

INTRODUÇÃO

“If speech is a gift, then unfortunately the gift is sometimes taken away, and we should do all we can first to understand and then try to restore such a precious facility.”

(John LAVER, 1991)

"Voice is generally thought of as a purely individual matter, yet is it quite correct to say that voice is given us at birth and maintained unmodified throughout life? Or has the voice a social quality as well as an individual one?"

(Edward SAPIR, 1927)

Início este trabalho por meio da citação acima, extraída do artigo "*Speech as personality trait*", pois, embora antiga, levanta questões fundamentais que deram origem a presente investigação. Há algum tempo venho escutando e analisando diferentes vozes de diversos sujeitos e ouvindo diversas queixas sobre voz no decorrer da minha prática profissional como fonoaudióloga.

Uma das queixas mais freqüentes refere-se ao "**não gostar**"¹ da própria voz e considerar que não é possível modificá-la, como se fosse uma característica imutável herdada ao nascer, um dom inato, próprio de cada ser humano. Contudo, pude acompanhar transformações vocais desses mesmos sujeitos, que antes não acreditavam na possibilidade de modificar suas vozes, mas que acabaram por conquistar essas mudanças por meio da exploração de potencialidades vocais adormecidas e dentro dos seus limites de produção vocal, participando de *Grupos de Vivência Voz*².

Tais mudanças mostravam que as possibilidades de produção vocal humana não poderiam estar restritas somente às condições impostas pelo organismo, ou seja, a voz não pode ser vista como algo "**dado**" ao nascimento,

¹ Os termos grafados em **negrito** no decorrer do trabalho correspondem a destaques de minha autoria.

² Esses grupos serão explicados no decorrer do trabalho.

nem como um processo **invariável** ao longo da vida.

Quero dizer, com isso, que a voz não pode ser vista simplesmente como um produto acabado, como **mera atividade laríngea**; mas como um processo flexível e dinâmico que, além dos aspectos biológicos, sofre influência de diversos outros, como os psicológicos, históricos, culturais e sociais.

Essas considerações anunciam pontos centrais que pretendo tratar nesta tese. Inicialmente, interessa-me mostrar que o fenômeno vocal não é uma questão puramente individual e invariável, o que responde, por ora, parte do questionamento de SAPIR no trecho citado.

Embora a voz de cada um tenha características individuais e relativamente fixas, em função das condições biológicas (como o sexo, a idade, o tipo e o porte físico, os aspectos anatômicos e fisiológicos da laringe e dos demais órgãos e estruturas envolvidos na fonação, entre outras), o ser humano pode também fazer ajustes no trato vocal, algumas vezes voluntários, outras vezes, involuntários, que resultam em significativas mudanças na voz. Diversos outros fatores influenciam e geram essas variações, como o padrão social, cultural e lingüístico do sujeito, o que conduz ao segundo ponto questionado por SAPIR na citação que introduz este trabalho: "*A voz tem uma qualidade social além da individual?*"

O autor estava preocupado em mostrar a presença do *background social* na fala de cada indivíduo e, desta forma, demonstrar a qualidade **social** além da **individual**, na voz de cada um.

Quando um sujeito fala, causa uma certa impressão no seu interlocutor e, embora possa parecer original, o falante está, na verdade, reproduzindo um

padrão social com um toque de individualidade, como afirma SAPIR (op. cit.). Acompanho ainda o autor, quando acrescenta que essa estrutura vocal básica, puramente individual e natural, constitui algo que não pode ser desfeito, mas que deve ser desvelado, decompondo-se as estruturas sociais e individuais sobrepostas. Tarefa, que, em parte (dentro dos limites desta investigação), me proponho a realizar nesta pesquisa.

Em outras palavras, a voz produzida por uma pessoa não é totalmente pura e imutável. Cada produção vocal é única e singular, mas nunca é exatamente a mesma, ainda que se possa reconhecer as características inerentes de quem a produziu. Uma mesma palavra, dita por uma mesma pessoa em momentos diversos, nunca será produzida de forma idêntica. Há componentes individuais e sociais que são influenciados pela cultura e pela sociedade em que o sujeito está inserido, bem como pelas condições de produção vocal, ou seja, pelo contexto do momento (a situação, os interlocutores, o estado de humor do falante, entre outras). Dentro deste complexo e fértil campo de pesquisa, meu foco volta-se a este último ponto - analisar a produção da voz de um mesmo sujeito em diferentes contextos.

Em parte, o sujeito carrega e transmite padrões da sociedade em que vive. Assim, a maneira como uma pessoa de descendência italiana, por exemplo, produziria sua voz seria, provavelmente, bastante distinta da forma como alguém de origem chinesa. Produções diversas também ocorreriam se fossem pessoas de sexos diferentes. Por outro lado, ainda que fossem sujeitos pertencentes ao mesmo grupo social, cultural e do mesmo sexo, o padrão vocal não seria o mesmo, uma vez que além dessas características relativamente fixas ou *quase-*

*permanentes*³, termo empregado por ABERCROMBIE (1967), existem os componentes individuais.

O autor considera aspectos *quase-permanentes*, aqueles que correspondem aos elementos **extralingüísticos** da linguagem, isto é, que transmitem informação acerca da identidade do falante (sexo, idade, características físicas), expressos, particularmente, na **qualidade de voz**⁴ de quem fala.

A parcela **individual** e **variável** da voz situa-se entre os aspectos que ABERCROMBIE denominou de **paralingüísticos**⁵. Estes correspondem aos ajustes temporários no trato vocal, que um sujeito realiza em função do contexto de produção vocal, e que resulta em mudanças na qualidade de voz habitual. Assim, a voz pode sofrer variações de acordo, por exemplo, com o interlocutor. A voz que um mesmo falante utilizaria para ninar uma criança, certamente, seria distinta daquela que usaria para defender um ponto de vista profissional, com um adulto, ou mesmo da que ele utilizaria em uma conversa coloquial com este último interlocutor. Diversos outros exemplos, deste tipo de variação, podem ser extraídos do cotidiano de qualquer pessoa.

O adjetivo *paralingüístico* refere-se aos componentes não verbais da fala.

³ Os termos grafados em *itálico* no decorrer do trabalho correspondem à terminologia adotada pelo autor citado.

⁴ Utilizo o termo **qualidade de voz** como proposto por ABERCROMBIE (1967), e desenvolvido por LAVER (1980), ou seja, como expressão dos traços que se fazem presentes na fala de um indivíduo. Explicarei, mais detalhadamente, esse conceito no Capítulo I, item 2.2.

⁵ O termo **paralingüístico** será retomado adiante, no Capítulo I, item 2.1.

Os gestos, os movimentos corporais, as expressões faciais, além do *tom de voz*⁶, são exemplos dessa categoria de análise como esclarece LAVER (1994).

Interessa-me, portanto, tratar, particularmente, das **variações da voz** na interação social e, assim, esta tese se caracteriza como um estudo dos aspectos **paralingüísticos** da fala, expressos na **voz** do falante num dado **contexto**.

Diante dessas colocações, retomo o início desta introdução. No desenvolvimento do meu trabalho como fonoaudióloga, fiquei impressionada com as **variações nas vozes** de profissionais diversos, que pude observar em grupos voltados à promoção de saúde vocal, inicialmente, em uma Unidade Básica de Saúde da Prefeitura Municipal de São Paulo e, posteriormente, como docente num Curso de Fonoaudiologia no interior de São Paulo⁷, no qual eram realizados Grupos de Vivência de Voz abertos à comunidade⁸.

No início dos grupos, cuja duração aproximada era de 15 encontros semanais, a grande maioria dos participantes, não tinha consciência de que suas vozes variavam segundo o contexto. Após o trabalho vocal desenvolvido, entretanto, eles demonstravam ter maior domínio e consciência dos usos de suas vozes em diferentes situações, e diversos componentes dos Grupos de Vivência de Voz relatavam estar escutando e apreciando mais a própria voz. Apresentavam maior consciência do processo vocal e dos fatores relacionados

⁶ Esclareço que *tom de voz* é a versão que utilizo para a expressão em inglês "*tone of voice*". Representa a qualidade de voz – o modo do sujeito de se expressar pela voz, que se traduz nos ajustes vocais por ele adotados e portanto, este termo não deve ser entendido como *tonalidade de voz*, isto é, a altura (frequência) do som, - grave/agudo e nem como *pitch* (correlato psicoacústico da frequência), termo corrente na literatura especializada, e que será explicitado no decorrer do trabalho. "*Tone of voice*" corresponderia, aproximadamente, ao que BAKHTIN (1990) chamou de "*entoação expressiva*".

⁷ Refiro-me ao Curso de Fonoaudiologia da UNIMEP, Piracicaba, São Paulo.

⁸ Cabe esclarecer que esses grupos não tinham caráter terapêutico e portanto, não era esperado que seus participantes apresentassem problemas vocais.

em sua produção, e demonstravam maior conhecimento dos seus limites e potenciais vocais.

Os resultados foram surpreendentes, instigantes, e muito gratificantes. Ratificaram os aspectos até aqui evidenciados de que, conforme a situação, há mudanças na fonação, na articulação e/ou nas demais características da voz e, em última instância, na **qualidade vocal**, conduzindo a mudanças no sinal acústico (PITTAM, 1994). Fica, portanto, cada vez mais necessária uma abordagem analítica que permita compreender como a voz emerge segundo a estrutura de participação social, ou seja, o contexto de produção, e não apenas como mera atividade laríngea.

Assim, a motivação inicial desta pesquisa surgiu a partir da constatação de que **havia** transformações significativas nas possibilidades de produção vocal dos componentes desses grupos (Grupos de Vivência de Voz), o que demonstrava que a voz não era uma característica individual e imutável e, portanto, tornava-se necessário compreender o fenômeno vocal, além de seu formato físico acústico, isto é, relacionado às condições sociais de produção. Essa constatação gerou diversas indagações e reflexões que me impulsionaram fortemente a direcionar meus estudos nesta linha de pesquisa, quer seja, investigar as mudanças da voz na interação social.

A literatura existente nesse campo, em particular, na Fonoaudiologia, revelou-se insuficiente para explicar a produção vocal sob essa ótica, o que implicou na busca de referenciais teóricos em outras áreas de conhecimento - na Lingüística, em especial, na Fonética e na Sociologia.

Da Lingüística, resgato autores como SAPIR (1927), ABERCROMBIE

(1967), LAVER & TRUDGILL (1979), LAVER (1968, 1980, 1981, 1991, 1994; 1995 e 2000) e PITTAM (1994), principalmente, entre outros que citarei no decorrer do trabalho, por abordarem conceitos centrais nesta análise, como aspectos paralingüísticos, qualidade e dinâmica vocais e por discutirem a necessidade de uma perspectiva de análise social da voz (LAVER, 1991 e PITTAM, 1994).

SAPIR faz referência da voz como um traço essencial na construção do estilo, no discurso oral. O autor impressionou-me pois, em 1927, levanta questões atuais e que coincidem, como apontei anteriormente, com reflexões que originaram a presente investigação.

ABERCROMBIE (1967) influenciou largamente as pesquisas na área da Fonética e, de seu importante trabalho, focalizo, particularmente, sua contribuição aos estudos da paralingüística, como também, aos aspectos que permitem identificar as características de voz do falante, ou seja, aos marcadores lingüísticos e aos conceitos de qualidade e dinâmica de voz.

LAVER, na introdução de uma coletânea de artigos que organizou em 1991 - *"The Gift of Speech"*, esclarece que seus estudos sobre qualidade de voz (LAVER, 1980, 1991 e 1994) e a perspectiva de descrição da fala presente em sua obra sofreram forte influência de ABERCROMBIE, além de dois outros autores⁹. LAVER teve grande impacto nos estudos da Fonética na década de 80 ao propor um modelo de análise fonética fundado na qualidade de voz. Além disso, introduz a necessidade de estudos numa abordagem social, questão aprofundada por PITTAM (1994). Ambos discutem aspectos necessários à uma

⁹ LADEFODEG e BOOMER, os quais não abordarei, por fugirem ao escopo deste trabalho. Para maiores detalhes, ver LAVER (1991).

investigação da materialidade fônica numa perspectiva social, que abordarei no decorrer deste trabalho.

Da Sociologia, utilizo GOFFMAN (1964, 1974, 1979 e 1981), cujos estudos despertaram meu interesse por analisar as estruturas emergentes nas situações interacionais do cotidiano. A partir de 1971 surgem seus primeiros trabalhos voltados à análise da conversação e sua obra assume uma direção sociolinguística. O autor introduz noções importantes para análise do discurso na interação, como "*footing*" (alinhamento/posicionamento) e "*frame*" (enquadre), entre outras, que serão esclarecidas ao longo deste estudo¹⁰.

Dessa forma, dadas as características desta pesquisa - de análise da materialidade fônica, ou seja, das ***variações do fenômeno vocal que ocorrem segundo o contexto de produção*** - as questões de ordem teórica seguirão fundamentalmente dois eixos temáticos básicos: um, voltado para o detalhamento fônico, abrangendo aspectos descritivos, históricos e funcionais da voz ao longo dos tempos, e o outro voltado para o estudo da organização social do discurso no contexto interacional, investigando noções essenciais para análise dos aspectos paralinguísticos da fala expressos na voz.

A preocupação em compreender o fenômeno vocal não é, certamente, recente, a forma de abordá-lo, porém, vem sofrendo transformações ao longo do tempo. Durante séculos, o estudo da qualidade de voz recebeu um tratamento considerado "impressionista", ou seja, mais subjetivo, baseado em "*rótulos*" (como "*voz aveludada*", "*voz afeminada*", por exemplo). Despertou o interesse de

¹⁰ As traduções da obra de GOFFMAN apresentam versões para o português dos seus conceitos, conforme indico entre parênteses, no entanto, utilizarei tanto a expressão em inglês como o correspondente em português.

estudiosos de épocas remotas, como o Período Clássico, em que a importância da Retórica e do Drama levaram ao desenvolvimento dessa área para fins pragmáticos de ensino e, também, para fins de diagnóstico e tratamento de doenças, impulsionando desta forma o desenvolvimento da Medicina no campo da Laringologia (LAVER, 1981; LEDEN, 1982 e PITTAM, 1994)

No entanto, somente a partir de ABERCROMBIE (1967) é que a *qualidade de voz* começa a receber um tratamento mais científico, em bases fonéticas (LAVER, 1981) e, na atualidade, o grande avanço tecnológico e o desenvolvimento dos laboratórios de voz e de síntese de fala vieram contribuir para o desenvolvimento de medidas e análises mais objetivas da produção vocal, promovendo um crescimento vertiginoso dos estudos em voz. Apesar disso, há ainda poucas pesquisas dentro de uma abordagem social (MARTZ, 1990; FERREIRA, 1993; PITTAM, 1994 ; SIMÃO & CHUN, 1997 e SERVILHA, 2000), como afirma FERREIRA (op. cit., p. 31): “...sendo a voz produto de fatores biológicos, psicológicos e sociais, parecem ser, estes últimos, os menos focalizados na literatura.”.

A descrição da voz pode ser realizada do ponto de vista perceptual, articulatorio e acústico. Entretanto, embora a descrição física, pertencente ao âmbito da Fonética, seja necessária para descrever o som e a maneira como ele é formado, essa perspectiva de análise não é suficiente para explicar como a voz funciona na interação social. Impõem-se, portanto, a necessidade de que as possibilidades analíticas para compreensão do fenômeno vocal sejam ampliadas, dentro de uma abordagem social, para que se possa apreender o papel que a voz representa nos processos de interação - sua natureza social. É de grande importância que os aspectos subjacentes a ela, e não apenas seu formato físico-

acústico, recebam atenção.

O texto abaixo é um claro exemplo de que a voz configura e transforma a palavra, acrescenta a ela o conteúdo emocional, o colorido, a expressão e a história do próprio sujeito, revelando e representando seu dono:

"Uma noite de Domingo aconteceu ter-me abeirado de um grupo de seis jovens trabalhadores bêbados, tendo ficado a uns quinze passos deles. Subitamente apercebi-me de que conseguiam exprimir todos os seus pensamentos, sentimentos e até todo um encadeado de raciocínios por meio dessa única palavra, que, ainda por cima, é extremamente breve. Um dos jovens disse-a de uma forma rude e enérgica para exprimir o seu completo desacordo com algo de que todos tinham estado a falar. Outro responde com o mesmo nome, mas num tom e num sentido totalmente diferentes - exprimindo as suas dúvidas sobre os fundamentos da atitude negativa do primeiro. Eis, senão quando um terceiro se exalta contra o primeiro, irrompendo abruptamente na conversação e gritando excitadamente a mesma palavra. Mas desta vez como se fora uma praga ou uma obscenidade. Aqui o segundo parceiro voltou a interferir, zangado com o terceiro, o agressor, retendo-o, como querendo dizer: "Tens alguma coisa que te pôr às marradas? Estávamos a discutir os assuntos calmamente e logo vens tu, metes-te, e comes logo a praguejar!" E disse todo este pensamento numa só palavra, a mesma venerável palavra; só que desta vez também levantou a mão, pondo-a sobre o ombro do companheiro.

