UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA CIÊNCIAS E LETRAS DE RIBEIRÃO PRETO
DEPARTAMENTO DE PSICOLOGIA E EDUCAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM PSICOBIOLOGIA

***Aspectos psicológicos da voz e seus correlatos
acústicos***

Adriana Campos Balieiro Panico

Dissertação apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto–USP, como parte das exigências para a obtenção do título de Mestre. Área: Psicobiologia.

**Ribeirão Preto-SP
2001**

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA CIÊNCIAS E LETRAS DE RIBEIRÃO PRETO
DEPARTAMENTO DE PSICOLOGIA E EDUCAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM PSICOBIOLOGIA

***Aspectos psicológicos da voz e seus correlatos
acústicos***

Adriana Campos Balieiro Panico

Dissertação apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto–USP, como parte das exigências para a obtenção do título de Mestre. Área: Psicobiologia.

**Ribeirão Preto-SP
2001**

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA CIÊNCIAS E LETRAS DE RIBEIRÃO PRETO
DEPARTAMENTO DE PSICOLOGIA E EDUCAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM PSICOBIOLOGIA

***Aspectos psicológicos da voz e seus correlatos
acústicos***

Adriana Campos Balieiro Panico

Dissertação apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto–USP, como parte das exigências para a obtenção do título de Mestre. Área: Psicobiologia.

Orientador: Prof. Dr. Sérgio Sheiji Fukusima

**Ribeirão Preto-SP
2001**

*Para Gabriela e Isabela,
vocês são o encanto da minha vida.*

AGRADECIMENTOS

Ao Prof. Dr. Sérgio Seiji Fukusima, pela disponibilidade, paciência e confiança.

Ao Prof. Dr. Rui Celso Mamede, pelo acesso ao Serviço de Cirurgia de Cabeça e Pescoço.

Ao Prof. Dr. Luiz Fernando Coláfêmima, pelo apoio oferecido quando cheguei ao Hospital das Clínicas.

Ao Dr. Walther Luis Jaeger, por repartir seus livros, seus programas e seu saber.

Ao Marcos Frateschi e à Patrícia Lauretti, pelas gravações feitas e, especialmente ao Marcos pelas “dicas” no início do trabalho.

Aos repórteres e apresentadores, que “emprestaram” suas vozes e tornaram este trabalho possível, ele é de vocês!

Aos profissionais, que gentilmente responderam ao questionário.

Aos amigos Fernanda e Fernando, companheiros que nunca me faltaram e resolveram coisas que me pareciam impossíveis.

Aos meus pais, Adalgiza e José Antônio, e à minha irmã Luciana, por me substituírem tantas vezes no papel de mãe.

Ao meu irmão Junior, por ter me emprestado seus olhos, quando os meus já não me eram fiéis, na revisão de dados, contas, nomes e tudo o que se fizesse necessário.

À amiga Marisa, por dividir este sonho desde o início, quando nem todos acreditavam que ele poderia ser nosso. O esforço valeu a pena!

À amiga Ana Irene, pela capacidade de simplificar as coisas que me pareciam mais difíceis.

À amiga Renata, presente em todos os momentos.

À amiga Ângela, pela revisão cuidadosa deste trabalho.

À Capes, pelo apoio financeiro a este trabalho.

ÍNDICE

1. RESUMO	01
2. ABSTRACT	02
3. INTRODUÇÃO	03
4. OBJETIVOS	17
5. MÉTODO	18
5.1. Participantes	18
5.2. Material e Equipamentos	18
5.3. Procedimento	19
6. RESULTADOS	25
6.1. Análise Acústica	25
6.2. Julgamentos de Agradabilidade	27
6.3. Julgamentos de Confiabilidade	32
6.4. Julgamentos de Naturalidade	37
7. DISCUSSÃO	44
8. CONCLUSÃO	62
9. REFERÊNCIAS	63
10. ANEXOS	68

Neste estudo investigou-se a relação entre aspectos psicológicos (agradabilidade, confiabilidade e naturalidade) de vozes profissionais e parâmetros acústicos (intensidade, frequência e duração). Os aspectos psicológicos foram previamente definidos através de um questionário aplicado a profissionais da voz. Trechos de fala de repórteres e apresentadores de rádio e televisão (masculinas e femininas) foram gravados em estúdio profissional e analisados acusticamente por meio do programa de computador *Dr. Speech 3.0*. Posteriormente as gravações foram apresentadas pelo método de comparação aos pares a ouvintes leigos para que estimassem o quão agradável, confiável e natural eram as vozes preferidas em relação às preteridas. As somas das pontuações dos julgamentos estabeleceram escalas decrescentes que ordenavam as vozes mais agradáveis, mais confiáveis e mais naturais; e os valores destas escalas foram correlacionadas com os parâmetros acústicos. Também, para cada aspecto psicológico, foi determinada uma matriz de índices de correlação de *Pearson* entre as somatórias das pontuações das vozes, que foi submetida a uma análise de escala multidimensional (*MDS*). Esta análise indicou que as vozes podem ser descritas adequadamente por três dimensões em cada aspecto psicológico. Os valores das escalas dessas dimensões também foram correlacionados com os parâmetros acústicos. Concluiu-se que os julgamentos de agradabilidade, confiabilidade e naturalidade das vozes guardam interdependência entre si. O aspecto confiabilidade teve duas dimensões fortemente correlacionadas com frequência e intensidade. Para agradabilidade e naturalidade houve concorrência de vários parâmetros acústicos. Sugere-se que no trabalho de imitação de voz se considerem os parâmetros acústicos que se correlacionam fortemente com as dimensões dos aspectos psicológicos.

ABSTRACT

This study investigated the relation between the psychological aspects (pleasantness, reliability and naturalness) of professional voices and acoustic parameters (intensity, frequency and duration). The psychological aspects were previously determined through a questionnaire applied to those who use their voices professionally. Parts of the speeches of reporters and radio and television presenters (male and female) were recorded in a professional studio and analyzed acoustically by means of the computer program *Dr. Speech 3.0*. Subsequently the recordings were presented, by the comparison method, in pairs, to laymen so that they could estimate how pleasant, reliable and natural were the preferred voices in relation to the neglected ones. The sum of the ratings established increasing scales that assorted the most pleasant, most reliable and most natural voices; and the values of those scales were correlated with the acoustic parameters. For each psychological aspect, as well, a *Pearson* correlation index matrix was determined between the total amounts of the voice ratings, and was submitted to a multidimensional scale analysis (*MDS*). This analysis indicated that three dimensions in each psychological aspect might adequately describe the voices. The values of the scales of these dimensions were also correlated with the acoustic parameters. The conclusion was that the pleasantness, reliability and naturalness ratings of the voices maintain interdependence among them. The reliability aspect had two dimensions that were strongly correlated with frequency and intensity. For pleasantness and naturalness aspects there was a balance among several of the acoustic parameters. It is suggested that the acoustic parameters that strongly correlate with the dimensions of the psychological aspects should be considered on the voice placement work.

O desenvolvimento de pesquisas em ciências jovens como a Fonoaudiologia, na tentativa de construção de conhecimentos próprios, traz desafios na determinação do melhor caminho a trilhar, uma vez que referências nem sempre são suficientes para indicá-lo.

Dentro de uma ciência nova, um interesse ainda mais recente é o estudo da voz profissional. De uma ciência que em sua concepção nasceu para tratar patologias, a intervenção com os indivíduos que usam a voz profissionalmente trouxe a necessidade de um novo pensamento que nos remetesse à demanda dessas pessoas que procuram o fonoaudiólogo na intenção de aperfeiçoar ou adequar sua voz e fala a suas profissões.

O número de pesquisadores envolvidos em estudar a voz, possivelmente ainda constitui um contingente muito pequeno, se comparado com aqueles envolvidos em pesquisar a linguagem (Chun & Navarro, 1998).

Os avanços na compreensão da produção vocal, dos aspectos fisiológicos envolvidos no processo de fonação muito têm contribuído para os estudos da voz. No entanto, analisar a produção vocal do profissional da voz envolve também considerações dos aspectos envolvidos na história do sujeito e na sua profissão específica.

Behlau e Ziemer (1987) disseram ser a voz uma das extensões mais fortes de nossa personalidade e se aguçassemos nossos sentidos reconheceríamos que esta extensão é mais profunda em sua dimensão não-verbal (altura, intensidade, qualidade vocal, entre outros) do que na verbal (estrutura lingüística). Além disso, em todas as situações de emissão pode-se ter vários níveis de análise, de leitura vocal: leitura dos parâmetros físicos, psicológicos, sociais, culturais e educacionais de um determinado falante.

O fato de ouvirmos desencadeia automaticamente projeções, sentimentos, emoções e julgamentos. A fala que comunica envolve o outro, considera o outro. Esse outro, o ouvinte, entende o que se fala e interpreta conteúdos também pela análise da qualidade vocal do falante.

Qualidade vocal é o termo empregado para designar o conjunto de características que identificam uma voz humana. Ela se relaciona à composição dos harmônicos da onda sonora e à impressão total criada por uma voz. Embora ela varie de acordo com o contexto da fala e com as condições físicas e psicológicas do indivíduo, há sempre um padrão básico de emissão que a identifica (Behlau & Pontes, 1995).

Embora a qualidade vocal seja um tópico de grande relevância e interesse não só para lingüistas e fonoaudiólogos como para todos aqueles interessados em estudar os aspectos comunicativos da linguagem, não registra um grande número de pesquisas sobre ela. A explicação para tal fato pode ser a ausência até a década de 70, de modelos apropriados e de aparelhos para análise instrumental de voz, segundo Madureira (1992), *apud* (Chun & Navarro, 1998).

A análise perceptiva da qualidade vocal, que envolve necessidade de treino para desenvolver atenção auditiva para que se avalie o sistema de ressonância, a sensação psicoacústica de intensidade (*loudness*), a sensação psicoacústica de frequência (*pith*), a articulação, o ritmo e a respiração na produção da voz, é a mais referida no meio fonoaudiológico e reconhecida como fundamental por unanimidade pelos que atuam na intervenção em vozes profissionais.

Kreiman, Gerrat, Precoda e Berke (1992), Kreiman, Gerrat, Kempter, Erman, e Berke (1993) e Wuyts, De Bodt, e Van de Heyning, (1999) sustentam que embora a qualidade vocal possa ser medida de forma objetiva, ela é perceptual por natureza. No entanto, mesmo os mais treinados ouvintes podem discordar completamente ao julgar o que ouvem. Quando os julgamentos subjetivos se juntam aos objetivos, por exemplo, à análise acústica, os dados tornam-se mais fidedignos.

Gerrat, Kreiman, Barroso e Berke (1993) compararam padrões internos e externos para o julgamento de qualidade de voz. Usaram para esse estudo vozes sintetizadas que foram julgadas por doze clínicos em relação à rouquidão. Eles concluíram que as medidas objetivas podem elucidar várias questões que, sozinhas, tornam-se difíceis de serem interpretadas, para diagnosticar, tratar e pesquisar sobre qualidade vocal.

Portanto, embora a análise subjetiva da voz feita através da percepção auditiva seja valiosa, ainda que forneça dados consensuais, nem sempre possibilita uma interpretação mais exata do estudo. Isto pode gerar problemas de terminologia, de confusões na interpretação dos resultados e não permitir comparação de trabalhos de diferentes pesquisadores. A análise acústica é um parâmetro objetivo que tem se demonstrado extremamente útil para o estudo da voz, por permitir que se estabeleçam dados normativos de avaliação (Rodrigues, Behlau & Pontes; 1994).

Os laboratórios de voz e fala permitem a compreensão dos dados da esfera perceptiva, fornecendo dados objetivos por meio de representações visuais de energia acústica obtidos e valores matemáticos. A dimensão acústica permite a mediação entre os níveis fisiológico e perceptivo da produção da voz e da fala (Pinho & Camargo; 1998).

O primeiro interesse dos pesquisadores no estudo da análise acústica tem sido estabelecer dados diferenciais entre patologia e normalidade. Para tanto, faz-se uso de vogal sustentada como amostra.

Em 1984 Behlau estabeleceu por análise computadorizada padrões de frequência fundamental para homens, mulheres e crianças, numa amostra de 90 sujeitos da população da cidade de São Paulo (Behlau & Pontes; 1995).

Pinho, Pontes e Ganança (1988), demonstraram que o comportamento do trato vocal na emissão das vogais orais apresenta modificações de acordo com a vogal emitida e mantém relações estreitas com o primeiro e segundo formante.

As pesquisas brasileiras mais recentes que envolvem parâmetros acústicos têm levantado dados que referem medidas específicas para mensurar a qualidade vocal. A proporção harmônico-ruído, medida através de amostras com vogais sustentadas e analisadas pelo programa *Soundscope II*, da *GW Instruments*, mostrou ser um parâmetro acústico promissor para se estabelecer dados normativos de avaliação (Rodrigues *et al*, 1994).

Outros parâmetros acústicos como *jitter* (perturbação de frequência) e *shimmer* (perturbação de amplitude) têm sido estudados em amostras de vogais sustentadas, utilizando-se o programa computacional de análise acústica *Dr. Speech* (Steffen & Moschetti, 1997).

Pegoraro-Krook e Castro (1994) obtiveram dados normativos de frequência fundamental para brasileiros investigando tarefas de fala.

Nepomuceno (1984) por meio de análise espectral relacionou qualidade dos sons da fala ao equilíbrio ideal proporcionado pelos órgãos moduladores (ressoadores).

Na literatura internacional, Bielańowicz, Kreiman, Gerratt, Dauer e Berke (1996) discutem a validade do uso dos programas acústicos *Cspeech* (ver. 4.0; Paul Milenkovic, Madison, WI), Computerized Speech Laboratory (*CSL*; Kay Elemetrics, Pine Brook, NJ) e *Soundscope* (ver. 1.09; GW Instruments, Cambridge, MA), considerando diferentes medidas. Todos os programas permitiam a importação de sinal digitalizado, o que evitou efeitos de ruídos introduzidos na amostra por transferência de sinais analógicos para digitais. Concluíram que a medida de frequência fundamental é a mais fiel em vozes normais, mas que nas vozes com disfonias médias e severas o *shimmer* é a melhor medida.

Eskenazi, Childers e Hicks (1990) observaram correlatos acústicos de qualidades vocais como rouquidão, sopro e aspereza, analisando o som emitido por vogal sustentada em sujeitos normais e portadores de distúrbios vocais. Concluíram que várias qualidades vocais podem ser preditas com medidas acústicas e o grau de confiabilidade aumenta quando se faz um grande número de medidas.

Rantala e Vilkman (1999) relacionaram queixas vocais de professores com uso de frequência fundamental e nível de pressão sonora, analisando não só vogal sustentada como também trechos de fala. Neste estudo a correlação entre as queixas vocais, ou seja, as sensações causadas pelo uso da voz, e as medidas acústicas ficam evidenciadas, o que levou os autores a sugerirem que haja a necessidade de pesquisas direcionadas aos profissionais da voz que complementem os exames de rotina.

Todos esses trabalhos têm sido desenvolvidos na busca de estabelecer padrões objetivos que possam ajudar na avaliação vocal e na conduta terapêutica com os pacientes portadores de disfunções vocais.

Nessas pesquisas há a preocupação principalmente em diferir normalidade de patologia vocal. Quando se trata de estudos que consideram correlações acústicas com aspectos subjetivos, o que mais se estuda é o aspecto naturalidade, uma vez que as vozes sintetizadas são uma realidade no mundo atual.

Nessa linha de pesquisa, Klatt e Klatt (1990) discutiram aspectos que diferissem vozes normais de vozes sintetizadas, com o objetivo de sintetizar vozes aparentemente mais naturais.

Ramos (1996) comparou a emissão de locutores de duas emissoras de rádio de São Paulo de estilos diferentes. Em sua análise usou um sonógrafo fabricado pela *Kay Elemetrics Inc.*, que permite analisar medidas de intensidade, duração dos segmentos, entoação e formantes. Embora tenha havido em seu trabalho medidas feitas através de percepção auditiva, seu trabalho foi pioneiro em correlacionar dados acústicos com produção de diferentes interpretações. Em suas conclusões apontou que para que os locutores obtivessem estilos diferentes, que deveriam estar de acordo com o perfil da emissora, faziam-se mudanças em frequência e duração de fonemas.

