

# Análise da prosódia afetiva na esquizofrenia a partir do aplicativo Exprosodia

---

DOI: <http://dx.doi.org/10.21165/el.v48i3.2200>

**Ana Cristina Aparecida Jorge<sup>1</sup>**

## Resumo

Esta pesquisa teve como objetivo realizar uma análise da entoação de voz de pacientes com esquizofrenia para, a partir de parâmetros acústicos advindos de variáveis linguísticas, examinar dados que caracterizem a prosódia como um possível indicativo diagnóstico. Para a realização dessa pesquisa, 16 pessoas diagnosticadas com esquizofrenia, de acordo com os critérios do DSM V (SECMs), e mais 16 indivíduos sem histórico médico anterior de transtorno mental (SCs), concordaram em realizar um diálogo gravado. Esse procedimento se dividiu em quatro etapas: entrevista de anamnese; relato empírico; descrição de uma imagem; e, por fim, a leitura de um trecho de uma história infantil. A análise dos dados coletados foi realizada pela rotina ExProsodia (FERREIRA-NETTO, 2016). Os resultados apresentados indicam a possibilidade de diferenciar pacientes acometidos pela esquizofrenia de sujeitos sem histórico anterior de transtornos psíquicos ou doenças mentais, baseada na análise de parâmetros acústicos de voz.

**Palavras-chave:** comunicação verbal; entoação; esquizofrenia; ExProsodia; prosódia afetiva.

---

<sup>1</sup> Universidade de São Paulo (USP), São Paulo, São Paulo, Brasil; [anacristinajorge@usp.br](mailto:anacristinajorge@usp.br); <https://orcid.org/0000-0001-5517-9747>

# Analysis of affective prosody in schizophrenia from the Exprosodia application

## Abstract

The aim of this study was to perform an intonation analysis of patients with schizophrenia to analyze data that characterize prosody as a possible diagnostic indicative of acoustic parameters derived from linguistic variables. To perform this research, 16 people diagnosed with schizophrenia according to the DSM V criteria (SECMs) and 16 individuals with no previous medical history of mental disorder (SCs) agree to perform a recorded dialogue. This procedure was divided into four stages: anamnesis interview; empirical report; description of one image; and, finally, the reading of an excerpt from a children's story. The analysis of the collected data was performed by routine ExProsodia (FERREIRA-NETTO, 2016). The results presented indicate the possibility of differentiating patients affected by schizophrenia from subjects with no previous history of psychic disorders or mental illness based on the analysis of acoustic parameters of voice.

**Key-words:** ExProsodia; intonation; pitch; schizophrenia; verbal communication.

## Introdução

A voz de cada indivíduo fornece muitos elementos para a identificação da personalidade e inferência sobre o estado mental. Indica características físicas do falante, por exemplo, sexo e idade; aspectos emocionais, como sentimentos de alegria, tristeza, euforia ou cansaço (CHHABRA *et al.*, 2012). Peres (2016) argumenta que uma das componentes mais fundamentais para a compreensão e interpretação da fala é a prosódia, fenômeno que apreende múltiplas funções no contexto comunicativo, inclusive demarcação sintática.

Para Castagna *et al.* (2012), a prosódia se refere à habilidade de reconhecer, compreender, produzir sentido e significado semântico baseado na entoação, acento e padrões rítmicos. Esses autores lembram que a prosódia pode ser classificada em emocional (ou afetiva), compreendida como o reconhecimento de elementos emocionais e afetivos provindos de informações prosódicas, e a prosódia não emocional que se caracteriza pela percepção de contornos entoacionais, ao discernir, por exemplo, frases interrogativas das exclamativas.

O estudo desenvolvido por Bachorowski e Owren (1995) retratou que propriedades acústicas da língua falada estão relacionadas à intensidade com que se expressa a emoção, a qual ainda é mediada pelo contexto em que a mensagem está sendo transmitida e uma possível consequência que esse enunciado possa prover ao falante.

Ferreira-Netto (2006) corroborou essa ideia; em suas pesquisas, o autor discorre sobre parâmetros acústicos de voz humana no português falado no Brasil, relacionando-os com elementos afetivo-emocionais e a sua influência no comportamento dos falantes. Para o autor, a linguagem é uma manifestação basicamente emocional, que procede de atributos biológicos do organismo somada à necessidade contextual da informação. Esse mesmo autor explica que, na língua portuguesa, a prosódia abrange ritmo, entoação e ênfase.

A prosódia emocional de frases caracteriza-se pela variação de tons abaixo de 1KHz em pontos da frase, de maneira a estabelecer uma curva melódica bem definida que os ouvintes associam a diferentes estados emocionais. Também fenômenos prosódicos menores, como duração vocálica, intensidade, pausas atuam como marcadores prosódicos capazes de expressar emoções. Trata-se, pois, de um fenômeno complexo que necessita ser avaliado como um conjunto de manifestações linguísticas cuja finalidade é expressar algum estado emocional. (FERREIRA-NETTO, 2006, p. 26).

Para Ferreira-Netto e seus colegas (2013), existe a possibilidade de diferenciar as manifestações verbais de certos sentimentos através do estudo estatístico de variáveis linguísticas, uma vez que concorda com a hipótese de que a entoação proceda de uma série temporal, composta por um conjunto numérico sequencial e interdependente (FERREIRA-NETTO, 2006; PERES, 2016). Tais estudos demonstram com efetividade que a análise da modulação da voz humana permite identificar que, ao exprimir certos conteúdos emocionais como a tristeza, o falante emprega uma tonalidade grave e a raiva é caracterizada com tom de voz agudo (FERREIRA-NETTO *et al.*, 2013). A partir dessas pistas prosódicas, também é possível o ouvinte compreender a manifestação emocional expressa pelo falante (LIEBERMAN; MICHAELS, 1962).

O entendimento sobre a codificação e a decodificação de pistas prosódicas foi abordado de forma mais detalhada por Banse e Scherer (1996), que sustentam a possibilidade de existir padrões específicos de perfis verbais para cada manifestação emocional. Por exemplo, para a expressão da raiva há um aumento significativo na taxa e na variabilidade da frequência fundamental do mesmo modo que há elevação na articulação dos enunciados; por outro lado, a tristeza é expressa através de um decréscimo na faixa média da frequência assim como os contornos entoacionais e a variabilidade também se encontram diminuídos.

