

MENEGON, Paulo Sérgio; FONTES, Mário A. S.; MADUREIRA, Sandra. Análise de interpretações orais de um poema quanto aos aspectos expressivos. *Revista Intercâmbio*, v.XLIX: 113-137. 2021. São Paulo: LAEL/PUCSP. ISSN 2237-759X

<https://doi.org/10.23925/2237.759X.2021V49.56747>

ANÁLISE DE INTERPRETAÇÕES ORAIS DE UM POEMA QUANTO AOS ASPECTOS EXPRESSIVOS

THE ANALYSIS OF ORAL INTERPRETATIONS OF A POEM ACCORDING TO ITS EXPRESSIVE ASPECTS

Paulo Sérgio MENEGON
(Faculdades Metropolitanas Unidas – FMU)
plsgmen@gmail.com

Mário A. S. FONTES
(Pontifícia Universidade Católica – PUC/SP)
fontes@pucsp.br

Sandra MADUREIRA
(Pontifícia Universidade Católica – PUC/SP)
sandra.madureira.liaac@gmail.com

RESUMO: A investigação da declamação suscita interesse por sua carga expressiva. O propósito deste trabalho é investigar, por meio de um experimento fonético, os efeitos impressionantes causados nos ouvintes por declamações do Soneto da Fidelidade. Na análise perceptiva, foram avaliados 4 descritores semânticos (agradabilidade, impacto emocional, projeção e interpretação vocais) com base em uma escala Likert. Como estímulos para a avaliação dos ouvintes foram utilizadas gravações de 8 locutores. O teste foi aplicado a 38 juízes. A análise acústica compreendeu parâmetros extraídos automaticamente. Os resultados são confrontados por meio de análise estatística multidimensional e o impacto impressionante das declamações é discutido.

PALAVRAS-CHAVE: expressividade da fala; análise perceptiva; análise acústica.

ABSTRACT: The investigation of the reciting speech style raises interest due to its expressive load. The objective of this work is investigating, by means of a phonetic experiment, the impressive effects caused on listeners by the "Sonnet on Faithfulness". For the perceptual analysis, 4 semantic descriptors (agreeability, emotional impact, vocal projection, and interpretation) were judged according to a Likert scale. Recordings of 8 speakers were used as stimuli. The test was applied to 38 judges. The acoustic analysis comprised parameters extracted automatically. The results were confronted by means of multidimensional statistical analysis and the impressive impact of the readings were discussed.

KEYWORDS: speech expressivity; perceptual analysis; acoustic analysis.

1. Introdução

A expressividade da linguagem oral se manifesta nos variados estilos de fala. Entre as formas de linguagem oral, a declamação suscita interesse na investigação pela sua função na comunicação e a sua capacidade de transmitir ao outro, estados afetivos tais como, tristezas, alegrias, dores, angústias e aflições, portanto a declamação é uma forma de expressão social (Lomba, Fontes e Madureira, 2016).

A finalidade do presente trabalho é investigar, por meio de um experimento fonético, os efeitos impressivos causados nos ouvintes ao escutarem declamações do poema de autoria de Vinicius de Moraes, intitulado "Soneto de Fidelidade", o qual integra a obra "Poemas, Sonetos e Baladas", também conhecida como "O Encontro do Cotidiano". O "Soneto de Fidelidade" foi publicado em 1946, e recebeu gravações de inúmeros intérpretes, inclusive do próprio Vinicius que o declamou em um show no meio da canção "Eu Sei que Vou Te Amar".

O experimento compreendeu dois tipos de análise: perceptiva e acústica. No experimento perceptivo desenvolvido foram avaliados 4 descritores semânticos (agradabilidade, impacto emocional, emissão vocal e interpretação vocal) em 3 graus de força, do mais fraco ao mais forte, dentro de uma escala Likert. Como estímulos para a avaliação dos ouvintes, os juízes do teste de percepção, foram utilizadas gravações do "Soneto de Fidelidade, por 8 locutores, 4 homens e 4 mulheres. A análise acústica compreendeu os parâmetros do SG ExpressionEvaluator (BARBOSA, 2009), script aplicável no software de livre acesso PRAAT, desenvolvido por Paul Boersma e David Weenink da Universidade de Amsterdam.

No poema Soneto de Fidelidade, Vinicius de Moraes remete aos sentimentos de amor e fidelidade em um relacionamento. Esse poema foi um dos publicados no livro Poemas, Sonetos e Baladas pelo referido autor em 1939 quando de sua permanência na freguesia do Estoril em Portugal.

Nascido no Rio de Janeiro em 19 de outubro de 1913, Vinicius de Moraes morreu na madrugada de 9 de julho de 1980, aos 67 anos. Apesar de ser conhecido por tantos sucessos musicais por meio de parcerias com Tom Jobim, Toquinho, Chico Buarque, entre outros, a sua especialidade é a poesia. Seu primeiro poema publicado foi "A Transfiguração da Montanha" em 1932.

Muitos de seus poemas foram transformados em letras de canções, entre elas, podemos citar: 'Eu sei que vou te amar', 'Garota de Ipanema', 'Tarde em Itapoã' e a 'Rosa de Hiroxima'. Esta última, lançada em 1973, com música composta por Gerson Conrad e

interpretação de Ney Matogrosso que integravam o grupo "Secos & Molhados. O início de sua trajetória como escritor de letras de canções se deu no fim da década de 1920, com a composição "Loira ou Morena", peça musical de Haroldo Tapajós e gravação dos irmãos Tapajós.

Uma das chancelas para legitimar uma obra e um talento criador é o impacto causado no ouvinte. Um dos mais conhecidos sonetos de Vinicius é o "Soneto de Fidelidade". Seu impacto pode ser verificado na quantidade de interpretações que recebeu por locutores profissionais e não profissionais. O talento do autor, por sua vez, é reconhecido por seus pares, entre os quais o poeta Carlos Drummond de Andrade, que foi muito amigo de Vinicius e demonstrava grande admiração pelo "poetinha" como era conhecido Vinicius. A admiração entre Carlos Drummond e Vinicius de Moraes era mútua. Ao se referir a Vinicius, Carlos Drummond o descreve como o "único poeta brasileiro que ousou viver sob o signo da paixão", enquanto Vinicius retratou o amigo "Considero Drummond o único poeta brasileiro de caráter universal".

Na declamação de poemas, dependendo dos elementos prosódicos empregados pelos locutores para veicular os sentidos produzidos a partir de suas interpretações do conteúdo do texto, ocorrerão impressões diferenciadas por parte dos ouvintes. De acordo com interpretação dada, variações de qualidade de voz e de dinâmica vocal (padrões entoacionais, de taxas de elocução, rítmicos, pausas e padrões de acentuação), que produzem consequências acústicas passíveis de serem mensuradas.

Dada a importância dos elementos prosódicos na expressão dos sentidos na fala, consideramos, a seguir, a sua caracterização, destacando os componentes de melodicidade, temporalidade, proeminência e dinamicidade.

2. Prosódia

O termo prosódia se origina do grego "prosódia", sendo que "pros", significa junto, e "odé" significa canto, o que confere à prosódia o sentido de "a melodia que acompanha o discurso". Na fala, a prosódia compreende elementos que segundo Barbosa (2019, p. 37) têm o papel de estruturar os enunciados:

"A prosódia é o componente de nossa fala que organiza nossos enunciados, moldando nossa maneira de falar através do concurso de modificações articulatórias que se manifestam acusticamente em unidades prosódicas. Essas unidades organizam a nossa fala em níveis que vão da sílaba ao enunciado entoacional".

Quando realizamos um estudo prosódico da fala, dois tipos de análise são de interesse: a acústica que nos fornece mensuração de parâmetros de duração, frequência e intensidade e a perceptiva que

nos possibilita avaliar de oitiva sensações auditivas correlatas e julgamentos de efeitos de sentido.

2.1 Os elementos prosódicos e seus correlatos acústicos e perceptivos

Os elementos prosódicos compreendem a sílaba, a acentuação, a entoação, a qualidade e a dinâmica da voz, o ritmo, a pausa, e a taxa de elocução. Na construção da expressividade da fala, têm especial relevância as variações entoacionais, a taxa de elocução, as pausas e a qualidade e a dinâmica da voz, como apontado nos trabalhos desenvolvidos por Weiss et al (2020), Barbosa e Madureira (2016 e 2018), Barbosa, Madureira e Mareuil (2017), Fontes e Madureira (2019), Madureira (2016 e 2020). Esses elementos apresentam correlatos acústicos e perceptivos, conforme apontamos a seguir.

2.1.1 A entoação

A entoação é definida por Bolinger (1986) como um gesto laríngeo, com uma compreensão centrada no sistema de produção da fala, um sistema apto a veicular informações linguísticas, paralinguísticas e extralinguísticas.

O principal correlato acústico da entoação é a frequência fundamental, a qual corresponde, em nível perceptivo, ao *pitch*, isto é, à sensação auditiva de variação entre grave e agudo referente à altura tonal. Modulações de intensidade e duração têm menor importância do que a variação de frequência fundamental na caracterização dos padrões entoacionais.