Não foram encontradas outras pesquisas brasileiras que tenham tido como objeto de estudo correlacionar características acústicas com aspectos psicológicos da voz humana.

Embora a fala forneça conteúdos semânticos em suas palavras e frases, parece claro que a comunicação verbal carrega em si informações que podem enfatizar ou contradizer sua mensagem. A impressão que a voz causa no ouvinte é baseada no seu som e constitui a imagem vocal do falante. E esta é a impressão que os ouvintes formam do indivíduo baseado no som de sua voz;

como se diz alguma coisa é por vezes mais importante e mais poderoso do que **o quê** se diz. (MacCoy, 1996).

A forma como é feita a impositação (emissão correta da voz), pode fazer com que uma mesma frase assumia sentidos diferentes.

Questões como se os ouvintes inferem emoção através da voz e quais seriam os padrões de expressão vocal para emoções específicas foram estudadas por Banse e Scherer (1996). Analisando os julgamentos da performance de atores em relação a voz, gestos e, por fim, voz e gestos simultaneamente, eles concluíram que os parâmetros vocais não só indicam a intensidade das emoções como também seus aspectos qualitativos.

As características da voz que podem ser responsáveis por inferência de emoções parecem ser as diferenças na intensidade, na altura e na entonação.

Profissionais da voz, como cantores, professores, atores, radialistas, repórteres, apresentadores, tentam controlar essa variáveis com treino, buscando a voz mais indicada para sua profissão. A grande dificuldade que se encontra é que, o que se busca como resultado final, são aspectos psicológicos que aquela voz vai proporcionar ao ouvinte. O profissional da voz espera que sua voz agrade ao ouvinte, faça com que ele esteja atento na escuta daquela que será sua mensagem.

Atentos a essa necessidade, Rossiter e Howard (1996) desenvolveram um novo programa computacional de análise acústica específico para vozes profissionais, especialmente cantores, portanto as medidas acústicas possibilitam análise especial para vibratos e formantes do cantor, por exemplo. Este programa é usado para treino de cantores, não havendo no estudo correlações dos parâmetros acústicos com aspectos psicológicos. No entanto,

o mais comum é que os programas sejam mais gerais e que o profissional que o manuseia faça as adaptações necessárias ao seu uso.

Novak e Vokrail (1995) apontaram parâmetros acústicos para avaliar a voz de futuros profissionais da voz, pesquisando sujeitos que iniciariam seus estudos em canto. Propuseram normas rigorosas para que se avalie vozes de futuros cantores; considerando-se que para atores as normas podem não ser tão severas, entretanto a indulgência pode ser perigosa para o futuro de profissionais da voz.

O estabelecimento de parâmetros acústicos que dirijam o trabalho de imitação vocal de indivíduos que precisam usar a voz profissionalmente, respeitando é claro os limites individuais, é um facilitador na abordagem com esses profissionais.

Como a voz profissional nos remete a uma série de profissões que não apresentam entre si o mesmo tipo de demanda, como por exemplo, ter ou não o aspecto artístico envolvido (cantores, atores), ter ou não o aspecto volume de voz como importante (professores, trabalhadores do pregão da bolsa) e tantas outras peculiaridades de cada profissão, o interesse desse estudo foi dirigido a um tipo específico de profissional da voz, mais precisamente a repórteres e apresentadores de televisão.

Qual é o tipo de impressão que os repórteres e apresentadores precisam causar nos seus ouvintes para que sua comunicação seja efetiva? Através de experiência clínica de sete anos atuando nesse campo, observou-se que alguns aspectos eram buscados por esses profissionais de forma muito constante. Assumiu-se então que esses aspectos psicológicos pareciam ser fundamentais para que esses profissionais fossem bem sucedidos: a

agradabilidade, a naturalidade e a confiabilidade que suas vozes imprimiam em seus ouvintes.

Embora não tenha sido encontrado na literatura estudada dados sobre aspectos psicológicos necessários a um bom desempenho vocal, a experiência clínica tem demonstrado que a voz de um repórter e/ou apresentador deve ser agradável ao ouvinte, o que parece ser senso comum; confiabilidade é condição para qualquer trabalho jornalístico bem sucedido e por fim naturalidade parece ser o que reitera as duas anteriores.

Encontrar esses aspectos não foi difícil, mas saber o significado de cada um deles foi um desafio. Relacioná-los com aspectos acústicos foi ainda mais instigante, pois possibilitaria o treino vocal para profissionais da voz, tendo como retroalimentação os dados acústicos apontados pelo computador.

Não se pretende com esses achados desvalorizar o trabalho clínico subjetivo, mas sim acrescentar a ele. Hollien (1993) ao apontar a necessidade de que o “professor” de voz tenha dentro de suas habilidades conhecimentos de fisiologia, psicologia, um “excepcional treinamento auditivo”, entre outros, não deixou de mencionar a importância do conhecimento de acústica. Segundo ele, é preciso aprender a conviver e entender novos dados que a acústica nos traz, por exemplo, em relação aos registros vocais, antes definidos somente por “sensações” dos cantores.

Não se trata, portanto, de um debate entre análises subjetivas e objetivas, mas sim do trabalho conjunto dessas duas vertentes, que se complementam elucidando uma a outra.

Há bem pouco tempo isso seria impossível, já que todo o trabalho fonoaudiológico era feito somente através de dados subjetivos. Os

conhecimentos trazidos pelo estudo da teoria acústica de produção de fala têm trazido a cientistas, professores e clínicos que trabalham com a voz humana a possibilidade de medidas objetivas tanto da voz como da função do trato vocal (Kent, 1993).

Atualmente, o laboratório de análise acústica tem permitido melhor compreensão dos achados perceptivos auditivos, constituindo-se num ponto de referência para o tratamento dos distúrbios de voz e fala (Pinho & Camargo, 1998).

A utilidade do laboratório de fala e voz na prática otorrinolaringológica e fonoaudiológica tem sido amplamente discutida na literatura internacional.

Crevier-Buchman, Monfrais-Pfauwadel, Begue, Lauga-Houdoyer, Laccourreye e Brasnu (1993) fizeram uma revisão de três programas computacionais (*Signalize*, *Mac Speech Lab* (MIT) e *Sound Edit*) e concluíram que eles são limitados para análise de vozes com problemas severos, mas que são particularmente interessantes para vozes profissionais como as de cantores e atores, e que a análise computadorizada é um bom complemento à avaliação clínica.

Nieto, Del Palacio, Lorenzo, Vegas e Cobeta (1995) discutiram noções básicas para o uso de análise acústica e que fatores deveriam ser considerados para a escolha de equipamento e suas aplicações clínicas.

Casmerides (1988) caracterizou um grupo formado por fonoaudiólogos brasileiros, formadores de opinião e usuários de programas acústicos e concluiu que os critérios de escolha do equipamento eram preço, facilidade de manuseio e necessidades particulares.

Todos consideraram a análise acústica complementar à avaliação; e o uso não era padronizado.

Entende-se que o uso de equipamentos sofisticados e caros não é uma realidade brasileira, mas a cada dia os laboratórios de voz ficam mais fáceis em seu manuseio e mais acessíveis em sua aquisição. Portanto, faz-se urgente a necessidade da investigação para que seu uso não seja banalizado ou feito de maneira inadequada.

Van Lierde, Moerman, Vermeersch e Van Cauwenberge (1996) referiram que o acúmulo de informações geradas pela análise acústica computadorizada da voz parece ser esmagador, levando os clínicos freqüentemente a se perderem em porcentagens, milisegundos e semitons. Por isso, há que se estabelecer padrões para se avaliar as medidas objetivas, que em conjunto com as medidas perceptuais podem responder a questões do diagnóstico do paciente; tendo-se como maior benefício a determinação de direções apropriadas para o tratamento e evolução do paciente.

Por que não pensar também em usar essas medidas para auxiliar na atuação com profissionais de voz, relacionando-as com os aspectos psicológicos almejados?

É objetivo do estudo o uso do laboratório de fala e voz, também para o trabalho de imitação vocal, de uma forma objetiva, com parâmetros acústicos que caracterizem as qualidades vocais necessárias para que o ouvinte julgue aquela voz agradável, confiável e natural.

Levantado o problema, trabalhou-se com aspectos psicológicos, para os quais fez-se um levantamento de definições; e qualidade vocal relacionada à efetividade da comunicação, e que foi definida por parâmetros acústicos como amplitude, freqüência fundamental e tempo da amostra.

O programa escolhido para a análise foi o *Dr. Speech* 3.0., fabricado pela *Tiger Eletronics* e o módulo utilizado o *Speech Analysis* (Huang, Lin, & O'Brien, 1995; Huang, 1995)

A amostra foi composta por fala encadeada uma vez que é assim que os repórteres e apresentadores trabalham, não interessando para esse estudo a análise de vogal sustentada.

Portanto, depois de escolhido o estímulo, gravaram-se as amostras que foram analisadas acusticamente quanto a amplitude, frequência fundamental e tempo da amostra, através do programa *Dr. Speech*. Julgamentos dos aspectos psicológicos dessas vozes geraram dados que foram submetidos a uma MDS (análise de escala multidimensional).

Análises de escalas multidimensionais têm sido usadas para se investigar qualidade vocal em vários estudos. A fisiologia da produção vocal é complexa, multidimensional e seus vários fenômenos têm muitos fatores interdependentes (Köster, Marx, Gemmar, Hess & Künzel, 1999).

Assim como a compreensão fisiológica da produção vocal, o entendimento da qualidade vocal envolve dimensões que remetem à parâmetros diversos não muito claros na literatura, principalmente se considerarem-se aspectos psicológicos.

Singh e Murry (1978) apresentaram segmentos de vozes normais, masculinas e femininas, a dez juízes leigos para julgamento de similaridade. Para cada amostra havia oito medidas acústicas e quatro apontamentos de termos descritivos feitos por dez patologistas da fala. Seus dados foram submetidos a uma análise de escala multidimensional, INDISCAL, e eles concluíram que vozes normais em relação aos julgamentos de similaridade podem ser interpretadas por três dimensões. Dessas, a segunda estava

correlacionada com os julgamentos de *pitch* e de duração dos segmentos e a terceira, interpretada somente para as vozes femininas, correlacionada com a duração total do segmento de fala. A primeira dimensão separou totalmente vozes femininas de masculinas e se correlacionou com quase todas as medidas acústicas, refletindo o fato de que elas são diferentes em essência.

Murry e Sing (1980) continuaram seu estudo separando os sujeitos masculinos e femininos. O interesse continuou sendo julgamento de similaridade, as amostras, vogais sustentadas e frases curtas. Os julgamentos de similaridade, que foram submetidos à análise de escala multidimensional, pareceram ter influências diferentes para vozes masculinas e femininas em alguns aspectos. No entanto a frequência fundamental foi importante para todos os julgamentos. Os resultados foram interpretados por parâmetros acústicos previamente analisados e por valores levantados com seis experientes patologistas da fala em relação a cinco qualidades vocais: *pitch*, sopro, esforço, nasalidade e rouquidão.

A MDS (escala de análise multidimensional) foi utilizada pelos estudos descritos a seguir.

Kreiman *et al* (1992) analisaram vozes normais e patológicas, usando juízes leigos e experientes. Os julgamentos analisados pela MDS demonstraram que os juízes variaram menos suas estratégias perceptuais para julgar vozes normais do que para as patológicas. Juízes experientes variaram muito sua estratégia perceptual para vozes patológicas, dependendo de seu treinamento e de sua prática clínica, o que sugeriu a necessidade de protocolos para padronizar as avaliações de vozes patológicas.

Gelfer (1993) usou a análise de escala multidimensional (MDS) para investigar a discriminação de vozes femininas normais por patologistas da fala e por ouvinte leigos. O julgamento de similaridade foi feito com sentenças apresentadas aos pares. Os dados submetidos a uma MDS apontaram cinco dimensões que estavam correlacionadas com o julgamento dos patologistas da fala: percepção de *pitch*, percepção de *loudness*, percepção de duração das sentenças usadas como estímulo, percepção de variabilidade e percepção de qualidade. Para os juízes leigos, somente duas dimensões foram correlacionadas.

O uso da técnica da análise de escala multidimensional foi interessante para que agrupássemos as vozes de acordo com os julgamentos feitos para os diferentes aspectos psicológicos (agradabilidade, confiabilidade e naturalidade). Verificamos então se as dimensões encontradas para esses diferentes aspectos psicológicos estudados, poderiam ser identificadas por características acústicas específicas.

Definir aspectos subjetivos sobre qualquer assunto é uma tarefa delicada. Como toda a investigação é permeada por termos não muito bem definidos na literatura e com variações de um trabalho para outro, procurou-se padronizar os termos que foram usados durante o experimento.

Após o levantamento desses termos com auxílio de questionários aplicados a profissionais da voz, para padronização do experimento, foram feitos os julgamentos, as correlações e a análise de escala multidimensional.

OBJETIVOS

Os objetivos do presente trabalho foram:

- Correlacionar o julgamento dos aspectos psicológicos às medidas acústicas feitas por programa de análise acústica.
- Verificar as dimensões em que se encontram as vozes, para cada aspecto psicológico, e se elas poderiam ser definidas por parâmetros acústicos.
- Interpretar clinicamente as correlações para intervenção em vozes profissionais.

5.1. Participantes

Houve dois grupos de participantes nesta pesquisa. O primeiro grupo consistiu de quarenta e cinco profissionais da voz (22 F e 23 M), sendo dezoito (13 F e 5 M) profissionais na faixa etária de 18 a 30 anos, vinte e dois (8 F e 14 M) profissionais na faixa de 31 a 45 anos e cinco (1 F e 4 M) profissionais na faixa de 45 a 60 anos, contatados pessoalmente em estúdios, consultórios e universidades. Estes profissionais eram fonoaudiólogos, atores, locutores, cantores, professores e repórteres de televisão e rádio. O segundo grupo consistiu de sessenta e sete universitários (35 F e 32 M), com idade entre 18 e 25 anos, escolhidos aleatoriamente nas universidades de Ribeirão Preto. Todos os participantes deste grupo não tinham histórico de qualquer problema ou limitação auditiva. Todos assinaram uma declaração de consentimento (ver anexo 1) para participar da pesquisa, de acordo com a Declaração de Helsinki.

5.2. Material e Equipamentos

Um questionário, tomando-se como referência os protocolos de avaliação global da voz da Escola Paulista de Medicina e Centro de Estudos da Voz (Behlau & Pontes, 1995) e uma tabela de termos descritivos para voz (Boone, 1996), foi elaborado para indicar termos de definição dos aspectos psicológicos: agradabilidade, confiabilidade e naturalidade quando aplicados à voz (ver anexo 2).

Uma fita cassete TDK AS60 *super high resolution* para gravar os estímulos em estúdio profissional e um aparelho *stereo radio cassette recorder* PANASONIC RX-FT 500 para reproduzi-la foram utilizados para a seleção e definição do estímulo. Os estímulos produzidos por um locutor profissional

masculino e um feminino, foram quatro opções de fala: (1) um trecho escrito por Carlos Drummond de Andrade, (2) uma simulação de recado de secretária eletrônica em que se procurou usar o maior número possível de fonemas da língua portuguesa, (3) frases com sentido balanceadas foneticamente e (4) frases sem sentido balanceadas foneticamente. Estas duas últimas opções foram retiradas de listas de sentenças em português desenvolvidas por Costa (1998). (ver anexo 3)

Dois CDs de 650MB da marca MITSUI, sendo um destinado para a análise acústica e outro para a apresentação aos participantes, foram utilizados na gravação do estímulo selecionado produzidos por 5 repórteres masculinos e 5 femininos.

Um computador Pentium 133MHz com 4 MB de RAM, *kit* multimídia *Sound Blaster*, sistema operacional *Windows 95* e com o módulo *Speech Analysis* do programa *Dr. Speech 3.0* (Huang *et al*; 1995) foi utilizado para a análise acústica de voz.