Uma análise acústica da prosódia requer a exploração atenta de quatro parâmetros: a frequência fundamental, medida que corresponde à melodia da fala, calculada em Hertz; a duração determina o tempo da elocução calculado em milissegundos; e a intensidade que diz respeito à energia vocal aplicada pelo falante (AGUIAR; MADEIRO, 2007). A esses Ferreira-Netto (2016) acrescenta o Tom Médio, que consiste na média aritmética

acumulada no tempo dos valores da frequência fundamental em uma dada elocução; esse parâmetro é mensurado em Hertz.

Pesquisas realizadas por Martins e Ferreira-Netto (2017, p. 3701, tradução nossa) apontam que em momentos de estresse intenso ou em casos de algumas psicopatologias é exibido padrão verbal "*Simulacrum of Neutral Intonation*" (SNI), definido pelos autores como "a fala monótona produzida pelos falantes com prevalência de ritmo e melodia da fala mais plana",<sup>2</sup> caracterizado pela presença de quatro parâmetros principalmente: "(i) menor valor de F0, (ii) menor dispersão positiva das variações de Foco/Ênfase, (iii) menor intervalo significativo entre as unidades da entoação, e (iv) menor mediana desses mesmos intervalos"<sup>3</sup>.

Para o *Manual Diagnóstico e Estatístico de Transtornos Mentais* (DSM V/APA, 2013), compêndio utilizado por profissionais da saúde para classificação de transtornos psíquicos e doenças mentais, a esquizofrenia se caracteriza por um misto de sinais e sintomas disformes. Todavia, para o seu diagnóstico, o indivíduo deve exibir pelo menos dois dos seguintes sintomas, sendo que pelo menos um deles deve ser um dos três primeiros: delírios, alucinações, discurso desorganizado, comportamento grosseiramente desorganizado ou catatônico e sintomas negativos, permanentes durante o período de um mês, com alguns sinais deste transtorno evidentes durante seis meses.

Esta versão do compêndio psiquiátrico contempla a inclusão da prosódia como um dos sintomas negativos desse transtorno mental, isto é, os que afetam a expressão emocional: "inclui reduções na expressão das emoções pelo rosto, pelo contato visual, na entoação da fala (prosódia) e no movimento das mãos" (DSM V, APA, 2013, p. 88), comportamentos que inferem ênfase ou facilitam o discurso. Outro sintoma negativo particular da doença é a diminuição ou perda da capacidade de sentir prazer em circunstâncias que eram anteriormente tidas como estímulos positivos ao paciente.

Trémeau (2006) defendeu que pessoas com esquizofrenia têm prejuízos nos três principais domínios emocionais: na expressão, no reconhecimento e na avaliação da experiência emocional. Na opinião do mesmo autor, apesar de esses indivíduos possuírem a mesma compreensão semântica sobre a expressão emocional, há evidências de déficits na forma de expressão, independente do canal de transmissão, seja expressão facial ou linguagem verbal. Há maior nível de interpretação da experiência subjetiva como sentimentos de valência negativos e mais alta sensibilidade a condições de estresse. Isso prejudica a sua atuação e interação social.

---

2 No original: "the monotone speech produced by the speakers with a prevalence of rhythm and a flat speech melody".

3 No original: "(i) lower F0 value, (ii) lower dispersion of positive variations of Focus/Emphasis, (iii) lower mean interval between units of bearing of intonation, and (iv) lower median of these same intervals".

## Objetivo

O objetivo dessa pesquisa é realizar uma análise da variação da entoação de voz de pacientes com esquizofrenia, visando examinar sinais prosódicos que a caracterizem como indicativo diagnóstico para a esquizofrenia.

## Procedimentos éticos e metodológicos

Inicialmente, para a realização dos procedimentos delineados por essa pesquisa foi solicitada a apreciação do Comitê de Ética em Pesquisa com Seres Humanos, através da Plataforma Brasil. O projeto de pesquisa teve os documentos encaminhados para o exame do Comitê de Ética da Faculdade de Saúde Pública da Universidade de São Paulo (FSP/USP), localizada na Avenida Doutor Arnaldo, 715; Cerqueira César; São Paulo; CEP: 01.246-904; telefone: (11) 3061-7779. A aprovação foi concedida através do Parecer nº 2.301.603; em 27/09/2017 às 17h53min. (ANEXO I). Na sequência, foi aplicada a proposta metodológica de coleta de dados na instituição de saúde mental, entre os meses de junho e julho de 2017.

Como um cuidado especial, a pesquisadora permaneceu por um período de 15 dias, antes de começar efetivamente a coleta de dados, observando a rotina da instituição de saúde mental em que os pacientes acometidos pela esquizofrenia são atendidos com o objetivo de estabelecer vínculo afetivo. Ao término dos procedimentos, o processo de desligamento da pesquisadora também foi gradual com o propósito de não despertar sentimentos negativos nos participantes. Sobretudo, é importante destacar que todo esse processo foi acompanhado pela equipe de saúde mental locada na unidade, através de reuniões diárias.

Antes da coleta, todos os participantes receberam explicações quanto aos objetivos e procedimentos a serem desenvolvidos. Foram esclarecidas possíveis dúvidas e demais questionamentos durante a aplicação da coleta de dados. Foram garantidos o sigilo e o respeito às informações coletadas. Todos os participantes, seus responsáveis legais, um profissional da equipe e a pesquisadora assinavam o Termo de Consentimento Livre e Esclarecido (TCLE). Esses documentos foram impressos em duas vias; uma delas foi entregue ao participante e a outra está sob a tutela da pesquisadora, bem como designa a Resolução 466/12 do Conselho Nacional de Saúde.

A coleta de dados foi realizada pela pesquisadora com 32 sujeitos: 16 pacientes diagnosticados com esquizofrenia (doravante, SECM) segundo os critérios do DSM V (APA, 2013) e 16 participantes sem histórico de transtorno mental que formaram o grupo de controles (doravante, SCs). Os participantes de ambos os grupos eram alfabetizados e falantes nativos do português brasileiro.

Os SECMs eram 13 homens e três mulheres, entre 25 e 65 anos. Desses, quatro eram pacientes que viviam em casas de repouso; dois residiam em Residências Terapêuticas sob a tutela da diretora geral do Centro de Atenção Integral à Saúde Professor Cantídio Moura de Campos; os demais estavam em tratamento ambulatorial e terapêutico frequente. Todos participavam ativamente das atividades promovidas pelo CAPS II “Espaço Vivo”, local no qual realizaram a coleta de dados. A esses foram emparelhados os SCs, quanto aos critérios de idade, gêneros sexuais e níveis de escolaridade.