Subitamente, um quarto, o mais novo do grupo, que até aquele momento se tinha mantido silencioso, como provavelmente tivesse encontrado repentinamente uma solução para a dificuldade inicial donde partira a discussão, levantou a mão num transporte de alegria e gritou ... Eureka, será isto? Terei encontrado a solução? Não, nem "Eureka", nem "encontrei a solução", repetiu a mesma palavra irreproduzível, uma palavra, uma simples palavra, mas com êxtase, numa explosão de comprazimento - manifestação essa provavelmente um pouco exagerada, porque o sexto membro do grupo, o mais velho deles, sujeito de aparência soturna, não gostou da coisa e cortou a alegria infantil do outro, dirigindo-se-lhe num tom de baixo solene e exortativo e repetindo ... , repetindo exatamente a mesma palavra, a mesma palavra proibida em presença de senhoras, mas que naquele momento queria dizer claramente "Para que são esses berros sem sentido?". Assim, sem terem proferido mais nenhuma palavra, nem uma sequer, repetiram aquela elocução querida seis vezes de enfiada, seis vezes sucessivas e entenderam-se perfeitamente."

(Diário de um escritor, ano de 1873. In: VYGOTSKY, L.S. – *Pensamento e Linguagem*. Lisboa:Antídoto,1979, p.187-8)

A respeito de tal relato VYGOTSKY (1979) lembra que a inflexão revela o contexto psicológico em que se deve compreender determinada palavra. Afirma que quando o contexto é claro, como no exemplo acima, torna-se possível transmitir pensamentos, sentimentos e inclusive toda uma cadeia de raciocínios

com apenas uma palavra. É importante, para os interesses desta pesquisa, acrescentar, aqui, que um dos fatores que permite essa múltipla possibilidade de interpretação e expressão são as **variações na voz** de quem fala.

O mesmo relato leva BAKHTIN (1990) a considerar que toda palavra tem um acento de valor, ou *apreciativo*, transmitido através do que ele chamou de *entonação expressiva*, mostrando que as seis "*falas*" são diferentes, entre si, apesar de todas se consistirem da mesma e única palavra. A entonação não traduziria, em si, o valor apreciativo, mas teria o papel de orientar a escolha e a distribuição dos elementos mais carregados de sentido na enunciação. Assim, cada elemento da enunciação conteria, ao mesmo tempo, um sentido e uma apreciação.

O autor esclarece que a compreensão é uma forma de *diálogo* e que a significação não estaria na palavra, no falante ou no interlocutor; mas seria, na verdade, "*efeito da interação do locutor e do receptor produzido através do material de um determinado complexo sonoro.*" (BAKHTIN, op. cit. p.132, grifo meu). Considero que esta afirmação de BAKHTIN reforça e justifica, ainda mais meu interesse em analisar e entender a voz na interação.

Mais recentemente, MADUREIRA (1992, p.20) também se envolve com a mesma questão e, em sua pesquisa, procura "*...resgatar e desenvolver a linha - tantas vezes interrompida no decorrer da história - que se ocupa da investigação da natureza do elo indissolúvel entre som vocal e sentido ...*", enfocando os efeitos dos sentidos produzidos pelos recursos fônicos.

Defende a posição de que o papel do som na construção do sentido no discurso oral pode ser descrito com base na relação entre sua manifestação no

discurso e sua manifestação física, ou, em outras palavras, afirma que a maneira como os fenômenos fônicos se apresentam no discurso pode refletir aspectos de natureza física que ocorrem na produção, transmissão e recepção dos sons vocais.

A autora destaca que uma investigação do papel do som na construção do sentido deve ser norteadada por três princípios, a saber:

“1) a epistemologia do som não pode ser desvinculada da epistemologia do sentido; 2) a inter-relação entre som e sentido dá-se de modo ambivalente e dinâmico; 3) a relação som/sentido não pode ser considerada fora do contexto situacional (conjunto de características extralingüísticas) e do contexto discursivo (conjunto de características discursivas).”

(op.cit., p.31)

Advoga também, em favor da ampliação do papel da fonética nos estudos do discurso oral, o que certamente, como ela própria afirma, poderia revelar aspectos interessantes e inéditos sobre as implicações discursivas dos traços fonéticos.

Para melhor explicitar a questão, recorro àPITTAM (1994), que reforça a necessidade de uma maior compreensão da relação entre as características físicas e a percepção do estado e da atitude do sujeito. Ele considera que estudos deste tipo necessitam uma estrutura teórica que contemple as várias dimensões na interação verbal.

A primeira dimensão, a física, necessitaria de um modelo fonético, centrado

na voz do sujeito, como o propiciado por LAVER, embora, ainda segundo PITTAM, ao relacionar os *settings* articulatórios¹¹ às medidas acústicas, LAVER afaste-se do individual como um fenômeno físico que os outros possam perceber. Essa perspectiva conduz à dimensão seguinte, a perceptual, que se refere à relação entre a medida física da voz e a percepção do outro. Para PITTAM, SCHERER (1979), assim como LAVER (embora sob óticas diferentes), contribui com um modelo teórico que possibilita esse tipo de análise. Ele estudou os marcadores sociais da fala como função da identidade do sujeito e, assim, se move do estudo da percepção da voz para as inferências e as atribuições sobre a voz. Essa abordagem nos leva à terceira dimensão, a social, como esclarece PITTAM:

“ At the third level we need models that show us how events such as social interactions develop and proceed. ... Finally, we need a broader perspective on the whole situation showing that social interactions are not just about individuals communicating, but are social events influenced by the institutional and generally sociostructural nature of society.”

(PITTAM, 1994, p.151-2)

A linha de investigação deste projeto segue essas dimensões, mas, enfoco, principalmente, a primeira e a terceira dimensões, pois meu objetivo é relacionar a medida física da voz, o nível acústico, ao nível social. Uma vez que o

¹¹ “*Setting*” é a unidade proposta por LAVER (1980), em seu modelo fonético, para análise da qualidade de voz, que corresponde a dois ou mais segmentos de unidade de fala.

fenômeno vocal não será investigado do ponto de vista individual, como marca da personalidade do sujeito, considero que a segunda dimensão não será contemplada da forma proposta por PITTAM. A dimensão perceptual servirá como ponto de referência para a análise acústica, e não me deterei, em detalhes, na mesma. Assim, o desenvolvimento deste trabalho implica num modelo fonético centrado na voz do sujeito, conforme o proposto por LAVER (1980), que possibilite uma caracterização científica, mais precisa e menos impressionista, da qualidade de voz, por meio da análise acústica da voz por um lado, e, por outro, uma abordagem que permita verificar a influência do contexto de produção vocal na qualidade e dinâmica de voz do sujeito.

Assim, o objetivo desta pesquisa é investigar as variações das vozes de sujeitos conforme o contexto de produção, em decorrência de mudanças do enquadre (*frame*) e das posições/alinhamentos (*footing*) do falante, nos termos de GOFFMAN, manifestas paralinguisticamente. Minha hipótese é que o contexto de produção modifica a voz e, também, que essas modificações são observáveis perceptual e acusticamente, ou seja, essas possibilidades de variância vocal se materializam e expressam conforme o contexto interacional e podem ser analisadas em sua manifestação acústica.

A investigação se pautará pela análise das produções vocais de dois sujeitos em contextos de fala diferentes, materialidade esta que pode ser decomposta em parâmetros físicos, mensuráveis por meio da análise computadorizada. Contudo, vale esclarecer que não seguirei o método, freqüentemente, utilizado nos estudos acústicos, que consiste no controle do conteúdo fonético (repetição do *corpus* pelos sujeitos e uso de frases

padronizadas).

O trabalho será norteado por uma abordagem naturalista/observacional (PERRONI, 1996), pois, como explica a autora, ao se recuperar a história do dado por meio da descrição das suas condições de produção, é possível identificar a constituição histórica do sentido; portanto, o não controle de variáveis pode possibilitar uma maior aproximação da naturalidade e o reconhecimento de que o que é dado é **o que acontece** e não o que **deveria** acontecer (PERRONI, 1996, grifos meus), o que coincide com os propósitos do meu trabalho.

Abordarei a preocupação com a qualidade de gravação e cuidados técnicos para evitar distorções do material (LABOV, 1972-8; PITTAM, 1994) na análise acústica, bem como o que LABOV considerou o "paradoxo do observador" - como coletar os dados sem interferir nos mesmos, uma questão de difícil solução. Explicitarei as condições de gravação e os procedimentos de coleta e análise dos dados.

Os resultados serão apresentados em três categorias de análise: contextual (contexto de produção), perceptual (auditiva)¹² e físico-acústica, e discutidos numa perspectiva social nos termos de GOFFMAN, principalmente.

O presente estudo representa um movimento em direção à análise da produção vocal numa abordagem social com repercussão nas áreas de conhecimento relacionadas ao campo da voz como a Lingüística, a Fonoaudiologia e a Medicina (Laringologia). Considero-o como um modo de mostrar uma possibilidade analítica e reflexiva distinta da abordagem tradicional

¹² Embora, no meio fonoaudiológico, esteja consagrado o uso da expressão “*análise perceptivo-auditiva*”, acompanho LAYER (1994) no emprego do termo **perceptual**. O autor esclarece que este tipo de análise compreende o registro do que é percebido sensorialmente, de diferentes formas, mas, principalmente, no nível **auditivo** e portanto, ele utiliza este termo em referência específica à percepção auditiva.

em voz no cenário brasileiro, preenchendo importante lacuna na área - um produto de uma etapa, parte de um processo em construção - como indicam os resultados que serão apresentados e discutidos na parte final deste trabalho.

Assim, divido a tese em cinco capítulos. No primeiro capítulo, recupero questões de ordem teórica para análise vocal quanto aos aspectos históricos, descritivos e funcionais. Resgato a voz através dos tempos e as mudanças de tratamento dado à qualidade e à dinâmica vocais em diferentes épocas - dos mitos à abordagem científica. Dedico um item para tecer considerações sobre a evolução do tratamento dado a voz no Brasil, no campo da Fonoaudiologia. Em seguida, trato de conceitos centrais para análise fonética da voz, quais sejam, a qualidade e a dinâmica de voz, os marcadores lingüísticos e os aspectos paralingüísticos, e retomo noções fundamentais para a análise fonética da voz.

O capítulo dois dedico às implicações de natureza social na análise vocal e, como afirma GOFFMAN, da situação social como *locus* de pesquisa - um lugar negligenciado. Abordo a possibilidade de análise do fenômeno vocal dentro da proposta de "*Frame Analysis*", proposta por este autor. Em seguida, discuto a voz na interação social – o espaço negligenciado na Fonoaudiologia em nosso país.

No capítulo três descrevo o método que utilizei nesta pesquisa, a constituição e delimitação do *corpus*, os procedimentos de coleta e análise adotados e os recursos materiais utilizados, entre eles, o *software* CSRE 4.5 (AVAAZ Innovations Inc., Canadá, 1995), "*Computerized Speech Research Environment*", um programa específico para análise computadorizada de voz.

No quarto capítulo apresento os resultados desta pesquisa e procedo a discussão dos mesmos à luz dos referenciais teóricos adotados. Por fim, no capítulo cinco apresento as considerações finais.

Aspectos importantes puderam ser revelados. Muitos outros, entretanto, ainda não que ser estudados, neste caminho nunca (ou pouco) antes navegado. A escassez de estudos em voz, fora do contexto patológico, indica o quanto ainda há que se avançar neste percurso.

Vamos, então, ao primeiro capítulo, de volta ao começo - uma breve retomada da evolução histórica e cultural da voz na Laringologia e na Fonoaudiologia, e do resgate de conceitos oriundos da Lingüística, particularmente da Fonética.

CAPÍTULO I

“Todo discurso consta ou de coisas, que são significadas, ou de coisas, que significam, quero dizer de Pensamentos e Palavras.”

(QUINTILIANO, século I d.C., traduzido por
Jerônimo Soares Barbosa, Tomo I, 1944)

ANÁLISE DA VOZ: ASPECTOS HISTÓRICOS, DESCRITIVOS E FUNCIONAIS

1. A VOZ ATRAVÉS DOS TEMPOS

“Tantos instrumentos podemos escolher:

Flauta, clarineta, corneta ou pistom,

Tímpano, pandeiro, chocalho, quantos são!

Viola, violino, violoncelo, violão.

Um já temos nós: É a nossa voz!

E que nosso canto

Seja o acalanto

A favor da vida,

Em nome do amor.

(Sheila MATOS, 1996)

A letra da música acima, extraída de um cancionário dedicado às crianças—*“Cantando, Brincando, Sonhando”*(MATOS, 1998, p.16), ilustra o quanto a voz é reconhecida como um precioso instrumento, em qualquer idade e a qualquer tempo.

Essa mesma analogia da voz a um instrumento musical foi feita, séculos antes, por Hermógenes, no “Crátilo”, descrito por PLATÃO (tradução de NUNES, 1973). Trata-se de um diálogo entre Sócrates, Hermógenes e um terceiro interlocutor, Crátilo, no qual são tecidas considerações sobre a justeza e os propósitos do uso dos nomes, além de serem feitas referências aos valores estéticos das qualidades sonoras. Hermógenes diz à Sócrates:

“- Teus nomes, Sócrates, apresentam todos os matizes imagináveis. Há pouco, quando pronunciaste a palavra ‘boulapterhoun’. Tive a impressão de dares à boca a forma de flauta, para cantares o prelúdio do nome de Atenas.”

(op.cit., p.164)

Como mostra a citação acima, o interesse por voz não é recente, remonta à Antiguidade e persiste até os dias atuais, nos quais assume importância cada vez maior. Vivemos num mundo sonoro, por excelência, em que há um número crescente de profissionais que não podem prescindir da voz como instrumento de trabalho. É muito importante, portanto, que se mostre esse percurso e se assinala as mudanças das bases descritivas da voz ocorridas no transcorrer dos tempos.

Os homens vêm utilizando esse valioso e, muitas vezes, desconhecido “instrumento” de diversas maneiras na história da humanidade. A preocupação com a voz pode ser encontrada em escritos antigos como em CÍCERO, QUINTILIANO, HIPÓCRATES, ARISTÓTELES e PLATÃO, por exemplo, citados por vários autores modernos como LAVER (1991), LEDEN (1982), MADUREIRA (1992), PITTAM (1994), STEMPLE & GLAZE (1995), SATALOFF (1996) e LOPES (1997).

O objetivo de resgatar a atenção despertada pela voz através dos tempos está em mostrar a evolução nos estudos dessa área, que começaram por interesses pragmáticos, visando o ensino (Retórica e o Drama), e partiram, primordialmente, de classificações impressionistas (como “voz afeminada”, por

exemplo).

Posteriormente, a partir dos estudos de ABERCROMBIE (1967), a qualidade vocal recebeu um tratamento mais científico, que identifica os processos de produção vocal, aborda o despertar da voz durante os séculos, correlacionando-a às impressões e às emoções causadas no ouvinte, e mostra como a forma de se ver a voz se modificou, a partir do desenvolvimento das teorias. Entretanto, conforme indica LAVER (1981), as origens de suas idéias remetem a QUINTILIANO, autor que viveu no século I d.C. e que antecipou as principais distinções entre a qualidade e a dinâmica de voz na Fonética corrente, bem como entre os aspectos extralingüísticos e os paralingüísticos. Tais conceitos foram analisados por ABERCROMBIE (1967) e serão por mim discutidos, mais detalhadamente, no item 2 deste capítulo.

Após o tratamento mais científico dado por ABERCROMBIE, mudanças de grande impacto no estudo da voz ocorreram somente na década de 80, com o modelo de LAVER (1980), que propõe um sistema fonético para descrição da qualidade de voz que, por ser fundamentado em bases articulatórias, acústicas, fisiológicas e auditivas, possibilita uma caracterização mais precisa e menos impressionista. O autor assinala que seus estudos sofreram forte influência de ABERCROMBIE (LAVER, 1991).

Na década de 90, o próprio LAVER (1991), assim como PITTAM (1994), indicam a necessidade de se preencher uma lacuna nos estudos de voz, incorporando-se, a eles, os aspectos sociais. PITTAM, inclusive, aprofunda essa discussão e realiza uma abordagem social, analisando a função vocal dentro de uma perspectiva comunicativa que se constitui em importante referência para

estudos que se realizem dentro dessa ótica.