Um aparelho SONY-CD *Radio Cassete-Corder* (1 BIT DUAL D/A Converter) tipo CFD-10S foi utilizado para apresentar as vozes dos repórteres aos pares aos juizes leigos.

Folhas de resposta previamente preparadas em papel sulfite, apresentadas com uma folha de instrução, foram utilizadas para se registrar os julgamentos dos participantes (ver anexo 4).

5.3. Procedimento

A execução desta pesquisa deu-se em cinco etapas. Na primeira etapa, aplicou-se o questionário aos profissionais de voz. Uma vez retornado o

questionário, as respostas foram analisadas e apresentadas em histogramas no anexo 5. Com os tópicos comuns apontados pelos profissionais pode-se definir cada um dos três aspectos psicológicos da voz. Para agradabilidade as definições foram voz limpa, clara, firme, alegre e harmoniosa; para confiabilidade as definições foram voz convincente, clara, firme, limpa, expressiva e autêntica e para naturalidade foram voz limpa, clara, estável e autêntica.

Embora esses termos pudessem variar entre indivíduos, supôs-se que eles serviriam para nortear o julgamento dos juizes leigos, uma vez que eles foram orientados a considerá-los nos seus julgamentos. Portanto, admitiu-se que o experimento foi "calibrado" por este procedimento prévio.

A segunda etapa consistiu em selecionar e definir como seria o estímulo a ser apresentado aos juizes leigos, pois não foram encontradas referências na literatura de quais seriam os estímulos mais adequados aos objetivos desta pesquisa. Para determiná-lo, gravou-se primeiramente em uma fita cassete a seqüência das quatro opções de fala encadeada (anexo 3), em estúdio profissional, com uma voz feminina e outra masculina, ambas de locutores profissionais.

Essas opções de fala encadeada foram apresentadas a vinte participantes (10 F e 10 M) do segundo grupo para que apontassem qual era a mais adequada para se avaliar somente a voz, desconsiderando o conteúdo semântico. Estes participantes estimaram essas opções numa escala de 1 a 10, sendo que quanto mais próximo de um, menos adequado seria o estímulo e quanto mais próximo de dez, mais adequado seria o estímulo. Estas pessoas não participaram de outras etapas da pesquisa para se evitar qualquer

influência de exposição prévia ao estímulo selecionado nos julgamentos dos aspectos psicológicos. Os resultados desta etapa estão no anexo 6.

Considerando que o estímulo gravado com a voz masculina mais pontuado foi o menos pontuado em relação àqueles gravados com a voz feminina, optou-se por escolher o estímulo que tinha pontuações similares para as vozes masculinas e femininas. Este estímulo escolhido teve pontuações medianas como indicado no histograma do resultado da seleção (ver anexo 6) e foi a lista de 10 frases desenvolvida por Costa (1998), a seguir:

Vou viajar às nove da manhã.

Meu irmão bateu o carro ontem.

Prometi a ele não contar o segredo.

Cheguei atrasada na aula.

Esta rua é perigosa.

Esqueci da bolsa na sua mesa.

Ela comprou os últimos pães.

A casa de campo já foi alugada.

Os preços não devem subir.

Não falei com sua filha.

Este estímulo foi produzido por 5 repórteres masculinos e 5 repórteres femininos e as produções gravadas nos dois CDs. As gravações foram feitas em estúdio profissional com tratamento acústico e com a postura dos repórteres em pé e com uma distância de 30 cm entre o microfone e a boca.

Em um CD as vozes dos locutores foram editadas em faixas de gravação individuais no formato *wave* para que pudessem ser submetidas a análise acústica. No outro CD as falas destes 10 locutores foram editadas em formato áudio de maneira que em cada faixa de gravação duas vozes distintas estavam sucessivamente emparelhadas. Houve 45 faixas de gravação, ou seja, todas as combinações possíveis de pares de vozes.

A terceira etapa consistiu na análise acústica das vozes dos locutores que produziram o estímulo selecionado, gravadas em formato *wave* em um dos CDs. Neste processo, a gravação foi convertida de estéreo para mono a 16 BIT para que a análise fosse feita pelo *Dr. Speech*. A análise acústica forneceu os seguintes parâmetros físico-acústicos: duração da amostra, contornos da frequência fundamental e da intensidade e espectogramas (ver anexo 7). Embora a literatura aponte os formantes mais altos (quarto em diante) como os que parecem estar relacionados com o timbre vocal, que é uma característica individual que compõe o som da voz (Kent, 1993), o programa usado não permitiu essa análise.

A quarta etapa consistiu em apresentar as vozes dos locutores gravadas aos pares em formato áudio no outro CD. Elas foram apresentadas pelo método de comparação aos pares aos 47 participantes restantes (25 F e 22 M) do segundo grupo. Dezesseis deles (10 M e 6 F) foram instruídos a escolher em cada par de vozes aquela que seria a mais agradável. Outros dezessete (10 F e 7 M) a mais confiável. E os catorze restantes (7 F e 7 M), a mais natural. Além disso, respectivo a cada grupamento, eles deveriam estimar numa escala de 1 a 10 o quão agradável (ou natural ou confiável) seria a voz escolhida em relação a preterida. Quanto mais próximo de 1, menos agradável (ou confiável ou natural) seria a voz escolhida. E quanto mais próximo de 10,

mais agradável (ou confiável ou natural) seria a voz escolhida. Para julgamentos de intensidade intermediária escolhiam-se números inteiros entre 1 e 10, cujos intervalos entre seus valores sucessivos foram assumidos como eqüidistantes na escala julgada (ver anexo 4). Cada um dos 45 pares de vozes foi apresentado uma única vez a cada participante em ordem aleatória para que as respostas não sofressem influência da seqüência apresentada.

As sessões foram individuais e cada uma durava aproximadamente quarenta minutos. No início das sessões os participantes leram as instruções e antes de começar o julgamento, o experimentador enfatizou verbalmente que a tarefa era para avaliar os aspectos psicológicos transmitidos pela voz e que deveria ser considerado a definição prévia dos profissionais da voz, apresentadas na instrução. Quando não restassem mais dúvidas em relação às instruções iniciavam-se os julgamentos deixando-se claro que o participante poderia interromper a sessão a qualquer momento se assim o desejasse.

O ambiente onde se realizaram as sessões teve sempre as condições de ruído e a posição dos participantes em relação ao aparelho reproduzidor de CD controladas. O volume da intensidade dos estímulos foi o médio do aparelho utilizado.

A quinta e última etapa consistiu na análise dos dados que está apresentada no capítulo resultados. Valores obtidos a partir dos julgamentos de agradabilidade, confiabilidade e naturalidade das vozes da etapa anterior foram correlacionados entre si e com os parâmetros da análise acústica. Além disso, matrizes de similaridade entre as vozes julgadas em cada aspecto psicológico, constituídas pelos coeficientes de correlação de Pearson entre as somas das pontuações individuais de cada voz julgada pelos juizes leigos, foram submetidas a uma análise de escala multidimensional (*MDS*), pelo

método de Kruskal (1964 a, 1964 b), disponível no pacote estatístico SYSTAT 7.0 (Wilkinson, 1997). A análise de *MDS* é uma técnica de redução da estatística multivariada que teve a finalidade nesta pesquisa de descrever as vozes em cada aspecto psicológico julgado, por dimensões que poderiam ser associadas a características subjetivas e/ou objetivas das vozes (Hair, Anderson, Tatham, & Black, 1998). Para se determinar o menor número de dimensões que descrevessem as vozes estabeleceu-se a quantidade de iterações que gerasse o menor *stress* (índice de bom ajustamento dos dados no método de Kruskal) que abrangesse pelo menos 95% da proporção de variância, indicada pelo quadrado da correlação entre as medidas de similaridade provenientes dos dados e daquelas geradas pelo modelo de *MDS*.

RESULTADOS

Esta seção foi dividida em quatro partes para facilitar a compreensão do leitor. Na primeira estão os resultados da análise acústica das vozes, na segunda, a análise dos julgamentos de agradabilidade, na terceira a análise dos julgamentos de confiabilidade e na última a análise dos julgamentos de naturalidade. Antes das análises, dados de 17 participantes (10 F e 7 M) foram desconsiderados porque estes reconheceram as vozes de alguns locutores nas sessões. Assim, os dados analisados foram de 30 participantes (15 F e 15 M), sendo 10 deles (5 F e 5 M) para cada aspecto psicológico.

6.1. Análise Acústica

A tabela 1 contém os parâmetros acústicos das vozes da amostra obtidos pelo programa *Dr. Speech* 3.0. As vozes femininas são as de números 1, 2, 4, 6 e 10 e as masculinas são as de números 3, 5, 7, 8 e 9.

A análise foi feita no *default* do programa que apresenta FFT: 1024 pontos; *display* de frequência entre 0 e 5500 Hz; *display* de amplitude entre -40-120dB; LPC *order* 14; análise de *pitch* limite inferior 80 Hz e limite superior 400Hz; análise de espectograma em banda larga de 300 Hz, que é a recomendada para análise de fala.

A análise na íntegra, com os contornos de frequência, intensidade e onda gerada, pode ser observada no anexo 7.

Tabela 1. Parâmetros acústicos das vozes obtidos com o programa *Dr Speech 3.0*.

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
duração da amostra (s)	27,68	27,22	23,41	25,28	29,54	24,75	23,36	22,26	21,36	24,1
Fo máxima (Hz)	367,5	380,17	400,91	400,91	370,59	400,91	400,91	383,48	324,26	400,91
Fo mínima (Hz)	117,91	84,97	99,1	98,44	93,43	93,23	100	89,27	83,05	125,28
Fo média (Hz)	214,75	202,68	192,78	239,06	188,52	238,31	165,16	173,39	123,07	241,1
intensidade máxima (dB)	78,61	81,51	81,21	80,63	80,53	79,22	81,86	80,24	81,32	75,6
intensidade mínima (dB)	21,58	12,04	13,98	16,9	15,56	19,08	25,11	13,98	20	12,04
Intensidade média (dB)	55,8	52,35	60,36	61,82	56,59	61,14	64,3	60,04	61,53	54,5
percentual tempo fala (%)	66,62	57,46	72,96	78,8	66,08	74,42	80,14	73,74	79,26	71,49
percentual tempo silêncio(%)	33,38	42,54	27,04	21,2	33,92	25,58	19,86	26,46	20,74	28,51
percentual tempo voz (%)	51,81	44,71	56,47	57,08	51,83	56,69	55,09	52,7	49,63	49,13
percentual tempo semvoz(%)	14,81	12,75	16,49	21,72	14,25	17,74	25,04	20,84	29,63	22,37
# max. Fo e min. Fo	249,59	295,2	301,81	302,47	277,16	307,68	300,91	294,21	241,21	275,63
# max Io e min. Io	57,03	69,47	67,23	63,73	64,97	60,14	56,75	66,26	61,32	63,56

Legenda

Vozes femininas

Vozes masculinas

Fo: frequência fundamental

Io: intensidade

As durações das amostras das vozes masculinas, exceto para a voz do locutor 5, foram menores que das vozes femininas.

As Fo máximas (frequências fundamentais máximas) de 5 locutores, 2 masculinos e 3 femininos foram iguais a 400.91 Hz.

Não houve demarcação específica para as Fo mínimas (frequências fundamentais mínimas) das vozes masculinas e femininas, sendo que algumas vozes femininas apresentaram valores mais baixos do que as masculinas.

As Fo médias (frequências fundamentais médias) das vozes masculinas foram as mais baixas.

Não houve característica que demarcasse limites entre intensidade de vozes femininas e masculinas.

O mesmo ocorreu para o percentual de tempo de fala e silêncio e para o percentual de tempo com voz e sem voz (com vibração e sem vibração de pregas vocais, ou seja, fonemas sonoros e surdos).

As diferenças entre Fo (frequência fundamental) máxima e mínima e entre intensidade máxima e mínima apresentaram padrões de variação similares para as vozes masculinas e femininas.

6.2. Julgamentos de agradabilidade

A preferência pela voz mais agradável em cada par e a atribuição de um valor entre 1 e 10 para sua agradabilidade resultaram para cada participante numa tabela de dupla entrada. Os valores nesta tabela, respectivos a cada voz, foram somados gerando uma escala individual de agradabilidade. As somatórias das pontuações destas escalas individuais resultaram na escala de agradabilidade ilustrada no gráfico da figura 1.

Figura 1. Somatórias das pontuações de agradabilidade.

A escala apresentada é decrescente, portanto a julgada mais agradável foi a voz do locutor 9 e a julgada menos agradável foi a voz do locutor 5.

Para gerar as medidas de similaridade entre as vozes que alimentaram a análise de escala multidimensional (MDS), as pontuações de agradabilidade de cada voz, indicadas na escala individual, foram correlacionadas entre si pelo método de correlação de *Pearson* (anexo 8). Essa matriz de coeficientes de correlação foi submetida a análise de MDS resultando em três dimensões. As coordenadas das vozes nestas dimensões estão na tabela 2. O número de dimensões foi determinado pela abrangência mínima de 95% de variância dos dados (anexo 9).

Tabela 2. Coordenadas para as dimensões de agradabilidade.

VOZ	DIMENSÕES		
	1	2	3
V1	-0.19	.76	.25
V2	.35	-.28	.49
V3	.40	-.73	.01
V4	.33	.35	.45
V5	-1.06	.16	-.76
V6	-1.03	.04	.55
V7	.11	-.81	-.52
V8	.45	.61	-.75
V9	1.37	.18	.04
V10	-.73	-.28	.25

Estas coordenadas estabeleceram as posições das vozes num espaço tridimensional representado na figura 2.

Figura 2. Configuração das vozes em função das dimensões para agradabilidade.

Prosseguiu-se a análise correlacionando os parâmetros da análise acústica com a escala de agradabilidade da figura 1. Na tabela 3 encontram-se os coeficientes destas correlações desconsiderando-se a distinção de gênero das vozes, considerando-se somente as vozes femininas e considerando somente as vozes masculinas.

Tabela 3. Coeficientes da correlação de *Pearson* entre os parâmetros acústicos e a escala de agradabilidade.

	V m+f	V fem	V masc
duração da amostra (s)	-0,38	0,09	-0,71*
Fo máxima (Hz)	-0,31	0,2	-0,57
Fo mínima (Hz)	-0,09	-0,6	-0,44
Fo média (Hz)	-0,09	-0,01	-0,93***
Intensidade máxima (dB)	0,06	0,76*	0,65
Intensidade mínima (dB)	0,35	-0,13	0,66
Intensidade média (dB)	0,2	0,44	0,71*
Percentual tempo de fala (%)	0,33	0,24	0,88***
Percentual tempo de silêncio(%)	-0,33	-0,24	-0,88***
Percentual tempo de voz(%)	-0,16	0,3	-0,37
Percentual tempo sem voz (%)	0,56**	0,1	0,96*****
# max Fo e min Fo (Hz)	-0,27	0,57	-0,59
# max Io e min Io (dB)	-0,35	0,42	-0,64

Legenda

P= 0.10	*	V m+ f = vozes femininas e masculinas
P= 0.05	**	V fem = vozes femininas
P= 0.025	***	V masc = vozes masculinas
P= 0.01	****	Fo: frequência fundamental
P= 0.005	*****	Io: intensidade

Nota-se que o maior número de coeficientes de correlação significativos da escala de agradabilidade com os parâmetros acústicos aparece para as vozes masculinas. Para as vozes femininas somente houve correlação significativa com a intensidade máxima. E para as vozes, desconsiderando-se a distinção de gênero, somente houve correlação significativa com o percentual de tempo sem voz.

A correlação de *Pearson* também foi aplicada entre os valores de cada dimensão determinada pela MDS e os parâmetros acústicos. A tabela 4 apresenta os coeficientes dessas correlações, desconsiderando-se a distinção de gênero das vozes, considerando somente as vozes femininas e somente as masculinas.

Tabela 4. Coeficientes da correlação de *Pearson* entre os parâmetros acústicos e as coordenadas das dimensões da MDS para agradabilidade.

