

Frequência Fundamental e Emoções: um estudo a partir da fala atuada em português brasileiro

FUNDAMENTAL FREQUENCY AND EMOTIONS: A STUDY BASED ON
ACTED SPEECH IN BRAZILIAN PORTUGUESE

Aline Mara de Oliveira **VASSOLER ***

Beatriz Raposo de **MEDEIROS ****

Resumo: Objetivo: analisar o comportamento da frequência fundamental (F_0) na fala emotiva e na fala neutra. Métodos: três atrizes profissionais participaram da pesquisa. As atrizes leram um texto com fala neutra e interpretando as emoções (*alegria, raiva, medo e tristeza*). Para cada situação, as atrizes repetiram o texto proposto cinco vezes, totalizando 75 gravações. O registro de amostra de fala ocorreu no laboratório Estúdio Multimeios do Centro de Computação Eletrônica da Universidade de São Paulo (CCE-USP) através de equipamentos específicos. Selecionou-se uma sentença do texto, a qual foi segmentada para unidades ainda menores (sílabas). A fim de obter os valores de F_0 , utilizou-se o *software Praat* e *scripts* específicos. Resultados: ao analisar a média de F_0 , foi possível constatar que a *alegria* se encontra nas faixas frequências mais altas, ao contrário da *tristeza*, que apresenta valores de F_0 baixo. A variação de F_0 na *raiva* manteve-se semelhante nas três atrizes, diferente do *medo*, que não manteve um padrão na variação de F_0 . Conclusão: os valores de F_0 , tanto da média, quanto da variação de F_0 são importantes na diferenciação das emoções: *alegria, raiva e tristeza*. A emoção

* Doutoranda do Departamento de Estudos Linguísticos (Oralidade e Letramento), da Universidade Estadual Paulista (UNESP – Rio Preto). Contato: alineoliveiravassoler@gmail.com.

** Doutorado (2002) em Linguística pela Universidade Estadual de Campinas. Professora Doutora do Departamento de Linguística da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo (FFLCH/USP). Contato: biarm@usp.br.

medo parece ser influenciada por outros comportamentos acústicos e fisiológicos – a serem estudados em estudo mais abrangente – além da F_0 .

Palavras-chaves: Fonética. Acústica da fala. Emoção.

Abstract: Purpose: Fundamental Frequency (F_0) in emotional and neutral speech in Brazilian Portuguese was analyzed. Methods: Three professional actresses participated in the survey reading a text in two conditions: neutral speech and acted emotion speech. Four emotions were taken into account for this study: joy, anger, fear and sadness. For each situation, actresses repeated five times, totaling 75 recordings. Speech sample recording took place in the Multimedia Lab Studio Electronic Computing Center of the University of São Paulo (USP-CCE) through specific equipment. A sentence was selected to be segmented and analyzed in smaller units (syllables). In order to obtain F_0 values, the software PRAAT and some of its scripts were used. Results: Analyzing F_0 average of each emotion, we found that *joy* shows higher frequency bands than *sadness* that has very low values of F_0 . Variation of F_0 in *anger* remained similar among the actresses speech. On the other hand, *fear* did not offer any F_0 range pattern. Conclusion: F_0 values, both mean and variation, show to be important in the differentiation of emotions: joy, anger and sadness. Fear seems to suffer other acoustic and physiological influences – a those could be analyzed in a more broad study – other than F_0 .

Key-words: Phonetics. Speech acoustics. Emotion.

Introdução

As emoções humanas manifestam-se no indivíduo a partir do nascimento e o acompanham ao longo de toda vida. As emoções podem ser interpretadas e expressas por meio de expressões faciais e vocais e por gestos corporais, que facilitam as relações interpessoais e auxiliam na adaptação do indivíduo com o meio no qual está inserido (CORREIA, 2007). Em decorrência dessa capacidade do homem de interpretar e enviar emoções, o estudo acerca dos estados emotivos despertou o interesse de diferentes ciências.

As emoções têm sido objeto de estudo de diversas áreas do conhecimento: as humanidades, as ciências sociais, as do comportamento e as tecnológicas. A Psicologia vem desenvolvendo diversos modelos teóricos

referentes às emoções e ao desenvolvimento emocional, sendo que a maioria das teorias recentes admite o reconhecimento da função adaptativa das emoções no desenvolvimento humano (MELO, 2005). A Antropologia ocupa-se em estudar a ligação entre as emoções e a dimensão cultural das sociedades. A História se empenha em descrever o curso do interesse pelas emoções, enquanto a Sociologia preocupa-se em perceber a relação entre o funcionamento emocional e os grupos sociais (KEMPER, 20041 apud MELO, 2005).

A Fonoaudiologia, por sua vez, preocupa-se com a expressividade vocal dos profissionais que trabalham com a voz (atores, cantores, advogados e jornalistas, dentre outros). O fonoaudiólogo deve estar atento a todos os recursos vocais, os quais sofrem variação de acordo com o contexto, com o interlocutor, com a mensagem, com a dinâmica vocal específica e com as emoções (OLIVEIRA; RAPOSO DE MEDEIROS, 2010). Tais mudanças vocais auxiliam esses profissionais a utilizar a expressão vocal de maneira eficaz em um discurso. Na integração mensagem e voz, surge inevitavelmente o fenômeno do discurso, que leva, tanto os profissionais da fonoaudiologia, como os estudiosos, a terem a linguagem como objeto de investigação. Assim, no domínio da linguagem, os estudos fonoaudiológicos podem se irmanar aos estudos linguísticos voltados para os fenômenos da fala. É o caso da fonética acústica, que estuda as propriedades físicas dos sons da fala, isto é, aquelas propriedades contidas nas ondas sonoras oriundas do aparelho fonador durante a produção da fala. A altura e a duração do som da fala são elementos importantes para que se conheçam mais a fundo a organização sonora do discurso, incluindo aí, ritmo e contornos melódicos, que podem ser medidos e delimitados em partes menores e maiores do discurso – desde a sílaba até um enunciado constituído de várias sentenças.

A relação entre as emoções e os parâmetros acústicos é estudada há muito tempo na literatura. Segundo Scherer (1989) e Banse e Scherer (1996), a interferência das emoções no sinal acústico tem sido investigada há mais de um século e envolve o comportamento biológico, a psicologia, a fala e a comunicação. Tais estudos confirmaram as relações entre os diferentes estados emocionais e motivacionais por meio de parâmetros acústicos vocais, já que existem evidências de que, ao mudar a respiração, a fonação e a articulação, o padrão do sinal acústico também é alterado. No entanto, ainda não se tem o conhecimento sistemático dessa relação.

A literatura das emoções demonstrou que, em geral, as medidas utilizadas são: frequência da vibração das pregas vocais (denominada de frequência fundamental – F_0), nível de variação e contorno da F_0 , energia vocal; distribuição do espectro de energia (relativa à relação entre as frequências baixas e altas, que afeta a qualidade de voz); localização dos formantes ($F_1, F_2, \dots F_n$, referente à percepção da articulação), variedade do fenômeno temporal, incluindo pausas e duração.

Medidas como a F_0 , a intensidade de voz, a velocidade de fala (LAUKKA, 2004) e a energia espectral (BANSE; SCHERER, 1996) são os parâmetros acústicos mais investigados na fala emotiva, já que são valores que expressam as mudanças respiratórias, fonatórias e articulatórias do indivíduo.