Neste trabalho foi adotada a definição de entoação de Barbosa (2019), a qual remete à variação de grave a agudo ao longo dos enunciados, ou seja, ao correlato perceptivo da entoação, o *pitch*, que é a sensação auditiva que leva à categorização da variação entre grave e agudo.

2.1.2 Sobre a taxa de elocução e as pausas

A taxa de elocução da fala se refere ao aspecto temporal. Do ponto de vista perceptivo, corresponde à variação percebida pelo ouvinte, em termos de rápida a lenta. Em termos acústicos se refere à duração dos enunciados na fala. O cálculo da taxa de elocução, diferentemente do da taxa de articulação inclui as pausas silenciosas, as quais podem ser empregadas com valor expressivo.

2.1.3 Sobre a qualidade de voz

O conceito de qualidade de voz segundo Laver (1979, 1980, 2000) é fundamentado em um modelo fonético descritivo que adota como unidade analítica o ajuste (setting). Os ajustes de qualidade de voz são descritos a partir de uma referência, o ajuste neutro, e se caracterizam por serem de curto ou longo termo. (ABERCROMBIE, 1967; LAVER, 1978, 2000).

Nesta perspectiva do trabalho de Laver, o atributo fonético da qualidade vocal, é o resultado de dois tipos de aspectos vocais, designados intrínsecos e extrínsecos, com a função de comunicação e informação sobre algo do falante (MACKENZIE-BECK, 2005). Os intrínsecos concernentes à qualidade auditiva e oriundo da anatomia constante do aparelho fonador do falante. E os extrínsecos, são provenientes da qualidade auditiva dos ajustes musculares de longo termo das estruturas do aparelho fonador.

No módulo fonético descritivo de qualidade de voz proposto por Laver (1980), dois princípios são considerados: o da susceptibilidade e o da compatibilidade. Esses princípios são relevantes de se considerar na identificação dos ajustes.

A relação entre ajustes e segmentos deve ser observada na identificação dos ajustes, quando os ajustes apresentam atributos não compartilhados pelo segmento, este último faz-se mais susceptível à influência dos primeiros (BONFIM et al, 2017).

O princípio da susceptibilidade pressupõe que certos segmentos fônicos são mais susceptíveis, do que outros, a determinados ajustes, por exemplo, a vogal anterior alta /i/, que se caracteriza por lábios estirados, é mais susceptível ao ajuste de protusão labial. O princípio de compatibilidade pressupõe à relação entre os ajustes de qualidade vocal, tratando das ações que são fisiologicamente compatíveis ou incompatíveis entre si, por exemplo, o ajuste de lábio estirado não é compatível com o ajuste de protusão labial.

Em relação à compatibilidade, alguns parâmetros podem ter combinações facilitadas, por ser a sua fisiologia é compatível (laringe abaixada e língua corporal recuada; ou laringe alta e lábios estirados), e dentro destas e outras configurações, um ou mais ajustes podem ser combinados. A combinação de ajustes de qualidade e de dinâmica de voz conferem expressividade à fala.

3. Sobre a expressividade na fala

A expressividade da fala remete aos efeitos de sentido que são veiculados a partir da materialidade fônica. Os ouvintes ao ouvirem os falantes atribuem a estas características psicológicas, físicas e sociais. São as relações motivadas entre som e sentido que impressionam os ouvintes.

A expressividade na fala está presente na comunicação cotidiana, na declamação e no canto. Por meio da fala, efeitos de sentido são expressos. Entre as formas de expressão veiculadas, temos, por exemplo, entre outros, os sentimentos de amor, hostilidade, frustração, desespero, culpa, ciúmes e as emoções de medo, alegria, tristeza, raiva, surpresa, desprezo e nojo.

As análises acústica e perceptiva são modalidades de investigação da expressividade de fala utilizadas. Como nos diz Madureira (2004), "O apoio na inspeção acústica do sinal de fala é, sem dúvida, um grande meio facilitador da escuta atenta." Investigar a expressividade de fala proporcionou diversos estudos e avanços, como a proposição do conceito de metáfora sonora (FÓNAGY, 1983) e simbolismo sonoro Hinton e Ohala (1994).

Fónagy (1983) conceitua o termo "metáfora sonora" em menção à relação de semelhança entre os ajustes motores (respiratórios, fonatórios e articulatórios) e suas consequências acústicas na expressão de efeitos de sentido. Os efeitos de sentido são pertinentes a sentidos e estímulos de múltipla natureza que sejam conferidos a partir das características do material fônico. Desse modo, se estabelece uma associação entre os gestos vocais e a expressão de sentidos.

O simbolismo sonoro diz respeito às conexões que se constituem entre as propriedades acústicas dos elementos da fala (os segmentos fônicos e a prosódia) e os efeitos de sentido percebidos pelos ouvintes. Hinton e Ohala (1994) apontam os seguintes tipos de simbolismo sonoro: simbolismo sonoro corpóreo, simbolismo sonoro imitativo, simbolismo sonoro sinestésico e simbolismo metacomunicativo. Entre os trabalhos de interesse para a presente investigação que se centra na percepção de efeitos impressivos a partir da voz em declamações de um poema, ressaltamos os de Menegon e Madureira (2013), Lomba (2013) e Madureira, Fontes e Camargo (2019). Sobre ao efeito impressivo da prosódia vocal destacamos Crochiquia et al (2020).

Crochiquia (2020) realiza uma pesquisa de natureza fonética experimental da fala expressiva com base em amostras de fala de quatro principais personagens do filme de animação intitulado "Zootopia" em dublagem realizada por atores brasileiros. A análise conjuga procedimentos metodológicos de naturezas perceptiva e acústica. O objeto dessa pesquisa é a qualidade de voz e a dinâmica vocal da fala de personagens com distintas personalidades. Os dubladores ao se defrontarem com os diferentes perfis de personalidade constroem perfis vocais que refletem características descritas nos códigos de frequência (OHALA, 1984), de esforço, de produção e sirênico (GUSSENHOVEN, 2002; 2004; 2015). Essas características vocais impressionam os ouvintes e os fazem atribuir, entre outras, características de personalidade aos personagens.

Crochiquia (op. cit) ressalta também, que no campo da expressividade vocal, como parte as produções de fala causam efeitos

impressivos nos ouvintes, que atribuem sentidos ao que ouvem. Estes efeitos impressivos causados no ouvinte com base na voz de um falante podem incluir traços de personalidade e atitude. Essas atribuições são reportadas em obras que abordam questões retóricas desde a Antiguidade (KREIMAN e SIDTIS, 2011). Assim, na fala expressiva, há uma variedade de sentidos que transmitem informações linguísticas, extralinguísticas e paralinguísticas, tornando-a um meio eficaz de comunicação.

Conforme relata Crochiquia (op.cit), a elaboração dos perfis vocais dos personagens reflete o trabalho de animação realizado pelo dublador no desempenho de seu papel como animador (Goffman, 1981). O falante, como animador, faz escolhas das variáveis segmentais e prosódicas para produzir os sentidos por ele interpretados, veiculando emoções e atitudes.

Este caminho de estudos sobre a expressividade da fala que se centra em características de simbolismo sonoro, entre elas as metáforas sonoras, tem sido trilhado no âmbito dos estudos fonéticos como apontado em Madureira, Fontes e Camargo (2019).

4. Metodologia

A metodologia de pesquisa deste trabalho é a experimental. Para a realização desta pesquisa, foi utilizado como corpus o poema "Soneto de Fidelidade" de autoria de Vinícius de Moraes.

4.1. Corpus

A escolha desse poema teve como motivação o fato de a poética de Vinicius ser caracterizada por forte carga emotiva que, em interpretações orais, suscita alterações vocais expressivas.

Segue o texto do poema:

Soneto de Fidelidade

De tudo ao meu amor serei atento
Antes, e com tal zelo, e sempre, e tanto
Que mesmo em face do maior encanto
Dele se encante mais meu pensamento.

Quero vivê-lo em cada vão momento
E em seu louvor hei de espalhar meu canto
E rir meu riso e derramar meu pranto
Ao seu pesar ou seu contentamento

E assim, quando mais tarde me procure

Quem sabe a morte, angústia de quem vive
Quem sabe a solidão, fim de quem ama

Eu possa me dizer do amor (que tive):
Que não seja imortal, posto que é chama
Mas que seja infinito enquanto dure.

A partir da coleta de 27 áudios disponíveis na internet, foram selecionadas declamações de oito locutores não profissionais da voz, 4 homens e quatro mulheres.

4.2. Procedimentos de análise de dados

O estudo compreendeu dois tipos de análise: a acústica e a perceptiva, sendo que esta última compreendeu a avaliação dos efeitos impressivos das interpretações dos locutores nos ouvintes, os juízes do teste de avaliação perceptiva e a avaliação perceptiva da qualidade e dinâmica vocais.