DIMENSÃO 1			
	V m/f	V fem	V masc
duração da amostra (s)	-0,6**	0,59	-0,94****
Fo máxima (Hz)	-0,48	-0,38	-0,44
Fo mínima (Hz)	-0,45	-0,37	-0,51
Fo média (Hz)	-0,66***	-0,55	-0,74*
Intensidade máxima (dB)	0,55**	0,69*	0,32
Intensidade mínima (dB)	0,08	-0,24	0,18
Intensidade média (dB)	0,3	-0,2	0,58
Percentual tempo de fala (%)	0,3	-0,33	0,77*
Percentual tempo de silêncio(%)	-0,29	0,33	-0,76*
Percentual tempo de voz (%)	-0,15	-0,29	-0,22
Percentual tempo sem voz (%)	0,5	-0,3	0,79*
# max Fo e min Fo (Hz)	-0,25	0,02	-0,41
# max Io e min Io (dB)	0,15	0,51	-0,15
DIMENSÃO 2			
	V m/f	V fem	V masc
duração da amostra (s)	0,22	0,46	0,04
Fo máxima (Hz)	-0,42	-0,48	-0,56
Fo mínima (Hz)	0	0,25	-0,8**
Fo média(Hz)	0,09	-0,06	-0,25
Intensidade máxima (dB)	-0,22	0,12	-0,81**
Intensidade mínima (dB)	0,01	0,87**	-0,46
Intensidade média (dB)	-0,16	0,39	-0,53
Percentual tempo fala (%)	-0,01	0,27	-0,31
Percentual tempo silêncio (%)	0,1	-0,27	0,32
Percentual tempo de voz (%)	-0,06	0,51	-0,76*
Percentual temposem voz (%)	-0,09	-0,11	0,03
# max Fo e min Fo (Hz)	-0,44	-0,49	-0,47
# max Io e min Io (dB)	-0,12	-0,76*	0,39
DIMENSÃO 3			
	V m/f	V fem	V masc
duração da amostra (s)	0,07	-0,09	-0,53
Fo máxima (Hz)	0,13	0,35	-0,32
Fo mínima(Hz)	0,19	-0,94****	-0,14
Fo média(Hz)	0,57**	-0,02	-0,42
Intensidade máxima (dB)	-0,28	0,72*	0,54**
Intensidade mínima (dB)	-0,13	-0,06	0,09
Intensidade média(dB)	-0,25	0,43	0,34
Percentual tempo fala(%)	-0,15	0,06	0,46
Percentual tempo silêncio (%)	0,15	-0,06	-0,46
Percentual tempo com voz (%)	-0,08	0,26	0,05
Percentual tempo sem voz (%)	-0,13	-0,21	0,40
# max Fo e min Fo (Hz)	0,03	0,90**	-0,36
# max Io e min Io (dB)	-0,01	0,38	-0,02

Legenda

P= 0.10	*	V m+ f = vozes femininas e masculinas
P= 0.05	**	V fem = vozes femininas
P= 0.025	***	V masc = vozes masculinas
P= 0.01	****	Fo: frequência fundamental
P= 0.005	*****	Io: intensidade

Nota-se também na tabela 4 que as vozes masculinas apresentam o maior número de coeficientes de correlação significativos.

6.3. Julgamentos de Confiabilidade

Um tratamento similar para analisar os julgamentos de agradabilidade foi feito para os julgamentos de confiabilidade das vozes. As somatórias das pontuações de das escalas individuais respectivas a cada voz geraram a escala figura 3.

Figura 3. Somatória das pontuações de Confiabilidade.

A escala apresentada é decrescente, portanto a julgada mais confiável foi a voz do locutor 4 e a julgada menos confiável foi a voz do locutor 5.

Os coeficientes de correlação de *Pearson* entre as pontuações de confiabilidade das vozes estão na segunda tabela do anexo 8. Esses dados submetidos à análise de MDS resultaram também em três dimensões, cujas coordenadas para cada voz podem ser observadas na tabela 5.

Tabela 5. Coordenadas para as dimensões de confiabilidade.

VOZ	DIMENSÕES		
	1	2	3
V1	-.27	-.15	.17
V2	1.00	.37	.67
V3	-.21	.22	-.63
V4	.12	1.02	-.32
V5	-1.14	-.66	-.10
V6	.96	-.28	.16
V7	-.48	.66	-.04
V8	.61	-.39	-.63
V9	-.97	.07	.66
V10	.39	-.86	.06

Estas coordenadas determinaram as posições no espaço tridimensional na figura 4.

Figura 4. Configuração das vozes em função das dimensões para confiabilidade.

Prosseguiu-se a análise com as correlações entre os parâmetros da análise acústica e a escala de confiabilidade da figura 3. Na tabela 6 encontram-se os coeficientes dessas correlações desconsiderando-se a distinção de gênero das vozes, considerando somente as vozes femininas e considerando somente as vozes masculinas.

Tabela 6. Coeficientes da correlação de *Pearson* entre os parâmetros acústicos e a escala de confiabilidade.

	V m/f	V fem	V masc
duração da amostra (s)	-0,13	0,57	-0,84**
Fo máxima (Hz)	-0,08	-0,39	-0,33
Fo mínima (Hz)	0,13	-0,18	-0,25
Fo média (Hz)	0,18	-0,42	-0,80**
intensidade máxima (dB)	-0,01	0,58	0,71*
intensidade mínima (dB)	0,22	0,01	0,66
intensidade média (dB)	-0,03	-0,06	0,86**
percentual tempo de fala (%)	0,08	-0,15	0,97*****
percentual tempo de silêncio(%)	-0,08	0,15	-0,97*****
percentual tempo de voz(%)	-0,18	-0,11	-0,12
percentual tempo sem voz (%)	0,24	-0,16	0,93*****
# max Fo e min Fo (Hz)	-0,16	-0,13	-0,34
# max Io e min Io (dB)	-0,23	0,28	-0,63

Legenda

P= 0.10	*	V m+ f = vozes femininas e masculinas
P= 0.05	**	V fem = vozes femininas
P= 0.025	***	V masc = vozes masculinas
P= 0.01	****	Fo: frequência fundamental
P= 0.005	*****	Io: intensidade

Nota-se que os coeficientes de correlação significativos apareceram somente para as vozes masculinas.

Os coeficientes de correlação entre os parâmetros acústicos e as coordenadas das dimensões determinadas pela MDS estão na tabela 7.

Tabela 7. Coeficientes da correlação de *Pearson* entre os parâmetros acústicos e as coordenadas das dimensões da MDS para confiabilidade.

DIMENSÃO 1			
	V m/f	V fem	V masc
duração da amostra (s)	-0,07	-0,24	-0,49
Fo máxima (Hz)	0,52	0,36	0,50
Fo mínima (Hz)	0,01	-0,68*	0,11
Fo média (Hz)	0,57**	-0,09	0,27
intensidade máxima (dB)	-0,27	0,29	-0,35
intensidade mínima (dB)	-0,43	-0,48	-0,36
intensidade média (dB)	-0,27	-0,11	0,20
percentual tempo de fala (%)	-0,28	-0,29	0,13
percentual tempo de silêncio(%)	0,29	0,29	-0,11
percentual tempo de voz (%)	-0,10	-0,26	0,40
percentual tempo sem voz (%)	-0,30	-0,24	-0,07
# max Fo e min Fo (Hz)	0,55**	0,72*	0,59
# max Io e min Io (dB)	0,32	0,60	0,35
DIMENSÃO 2			
	V m/f	V fem	V masc
duração da amostra (s)	-0,15	0,33	-0,56
Fo máxima (Hz)	0,14	-0,04	0,29
Fo mínima (Hz)	-0,31	-0,62	0,44
Fo média(Hz)	-0,10	-0,21	-0,26
intensidade máxima (dB)	0,66**	0,85**	0,93***
intensidade mínima (dB)	0,32	0,03	0,71*
intensidade média (dB)	0,46	0,37	0,92***
percentual tempo fala (%)	0,31	0,10	0,81**
percentual tempo silêncio (%)	-0,31	-0,10	-0,82**
percentual tempo de voz (%)	0,30	0,25	0,50
percentual temposem voz (%)	0,18	-0,11	0,53
# max Fo e min Fo (Hz)	0,34	0,42	0,24
# max Io e min Io (dB)	-0,07	0,33	-0,66
DIMENSÃO 3			
	V m/f	V fem	V masc
duração da amostra (s)	0,16	0,53	-0,12
Fo máxima (Hz)	-0,56**	-0,53	-0,81**
Fo mínima(Hz)	-0,17	-0,38	-0,56
Fo média(Hz)	-0,17	-0,80**	-0,85**
intensidade máxima (dB)	0,00	0,26	0,43
intensidade mínima (dB)	0,19	-0,27	0,60
intensidade média(dB)	-0,36	-0,75*	0,23
percentual tempo fala(%)	-0,32	-0,94****	0,43
percentual tempo silêncio (%)	0,31	0,94****	-0,44
percentual tempo com voz (%)	-0,67***	-0,80*	-0,71*
percentual tempo sem voz (%)	0,09	-0,83**	0,70*
# max Fo e min Fo (Hz)	-0,49	-0,08	-0,86**
# max Io e min Io (dB)	-0,18	0,43	-0,61

Legenda

P= 0.10	*	V m+ f = vozes femininas e masculinas
P= 0.05	**	V fem = vozes femininas
P= 0.025	***	V masc = vozes masculinas
P= 0.01	****	Fo: frequência fundamental
P= 0.005	*****	Io: intensidade

Nota-se na tabela 7 que a dimensão 1 não se correlacionou significativamente com os parâmetros acústicos das vozes masculinas. Porém, ela se correlaciona significativamente com Fo média e a diferença entre Fo máxima e mínima quando se desconsiderou a distinção de gênero. Para as vozes femininas, a dimensão 1 correlacionou-se significativamente com os parâmetros Fo mínima e diferença entre Fo máxima e mínima.

A dimensão 2 correlacionou-se significativamente em predomínio com os parâmetros acústicos das vozes masculinas. Para as vozes femininas e para as vozes quando se desconsiderou a distinção de gênero, a dimensão 2 somente se correlacionou significativamente com a intensidade máxima.

Em prevalência, a dimensão 3 correlacionou-se significativamente com vários parâmetros acústicos das vozes femininas (Fo média, intensidade média, percentual de tempo de fala e silêncio e percentual de tempo com e sem voz); e masculinas (Fo máxima, Fo média, percentual de tempo com e sem voz e diferença entre Fo máxima e mínima). Para as vozes desconsiderando-se a distinção de gênero as correlações foram significativas somente para os parâmetros acústicos de Fo máxima e percentual de tempo com voz.

6.4. Julgamentos de naturalidade

Um tratamento similar àqueles dados à agradabilidade e à confiabilidade das vozes também foi aplicado para os julgamentos de naturalidade. Este

tratamento resultou na escala da figura 5, a partir das somatórias das pontuações individuais de naturalidade de cada voz.

Figura 5. Somatória das pontuações de naturalidade.

A escala é decrescente, portanto a julgada mais natural foi a voz do locutor 9 e a julgada menos natural foi a voz do locutor 5.

De maneira análoga à análise para agradabilidade e confiabilidade, uma matriz indicando as similaridades entre as vozes em relação a naturalidade por meio dos coeficientes de correlação de *Pearson* (anexo 8) foram submetidos a análise de MDS. Esta análise resultou em três dimensões para configurar a naturalidade das vozes, cujas coordenadas estão na tabela 8.

Tabela 8. Coordenadas para as dimensões de naturalidade.

VOZ	DIMENSÕES		
	1	2	3
V1	-.98	-.20	-.70
V2	.58	-.73	.12
V3	-.61	-.63	-.11
V4	-1.13	.19	.19
V5	-.35	1.19	.01
V6	.66	-.25	-.25
V7	.09	-.65	.43
V8	.47	.35	.24
V9	.02	.30	.54
V10	1.25	.43	-.47

Estas coordenadas determinaram as posições das vozes no espaço tridimensional na figura 6.

Figura 6. Configuração das vozes em função das dimensões para naturalidade.

Prosseguiu-se a análise com as correlações entre os parâmetros da análise acústica e a escala de naturalidade mostrada na figura 8. Os coeficientes dessas correlações para as vozes desconsiderando-se a distinção de gênero, para as vozes femininas e para as vozes masculinas estão na tabela 9.

Tabela 9. Coeficientes da correlação de *Pearson* entre os parâmetros acústicos e a escala de naturalidade.

	V m/f	V fem	V masc
duração da amostra (s)	-0,64***	-0,74*	-0,85**
Fo máxima (Hz)	0,08	0,88***	0,18
Fo mínima (Hz)	-0,05	-0,45	-0,11
Fo média (Hz)	0,05	0,67	-0,57
intensidade máxima (dB)	0,02	0,13	0,66
intensidade mínima (dB)	0,16	-0,1	0,35
intensidade média (dB)	0,37	0,70*	0,74*
percentual tempo de fala (%)	0,46	0,62	0,82**
percentual tempo de silêncio(%)	-0,47	-0,62	-0,82**
percentual tempo de voz(%)	0,19	0,58	0,13
percentual tempo sem voz (%)	0,48	0,49	0,69*
# max Fo e min Fo (Hz)	0,11	0,90***	-0,19
# max Io e min Io (dB)	-0,15	0,16	-0,29

Legenda

P= 0.10	*	V m+ f = vozes femininas e masculinas
P= 0.05	**	V fem = vozes femininas
P= 0.025	***	V masc = vozes masculinas
P= 0.01	****	Fo: frequência fundamental
P= 0.005	*****	Io: intensidade

A escala de naturalidade correlacionou-se significativamente com a duração da amostra, quando se desconsiderou o gênero de voz. Entretanto, quando a correlação foi feita separando-se os gêneros, aumentou-se o número de correlações significativas.

A correlação de *Pearson* também foi aplicada entre os parâmetros acústicos e as coordenadas das três dimensões de naturalidade. A tabela 10 apresenta os coeficientes dessas correlações.

Tabela 10. Coeficientes da correlação de *Pearson* entre os parâmetros acústicos e as coordenadas das dimensões da MDS para naturalidade.

DIMENSÃO 1			
	V m/f	V fem	V masc
duração da amostra (s)	-0,29	-0,49	-0,47
Fo máxima (Hz)	0,13	0,38	-0,16
Fo mínima (Hz)	0,00	0,02	-0,44
Fo média (Hz)	0,04	0,13	-0,45
intensidade máxima (dB)	-0,40	-0,43	-0,21
intensidade mínima (dB)	-0,31	-0,60	0,25
intensidade média (dB)	-0,27	-0,42	0,35
percentual tempo de fala (%)	-0,14	-0,26	0,43
percentual tempo de silêncio(%)	0,14	0,26	-0,42
percentual tempo de voz (%)	-0,45	-0,46	-0,37
percentual tempo sem voz (%)	0,15	0,07	0,54
# max Fo e min Fo (Hz)	0,13	0,23	-0,08
# max Io e min Io (dB)	0,18	0,39	-0,31
DIMENSÃO 2			
	V m/f	V fem	V masc
duração da amostra (s)	0,21	-0,71*	0,60
Fo máxima (Hz)	-0,32	0,57	-0,52
Fo mínima (Hz)	0,05	0,72*	-0,56
Fo média(Hz)	-0,04	0,83**	-0,01
intensidade máxima (dB)	-0,32	-0,72*	-0,72*
intensidade mínima (dB)	-0,23	-0,08	-0,38
intensidade média (dB)	-0,12	0,34	-0,81**
percentual tempo fala (%)	0,05	0,75*	-0,66
percentual tempo silêncio (%)	-0,05	-0,75*	0,66
percentual tempo de voz (%)	-0,07	0,41	-0,74*
percentual tempo sem voz (%)	0,12	0,94****	-0,28
# max Fo e min Fo (Hz)	-0,37	-0,14	-0,50
# max Io e min Io (dB)	0,08	-0,31	0,32
DIMENSÃO 3			
	V m/f	V fem	V masc
duração da amostra (s)	-0,46	-0,07	-0,56
Fo máxima (Hz)	-0,29	0,38	-0,53
Fo mínima(Hz)	-0,77*****	-0,79*	-0,50
Fo média(Hz)	-0,71***	-0,05	-0,90***
intensidade máxima (dB)	0,76****	0,74*	0,42
intensidade mínima (dB)	0,17	-0,48	0,74*
intensidade média(dB)	0,54**	0,23	0,65
percentual tempo fala(%)	0,43	0,04	0,80**
percentual tempo silêncio (%)	-0,43	-0,04	-0,80**
percentual tempo com voz (%)	0,01	0,03	-0,52
percentual tempo sem voz (%)	0,56**	0,03	0,95*****
# max Fo e min Fo (Hz)	0,14	0,81**	-0,52
# max Io e min Io (dB)	0,15	0,75*	*-0,76

Legenda

P= 0.10	*	V m+ f = vozes femininas e masculinas
P= 0.05	**	V fem = vozes femininas
P= 0.025	***	V masc = vozes masculinas
P= 0.01	****	Fo: frequência fundamental
P= 0.005	*****	Io: intensidade

Nota-se que a dimensão 1 não se correlacionou significativamente com nenhum parâmetro acústico, nem para as vozes femininas, nem para as masculinas, nem desconsiderando-se a distinção de gênero.