Além desses parâmetros, a entonação da fala é também utilizada frequentemente na diferenciação das emoções, uma vez que está relacionada com a expressão vocal do falante. A entonação da fala é representada acusticamente pela variação da F_0 , a qual é produzida pelas mudanças relativas à altura e tensão das pregas vocais, controladas pela musculatura intrínseca e extrínseca da laringe. A pressão de ar oriunda abaixo da laringe é considerada influência secundária na taxa de vibração das pregas vocais (CRUTTENDEN, 1986).

A F_0 é o principal parâmetro acústico utilizado para diferenciar as emoções na fala (SCHERER, 1989; BANSE; SCHERER, 1996; LAUKKA, 2004; RODRIGUES, 2007), por isso, será o componente acústico analisado neste trabalho. Assume-se a hipótese de que os valores de F_0 diferenciam as quatro emoções – *alegria, raiva, medo e tristeza* – da fala neutra¹.

A hipótese de que um F_0 médio da fala está relacionado a uma dada emoção já foi investigada e verificada ou parcialmente verificada, em outras línguas como no Inglês (LAUKKA, 2004), no Alemão (PAESCHKE; SENDLMEIER, 2000) e no Português Europeu (RODRIGUES, 2007), porém no Português Brasileiro (PB), pesquisas dessa natureza ainda são escassas. Dessa forma, esse trabalho busca contribuir com dados e análises de F_0 na fala emotiva e na fala neutra do PB.

¹ Para esse trabalho, o termo “fala neutra” será utilizado quando se referir à forma habitual da leitura em voz alta, ou seja, sem inflexões emoções intencionais.

1 As Emoções na Fala

As emoções têm sido objeto de estudo há muitos anos e um dos pioneiros nos estudos das expressões emotivas foi Charles Darwin (1872), que na obra *The expression of the emotions in man and animals*, analisou as relações das expressões faciais dos seres humanos e dos animais em conjunto com a expressão funcional das emoções durante a intenção dos movimentos. A preocupação do autor estava voltada para as características biológicas e para os componentes inatos das expressões faciais. O autor defende que os comportamentos expressivos dos humanos e dos primatas podem ser consequências da adaptação ao meio no qual estão inseridos ou de respostas funcionais propriamente ditas.

A definição e a classificação das emoções humanas são bastante controversas na literatura. A teoria discreta das emoções, de Darwin, defendida posteriormente por Ekman (1992), propõe a existência de nove emoções básicas ou fundamentais. Tais emoções são caracterizadas por padrões fisiológicos específicos externalizados pela expressão facial e pela vocal (LAUKKA, 2004). São ditas emoções “básicas” aquelas ligadas às situações da vida, tais como *felicidade, tristeza, medo, raiva e surpresa* (EKMAN, 1992). A *raiva* corresponderia às situações de competição, o *medo* estaria ligado ao perigo, a *alegria* estaria envolvida com a cooperação, e a perda despertaria a *tristeza* (SCHERER, 1989).

No século XX, iniciaram-se os primeiros estudos empíricos que investigavam o efeito das emoções na voz. Os psiquiatras tentavam diagnosticar distúrbios ligados à emoção por meio de métodos da análise eletroacústica (LAUKKA, 2004). A primeira diferenciação de que temos notícia, entre fala triste e fala alegre, medindo-se a frequência fundamental, foi obtida por Skinner (1935). Em estudo experimental, o autor indicou que indivíduos tendem a usar uma frequência fundamental mais baixa para a fala triste (*sad*) e elevá-la na fala alegre (*happy*).

Diversos autores (SCHERER, 1989; BANSE; SCHERER, 1996) confirmaram, por meio de diversas pesquisas envolvendo os diferentes estados emocionais e as mudanças dos parâmetros acústicos vocais, que, para cada emoção, modifica-se a respiração, a fonação e a articulação do indivíduo, e, conseqüentemente, tem-se uma modificação no padrão do sinal acústico.

No artigo *O que são as emoções? E como elas são medidas? (What are emotions? And how can they be measured?)*, Scherer (2005) tenta discutir a definição do que é emoção e argumenta como esse conceito tem se tornado banalizado. O autor acredita que a emoção seria uma resposta mediante um estímulo interno ou externo. Scherer (2005) demonstra as relações entre os substratos neuronais e o componente emocional.

Quadro 1 – Tradução da tabela de Scherer (2005)

Função da emoção	Principais substratos neuronais	Componente emocional
Avaliação dos objetos e eventos	Processamento de informação (SNC)	Componente cognitivo (avaliação)
Regulação do sistema	Suporte (SNC, SNE, SNA)	Neurofisiológicos (componentes corporais)
Preparação e direção da ação	Executivo (SNC)	Componente motivacional (tendências de ação)
Comunicação de reação e intenção comportamental	Funcionamento (ação) (SNS)	Componente da expressão motora (facial e expressão vocal)
Monitoramento do estado interno e interação organismo-ambiente	Monitor (SNC)	Percepção subjetiva (componente emocional experiência)

A sigla SNC refere-se ao Sistema Nervoso Central, SNE refere ao sistema nervoso neuroendócrino, SNA é referente ao Sistema Nervoso Autônomo, e SNS está ligado ao Sistema Nervoso Somático.

Essas modificações nos componentes emocionais estão integradas ou sincronizadas aos episódios das emoções. No momento em que determinado componente é alterado, todos os outros componentes interagem entre si, causando, assim, alterações nos outros componentes (SCHERER, 1984).

As emoções são acompanhadas por várias respostas adaptativas do sistema nervoso autônomo e somático (JOHNSTONE; SCHERER, 2000). Essas respostas proporcionam modificações no funcionamento parcial ou total do sistema de produção de fala, como na respiração, na vibração das pregas vocais e na articulação.

Patrik e Västfjäll (2008) propõem três diferentes métodos para investigar as emoções na fala: o nível psicológico, o nível fonatório e articulatório, e o nível acústico. O nível psicológico busca descrever os impulsos nervosos e intervenção dos músculos; o nível fonatório e articulatório investiga a posição e os movimentos das principais estruturas das pregas vocais. Já o nível acústico analisa as ondas sonoras provenientes da fala, inferindo as mudanças articulatórias e fisiológicas do trato vocal.

Assim, o que determina as variações vocais que acompanham as emoções são as modificações fisiológicas que, por sua vez, induzem a alterações nos sistemas de produção vocal. A alteração de um dos desses componentes produz alterações no outro componente. Por exemplo, numa situação em que se exige um padrão respiratório maior, aumenta a necessidade do suporte de oxigênio, que vai afetar a expressão facial (forma da boca) e a expressão vocal (alterações na pressão subglotal), bem como um número de parâmetros fisiológicos periféricos (CORREIA, 2007).

O ouvinte não atenta somente para o conteúdo linguístico do discurso do falante, mas também para as mudanças acústicas que o falante pode provocar na sua voz. Além das informações sintáticas e semânticas, a comunicação verbal transmite pistas prosódicas que carregam conteúdos fundamentais para a compreensão da mensagem. Essas informações permitem que o ouvinte, não só codifique e decodifique toda a mensagem falada, mas também compreenda importantes funções linguísticas como a estrutura do discurso, as informações sobre a personalidade do indivíduo, a idade do falante, o sexo e até a condição física, emocional e atitudinal com relação ao tema, ao parceiro da conversa ou à situação propriamente dita. Torna-se, então, necessária a realização de estudos linguísticos aprofundados que envolva a F_0 , a duração, a intensidade, o tom de voz, o ritmo, e os outros elementos da prosódia (ERICKSON et al., 2006).