Os SCs foram recrutados nas intermediações da Universidade de São Paulo (USP), local em que realizaram a coleta de dados. Os critérios de exclusão desse grupo foram: histórico anterior de transtornos mentais e/ou doenças psíquicas assim como o uso de qualquer medicamento psicofarmacológico.

Os medicamentos mais utilizados pelos pacientes SECMs eram: o Biperideno (anticolinérgico); a Prometazina (anti-histamínico); e o Hadol/Haloperidol (neuroléptico, utilizado no tratamento de manutenção de pacientes psicóticos crônicos estabilizados).

Em especial, nessa instituição de saúde mental, por um cuidado particular com os atendidos, os procedimentos foram realizados com o acompanhamento de um dos profissionais da equipe de técnicos multiprofissionais (psicólogos, assistentes sociais ou pertencentes à enfermagem). A única ressalva feita pela pesquisadora foi que não interrompessem enquanto o paciente falava ou respondessem por ele.

Os procedimentos metodológicos consistiram na gravação da voz dos participantes com o auxílio do gravador de voz digital portátil da marca H4 ZOOM em uma sala agradável, sem isolamento acústico, com a presença do participante e da pesquisadora para se assemelhar a um atendimento psicoterapêutico. As gravações foram realizadas individualmente e não foi permitida a reaplicação dos procedimentos.

A coleta de dados para ambos os grupos se dividiu em quatro etapas. Na primeira, foi realizada uma entrevista sucinta de anamnese, com o objetivo de coletar dados sócio demográficos da população analisada. Os participantes abordaram informações pessoais, dados relativos à vida acadêmica e laboral. No caso específico das pessoas com esquizofrenia, foi comentado sobre a rotina farmacológica e demais tratamentos efetivados na instituição de saúde mental em que estão acolhidos.

Na segunda etapa, solicitou-se ao participante que contasse uma recordação triste e uma feliz entre suas experiências de vida. Nesse caso, o intuito foi promover uma narrativa de um momento emocional mais subjetivo, o qual demandaria do participante a expressão afetiva emocional, logo maior índice de variações entoacionais. A isso se soma a possibilidade de apreender o imprevisto durante a transmissão da fala, o que, conseqüentemente, requer maior envolvimento do falante durante a enunciação.

Baseando-se no proposto por Silveira (1992), ao afirmar que pessoas com esquizofrenia têm uma grande dificuldade no uso da linguagem, concluiu-se que o trabalho através da língua oral pode ser difícil, principalmente nos casos mais graves. Com o auxílio de uma imagem, a comunicação pode se tornar mais acessível e proporcionar resultados promissores. Portanto, do mesmo modo como fez Jorge (2018) em seu trabalho, foi proposto aos participantes deste estudo que discorressem sobre a Figura 1.



**Figura 1.** Procedimento

**Fonte:** <http://www.1freewallpapers.com/road-summer-background/pt>

Na última etapa, convidamos o participante a ler em voz alta um trecho de uma história infantil sem conotação emocional. Essa recomendação está fundamentada através da pesquisa efetivada por Rapcan e outros (2010) que constatou com efetividade, através de uma tarefa de leitura, a possibilidade de diferenciar pessoas com esquizofrenia de sujeitos controle sem histórico anterior de transtorno mental. A leitura envolve uma atividade programada em que o imprevisto é unicamente na expressão prosódica. Ao envolver um treino mais intenso, torna-se uma atividade mais tensa.

Nesta pesquisa, a sugestão de texto para leitura foi uma adaptação de um trecho da história infantil *O mágico de Oz* de L. Frank Baum (2011, p. 95). Este trecho a seguir foi especialmente selecionado por ser um estrato literário emocionalmente neutro com simplicidade verbal e semântica.



*Quando todos estavam sentados em fila no alto do muro, olharam para baixo e viram uma cena muito estranha.*

*À frente deles se estendia uma vasta região em que o solo era liso, reluzente e branco como o fundo de uma travessa de louça. Espalhadas por ele, havia muitas casas, todas feitas de porcelana e pintadas das cores mais vivas. As casas eram todas bem pequenas, e a maior delas mal chegava à cintura da menina.*

*Ao lado, havia pequenos celeiros rodeados por cercas de louça, reunindo grupos de vacas, carneiros, cavalos, porcos e galinhas, todos feitos de porcelana. Mas o mais diferente de tudo eram as pessoas que viviam naquela terra estranha. Eram camponesas que tiravam leite das vacas, ou pastorinhas com vestidos de cores vivas e aventais de bordas douradas; e princesas com lindos trajes prateados, dourados e púrpura; e pastores usando suspensórios e calças curtas com listras cor-de-rosa, amarelas e azuis, e fivelas de ouro nos sapatos; e príncipes com coroas na cabeça, cravejadas de pedras preciosas, usando mantos de arminho e calções de seda; e palhaços engraçados com roupas enfeitadas por muitas rendas, com manchas vermelhas pintadas nos rostos e chapéus compridos e pontudos. E, o mais estranho de tudo, essas pessoas eram todas feitas de porcelana, inclusive as suas roupas, e eram tão pequenas que a maior delas não passava da altura do joelho da menina.*

Cada sujeito produziu um áudio diferente. O tempo das gravações variou conforme a necessidade de cada um. A pesquisadora permitiu que os participantes se expressassem livremente. Não houve resistência ou oposição dos grupos para a realização das tarefas propostas. Ambos os grupos a efetivaram sem pressão e com liberdade de expressão.

Apenas na tarefa de leitura, alguns pacientes apresentaram maior embaraço, dois não leram afirmando terem dificuldades visuais e dois realizaram a atividade de leitura, porém não foi lido o que estava disposto no papel entregue ao participante e sim, palavras aleatórias. Particularmente, esses trechos foram descartados para a análise. Para igualar aos SECMS, foram excluídos os extratos de leitura realizados por quatro dos SCs.

## **Análise dos dados**

O presente trabalho manteve as considerações de Ferreira-Netto (2016) e Martins e Ferreira-Netto (2017) para analisar a entoação de voz dos SEIs e dos SCs. Foi realizada a análise dos parâmetros acústicos referentes à entoação com o auxílio da rotina de dados ExProsodia (FERREIRA-NETTO, 2008, 2010, 2016), aplicativo desenvolvido por Ferreira-Netto (2010) em 2008, registrado no INPI (RS08992-2). Sobre ele, o seu criador explica: “o programa de pesquisa ExProsodia tem por objetivo propor uma interpretação para a relação entre a entoação e a fala” (FERREIRA-NETTO, 2010, p. 06).