A dimensão 2 não se correlacionou significativamente com os parâmetros acústicos quando se desconsiderou o gênero de voz. Entretanto, quando houve esta distinção, as vozes femininas apresentaram um maior número de correlações significativas com os parâmetros acústicos que as vozes masculinas.

A dimensão 3 apresentou correlações significativas com alguns parâmetros acústicos quando se desconsiderou a distinção de gênero das vozes e também quando se considerou esta distinção. Porém, houve uma prevalência destas correlações para as vozes masculinas.

Houve nos primeiros procedimentos desta pesquisa dificuldade para estabelecer os significados dos termos comumente usados, e nem sempre bem definidos, aos aspectos a que se referiam.

Na elaboração do questionário, por exemplo, não foi encontrado na literatura protocolos onde não aparecessem termos como "normal" ou "adequado" para se definir observações feitas a partir do comportamento vocal.

Para garantir que as respostas tivessem consistência, aplicou-se o questionário somente em profissionais da voz, que consensualmente usam tal terminologia (normal e adequada) para definir a voz sem patologias aparentes (funcional ou orgânica) e sem características de estranheza entre si, segundo Moore (1971), (*apud* Pereira, 1999).

A percepção auditiva é uma habilidade natural que acaba sendo desenvolvida por profissionais que trabalham com a voz humana para ser utilizada como instrumento de análise vocal. Embora esta análise seja importante, ficou clara a necessidade de procedimentos objetivos, no caso deste estudo a análise acústica, que indiquem uma linguagem única para analisar o comportamento vocal.

Behlau (1997) em suas considerações sobre análise acústica apontou que ela é especialmente útil em vozes normais e que ajuda no acompanhamento de vozes profissionais ao longo de um período. No entanto, disse que a análise acústica tem que estar associada a uma análise perceptivo auditiva. Considerou ainda que um laboratório computadorizado de voz ideal ainda não foi definido.

Há ainda que se considerar que quando se relaciona critérios perceptuais a medidas objetivas extraídas de análise acústica, estabelece-se

um elo entre as diversas pessoas que se ocupam de vozes profissionais, como por exemplo professores de canto e cientistas vocais (Ekholm, Papagiannis & Chagnon, 1998).

A escolha do programa usado para análise acústica da fala, o *Dr Speech* 3.0, deu-se por ser ele o disponível à época do experimento. Apresentou também a seu favor ser o de mais fácil manuseio e aparentemente o mais adotado em clínicas particulares, além de apresentar preço acessível.

Casmerides (1988) levantou características de um grupo de usuários de programas acústicos, em que preço e facilidade de manuseio eram determinantes na escolha do programa, o que corroborou com a escolha neste trabalho, já que é objetivo que os dados possam ser usados em clínica.

Outros programas acústicos interessantes apareceram depois de concluída a pesquisa, como por exemplo o *Multi-Speech* da *Kay Elemetrics*, que apesar de ter se tornado mais acessível, ainda guardava a desvantagem de conter um único manual na ajuda do próprio programa, que não foi suficiente para sanar todas as dúvidas que apareceram e ainda houve falta de suporte técnico de seus representantes no Brasil.

Além desses, sabe-se ser usado em pesquisas o *Soundscope*, o *Sound Edit*, e ainda outras versões mais sofisticadas da *Kay Elemetrics*.

Crevier-Buchaman, Monfrais-Pfauwadel, Begue, Lauga-Houdoyer, Laccourreye e Brasnu (1993) revisam em seu trabalho alguns desses programas apontando a importância deles como complemento à avaliação clínica. Berke, Hanson, Trapp, Moore, Gerrat, e Natividad (1989) concluíram que medidas acústicas podem ser relacionadas à qualidade de voz.

Alguns programas desenvolvidos no Brasil, como por exemplo o DSP da Universidade de São Carlos, apesar dos preços convidativos, ainda não

possuíam pesquisas referidas com seu uso para análise de trechos de fala. Pesquisadores do Departamento de Física da Universidade Federal de Minas Gerais trabalham no desenvolvimento de outro programa nacional, mas que ainda não está disponível no mercado (Souza, 1999).

Sabia-se à época do experimento da versão nova do *Dr Speech*, no entanto a que foi usada era aquela disponível no Serviço de Cirurgia de Cabeça e Pescoço do Hospital das Clínicas de Ribeirão Preto. É desejo então deixar claro que estava-se ciente da possibilidade do uso de outros programas, com os quais poderiam ter sido achados resultados diferentes para os parâmetros avaliados.

Lee, Stemple, e Kizer (1999) disseram que se as variáveis no estudo forem controladas, as medidas acústicas tornam-se mais fidedignas.

Bielamowicz *et al* (1996), ao comparar medidas de perturbação em vários sistemas de análise de voz, concluiu que as medidas se mantêm em vozes normais. Portanto, não se acredita que mudanças de programas resultariam em correlações diferentes das encontradas, uma vez que a análise global da fala seria a mesma. No entanto, deve-se considerar que comparações são válidas somente entre mesmos programas e mesmas condições de gravações.

Estender os métodos de avaliação acústica de vogais sustentadas à fala encadeada é um desafio (Hillebrand, Cleveland & Erickson, 1994), mas necessário quando se trata de voz profissional, uma vez que a fala encadeada é a forma natural de conversação.

Para o objetivo inicial de estabelecer parâmetros acústicos que servissem de base para a imitação vocal de repórteres e apresentadores de

televisão, era importante que vozes masculinas e femininas fossem julgadas juntas, uma vez que os principais jornais no Brasil são apresentados por casais.

Além disso, nas reportagens que se seguem à apresentação, nunca se observa a presença exclusiva de profissionais de um mesmo sexo, independente do assunto a ser abordado. Pareceu claro então, que não seria adequado, ao objetivo desse trabalho, separar o experimento considerando o sexo do locutor que gravou o estímulo. No entanto, nas análises houve o cuidado em fazer essa separação, para que se verificasse possíveis diferenças de estratégias envolvidas no processo de julgamento dos juizes.

Na análise acústica, dentro do *Dr. Speech* usou-se o módulo *Speech Analysis* no *default* do programa, o que justifica os valores idênticos que apareceram na medida frequência fundamental máxima para as vozes 3, 4, 6, 7, e 10. Se fossem aumentados os parâmetros de análise para o máximo permitido de 1200 Hz, essa medida aumentaria para todas as vozes. Para vozes masculinas o limite do *range* para frequência fundamental máxima fica entre 200 e 300 Hz; e para vozes femininas entre 200 e 350 Hz, segundo o manual de uso do programa. Embora tivessem sido encontrado limites superiores a esses, acredita-se ser o limite adequado para esta análise uma vez que acima disso, estariam sendo observados os formantes e não a frequência fundamental.

Yumoto, Gould e Baer (1982) consideram que avaliar o aspecto do espectograma é um método visual e portanto subjetivo. Para que se evitasse tal confusão, considerou-se somente medidas fornecidas por estatística do programa *Dr. Speech*.

Ao se usar o *default* do programa, tentou-se simplificar a repetição do experimento por outros profissionais, evitando eventuais problemas com calibração de medidas.

Com as medidas fornecidas pelo programa, fez-se duas outras manualmente, que foram as diferenças entre Fo máxima e Fo mínima; e a diferença entre intensidade máxima e intensidade mínima. Com essas medidas pretendeu-se avaliar a importância das modulações de frequência e intensidade para a avaliação dos sujeitos.

A possível interdependência dos aspectos psicológicos estudados foi verificada e encontra-se demonstrada pelas correlações na tabela 11.

Tabela 11. Coeficientes da correlação de Pearson entre os aspectos psicológicos julgados.

	Vozes Fem + Masc	Vozes Fem	Vozes Masc
Agrad x Conf	0.87 *****	0.72 *****	0.95 *****
Agrad x Nat	0.70 *****	0.35	0.79 *****
Conf X Nat	0.63 ***	-0.4	0.89 *****

Legenda

P= 0.10	*	V m+ f = vozes femininas e masculinas
P= 0.05	**	V fem = vozes femininas
P= 0.025	***	V masc = vozes masculinas
P= 0.01	****	
P= 0.005	*****	

Estas correlações indicam uma forte interdependência entre os aspectos psicológicos para as vozes masculinas. No entanto, para as vozes femininas houve correlação significativa somente entre agradabilidade e confiabilidade e para as outras não houve correlação significativa. Esta observação chamou a

atenção para o fato de que as estratégias de avaliação diferiam para as vozes masculinas e femininas.

Klatt e Klatt (1990) afirmaram que as estratégias perceptuais para se avaliar vozes femininas e masculinas são diferentes, o que vem confirmar esse achado.

Emanuel, Lively e McCoy (1973) também encontraram diferenças em julgamentos de vozes femininas e masculinas, mesmo para juizes treinados.

Walker-Andrews (1997) discutiu que as preferências para vozes são construídas por experiências vividas desde muito cedo no desenvolvimento do bebê, considerando inclusive o período pré-natal, e que informações auditivas da voz como frequência, intensidade e estrutura temporal informam afeto. Em seu experimento, bebês demonstraram preferir a voz da mãe a outras.

A colocação desse estudo trouxe a idéia de que a avaliação que se faz de vozes femininas estaria ligada a qualidade dessas primeiras experiências vividas com a mãe. Este seria o motivo de haver maior consistência nas respostas relacionadas às vozes masculinas, que não sofreriam essa influência.

Duas das vozes masculinas a 5 e a 8 apareceram sempre em nono e décimo lugares respectivamente, para todos os aspectos julgados. Tinham como semelhança entre si os valores de intensidade (máxima, mínima e diferença). Além disso, as duas, em relação as mais preferidas tinham a frequência fundamental média superior em pelo menos 50 Hz, o que as define como mais agudas.

A voz 9, também masculina apareceu por duas vezes na primeira posição, nos aspectos agradabilidade e naturalidade. É uma voz com características que denotam uma voz grave, com frequência fundamental

média de 123.07 Hz e percentual de fala alto, 79.26%, o que confere menor número de pausas e/ou pausas mais rápidas.

A voz 4, feminina, ocupou a primeira posição em confiabilidade e nos outros dois aspectos ficou também entre os primeiros colocados. É uma voz aguda, a segunda entre as femininas e a de percentual de fala mais alto, 78.8%.

A voz 2, feminina, que ocupou terceiro e segundo lugares para agradabilidade e confiabilidade, respectivamente, em naturalidade caiu para sétimo lugar. A medida da média de frequência fundamental dessa voz é 202.68 Hz, a mais grave dentre as femininas, talvez esse seja o fator determinante para sua baixa pontuação em naturalidade, já que se espera de uma voz feminina que ela seja aguda.

A voz 3, masculina, que ocupou oitavo lugar na preferência dos sujeitos para agradabilidade e confiabilidade, subiu para a terceira posição no aspecto naturalidade. A diferença entre frequência fundamental máxima e mínima para essa voz é a maior dentre as masculinas, assim como a diferença entre a intensidade máxima e mínima. Isto talvez confira a voz um excesso de modulação, o que não agrada em vozes masculinas.

Na correlação entre os dados acústicos e as preferências para agradabilidade, confiabilidade e naturalidade, os resultados para as vozes masculinas foram mais consistentes, conforme discutido anteriormente, exceção feita à naturalidade, que apresentou correlações fortes para as vozes femininas também.

Para o aspecto agradabilidade, a correlação mais significativa aconteceu no parâmetro percentual de tempo sem voz nas vozes masculinas, o que indicou que quanto maior fosse o parâmetro, mais agradável seria a voz. A

voz com alto grau de sonoridade confere a voz masculina virilidade, o que não corrobora esse achado.

Parâmetros relacionados ao percentual de tempo de fala nas vozes masculinas indicou que quanto menos silêncio mais agradável é a voz, ou seja, menos pausas durante a locução ou pausas mais rápidas, o que confere à locução maior consistência. No parâmetro relativo à intensidade média usada durante a locução para as vozes masculinas, com $p=0.10$, foi observada a preferência por médias mais altas. Em relação à média de F_0 para as vozes masculinas, com $p=0.025$, encontramos que quanto menor a média de F_0 , maior a preferência, ou seja, preferência pelas vozes mais graves, em relação à agradabilidade.

Por fim, também relativo às vozes masculinas, com $p=0.10$, verificou-se que quanto menor o tempo de locução maior a agradabilidade.

Para as vozes femininas em relação ao aspecto agradabilidade, o único parâmetro com correlação significativa foi a intensidade máxima, com $p=0.10$, indicando que quanto maior for o valor para a intensidade máxima usada mais agradável é a voz feminina.

Intensidade e frequência são parâmetros interdependentes, e vozes agudas tendem a ser mais intensas por facilitação da dinâmica laríngea, ou seja, o aumento da tonicidade laríngea gera maior resistência glótica e, portanto, maior intensidade (Titze, 1980; Titze, 1992; Behlau & Pontes, 1995).

Sabendo-se que intensidade e frequência são parâmetros que perceptualmente estão associados, concluiu-se que deveria haver uma preferência por vozes agudas. Embora as correlações feitas em parâmetros de medidas de frequência não tenham sido significativas, houve uma tendência

para a preferência por vozes de frequência fundamental mínima mais baixa, o que não estaria de acordo com tal observação. Tal fato poderia ser explicado também pela tendência de preferência pelas vozes femininas com maior diferença entre as frequências fundamentais máxima e mínima, o que conferiria a voz maior modulação.

Para o aspecto confiabilidade as correlações só se mostraram significativas para as vozes masculinas. As vozes apontadas como as de maior confiabilidade, por ordem de maior significância, foram as com maior percentual de fala (menos pausas ou pausas mais curtas); maior percentual de tempo sem voz; as com intensidade média mais alta; as com tempo de locução menor; as com frequência fundamental média menor e as com intensidade máxima maior.

Para o julgamento de naturalidade, o tempo da mostra é fator determinante, quanto menor, mais natural é a voz. Neste aspecto apareceram várias correlações significativas para vozes femininas, sendo as mais naturais as de frequência fundamental máxima mais alta; as com diferença entre frequência fundamental máxima e mínima maior e as de intensidade média maior. Para as masculinas, as com maior percentual de tempo de fala, as com maior percentual de tempo sem voz e as de intensidade média maior são as mais naturais.

Pareceu pelo exposto acima que fatores determinantes para o julgamento dos aspectos aqui estudados são o tempo da locução, a forma das pausas dentro desse tempo, que devem ser curtas ou em menor número, o uso da intensidade e da frequência.

Walden, Montgomery, Gibeilly, Prosek, e Schwartz (1978) em julgamentos de similaridade também encontraram parâmetros acústicos relacionados à frequência e duração como determinantes.

Russo e Behlau (1993) disseram que os fatores suprasegmentais da fala como aumento na duração da última sílaba tônica, queda na frequência fundamental e pausas do discurso oferecem informações adicionais que podem alterar completamente uma mensagem.

Uma palavra acentuada costuma ser mais intensa, mais aguda e mais longa. A ênfase é expressa principalmente através do aumento de intensidade, quer seja em certas palavras ou mesmo em frases completas (Behlau & Pontes, 1995).