2 Emoções e os Correlatos Acústicos da Fala

Abordaremos aqui alguns achados existentes na literatura sobre a relação entre aspectos acústicos da fala e as quatro emoções que constituem o objeto do estudo que ora apresentamos, a saber: a *raiva*, a *alegria*, a *tristeza* e o *medo*.

2.1 A *raiva*

Scherer (1985) afirma que as condições que geram a *raiva* estão relacionadas às situações que geram obstáculos e impedem o sujeito de satisfazer alguma necessidade ou de atingir algum objetivo. O indivíduo pode se sentir confiante para remover os obstáculos por meio da força ou para prevenir as consequências negativas. Além disso, essa emoção pode estar relacionada ao sentimento de frustração.

A *raiva* também pode ser dessa irritação ou cólera, além de outros termos relacionados com sentimentos interpessoais hostis. As características vocais dessa emoção foram consideradas moderadamente tensas e com registro de peito. Em se tratando dos aspectos acústicos, a intensidade vocal foi forte, com evidências de pulsos fortes no espectro e presença de alguns ruídos inter-harmônicos. Para a F_0 , a média e a extensão foram baixas com algumas perturbações, e no primeiro formante (F1) a banda estava estreita (SCHERER, 1985).

Já Stibbard (2001) e Laukka (2004) investigaram a F_0 média alta, e nela encontraram grande variação a longo termo e contorno entoacional crescente para a *raiva*. A média de F1 também era alta, a intensidade vocal era forte, a articulação estava precisa e a velocidade de fala era rápida (com poucas pausas). Na imagem espectral, observaram-se poucas irregularidades microestruturais e a rampa espectral encontrava-se íngreme. Verifica-se que, para a F_0 , não existe homogeneidade no resultado dos achados acústicos, no entanto a intensidade vocal prevalece forte na fala raivosa.

2.2 A *alegria*

A *alegria* caracteriza-se pelo encontro de um estímulo intrínseco agradável e/ou a satisfação de uma necessidade ou a sensação prazerosa sentida ao atingir um objetivo. As modificações do aparelho vocal podem

ser caracterizadas por sons respiratórios intensos, por fonação relaxada e ampla e por configurações do trato vocal também relaxadas. As vocalizações podem ser breves e suaves e possuir frequência fundamental relativamente baixa, bem como pequena variação ao longo da fala, contornos entoacionais ascendentes ou descendentes e largura de banda do F1 um pouco mais ampla (SCHERER, 1985).

Outros estudos que avaliaram a *alegria* encontraram alta taxa de vibração das pregas vocais, sendo que a F_0 sofreu aumento na média e houve alta variação durante a fala. Notaram-se algumas evidências de aumento de energia nas altas frequências e na taxa de elocução (BANSE; SCHERER, 1996; LAUKKA, 2004). O contorno entoacional também foi crescente, assim como ocorrido na *raiva* (LAUKKA, 2004; OLIVEIRA; RAPOSO DE MEDEIROS, 2010), a intensidade vocal foi forte, a variabilidade da intensidade foi alta, a média do F1 encontrava-se elevada, com onda espectral íngreme, velocidade de fala rápida e poucas pausas (LAUKKA, 2004).

2.3 A *tristeza*

A fala triste pode surgir quando ocorre um evento que impeça o sujeito de alcançar um importante objetivo ou de realizar um forte desejo, cujas consequências não podem ser controladas pelo organismo de nenhuma forma ou não podem ser evitadas. Esse estado emocional reflete o interior do sujeito e apresenta comportamentos da *tristeza*, momentos de desesperança e de isolamento (SCHERER, 1985).

De acordo com Scherer (1979), pode haver dois tipos de fala triste: a fala triste passiva e a fala triste ativa. A fala triste passiva está relacionada com a *tristeza* calma e tranquila. A fala triste ativa ocorre em situações de luto e afeta as coordenações motoras envolvidas no controle das vibrações das pregas vocais e nos movimentos articulatorios supraglóticos. As características da voz triste provavelmente incluíram respiração passiva e tranquila assim como musculatura da fonação e do trato vocal relaxados. A F_0 obteve média baixa, variações restritas, perturbações fortes (dada à flacidez das pregas vocais) e contorno decrescente. A intensidade vocal apresentou-se muito fraca, com presença de ruído espectral (devido à soprosidade), faixa de frequência estreita no espectro e energia forte nas altas frequências.

Os resultados da acústica da fala característica da *tristeza* ainda são incongruentes (BANSE; SCHERER, 1996). Laukka (2004) encontrou a F_0 baixa, com pouca variação, contornos de F_0 decrescentes, *jitter*² baixo e intensidade vocal fraca. Além de baixa frequência de alta energia, baixa média de F1, onda glótica arredondada, velocidade de fala lenta e muitas pausas na fala.

Erickson et al. (2006) estudaram acústica e articulatoriamente a fala triste, comparando dados da fala espontânea com a atuada. Os resultados mostraram que a fala triste espontânea e imitada parecem ter características acústicas semelhantes (alta F_0 , mudança de F1 e na qualidade da voz), porém a articulação é diferente em termos de posições dos lábios, da língua e da mandíbula.

Os autores defendem que, ao interpretar a *tristeza*, as atrizes foram sensíveis a algumas características da fala triste espontânea. Durante a interpretação da emoção, notou-se que o falante aproximou-se dos parâmetros da fala espontânea na F_0 e na qualidade de voz, porém, a duração da fala mostrou-se divergente da duração da fala habitual. Os autores acreditam que as atrizes utilizaram, como recurso, um conjunto de estereótipos da sua memória que poderia “lembrar” a *tristeza*.

2.4 O medo

O *medo* pode ocorrer (com ou sem consequências negativas) após outros eventos que colocam em risco as necessidades ou os objetivos importantes do indivíduo, tais como a sobrevivência e a integridade física. A resposta vocal para essas transformações emocionais é uma aspiração súbita, com pouca sopro, tensão na fonação e no trato vocal e registro de cabeça. No que diz respeito aos sinais acústicos, a F_0 foi alta, com grandes perturbações, altas variações e decrescente contorno entoacional. A banda do F1 apresentou-se estreita com o aumento da faixa do espectro de frequência (amplo espaço entre os harmônios) e com forte energia nas regiões altas do espectro.

² O *jitter* é uma medida de curto termo (ciclo a ciclo) e está ligada às pequenas variações involuntárias na F_0 , logo, identifica o grau de estabilidade do sistema fonatório. É uma medida utilizada na investigação da qualidade vocal.

Os sinais acústicos associados ao *medo* ainda não estão consolidados na literatura, todavia, pode-se afirmar que há altos níveis de vibração das pregas vocais e a taxa de articulação é, na maioria das vezes, muito rápida. Os autores também encontraram um aumento na média da F_0 para formas mais leves da emoção, tais como a preocupação e a ansiedade (BANSE; SCHERER, 1996).