Embora pareça que a voz tenha sido sempre um instigante objeto de estudo através da História, as possibilidades de estudos mais amplos e profundos, nos dias de hoje, são indiscutíveis, por duas razões principais: o crescente avanço tecnológico, produzindo laboratórios que permitem análises acústicas da voz cada vez mais sofisticadas e a maior preocupação com o papel social da voz. Entretanto, como o conhecimento é sempre cumulativo e um estudo sempre se apoia em estudos anteriores, achei por bem fazer, a seguir, uma breve revisão histórica e cultural da voz, baseada, principalmente, na trajetória delineada por Hans Von LEDEN. Porém, como o enfoque do autor está relacionado ao desenvolvimento da Medicina em Voz, eu, mesmo seguindo a trajetória por ele proposta, estarei dando maior destaque àqueles autores que contribuíram para a compreensão dos aspectos da qualidade e da dinâmica da voz.

1.1.Dos Mitos à Abordagem Científica da Voz – Revisão Histórica e Cultural

LEDEN (1982)¹³, médico laringologista, autor de clássica e fascinante revisão histórica da voz e da laringe - *"A Cultural History of the Larynx and Voice"*, considera que, como outros aspectos do desenvolvimento humano, a história da voz passou por quatro fases culturais: a) *fictícia* ou *mítica* (na qual o homem explica os fenômenos naturais como mágicos, como religiosos ou como eventos

¹³ O texto original de Hans Von LEDEN foi escrito em 1982, e apresentado no " *Annual Symposium on Care of the Professional Voice in New York City*", organizado pela "Voice Foundation". Foi publicado em 1991 e 1996, em SATALOFF, R.T. - **Professional Voice: The Art and Science of Clinical Care** e reimpresso em 1998, em SATALOFF, R.T. - **Voice Perspectives**.

sobrenaturais); b) *metafísica* (em que o conhecimento é baseado na observação e, principalmente, na especulação - os médicos de então, consideravam-se filósofos); c) *tradicional* (assim denominado pelo autor porque toda informação está baseada na tradição ou na revelação das grandes autoridades do mundo antigo ou da igreja), e a que o autor considera como d) *realista* (inicia-se na Renascença, quando a Medicina é vista como ciência).

A fase inicial, a *mítica*, é bem representada nos relatos de LEDEN (op.cit.) através de exemplos como o dos egípcios, que acreditavam que o poder mágico da voz estava nos pulmões e, por isso, utilizavam amuletos de pulmão para se proteger dos "*poderes do mal*", ou o da Índia Antiga, onde se associava a criação da voz e da fala aos movimentos do vento e do fogo. Cito estes exemplos por considerá-los interessantes para marcar "o ponto de partida" dessa trajetória.

HIPÓCRATES, o "Pai da Medicina", viveu na ilha de Cos, no século V a.C., e é o principal representante da segunda fase, a *metafísica*. Sua obra, *Corpus Hippocraticum*, inclui a produção de outros autores contemporâneos daquela época. Como mostra LEDEN, a voz recebe uma atenção especial em seu trabalho, sendo distinguida através de termos como "*clarity*", "*hoarseness*", "*shrillness*" e outros, tanto no diagnóstico quanto no prognóstico. Ainda que HIPÓCRATES, e seus contemporâneos, empregassem os termos faringe e laringe indistintamente, ele é das primeiras referências escritas sobre a observação da qualidade vocal, reconhecendo a importância dos pulmões e da traquéia na produção vocal, e a dos lábios e da língua, na articulação.

Outro grande filósofo também interessado em voz, foi ARISTÓTELES, que viveu em Atenas no século III a.C. LEDEN ressalta que este filósofo deu um

passo adiante, em relação a HIPÓCRATES, por compreender que as vogais e as consoantes eram produzidas de diferentes formas e por postular que a voz era produzida na traquéia e na laringe, pelo impacto do ar. Para ARISTÓTELES, esse processo recebia inspiração do espírito, que ele acreditava estar localizado no coração e nos pulmões, dando origem a frase *“the voice is the mirror of the soul”*- expressão que ainda é utilizada no século XX. ARISTÓTELES expandiu o conhecimento científico dos trabalhos em voz, teceu comentários sobre a relação entre voz e espírito e também, reconheceu a importância da voz na expressão emocional. Foi, talvez, o primeiro escritor a se referir à laringe como o órgão do qual a voz emana (STEMPLE & GLAZE, 1995).

Embora Roma dominasse a Grécia politicamente, a cultura grega predominava em Roma. Na Medicina, as atitudes, os métodos e as práticas tiveram, principalmente, origens gregas. Claudius Galen (130-200 d.C.), conhecido como *“Clarissimus”*, grande exemplo desse período, é aclamado como o fundador da laringologia e o padrinho da foniatria e da ciência da voz. Grego, recebeu formação em filosofia e ciências naturais, viveu grande parte de sua vida em Atenas, tendo sido médico do imperador Marco Aurélio. Foi o primeiro a identificar as cartilagens da laringe e, corretamente, identificou-a como o instrumento de produção da voz. Sua obra é constituída de mais de 300 livros, cerca de 120 dos quais encontram-se disponíveis para estudo e os demais perderam-se ao longo do tempo. O conhecimento dessa extensa produção é decorrente de referências em seus próprios trabalhos (LEDEN, 1982).

Dois escritores romanos do período clássico teceram comentários analíticos sobre a qualidade e tonalidade de voz, ao tratar especificamente da

Retórica e do treinamento de oradores: – CÍCERO, que viveu no primeiro século antes de Cristo, em suas duas obras, *De Oratoria* e *Brutus*, e QUINTILIANO, do primeiro século depois de Cristo, em *Institutiones Oratoriae* (LAVER, 1981 e PITTAM, 1994). Ambos descreviam a voz de forma impressionista, CÍCERO, por exemplo, utilizava rótulos como “*rough*” e QUINTILIANO, outros como “*effeminate*” (LAVER, 1981). CÍCERO também discutiu a comunicação da atitude e da emoção na fala, conforme demonstra trecho extraído de *De Oratoria* (traduzido por Watson, 1889, copilado de LAVER, 1981, p.353-4) :

“For every emotion of the mind has from nature its own peculiar look, tone and gesture; and the whole frame of a man, and his whole countenance, and the variations of his voice, sound like strings in a musical instrument, just as they are moved by the affections of the mind. For the tones of the voice, like musical chords, are so wound up as to be responsive to every touch, sharp, flat, quick, slow, loud, gentle; and yet, among all these, each in its kind has its own middle tone. From these tones, too, are derived many other sorts, as the rough, the smooth, the contracted, the broad, the protracted, and interrupted; the broken and divided, with varieties of modulation; for there is none of these, or those that resemble them, which may not be influenced by art and management; and they are presented to the orator, as colors to the painter, to produce variety.”

(De Oratoria, Book III, C.LCIII)

LAVER aponta QUINTILIANO (*Institutiones Oratoriae*) como o primeiro

escritor a tentar distinguir as características da **qualidade vocal** daquelas da **dinâmica vocal**, obtendo sucesso parcial. Suas idéias, contudo, podem ser consideradas como precursoras das principais distinções delineadas nas descrições de qualidade vocal nos estudos correntes de Fonética e, como já foi comentado, da distinção entre os aspectos da voz natural do falante (características decorrentes da anatomia do indivíduo) e os aspectos passíveis de controle voluntário pelo falante, aspectos estes analisados por ABERCROMBIE (1967, p.92) e LAVER (1968 e 1991).

No início de sua obra, (*Instituições Oratórias*, traduzida por Jerônimo Soares Barbosa, Tomo I, 1944), QUINTILIANO discorre sobre as várias definições de Eloquência, considerada por ele como a *Arte de falar*, e comenta sobre a posição de CÍCERO com relação a essa questão.

Além de arte, QUINTILIANO considera a Eloquência como ciência, conforme deixa claro no título de sua obra, *“Eloquência é a Ciência de falar bem”* (traduzido por BARBOSA, 1944), No Artigo II, parágrafo I, p.39, o autor afirma que o fim da Eloquência é o último termo a que toda arte se encaminha, ou seja, se esta é uma ciência de falar bem, o seu fim será *“falar bem”* (grifo do autor).

Já naquela época, QUINTILIANO trata de questões importantes sobre o uso, a dinâmica e a qualidade de voz, conforme pode ser visto em trecho extraído da sua obra (C.III, Book X, traduzido por Watson, 1899 e copilado de LAVER (1981, p.354, grifos meus)¹⁴:

¹⁴ LAVER vislumbra nesse trecho uma aproximação da descrição de *qualidade de voz*, na qual, o *“infinitamente diversificado”*, deste clássico escritor, corresponderia a um pequeno número de componentes básicos, cada qual divisível em várias graduações; porém, LAVER considera mais apropriado considerar essa diversificação como correspondendo à variabilidade das qualidades de vozes individuais esperada, em parte, pelo envolvimento das variáveis anatômicas.

*“The first thing to be considered is **what sort of voice we have**, and the next, **how we use it**._The natural power of the voice is estimated by its **quantity** and **quality**. Of these, the quantity is the more simple consideration, for it may be said in general that it is either much or little; but between the extremes of these quantities there are many diversities, and many gradations from the lowest tone to the highest, and from the highest to the lowest. Quality is more varied; for the voice is either clear or husky, full or weak, smooth or rough, of smaller or larger compass, hard or flexible, sharp or flat ... The general tone of the voice, however, ought to be sweet, not grating. In the **management of the voice** there are many particulars to be observed: for besides the three main distinctions of acute, grave, and intermediate, there is need of many other kinds of intonation, as the forcible, and the gentle, the higher and the lower; and of slower or quicker time. But between these varieties there are other intermediate varieties: and as the face, though it consists of very few features, is infinitely diversified, so the voice, though it has very few variations that can be named, has yet a peculiar tone in each individual; and the voice of a person is as easily distinguished by ear as the face by the eye.”*

Discorre, também, sobre os fatores dos efeitos na voz em função de diferentes tipos de abuso e fadiga vocais. O autor assinala que a qualidade de voz como qualquer outra faculdade, melhora quando se está atento a ela e

deteriora quando é negligenciada. Esclarece que, embora a atenção que um orador deva despender à sua voz seja diferente da exigida a um cantor, alguns aspectos são comuns a ambos, como as condições de saúde, uma "*garganta*" em boas condições, isto é, "*macia e flexível*", nos termos do autor. Qualquer deficiência, a voz poderá ser interpretada como áspera, rouca ou rangida. A fadiga afeta não só a voz como todo corpo.

E, assim como CÍCERO (outro escritor do período clássico, anterior a QUINTILIANO), comenta a função da voz na comunicação dos estados emocionais, afirmando que a voz é um índice da mente que apresenta muitas variações como a própria mente, como por exemplo, ao se falar sobre temas alegres, a voz flui "*num tom cheio e claro*", como se estivesse "*alegre*"; ou num momento de raiva, a voz se torna "*áspera, grossa, ameaçadora*".

CÍCERO parece ter sido um dos primeiros a sugerir a universalidade da expressão da emoção, aspecto este, ainda em discussão, dois milênios depois. Suas obras alertam para o fato de que a voz pode ser usada para comunicar emoção. Entretanto, o emprego de terminologia *impressionista*, sem uma definição clara dos termos empregados, dificulta uma compreensão mais precisa do que esse autor pretendia expressar. Para CÍCERO, todas as emoções eram consideradas associadas a um tom particular da voz, embora seus exemplos indicassem que ele, na verdade, pensava em tons complexos ou compostos, como afirma PITTAM (1994).

O interesse pela voz pode ter se originado, em parte, com o desenvolvimento do Drama e da Retórica durante o período clássico, particularmente na Grécia e, posteriormente, no Império Romano. Nas duas

culturas, o “Drama” necessitava de atores “*treinados*” para interpretar os grandes autores da época; e a “Retórica”, a partir do trabalho de ARISTÓTELES, entre outros, necessitava de oradores profissionais e de professores de oratória. Além disso, no sistema educacional grego, a arte de declamar era ensinada aos estudantes. Todo esse interesse pela expressão oral fez com que houvesse, nessa época, um grande intercâmbio entre estas áreas: educação, teatro e profissionais de oratória (RUSSELL, 1983, apud PITTAM, 1994). Entretanto, embora em “*Rethorica*” (apud PITTAM, 1994), ARISTÓTELES tenha escrito que a manipulação de três qualidades vocais, *volume*, *ritmo* e *harmonia*, fosse necessária para a produção de uma peça de retórica persuasiva, existem poucas referências na literatura clássica sobre os métodos para treinamento da voz, utilizados pelos professores desse período.

Como esclarece PITTAM, os relatos de Aristóteles demonstram que a importância da voz como “portadora” das características emocionais (“*affect*”) e da personalidade estabeleceu-se cedo, e a comunicação dessas características através da voz eram utilizadas, especialmente, nas oratórias política e legal. Como continua ocorrendo nos dias de hoje, a Lei e a Política eram as duas instituições nas quais as *técnicas retóricas* costumavam ser mais utilizadas e, desde aquela época, a preocupação com a voz sempre esteve mais ligada a profissionais de oratória ou de canto conforme aponta NAVARRO (1996), entre outros autores.

A terceira fase, que LEDEN chamou de *tradicional*, corresponde ao período de primazia da Igreja. O Novo Testamento mostra as curas milagrosas de Cristo, entre as quais existem referências explícitas aos órgãos da

comunicação, o que pode ser constatado nas palavras do Apóstolo Marcos (7, p.32-34, citado por LEDEN, 1982, p.23):

“Then they brought to him one who was deaf and had an impediment in speech and they begged him to put his hand on him. And he took him aside from the multitudes and put his finger in his ear and he spat and touched his tongue. Then, looking up to heaven, he sighed and said to him ‘Ephphatha’, that is ‘He is opened’; immediately his tongue was loosed and he spoke plainly.”

No século VI d.C, um desconhecido artista italiano criou um mosaico que é, possivelmente, a primeira pintura do órgão fonador. Esse mosaico da Igreja do Santo Apolinário, em Ravena, concluída aproximadamente em 535-538 d.C., mostra a imagem do Redentor, na qual podem ser vistas, acima do coração, a laringe e a traquéia, mantendo a crença tradicional de que a voz viria do coração.

Prosseguindo na revisão desse período, LEDEN relata que, durante os séculos seguintes, as autoridades da Igreja permaneceram nesse estado de confusão, sem conseguir concluir se a voz era, ou não, uma função do corpo humano. Teólogos eminentes preocupavam-se, além de outros mistérios, com questões do tipo: *“Deus fala?”*, *“Os anjos e os demônios falam?”*, *“Quem falou primeiro, Adão ou Eva?”*.

Como o controle intelectual da Igreja e as condições políticas instáveis da Europa desencorajavam o progresso científico, os médicos árabes passaram a imperar, e Bagdá, Cairo e Cordova tornaram-se importantes centros de referência médica desse período. Segundo afirma LEDEN, o século IX revelou o que foi,

provavelmente, uns dos maiores e mais originais *experts* na área: o médico Rhazes the Experienced (850 – 923, 237 livros) - importante contribuição árabe para o desenvolvimento da Medicina em voz, através de seus relatos sobre mudanças na voz causadas pelo alinhamento da laringe, dos nervos recorrentes, dos músculos laríngeos, do sistema respiratório e do cérebro. Indicava, como terapia, exercícios respiratórios e treino vocal com escalas musicais.

Outro autor importante daquela mesma época foi o médico Haly Abass, falecido em 994, que escreveu *Liber Regius*, considerado como o “*perfect book of the Art of Medicine*” (LEDEN, 1982), no qual há uma concisa revisão da anatomia e da fisiologia da laringe. Nessa obra, o autor faz referências à laringe como uma membrana mucosa, composta por cartilagens articuladas e por ligamentos. Aborda também sua dupla função – respiração e fonação. Além disso, declara que a voz é produzida “soprando o ar do peito”, no fechamento da laringe.

LEDEN aponta, ainda, que Avicenna (980 – 1037), o mais famoso dos médicos árabes e autor de “*Quanun*”, incluiu importantes informações sobre a anatomia e a fisiologia da laringe, além de um capítulo inteiro sobre a produção e os distúrbios de voz.

Apesar dessas preocupações, SATALOFF (1996) assinala que não houve muito avanço nos estudos da Medicina em voz até a Renascença — início da 4ª fase, a *científica* —, época dos escritos de Leonardo Da Vinci (1452 -1519) - “*Quaderni D’Anatomia*” (1500). Da Vinci, considerado um gênio versátil, propiciou substancial contribuição por acrescentar novas informações sobre anatomia, fisiologia e patologia da voz humana. Em seu livro, incluiu diversos desenhos da laringe como órgão produtor da voz, além de, entre outras

novidades, especificar a articulação, as estruturas dos lábios, língua e dentes, e atribuir diversos termos fonéticos aos sons. Segundo LEDEN (op. cit.), Leonardo DA VINCI, em seus experimentos e observações, situava-se vários séculos adiante de sua época.