No aspecto confiabilidade, em que o percentual de fala foi o parâmetro mais significativamente correlacionado com as preferências, viu-se a importância do uso das pausas que, segundo Russo e Behlau (1993), servem para enfatizar uma palavra ou expressão, ênfase que confere maior credibilidade ao discurso.

O uso da análise de MDS ajudou a verificar, que os dados dos juízes levavam as diversas vozes da amostra a se agrupar de uma determinada maneira que poderia ser representada por três dimensões distintas.

Em agradabilidade, a dimensão 1 parece ser determinada pelos parâmetros de frequência (F_0 média), mais fortemente em relação às vozes masculinas, indicando que quanto mais baixa for a F_0 média, mais agradável é a voz; o tempo da mostra (quanto menor, mais agradável), mais fortemente para as vozes masculinas; a intensidade máxima (quanto maior, mais agradável), mais significativamente para as vozes femininas; e por fim o

percentual de fala e de tempo sem voz, somente em relação às vozes masculinas.

A dimensão 2 de agradabilidade tem dois parâmetros importantes, o de intensidade, intensidade máxima para as vozes masculinas e intensidade mínima e diferença entre intensidade máxima e mínima para as femininas; e o tempo de percentual de fala, e frequência fundamental mínima para as vozes masculinas.

A dimensão 3, aparece com os seguintes parâmetros: frequência fundamental mínima e diferença entre frequência fundamental máxima e mínima para as vozes femininas e frequência fundamental média para as vozes sem distinção de gênero; e a intensidade, especificamente a máxima, para vozes masculinas e femininas.

Para melhor compreensão dos fatores que determinam as diferentes dimensões para agradabilidade, a tabela 12 resume esta discussão.

Tabela 12. Correlação dos parâmetros acústicos com as dimensões de agradabilidade.

	DIMENSÃO 1			DIMENSÃO 2			DIMENSÃO 3		
	M/F	F	M	M/F	F	M	M/F	F	M
duração da amostra (s)	**		****						
Fo máxima									
Fo mínima						**		****	
Fo média	***		*				**		
intensidade máxima	*	*				**		*	**
intensidade mínima					**				
intensidade média									
percentual tempo fala			*						
percentual tempo silêncio			*						
percentual tempo voz						*			
percentual tempo sem voz			*						
# Fo máxima e Fo mínima								**	
# Io máxima e Io mínima					*				

Legenda

P= 0.10	*	V m+ f = vozes femininas e masculinas
P= 0.05	**	V fem = vozes femininas
P= 0.025	***	V masc = vozes masculinas
P= 0.01	****	Fo: frequência fundamental
P= 0.005	*****	Io: intensidade

No aspecto confiabilidade encontrou-se para a dimensão 1 somente parâmetros relativos à frequência: frequência fundamental mínima relacionado às vozes femininas, frequência fundamental média relacionado às vozes sem distinção de gênero e diferença entre frequência fundamental máxima e mínima para as vozes sem distinção de gênero e para as vozes femininas em separado.

Na dimensão 2 de confiabilidade, a intensidade e o tempo de fala parecem ser os fatores determinantes, sendo que a intensidade máxima está correlacionada com os três e a intensidade mínima, a intensidade média e os percentuais de fala e de silêncio somente com as vozes masculinas.

Na dimensão 3, frequência fundamental, máxima para as vozes sem distinção de gênero e para as masculinas em separado estão correlacionados, assim como a frequência fundamental média para as femininas e masculinas, a diferença da frequência fundamental máxima e mínima é correlacionada para as vozes masculinas. A intensidade média e os percentuais de fala e silêncio estão correlacionados com as vozes femininas. O percentual de tempo sem voz está correlacionado com as vozes femininas e masculinas. Nesta dimensão parece ser um conjunto de parâmetros que a determina, o que faz crer que seja o timbre, que é característico de cada voz, que o nomeie, portanto não parece ser possível determinar essa dimensão com parâmetros acústicos, objetivos em sua natureza.

Na tabela 13 resume-se esta discussão.

Tabela 13. Correlação dos parâmetros acústicos com as dimensões de confiabilidade.

	DIMENSÃO 1			DIMENSÃO 2			DIMENSÃO 3		
	M/F	F	M	M/F	F	M	M/F	F	M
duração da amostra (s)									
Fo máxima							**		***
Fo mínima		*							
Fo média	**							**	**
intensidade máxima				***	**	***			
intensidade mínima						*			
intensidade média						***		*	
percentual tempo fala						**		****	
percentual tempo silêncio						**		****	
percentual tempo voz							***	*	*
percentual tempo sem voz								**	*
# Fo máxima e Fo mínima	**	*							**
# Io máxima e Io mínima									

Legenda

P= 0.10	*	V m+ f = vozes femininas e masculinas
P= 0.05	**	V fem = vozes femininas
P= 0.025	***	V masc = vozes masculinas
P= 0.01	****	Fo: frequência fundamental
P= 0.005	*****	Io: intensidade

Na análise do aspecto naturalidade nenhum parâmetro acústico foi determinante para definir a dimensão 1.

Na dimensão 2, o tempo da amostra, a intensidade mínima e a intensidade média foram determinantes em relação às vozes femininas. Os parâmetros de intensidade máxima para as vozes femininas e masculinas e de intensidade média para as vozes masculinas estão correlacionados. O percentual de tempo de fala e de silêncio estão correlacionados para as vozes

femininas. O percentual de tempo com voz está correlacionado com as vozes masculinas e o de sem voz com as vozes femininas.

Para a dimensão 3, repete-se o que se observa na dimensão 3 de confiabilidade. Há uma concorrência entre os parâmetros que a determinam, o que novamente faz referir o timbre como determinante dessa dimensão.

Na tabela 14 apresentam-se os parâmetros que definem as dimensões do aspecto naturalidade.

Tabela 14. Correlação dos parâmetros acústicos com as dimensões de naturalidade.

[illegible]

Legenda

P= 0.10	*	V m+ f = vozes femininas e masculinas
P= 0.05	**	V fem = vozes femininas
P= 0.025	***	V masc = vozes masculinas
P= 0.01	****	Fo: freqüência fundamental
P= 0.005	*****	Io: intensidade

Na análise do conjunto percebe-se que para os diferentes aspectos em suas distintas dimensões há quase sempre uma concorrência de parâmetros que as definem, o que leva a acreditar que se houvesse que nomear esse conjunto teria que se referir ao timbre da voz, que é individual e, portanto, subjetivo. Como o objetivo é delimitar parâmetros objetivos, por ora a atenção será dada às dimensões que são claramente definidas por um determinado parâmetro.

Para os aspectos agradabilidade e naturalidade não se nomeou portanto as dimensões.

No aspecto confiabilidade o parâmetro determinante da dimensão 1 é a freqüência, em especial para as vozes femininas, já que na análise em separado das vozes, a correlação também foi significativa. A dimensão 2 é determinada pela intensidade em todos os seus parâmetros, intensidade máxima, média e mínima, neste caso em especial para as vozes masculinas. Nesta dimensão também aparece o percentual de tempo de fala, que diz respeito às pausas feitas durante a locução. Para os dois parâmetros a respiração é fator fundamental dentro do trabalho de imitação, uma vez que é através dela que se conseguem os diferentes controles de intensidade, além da agilidade nas pausas, conseguida através de boa coordenação pneumofônica.

Embora a análise de MDS tenha deixado claro somente dimensões de confiabilidade, através dos parâmetros apontados nas outras dimensões, quando comparados com as preferências, confirmam o que deve ser trabalhado clinicamente com o profissional da voz em questão.

O objetivo do trabalho foi a busca de fatores objetivos, por isso não faz parte do interesse dissertar sobre o aparecimento de dimensões determinadas pelo timbre, no entanto, não se pode deixar de observar que este é um achado que corrobora com o fato de que a análise objetiva, nesse estudo feita através da análise acústica computadorizada, deve ser parte que complementa a análise da voz, importante, mas não única. Afirmação esta que já foi feita no início desta dissertação e reiterada neste momento.

Importante dizer também, que o uso dos procedimentos computacionais em análises de fala e voz exigem que se estude com cuidado os conceitos vinculados à teoria acústica da produção da fala (Camargo, 1999).

A realização de trabalhos que envolvam correlações entre aspectos de origens distintas, como os de natureza psicológica, fundamentalmente subjetivos, e os parâmetros acústicos de análise da fala, de ordem objetiva, trazem como primeira dificuldade diferenças de definição existentes entre os sujeitos. No questionário feito com profissionais da voz, procurou-se através de um levantamento prévio, direcionar a avaliação dos sujeitos para definições comuns; no entanto, a alta correlação verificada entre os três diferentes aspectos indicou a interdependência entre eles, o que fez acreditar-se que as estratégias usadas pelos juizes para a avaliação tenham sido similares para todos eles.

A importância do trabalho vocal com profissionais de voz é amplamente discutida e aprovada na literatura.

Awan (1993) comparou os fonetogramas da voz falada de sujeitos treinados e não treinados em voz cantada. Os treinados apresentaram um *range* maior do que os não treinados, o que sugere que o treino para a voz cantada influa na voz falada. Pode-se extrapolar esse conceito para o trabalho de impostação, uma vez que também existe um treino direcionado.

Ramos (1996) apontou diferenças de estratégias para produção de diferentes interpretações, e sugeriu o trabalho com articulação, práticas desenvolvidas por cantores e conhecimento da função das pausas; como intervenção com locutores de rádio.

Os fatores suprasegmentais da fala, tais como aumento na duração da última sílaba tônica, queda na frequência fundamental e pausas do discurso, oferecem informações adicionais. Se uma pausa é introduzida no meio de uma sentença, como uma vírgula, pode alterar completamente uma mensagem, passando-a de negativa a afirmativa (Russo & Behlau, 1993).

Essas afirmações estão de acordo com os achados dessa investigação que mostraram que as medidas acústicas correlacionadas às preferências dos sujeitos dizem respeito principalmente à intensidade, frequência e duração de fala. São essas medidas que também dão nome às dimensões que foram definidas para o aspecto confiabilidade.

O conhecimento dos parâmetros acústicos que se relacionaram com determinado aspecto psicológico pode aumentar a chance de sucesso da intervenção fonoaudiológica. Hollien (1993) em seu artigo que discute se a “voz de ouro” é talento ou treino, afirmou que se a educação vocal tiver boa qualidade, pode-se aumentar a qualidade de voz do aluno, especialmente dos profissionais.

Portanto, parece claro que a análise acústica é de grande valia no trabalho para imitação vocal de repórteres e apresentadores de televisão, servindo como avaliação e direcionamento do trabalho, além de *feedback* visual e auditivo para novos padrões de uso da voz.

CONCLUSÃO

Ao término desta pesquisa, alguns pontos importantes foram levantados:

- Agradabilidade, confiabilidade e naturalidade, enquanto aspectos psicológicos percebidos na voz humana, guardam interdependência entre si.
- Confiabilidade pode ser definida pela análise de MDS por três dimensões, sendo que duas delas estão associadas a parâmetros acústicos: uma com frequência e a outra com intensidade.
- Pode-se considerar no trabalho de impostação os fatores acústicos associados às dimensões dos aspectos.

No início deste trabalho, sabia-se do desafio que seria a tentativa de estabelecer parâmetros objetivos para aspectos de natureza psicológica, no entanto, ao término desse experimento reafirma-se a importância dessas tentativas para a construção do pensamento científico dentro da fonoaudiologia. Acredita-se que o caminho aqui trilhado poderá direcionar futuras pesquisas na área de voz profissional e nas demais áreas em que se fizer uso de programas de análise acústica.

REFERÊNCIAS

- Awan, S. N. (1993). Superimposition of speaking voice characteristics and phonetograms in untrained and trained vocal groups. Journal of Voice, 7(1), 30-37.
- Banse, R. & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. Journal of Personality and Social Psychology, 70(3), 614-636.
- Behlau, M. (1997). Considerações sobre a análise acústica em laboratórios computadorizados de voz. In: Araújo, R. B., Pracownik, A. & Soares, L. S. D., Fonoaudiologia Atual. Cap. 12, p.93-115. Rio de Janeiro: REVINTER Editora.
- Behlau, M. e Pontes, P. (1995). Avaliação e tratamento das disfonias. São Paulo: Editora Lovise.
- Behlau, M. S. & Ziemer, R. (1987). Psicodinâmica vocal. In: Ferreira, L. P., Trabalhando a voz – Vários enfoques em fonoaudiologia. Parte II, p.71-88. São Paulo: Summus Editorial Ltda.
- Berke, G., Hanson, D., Trapp, T., Moore, D., Gerrat, B. & Natividad, M. (1989). Office-based system for voice analysis. Archives of Otolaryngology Head Neck Surgery, 115, 74-77.
- Bielamowicz, S., Kreiman, J., Gerrat, B. R., Dauer, M. S. & Berke, G. S. (1996). Comparison of voice analysis systems for perturbation measurement. Journal of Speech and Hearing Research, 39(1), 126-134.
- Boone, D. R. (1996). Sua voz está traindo você? Porto Alegre: Artes Médicas.
- Camargo, Z. (1999). Da fonação à articulação: Princípios fisiológicos e acústicos. Fonoaudiologia Brasil, 2 (2), 14-19.
- Casmerides, M. C. B. (1988). Laboratório computadorizado de voz: Caracterização de um grupo de usuários. Dissertação de Mestrado. PUC-SP.
- Chun, R. Y. S. & Navarro, C. (1998). Voz profissional – A formação acadêmica. In: Lacerda, C. B. F. & Panhoca, I. Tempo de Fonoaudiologia. Cap 5, p.75-93. Taubaté: Cabral Editora Universitária.
- Costa, Maristela J. (1998). Listas de sentenças em português. Santa Maria: Pallotti.
- Crevier-Buchaman, L., Monfrais-Pfauwadel, M. C., Begue, D., Lauga-Houdoyer, L., Laccourreye, O. & Brasnu, D. (1993). The acoustic examination and

- use of computers. Revue de Laryngologie Otologie Rhinologie, 114(4), 311-314.
- Ekholm, E., Papagiannis, G., & Chagnon, F. P. (1998). Relating objective measurements to expert evaluation of voice quality in western classical singing: Critical perceptual parameters. Journal of Voice, 12 (2), 182-196.
- Emanuel, F. W., Lively, M. A. & McCoy, J. F. (1973). Spectral noise levels and roughness ratings for vowels produced by males and females. Folia Phoniatic, 25, 110-120.
- Eskenazi, L., Childers, D. G. & Hicks, D. M. (1990). Acoustic correlates of vocal quality. Journal of Speech Hearing Research, 33(2), 298-306.
- Gelfer, M. P. (1993). A multidimensional scaling study of voice quality in females. Phonetica, 50(1), 15-27.
- Gerrat, B. R., Kreiman, J., Barroso, N. A. & Berke, G. S. (1993). Comparing internal and external standards in voice quality judgments. Journal of Speech and Hearing Research, 36, 14-20.
- Hair, J. F. Jr., Anderson, R. E., Tatham, R. L. & Black, W. C. (1998). Multivariate data analysis. Upper Saddle River, NJ: Prentice Hall.
- Hillenbrand, J., Cleveland, R. A. & Erickson, R. L. (1994). Acoustic correlates of breathy vocal quality. Journal of Speech and Hearing Research, 37, 769-778.
- Hollien, H. & Moore, G. P. (1960). Measurements of the vocal folds during changes in pitch. Journal of Speech and Hearing Research, 3(2), 157-165.
- Hollien, H. (1993). That golden voice-talent or training? Journal of Voice, 7(3), 195-205.
- Huang, D. Z. (1995). Avaliação clínica da voz – Dr. Speech: Um laboratório de voz portátil. Tiger Eletronics Inc.
- Huang, D. Z. (1995). Um laboratório de voz acessível na prática clínica. Tiger Eletronics Inc.
- Huang, D., Lin, Sarah & O'Brien, R. (1995). Dr Speech for Windows Version 3. User's Guide. Tiger Eletronics, Inc.
- Kent, R. D. (1993). Vocal tract acoustics. Journal of Voice, 7(2), 97-117.