Laukka (2004) encontrou a intensidade de voz fraca (salvo nos episódios em que o *medo* era parecido com o pânico), com momentos de variabilidade na intensidade de voz para forte. Já Williams e Stevens (1972) encontraram a média da F_0 baixa, com onda espectral arredondada e microestruturas irregulares. A articulação desse tipo de fala é, geralmente, mais precisa do que na fala neutra.

3 Algumas Considerações sobre Métodos de Coleta de Dados da Fala Emotiva

É importante destacar as metodologias comumente utilizadas para capturar as emoções na fala. A literatura descreve três métodos para capturar a voz do sujeito, a saber: a fala espontânea, a fala atuada e a emoção evocada.

Os pesquisadores que utilizam a fala espontânea para análise das emoções defendem que esse é o método mais direto e autêntico de capturar as emoções humanas, porém ainda existem muitas dificuldades (BANSE; SCHERER, 1996), já que não garante ao pesquisador a captura da emoção esperada. Além disso, os falantes devem ser gravados desconhecendo a gravação, para que eles se comportem de maneira natural, o que compromete os aspectos éticos da pesquisa (CAMPBELL, 2000, 2001).

A emoção denominada evocada é aquela para cuja gravação o pesquisador prepara um local adequado e induz o sujeito para determinada situação a fim de provocar a emoção. No entanto, o estímulo que se pretende evocar pode provocar reações diferentes em cada sujeito. Geralmente, os pesquisadores não optam por esse tipo de metodologia por barreiras éticas. Um exemplo de falta de ética na indução de emoção seria assustar alguém e, em seguida, gravar a voz desse sujeito sem o seu conhecimento (GUSTAFSON-CAPKOVÁ, 2001).

Na fala atuada, o ator interpreta a emoção específica que o pesquisador deseja investigar, possibilitando o controle mais efetivo da emoção que se pretende estudar, o que não acontece com os demais métodos.

No entanto, o grau de naturalidade e a qualidade da atuação são frequentemente questionados. Alguns pesquisadores utilizaram atores profissionais (BANSE; SCHERER, 1996; VASSOLER, 2011), em outros casos, os atores amadores foram os sujeitos (FAIRBANKS; HOAGLIN, 1941 apud GUSTAFSON-CAPKOVÁ, 2001).

Além da fala atuada permitir o controle da emoção que se pretende analisar, também possibilita ao pesquisador ter, de antemão, o conteúdo verbal que os atores irão gravar. Esse controle possibilita a comparação entre os atores participantes da pesquisa, o que diminui o grau de variabilidade da amostra da fala emotiva.

Considerando que os três grupos têm os prós e os contras para o mapeamento das emoções por meio da voz, faz-se necessário que o pesquisador escolha o melhor método que corresponde aos objetivos da pesquisa. Os debates giram em torno de qual seria a metodologia mais adequada para as investigações dos estados emotivos. Essa problemática envolve desde a escolha do método de gravação das vozes, os sujeitos que participarão da pesquisa, o equipamento a ser utilizado para a gravação até o *corpus* que servirá como base para as gravações das vozes.

4 Material e Métodos

A fim de medir a F_0 das falas emotivas, escolhemos um método bastante simples, que consistiu em solicitar a interpretação de quatro emoções a atrizes, acreditando, como já explicamos, ser este um método válido. Assim, optou-se, neste trabalho, pela utilização da fala atuada por atrizes profissionais, que interpretam a emoção específica que o pesquisador deseja investigar. Esse método possibilita o controle mais efetivo da emoção que se pretende estudar, o que não acontece com os demais métodos (fala espontânea e emoção evocada) (BANSE; SCHERER, 1996).

4.1 Sujeitos

A amostra de fala foi obtida por meio da gravação das vozes de três atrizes brasileiras (atriz 1, atriz 2 e atriz 3) com experiência profissional entre 20 e 30 anos.

Os critérios de inclusão para essa pesquisa foram: a aceitação do Termo de Consentimento Livre e Esclarecido (TCLE); ser atriz há pelo menos 10 anos e ser do sexo feminino³. Os sujeitos leram um texto no Português Brasileiro e variante dialetal de São Paulo, interpretando as emoções solicitadas e descritas anteriormente.

Além das emoções, todas as atrizes leram o texto de forma habitual, ou seja, sem inflexões emoções intencionais, que nesse trabalho denominaremos *fala neutra*. A fala neutra é comumente utilizada nos estudos acerca da fala emotiva para servir de situação controle.

4.2 *Corpus*

O *corpus*, previamente utilizado em outros trabalhos (FIGUEIREDO, 1993), foi escolhido com o objetivo de oferecer às atrizes um texto de estilo científico e o mais árido possível, a fim de evitar emoções já implícitas no texto. Segue abaixo o texto de Vaz (1983):

A reatividade dos linfócitos, as células do sangue que fabricam anticorpos são individualizadas. Em cada organismo, as células do fígado são provavelmente iguais entre si, as da pele também, mas os linfócitos são diferentes uns dos outros. Cada um difere do seguinte por possuir na membrana diferentes receptores, moléculas que garantem a aderência a certas estruturas (ou a capacidade de fixar certas substâncias). Assim, o linfócito seguinte adere às estruturas diferentes. Para ser mais exato, as diferenças existem entre clones de linfócitos. Quando um determinado linfócito se multiplica e gera duas, quatro, oito milhares de cópias idênticas, esse conjunto constitui um clone linfocitário. Dentro de um mesmo clone, os linfócitos são iguais: têm os mesmos receptores de membrana, aderem às mesmas coisas, participam das mesmas interações.

Para a extração das medidas de F_0 , elencou-se uma sentença do texto acima: **“As células do fígado são provavelmente iguais entre si”**.

³ Este trabalho foi aprovado pelo Comitê de Ética em Pesquisa da Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, sob o número 535/2010.

Cada uma das atrizes leu o texto interpretando quatro emoções: *raiva*, *alegria*, *medo* e *tristeza*, além da fala neutra, totalizando 75 gravações (4 emoções + 1 controle (fala neutra) x 3 atrizes x 5 repetições).

4.3 Coleta de dados

As gravações das vozes foram realizadas no laboratório Estúdio Multimeios do Centro de Computação Eletrônica da Universidade de São Paulo (CCE-USP). Para a captação dos dados, foram utilizados equipamentos de alta qualidade: computador *Pentium 4* de 3,2GHz, com 1GB de RAM e 180GB de HD, acoplado a uma placa de som *Creative Sound Blaster Audigy2*, microfone *Shure SM58* e fone de ouvido *Philips SBC,HP195*. O registro dos dados foi realizado por meio do *software* de edição de som *Sound Forge 9*, com a mesa de som *Mackie CFX16*, com equalizador de 80Hz a 12KHz e taxa de amostragem de 44 KHz e taxa de quantização de 32 bits.

A gravação da fala de cada atriz obedeceu à seguinte ordem:

1. Aceitação ou negação por parte do sujeito na participação da pesquisa mediante assinatura do Termo de Consentimento Livre e Esclarecido;
2. Solicitou-se que as atrizes lessem o texto e interpretassem as emoções estudadas: *alegria*, *raiva*, *medo*, *tristeza*, e realizassem também uma leitura neutra. A interpretação de cada emoção foi exclusiva da atriz, sem interferência da pesquisadora. Não houve orientações ou sugestões com relação à interpretação das atrizes;
3. Todas as atrizes repetiram cinco vezes cada leitura para cada emoção e para a fala neutra. Elas interpretaram uma emoção de cada vez, realizando as cinco repetições solicitadas.