A entoação é entendida como uma sucessão de tons iguais ou diferentes que são produzidos pela voz durante a produção da fala. Consiste em uma série temporal na qual atuam diversas componentes para a sua configuração periodicamente. Para decomposição dessa determinante, Ferreira-Netto (2006) cita a proposta de Xu e Wang (1997), em que a entoação é compreendida como a composição de dois elementos principais: composição mecânica-fisiológica do organismo e as necessidades expressivas do seu falante.

O *software* ExProsodia é programado em Visual-Basic para Excel com a finalidade de examinar automaticamente os elementos constituintes da prosódia. Esse equipamento analisa as variações entre o Tom Médio (TM), da frequência fundamental (F0) e as alterações no ritmo tonal geradas pela voz humana. Tom médio ideal de F0 consiste no esforço para a manutenção da voz com uma frequência relativamente estável. É obtido a partir de uma média aritmética acumulada no tempo em que a linguagem vocal foi produzida. O ritmo tonal, entretanto, é obtido como consequência de dois componentes da fala: a sustentação (S) e a finalização (F) (FERREIRA-NETTO, 2008, 2010, 2016).

A sustentação (S) é o elemento que, conforme explica Ferreira-Netto (2016), possibilita a produção do som emitido durante a fala, denominado de UBI (termo proposto a partir do inglês *Unit of Base of Intonation*). O mesmo autor compreende que, para o estabelecimento desse parâmetro, é necessário que a intensidade verbal seja maior do que zero, duração igual ou maior do que 20 milissegundos e frequência superior a 50 Hz. Por sua vez, a Finalização (F) consiste um decréscimo gradual que há no final do enunciado.

Em suma, os relatos foram gravados em arquivos de áudio, convertidos em parâmetros acústicos através da rotina ExProsodia (FERREIRA-NETTO, 2016). Em especial, neste trabalho é apresentada a análise dos parâmetros acústicos obtidos pelo aplicativo desenvolvido por Ferreira-Netto (2016) apenas para a frequência. São eles:

skew\_F0\_UBI – assimetria ou inclinação da curva de F0;

mediana\_F0\_UBI – mediana de F0;

cv\_F0\_UBI – coeficiente de variação de F0. Cálculo do desvio padrão de F0 pela média de F0 (TM);

kurt\_F0\_UBI – curtose da variação de F0;

TM-mUBI – diferença entre o TM e o menor valor de UBI. Esse parâmetro é obtido através do cálculo de subtração do TM pelo menor\_F0\_UBI;

skew\_TM – assimetria do tom médio;

mediana\_TM – valor relativo à mediana de TM;

cv\_TM – coeficiente de variação do Tom Médio (TM). Utilizado para averiguar a dispersão, esse parâmetro é estabelecido através do cálculo de divisão do desvio padrão do TM pela média do TM;

dp\_FEpos\_UBI – desvio padrão de foco/ênfase positivo;

skew\_FEpos\_UBI – assimetria ou inclinação do foco/ênfase positivo;

cv\_FEpos\_UBI – coeficiente de variação de foco/ênfase;

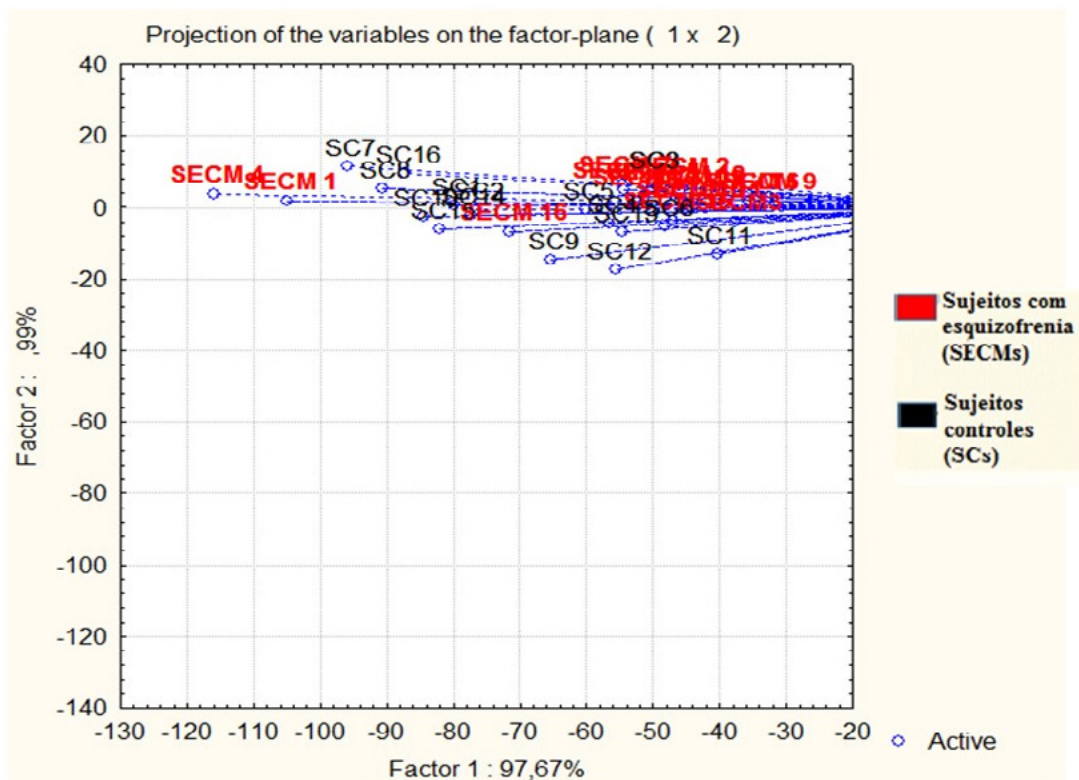
kurt\_FEpos\_UBI – curtose de foco/ênfase.

Posteriormente, os valores obtidos especificamente para esses parâmetros acústicos foram transformados em dados estatísticos para efetivar a Análise Multivariada com o auxílio do *software* Statistica, especialmente, para concretizar o cálculo de Análise de Componentes Principais (Covariância).

## Resultados

### Anamnese

Nessa etapa, a partir de dados iniciais dos diálogos estabelecidos com os participantes, foi possível verificar que houve uma divisão gráfica entre os grupos de amostras. Esse fato é ilustrado com o auxílio do Gráfico 1, no qual, em sua maioria, os SECMs representados na cor vermelha permaneceram mais próximos, em oposto aos SCs (preto) que apresentaram maior dispersão. Tais informações permitem argumentar que, nessa etapa, houve uma diferença estatisticamente significativa entre a voz de pessoas sem transtorno mental e pessoas diagnosticadas com esquizofrenia em seus diferentes graus.