LEDEN cita diversos outros autores de menor impacto, até o nascimento de Manuel Garcia em 1805, que foi reconhecido mundialmente como um importante precursor da arte médica em voz. Segundo relatos de SATALOFF (1996) e LEDEN (1982) Garcia (1805 - 1906) foi um famoso cantor de óperas. Porém, um possível estresse vocal, decorrente da sobrecarga de ensaios e apresentações, destruiu sua voz e obrigou-o a encerrar sua carreira de cantor com apenas 20 anos de idade. Tornou-se, então, um famoso professor de canto, (talvez, para evitar que outros profissionais passassem pela mesma situação que ele) no Conservatório de Paris.

No ano seguinte instalou-se em Londres. Em 1854, visitando Paris, deparou-se com o reflexo do sol nas janelas do Palácio Real. Este simples fato serviu como fonte inspiradora para que Garcia inventasse a revolucionária técnica da laringoscopia indireta: utilizando um espelho de dentista e a luz solar como fonte luminosa, ele pode ver sua própria laringe refletida num espelho de mão. Essa técnica é, até hoje, um recurso básico dos otorrinolaringologistas para a visualização das pregas vocais. Seu autor permaneceu modesto, apesar de sua valiosa descoberta e, na celebração do seu aniversário de 100 anos, evento que reuniu celebridades de diversos países, teria dito: *"What is all the fuss about; the little mirror cost only six francs."* (LEDEN, 1982, p.85)¹⁵.

¹⁵ Esta citação foi extraída da versão publicada em 1998.

Gostaria de encerrar esta retrospectiva histórica reproduzindo as palavras¹⁶ com as quais o médico otorrinolaringologista Hans Von LEDEN, encerra sua revisão histórica e cultural:

"Garcia might have added that the greatest gift of all, an object beyond price, all of us have received free of charge from our God and Creator - the glory of the human voice."

(LEDEN, 1982, p.85).

1. 2. Da Evolução da Abordagem Científica da Voz

LAVÉ (1981, p.369) utiliza uma citação de ABERCROMBIE, de um artigo escrito em 1949 - *"Forgotten Phoneticians"* , no qual este segundo autor escreveu que *"nossos antecedentes são mais velhos e melhores do que nós pensamos"*, para demonstrar que a incursão exploratória sobre os aspectos da qualidade de voz na literatura deixa claro o quanto os foneticistas modernos tinham a ganhar com os *insights* de seus predecessores.

Quase duas décadas depois que LAVÉ fez essa afirmação, pode-se constatar que, de fato, as origens de diversos conceitos adotados na Fonética atual têm, como precursores, autores como o próprio ABERCROMBIE (1967) que, por sua vez, se inspirou em outros mais antigos. Deriva daí, entre outros motivos, minha preocupação em abordar a evolução da abordagem científica da voz.

Outra razão, esta de ordem particular, são minhas raízes profissionais, que

¹⁶ Citação extraída também, da versão publicada em 1998.

me conduziram por esse caminho. Por se tratar de uma profissão relativamente nova (os primeiros cursos no Brasil datam da década de 60), a Fonoaudiologia realiza "empréstimos" de diversas outras disciplinas, como da Medicina e da própria Lingüística, por exemplo. Utiliza termos como ***qualidade de voz***, entre outros, os quais, no decorrer do curto tempo de existência desta profissão, acabam por "perder" suas origens, ou então adquirem significados consagrados pela prática clínica. Diante disso, tenho interesse, particular e especial, em resgatar a evolução da abordagem da voz, de forma a contribuir para a elucidação desses conceitos e de suas raízes.

Cabe reiterar que o conceito de qualidade de voz é fundamental para profissionais de diferentes áreas (lingüistas, laringologistas, fonoaudiólogos, entre outros), embora não haja um consenso geral sobre o mesmo. HOLLIEN (2000), em interessante discussão sobre a *"qualidade de voz ideal"*, considera que seria necessário um esforço conjunto e interdisciplinar para responder essa questão e atender as necessidades de muitos profissionais que, por sua vez, se beneficiariam com os resultados desta discussão.

O autor afirma que a maioria dos fonoaudiólogos atuais, ou não se manifesta sobre o que seria uma *"qualidade de voz ideal"*, ou a define por meio de uma lista do que *"não deveria"* se constituir como um elemento da voz *"normal"*. As discussões, nesse sentido, são ainda bastante incipientes, na literatura nacional e não considero, portanto, que haja dados suficientes que possibilitem identificar se essa afirmação corresponde também à realidade brasileira, embora a produção científica, na área de voz, em nosso país, tenha aumentando substancialmente nos últimos anos.

No sub-item seguinte faço breves considerações sobre a evolução da abordagem da voz na Fonoaudiologia no cenário brasileiro, que poderão contribuir para elucidar alguns aspectos dessa importante questão.

Feitas essas considerações, passo aos aspectos históricos da evolução da voz ao longo dos tempos, para tratar de sua abordagem científica. Recorro, novamente, a LAYER (1981), que afirma que a classificação da voz em bases fonéticas só começou em torno do século XIX, apesar de os escritores clássicos estarem atentos para o fato de que a voz poderia funcionar, social e individualmente, de diferentes maneiras.

Uma análise dos termos empregados por seguidores de Aristóteles, provavelmente no século IV d.C., por exemplo, mostra uma tentativa de defini-la em termos de “frouxidão” (*“laxness”*), nível e extensão de *pitch*, correlato psicoacústico da frequência¹⁷. É uma definição que se aproxima de uma descrição fonética o que, para PITTAM (1994), poderia ser sugestivo da competência desses autores em realizar uma avaliação auditiva da voz.

Além destes, LAYER (op.cit.) observa que CÍCERO e QUINTILIANO já demonstravam um certo interesse na fonética, uma vez que, em seus escritos, indicavam preocupação com os componentes paralingüísticos da fala, ou seja, com o fato de que o estado emocional do falante pudesse ser expresso no *“tom”* da sua voz. LAYER assinala que o estudo de aspectos paralingüísticos são fundamentais para uma análise da qualidade de voz em bases fonéticas, podendo-se, portanto, considerar a preocupação desses escritores clássicos como uma das raízes precursoras da abordagem científica.

¹⁷ Este conceito será melhor explicitado em outra parte deste trabalho. Opto por manter o termo *“pitch”* na língua inglesa, em função do uso consagrado na literatura especializada.

Na década de 80, o próprio LAYER apresenta um modelo de impacto para o estudo da qualidade vocal, descrito em "*The Phonetic Description of Voice Quality*" (LAYER, 1980). Sua proposta é a de um sistema fonético, fundamentado em dados de análise articulatória, acústica, fisiológica e auditiva, para a descrição da qualidade de voz. Conseguiu, assim, como analisa MADUREIRA (1992), uma melhor caracterização da qualidade de voz, utilizando não apenas os recursos de análise instrumental, mas também as poucas análises sobre qualidade de voz produzidas por foneticistas entre o final do século XIX e início do século XX.

O autor esclarece que a sua obra é uma tentativa de aplicação dos princípios de análise fonética na descrição da qualidade de voz. Ele parte da concepção de voz entendida num sentido mais abrangente do que o de mera atividade laríngea, ou seja, como uma característica auditiva da voz de um falante. Acredita, portanto, que tanto as características laríngeas como as supralaríngeas contribuem para a qualidade de voz. Adota a mesma conceituação de ABERCROMBIE, que considera qualidade vocal como as características presentes quase o tempo todo em que uma pessoa está falando:

"... It is a quasi-permanent quality running through all the sounds that issues from mouth. These characteristics do, naturally, include some that have their origin in the anatomy of the larynx, and are therefore concerned with phonation, but as we shall see they also include many other characteristics which have their origin elsewhere."

(ABERCROMBIE, 1967, p. 91)

Para LAYER (1980), o estudo da qualidade de voz tem maior relevância para a Lingüística por ter sido proveniente de considerações perceptuais. Ele acredita que, para decidir que aspectos da produção vocal em uma determinada língua podem ser considerados como lingüisticamente pertinentes, o lingüista foneticista terá que ser capaz de isolá-los e compará-los a outros aspectos da produção vocal, em termos perceptuais. Acrescenta ainda que, para que sua proposta seja entendida, é necessário que se adote uma visão de linguagem mais abrangente, na qual o material fonético da linguagem falada e o componente fonético da qualidade de voz se sobreponham de alguma forma. Ele explica:

“In this broader approach, the view that is taken of the linguistic accountability of phonetic theory is that phonetic theory should be responsible for describing all recurrent, patterned, phonetic activity that characterizes the spoken language of the speech community concerned.”

(LAYER, 1980, p.5)

Além de reiterar que o estudo fonético dos componentes da qualidade de voz não pode ser considerado irrelevante quando se estuda a linguagem falada, ele afirma ainda que, embora os estudos da qualidade de voz tenham muito a oferecer para as diversas disciplinas interessadas na fala, o interesse dessas disciplinas é apenas parcial. Para o foneticista, a totalidade de seu assunto se define pela fala, em todos os seus aspectos e, por isso, propõe um modelo descrito em bases fonéticas.

Segundo LAYER (1980), existem quatro características essenciais para

que um modelo possa ser qualificado como um modelo fonético. Em primeiro lugar, que tenha bases científicas e não dependa de descrições impressionistas. Para isto, as vantagens da padronização são evidentes por si. O segundo ponto é que ele permita uma análise descritiva da voz que possa ser expressa por escrito, independente de uma demonstração auditiva. Em terceiro lugar, que a análise que esse modelo propiciar possa ser reproduzível e, por último, que o modelo seja integrativo, delineado num trabalho baseado nas mesmas áreas de uma análise fonética segmental, quais sejam, análise auditiva e articulatória, acústica e fisiológica.

A partir desses critérios, aponta o autor, o modelo descritivo da qualidade de voz proposto é aplicável para a *performance* vocal de todo ser humano normal, do ponto de vista anatômico e fisiológico.

Além disso, LAVER esclarece que, embora seu sistema descritivo seja proposto em bases auditivas, todos os componentes auditivos identificados apresentam correlatos específicos em cada um dos três níveis de análise, todos passíveis de verificação instrumental, isto é, os níveis articulatório, fisiológico e acústico. Ele utiliza, como unidade analítica, o “*articulatory setting*”, termo originado de HONIKMAN (1964, apud LAVER, 1980), esclarece, porém, que o conceito geral desse termo veio de escritores anteriores¹⁸. Assinala, ainda, que adota uma visão mais simplificada desse conceito, entendido como a *postura articulatória*¹⁹, que se mantém através da fala e que pode caracterizar parte da

¹⁸ LAVER (1980) cita autores que escreveram sobre fonética nos séculos XVII (WALLIS, 1653; WILKINS, 1668; HOLDER, 1669) e XVIII (HERRIES, 1773) que introduziram o conceito geral de “*setting*” e acrescenta que foi preciso esperar três séculos até a segunda metade do século XIX para que foneticistas como SWEET, SIEVERS, STORM e JESPERSEN se interessassem pelo assunto.

¹⁹ Versão para o português de Sandra MADUREIRA (1992).

fala de um sujeito, ou ainda, toda a fala do sujeito, do grupo social ou do grupo regional.

A função dos “*settings*” seria a de estabelecer distinções fonológicas ou paralingüísticas, e todos os “*settings*” são definidos a partir de um *setting* padrão de referência, o “*setting* neutro”. Em alguns casos o “*setting* neutro” pode coincidir com aquele usado habitualmente pelo sujeito, ou com a posição de seus órgãos vocais. Se a coincidência ocorre, ou não, é irrelevante para a definição de “*setting* neutro”, cujo *status* é servir como uma linha de base que possibilite uma descrição conveniente dos demais *settings*.

LAVIER (op. cit., p.14) considera mais apropriado pensar no “*setting* neutro” como um conjunto de “*settings*” que ocorrem simultaneamente em diferentes locais do aparato vocal. Esse conjunto abrange as seguintes especificações:

- os lábios não estão protruídos,
- a laringe não se encontra levantada nem abaixada
- o diâmetro do trato vocal supralaríngeo mantém-se aproximadamente uniforme ao longo de todo seu comprimento,
- as articulações orais anteriores são realizadas pela lâmina da língua,
- a raiz da língua não se encontra avançada nem recuada,
- a fauce não exerce constrição sobre o trato vocal,
- os músculos da faringe não contraem o trato vocal,
- a mandíbula não está acentuadamente aberta nem fechada,
- o uso do sistema velofaríngeo causa nasalidade audível somente quando necessário para fins lingüísticos;

- a vibração das pregas vocais é periódica, regularmente com uso eficiente do ar, sem fricção audível, com as pregas em total vibração glotal sob tensão longitudinal moderada, tensão adutora moderada e compressão medial moderada.

As características acústicas do “*setting* neutro”, também, estão bem estabelecidas: considerando-se um indivíduo adulto do sexo masculino com o comprimento do trato vocal de aproximadamente 17 cm e a laringe com dimensões normais, quando o trato vocal apresentar aproximadamente um diâmetro equivalente ao longo de toda sua extensão e não houver a ocorrência de acoplamento com a cavidade nasal, a frequência do primeiro formante deverá ser 500 Hz e formantes mais altos com valores múltiplos ímpares deste. Em média a frequência fundamental para mulheres é de 220 Hz e para homens, 120Hz.

O autor descreve ainda, diversos “*settings*”, simples e compostos, que permitem uma caracterização mais precisa e menos impressionista da qualidade vocal. Como avalia PITTAM (op.cit.), o trabalho de LAYER introduz pontos-chaves para a descrição articulatória da voz, tais como, a padronização da terminologia, o alcance dos termos individuais e a necessidade de um modelo coerente baseado na prática fonética.

Essas mudanças - de um tratamento impressionista no Período Clássico para uma abordagem mais científica, a partir de ABERCROMBIE, precursor do modelo fonético de LAYER, e, então, para uma abordagem da voz na interação social, como a proposta por PITTAM - contribuem para uma visão abrangente da materialidade fônica, capaz de contemplar as necessidades atuais de pesquisa

do complexo fenômeno vocal, entendendo-se, como argumenta LAYER (1995), que a interação social humana transcende a natureza lingüística na troca de mensagens faladas.

Torna-se necessário, portanto, que se investigue esse fenômeno, compreendendo-se que a voz representa o falante. Nesse enfoque, cada uma das produções da linguagem falada carrega, não só a própria mensagem mas, ao mesmo tempo, a qualidade de voz habitual do sujeito. Transmite, assim, a expressão auditiva do grupo social e regional ao qual o sujeito pertence, sua identidade física e psicológica, e seu estado de humor e emocional naquele dado momento de realização do ato de fala. Uma grande parcela de informação social e pessoal sobre o falante está, portanto, contida e expressa em cada produção vocal. Estes aspectos serão discutidos mais detalhadamente adiante.

A enorme expansão dos horizontes da comunicação humana através da linguagem e da voz, ocorrida neste século, torna ainda mais relevante a busca de uma maior compreensão da natureza desse processo. Para que isto seja possível, entretanto, há que se recorrer não apenas aos fundamentos da Lingüística, mas também aos de todas as diversas ciências - Humanas, Biológicas e Exatas, acompanhando, criteriosamente, seus avanços históricos, sociais, tecnológicos e filosóficos. Por conseguinte, adiante no capítulo II abordo as implicações de natureza social para análise da voz e, para tal, recorro a autores, cujas obras fundamentam-se em bases sociais, como PITTAM e GOFFMAN.

Para finalizar estes aspectos históricos da voz teço a seguir breves considerações sobre a mesma na Fonoaudiologia²⁰.

1.3. Da Consideração da Voz na Fonoaudiologia no Brasil

“Em suma: o dizer tem sua história.”

(Eni ORLANDI, 1988)

Para que se possa compreender a evolução da abordagem da Voz na Fonoaudiologia em nosso país, abordo brevemente a história²¹ do surgimento da profissão no Brasil, resgate este que não pode ser feito dissociado do contexto sócio-econômico, político e ideológico do cenário nacional.

Como tantas outras profissões, o processo de constituição da Fonoaudiologia como atividade profissional surgiu da prática e não do ensino institucionalizado. Os ambientes político e social das décadas de 30 a 50 abriram caminho para a consolidação do fonoaudiólogo e justificaram a necessidade social de seu aparecimento.

A fase de ideação do profissional teve suas origens na década de 30 a partir de ideais de uniformização da língua do Movimento Nacionalista. Esse Congresso foi realizado em 1937, sob a organização do escritor Mário de

²⁰ Agradeço às professoras - membros da banca de qualificação II - Prof.a. Dra. Maria Cecília Perez de Souza e Silva e Prof.a. Dra. Lésle Piccolotto Ferreira, dentre outros preciosos comentários, esta sugestão para introduzir este item, de forma a contribuir para a dimensão interdisciplinar deste trabalho.