É importante ressaltar que os sujeitos foram gravados separadamente. Portanto, não existiu contato entre eles, evitando a influência da interpretação de uma atriz sobre a outra.

4.4 Segmentação e extração de F_0 dos dados

O texto foi segmentado em sentenças a fim de facilitar a análise dos dados. Em seguida, as sentenças foram segmentadas em unidades menores, ou seja, em sílabas. Os valores da F_0 foram extraídos em um programa livre,

o Praat 5.3.52, desenvolvido por Boersma e Weenink (2013) e *scripts* necessários. A segmentação dos fonemas foi realizada automaticamente, porém com ajustes e inspeções visuais.

A figura 1 (na página seguinte) tem uma imagem ilustrativa da segmentação realizada no espectrograma.

A F_0 é o correlato acústico do número de vezes que uma onda se repete em determinada unidade de tempo, sendo que a medida foi padronizada como Hertz (Hz) e um ciclo por segundo refere-se a um Hz (ORLIKOFF; KAHANE, 1996).

A extração de F_0 compreendeu a extração da F_0 média, da F_0 mínima e máxima, em Hertz, da sentença escolhida. Tais medidas da frequência fundamental são consideradas parâmetros globais. Obteve-se uma média geral de cada sentença a partir dos valores de F_0 , repetindo esse procedimento em cada repetição. Por fim, foi gerada uma única medida de F_0 .

5 Resultados e Discussões

Como discutido anteriormente, diversos parâmetros acústicos são utilizados nos estudos da fala emotiva, porém o mais utilizado para a diferenciação das emoções é a F_0 . Os resultados foram descritos, utilizando as medidas da fala neutra, como um ponto de referência. O gráfico 1 permite visualizar o comportamento das médias de F_0 nas quatro emoções (*alegria, raiva, medo e tristeza*), além da fala neutra (referência para comparações):

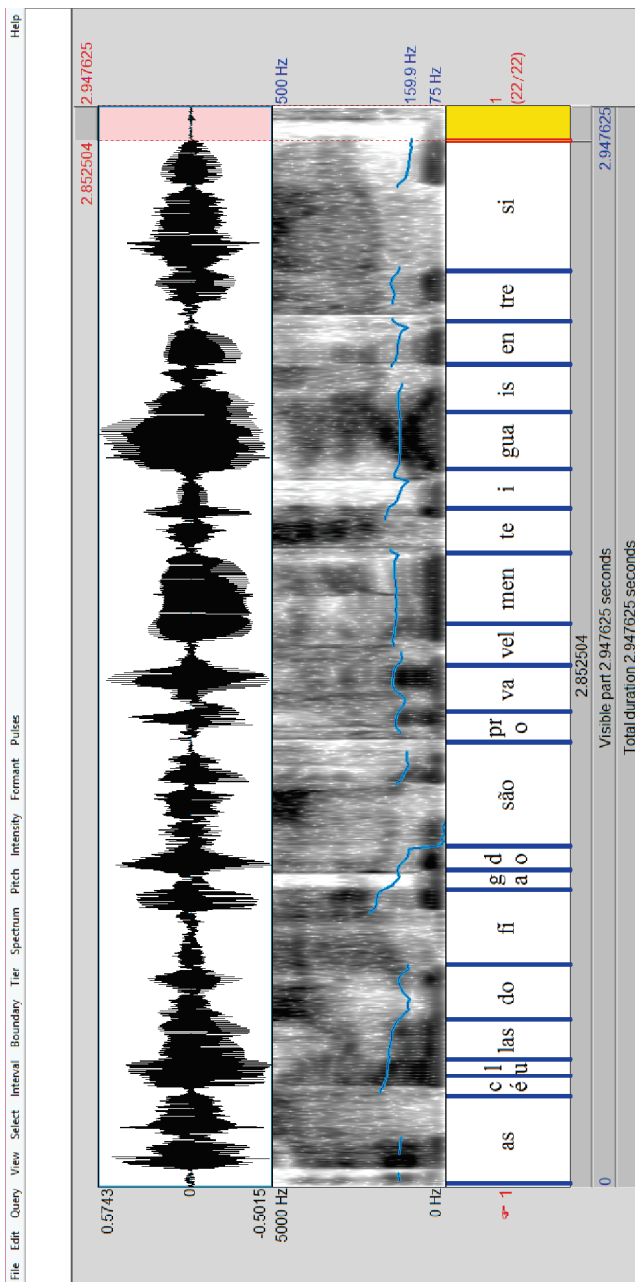


Figura 1 – Espectrograma da sentença analisada na fala neutra

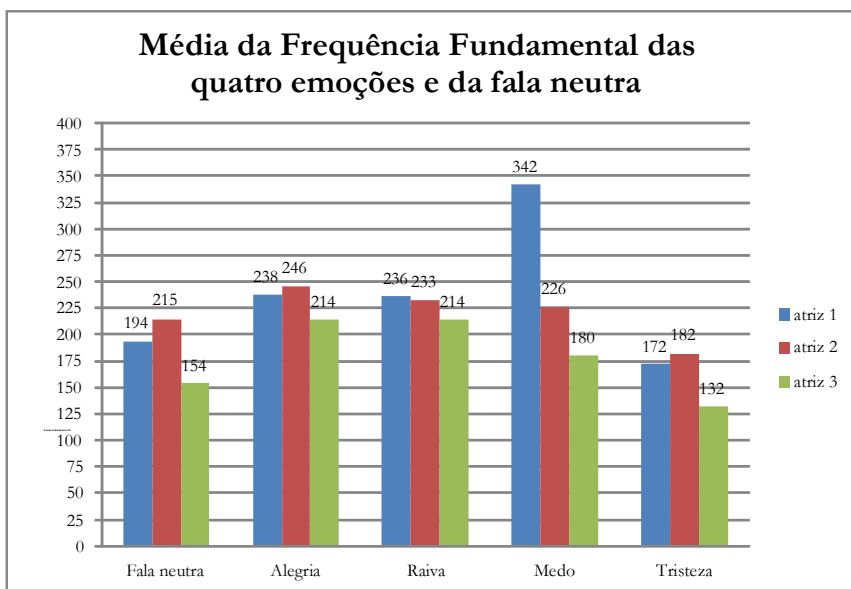


Gráfico1 – Valores da média de F_0 (média das cinco repetições) das três atrizes nas quatro emoções e na fala neutra

A disposição geral das emoções vinculadas às medidas de F_0 foi a seguinte: a *alegria* está na faixa de valores maiores de F_0 , ao passo que a *tristeza* apresentou os menores valores de F_0 , se comparado com as demais emoções; os valores da média de F_0 da *alegria* e da *raiva* são próximos, dificultando a comparação dessas emoções por meio dessa medida apenas.

Em geral, a média da F_0 da voz feminina no Português Brasileiro é de 215,42 Hz (para os homens entre os 127,61 Hz e os 142,63 Hz), o que também se observa nos dados da fala neutra dessa pesquisa (ARAÚJO; GRELLET; PEREIRA, 2002). Em geral, a *alegria* apresentou F_0 alto se comparado às demais emoções e à fala neutra. Em contrapartida, a *tristeza* foi interpretada em uma faixa de F_0 mais baixa e em estreita variação de F_0 , o que corroboram os resultados encontrados em trabalhos anteriores (BANSE; SCHERER, 1994; SOBIN; ALPERT, 1999; PAESCHKE; SENDLMEIER, 2000; LAUKKA, 2004).

