Mouton; 1960.
- Lehiste I. Suprasegmentals. Cambridge: MIT Press, 1970.
- Kent RD, Read C. Análise acústica da fala. 1a. ed. São Paulo: Cortez Editora; 2015.
- Escudero P, Boersma P, Rauber AS, Bion RAH. A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *J Acoust Soc Am*. 2009;126(3):1379-93.
- Sundberg J. Ciência da voz: fatos sobre a voz na fala e no canto. 2a. ed. São Paulo: EdUSP; 2018.
- Maryn Y, Roy N. Sustained vowels and continuous speech in the auditory-perceptual evaluation of dysphonia severity. *J Soc Bras Fonoaudiol*. 2012;24(2):107-12.
- Lu FL, Matteson S. Speech tasks and interrater reliability in perceptual voice evaluation. *J Voice*. 2014;28(6):725-32.
- Gerratt BR, Kreiman J, Garellek M. Comparing measures of voice quality from sustained phonation and continuous speech. *JSLHR*. 2016;59(5):994-1001.
- Spazzapan EA, Cardoso VM, Fabron EMG, Berti LC, Brasolotto AG, Marino VC de C. Acoustic characteristics of healthy voices of adults: from young to middle age. *CoDAS*. 2018;30(5):1-7.
- De Felício CM, Folha GA, Ferreira CLP, Medeiros APM. Expanded protocol of orofacial myofunctional evaluation with scores: validity and reliability. *Int J Pediatr. Otorhinolaryngol*. 2010;74(11):1230-9.
- Isshiki N, Okamura H, Tanabe M, Morimoto M. Differential diagnosis of hoarseness. *Folia Phoniatr. Logoapédica*. 1969;21(1):9-19.
- Viegas D, Viegas F, Atherino CCT, Baeck H. Parâmetros espectrais da voz em crianças respiradoras orais. *Rev. CEFAC*. 2010;12(5):820-30.
- Loureiro LMJ, Gameiro MGH. Interpretação crítica dos resultados estatísticos: para lá da significância estatística. *Rev Enferm Ref*. 2011;3(3):151-62.
- Moon KR, Chung SM, Park HS, Kim HS. Materials of acoustic analysis: sustained vowel versus sentence. *J Voice*. 2012;26(5):563-5.
- Honda K. Relationship between pitch control and vowel articulation. Haskins Laboratories Status Report on Speech Research. 1983;73:269-82.
- Shaw JA, Chen W, Proctor MI, Derrick D. In luences of tone on vowel articulation in mandarin chinese. *JSLHR*. 2016;59(6):S1566-74.
- Kent RD, Vorperian HK. Static measurements of vowel formant frequencies and bandwidths: a review. *J Commun Disord*. 2018;74:74-97.