²¹ Para uma revisão mais detalhada do tema, ver FIGUEREDO NETO, L. E. – “O Início da Prática Fonoaudiológica na Cidade de São Paulo. Dissertação de Mestrado, PUC/SP, 1988.

Andrade²², e tinha como objetivo adotar uma língua padrão na pronúncia artística da língua nacional, para combater a “impureza” da língua nas manifestações culturais e manter a “ordem social”. Pretendia-se alcançar a unidade nacional, a disciplina e a ordem, por meio da uniformização da língua, intenção claramente expressa nos Anais do I Congresso Da Língua Nacional Cantada, como indica FIGUEREDO NETO (1988). De ordem mais geral – a uniformização da língua era vista como importante instrumento de unificação, fortalecimento e defesa do país, tendo, portanto, um importante papel na formação da consciência nacional. A maioria dos trabalhos apresentados no Congresso preocupava-se com os aspectos dialetais das várias regiões brasileiras, embora uma abordagem sobre “os *desvios*” da língua tenha aparecido em uma pesquisa realizada nos Parques Infantis da Prefeitura de São Paulo.

Interessante assinalar que nesse estudo - ‘*Vícios e Defeitos nas Falas das Crianças dos Parques Infantis*’ – os autores²³ já destacavam a importância em se tratar da voz, considerando-a um elemento tão relegado quanto a língua nacional. Assim, além dos “*distúrbios na articulação*”, os autores, também, apontavam os “*distúrbios na voz*” e indicavam a necessidade de medidas profiláticas, terapêuticas e pedagógicas para “*correção*” desses “*vícios e defeitos das palavras*”.

Os autores dessa pesquisa justificavam a necessidade de se cuidar da fala e da voz, considerando que suas perturbações levariam o Homem “*a uma posição*

²² Na época, Diretor do Departamento de Cultura do Teatro Municipal de São Paulo.

²³ MIRANDA, NICANOR e REIS, J. D. BUENO (1937, apud FIGUEREDO NETO) – Separata do I Congresso de Língua Nacional Cantada).

de visível inferioridade social” , que o privaria de felicidade e êxito (MIRANDA, NICANOR e REIS, J. D. BUENO, 1937, apud FIGUEREDO NETO, 1988). FIGUEREDO NETO (op.cit.) esclarece a posição dos autores da seguinte forma:

“O ‘enfoque social’ dado pelos autores ao tratarem de manifestações da voz e da fala está estritamente ligado a uma conotação valorativa do uso da língua, onde o Homem vale pela forma como se expressa.”

(FIGUEREDO NETO, 1988, p.30)

Entretanto, como afirma a autora (grifos meus), na descrição dos processos de produção e perturbações da voz e da fala, uma preocupação social real foi desconsiderada nessa pesquisa e as explicações estavam associadas a um referencial exclusivamente **organicista**, em que a fala era vista como mero resultado do **funcionamento de vários órgãos**.

Os estudos dialetais e os dos desvios da língua apresentados no Congresso enfocavam a análise e a descrição da produção dos sons da língua. A concepção de língua defendida nesse evento conduzia à sistematização de uma língua padrão. Na época, a vigência das idéias nacionalistas levavam à eleição de um único substrato da língua que deveria ser imposto e imperar sobre os demais, procurando-se justificar *“lingüísticamente”* a supremacia do dialeto carioca perante os outros, considerando-se que, na época, o Rio de Janeiro era a capital do país. Uma opção que não tinha um caráter valorativo, mas estava condicionada ao peso do poder político. Assim, como analisa FIGUEREDO NETO (op. cit. p.54), *“o significado (signo) perde valor de análise e o que importa é a uniformidade do uso do*

senal". Dessa forma, o conteúdo ideológico, a historicidade da língua e a consciência individual na produção da língua são descartados, ou seja, em nome da uniformidade, relegam-se as variações do uso concreto da língua.

A autora concluiu que o "Congresso Nacional da Língua Cantada" continha uma visão de língua como sistema único e fechado e o desvio era concebido dentro dessa mesma lógica, ou seja, tudo que escapasse ao sistema de língua eleito era considerado como "erro". Nesse contexto, a Medicina e a Educação assumiam a tarefa de eliminação dos "*desvios da língua*", tendo em vista a preservação da norma. Para tal, apontava-se a necessidade de um especialista, cujo perfil, ainda não estava claramente delineado.

Originaram-se assim, os "*ortofonistas*" (*orto* = perfeição e *fono* = fala). O profissional foi concebido como um professor especializado que deveria atuar na profilaxia e correção dos "*erros da língua*".

O uso do termo "*ortofonista*", precursor de fonoaudiólogo, expressava o profissional que lidava com o "*erro*", distanciando-o do global do indivíduo, ou seja, da linguagem e seu funcionamento e da atividade educacional, perdendo-se assim, a dimensão de desenvolvimento global e focalizando-se apenas, a "*doença*", o "*erro*".

Iniciativas concretas de atuação do fonoaudiólogo ocorreram dentro de instituições, nas décadas de 40 e 50, distanciando-se do enfoque educacional e passando para um perfil mais *clínico*. Em consequência das guerras, passou-se a enfatizar a área de *reabilitação*, conferindo à Fonoaudiologia o perfil de *paramédica*, condicionada à figura do médico.

Somente na década de 60 foram criados os primeiros cursos de Fonoaudiologia em nível superior, em São Paulo. Um deles, vinculado inicialmente, à Clínica de Otorrinolaringologia da Faculdade de Medicina da USP (1961), e somente, na década de 70 passou a se constituir como curso oficial da Universidade de São Paulo, permanecendo até hoje ligado à Faculdade de Medicina da USP e, portanto, com uma formação mais voltada à área de Saúde. O outro curso, da Pontifícia Universidade Católica de São Paulo (1962), era vinculado à Psicologia. Efetivou-se, nesse período, uma atuação profissional prática, técnica e especializada, tutelada pela Medicina e pela Psicologia. Sob essa tutela estabeleceram-se vínculos de dependência e subordinação. A atividade do fonoaudiólogo voltava-se para a **doença**, alimentando a dicotomia **saúde x doença e normalidade x anormalidade**.

Assim, a década de 60 foi marcada por uma atuação em voz um pouco "científica e artística", em função da presença de médicos na direção dos primeiros cursos de Fonoaudiologia e das primeiras iniciativas junto aos laboratórios de fonética acústica e junto a professores de canto, oratória e dicção (FERREIRA, 2000). Nas palavras desta autora:

"O trabalho do fonoaudiólogo, diretamente atrelado ao do médico, surge desde os seus primórdios: não é possível atender um disfônico que nos procure, sem encaminhá-lo antes a um exame médico, para se saber como estão as suas cordas vocais. A maioria dos formados na época fugia dos disfônicos (assim como dos gogos...)"

(FERREIRA, 2000, p.128)

Acompanho a divisão por décadas, utilizada por FERREIRA (op. cit.), para seguir descrevendo a evolução da abordagem da voz na Fonoaudiologia.

Foi a partir do vínculo estabelecido com a Argentina, que a denominação *Fonoaudiologia* passou a ser utilizada no Brasil. Surgiram, então, outros cursos, como o da Pontifícia Universidade Católica de Campinas (1971) e o da Escola Paulista de Medicina, hoje Universidade Federal de São Paulo.

Na década de 70, o tratamento dado à voz possuía um cunho eminentemente **clínico e curativo**. O procedimento terapêutico era hierarquizado e apoiado em exames realizados com o "*espelho*", técnica idealizada por Manuel Garcia em 1855, um exame que, por vezes, deixava os fonoaudiólogos inseguros e à mercê do que o profissional (médico) visualizasse. Destaco algumas considerações que FERREIRA faz em referência à essa época:

*"Poucas iniciativas davam conta de questões relacionadas à voz sem ser no contexto patológico. ... Enveredando por caminhos em que predominava a prática exclusiva de exercícios, o foco era pura **manifestação orgânica**. Nesse período havia uma primazia da técnica... . O olhar do fonoaudiólogo **focalizava a voz do sujeito em sua dimensão orgânica** sem qualquer consideração de outra natureza."*

(FERREIRA, 2000, p.128, grifos meus)

A análise da autora coincide com a preocupação que venho expressando desde o início deste trabalho, ou seja, compreender a produção vocal além da

esfera biológica, tendo em vista contemplar aspectos **relegados** ao longo da história da profissão.

Em função de suas origens e dos "*empréstimos*", marcadamente influenciados por uma visão organicista e positivista, que a Fonoaudiologia foi absorvendo ao longo de sua curta existência, a linguagem, neste meio profissional, sempre foi, primordialmente, vista sob o prisma patológico, o que levava o fonoaudiólogo a atuar de uma forma limitada, fundado numa concepção de linguagem, e de sujeito, fragmentada e reducionista, o que se refletia na prática clínica em voz, aspecto este apontado por alguns autores, como FERREIRA; ALGODOAL & ANDRADA e SILVA (1998, p.394) :

“Desde os primórdios da Fonoaudiologia a idéia de que a voz é o resultado de fatores orgânicos (uma vez que para produzi-la músculos, cartilagens, mucosas, inervações... estão em ação, determinados por fatores como sexo, idade, constituição física), psicológicos (quem ainda não percebeu sua voz diferente num momento de irritação...) e sociais (quantas vezes, inconscientemente até, nossa voz não se torna mais grave ou intensa quando queremos transmitir autoridade a quem nos ouve...), sempre foi compreendida e aceita. No momento da efetivação do processo terapêutico, contudo, provavelmente em função da forte herança que recebemos da Medicina, privilegiávamos os aspectos puramente orgânicos.”

Posteriormente, mesmo com a conquista de outros espaços, como escolas e creches, por exemplo, a atuação do fonoaudiólogo não se modificou. Limitou-se a estender sua abordagem clínica para esses novos locais. Prevalecia assim, uma visão **patológico/tecnicista** e, assim, a consciência e o papel social do profissional não se alteraram significativamente. Este, muitas vezes sem ter consciência disto, contribuiu para a defesa dos interesses da classe dominante e para a legitimação de um modelo de língua determinado.

Gradativamente, o fonoaudiólogo passou a se integrar na equipe de profissionais que atuavam na área da Saúde Pública, o que impôs a necessidade de uma revisão de sua atuação, uma vez que o instrumental adquirido na universidade, voltado fundamentalmente para uma atenção individual, não atendia aos interesses da coletividade, uma das demandas da Saúde Pública.

Privilegiando as ações de promoção e prevenção da saúde, o fonoaudiólogo ascendeu para um novo tipo de abordagem: o de compreender a linguagem extrapolando seu aspecto patológico; começou, então, em algumas instâncias, a percebê-la como meio de aperfeiçoamento intelectual e como instrumento de luta, de trabalho e de adaptação do homem ao seu meio. Paralelamente, a Fonoaudiologia vivia, em seu interior, um intenso movimento de discussão das concepções de linguagem, sofrendo forte influência das chamadas vertentes “construtivistas” e “interacionistas” de linguagem.

Em função desse intenso debate, particularmente, a partir da década de 80, alguns cursos de Fonoaudiologia começaram a rever os conteúdos de linguagem em suas grades curriculares. A área de voz, contudo, não acompanhou essa reflexão no mesmo compasso.

FERREIRA (1993, p.31) assinala que, quando a Fonoaudiologia iniciou este debate em linguagem, houve profissionais que acreditavam que uma discussão, em uma perspectiva dialética era possível em linguagem, mas não, em voz.

A autora levanta importante questão sobre esse posicionamento:

“Mas, será que é possível a um terapeuta ter uma determinada postura frente a um paciente, e outra, completamente diversa, frente a outro paciente, só porque a patologia dos dois é diferente? E a concepção de homem que subjaz a essas posturas, como ficaria?”

Esses questionamentos eram pouco freqüentes no cotidiano do fonoaudiólogo que atuava em voz. As modificações durante esse período, eram mais lentas e graduais, se comparadas ao avanço da reflexão sobre linguagem. Alguns profissionais começaram a questionar a quantificação nos achados da avaliação vocal e a considerar a dinâmica psicossocial do sujeito disfônico. Um movimento, contudo, ainda restrito. Aspectos psicossociais, começaram, então, a ser valorizados na prática terapêutica, além dos, já privilegiados, componentes orgânicos.

O advento da fibra ótica permitiu a visualização das pregas vocais em movimento, possibilitando maior compreensão da fisiologia vocal. Para alguns profissionais isso permitiu reafirmar a importância dos aspectos orgânicos, que já podiam ser vistos e discutidos interdisciplinarmente. Para outros, promoveu-se o questionamento acerca do que a fisiologia não podia explicar no padrão vocal

apresentado pelo sujeito, e incentivou-se a busca de respostas, além dos limites da esfera biológica (FERREIRA, 2000).

A área de voz viveu seu momento de efervescência na década de 90, por forte influência do mercado de trabalho e do grande desenvolvimento científico e tecnológico da Fonoaudiologia. Surgiu um maior número de publicações nacionais e de profissionais inseridos no meio acadêmico, resultando em mais pesquisas sobre voz.

Por outro lado, no mercado de trabalho aumentou o número e o tipo de profissionais que necessitam da voz como instrumento de trabalho, como os operadores de *telemarketing*, os leiloeiros, os locutores de rodeio, os operadores de Bolsas de Valores e os religiosos (padres e pastores), entre vários outros. Segundo dados do manual elaborado pela Sociedade Brasileira de Laringologia e Voz (2000), em comemoração ao Dia Nacional da Voz, 25% da população ativa utiliza a voz profissionalmente. Diversos fonoaudiólogos vêm contribuindo para a ampliação do raio de ação nesse campo a partir da atuação com outras categorias, como por exemplo, advogados (CAVALVANTI & BOMPET, 2000) e professores de academia de ginástica (RAMECK, 2000).

Alguns outros profissionais, como atores, cantores e professores que, tradicionalmente buscavam o fonoaudiólogo, somente quando já apresentavam uma alteração vocal instalada, procuram, atualmente, um trabalho voltado para o aprimoramento vocal e para a prevenção de problemas de voz. Não querem perder espaço no mercado de trabalho e, ao contrário, necessitam se aperfeiçoar para poder competir em melhores condições.

A título de ilustração, cito uma grande emissora de TV brasileira, que possui, desde o início dos anos 90, um fonoaudiólogo em cada rede associada sua, no país, acompanhando essa exigência de mercado. Antes, porém, já haviam fonoaudiólogos atuando, de forma mais esparsa, junto à essa mesma emissora, contudo, vale lembrar que os trabalhos pioneiros em promoção da Saúde Vocal ocorreram na década de 80, como os de PINTO & FURCK (1988) voltado para professores, OLIVEIRA (1988) para locutores de rádio e QUINTEIRO (1989) para atores.

As necessidades e tendências do mercado e conseqüentes exigências na formação acadêmica do fonoaudiólogo, associadas ao vertiginoso crescimento tecnológico provocou um grande desenvolvimento da área de voz em nosso país e implicou em rever conceitos e formas de atuar na última década.

Surgiu, assim, o termo ***Voz Profissional***, embora não seja possível precisar quando essa expressão começou a ser utilizada pelos especialistas para definir a área de atuação com profissionais que dependem da voz como instrumento de trabalho, como afirmam FERREIRA et al. (1995) e CHUN et al. (2000). Na literatura especializada, nacional e estrangeira, há diversas referências de uso dessa expressão, como por exemplo: SATALOFF (1996), STEMPLER & GAZE (1995), FERREIRA et al. (1995), RODRIGUES et al. (1996), LOPES FILHO et al. (1997), FERREIRA & COSTA (2000), entre outras. Uma das publicações nacionais, pioneira sobre a atuação fonoaudiológica com os usuários de voz profissional, foi organizado por FERREIRA em 1988.

Em consequência dessas transformações, o **normal** e o **patológico** se confrontam o tempo todo nesta década, como esclarece FERREIRA (2000, p.129):

"A plasticidade vocal presente em alguns profissionais desmascara o determinismo e o reducionismo da explicação orgânica, evidenciando o quanto a voz é um processo dinâmico e o como é processada nas relações com o outro".

Contudo, poucos são os trabalhos nacionais que buscam transcender essa dimensão orgânica e, entre eles, cito os de MARTZ (1990, 1999); FERREIRA (1993 e 2000); SIMÃO & CHUN (1997); FERREIRA, ALGODOAL e ANDRADA & SILVA (1998) e SERVILHA (2000). Dedicarei maior atenção a alguns desses estudos no Capítulo II, com a intenção de discutir a análise da voz na interação social.

Pesquisas de tal tipo são importantes, porém escassas, pois, a produção vocal traz em seu bojo a produção de sentidos, expressa na relação entre o dito e o compreendido e entre os discursos e as práticas. (FERREIRA, 1993).