A fala alegre tende a ter a F_0 tanto em valores de F_0 médio, quanto de variação, portanto se encontra em frequências mais altas que a fala neutra.

Além disso, apresenta uma maior oscilação entre os valores máximos e mínimos de F_0 (SCHERER, 1989; PAESCHKE; SENDLMEIER, 2000; ERICKSON et al., 2006).

No Português Europeu (PE), Correia (2007) avaliou as emoções evocadas em um grupo com alterações vocais (rouquidão, aspereza, entre outros) e outro grupo sem alterações vocais. A média da F_0 no grupo controle (sem patologias vocais associadas) foi de 252,8 Hz e o desvio padrão de 25,3 Hz, valores próximos aos encontrados no Português Brasileiro (gráfico 1). Já o grupo experimental executou a evocação da emoção alegre com medidas superiores, tanto em média de F_0 , quanto na intensidade.

A maioria das pesquisas aponta que a média e a variação de F_0 da *tristeza* comumente são as mais baixas, se comparadas às demais emoções. Um estudo investigou as relações entre a fala triste interpretada e a espontânea e mostrou que a primeira tem sempre a tendência de mudar a qualidade de voz, tal como a “*creaky*” voz. Ao contrário da maioria dos casos, nesse mesmo estudo, os autores encontraram valores altos para a média da F_0 na *tristeza* (ERICKSON et al., 2006).

No PE, a média da F_0 foi de 235,6 Hz com desvio padrão de 26,0 Hz. Se comparado com o grupo com patologias vocais, o grupo controle executou essa tarefa recorrendo a uma média de F_0 e uma intensidade média inferior.

Tanto a *alegria* quanto a *tristeza* são emoções que se diferenciaram dos valores da média e da variação de F_0 , sendo ambas contrárias entre si, o que aponta para uma diferenciação categórica dessas duas emoções. No entanto, na média de F_0 , a *alegria* e a *raiva* parecem não se diferenciar. A *alegria* e a *raiva* são caracterizadas por altos parâmetros de F_0 , de extensão vocal, de intensidade e de velocidade de fala (BEZOOYEN, 1984). Os altos valores da média de F_0 , na raiva, confirmam os dados de Laukka (2004), no inglês e no sueco. Laukka (2004) também identificou alto valores de *jitter* e de intensidade vocal, se comparados às demais emoções e à fala neutra. As medidas de *jitter* altas significam que, nessas emoções, apresentam rápido fechamento de glote. Essa rapidez causa maior esforço vocal e maior tensão muscular (*alegria*, irritação, ansiedade) (BEZOOYEN, 1984).

Para o *medo*, o comportamento de F_0 foi atípico, pois cada atriz interpretou de uma maneira diferente, quanto à média de F_0 , como podemos observar nos gráfico 1. Em uma pesquisa, Paeschke e Sendlmeier (2000) obtiveram média de F_0 tão baixa como na *tristeza* (assim como a atriz 1) e a

variação de F_0 quase tão alta quanto a *alegria* e a *raiva* (como as atriz 1 e 3). Já em outro estudo, Laukka (2004) encontrou baixa variação de F_0 para o *medo*.

Ao interpretar a emoção do *medo*, nota-se que cada atriz flexionou um padrão de F_0 distinto: enquanto a F_0 média da atriz 1 foi de 342 Hz, da atriz 2 foi 226 Hz e da atriz 3 foi de 180 Hz, ou seja, em faixas de frequências variadas. Nos dados de PE, a média da F_0 foi ligeiramente mais baixa que *alegria* (NUNES, 2009).

No gráfico 1, observa-se que a fala neutra ancora o comportamento da F_0 com relação ao padrão da F_0 das emoções, ou seja, assim como na fala neutra, a atriz 3 mantém os menores valores de F_0 em todas as emoções, o mesmo fenômeno ocorre com a atriz 1 (salvo no *medo*) e a atriz 3. Esse padrão de comportamento indica uma tendência fisiológica individual, uma vez que cada atriz mantém um padrão de vibração de pregas vocais. No entanto, essa variação individual não interfere em um padrão mais geral característico de cada emoção.

Quanto à variação⁴ de F_0 (diferença entre o valor máximo e mínimo da F_0 da curva), é possível visualizar no gráfico 2:

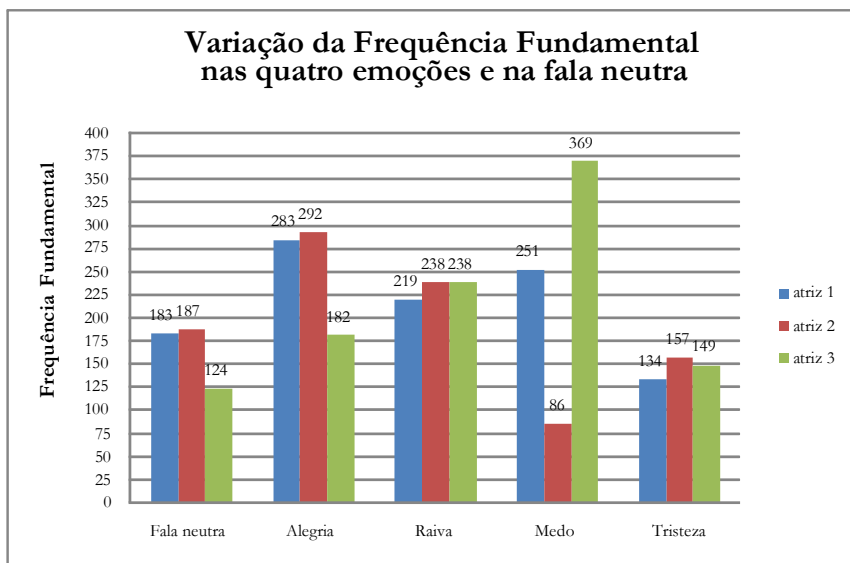


Gráfico 2 – Variação da F_0 das três atrizes nas quatro emoções e na fala neutra

⁴ Essa medida também é denominada de gama tonal e de extensão de F_0 .

Observa-se que a variação de F_0 na fala neutra das atrizes 1 e 2 são semelhantes. Apenas a atriz 3 produz uma extensão de F_0 mais baixa que as demais. A variação de F_0 na *alegria*, mantém os mesmos padrões que a da fala neutra, porém em faixas de frequência mais altas.

A *alegria* e a *raiva* apresentam alta variação de F_0 , sendo que na *raiva*, as três atrizes produzem variação de F_0 próximas entre si (219, 238 e 238 Hz, respectivamente). Na *alegria*, a atriz 1 teve uma variação de F_0 de 283 Hz, a atriz 2 de 292 e a atriz 3, com uma variação mais baixa, de 182 Hz.

Na investigação de Vassoler e Martins (2013), ao comparar a *raiva* com a fala neutra, vê-se que a primeira apresenta-se em faixas de frequência mais altas, ao passo que a segunda mantém-se em faixas de frequências médias, em conformidade com os dados do presente estudo. Os autores também apontaram que o coeficiente de variação entre as repetições realizadas pelas atrizes foi menor na *raiva*, se comparado com a fala neutra, mostrando mais consistência entre as repetições da *raiva*. O desvio padrão de ambas as situações também confirma essa consistência. Os valores absolutos médios da F_0 são semelhantes entre a *raiva* e a fala neutra, entretanto a diferença parece estar na variação de F_0 – *pitch ranges* – corroborando os dados desta pesquisa (gráficos 1 e 2).