4.2.1 A análise acústica

Para a análise acústica foi utilizado o *script* SG - ExpressionEvaluator desenvolvido por Barbosa (2009). O ExpressionEvaluator roda no ambiente do software de livre acesso, PRAAT, desenvolvido por Paul Boersma e David Weenink da Universidade de Amsterdam. No presente trabalho foi utilizada a versão 6.2.01.

Os parâmetros extraídos automaticamente pelo ExpressionEvaluator consideraram medidas de frequência fundamental - f_0 (mediana, semi-amplitude entre quartis, assimetria e quantil 99,5%) e primeira derivada de f_0 (média, desvio padrão e assimetria), intensidade (assimetria), declínio espectral (média, desvio padrão e assimetria) e Slope LTAS (espectro de longo termo).

As medidas extraídas do script são esclarecidas, como segue:

- mediana de f_0 (frequência fundamental): medida que pode ser considerada como a f_0 habitual do falante. Sua extração prevê a suavização de valores. Trata-se de uma medida que evita erros de detecção do valor de f_0 que podem ser frequentes, sobretudo os que mudam o valor para uma oitava abaixo ou acima;

- semi-amplitude entre quartis de f_0 : medida de variação dos valores de f_0 , excluindo-se valores espúrios (por motivos de normalização, foi dividida pela metade, justificando a atribuição do termo "semi-amplitude"). O termo amplitude refere-se à variação da medida;

- quantil 99,5% de f_0 : valores de frequência fundamental normalizados por meio do cálculo de Z score da mediana (dobro),

acrescido do valor de desvio padrão. Essa medida busca, sem valores espúrios, detectar o limite superior de f_0 do falante;

- média de derivada de f_0 : representa a taxa de variação de uma função. Medida que considera não somente a variação, mas também a “velocidade” dessa alteração/variação;

- desvio padrão de derivada de f_0 : medida que infere a dispersão estatística conforme a variância de f_0 contida na amostra;

- assimetria de derivada de f_0 : infere relações de como aconteceram as mudanças de taxa de variação de f_0 na função;

- assimetria de f_0 : medida da assimetria de distribuição de f_0 , baseada na razão de diferença entre média e mediana/semi-amplitude entre quartis de f_0 . Essa medida busca estimar a tendência a ter valores à esquerda (negativos) ou à direita da média (positivos);

- assimetria de intensidade: medida de intensidade normalizada, baseada na proporção de intensidade no intervalo de frequências de 0,0-1250 Hz/1250-4000 Hz. Permite que as amostras possam ter suas intensidades estimadas, mesmo aquelas que não tenham sido captadas com fixação da distância entre boca e microfone;

- média de declínio espectral: média de valores da proporção de intensidade nos intervalos de 0-1,0 kHz/1,0-4,0 kHz. Considerada como importante medida do nível de tensão laríngea do estímulo aferido.

- Assimetria da derivada de declínio espectral: medida do desvio padrão de declínio espectral

- assimetria de declínio espectral: medida de assimetria de distribuição das medidas de declínio espectral;

- *SlopeLTAS*: medidas normalizadas de intensidade ao longo de intervalos de frequências do espectro sonoro.

4.2.2. A análise perceptivo-auditiva dos descritores semânticos

Para a confecção dos estímulos do teste perceptivo, os áudios foram editados no PRAAT.

Para a avaliação semântico-perceptiva foi utilizada uma escala Likert com 3 graus. Nesta pesquisa, os descritores escolhidos foram fundamentados em quatro aspectos vocais: agradabilidade da voz, impacto emocional da voz, emissão vocal e a qualidade da interpretação vocal.

O teste foi aplicado a 38 juízes, 26 do sexo feminino e 12 do sexo masculino na faixa etária entre 19 aos 65 anos, sendo que 35 possuíam grau de escolaridade superior, 2 com grau de escolaridade média e 1 com grau de escolaridade fundamental. Todos os juízes receberam as mesmas instruções para avaliar os efeitos impressivos causados pelas interpretações dos locutores a partir dos descritores semânticos chaves.

Ao ouvirem os estímulos, os juízes avaliaram os descritores dentro da escala Likert com graus de 1 a 3, sendo o grau 1 correspondente a pouco e o 3 a muito. Os estímulos foram apresentados em forma de áudios no formulário do Google (Google Forms).

4.2.3. Análise estatística dos dados

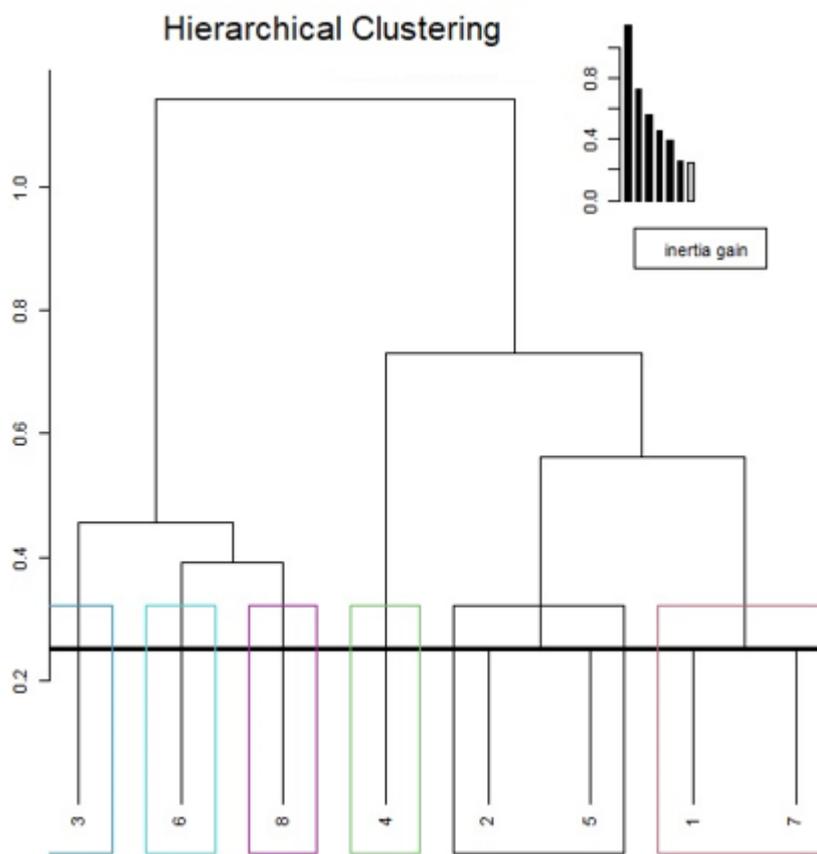
Para a análise estatística dos dados, realizada com o método estatístico MFA (Multiple Factor Analysis), foi preparada uma planilha da qual constavam: a identificação das gravações, as variáveis correspondentes aos descritores semânticos, os julgamentos dos juízes em relação aos descritores semânticos e as medidas acústicas extraídas com a aplicação do ExpressionEvaluator (Barbosa, 2009). As medidas acústicas formaram o conjunto denominado Gc1 e os valores médios dos julgamentos dos juízes o conjunto denominado Gc2.

Todas as variáveis foram normalizadas por meio da aplicação de z-score. Os dados foram exportados em arquivo .txt e analisados nos ambientes do software R e do FactoMineR (HUSSON, LÊ e PAGÉS (2009).

5. Resultados

Os resultados obtidos pela aplicação do método MFA indicaram que os estímulos compreendidos pelas 8 declamações do poema "Soneto de Fidelidade", identificadas com números pares para as vozes femininas e números ímpares para as vozes masculinas, distribuíram-se em 6 *clusters*, como se pode verificar na Figura 1.