**Gráfico 1.** Análise de Componentes Principais (Covariância) - Anamnese

**Fonte:** Elaboração própria

Abaixo estão representadas pela Tabela 1 as variáveis acústicas que foram determinantes para esse resultado.

**Tabela 1.** Anamnese

| Case           | Factor 1 | Factor 2 | Factor 3 |
|----------------|----------|----------|----------|
| skew_TM        | 226,522  | 14,727   | 5,5580   |
| cv_TM          | 224,142  | 16,244   | 7,5139   |
| cv_F0_UBI      | 222,824  | 15,632   | 7,2451   |
| skew_F0_UBI    | 220,908  | 12,106   | 1,6342   |
| cv_FEpos_UBI   | 210,132  | 14,930   | 1,7641   |
| skew_FEpos_UBI | 208,854  | 15,267   | 0,8565   |
| kurt_F0_UBI    | 202,726  | 3,673    | -15,8761 |
| kurt_FEpos_UBI | 165,084  | 12,722   | -31,9000 |
| dp_FEpos_UBI   | 58,082   | -104,938 | -30,6936 |
| TM-mUBI        | -273,847 | -36,556  | 95,6441  |
| mediana_F0_UBI | -725,382 | 38,449   | 6,6109   |
| mediana_TM     | -740,044 | -2,256   | -48,3573 |

**Fonte:** Elaboração própria

Conforme esboça a Tabela 1, as variáveis acústicas mais relevantes nessa etapa foram: para o Fator 1, o skew\_TM, o cv\_TM, o skew\_F0\_UBI e cv\_F0\_UBI. As duas primeiras representam parâmetros acústicos referentes à assimetria e à dispersão do Tom Médio. Os dois parâmetros seguintes representam a assimetria e a dispersão da frequência fundamental. O único fator significativo para o Fator 2 foi a mediana da frequência fundamental.

Esses dados indicam que há diferenças significativas entre a entoação de voz de pacientes com esquizofrenia e sujeitos controle sem histórico de transtorno mental. A isto se soma que os parâmetros acústicos que mais influenciaram a composição desse resultado foram a assimetria e a dispersão da curva da frequência.

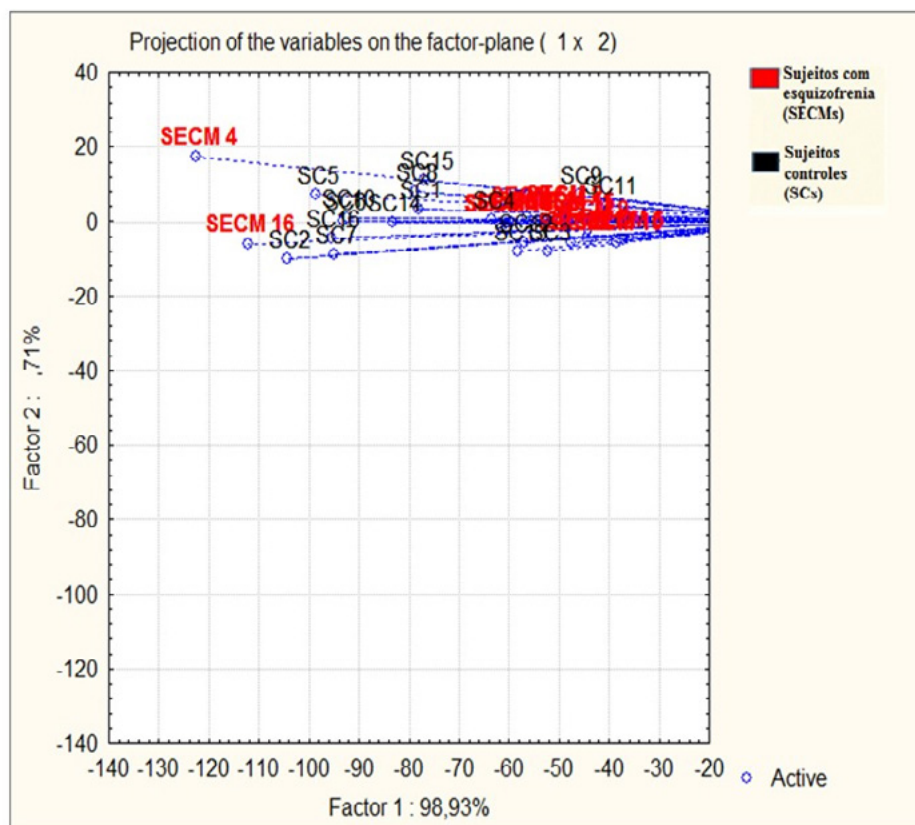
## **Relato empírico**

Em especial, nesta etapa, optou-se pelo uso da fala espontânea, partindo das mesmas considerações de Peres (2016), já que era nossa intenção promover manifestações emocionais decorrentes de situações reais relatadas pelos participantes, com a menor probabilidade de interferência da pesquisadora. Ainda assim ela estava exposta a uma diversidade de fatores envolvidos, devido a características pessoais distintas ou a influência do contexto clínico.

No ambiente clínico, tomar ciência das dores, dos medos e das raivas, do mesmo modo que reviver momentos felizes é uma experiência evolutiva. O analista pode ser palco de inveja, abandono, desamor ou mesmo uma figura que proporciona equilíbrio ao paciente. Esses sentimentos serão refletidos em seus comportamentos e em seu tom de voz (STEINBERG, 1990).

Os participantes de ambos os grupos se emocionaram, pois, para compor essa etapa a pesquisadora solicitava que fosse narrado um evento feliz. Ao final da explanação do participante, era proposto ao mesmo relatar uma circunstância triste dentre os seus momentos vividos. Foram narradas histórias de dores, alegrias, mortes e renascimentos; são biografias verbais que descrevem marcas na psique individual, garantindo a sua tonalidade emocional.

Nessa etapa, também foi possível perceber a divisão gráfica dos sujeitos em que há menor dispersão entre os SECMs, do mesmo modo como exprime o Gráfico 2.



**Gráfico 2.** Análise de Componentes Principais (Covariância) - Relato empírico

**Fonte:** Elaboração própria

Do mesmo modo, apresentamos abaixo na Tabela 2 as variáveis acústicas mais relevantes desse processo. Como é possível observar, foi significativo para o Fator 1 o skew\_TM, cv\_TM, cv\_F0\_UBI e o skew\_F0\_UBI. Logo, para o Fator 2 apenas o TM-mUBI denotou representatividade. Esse resultado sugere que as dificuldades entoacionais apresentadas pelos pacientes residam, principalmente, nesses parâmetros acústicos.