SERVILHA (2000) observa em seu estudo que, de forma marcante, as atuais pesquisas, tanto nacionais como estrangeiras²⁴, apresentam, primordialmente, um eixo voltado a medidas remediativas e preventivas, apoiado

²⁴ A autora realizou uma revisão de estudos sobre a voz do professor, mas que de uma certa forma refletem o estado das pesquisas em voz, como ela própria observa, ao analisar os anais de importante evento na área de voz ocorrido no Brasil, o qual me reporto um pouco mais a frente neste mesmo item.

em modelos teóricos influenciados por uma visão positivista da Medicina, como mencionei anteriormente.

Esse tipo de abordagem contribui para que se perpetue uma tipologia da qualidade de voz baseada na patologia vocal, isto é, centrada na “doença”. Uma tipologia que parte de valores morais e culturais, entre outros, considerados como ideais e dominantes para o profissional. *“Tudo que foge desta teia de características classificadas nesta tipologia ideal é deixada de lado.”* (FERREIRA, 1993, p.33).

Decorre dessa posição, uma classificação de voz, freqüentemente, encontrada na literatura especializada, seja nacional ou estrangeira, que considera a qualidade vocal como *“rouca”, “áspera”, “soprosa”,* entre outros tipos de voz. Essa forma de classificação e análise da qualidade de voz relaciona-se, essencialmente, à dimensão biológica, ainda que, alguns autores reconheçam que a voz carrega elementos de outras dimensões – psicológica, social e cultural, como afirmam, por exemplo, BEHLAU & PONTES (1995).

Concordo com SERVILHA (2000), que não se trata de negligenciar (e eu acrescentaria, nem de discutir), a relevância dessas abordagens, pois, elas ocupam papel crucial na prática clínica, e sua contribuição científica é inquestionável. Trata-se, sim, como indica a autora, da necessidade de se buscar um outro eixo de análise que:

“...revele o movimento das características da voz nos diferentes contextos interacionais, descobrindo outros processos que não aqueles vinculados à presença do distúrbio (remediação) ou a minimizar o potencial de seu surgimento (prevenção).

(SERVILHA, 2000, p.10)

Faltam pesquisas que contribuam para um estudo da voz que vá além dos aspectos orgânicos e patológicos. SERVILHA (op. cit.) reafirma essa posição, ao tecer comentários sobre os trabalhos apresentados no – “2nd World Voice Congress”, realizado no Brasil (São Paulo, 1999), os quais indicaram que a voz vem sendo estudada sob diferentes óticas, mas, que uma perspectiva voltada para a compreensão da voz na interação social, não ocupa, ainda, o devido espaço no âmbito das pesquisas na área (*Scientific Program*, 1999 e SERVILHA, 2000).

A autora assinala a carência de estudos que considerem a voz como variável determinante – um elemento de determinação recíproca, isto é, além de ser determinada pelo contexto, também, ocasiona mudanças vocais nos participantes de uma interação verbal.

É dessa relação de mútua determinação, da voz como processo dinâmico, ou ainda, das variações que a voz do sujeito sofre em função do contexto de produção vocal que trata esta tese. Esta é a tarefa que me coloco: buscar escapar das **“armadilhas da patologia”**, e evitar ficar à deriva por mares nunca dantes navegados, ou pouco conhecidos, mas já, apontados por sábios antecessores, como SAPIR e ABERCROMBIE, entre outros.

Quero dizer com isso, que embora a abordagem tradicional em voz na Fonoaudiologia, pareça estar respondendo aos anseios da prática clínica, além de ter seu espaço consagrado, ela ainda, se revela insuficiente para a compreensão da voz *normal* ou para uma conceituação da *“qualidade de voz ideal”* - como discute HOLLIEN (2000), que, inclusive, critica os fonoaudiólogos

pelo “*silêncio*” e pelas listas dos “*should nots*” - e, principalmente, insuficiente para a compreensão do fenômeno vocal em uma perspectiva social.

Gostaria de citar um estudo que, embora não tenha sido realizado em nosso país, traz importante contribuição para o delineamento das pesquisas científicas das últimas décadas. Vale esclarecer que o foco desse estudo era voltado para um aspecto específico, qual seja, a avaliação perceptual da qualidade de voz. Refiro-me ao artigo: “*Perceptual evaluation of voice quality: review, tutorial and a framework for future research*”, de autoria de KREIMAN, J. et al. (1993), cujos resultados são significativos para a caracterização da voz na Fonoaudiologia.

Os autores analisaram 57 artigos sobre o assunto, selecionados casualmente entre publicações de 1951 a 1990, e concluíram que a literatura não é suficientemente clara na identificação de diferentes medidas, escalas, testes e procedimentos confiáveis para uma avaliação vocal.

Além de analisar tais pesquisas, os autores fizeram dois experimentos controlados sobre uma única “qualidade vocal” (rouquidão) e, a partir dos resultados encontrados, sugeriram que os ouvintes desenvolvem “*padrões internos*” para as diferentes qualidades vocais de acordo com suas próprias experiências. Embora, os resultados apontem para uma ampla variabilidade intra e intersujeitos nos estudos analisados, essa variabilidade não se mostrou consistente em relação aos aspectos metodológicos pesquisados.

Diante desses dados, os autores levantam dúvidas sobre a confiabilidade e a consistência dos protocolos de avaliação vocal tradicionais, e apontam para a necessidade de um modelo teórico que supere essas dificuldades. KREIMAN et

al. fazem algumas proposições teóricas relacionadas ao controle das fontes de variabilidade, cujas discussões fogem ao escopo deste trabalho.

A importância deste estudo está em mostrar o estado em que se encontram as pesquisas e conseqüentemente, a prática em voz. Os resultados refletem as dificuldades conceituais em vários aspectos fundamentais, sejam metodológicos como mostram estes autores, sejam teóricos e ou práticos, como indicam o trabalho de outros autores, como FERREIRA (1993 e 2000), MARTZ (1990 e 1999), SERVILHA (2000) e HOLLIEN (2000).

O grande avanço no campo da voz nos últimos tempos, vem subsidiar um conhecimento mais aguçado e aprofundado quanto à fisiologia da voz normal e, principalmente, patológica. Entretanto, para que a materialidade fônica possa ser melhor compreendida como um processo que vá além da mera vibração das pregas vocais, há que se percorrer, ainda, um longo caminho, que possibilite transcender fronteiras disciplinares e conceituais já estabelecidas pela história da Fonoaudiologia no Brasil e em outros países.

2. ANÁLISE FONÉTICA DA VOZ

“... our social interaction through speech depends on much more than solely the linguistic nature of the spoken messages exchanged. The voice is the very emblem of the speaker, indelibly woven into the fabric of speech.”

(John LAVER, 1994)

LAVER inicia sua obra *“Principles of Phonetics”* (1994), com o esclarecimento acima, que mostra a importância do estudo da voz. O autor acrescenta que cada enunciado de fala carrega, não apenas, a mensagem em si, como também, por meio da entoação, do *“tom de voz”* e da qualidade de voz habitual do falante, uma *“declaração audível”* do grupo social, da característica física, da personalidade e da atitude de quem fala. Estas considerações introduzem os aspectos que abordarei neste item, ou seja, as diferentes dimensões da voz e as noções fundamentais para a análise do fenômeno vocal.

Os pontos mencionados por LAVER estão relacionados à dimensão funcional da voz, isto é, aos marcadores fonéticos e lingüísticos da fala, dos quais, me interessam, principalmente, os aspectos paralingüísticos. Eles não podem ser considerados dissociados da dimensão da produção fono-articulatória, ou seja, nem da qualidade e da dinâmica de voz, nem da dimensão física e acústica. Desta última dimensão, me restringirei aos parâmetros de análise fonética da voz utilizados nesta investigação.

Iniciarei pela dimensão funcional *da voz*, descrevendo, de forma breve, conceitos importantes para a semiótica vocal, quais sejam, os **marcadores fonéticos e lingüísticos** da fala. Eles correspondem aos aspectos relacionados

à *performance* do ouvinte em atribuir características particulares a um falante por meio desses próprios marcadores (LAVER & TRUDGILL, 1979). Tal revisão conduz à discussão dos elementos extralingüísticos e, em especial, dos **paralingüísticos** expressos na voz do falante.

Essa discussão, por sua vez, implica em uma focalização mais estreita dos conceitos de **qualidade e dinâmica de voz**, os quais tenho mencionado diversas vezes no decorrer do trabalho, dada sua importância para o tipo de estudo em questão e, portanto, analiso-os mais detalhadamente.

Por fim, como os aspectos paralingüísticos são analisáveis por meio da dimensão físico-acústica, esclareço os componentes físicos da voz e seus correlatos acústicos. Dentre os parâmetros fonéticos-acústicos mais estudados estão a frequência fundamental (Fo), correlato acústico do *pitch*, e a duração.

Outros parâmetros fonéticos também mereceram atenção, pois, como menciona BOLINGER (1985, p.15), no fluxo da fala há um contínuo mais ou menos estável, que é interrompido e percebido como proeminente. Essas proeminências são a realização das intenções do falante de enfatizar uma palavra ou qualquer outra unidade lingüística. Dentro do fluxo contínuo de fala, o fenômeno da interrupção é freqüente e resultante do modo de inscrição da linguagem falada no eixo temporal. As interrupções (com ou sem retomadas) sinalizam as estratégias do texto falado e, como afirmam SOUZA e SILVA & CRESCITELLI (1998, p.253):

“Como fenômeno sinalizador, elas apontam, ainda, para o caráter reflexivo da linguagem, isto é, para a possibilidade que a linguagem tem de poder olhar para si mesma, de se voltar sobre

aquilo que acabou de ser dito ou de antecipar aquilo que ainda vai ser proferido.”

Depreende-se do texto das autoras²⁵, a importância que este fenômeno (interrupção) representa para a investigação da análise da estrutura do discurso. Contudo, como o foco de atenção desta pesquisa é a materialidade fônica, não me deterei no estudo dos tipos de interrupção, mas somente, na verificação da ocorrência das pausas, através das quais, ela é introduzida.

Outros aspectos serão objeto de estudo, pois, várias são as maneiras de se enfatizar uma unidade no fluxo da fala por meio de mudanças na qualidade de voz, expressas em variações do *pitch*, da velocidade de fala ou, ainda, da duração da elocução. Manifestações estas, como assinala GOFFMAN (1974), que representam o *posicionamento* do falante na situação interacional. São, portanto, noções linguisticamente relevantes e fundamentais para a compreensão do fenômeno vocal na situação face a face e serão discutidas a seguir.

2.1. A Dimensão Funcional da Voz: os Aspectos Paralingüísticos da Fala e os Marcadores Fonéticos e Lingüísticos

A descrição dos aspectos paralingüísticos implica em rever algumas questões da Semiótica, pois, é uma análise que envolve não só o estudo dos sistemas de signos utilizados na comunicação, como também, os mecanismos pelos quais os signos são produzidos pelo *falante* e percebidos pelo *ouvinte*

²⁵ SOUZA e SILVA e CRESCITELLI realizaram diversos estudos sobre esse tópico, constituindo referência importante para aqueles interessados no tema.

(LAVÉ, 1994, grifos do autor). Compreender esse processo será bastante útil para a investigação da materialidade fônica.

A fala constitui-se no meio fônico da linguagem, e é priorizada pela humanidade por ser um meio econômico - com poucas unidades fônicas pode-se expressar um número infinito de idéias. Além disso, sua realização ocorre em três planos complementares e de naturezas diferentes. – acústico, perceptual e de produção.

Para a Semiótica é relevante a maneira como a produção e o uso dos signos podem carregar e transmitir informações sobre as características do falante. O signo em sua própria constituição traz características do produtor e por isso causa impressão no ouvinte.

Como personagens de um mundo social, um sujeito interage com o outro de diferentes formas, entre as quais, pela oralidade, que por sua vez, tem sido um dos aspectos mais estudados por pesquisadores das mais variadas áreas. Conhecer e desvendar seus mistérios tem fascinado inúmeros estudiosos ao longo dos tempos. O estudo dos aspectos lingüísticos, entretanto, não a abrange em todas as suas possibilidades, uma vez que, além dos lingüísticos, ela envolve também aspectos paralingüísticos e extralingüísticos.

Na interação verbal, além de compreender o significado das palavras ditas, pode-se também abstrair impressões e características de quem fala a partir das informações **paralingüísticas** expressas por meio da gestualidade, da expressão corporal, do contato ocular e da **voz** em si. Como afirma ABERCROMBIE (1967), a voz transmite informações indexicais que carregam dados do próprio falante, através das quais o ouvinte pode delinear inferências

sobre a identidade do seu interlocutor. Evidenciam, portanto, aspectos biológicos, psicológicos e sociais do falante.

Há um contínuo fluxo de informações entre os participantes de qualquer situação de conversação, expresso e percebido por meio de **marcadores fonéticos e lingüísticos de fala**, que revelam características do falante. Entre esses marcadores, os aspectos fônicos, que são passíveis de análise perceptual e acústica.

ABERCROMBIE (1967), que estuda os aspectos paralingüísticos relacionados à linguagem, assinala que é tão grande a importância do processo auditivo nas situações dialógicas, que ele sugere a inclusão do ouvido como parte dos órgãos fonadores, apesar de o ouvido não fazer parte da produção vocal em si. Ressalta, porém, que os sons da fala não são ouvidos como qualquer outro som, porque, na oralidade, o som carrega um significado lingüístico.

Geralmente, não se dá muita atenção à voz, seja quando se fala, seja quando se escuta alguém. Produz-se e escuta-se não só os próprios sons da linguagem, como também a voz de quem fala, embora como ouvintes, e também como falantes, nem sempre se tenha consciência da última. Durante a produção vocal, comumente não se está cômico de como a voz é produzida e não é também, habitual monitorá-la para controlar sua produção.

Normalmente, a atenção, tanto do falante quanto de seu interlocutor está voltada **ao** que é dito e não ao **como** é dito, ignorando-se, assim, o processo de produção vocal e seus efeitos na interação. É importante, portanto, que se analise como os usos da voz e o contexto em que ela é produzida interferem e se manifestam na produção vocal, expressando as informações relativas aos “donos”

da voz, ou seja, **como a interação transforma a voz** e quais as variações vocais decorrentes dos diferentes contextos de produção vocal.

As características vocais do falante, portanto, podem ser percebidas por meio dos chamados marcadores fonéticos e lingüísticos, terminologia introduzida por LAVER & TRUDGILL (1979), os quais tomaram como referência os estudos de ABERCROMBIE (1967). O termo marcador deriva do conceito de *índice* ("*index*"), introduzido por ABERCROMBIE (1967) no estudo semiótico da fala. Este autor estabeleceu este conceito, a partir do termo *indicial* ("*indexical*"), analisado por Charles Peirce, filósofo pragmático americano do século XIX, cujos estudos tiveram grande impacto na Semiótica. ABERCROMBIE define três grupos de índices:

- a) *aqueles que marcam as características físicas do indivíduo como sexo, idade e condições físicas e de saúde do falante;*
- b) *aqueles que indicam o grupo social a que o sujeito pertence, como o status social, o nível educacional, a ocupação e o papel social e*
- c) *aqueles que revelam características da personalidade e da atitude do falante, como o estado de humor e emocional do sujeito.*

LAVER & TRUDGILL (1979), posteriormente, e seguindo a tipologia proposta por ABERCROMBIE, descreveram três tipos de marcadores. Respectivamente, os **físicos**, os **sociais** e os **psicológicos**.

Os marcadores das **características físicas** do falante são percebidos na **qualidade de voz** do sujeito e estão relacionados aos aspectos anatômicos e aos ajustes habituais do trato vocal (LAYER &TRUDGILL, 1979).

Os marcadores **sociais** são percebidos por meio do sotaque característico do sujeito e pela escolha do vocabulário utilizado. São marcadores, em parte, quase-permanentes, marcando o grupo social a que pertencem o sujeito, e, em parte, numa perspectiva temporal, teriam uma variação circunstancial de média e curta duração, de acordo com os ajustes vocais temporários realizados pelo sujeito.

Os marcadores **psicológicos** revelam-se, freqüentemente, no "*tom de voz*" do sujeito, conforme indicado por ABERCROMBIE (1967 e 1968) e teriam uma perspectiva temporal de média duração, alcançada por meio de ajustes temporários do trato vocal. Segundo LAYER & TRUDGILL (1979), os marcadores sociais e psicológicos podem ser controlados pelo sujeito. Auditivamente, todo segmento de fala é composto de sons descritíveis, em termos de qualidade, *pitch*, *loudness* e duração, que possibilitam a percepção e a análise dos aspectos extralingüísticos e paralingüísticos.