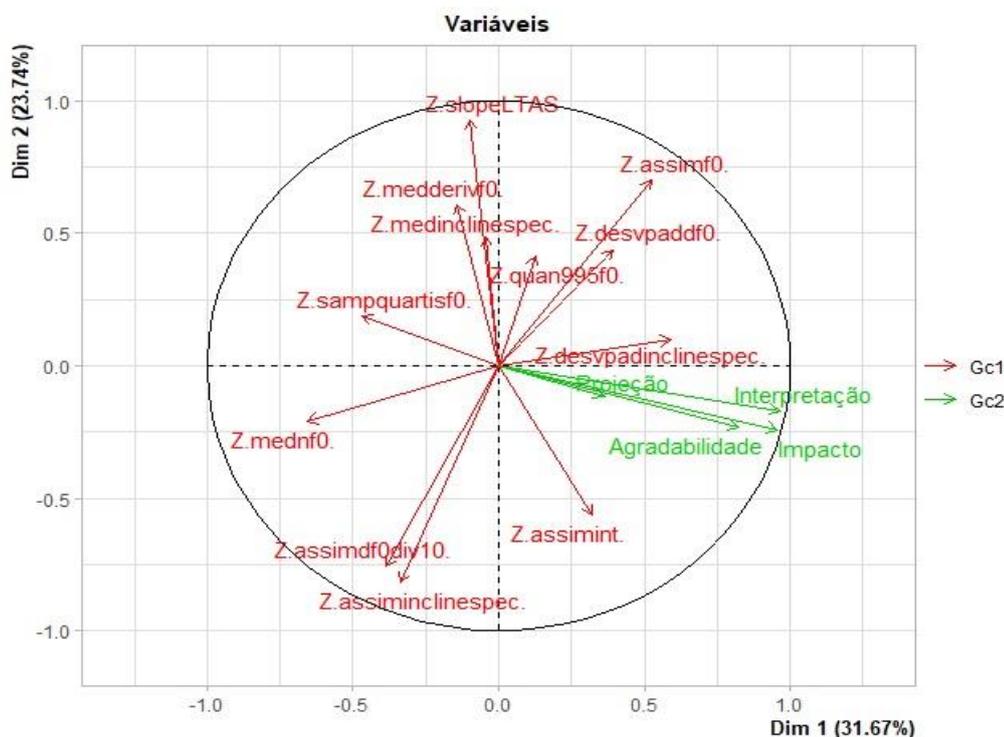
Figura 1



Clusters hierárquicos resultantes da aplicação do MFA (Multiple Factor Analysis) às análises perceptiva e acústica das declamações do poema "Soneto de Fidelidade" por 4 falantes do sexo feminino e 4 do sexo masculino.

O resultado da classificação das declamações, de maior para menor pontuação foi: 6, 8, 3, 4, 5, 7, 1 e 2. As três declamações melhor classificadas (6, 8 e 3) aparecem próximas no ramo esquerdo do gráfico de *cluster* hierárquico apresentado na Figura 1. As demais declamações aparecem em *clusters* no ramo direito do gráfico, sendo um subconjunto constituído pelas declamações que obtiveram as menores pontuações da maior para a menor 5, 7, 1, 2 e um *cluster* constituído pela declamação de número 4 que obteve a quarta classificação em termos de pontuação. A composição dos *clusters* foi guiada pela distribuição das variáveis no espaço vetorial, como pode ser observado na Figura 2. No gráfico da Figura 2, a relevância dessas variáveis pode ser visualizada pela extensão das setas: quanto mais perto da borda, mais relevantes.

Figura 2



Variáveis acústicas (Gc1) e variáveis concernentes aos descritores semânticos do teste de percepção (Gc2) nas dimensões 1 (DIM 1) e 2 (DIM2) do espaço vetorial.

Na Figura 2, as variáveis correspondentes aos descritores semânticos se distribuíram na parte inferior direita do gráfico e as acústicas, principalmente, na parte superior do gráfico. Entre as variáveis acústicas as que apresentaram valores com significância estatística na Dimensão 1 foram os descritores Interpretação, Impacto e Agradabilidade. Na dimensão 2 se destacaram: o *SlopeLTAS*, a Assimetria da derivada de declínio espectral (*Assimdf0div10*) e a Assimetria da inclinação espectral (*Assiminclinespec*).

O *SlopeLTAS* se refere ao desvio padrão das medidas normalizadas de intensidade ao longo de intervalos de frequências do espectro sonoro. O *Assimdf0div10* é a medida da assimetria de distribuição de f_0 , baseada na razão de diferença entre média e mediana/semi-amplitude entre quartis de f_0 . Essa medida busca estimar a tendência a ter valores à esquerda (negativos) ou à direita da média (positivos). O *Assiminclinespec*, se refere à medida de assimetria de distribuição das medidas da inclinação espectral.

A seguir, apresentamos a Tabela 1, com a especificação das dimensões envolvidas, as variáveis, os graus de correlação e de significância.

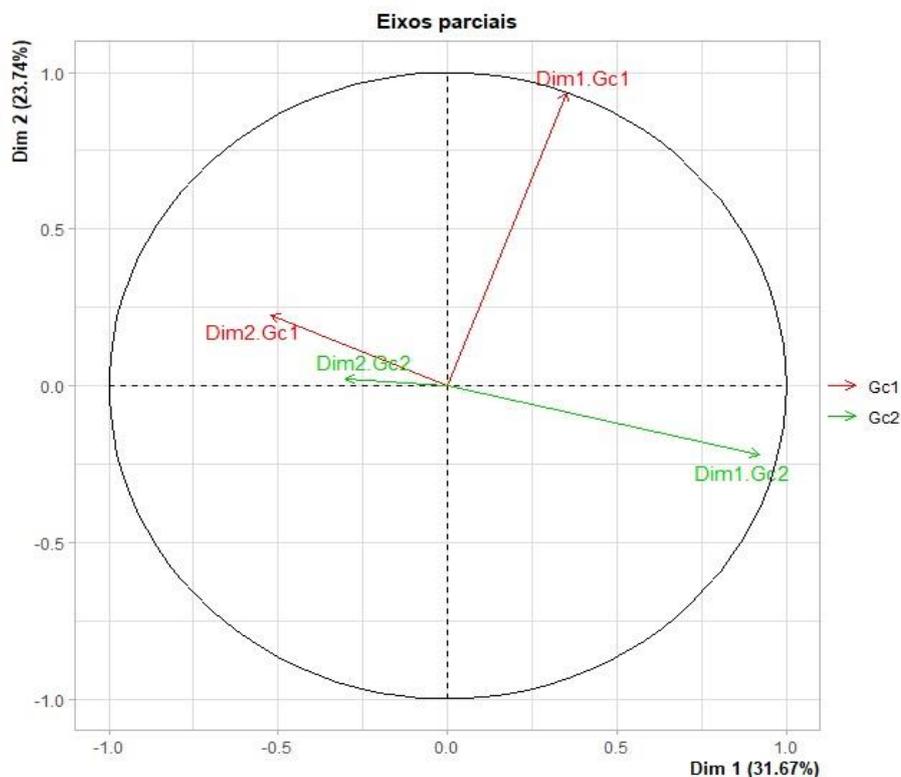
Tabela 1

| Dimensão | Variável | Grau de correlação | Grau de significância (p. value) |
|----------|-------------------------|--------------------|----------------------------------|
| 1 | Interpretação (Z.Mintr) | 96,73% | 0.0001 |
| 1 | Impacto (Z.Impcp) | 95,65% | 0.0002 |
| 1 | Agradabilidade (Z.Agrd) | 82,21% | 0.0123 |
| 2 | Slope LTAS | 92,86% | 0.0009 |
| 2 | Assimdf0div10 | - 75,65% | 0.0298 |
| 2 | Assiminclinespec | - 81,71% | 0.0133 |

Explicitação dos graus de correlação e de significância das variáveis nas dimensões 1 e 2.

A Figura 3 mostra a distribuição das variáveis acústicas (Gc1) e das semânticas (Gc2) nos eixos parciais das dimensões 1 e 2 do espaço vetorial. A dimensão 1 (Dim 1) explica 31,67 % dos dados, enquanto a dimensão 2 (Dim2) explica 23,74% dos dados.

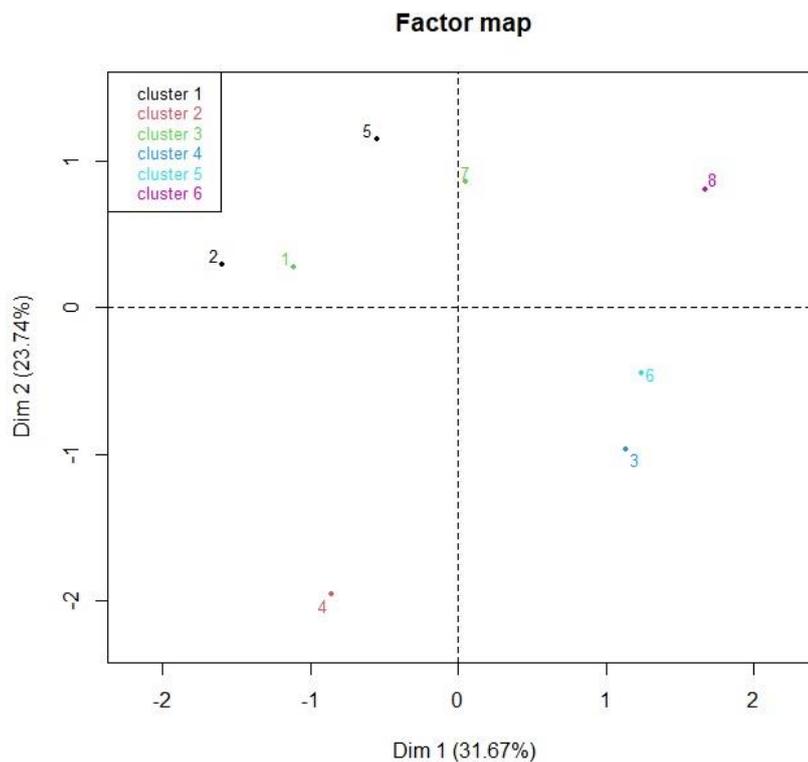
Figura 3



Distribuição das variáveis nos eixos parciais variáveis acústicas (Gc1) e das semânticas (Gc2) nos eixos parciais das dimensões 1 e 2 do espaço vetorial.