Já no *medo*, a variação de F_0 não teve um padrão entre as atrizes, do mesmo modo como ocorreu nos valores da média de F_0 . As variações de F_0 na *tristeza* se mantiveram em faixas de frequências baixas, se comparadas com as demais emoções. Laukka (2004) identificou como alta variação de F_0 nas emoções *alegria*, *medo* e *raiva*, nessa ordem. A *tristeza* apresentou baixa variação de F_0 , inclusive menor que a fala neutra, assim como a fala emotiva do PB.

Também no inglês, Morley et al. (2011) identificaram ampla variação de F_0 na fala alegre e *raiva* e baixa variação de F_0 na fala com *medo* e com *tristeza*. Os dados revelaram também que apesar das grandes diferenças na amplitude F_0 entre as emoções, a localização do pico de F_0 mantém-se relativamente estável em contextos diferentes.

O gráfico 3 expõe as médias da F_0 mínima e máxima nas cinco situações: *alegria*, *medo*, *tristeza*, *raiva* e fala neutra:

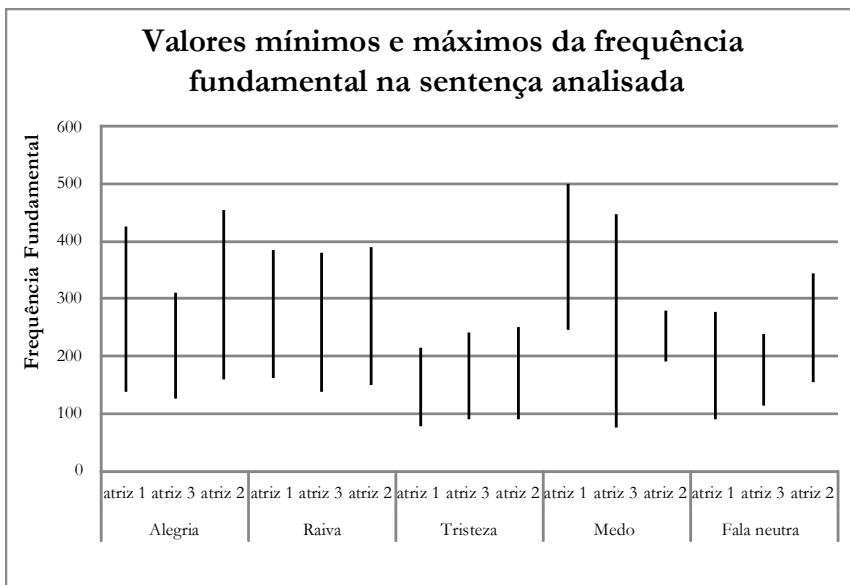


Gráfico 3: F₀ máximo e mínimo ao longo da sentença (referente às cinco repetições)

Os valores mínimos de F₀ da *alegria* (139, 159 e 125 Hz) estão na faixa dos 100 Hz, já a frequência fundamental máxima está acima de 400 Hz (com exceção da atriz 3). Para a *raiva*, todas as atrizes tiveram F₀ máxima e mínima semelhantes, assim como na *tristeza*. A interpretação do *medo* influiu diversas caracterizações de F₀ mínimo e máximo, não mostrando regularidades entre as atrizes.

Verifica-se que os valores mínimos de F₀ não permitem distinguir emoções entre si, assim como na fala do PE (RODRIGUES, 2007). Já nos valores da F₀ máxima, a *alegria* foi a mais alta (duas atrizes), em seguida, a *raiva*. No PE, a fala com *raiva* atingiu os valores de F₀ máxima mais elevada. Apesar de a *raiva* atingir valores altos de F₀, nos dados de Rodrigues (2007) houve maior variação nas emoções *alegria* e *medo*. No PE, Nunes (2009) comparou os dados da F₀ máxima, mínima e mediana na fala emotiva e verificou que existem diferenças significativas entre todos os pares de emoções, com exceção ao desespero-*medo*, desespero-*alegria*, *medo-alegria* e fala neutra-*tristeza*.

Portanto, as médias e a variação de F_0 foram importantes para diferenciar a *alegria* e a *tristeza*. A variação de F_0 é importante na *raiva*, pois parece afastá-la da *alegria*. Já o *medo* apresentou diferenças entre a interpretação das atrizes entre si, o que não permite inferir a importância da F_0 nessa emoção.

No PE, a *raiva* apresenta o valor mais alto de F_0 , seguido da *alegria*. O valor de F_0 mais baixo, no PE, é da *tristeza*, aproximando da F_0 da expressão neutra (NUNES, 2009). Emoções distintas como *tristeza*, *alegria*, e fala neutra se aproximam no PE, ou seja, acontece de forma semelhante, o que não difere do PB. A interpretação dessa autora para essa coincidência é que o povo português (residentes em Portugal) não é “povo muito efusivo, ou expansivo na expressão da alegria” (NUNES, 2009). Esses achados parecem estar vinculados a expressões culturais, parece que o povo brasileiro expressa com mais evidência suas emoções alegres e tristes.

Conclusões

O estudo da fala emotiva é de natureza multidisciplinar, agregando a Psicologia, as Ciências da Linguagem, a Medicina, o Processamento de Sinal, a Engenharia Computacional, as Ciências Cognitivas e a Fonoaudiologia.

Em geral, os estudos têm demonstrado que os estados emocionais influenciam as mudanças fisiológicas do organismo, tais como a respiração, a fonação e a articulação que, por sua vez, interferem nos parâmetros acústicos da fala, nos permitindo, assim, estudar as emoções por meio da fala. Por se tratar de fenômenos extremamente complexos, com a participação de diversos componentes cerebrais, o método da fala atuada mostrou-se mais eficaz para o objetivo proposto, já que possibilita minimamente determinar a emoção a ser analisada. Nunes (2009) comparou a emoção atuada com a espontânea e identificou que, em termos de F_0 , os valores apresentam ligeiras diferenças, portanto, a fala atuada pode inferir o que acontece nas expressões espontâneas.

Os valores de F_0 , nesse estudo, vislumbraram características emocionais entre a *alegria*, a *raiva* e a *tristeza*. No entanto, para o *medo*, esse parâmetro acústico não pareceu ser eficaz, pois parece que essa emoção sofre outras influências acústicas e fisiológicas. Portanto, a hipótese inicial desse trabalho foi parcialmente corroborada à medida que diferenciaram as emoções *alegria*, *tristeza* e na *raiva*.

Do ponto de vista da F_0 , a fala emotiva do PB e a do PE ocorre de forma semelhante na *raiva*, com valores altos de F_0 e na *tristeza*, com valores baixos de F_0 . No entanto, a *alegria* se comporta diferente, e está ligada, possivelmente, à expressividade dos falantes das línguas.

A F_0 parece ser um aspecto significativo nas emoções, porém, em trabalhos futuros, sugere-se que outros parâmetros acústicos devam ser investigados, tais como a qualidade vocal, a intensidade e a velocidade de fala, entre outros.