Os elementos extralingüísticos transmitem informação evidencial acerca da identidade do falante, como o sexo e a idade, e correspondem, portanto, aos denominados marcadores físicos. Estão, particularmente, relacionados à qualidade vocal do sujeito e são expressos na faixa de extensão habitual de *pitch* e *loudness* do falante. Como esses aspectos estão associados às características físicas individuais, não são, geralmente, manipulados nem controlados pelo sujeito. Estão diretamente relacionados ao comprimento do trato vocal e às

características anatômicas dos lábios, da língua, dos dentes, dos maxilares, da cavidade nasal, da faringe e das estruturas laringeas.

Por conseguinte, as condições anatômicas do sujeito determinam características vocais pessoais e permanentes (ABERCROMBIE, 1967). Naturalmente, servem para caracterizar apenas o indivíduo e não o grupo social a que pertence. Assim, por exemplo, conforme uma pessoa envelhece, suas mudanças físicas (orgânicas) são acompanhadas por mudanças na qualidade de voz.

Os elementos extralingüísticos são produtos da maneira como o falante **habitualmente** ajusta seu trato vocal e, precisamente, por serem baseados em aspectos que, em essência, são invariáveis e involuntários da *performance* vocal do falante, possibilitam, ao interlocutor, um julgamento provavelmente mais apurado dos atributos físicos do falante.

No entanto, esse ajuste **habitual** do trato vocal gera efeitos vocais *quase-permanentes* (ABERCROMBIE, 1967, grifo meu), pois, a qualidade vocal pode, também, sofrer mudanças circunstanciais (nem sempre conscientes por parte do falante). Pode ocorrer, por exemplo, uma variação em função do grau de tensão dos músculos envolvidos na produção vocal, a qual interfere no modo de vibração das pregas vocais e gera diferentes resultados na **qualidade de voz característica** da pessoa.

O processo de fonação pode se modificar em decorrência de outros fatores, como o grau de constrição da laringe, e o posicionamento dos articuladores, entre outras possibilidades, cujos resultados podem ser

relativamente permanentes se, por exemplo, se esses ajustes estiverem cristalizados no sujeito, sempre que ele fala.

O falante é capaz de realizar mudanças no trato vocal que provocam variações temporárias, de maior duração ("*medium-term*") do que as citadas anteriormente e, que resultam em mudanças na qualidade de voz. Elas constituem os componentes paralingüísticos da fala, parcela **individual** e **variável** da voz, elementos não verbais, que expressam o estado emocional do sujeito e que teriam, também, a função de regular os turnos da conversação em curso (ABERCROMBIE, 1967, LAVER & TRUDGILL, 1979 e LAVER, 1994).

O adjetivo **paralingüístico** tem sua origem no termo *paralinguagem* ("*paralanguage*"), sugerido pelo lingüista americano Archibald Hill. Esta última expressão, contudo, foi questionada por ABERCROMBIE (1968), segundo o qual, sua utilização poderia levar a equívocos quanto à sua compreensão, pois se há uma "*entidade*" chamada *linguagem*, poderia haver outra, homogênea, denominada "*paralinguagem*". Por conseguinte, ele afirma que **paralingüístico** seria um termo mais inócuo e com menos chance de provocar mal entendidos. Propõe, assim, a utilização de "*fenômenos paralingüísticos*".

ABERCROMBIE indica a relevância do estudo paralingüístico como parte dos estudos da conversação, assinalando que, se os fenômenos paralingüísticos são elementos da conversação que interagem, ocorrem simultaneamente à fala e são co-produtores de um sistema total de comunicação, a linguagem falada não poderia ser apropriadamente compreendida, a menos que esses elementos fossem considerados.

Para este lingüista, as atividades paralingüísticas têm função comunicativa e constituem parte da interação conversacional. Tais requisitos permitem diferenciá-las de outros tipos de atividades como tiques nervosos, por exemplo. Para ABERCROMBIE (1968), como os elementos paralingüísticos, pelo menos potencialmente, deveriam ser conscientemente passíveis de controle pelo sujeito, uma *rouquidão* de origem orgânica, embora possa transmitir informação sobre o falante, não pode ser considerada como um elemento paralingüístico na situação de conversação. Neste caso, trata-se de uma informação extralingüística e relaciona-se às condições gerais de saúde do falante.

Segundo ABERCROMBIE (1968), os componentes paralingüísticos auditivos são caracterizados pelo fato de não seguirem as regras fonéticas e fonológicas formais da língua e são conjuntamente, abarcados sob uma terminologia popular - "*tons de voz*". São elementos produzidos por variações da norma lingüística social e expressos pela ***dinâmica de voz***, ou seja, pelas variações de *loudness*, *tempo* (duração) e *tessitura* (faixa de extensão de *pitch*), entre outros.

Até esta parte, tenho afirmado que este trabalho se insere na categoria de análise de ***estudos paralingüísticos***, em particular, dos aspectos expressos na ***qualidade e dinâmica*** de voz; recorro, contudo, a LAYER (1994) para fazer uma importante ressalva. O autor assinala que não se deve tirar conclusões fechadas como, por exemplo, atribuir as características sociais do falante somente às origens lingüísticas e paralingüísticas ou igualmente, considerar os atributos físicos do sujeito, somente a partir da informação extralingüística.

A atribuição das características sociais, psicológicas e físicas não devem ser, direta e simplesmente, correlacionadas às informações lingüísticas, paralingüísticas e extralingüísticas, respectivamente. Para atribuir qualquer característica a um determinado falante, o ouvinte deve considerar as evidências decorrentes dos três tipos de informação indiciais, de forma a não se incorrer em equívocos, como já indicava SAPIR (1927) na década de 20.

Considerando-se esta ressalva, situo esta pesquisa na esfera **paralingüística**, pois, o que me importa são as variações de voz decorrentes da situação de interação o que, possivelmente, leva o falante a realizar ajustes temporários (alguns sob seu controle, outros não) no seu trato vocal, resultando, por exemplo, em mudanças no *pitch* ou na duração do segmento fônico, que em última instância alteram a **qualidade de voz habitual** do sujeito em função da **dinâmica de voz**.

Isto não significa que os outros marcadores fonéticos e lingüísticos sejam menos evidentes ou importantes. Opto, contudo, por centralizar minha atenção nos aspectos paralingüísticos frente à complexidade de uma investigação que envolvesse os três tipos de marcadores na perspectiva teórico-metodológica adotada.

2.2. A Dimensão da Produção Fono-articulatória: a Qualidade e a Dinâmica Vocal

“Qualidade de voz é o principal veículo de informação sobre as características físicas, psicológicas e sociais do

falante e tem um papel semiótico vital na interpretação falada.”

(John LAVER, 1980, traduzido por
MADUREIRA, 1992)

Embora, os termos **qualidade** e **dinâmica de voz** tenham sido mencionados várias vezes, até esta parte do trabalho, gostaria de discuti-los de forma um pouco mais detalhada. Esta retomada é fundamental, pois são conceitos de crucial importância, não apenas para a compreensão do fenômeno vocal na interação face a face, especialmente, em uma perspectiva social de análise, mas também, para a compreensão e a investigação da voz “normal”.

Além desses aspectos, há que se considerar a necessidade de se transpor as fronteiras conceituais consagradas pela prática clínica da Fonoaudiologia e da Laringologia, que levaram a cristalização do conceito de **qualidade vocal** fundamentado em uma visão organicista da produção da voz, ou seja, em uma dimensão biológica, centrada na patologia (doença).

Dinâmica de voz é uma expressão pouco empregada na literatura especializada, das áreas de Fonoaudiologia e de Laringologia, principalmente da forma como é concebida por ABERCROMBIE (1967), e será explicitada no decorrer deste item.

Alguns autores, como BEHLAU & PONTES (1995), utilizam o termo “*Psicodinâmica Vocal*” para relacionar a voz aos aspectos psicoemocionais do falante. Eles consideram que a *psicodinâmica vocal* é um procedimento básico que possibilita levantar hipóteses sobre atributos “*simples*” do falante como sexo, idade e procedência, até outros, como tipo de estrutura física, expressões faciais

e “...até mesmo, a cor dos cabelos do interlocutor.”(BEHLAU & PONTES, 1995, p. 127).

Esclarecem que essa percepção auditiva é uma habilidade natural do ser humano que pode e deve ser desenvolvida como instrumento de análise na **terapêutica vocal**, pois, cada parâmetro avaliado (respiração, altura, ressonância, etc.) pode fornecer informações sobre a psicodinâmica vocal do **paciente** e sobre a repercussão de sua voz nos ouvintes, dados importantes para a **terapia** (grifos meus).

BEHLAU & PONTES acrescentam que é muito difícil traçar diretrizes para uma avaliação da psicodinâmica vocal, dado o caráter individual do uso da voz e a modificação que o falante realiza em função da situação e do contexto de comunicação. Afirma, entretanto, que apesar dessa limitação, existem algumas relações básicas que podem ser traçadas entre os ajustes motores, certas características da personalidade do falante e os efeitos causados no ouvinte. Baseados nesses aspectos apresentam uma tipologia de voz, em que relacionam o tipo de voz e seus efeitos sobre os ouvintes. Uma classificação composta pelos 23 tipos de vozes, encontrados com maior frequência na **clínica** (grifo meu).

Destaco que a utilização da expressão psicodinâmica vocal volta-se, fundamentalmente, à **prática clínica** (vide grifos anteriores) e, portanto, seu emprego, embora, útil para a terapêutica em voz, é muito restrito para investigação da voz de falantes que não apresentem alterações vocais ou as chamadas disfonias .

Diante dessas considerações, torna-se clara a necessidade de recorrer a outros referenciais teóricos, e portanto, retomo ABERCROMBIE (1967). Ainda

que, como menciona o autor, pelo menos, a metade do corpo humano, da cabeça ao abdômen, esteja diretamente relacionada com a produção vocal, envolvendo basicamente três sistemas: o *respiratório* (pulmões, traquéia, e outros órgãos do sistema respiratório), o *fonatório* (laringe) e o *articulatório* (nariz, lábios, língua, dentes e demais articuladores e ressoadores), é interessante ser observado que nenhum dos órgãos envolvidos tem a **função específica de produzir a voz**, mas sim, a **de sobrevivência**, ou seja, de respiração e alimentação.

A voz é, portanto, resultante de um esforço histórico de adaptação dessas funções primárias e vitais do ser humano, decorrente da necessidade do ser humano em se comunicar. Há, portanto, na produção da voz, um caráter eminentemente social que não pode ser desconsiderado.

Como uma pessoa reage ao ouvir uma voz e o que esta voz carrega e transmite são questões importantes para o estudo da materialidade fônica, o que, entretanto, não tem recebido atenção suficiente por parte dos pesquisadores. Quando alguém fala, causa sempre uma certa impressão no outro e, certamente, sua voz contribui para essa impressão. Entretanto, embora a voz seja um importante aspecto da interação, geralmente não se tem consciência do papel que ela ocupa nas relações humanas e, principalmente, de como ocorre o processo de produção da voz na situação interacional.

Tal fato é decorrente, de um lado, do tipo de prática clínica consagrada em voz e, por conseguinte, da maior valorização e desenvolvimento da dimensão biológica. Por outro lado, embora, nos estudos lingüísticos, muita atenção tenha sido dada aos aspectos segmentais da fala, através do estudo de sílabas, vogais

e consoantes, outros componentes não-segmentais têm merecido uma atenção menor.

ABERCROMBIE dividiu esses componentes em dois grupos, um concernente à **qualidade vocal** (contribui para a qualidade geral da voz) e o outro resultante de como a voz é produzida, isto é, a **dinâmica vocal**. O autor assinala, entretanto, que os três grupos - aspectos segmentais, características da qualidade de voz e características da dinâmica de voz ocorrem simultaneamente e estão continuamente presentes e juntos.

A análise desses aspectos segmentais e não-segmentais são essenciais para o estudo do funcionamento da língua, pois são de tal forma idiossincráticos e característicos do indivíduo, que a identificação do falante através da voz está na dependência deles. Os aspectos de qualidade e dinâmica de voz, definidos por ABERCROMBIE, vêm despertando maior interesse dos lingüistas e, em especial, dos foneticistas e de outros estudiosos preocupados com a linguagem.

Qualidade vocal, como afirma ABERCROMBIE, embora, seja um termo tradicional, é possivelmente, muito mal compreendido. O autor questiona a “*bem definida conceituação técnica*” dessa expressão no campo fonético, como, por exemplo, “*som resultante da fonação*” ou como “*vibração das cordas vocais*”, e acrescenta que a abrangência e o significado desse termo vai muito além dessa definição.

O autor assinala que **qualidade de voz** refere-se às características presentes quase o tempo todo em que uma pessoa fala. Nas palavras de ABERCROMBIE (1967, p. 91) “... *it is a quasi-permanent quality running through all the sounds that issues from his mouth.*”.

Como afirma o autor, essas características, naturalmente, incluem aquelas que têm sua origem nos aspectos anatômicos da laringe e, portanto, pertencem ao domínio da fonação. Incluem, contudo, também outras características que têm origem em algum *“outro lugar”* (*“elsewhere”*), que merece ser melhor investigado.

Alguns aspectos da qualidade de voz estão sob controle do falante, outros não. Alguns dos componentes fora do controle do sujeito são efeito de condições inatas, ou seja, das características físicas da pessoa, como por exemplo, a estrutura óssea do crânio e da região peitoral, o comprimento do trato vocal - da laringe aos lábios - , o tamanho das pregas vocais, a morfologia dos articuladores, e outras. Todas elas serão determinantes na qualidade de voz do falante. São evidentes, por exemplo, as diferenças entre as vozes de um homem e de uma mulher, e também, as mudanças numa voz infantil, conforme a criança vai crescendo, uma vez que as transformações em seus aspectos físicos são acompanhadas por mudanças na qualidade vocal. São aspectos que naturalmente permitem caracterizar um indivíduo e podem ser indiciais de grupos que apresentem as mesmas características físicas como, por exemplo, um grupo de mulheres.

Outros componentes que também não estão sob controle do sujeito e que são, às vezes, temporários, são aqueles decorrentes, por exemplo, de uma faringite, de uma laringite ou de um simples resfriado.

Os demais componentes da qualidade vocal estão sob controle do sujeito e não derivam de suas características físicas. São decorrentes dos ajustes musculares mantidos habitualmente pelo sujeito durante o tempo da sua

produção vocal, mantendo os órgãos fonadores ajustados de forma diferente à posição de repouso. Estes ajustes configuram o trato vocal de tal forma que, inevitavelmente, afetam a qualidade da voz de quem a produz. Por isso, ABERCROMBIE considerou que além do caráter **permanente**, a qualidade vocal apresenta também aspectos de caráter “**quase permanentes**”, em uma perspectiva temporal. Ela pode ser modificada circunstancialmente, sofrendo alterações temporárias. O falante, embora nem sempre conscientemente, realiza ajustes no seu trato vocal que se podem refletir , temporariamente, na sua qualidade de voz.

Essas mudanças derivam da **dinâmica de voz** utilizada pelo falante, expressão adotada por ABERCROMBIE (1967), em referência às mudanças das características vocais que são **controláveis** pelo sujeito. A influência dos aspectos anatômicos na dinâmica vocal decorre de fatores como a dimensão e a massa das pregas vocais e o volume respiratório, que exercem influência nas faixas de extensão de *pitch* (correlato psicoacústico da frequência) e *loudness* (correlato psicoacústico da intensidade). Tais condições delimitariam as faixas de variação de extensão da frequência fundamental e da amplitude do som que o falante poderia produzir (LAVER & TRUDGILL, 1979).

ABERCROMBIE engloba como dinâmica vocal os seguintes aspectos: *loudness* (intensidade), “*tempo*” (velocidade de fala), “*continuity*”, (continuidade), ritmo, tessitura, variação de *pitch* e registro. Essas são características que estão sob controle do falante e podem caracterizar tanto grupos sociais como um único falante.

O autor menciona que *loudness* é um aspecto de fácil controle pelo falante, uma vez que a faixa de variação da voz humana, nesse aspecto, é muito grande, e os ajustes ocorrem de acordo com a situação. Segundo este importante lingüista, trata-se, porém, de um aspecto de pouco significado lingüístico, pois um falante com uma qualidade vocal “*pobre*” não reproduziria os contrastes de *loudness* com fidelidade. Além disso, é uma característica que passa despercebida pela maioria das pessoas, quando se reproduz a fala de alguém.

Tempo significaria a velocidade da fala, que poderia ser melhor medida pela sucessão do número de sílabas. Pode variar de indivíduo para indivíduo, mas as pessoas tendem a seguir uma norma conforme o estilo conversacional usual. Assim como o *loudness*, pode caracterizar uma língua, mas também, seria de difícil demonstração.

“*Continuity*” estaria associado ao tempo, referindo-se à incidência de pausas num determinado trecho de fala em que aparecem, e a freqüência e duração com que ocorrem. Para ABERCROMBIE, apesar das hesitações ou das interrupções deliberadas para respirar, por exemplo, as pausas constituem uma matéria altamente idiossincrática com grande variação de um falante para outro. Passariam quase despercebidas numa situação normal de conversação.