Dada a distribuição das variáveis nas dimensões referentes aos eixos parciais, pode ser captada a força que determinou o arranjo dos estímulos das declamações dos locutores, como visualizado na Figura 4, na qual cada ponto representa um estímulo, ou seja, uma declamação de cada locutor numerados de 1 a 8.

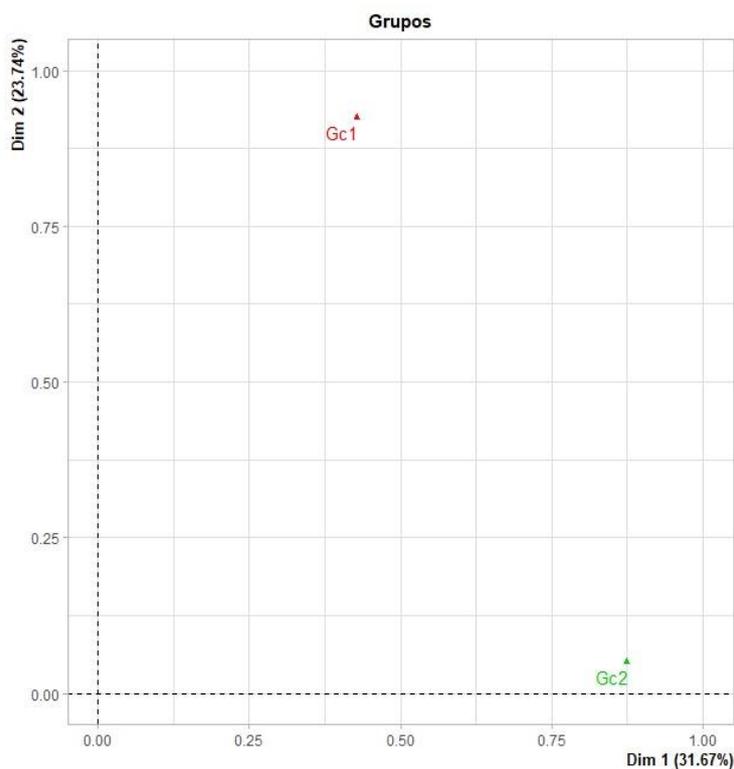
Figura 4



Distribuição dos estímulos referentes às declamações dos locutores no espaço vetorial. À direita do gráfico os locutores melhor avaliados e à esquerda os pior avaliados.

A contribuição das variáveis acústicas e semânticas foram semelhantes, como pode ser observado a partir da distância da localização dos grupos de variáveis em relação à intersecção (eixos 0:00) na Figura 5.

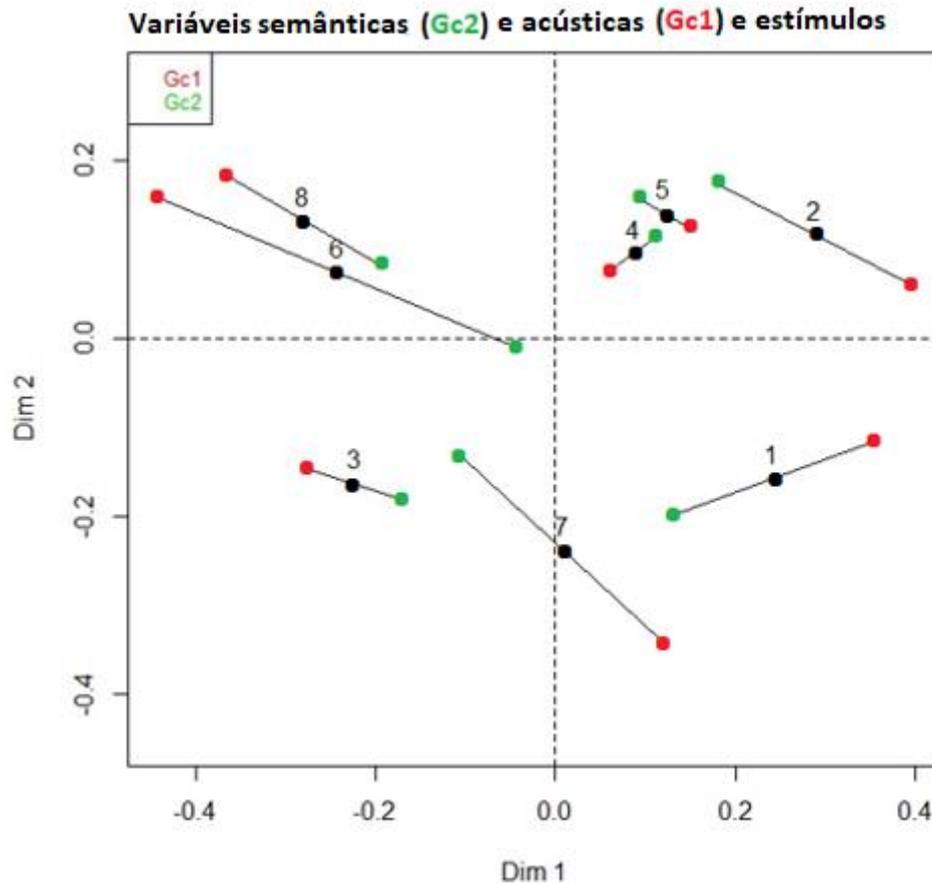
Figura 5



Contribuição do grupo de variáveis para a explanação dos dados.

Para considerarmos a influência das variáveis acústicas e semânticas na distribuição dos pontos relativos às dimensões 1 e 2 do espaço vetorial (Dim1 e Dim2) apresentamos a Figura 6 que mostra a direcionalidade dos fatores de influência e a extensão da variação.

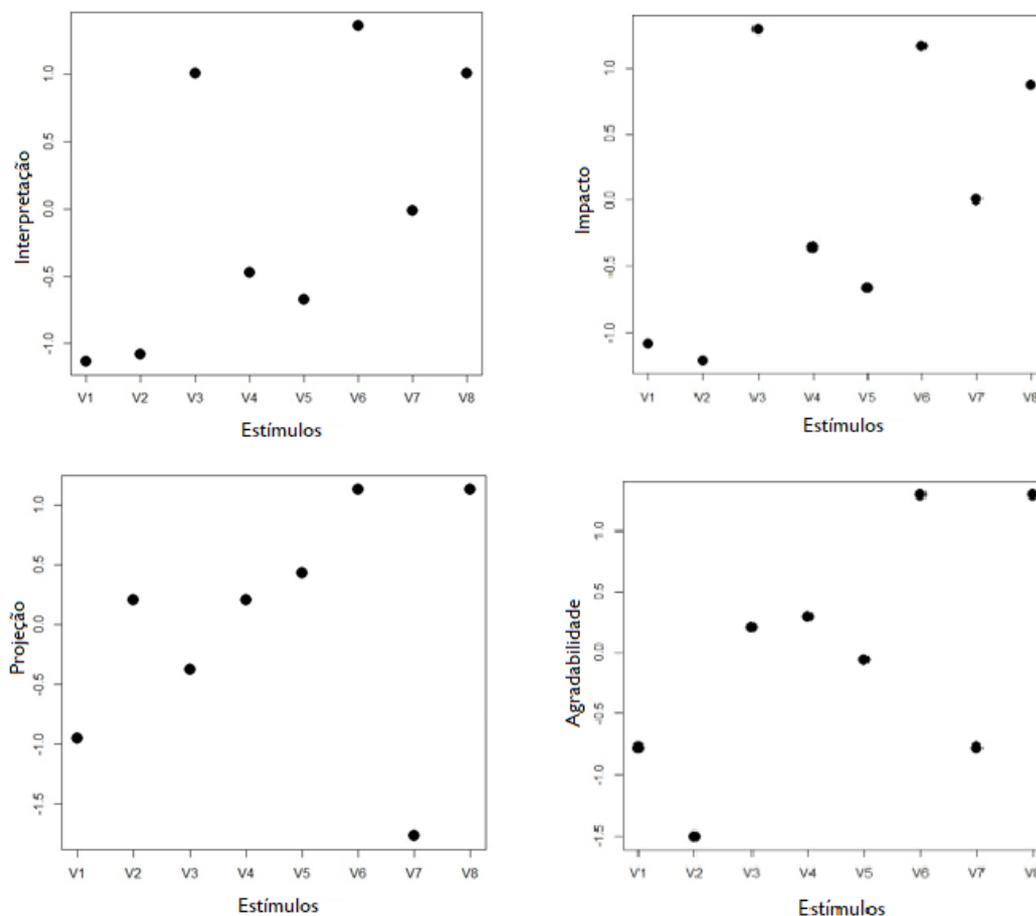
Figura 6



A influência das variáveis acústicas e semânticas sobre os estímulos correspondentes às declamações dos locutores de 1 a 8.