- Klatt, D. H. & Klatt, L. C. (1990). Analysis, synthesis and perception of voice quality among female and male talkers. Journal of The Acoustical Society of America, 87(2), 820-857.
- Köster, O., Marx, B., Gemmar, P., Hess, M. M. & Künzel, H. J. (1999). Qualitative and quantitative analysis of voice onset by means of a multidimensional voice analysis system (MVAS) using high-speed imaging. Journal of Voice, 13 (3), 355-374.
- Kreiman, J., Gerrat, B. R., Precoda, K., & Berke, G. S. (1992). Individual differences in voice quality perception. Journal of Speech and Hearing Research, 35, 512-520.
- Kreiman, J., Gerrat, B. R., Kempter, G. B., Erman, A. & Berke, G. S. (1993). Perceptual evaluation of voice quality: Review, tutorial, and a framework for future research. Journal of Speech and Hearing Research, 36, 21-40.
- Kruskal, J.B. (1964a). Multidimensional scaling by optimizing goodness of fit to a nometric hypothesis. Psychometrika, 29, 1-27.
- Kruskal, J.B. (1964b). Nonmetric multidimensional scaling: A numerical method. Psychometrika, 29, 115-129.
- Lee, L., Stemple, J. C. & Kizer, M. (1999). Consistency of acoustic and aerodynamic measures of voice production over 28 days under various testing conditions. Journal of Voice, 13(4), 477-483.
- McCoy, L. A. (1996). The power of your vocal image. Journal of the Canadian Dental Association, 62(3), 231-234.
- Moore, G. P. (1971). Organic voice disorders. Englewood Cliffs.: Prentice Hall.
- Murry, T. & Singh, S.(1980). Multidimensional analysis of male and female voices. Journal. Acoustical Society of America, 68(5), 1294-1300.
- Nepomuceno, L. A. (1984). Acústica e fonação-ressonância nasal. Acta AWHO, 3(2), 31-36.
- Nieto, A., Del Palacio, A. J., Lorenzo, F. J., Vegas, A & Cobeta, I. (1995). Computer use in voice analysis: Clinical applications. Acta Otorrinolaringologica Espanola, 46(4), 241-245.
- Novák, A. & Vokrail, J. (1995). Acoustic parameters for the evaluation of voice of future voice professionals. Folia Phoniatic et Logopaedic, 47, 279-285.

- Pegoraro-Krook, M. I. e Castro, V. C. (1994). Normative speaking fundamental frequency (SFF) characteristics of Brazilian male subjects. Brazilian Journal Medical Biological Research, 27(7), 1659-1661.
- Pereira, E. S. (1999). Medidas de ruído espectral da voz: Poder de discriminação patológica e correlatos acústico-perceptíveis. Dissertação de Mestrado, FMRP – USP.
- Pinho, S. M. R. & Camargo, Z. (1998). Laboratório de voz e fala. In: Foz, F. B., Piccarone, M. L. C. D. e Bursztyn, C. S. A Tecnologia informática na fonoaudiologia. Cap. 10, p.135-146. São Paulo: Plexus Editora.
- Pinho, S. M. R., Pontes, P. A. L & Ganança, M. M. (1988). Configurações do trato vocal nas vogais orais do português. Acta Awho, 7(2), 124-136.
- Ramos, A. L. N. F. (1996). Análise da constituição do estilo oral por radialistas: Um estudo fonético-acústico comparativo. Dissertação de Mestrado, PUC-SP.
- Rantala, L. & Vilkman, E. (1999). Relationship between subjective voice complaint and acoustic parameters in female teachers'voices. Journal of Voice, 13 (4), 484-495.
- Rodrigues, S., Behlau, M. e Pontes, P.(1994). Proporção harmônico-ruído: valores para indivíduos adultos brasileiros. Acta AWHO, 13(3), 112-116.
- Rossiter, D. & Howard, D. M. (1996). ALBERT: A real time visual feedback computer tool for professional vocal development. Journal of Voice, 10, 321-336.
- Russo, I. & Behlau, M. (1993). Percepção da fala: Análise acústica do português brasileiro. São Paulo: Editora Lovise.
- Singh, S. & Murry, T. (1978). Multidimensional classification of normal voice qualities. Journal Acoustical. Society of America, 64(1), 81-87.
- Souza, H. S. (1999). Avaliação por computador de disfonias em voz cantada. (On-line) disponível em: <http://dcc.ufmg.br>.
- Steffen, N. e Moschetti, M. B. (1997). Parâmetros acústicos de *jitter* e *shimmer* de 248 crianças de 6 a 8 anos, estudantes de Porto Alegre. Revista da Sociedade Brasileira de Otorrinolaringologia (On-line) disponível: <http://www.sborl.com.br/html/body-materia3>.
- Titze, I. R. (1980). Comments on the myoelastic-aerodynamic theory of phonation. Journal of Speech and Hearing Research, 23, 490-494.

- Titze, I. R. (1992). Acoustic interpretation of voice range profile (phonetogram). Journal of Speech and Hearing Research, 35, 21-34.
- Van Lierde, K., Moerman, M., Vermeersch, H. & Van Cauwenberge, P. (1996). An introduction to computerised speech lab. Acta Oto-Rhino-Laryngologica Belgica, 50, 309-314.
- Walker-Andrews, A. S. (1997). Infants perception of expressive behaviors: differentiation of multimodal information. Psychological Bulletin, 121(3), 437-456.
- Walden, B. E., Montgomery, A., Gibeily, G. J., Prosek, R. A. & Schwartz, D. M. (1978). Correlates of psychological dimensions in talker similarity. Journal of Speech and Hearing Research, 21, 265-275.
- Wilkinson, L. (1997). SYSTAT: 7.0 New Statistics. Chicago, IL.: SPSS, Inc.
- Wuyts, F. L., De Bodt, M. S. & Van de Heyning, P. H. (1999). Is the reliability of a visual analog scale higher than an ordinal scale? An experiment with the GRBAS scale for the perceptual evaluation of dysphonia. Journal of Voice, 13(4), 508-517.
- Yumoto, E., Gould, W. J. & Baer, T. (1982). Harmonics-to-noise ratio as an index of the degree of hoarseness. Journal of Acoustic Society of America, 71 (6), 1544-1549.

Anexo 1. Declaração de Consentimento.

DECLARAÇÃO DE CONSENTIMENTO

O objetivo desse trabalho é estudar correlatos acústicos de qualidades vocais. Para o estudo elaboramos um experimento em que você ouvirá, em *compact disc* (CD), quarenta e cinco pares de vozes que deverão ser julgadas.

No experimento você participará de uma sessão.

O experimento não apresenta risco a sua integridade física ou mental. A sessão pode durar em torno de uma hora, com a possibilidade de algumas pausas durante a sessão, caso você julgue necessário.

É importante que você finalize a sessão, porém, caso você deseje interromper o experimento, poderá fazê-lo em qualquer momento e teremos o experimento como encerrado.

Os resultados desta pesquisa poderão ser usados para publicação ou apresentados em reuniões científicas, no entanto, a identificação dos participantes será mantida em sigilo.

Declaro que estou ciente das informações acima para participar deste experimento.

Local e data: _____

Nome do participante: _____

Assinatura do participante: _____

DN: _____ Curso: _____

Obs: Você já teve ou têm algum problema de audição?

() sim () não

Anexo 2. Questionário

A-Carta de apresentação do questionário.

Ribeirão Preto, 15 de Agosto de 1998.

Você, como profissional da voz, está sendo convidado a participar de uma pesquisa que está sendo desenvolvida na Faculdade de Filosofia Ciências e Letras de Ribeirão Preto da Universidade de São Paulo, Departamento de Psicologia, para uma dissertação de mestrado.

Nesta parte da pesquisa, estaremos fazendo um levantamento de características psicológicas da voz, na tentativa de definir parâmetros para ***agradabilidade, confiabilidade e naturalidade*** da voz.

Embora estejamos pedindo uma identificação, queremos deixar claro que ela não será divulgada. Usaremos os dados de identificação somente para análise estatística e eventuais contatos que se fizerem necessários para elucidar alguma dúvida relativa às suas respostas.

Contamos com seu apoio e atenção ao responder nosso questionário. Suas informações serão essenciais para a conclusão de nosso trabalho. Colocamo-nos a disposição para quaisquer dúvidas ou esclarecimentos.

Antecipadamente, agradecemos,

Adriana C. Balieiro Panico
Fonoaudióloga

Aluna do programa de pós-graduação
em Psicobiologia, nível mestrado

Prof. Dr. Sérgio Sheiji Fukusima
Orientador

Docente do Departamento de Psicologia
e Educação da FFCLRP - USP

B- Instrução do Questionário

O questionário tem por objetivo o levantamento de características vocais relacionadas a três aspectos psicológicos da voz: agradabilidade, confiabilidade e naturalidade.

Para cada questão assinale apenas uma alternativa, com exceção da questão 7, em que serão assinaladas pelo menos 10 alternativas.

No fim do questionário, há um espaço para que você complemente sua informação com dados que julgue importantes, que não tenham sido mencionados anteriormente.

Qualquer dúvida, contacte-nos para esclarecimentos.

Mais uma vez agradecemos sua colaboração.

Após responder o questionário, remeta-o para o endereço abaixo:

a/c Prof. Dr. Sérgio S. Fukusima
Departamento de Psicologia - FFCLRP
Universidade de São Paulo
Av. Bandeirantes, 3900
CEP: 14040-901- Ribeirão Preto - São Paulo
Telefone: (016) 602.3729
e-mail: acbali@convex.com.br

C- Questionário

Assinale quais são as características que uma voz deve ter para que você a julgue **AGRADÁVEL (ou CONFIÁVEL ou NATURAL)**:

1.Intensidade vocal:

☐ adequada ☐ aumentada ☐ reduzida

2.Modulação (entonação):

☐ adequada ☐ excessiva ☐ restrita ☐ repetitiva

3.Velocidade da fala (ritmo):

☐ normal ☐ reduzida ☐ aumentada ☐ variada

4.Frequência (tom):

☐ normal ☐ agudo ☐ grave

5.Ressonância:

☐ equilibrada ☐ nasal ☐ oral ☐ laringo-faríngea

6.Qualidade vocal (timbre):

☐ rouco ☐ áspero ☐ soproso ☐ fluída ☐ feminilizada ☐ masculinizada

☐ virilizada ☐ clara ☐ infantil ☐ velada ☐ gutural ☐ impostada

Outros: _____

7.A seguir você encontrará uma lista com termos descritivos para voz. Assinale pelo menos 10 que estejam relacionados com uma voz **AGRADÁVEL (ou COMFIÁVEL, ou NATURAL)**:

abafada	arrogante	colorida	grossa	meticulosa	rouca
aberta	áspera	convicente	harmoniosa	mole	rude
afetada	assobiada	dócil	imponente	monótona	ruidosa
agitada	autêntica	estável	incisiva	ofensiva	seca
agressiva	aveludada	estridente	lenta	pesada	sedosa
aguda	baixa	expressiva	leve	potente	sensual
alegre	boa	firme	limpa	prazerosa	simpática
amável	bonita	flutuante	macia	profunda	sofisticada
ameaçadora	brilhante	forte	madura	quente	solta
anasalada	charmosa	fraca	meiga	relaxada	soprosa
apagada	chorosa	gostosa	melodiosa	ressonante	suave
apertada	clara	grave	metálica	rica	vigorosa

Outros comentários para **AGRADABILIDADE (ou CONFIABILIDADE, ou NATURALIDADE)** da voz:

Anexo 3. Estímulos usados para seleção da amostra.

1-Quadrilha, Carlos Drumond de Andrade

João amava Teresa que amava Raimundo
que amava Maria que amava Joaquim que amava Lili
que não amava ninguém.

João foi para os Estados Unidos, Teresa para o convento, Raimundo morreu de desastre, Maria ficou para tia, Joaquim suicidou-se e Lili casou-se com J. Pinto Fernandes que não tinha entrado na história.

2-Lista 6B de Maristela J. Costa, escolhida por ser a que contém maior número de fonemas (214)

Vou viajar às nove da manhã.
Meu irmão bateu o carro ontem.
Prometi a ele não contar o segredo.
Cheguei atrasada na aula.
Esta rua é perigosa.
Esqueci da bolsa na sua mesa.
Ela comprou os últimos pães.
A casa de campo já foi alugada.
Os preços não devem subir.
Não falei com sua filha.

3-Mensagem de secretária eletrônica

Você ligou meia dois cinco, três sete, quatro zero. Toalhas de banho Gentil. No momento, ninguém está disponível para atendê-lo. Para deixar seu recado, fale após o sinal, para transmitir seu fax inicie-o agora.

4-Teste SSI em português

Que ignora o fim principal é ganhar.
A porta larga para ser mais rápido.
Gosta muito crer te dá muito para.
Quarto golpe de estado e o campo.
Sempre e corre muito mais bonito que.
Confiança em minha alma cai dentro de.
Ação humilde é bem claro o céu.
Sobre minha cabeça está de Deus pai.
Grande general chega já e não creias.
Assista a aula de papel branco na.

Instrução e Aplicação da seleção de estímulo da amostra.

NOME: _____ IDADE: _____

ESCOLARIDADE: _____ SEXO: _____

VOCÊ OUVIRÁ POR DUAS VEZES UMA FITA COM 04 (QUATRO) ESTÍMULOS SONOROS GRAVADOS POR UM HOMEM E EM SEGUIDA, OS MESMOS ESTÍMULOS GRAVADOS POR UMA MULHER. SUA TAREFA SERÁ APONTAR QUAL DOS ESTÍMULOS LHE PARECE SER MAIS ADEQUADO PARA SE OBSERVAR A QUALIDADE DA VOZ QUANTO À AGRADABILIDADE, CONFIABILIDADE E NATURALIDADE; PARA TANTO, NA SEGUNDA VEZ QUE OUVIR OS ESTÍMULOS, FAREMOS PAUSAS PARA QUE VOCÊ PONTUE DE 01 (UM) A 10 (DEZ) CADA UM DELES. QUANTO MAIS PRÓXIMO DO 01 (UM) MENOS ADEQUADO SERÁ O ESTÍMULO, QUANTO MAIS PRÓXIMO DO DEZ, MAIS ADEQUADO SERÁ O ESTÍMULO.

O JULGAMENTO DEVE SE ATER AOS ASPECTOS ACÚSTICOS DA VOZ, E NÃO AO CONTEÚDO SEMÂNTICO DO TRECHO GRAVADO.

VOZ MASCULINA

ESTÍMULO 1

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

ESTÍMULO 2

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

ESTÍMULO 3

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

ESTÍMULO 4

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

VOZ FEMININA

ESTÍMULO 1

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

ESTÍMULO 2

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

ESTÍMULO 3

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

ESTÍMULO 4

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

Anexo 4. Folhas de respostas - instrução e aplicação

INSTRUÇÃO

Serão apresentadas 45 pares de vozes que você deverá comparar, analisando três aspectos:

agradabilidade, confiabilidade e naturalidade. Você deverá considerar as seguintes qualidades de voz para sua análise:

- **Voz agradável:** voz limpa, clara, firme, alegre e harmoniosa.
- **Voz confiável:** voz convincente, clara, firme, limpa, expressiva e autêntica.
- **Voz natural:** voz limpa, clara, estável e autêntica.

O julgamento, que será feito aos pares, deverá apontar qual das duas vozes é a mais agradável (confiável, natural). A seguir, você deverá atribuir uma nota entre 1 e 10, o quão agradável (confiável, natural) é a voz escolhida em relação a outra. O valor 1 indica uma agradabilidade (confiabilidade, naturalidade) mínima e o valor 10 indica uma agradabilidade (confiabilidade, naturalidade) máxima. As distâncias entre os números deverão ser consideradas equitativas, ou seja, os intervalos serão sempre idênticos.