**Tabela 2.** Relato empírico

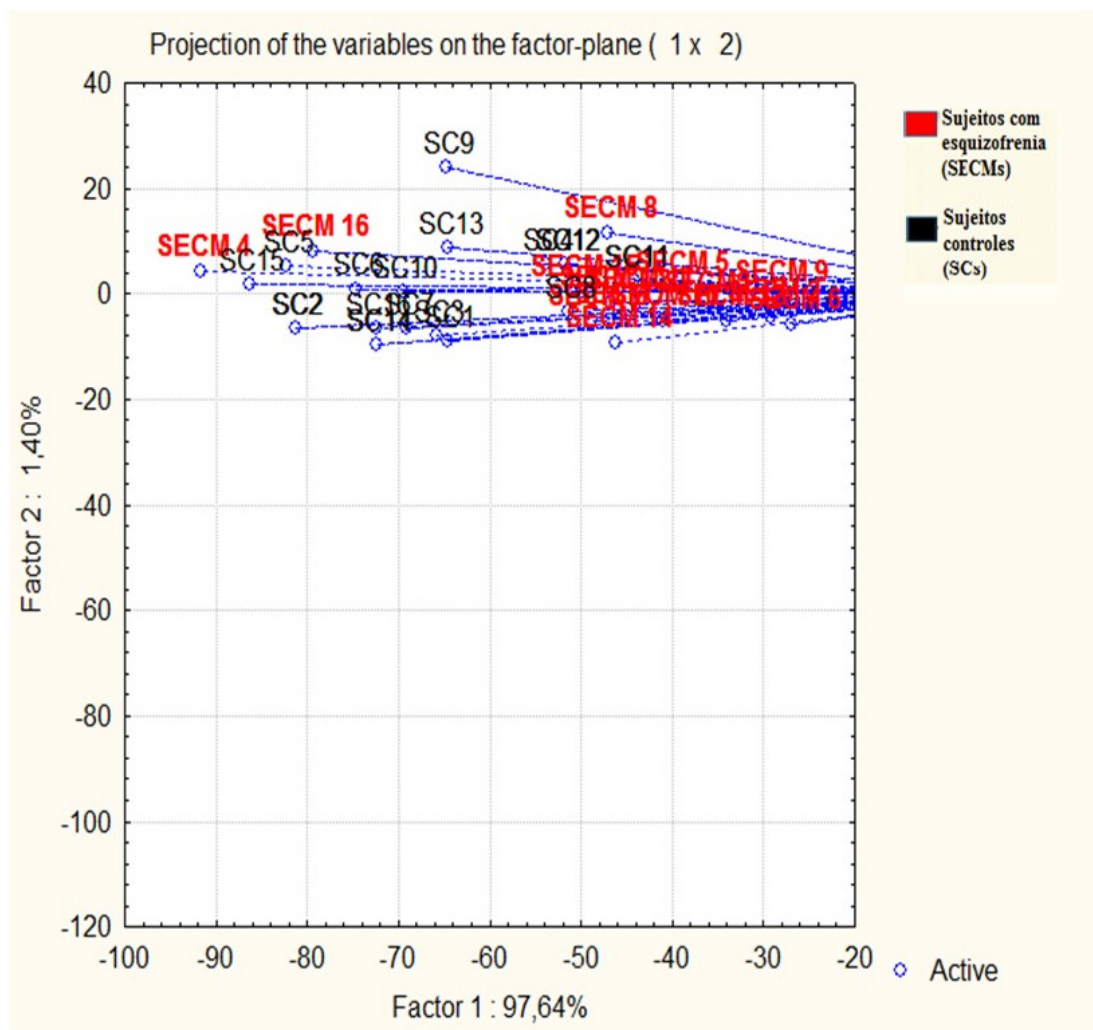
| Case           | Factor 1 | Factor 2 | Factor 3 |
|----------------|----------|----------|----------|
| skew_TM        | 237,365  | -1,7819  | 3,5331   |
| cv_TM          | 234,719  | -1,3700  | 4,5105   |
| cv_F0_UBI      | 233,483  | -1,2911  | 4,0357   |
| skew_F0_UBI    | 233,205  | -4,9220  | 1,1224   |
| kurt_F0_UBI    | 221,649  | -11,7599 | -4,6774  |
| cv_FEpos_UBI   | 221,168  | -4,4685  | 2,5889   |
| skew_FEpos_UBI | 219,724  | -4,5446  | 2,4924   |
| kurt_FEpos_UBI | 182,942  | -17,0672 | -5,6190  |
| TM-mUBI        | -278,846 | 94,3959  | -20,0862 |
| mediana_F0_UBI | -737,506 | -0,2098  | 50,4523  |
| mediana_TM     | -767,903 | -46,9809 | -38,3526 |

**Fonte:** Elaboração própria

### Descrição

A tarefa descritiva propiciou a seguinte distribuição gráfica dos grupos, conforme esboça o Gráfico 3, referente à Análise de Componentes Especiais (Covariância). Nessa ilustração, os SECMs (vermelho) permanecem mais agrupados entre si. Por contraste, os SCs (preto) encontram-se mais dispersos.





**Gráfico 3.** Análise de Componentes Principais (Covariância) - Descrição

**Fonte:** Elaboração própria

Na Tabela 3 fica evidente que as variáveis acústicas que mais se destacam são iguais às outras duas anteriores de fala espontânea. Para o Fator 1 obteve representatividade o skew\_TM, cv\_TM, cv\_F0\_UBI e skew\_F0\_UBI. A diferença foi no parâmetro do Fator 2, o qual nessa fase foi o dp\_FEpos\_UBI.