A Figura 7 apresenta a pontuação geral, normalizada por z-score, em relação à avaliação dos quatro descritores semânticos. Tal pontuação resulta da soma de pontos atribuídos pelos juízes às declamações dos locutores. Em referência ao descritor "interpretação", tivemos o locutor 6 como o melhor avaliado, seguido em ordem decrescente pelos locutores 8 e 3; em referência ao impacto, o melhor locutor foi considerado o de número 3, seguido em ordem decrescente pelos locutores 6 e 8; quanto à projeção, os locutores melhor avaliados foram os de número 6 e 8, e em segundo lugar pontuou o locutor 5; em relação à agradabilidade, os dois melhor avaliados foram os locutores de número 6 e 8, seguidos do locutor 4 em segundo lugar.

Figura 7



Gráficos demonstrativos da totalização das pontuações atribuídas a cada locutor em relação a cada um dos descritores semânticos do teste de percepção.

A seguir, nas Figuras de 8 a 13, podem ser observadas as correlações entre as variáveis semânticas pareadas e pode se visualizar a força das variáveis acústicas e o impacto dessas variáveis na distribuição dos estímulos.

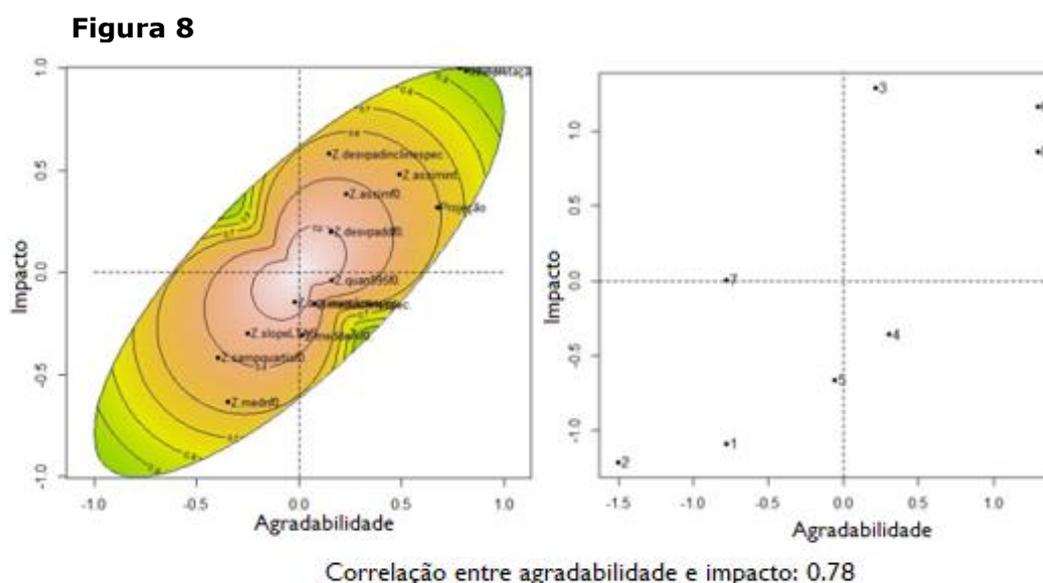
Cada uma das figuras apresenta a correlação de duas variáveis semânticas, ou seja, encontram-se pareados os seguintes descritores semânticos: impacto e agradabilidade (Figura 8); interpretação e agradabilidade (Figura 9); projeção e agradabilidade (Figura 10); interpretação e impacto (Figura 11); interpretação e projeção (Figura 12); impacto e projeção (Figura 13).

No interior de cada gráfico dessas figuras elípticas, estão dispostas as linhas de campo que explicitam a influência das medidas acústicas: quanto mais afastadas do centro, menor a força de influência. As medidas com maior relevância para a explanação dos dados se encontram sinalizadas na parte central desses gráficos.

As Figuras contêm dois gráficos. Na representação gráfica à esquerda, são apresentadas, nos eixos vertical e horizontal, as variáveis semânticas e, dentro dos gráficos de dispersão (figuras elípticas), as variáveis de influência referentes às medidas acústicas.

Na representação gráfica à direita, pode se visualizar a distribuição resultante das variáveis de influência dispostas no gráfico à esquerda. Cada ponto representa o resultado do julgamento da declamação de cada locutor.

Explicitaremos, a seguir, em relação a cada figura a influência das variáveis na classificação das declamações. Na Figura 8 se encontram pareadas as variáveis de agradabilidade e impacto.

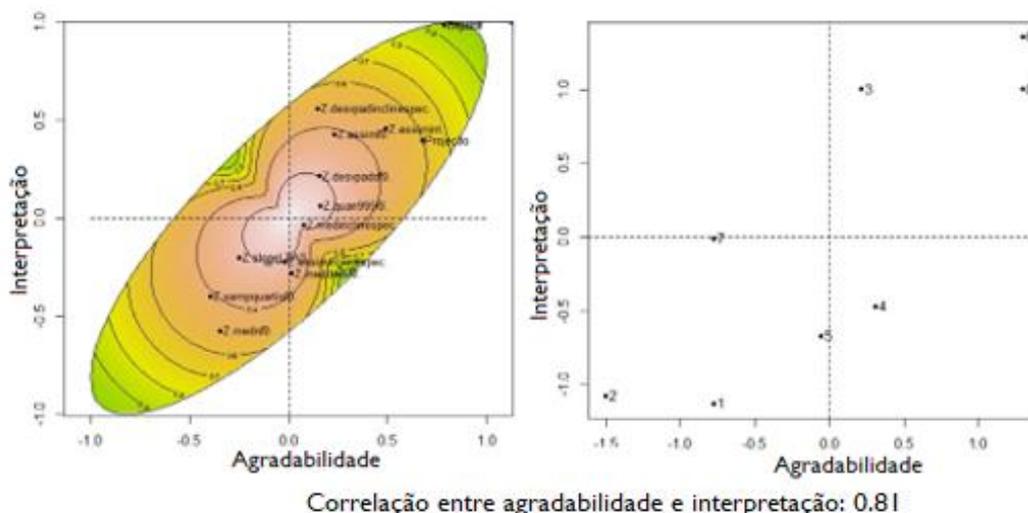


Correlação entre agradabilidade e impacto: 0,78

Pode-se visualizar pela distribuição dos pontos, que na figura 8, em relação à correlação entre os descritores de impacto vocal (eixo vertical) e agradabilidade (eixo horizontal), as declamações dos locutores 3, 6 e 8 se destacaram. Se observarmos a localização dos pontos nos eixos vertical e horizontal, percebemos que a declamação do locutor 3 foi a melhor classificada quanto ao impacto, pois ocorre na posição mais alta do eixo vertical, seguida das declamações dos locutores 6 e 8. Em termos de agradabilidade, as duas declamações mais bem avaliadas foram as dos locutores 6 e 8, pois se situam na extremidade do eixo horizontal. Na sequência, aparecem as dos locutores 4 e 3. A porcentagem de correlação entre as variáveis de impacto e agradabilidade é de 78%. As variáveis acústicas de maior relevância foram: assimetria da inclinação espectral (assiminclinespec), quantil 99,5% de f0 e desvio padrão de derivada de f0 (desvpaddf0).

Na Figura 9 se encontram pareadas as variáveis de agradabilidade e interpretação.

Figura 9

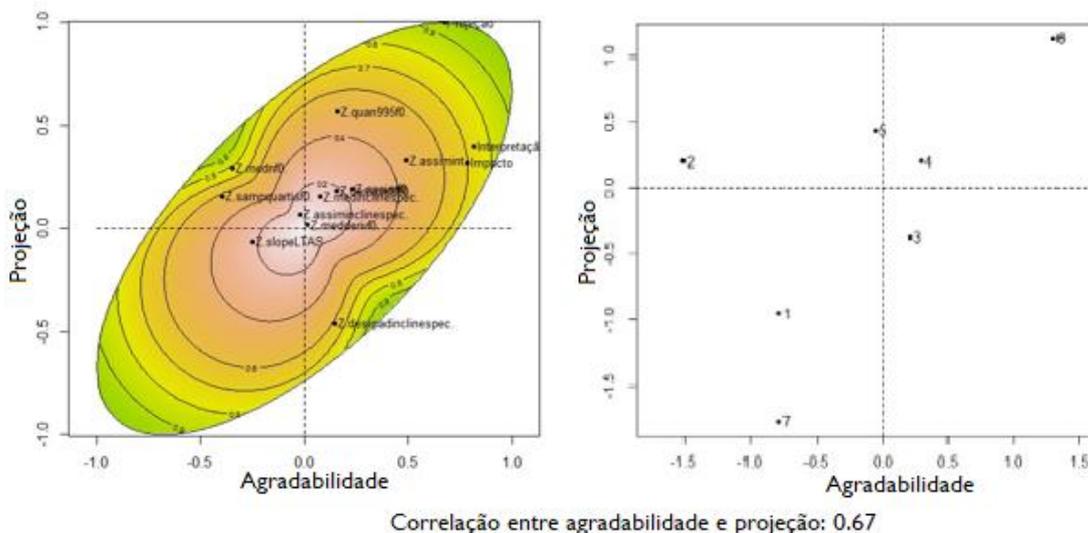


Correlação entre agradabilidade e interpretação: 0,81

Na Figura 9, na correlação entre os descritores de interpretação vocal e agradabilidade destacam-se as declamações dos locutores, 6, 8 e 3. A porcentagem de correlação é de 81%. As variáveis acústicas de maior relevância foram: média de declínio espectral (medinclinespec), quantil 99,5% de f0 e desvio padrão de derivada de f0 (desvpaddf0).