Referências

ARAÚJO, S. A.; GRELLET, M.; PEREIRA, J. C. Normatização de medidas acústicas da voz normal. *Revista Brasileira de Otorrinolaringologia*, v. 68, n. 4, p. 540-544, 2002.

BANSE, R.; SCHERER, K. R. Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, v. 70, n. 3, p. 614-636, 1996.

BEZOOYEN, R. van. *Characteristics and recognizability of vocal expressions of emotion*. Dordrecht, Netherlands: Gruyter GmbH; Co. KG, Walter (Netherlands Phonetic Archives Ser.)

BOERSMA, P.; WEENINK, D. *Praat: doing phonetics by computer* (versão 5.3.52). Disponível em: <<http://www.praat.org/>>. Acesso em: 12 jul. 2013.

CAMPBELL, N. Databases of emotional speech. In: COWIE, R.; DOUGLAS-COWIE, E.; SCHRÖDER, M. (Orgs.). *Proceedings of the ISCA Workshop on Speech and Emotion*. Belfast: Ireland, 2000.

CAMPBELL, N. Building a corpus of natural speech – and tools for the processing of expressive speech – the JST CREST ESP Project. *Proceedings of Eurospeech 2001*. Aalborg: Denmark, 2001.

CRUTTENDEN, A. *Intonation*. London: Cambridge University Press, 1986.

CORREIA, P. C. G. *Sob o signo das emoções: expressões faciais e prosódia em indivíduos com perturbação vocal*. 2007. Dissertação. (Mestrado em Ciências da Fala) – Universidade Católica Portuguesa, Alcoitão.

DARWIN, C. R. *The expression of the emotions in man and animals*. London: John Murray, 1872.

EKMAN, P. An argument for basic emotions. *Cognition and Emotion*, Palo Alto, CA, v. 6, n. 3/4, p. 169-200, 1992.

FIGUEIREDO, R. M. A eficácia de medidas extraídas do espectro de longo termo para a identificação de falantes. *Caderno de Estudos Linguísticos*, Campinas, v. 25, p. 129-160, 1993.

ERICKSON, D.; YOSHIDA, K.; MENEZES, C.; FUJINO, A.; MOCHIDA, T.; SHIBUYA, Y. Exploratory study of some acoustic and articulatory characteristics of sad speech. *Phonetica*, Freiburg, v. 63, p.1-25, 2006.

GUSTAFSON-CAPKOVÁ, S. Emotions in speech: Tagset and acoustic correlates. *Speech Technology 1*, Stockholm, p. 1-11, 2001.

JOHNSTONE, T.; SCHERER, K. R. Vocal communication of emotion. In: LEWIS, M.; HAVILAND, J. (Orgs.). *Handbook of emotions*. 2. ed. New York: Guilford Press, 2000. p. 220-235.

LAUKKA, P. Vocal expression of emotion: Discrete-emotions and dimensional accounts. In: *Comprehensive summaries of Uppsala dissertations from the faculty of social sciences*, Acta Universitatis Upsaliensis, Uppsala, Sweden, p. 1-80, 2004.

MELO, A. I. M. T. *Emoções no período escolar: estratégias parentais face à expressão emocional e sintomas de internalização e externalização da criança*. 2005. Dissertação (Mestrado em Psicologia Clínica) – Universidade do Minho, Braga.

MORLEY, E.; VAN SANTEN, J.; KLABBERS, E.; KAIN, A. F₀ range and peak alignment across speakers and emotions. *Proceedings of Acoustics, Speech and Signal Processing*, 2011.

NUNES, A. M. B. *Voz e emoção em português europeu*. 2009. Dissertação (Mestrado em Ciências da Fala) –Universidade Católica Portuguesa, Alcoitão.

OLIVEIRA, A. M.; RAPOSO DE MEDEIROS, B. R. *Aspectos prosódicos de quatro emoções na voz falada*. In: SIMPÓSIO DE COGNIÇÃO E ARTES MUSICAIS, 6., 2010, Rio de Janeiro.

ORLIKOFF, R. F.; KAHANE, J. C. Laryngeal structure and function. In: LASS, N. J. (Orgs.). *Principles of experimental phonetics*. St. Louis, MO: Mosby, 1996. p. 112-181.

PAESCHKE, A.; SENDLMEIER, W. F. Prosodic characteristics of emotional speech: Measurements of fundamental frequency movements. *SpeechEmotion* 2000, p. 75-80, 2000.

PATRIK, N. J.; VÄSTFJÄLL, D. All emotions are not created equal: Reaching beyond the traditional disputes. *Behavioral and Brain Sciences*, v. 31, n. 5, p. 600-621, 2008.

RODRIGUES, A. J. M. *As emoções na fala*. 2007. Dissertação (Mestrado em Ciências da Fala) – Universidade Católica Portuguesa, Alcoitão.

SCHERER, K. R. Nonverbale Kommunikation. *Nonverbal communication*. Zürich: Kindler, 1979.

SCHERER, K. R. On the nature and function of emotion: a component process approach. In: SCHERER, K. R.; EKMAN, P. (Orgs.) *Approaches to emotion*. Hillsdale, NJ: Lawrence Erlbaum, 1984. p. 293-318.

SCHERER, K. R. Vocal affect signalling: A comparative approach. In: ROSENBLATT, J.; BEER, C.; BUSNEL, M.; SLATER, P. J. B. (Orgs.). *Advances in the study of behavior*. New York: Academic Press, 1985. p. 189-244.

SCHERER, K. R. Vocal affect expression: A review and a model for future research. *Psychological Bulletin*, Washington, v. 99, p. 143-165, 1986.

SCHERER, K. R. Vocal measurement of emotion. In: PLUTCHIK, R.; KELLERMAN, H. (Orgs.). *Emotion: Theory, research, and experience*.

The measurement of emotion. New York: Academic Press, 1989.
p. 233-260.

SCHERER, K. R. Appraisal considered as a process of multi-level sequential checking. In: SCHERER, K. R.; SCHORR, A.; JOHNSTONE, T. (Orgs.). *Appraisal processes in emotion: Theory, methods, research*. New York; Oxford: Oxford University Press, 2001. p. 92-120.

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

SOBIN, C.; ALPERT, M. Emotions in speech: the acoustic attributes of fear, anger, sadness and joy. *Journal of Psycholinguistic Research*, v. 28, n. 4, p. 347-365, 1999.

STIBBARD, R. M. Vocal expression of emotions in non-laboratory speech: An investigation of the reading/leeds emotion in Speech Project Annotation Data. 2001. University of Reading, UK. Unpublished PhD Thesis.

VAZ, N. M. Idéias para uma nova imunologia. *Ciência Hoje II*, v. 7, p. 32-38, 1983.

VASSOLER, A. M. O. *O papel da frequência fundamental da fala na representação teatral das emoções*. 2011. Dissertação (Mestrado em Semiótica e Linguística Geral) – Universidade de São Paulo, São Paulo.

VASSOLER, A. M. O. ; MARTINS, M. V. M. A entoação em falas teatrais: uma análise da raiva e da fala neutra. *Estudos Linguísticos*, São Paulo, v. 1, p. 9-18, 2013.

WILLIAMS, C. E.; STEVENS, K. N. Emotions and speech: Some acoustical correlates. *Journal of Acoustics Society of America*, v. 52, p. 1238-1250, 1972.