**Tabela 3.** Descrição

| Case           | Factor 1 | Factor 2 | Factor 3 |
|----------------|----------|----------|----------|
| skew_TM        | 177,360  | -3,8788  | 14,2595  |
| cv_TM          | 177,257  | -5,8522  | 14,8891  |
| cv_F0_UBI      | 175,838  | -5,5229  | 14,3692  |
| skew_F0_UBI    | 175,182  | -4,3700  | 9,0391   |
| cv_FEpos_UBI   | 162,341  | -9,8164  | 8,3346   |
| kurt_F0_UBI    | 160,882  | -14,5736 | -7,1052  |
| skew_FEpos_UBI | 160,682  | -11,2380 | 6,6143   |
| kurt_FEpos_UBI | 106,962  | -45,9721 | -44,4934 |
| dp_FEpos_UBI   | -20,903  | 96,8929  | -47,2094 |
| TM-mUBI        | -425,834 | 46,7737  | 46,7544  |
| mediana_TM     | -849,767 | -42,4427 | -15,4522 |

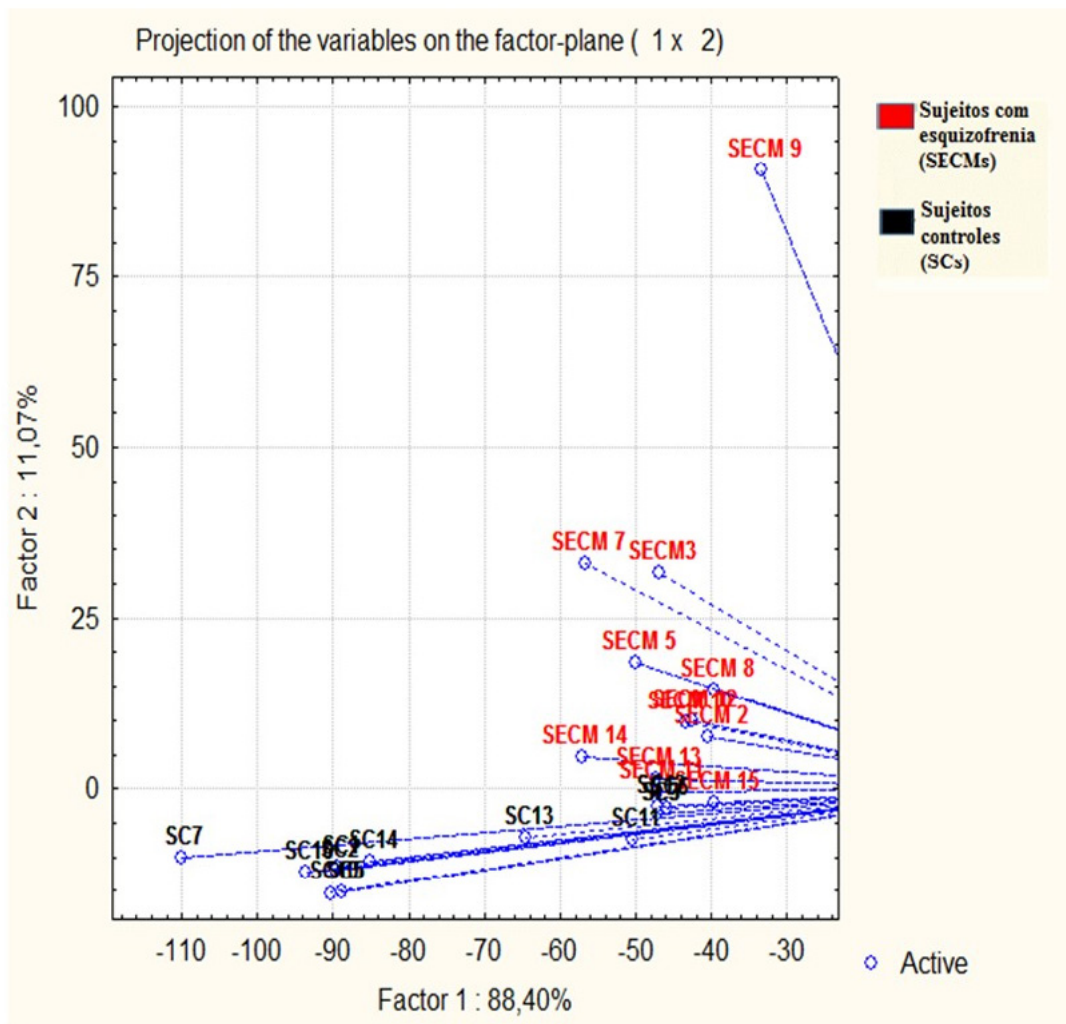
**Fonte:** Elaboração própria

### **Leitura de texto**

Essa etapa da pesquisa corrobora a hipótese apontada por Ferreira-Netto e Consoni (2008) ao sugerir que cada sujeito estabeleceria uma prosódia para ler um texto.

O mesmo foi apontado por Rapcan e outros (2010) ao realizar uma atividade de leitura com pessoas acometidas por esquizofrenia e sujeitos controles. Para os autores, o exame acústico possibilitado por esse teste fornece ferramentas fundamentais para o diagnóstico desse transtorno mental.

Tais princípios orientaram a elaboração dessa tarefa na presente pesquisa. Ambos os grupos eram escolarizados. Então, atendiam os requisitos necessários para executarem o proposto.



**Gráfico 4.** Análise de Componentes Principais (Covariância) - Leitura de texto

**Fonte:** Elaboração própria

Nessa etapa, apresentamos os resultados mais significativos, já que em sua maioria, os SECMs (vermelho) permaneceram na parte de cima da representação gráfica, em oposição aos SCs (preto).

**Tabela 4.** Leitura de texto

| Case           | Factor 1 | Factor 2 | Factor 3 |
|----------------|----------|----------|----------|
| cv_TM          | 215,211  | -42,8268 | 3,3554   |
| cv_F0_UBI      | 213,986  | -42,5856 | 2,8383   |
| skew_TM        | 213,455  | -47,7587 | 0,2484   |
| skew_F0_UBI    | 208,053  | -36,7955 | -4,2885  |
| cv_FEpos_UBI   | 195,765  | -34,7571 | -2,8495  |
| skew_FEpos_UBI | 188,513  | -23,3417 | -1,8462  |
| kurt_F0_UBI    | 167,012  | 24,5892  | -27,3415 |
| kurt_FEpos_UBI | -6,886   | 326,2038 | 12,4568  |
| TM-mUBI        | -273,815 | -81,2178 | 50,9361  |
| mediana_F0_UBI | -553,432 | -19,9065 | -0,9540  |
| mediana_TM     | -567,861 | -21,6033 | -32,5552 |

**Fonte:** Elaboração própria

Baseado na reprodução da Tabela 4, constatamos que os parâmetros acústicos mais relevantes para a consecução desse resultado no Fator 1 foram os mesmos das etapas anteriores, isto é, o cv\_TM, cv\_F0\_UBI, o skew\_TM e o skew\_F0\_UBI. Esse fato demonstra que é provável que a diferença entoacional na população analisada por este estudo esteja na assimetria e/ou na dispersão da curva da frequência.

Nessa fase, conforme exposto na Tabela 4, a kurt\_FEpos\_UBI também expressou grande representatividade.