PAR 1 (avaliar somente a voz escolhida)

Agradabilidade: a mais agradável é () voz 1 () voz 2

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

Confiabilidade: a mais confiável é () voz 1 () voz 2

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

Naturalidade: a mais natural é () voz 1 () voz 2

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

PAR 2 (avaliar somente a voz escolhida)

Agradabilidade: a mais agradável é () voz 1 () voz 2

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

Confiabilidade: a mais confiável é () voz 1 () voz 2

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

Naturalidade: a mais natural é () voz 1 () voz 2

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

PAR 3 (avaliar somente a voz escolhida)

Agradabilidade: a mais agradável é () voz 1 () voz 2

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

Confiabilidade: a mais confiável é () voz 1 () voz 2

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

Naturalidade: a mais natural é () voz 1 () voz 2

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

E assim sucessivamente até o par 45. Quando os juizes estivessem avaliando agradabilidade, na sua folha haveria somente as marcações para agradabilidade, o mesmo para confiabilidade e naturalidade.

Anexo 5. Respostas dos questionários

Histogramas das respostas para os parâmetros de Agradabilidade; (A) Intensidade, (B) Modulação, (C) Ritmo, (D) Frequência, (E) Ressonância, (F) Qualidade Vocal e (G) Termos Descritivos.

Histogramas das respostas para os parâmetros de Confiabilidade: (A) Intensidade, (B) Modulação, (C) Ritmo, (D) Frequência, (E) Ressonância, (F) Qualidade Vocal e (G) Termos Descritivos.

Histogramas das respostas para os parâmetros de Naturalidade: (A) Intensidade, (B) Modulação, (C) Rítmico, (D) Frequência, (E) Ressonância, (F) Qualidade Vocal e (G) Termos Descritivos.

Anexo 6. Resultados da seleção do estímulo da amostra.

Sujeitos	sexo	M-E1	M-E2	M-E3	M-E4	F-E1	F-E2	F-E3	F-E4
1	F	8	6	7	9	5	9	1	6
2	F	8	9	10	2	9	7	2	8
3	F	8	5	10	6	9	7	2	10
4	F	8	6	10	4	9	5	2	6
5	F	8	4	9	7	8	7	2	7
6	F	9	6	10	7	10	7	5	6
7	F	9	10	5	6	3	8	8	6
8	F	9	9	6	7	4	6	5	6
9	F	4	7	4	4	6	5	7	7
10	F	4	7	4	7	3	4	2	6
11	F	9	4	8	2	3	7	1	6
12	F	3	3	5	7	8	9	5	7
TOTAL F		87	76	88	68	77	81	42	81

1	M	7	5	10	4	10	6	3	2
2	M	8	8	10	9	9	6	5	8
3	M	6	5	8	5	9	4	1	8
4	M	3	9	1	10	4	7	1	10
5	M	7	8	9	7	4	5	4	3
6	M	5	8	10	8	3	5	7	7
7	M	10	4	8	3	5	9	5	9
8	M	9	7	10	8	8	5	7	6
9	M	8	8	9	5	6	5	5	4
10	M	8	3	9	7	3	10	10	9
11	M	6	9	5	4	7	10	4	10
12	M	9	6	5	2	7	3	5	9
TOTAL M		86	80	94	72	75	75	57	85

		M-E1	M-E2	M-E3	M-E4	F-E1	F-E2	F-E3	F-E4
TOTAL F		87	76	88	68	77	81	42	81
TOTAL M		86	80	94	72	75	75	57	85
TOTAL GERAL		173	156	182	140	152	156	99	166

Anexo 7. Dados da análise acústica Dr. Speech 3.0.

Name:FB

Sex:F

ACOUSTIC MEASURE OF SPEECH

Audio FileName = C:\VOZES\WAV_16~1\TESTE1.WAV

Starting Time = 0.00 (s) (or 0 Frames)

Ending Time = 27.68 (s) (or 2768 Frames)

Sampling Frequency (Hz) = 44100

Maximum F0 (Hz) = 367.50

Minimum F0 (Hz) = 117.91

Mean F0 (Hz) = 214.75

Standard Deviation (Hz) = 49.99

Maximum Intensity (dB) = 78.61

Minimum Intensity (dB) = 21.58

Mean Intensity (dB) = 55.80

Standard Deviation (dB) = 18.46

Percent Speech Time (%)= 66.62

Percent Silence Time (%)= 33.38

Percent Voice Time (%)= 51.81

Percent Voiceless Time (%)= 14.81

Name:FC
Sex:F

ACOUSTIC MEASURE OF SPEECH

Audio FileName = C:\VOZES\WAV_16~1\TESTE2.WAV

Starting Time = 0.00 (s) (or 0 Frames)
Ending Time = 27.22 (s) (or 2722 Frames)

Sampling Frequency (Hz) = 44100

Maximum F0 (Hz) = 380.17
Minimum F0 (Hz) = 84.97
Mean F0 (Hz) = 202.68
Standard Deviation (Hz) = 44.69

Maximum Intensity (dB) = 81.51
Minimum Intensity (dB) = 12.04
Mean Intensity (dB) = 52.35
Standard Deviation (dB) = 22.18

Percent Speech Time (%)= 57.46
Percent Silence Time (%)= 42.54
Percent Voice Time (%)= 44.71
Percent Voiceless Time (%)= 12.75

Name:MP
Sex:M

ACOUSTIC MEASURE OF SPEECH

Audio FileName = C:\VOZES\WAV_16~1\TESTE3.WAV

Starting Time = 0.00 (s) (or 0 Frames)
Ending Time = 23.41 (s) (or 2341 Frames)

Sampling Frequency (Hz) = 44100

Maximum F0 (Hz) = 400.91
Minimum F0 (Hz) = 99.10
Mean F0 (Hz) = 192.78
Standard Deviation (Hz) = 45.16

Maximum Intensity (dB) = 81.21
Minimum Intensity (dB) = 13.98
Mean Intensity (dB) = 60.36
Standard Deviation (dB) = 19.63

Percent Speech Time (%)= 72.96
Percent Silence Time (%)= 27.04
Percent Voice Time (%)= 56.47
Percent Voiceless Time (%)= 16.49

Name:RC
Sex:F

ACOUSTIC MEASURE OF SPEECH

Audio FileName = C:\VOZES\WAV_16~1\TESTE4.WAV

Starting Time = 0.00 (s) (or 0 Frames)
Ending Time = 25.28 (s) (or 2528 Frames)

Sampling Frequency (Hz) = 44100

Maximum F0 (Hz) = 400.91
Minimum F0 (Hz) = 98.44
Mean F0 (Hz) = 239.06
Standard Deviation (Hz) = 64.79

Maximum Intensity (dB) = 80.63
Minimum Intensity (dB) = 16.90
Mean Intensity (dB) = 61.82
Standard Deviation (dB) = 17.40

Percent Speech Time (%)= 78.80
Percent Silence Time (%)= 21.20
Percent Voice Time (%)= 57.08
Percent Voiceless Time (%)= 21.72

Name:LHP
Sex:M

ACOUSTIC MEASURE OF SPEECH

Audio FileName = C:\VOZES\WAV_16~1\TESTE5.WAV

Starting Time = 0.00 (s) (or 0 Frames)
Ending Time = 29.54 (s) (or 2954 Frames)

Sampling Frequency (Hz) = 44100

Maximum F0 (Hz) = 370.59
Minimum F0 (Hz) = 93.43
Mean F0 (Hz) = 188.52
Standard Deviation (Hz) = 33.51

Maximum Intensity (dB) = 80.53
Minimum Intensity (dB) = 15.56
Mean Intensity (dB) = 56.59
Standard Deviation (dB) = 20.70

Percent Speech Time (%)= 66.08
Percent Silence Time (%)= 33.92
Percent Voice Time (%)= 51.83
Percent Voiceless Time (%)= 14.25

Name:VC

Sex:F

ACOUSTIC MEASURE OF SPEECH

Audio FileName = C:\VOZES\WAV_16~1\TESTE6.WAV

Starting Time = 0.00 (s) (or 0 Frames)

Ending Time = 24.75 (s) (or 2475 Frames)

Sampling Frequency (Hz) = 44100

Maximum F0 (Hz) = 400.91

Minimum F0 (Hz) = 93.23

Mean F0 (Hz) = 238.31

Standard Deviation (Hz) = 40.56

Maximum Intensity (dB) = 79.22

Minimum Intensity (dB) = 19.08

Mean Intensity (dB) = 61.14

Standard Deviation (dB) = 16.39

Percent Speech Time (%)= 74.42

Percent Silence Time (%)= 25.58

Percent Voice Time (%)= 56.69

Percent Voiceless Time (%)= 17.74

Name:DM
Sex:M

ACOUSTIC MEASURE OF SPEECH

Audio FileName = C:\VOZES\WAV_16~1\TESTE7.WAV

Starting Time = 0.00 (s) (or 0 Frames)
Ending Time = 23.36 (s) (or 2336 Frames)

Sampling Frequency (Hz) = 44100

Maximum F0 (Hz) = 400.91
Minimum F0 (Hz) = 100.00
Mean F0 (Hz) = 165.16
Standard Deviation (Hz) = 39.53

Maximum Intensity (dB) = 81.86
Minimum Intensity (dB) = 25.11
Mean Intensity (dB) = 64.30
Standard Deviation (dB) = 13.90

Percent Speech Time (%)= 80.14
Percent Silence Time (%)= 19.86
Percent Voice Time (%)= 55.09
Percent Voiceless Time (%)= 25.04

Name:RS
Sex:M

ACOUSTIC MEASURE OF SPEECH

Audio FileName = C:\VOZES\WAV_16~1\TESTE8.WAV

Starting Time = 0.00 (s) (or 0 Frames)
Ending Time = 22.26 (s) (or 2226 Frames)

Sampling Frequency (Hz) = 44100

Maximum F0 (Hz) = 383.48
Minimum F0 (Hz) = 89.27
Mean F0 (Hz) = 173.39
Standard Deviation (Hz) = 37.27

Maximum Intensity (dB) = 80.24
Minimum Intensity (dB) = 13.98
Mean Intensity (dB) = 60.04
Standard Deviation (dB) = 18.36

Percent Speech Time (%)= 73.54
Percent Silence Time (%)= 26.46
Percent Voice Time (%)= 52.70
Percent Voiceless Time (%)= 20.84

Name:AR
Sex:M

ACOUSTIC MEASURE OF SPEECH

Audio FileName = C:\VOZES\WAV_16~1\TESTE9.WAV

Starting Time = 0.00 (s) (or 0 Frames)
Ending Time = 21.36 (s) (or 2136 Frames)

Sampling Frequency (Hz) = 44100

Maximum F0 (Hz) = 324.26
Minimum F0 (Hz) = 83.05
Mean F0 (Hz) = 123.07
Standard Deviation (Hz) = 25.19

Maximum Intensity (dB) = 81.32
Minimum Intensity (dB) = 20.00
Mean Intensity (dB) = 61.53
Standard Deviation (dB) = 14.68

Percent Speech Time (%)= 79.26
Percent Silence Time (%)= 20.74
Percent Voice Time (%)= 49.63
Percent Voiceless Time (%)= 29.63

Name: PL
Sex: F

ACOUSTIC MEASURE OF SPEECH

Audio FileName = C:\VOZES\WAV_16~1\TESTE10.WAV

Starting Time = 0.00 (s) (or 0 Frames)
Ending Time = 24.10 (s) (or 2410 Frames)

Sampling Frequency (Hz) = 44100

Maximum F0 (Hz) = 400.91
Minimum F0 (Hz) = 125.28
Mean F0 (Hz) = 241.10
Standard Deviation (Hz) = 56.08

Maximum Intensity (dB) = 75.60
Minimum Intensity (dB) = 12.04
Mean Intensity (dB) = 54.50
Standard Deviation (dB) = 18.38

Percent Speech Time (%) = 71.49
Percent Silence Time (%) = 28.51
Percent Voice Time (%) = 49.13
Percent Voiceless Time (%) = 22.37

Anexo 8. Coeficientes de correlação entre as pontuações das vozes.

Agradabilidade

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
V1	1									
V2	0,223	1								
V3	-0,061	0,697	1							
V4	0,587	0,822	0,359	1						
V5	0,092	-0,205	-0,139	-0,162	1					
V6	0,243	0,244	0,057	0,207	0,441	1				
V7	0,029	0,402	0,649	0,146	0,094	-0,163	1			
V8	0,264	0,162	0,16	0,385	0,214	-0,246	0,076	1		
V9	-0,036	0,242	0,277	0,252	-0,641	-0,478	-0,073	0,282	1	
V10	0,542	0,231	0,402	0,24	0,386	0,645	0,261	-0,129	-0,316	1

Confiabilidade

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
V1	1									
V2	-0,001	1								
V3	0,352	-0,035	1							
V4	0,29	0,2	0,594	1						
V5	0,209	-0,417	0,284	-0,168	1					
V6	0,156	0,635	0,227	0,017	-0,264	1				
V7	0,476	0,084	0,6	0,86	0,237	-0,037	1			
V8	0,274	0,09	0,543	0,134	-0,005	0,663	0,048	1		
V9	0,352	-0,136	0,254	-0,019	0,453	-0,111	0,351	-0,36	1	
V10	0,55	0,231	0,084	-0,139	0,21	0,611	0,054	0,638	-0,114	1

Naturalidade

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
V1	1									
V2	-0,399	1								
V3	0,555	0,165	1							
V4	0,451	-0,289	0,448	1						
V5	-0,175	-0,589	-0,422	0,37	1					
V6	-0,217	0,527	0,116	-0,459	-0,359	1				
V7	-0,116	0,677	0,521	-0,004	-0,434	0,233	1			
V8	-0,193	0,12	-0,037	-0,321	0,154	0,57	0,375	1		
V9	-0,148	0,097	0,14	0,203	0,273	0,374	0,432	0,67	1	
V10	-0,667	0,164	-0,593	-0,867	-0,146	0,295	-0,159	0,108	-0,117	1

Anexo 9. Dados da MDS (SYSTAT 7.0)

Agradabilidade

Monotonic Multidimensional Scaling

The data are analysed as similarities

Fitting is split between data matrices

Minimizing Kruskal STRESS (form 1) in 3 dimensions

Iteration	STRESS
0	0.051975
1	0.045279
2	0.041619
3	0.039328

Stress of final configuration is: 0.03933

Proportion of variance (RSQ) is: 0.98232

Coordinates in 3 dimensions

Variable	Dimension		
	1	2	3
V1	-0.19	.76	.25
V2	.35	-.28	.49
V3	.40	-.73	.01
V4	.33	.35	.45
V5	-1.06	.16	-.76
V6	-1.03	.04	.55
V7	.11	-.81	-.52
V8	.45	.61	-.75
V9	1.37	.18	.04
V10	-.73	-.28	.25

Confiabilidade

Monotonic Multidimensional Scaling

The data are analysed as similarities

Fitting is split between data matrices

Minimizing Kruskal STRESS (form 1) in 3 dimensions

Iteration	STRESS
0	0.080107
1	0.073781
2	0.070329
3	0.067933
4	0.066031
5	0.064497
6	0.063319
7	0.062355
8	0.061521

Stress of final configuration is: 0.06152

Proportion of variance (RSQ) is: 0.95658

Coordinates in 3 dimensions

Variable	Dimension		
	1	2	3
V1	-.27	-.15	.17
V2	1.00	.37	.67
V3	-.21	.22	-.63
V4	.12	1.02	-.32
V5	-1.14	-.66	-.10
V6	.96	-.28	.16
V7	-.48	.66	-.04
V8	.61	-.39	-.63
V9	-.97	.07	.66
V10	.39	-.86	.06

Naturalidade

Monotonic Multidimensional Scaling

The data are analysed as similarities

Fitting is split between data matrices

Minimizing Kruskal STRESS (form 1) in 3 dimensions

Iteration	STRESS
0	0.074478
1	0.065134
2	0.060995
3	0.058350
4	0.056405

Stress of final configuration is: 0.05641

Proportion of variance (RSQ) is: 0.96825

Coordinates in 3 dimensions

Variable	Dimension		
	1	2	3
V1	-.98	-.20	-.70
V2	.58	-.73	.12
V3	-.61	-.63	-.11
V4	-1.13	.19	.19
V5	-.35	1.19	.01
V6	.66	-.25	-.25
V7	.09	-.65	.43
V8	.47	.35	.24
V9	.02	.30	.54
V10	1.25	.43	-.47

Anexo 9. Diagramas de Shepard da MDS.

Shepard Diagram

Confiabilidade

Shepard Diagram

Naturalidade