Na Figura 10 se encontram pareadas as variáveis de agradabilidade e projeção.

Figura 10

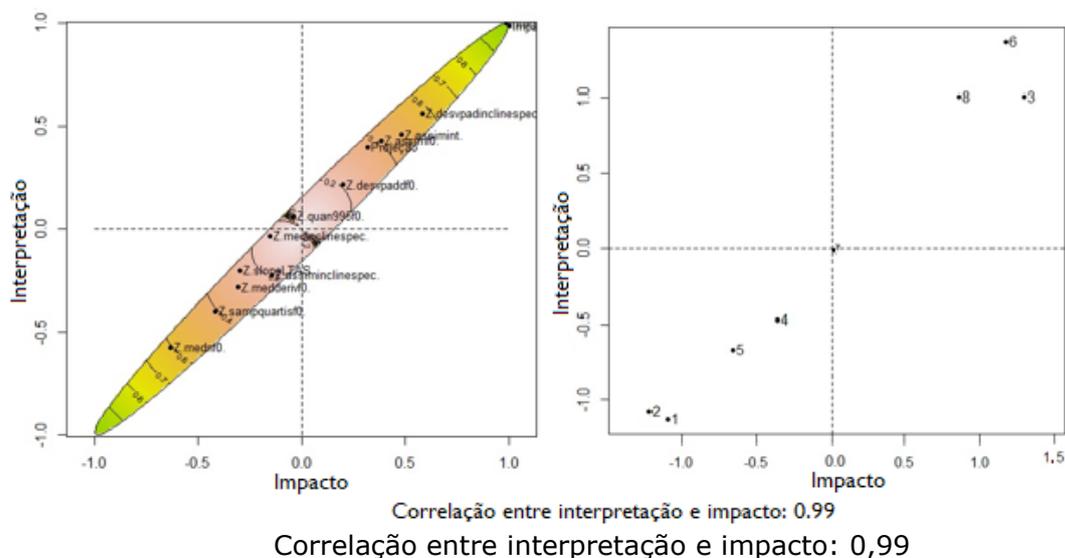


Correlação entre agradabilidade e projeção: 0,67

Na correlação entre os descritores de projeção vocal e agradabilidade mostrada na Figura 10, destacam-se as declamações dos locutores, 6, 8 e 3. A porcentagem de correlação é de 67%. As variáveis acústicas de maior relevância foram: desvio padrão de derivada de f0 (desvpaddf0) e média de declínio espectral (medinclinespec).

Na Figura 11 se encontram pareadas as variáveis de interpretação e impacto.

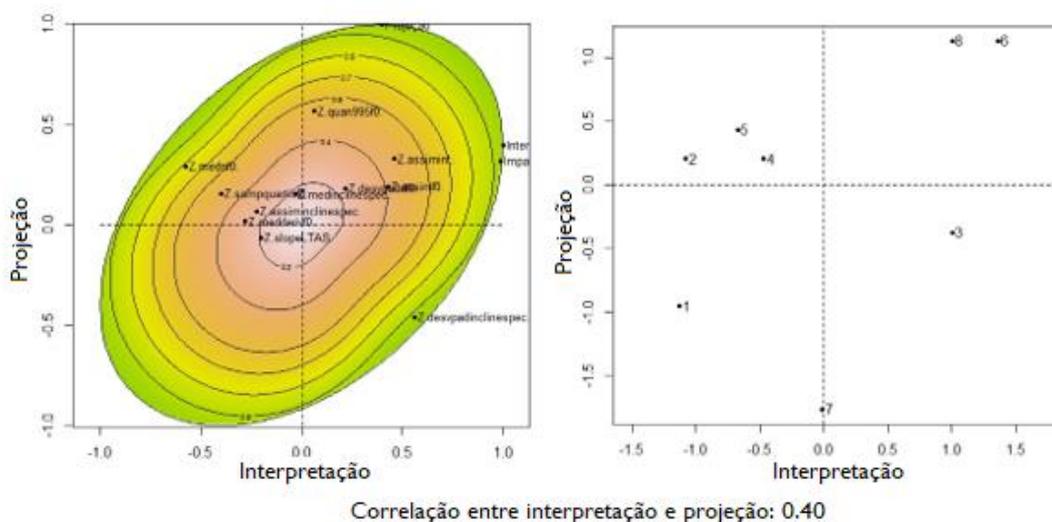
Figura 11



Na figura 11, a correlação entre os descritores de interpretação vocal e impacto é forte (99%). Com pontuações mais altas aparecem os locutores 6, 8 e 3. As variáveis acústicas de maior relevância foram: assimetria da inclinação espectral (assiminclinespec), média de declínio espectral (medinclinespec) e quantil 99,5% de f0.

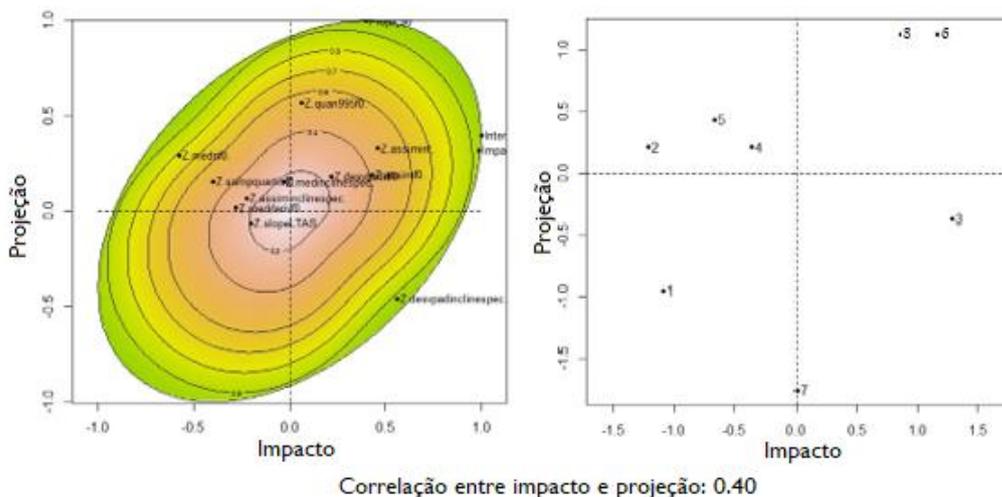
A seguir, nas Figura 12 e 13 se encontram pareadas, respectivamente, as variáveis de interpretação e projeção e de impacto e projeção.

Figura 12



Correlação entre interpretação e projeção: 0,40

Figura 13



Correlação entre impacto e projeção: 0,40

Nas figuras 12 e 13, a correlação entre os descritores de interpretação vocal e impacto com o descritor projeção se revelou fraca (40%). Os locutores com maior pontuação foram 6 e 8. As variáveis acústicas de maior relevância foram: espectro de longo termo (Slope LTAS) e média de declínio espectral (medinclinespec).

Os gráficos das Figuras de 8 a 13 permitem avaliar a força da correlação entre as variáveis. O gráfico da Figura 11 é estreitado, o que demonstra que as variáveis são mais fortemente correlacionadas (99%). Os demais são mais alargados, especialmente os de número 12 e 13 cujas porcentagens de correlação são de 40%. Os demais gráficos das figuras 8, 9 e 10 apresentam variáveis com 78%, 81% e 67 % respectivamente.

As variáveis mais fortemente correlacionadas são, portanto, a interpretação e o impacto. A interpretação e o impacto dependem das escolhas prosódicas que o locutor utiliza. A agradabilidade e a projeção referem-se mais especificamente à qualidade da voz do locutor.

6. Conclusão

A investigação da expressividade da fala, que neste trabalho buscou os efeitos impressionantes causados nos ouvintes ao escutarem declamações do poema "Soneto de Fidelidade de Vinicius de Moraes, utilizou uma escala Likert em 3 graus, de fraco a forte, para avaliação perceptiva de 4 descritores semânticos (agradabilidade, impacto emocional, emissão vocal e interpretação vocal) por 38 juízes e medidas acústicas de frequência fundamental intensidade e Espectro de Longo Termo.

A análise perceptiva dos dados revelou que dentre as variáveis que influenciaram os julgamentos das declamações a que teve maior relevância foi a interpretação vocal seguida do impacto vocal. A correlação entre esses dois descritores foi de 99%. Os descritores de agradabilidade e projeção vocal foram os que menos influenciaram nas escolhas dos juízes. A projeção teve baixa correlação com o impacto (40%).