## Considerações finais

Os resultados dos testes realizados por essa pesquisa apontaram a possibilidade de diferenciação de pacientes com esquizofrenia de sujeitos controles sem histórico médico de transtornos mentais baseada na análise de parâmetros prosódicos. Assim, como propuseram Mota e seus colegas (2014), avaliar como uma informação é transmitida fornece pistas salutares para a composição de uma perspectiva de diagnóstico mais acurado.

Do mesmo modo que apresentado nos gráficos de Análise de Componentes Principais (por covariância), foi possível constatar que os parâmetros acústicos da frequência mais

representativos para o resultado obtido foram da assimetria e da dispersão tanto do Tom Médio como da frequência fundamental nas tarefas de fala espontânea e na leitura narrativa.

À luz dessas considerações, é possível validar a hipótese do *Simulacrum of Neutral Intonation* (SNI) formulada por Martins e Ferreira-Netto (2017). Ao discorrer sobre uma "monotone speech" com prevalência de ritmo e "flat affect", no qual o sujeito tem uma redução quantitativa singular na manifestação emocional, tais condutas seriam evidentes em momentos de estresse intenso e/ou em alguns quadros clínicos de diferentes patologias psíquicas do mesmo modo como descrito neste trabalho ao versar sobre a prosódia afetiva na esquizofrenia.

Essa perda atípica da expressão da emoção através da linguagem pode estar associada à possibilidade citada por Behlau (2001) ao versar que a pouca variação prosódica pode indicar o controle emocional do falante. Huang e seus colegas (2011) sustentam a hipótese de que esse fato pode estar relacionado a um esforço consciente ou inconsciente do organismo dos indivíduos acometidos pela esquizofrenia para tentar manter a homeostase psíquica.

## REFERÊNCIAS

AGUILAR, M. A. de M.; MADEIRO, F. *Em-Tom-Ação: a prosódia em perspectiva*. Recife: Ed. Universitária da UFPE, 2007.

AMERICAN PSYCHIATRY ASSOCIATION. *Diagnostic and Statistical Manual of Mental disorders – DSM-5*. 5th.ed. Washington: American Psychiatric Association, 2013.

BACHOROWSKI, J.; OWREN, M. J. Vocal expression of emotion: Acoustic properties of speech are associated with emotional intensity and context. *Psychological Science*, v. 6, p. 219-224, jul. 1995.

BANSE, R.; SCHERER, K. Acoustic profiles in vocal emotion expression. *Journal of Personality and Society of Psychology*, v. 70, p. 614-636, mar. 1996.

BAUM, L. F. *O mágico de Oz*. São Paulo: Leya, 2011.

BEHLAU, M. *Voz: o livro do especialista*. v. I. Rio de Janeiro: Revinter, 2001.

CASTAGNA, F. *et al.* Prosody recognition and audiovisual emotion matching in schizophrenia: The contribution of cognition and psychopathology. *Psychiatry Research*, n. 205, p. 192-198, fev. 2013.

CHHABRA, S. *et al.* Voice identity discrimination in schizophrenia. *Neuropsychologia*, n. 50, p. 2730-2735, out. 2012.

FERREIRA-NETTO, W. *ExProsodia: resultados preliminares*. São Paulo: Ed. Paulistana, 2016.

FERREIRA-NETTO, W. *ExProsodia*. *Revista da Propriedade Industrial – RPI*, 2038, 26 out. 2010.

FERREIRA-NETTO, W. Decomposição da entoação frasal em componentes estruturadoras e em componentes semântico-funcionais. *In: CONGRESSO NACIONAL DE FONÉTICA E FONOLOGIA/CONGRESSO INTERNACIONAL DE FONÉTICA E FONOLOGIA, X/IV*. Niterói, 2008. Disponível em: <http://bit.ly/31YpzFM>. Acesso em: 22 jan. 2013.

FERREIRA-NETTO, W. *Variação de frequência e constituição da prosódia na Língua Portuguesa*. 2006. Tese (Livre Docência) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2006.

FERREIRA-NETTO, W. *et al.* Análise automática da entoação emotiva (colérica, triste e neutra) pelo aplicativo ExProsodia. *In: BRAZILIAN SYMPOSIUM IN INFORMATION AND HUMAN LANGUAGE TECHNOLOGY AND COLLOCATED EVENTS, 9*, Fortaleza, 2013. Proceedings-BRACIS 2013. Disponível em: <http://www.lbd.dcc.ufmg.br/colecoes/jdp/2013/001.pdf>. Acesso em: 22 jan. 2013.

FERREIRA-NETTO, W.; CONSONI, F. Estratégias prosódicas da leitura em voz alta e da fala espontânea. *Alfa*, n. 52, v. 2, p. 521-534, 2008.

HUANG, J. *et al.* Trying to be optimistic? The emotion perception of schizophrenia within conversation context. *Psychiatry Research*, v. 185, n. 1/2, p. 300-301, jan. 2011.

JORGE, A. C. A. *Prosódia afetiva na esquizofrenia*. 2018. Dissertação (Mestrado) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2018.

LIEBERMAN, P.; MICHAELS, S. B. Some aspects of fundamental frequency and envelope amplitude as related to the emotional content of speech. *Journal of the Acoustical Society of America*, v. 34, p. 922-927, jul. 1962.

MARTINS, M. V. M.; FERREIRA-NETTO, W. Proposal of description for an intonation pattern: The simulacrum of neutral intonation. *The Journal of the Acoustical Society of America*, n. 141, p. 3701, 2017.

MOTA, N. *et al.* Graph analysis of dream reports is especially informative about psychosis. *Scientific Reports*, v. 4, art. 3691, jan. 2014.

PERES, D. O. *A percepção da emoção na fala por nativos e não nativos*. 2016. Tese (Doutorado) – Faculdade de Filosofia, Ciências Humanas e Letras, Universidade de São Paulo, São Paulo, 2016.

RAPCAN, V. *et al.* Acoustic and temporal analysis of speech: a potential biomarker for schizophrenia. *Medical Engineering and Physics*, v. 32, n. 9, p. 1074-1079, jul. 2010.

SILVEIRA, N. da. *O mundo das imagens*. Rio de Janeiro: Ática, 1992.

SILVEIRA, N. da. *Imagens do Inconsciente*. Rio de Janeiro: Ed. Alhambra, 1981.

STEINBERG, W. *Aspectos clínicos da terapia Junguiana*. Tradução Pedro da Silva Dantas. São Paulo: Ed. Cultrix, 1990.

TRÉMEAU, F. A review of emotion deficits in schizophrenia. *Clinical Research: Dialogues in Clinical Neuroscience*, v. 8, n. 1, p. 59-70, 2006.