A análise acústica indicou que a variação de intensidade ao longo de intervalos de frequências do espectro sonoro (SlopeLTAS), a medida da assimetria de distribuição de f0 (Assimdf010) e a Assimetria da Inclinação espectral (Assiminclinespec) apresentaram significância. Essas medidas estão correlacionadas com a variabilidade da fala e as características de qualidade vocal.

A declamação é um estilo que busca transmitir uma mensagem aos ouvintes e passar a emoção que o conteúdo de um texto desperta no interpretante. A interpretação é a força motriz que, com a técnica de projeção vocal comumente utilizada por profissionais da voz, impressiona o ouvinte pela expressividade da voz. O estilo de declamação, sob a ótica deste trabalho, é fruto da força motriz da interpretação.

Referências Bibliográficas

ABELIN, A. **Studies in Sound Symbolism**. Doctoral dissertation. Gothenburg: Göteborg University, 1999.

ABERCROMBIE, D. **Elements of general phonetics**. Edinburgh: Edinburgh University Press, 1967.

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

MENEGON, Paulo Sérgio; FONTES, Mário A. S.; MADUREIRA, Sandra. Análise de interpretações orais de um poema quanto aos aspectos expressivos. *Revista Intercâmbio*, v.XLIX: 113-137. 2021. São Paulo: LAEL/PUCSP. ISSN 2237-759X

BARBOSA, P. A.; MADUREIRA, S.; MAREÜIL, P. B. Cross-Linguistic Distinctions Between Professional and Non-Professional Speaking Styles. In: *Interspeech 2017*, 2017, Stockholm. Interspeech 2017. Stockholm: Editors: Wolfgang Hess (Bonn, Germany) & Martin Cooke (University of the Basque Country, Spain), 2017. v. 1. p. 3921-3925.

BARBOSA, P.; MADUREIRA, S. Elicitation techniques for cross-linguistic research on professional and non-professional speaking styles. In: *Speech Prosody 2016*, 2016, Boston. Boston, USA: Boston University, 2016. v. 1. p. 503-507.

BARBOSA, P.; MADUREIRA, S. The Interplay between Speech and Breathing across three Brazilian Portuguese Speaking Styles. Instituto de Estudos da Linguagem, University of Campinas, Brazil, LIACC, Catholic University of São Paulo, Brazil, 2018.

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

BOLINGER, D. *Intonation and its Parts: Melody in Spoken English*. London: Edward Arnold, 1986.

BONFIM, L. S. M. F.; MADUREIRA, S.; RUSILO, L. C.; CAMARGO, Z. Avaliação de qualidade vocal: abordagem metodológica para análise de dados perceptivos *Rev. CEFAC*. 2017 Nov-Dez; 19(6):831-841.

CROCHIQUIA, A.; ERIKSSON, A.; FONTES, M. A.; MADUREIRA, S. Um estudo fonético das vozes de personagens do filme *Zootopia* na dublagem em português brasileiro: o papel dos estereótipos. In: *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada*, 36(3), 2020.

FÓNAGY, I. *La vive voix: Essais de psycho-phonétique*. Paris: Payot, 1983.

FONTES, M.; MADUREIRA, S. Vocal and facial expressions and meaning effects in speech expressivity. In: *10th International Conference of Experimental Linguistics*, 2019, Lisboa. Lisboa: Exling 2019, 2019. v. 1. p. 81-84.

GOFFMAN, ERVING.. *Forms of Talk*. Filadélfia: University of Pennsylvania Press, 1981.

GUSSENHOVEN, C. Intonation and Interpretation: Phonetics and Phonology. *Proceedings of the 1st International Conference on Speech Prosody*, 2002. Aix-en-Provence, 47-57.

MENEGON, Paulo Sérgio; FONTES, Mário A. S.; MADUREIRA, Sandra. Análise de interpretações orais de um poema quanto aos aspectos expressivos. *Revista Intercâmbio*, v.XLIX: 113-137. 2021. São Paulo: LAEL/PUCSP. ISSN 2237-759X

GUSSENHOVEN, C. *The Phonology of Tone and Intonation*. Cambridge University Press, 2004.

GUSSENHOVEN, C. (ms) *Foundations of intonation meaning: anatomical and physiological factors*, Faculteit der Letteren, Radboud University Nijmegen, 2015.

FONTES, M.; MADUREIRA, S. Vocal and facial expressions and meaning effects in speech expressivity. In: 10th International Conference of Experimental Linguistics, 2019, Lisboa. Lisboa: Exling 2019, 2019. v. 1. p. 81-84.

HINTON, L.; NICHOLS, J. & OHALA, J. J. (ed.). *Sound Symbolism*. Cambridge: Cambridge University Press, 1994.

HUSSON, F., LÊ, S., et PAGÈS, J. *Analyse de Données avec R* Livre (224 pages) Édition : Presses Universitaires de Rennes – Collection Pratique de la statistique – 2009

HUSSON, F., JOSSE, J., Lê, S., MAZET, J. 2013. *FactoMineR: Multivariate Exploratory Data Analysis and Data Mining with R* (version 1.25) [R package].

KREIMAN, J.; SIDTIS, D. *Foundations of voice studies: An interdisciplinary approach to voice production and perception*. Boston: Wiley-Blackwell, 2011.

LAVÉ, J. The concept of articulatory settings: an historical survey. *Historiographia Linguistica*, 5:1-14, 1978.

LAVÉ, J. The description of voice quality in general phonetic theory. *Edinburgh University Department of Linguistics Work in Progress*. 12:3-52, 1979.

LAVÉ, J. *The phonetic description of voice quality*. Cambridge: Cambridge University Press, 1980.

LAVÉ, J. Phonetic evaluation of voice quality. *Voice quality measurement*, p. 37-48, 2000.

LOMBA, SOUZA FONTES, M. A.; MADUREIRA, S.; *Sonoridades - Sonorities*. 1. ed. São Paulo: Edição da Pontifícia Universidade Católica de São Paulo, 2016. v. 1. 185p.

MACKENZIE-BECK, J. Perceptual analysis of voice quality: the place of vocal profile analysis. *A Figure of Speech. A Festschrift for John Laver*, p. 285-322, 2005.

MENEGON, Paulo Sérgio; FONTES, Mário A. S.; MADUREIRA, Sandra. Análise de interpretações orais de um poema quanto aos aspectos expressivos. *Revista Intercâmbio*, v.XLIX: 113-137. 2021. São Paulo: LAEL/PUCSP. ISSN 2237-759X

MADUREIRA, S. Intonation and Variation. *Dialectologia*, v.VI, p. 57-74, 2016.

MADUREIRA, S. Fala e expressividade. In: Verbetes LBASS. Disponível em: <http://www.lettras.ufmg.br/lbass/> , 2020.

MADUREIRA, S.; SOUZA FONTES, M. A. & CAMARGO, Z. Sound symbolism, speech expressivity and crossmodality. *Signifiances (Signifying)*, 2019, p. 19-26. DOI [10.18145/SIGNIFIANCES.V3I1.234](https://doi.org/10.18145/SIGNIFIANCES.V3I1.234). Disponível em: <https://revues.bu.uca.fr/index.php/Signifiances/article/view/234>.

MENEGON, PAULO.; MADUREIRA, S. Metáforas no Ensino de Canto e seus Efeitos na Qualidade Vocal: um Estudo Acústico-Perceptivo, in S. Madureira (org.) *Sonoridades - Sonorities*. 2016. São Paulo: Edição da Pontifícia Universidade Católica de São Paulo.

NOBILE, L. Introduction – Sound symbolism in the age of digital orality. A perspective on language beyond «nature» and «culture». In: *Signifiances (Signifying)*, 2019. DOI [10.18145/SIGNIFIANCES.V3I1.248](https://doi.org/10.18145/SIGNIFIANCES.V3I1.248). Disponível em: <https://revues.bu.uca.fr/index.php/Signifiances/article/view/248>.

OHALA, John J.. An Ethological Perspective on Common Cross-Language Utilization of F0 of Voice. *Phonetica*, 1984. v. 41, n. 1, p.1-16.

SALOMÃO, G. L. Expressão Vocal de Emoções: Metáfora Sonora, Fala e Canto. In: Sandra Madureira (org.). *Sonoridades – Sonorities*. 1ed. São Paulo: Pontifícia Universidade Católica de São Paulo, 2016, p. 31-43.

SCHERER, K. R. What are emotions? And how can they be measured?, *Social Science Information* 44/4, 2005, p. 695–729.

WEISS, B.; Trouvain, J.; Barkat-Defradas M.; Ohala J. J.; *Voice Attractiveness: Studies on Sexy, Likable, and Charismatic Speakers*. Published by the registered company Springer Nature Singapore Pte Ltd. 2020.

Recebido em 14/12/2021
Aprovado em 01/02/2022