O *ritmo* aparece claramente nos momento em que a fala é fluente e ininterrupta. Decorre de um movimento periódico, produzindo uma expectativa de que haverá uma regularidade em sua sucessão. ABERCROMBIE considera que há nele importantes implicações para a percepção, pois, o ritmo da fala é experienciado como um ritmo do movimento, que pode ser no tempo da sílaba como no francês, (“*syllable-timed rhythm*”), ou de acordo com o acento, como no

inglês (“*stress-timed rhythm*”). O ritmo é também “*foundation of verse*” e, por essa razão, o autor assinala a necessidade de se ouvir as línguas diferenciadamente, para que se possa analisar o ritmo de fala de uma língua, materna ou não, e descrevê-la foneticamente.

O componente seguinte – tessitura, foi convenientemente “tomado emprestado” da terminologia empregada na música, conforme explica ABERCROMBIE:

“... ‘Voice’, in its technical phonetic sense of sound resulting from phonation, is a musical tone which has a fundamental frequency, and therefore a recognizable pitch. The pitch of the voice is in continual fluctuation during speech, but the fluctuations tend to take place round a central point: if we disregard the occasional extremes, a speaker has a characteristic range of notes, or compass, within which the pitch fluctuation of his voice falls during normal circumstances. This range or compass is the tessitura.

(ABERCROMBIE, 1967, p.99)

A *tessitura* pode variar de pessoa a pessoa e guarda uma relação com a espessura, o tamanho e a condição das pregas vocais. Em algumas comunidades ela é uma característica institucionalizada, e cada cultura apresenta diversificações no uso da extensão vocal. A este respeito BEHLAU & PONTES (1995) afirmam que:

"A restrição ao uso potencial vocal é uma realidade de cada cultura e sociedade, e geralmente, chegamos à idade adulta com 'o que restou da nossa voz' (Behlau & Ziemer, 1986); a extensão potencial é determinada constitucionalmente, mas, o seu uso, a seleção de uma extensão mais ou menos ampla, depende de fatores ambientais, emocionais, educacionais e patológicos, que dizem respeito à história individual de cada ser."

(Ibidem, p. 90)

Esses autores lembram, ainda, que diferentes culturas selecionam distintamente suas tessituras, e citam como exemplo a proximidade entre a voz falada e cantada no italiano e ao contrário, o distanciamento que ocorre no japonês, que apresenta uma tessitura mais grave na voz falada do que na voz cantada.

Registro, como *tessitura*, é um termo “*emprestado*” da terminologia empregada para a voz cantada e é empregado de forma análoga ao canto, ou seja, ao modo de produzir os sons da tessitura. ABERCROMBIE assinala que há controvérsias quanto à nomenclatura utilizada e que, talvez, somente, os termos registro de “*cabeça*” e de “*peito*” não causem polêmica no meio musical²⁶.

O autor propõe o emprego do termo *registro* para descrever ajustes temporários que o falante realiza, e por isso sugere que o termo seja considerado como um dos aspectos da dinâmica vocal, isto é, ao modo como o falante varia a

²⁶ Na literatura especializada em Fonoaudiologia e Laringologia, BEHLAU & PONTES (1995) apontam o emprego de três registros vocais principais : *basal* (equivale às frequências mais graves da tessitura, varia de 10 a 70 Hz), *modal* (utilizado na fala habitual e subdividido em três categorias: peito, misto e de cabeça, varia de 80 a 560 Hz) e o *elevado* (corresponde às frequências mais agudas, varia de 160 a 800 Hz).

voz , conforme as circunstâncias. Cita, como exemplo, as mudanças de registro que podem ocorrer para sinalizar o estado emocional e a atitude do falante. Além disso, ABERCROMBIE menciona os contrastes de registro nas diferentes línguas.

O último aspecto apontado por ABERCROMBIE na dinâmica vocal é a *flutuação do pitch* presente nos atos de fala de qualquer comunidade, um aspecto lingüística e socialmente relevante. O autor esclarece que se trata de uma flutuação bem definida dos padrões melódicos característicos de cada língua. Para ele, a *flutuação do pitch* é, provavelmente, o elemento mais importante da dinâmica vocal.

ABERCROMBIE ilustra as diferenças melódicas e entoacionais das línguas, citando, como exemplo, o chinês, em que a variação tonal ocorre na palavra e o inglês, em que a variação de *pitch* ocorre na sentença.

Retomo alguns aspectos teóricos sobre **qualidade vocal** de forma a poder explicitar como adoto o conceito neste estudo.

O termo qualidade vocal, entretanto, não tem sido tratado dentro de um único conceito. Na verdade, tem sido empregado de diversas maneiras na literatura: como atributos específicos do som (OGDEN, 1924 & ARNOLD, 1967 apud CRYSTAL, 1969), como um único, identificável e convencional aspecto do som (LUCHINGER & ARNOLD, 1965, apud CRYSTAL, 1969); como um elemento prosódico (CARRELL & TIFFANY, 1960 apud CRYSTAL, 1969), ou, ainda, como *timbre*, cujo significado remete à impressão da voz ou “*tom de voz*”. Esta última forma constitui seu uso mais comum, nas diversas disciplinas que se ocupam

em estudar o fenômeno vocal, havendo dela referências mais antigas, como em SWEET (1878, citado por CRYSTAL, 1969).

A maneira como CRYSTAL concebe qualidade vocal, entretanto, coincide em diversos aspectos, com ABERCROMBIE (op.cit.), como se pode ver abaixo:

“... the phenomenon voice-quality or voice-set, i. e., that relatively permanent, non-institutionalized, idiosyncratic, background voice quality which accompanies a person when he speaks and is the main source of our ability to recognize personal identity vocally.”

CRYSTAL (1969, p.100)

Ambos, portanto, consideram qualidade vocal como relativamente permanente, idiossincrática, presente na fala da pessoa, e capaz de permitir o reconhecimento do indivíduo através da voz. Sobre esse tópico, CRYSTAL (1969, p.123) acrescenta, ainda, que:

“Voice quality is thus a single impression of a voice existing throughout the whole of a normal utterance (i.e. excluding deliberate simulation of a different quality, physiological change, or environmental distortion): it corresponds to a combination of independently varying acoustic and articulatory parameters, of which the most important attributes are pitch, loudness, duration and timbre.

Segundo ele, qualidade vocal não deveria ser equiparada à *timbre* (uma característica de qualquer som, referindo-se à ressonância inerente do som), pois o último teria uma aplicação mais restrita que a de qualidade vocal (que se refere, primariamente, às características individuais dos sons). Uma distinção, apontada por ele, potencialmente útil para os lingüistas, poderia ser o uso de *timbre* para referir-se às características inerentes dos sons segmentais, como por exemplo “*timbre vocálico*” e “*timbre fricativo*” e o uso de qualidade vocal ficaria restrito às características inerentes (“*background*”) da expressão verbal, por meio das quais, os contrastes lingüísticos poderiam ser percebidos.

Na atualidade, BEHLAU & PONTES (1995) assinalam que qualidade vocal é o termo utilizado em referência ao conjunto de características que identificam uma voz humana, e em substituição à *timbre* (vocal), cujo uso, segundo os autores, estaria restrito aos instrumentos musicais. Para eles, qualidade vocal relaciona-se à impressão total criada por uma voz variável segundo o contexto de fala e em função das condições físicas e psicológicas dos indivíduos, mas sempre com a manutenção de um padrão básico que possibilite a identificação do dono dessa voz. Acrescentam que esse padrão seria o “*índice mais completo dos atributos da emissão de um indivíduo, sendo capaz de fornecer desde informações sobre suas características físicas a dados de sua formação educacional.*” (1995, p.69)

Já em sua época, CRYSTAL apontava para a necessidade de uma maior compreensão da qualidade vocal para que fosse possível uma análise adequada das características prosódicas da língua, particularmente daquelas permanentes como *pitch* e *loudness*. Através do estudo independente dos parâmetros acústicos da fala e de seus correlatos perceptuais, ter-se-ia uma perspectiva das

características essenciais para a identificação e a definição dos aspectos prosódicos da fala.

Nesse sentido, o autor acredita que, para uma abordagem lingüística da descrição dos efeitos vocais relevantes, seria essencial haver um termo geral para as características permanentes da voz de um indivíduo, em contraposição às quais, os padrões lingüísticos convencionais seriam identificados. Aponta, assim, a relevância lingüística da qualidade vocal, considerando-a básica para qualquer estudo dos contrastes prosódicos de uma língua.

Essas afirmações remetem a LAVER (1968) para quem os conceitos disponíveis para uma descrição lingüística estavam bem desenvolvidos, mas, o mesmo não ocorria em relação a uma teoria descritiva da qualidade vocal. Embora este fato continue sendo realidade até os dias atuais, alguns dos autores citados, como SAPIR, 1927; ABERCROMBIE, 1967; CRYSTAL, 1969 e LAVER, 1968 e 1980 foram importantes para que esse quadro começasse a ser mudado.

As razões colocadas por LAVER para um modelo descritivo da voz foram: favorecer a discussão interdisciplinar da função indicial da qualidade de voz e incorporar esse modelo na teoria geral da fonética. A contribuição de LAVER para a análise da qualidade vocal é inquestionável. Assim como ABERCROMBIE, o autor critica os rótulos impressionistas empregados na classificação das vozes, considerando-os vagos do ponto de vista de significado, exceto no sentido metafórico ou como rótulos arbitrários de imitação. Sua tese principal é a de que a qualidade de voz pode ser descrita pelos seus componentes e que uma descrição em bases fonéticas pode suprir os conceitos necessários para uma descrição fisiológica e significativa de cada componente.

Diante dessas considerações, esclareço que emprego o termo **qualidade vocal** da maneira como proposta por LAYER. Para justificar meu posicionamento, utilizo o esclarecimento de MADUREIRA (1999, p.64) em seu artigo sobre “*Entoação e Síntese de Fala: Modelo e Parâmetros*”, citado a seguir:

“Para este autor, ‘qualidade de voz’ não se restringe apenas aos ajustes fonatórios como o vocábulo ‘voz’ pode induzir a pensar. Também, não se refere a avaliação qualitativa de alguma propriedade de voz. Para Laver, ‘qualidade de voz’ refere-se aos traços que se fazem presentes de forma intermitente na fala de um indivíduo e que derivam de ajustes fonatórios e articulatórios. A unidade proposta para a análise das qualidades de voz no modelo de Laver é o ‘setting’...”

Os aspectos discutidos sobre a análise fonética podem ser resumidos num quadro proposto por LAYER & TRUDGILL (1979, p.9, traduzido por mim), que mostra a relação entre as variáveis vocais e suas funções. Apresento uma versão adaptada:

QUADRO 1 – Relação entre as variáveis vocais e suas funções

Função	Informativa		Informativa e comunicativa	
Relação com a Linguagem	Características vocais extralingüísticas		Características Paralingüísticas	Realizações fonéticas das unidades lingüísticas
Perspectiva temporal	permanente	quase permanente	média duração	curta duração
Variáveis vocais	Características vocais derivadas das diferenças anatômicas entre os sujeitos, exercendo influência na <i>qualidade</i> e na <i>dinâmica</i> de voz	Ajustes vocais, por exemplo, ajustes musculares do aparato vocal, incluindo ajustes da <i>qualidade</i> e da <i>dinâmica</i> de voz	“Tom de voz” decorrente dos ajustes vocais temporários, incluindo ajustes paralingüísticos da <i>qualidade</i> e da <i>dinâmica</i> de voz	Realizações articulatórias, momentâneas, das unidades fonológicas, incluindo manipulações de curta duração dos aspectos fonéticos da <i>qualidade</i> e da <i>dinâmica</i> de voz
Marcadores	Físicos	sociais e psicológicos		
Controle potencial do falante	não controlável	Controlável		

2.3. A Dimensão Físico-acústica da Voz

“Voice quality is a concept that is at once widely recognized and very difficult to define in a way that is universally satisfactory. The term is used by many specialists for many different purposes. Voice quality can range from the ideal to the acceptable to the disordered, and numerous methods have been applied to its

measurement. Perceptual assessment of voice quality has a long history but also, we believe, a new horizon.”

(Raymond D. KENT & Martin J. BALL, 2000)

Assim como os autores, acredito que a avaliação da *qualidade de voz* apresenta uma longa história, contada, em parte nos itens anteriores. É uma história que apresenta um grande avanço em vários aspectos, controvérsias em vários outros e diferentes formas de abordagem e, também, um novo horizonte.

Este item tem por objetivo esclarecer os parâmetros físicos e acústicos utilizados para análise da qualidade e dinâmica vocal dos sujeitos que fizeram parte deste estudo. Cabe esclarecer que esta pesquisa não tem um caráter quantitativo nem experimental e, por isso, as medidas realizadas inserem-se numa perspectiva mais qualitativa, de forma a elucidar os referenciais analíticos adotados e contribuir para novos horizontes de investigação e compreensão da materialidade fônica no universo das pesquisas em voz numa perspectiva social.

Toda voz é única, pois, diversos fatores – anatômicos, fisiológicos, psicológicos, culturais, sociais e lingüísticos - contribuem para o produto final. Estes elementos não são estáticos, e a voz do falante varia dia a dia, palavra a palavra. As características acústicas da voz podem ser descritas perceptualmente, e analisadas instrumentalmente por meio de alguns parâmetros que refletem a estrutura e a função da laringe e do trato vocal infra e supra – glótico (KENT & READ, 1992). Reitero que refletem, também, os demais fatores, que exercem influência na qualidade de voz do falante, citados anteriormente.

Ao longo dos anos, a falta de consenso entre as pesquisas sobre análise perceptual da voz impulsionou o desenvolvimento de parâmetros considerados

como objetivos, justamente, por não serem considerados, em primeira instância como modalidades perceptivas ou subjetivas (CAMARGO, 2000). Como acrescenta a autora, a discussão desses parâmetros está longe de ser uma questão de fácil elucidação, pois, as várias possibilidades tecnológicas de estudo da função fonatória são passíveis de diferentes interpretações pelos avaliadores, o que aumenta o risco de subjetividade nas avaliações.

Além disso, KREIMAN & GERRAT (2000) assinalam que a qualidade de voz representa uma interação entre o sinal acústico da voz do falante e o ouvinte. O sinal acústico em si não possui uma qualidade, ela é evocada pelo ouvinte. Por esta razão, os autores consideram que as medidas acústicas somente apresentam um significado quando correspondem ao que os ouvintes escutam como eles indicavam em artigos anteriores (GERRAT & KREIMAN, 1995 e KREIMAN & GERRAT, 1996, apud KREIMAN & GERRAT, 2000).

Dessa forma, a aplicação dos métodos de avaliação da voz, seja perceptual ou acústico, requer um conhecimento não só da técnica e dos parâmetros em si, mas também, das peculiaridades das condições de produção vocal do falante. Os meios de avaliação subjetivo (perceptual) e objetivo (acústico) contribuem para que se desvende os ajustes vocais realizados pelo falante e suas variações, conforme o contexto; sua análise, porém, está sujeita à variabilidade de interpretações de quem avalia.

Os métodos de avaliação objetiva da voz vem sendo descritos ao longo das últimas décadas por vários autores. Um referência básica dos princípios acústicos de análise da produção vocal é o autor Gunnar FANT, que descreveu o

modelo *fonte-filtro para a produção das vogais* em “*Acoustic Theory of Speech Production*”, em 1960 (KENT & READ, 1992).

Essa teoria é útil para a compreensão da relação entre os aspectos articulatórios e acústicos e fornece os fundamentos para os procedimentos da análise acústica da fala. O modelo de FANT fundamenta-se em estudos que consideram o trato vocal como um tubo uniforme aberto de um lado, que corresponde à abertura da boca e fechado do outro, que representa uma membrana que simula as pregas vocais. O comprimento médio do trato vocal masculino é de 17,5 cm, da glote aos lábios e teria aproximadamente as mesmas frequências de ressonância que um tubo do mesmo comprimento e diâmetro. Assim, este aparato representa um modelo simples da produção de uma vogal pelo homem. Uma explicação dessa teoria também pode ser encontrada em KENT & READ (1992).

Dentre os eventos fonéticos-acústicos mais estudados estão a frequência fundamental, que é, convenientemente utilizada, em sua forma abreviada como F_0 e a duração (MADUREIRA, 1999). A frequência fundamental é medida em Hertz e a duração em milissegundos.

A **frequência fundamental** é um parâmetro acústico que designa o número de repetições de ciclos de uma onda vibratória, cujo correlato fisiológico corresponde ao número de vibrações das pregas vocais (abertura e fechamento). O correlato perceptual é o *pitch*.

Os sons com estrutura harmônica são sons periódicos, o que significa que o mesmo padrão de vibração se repete num intervalo fixo, e desta forma, podem ser medidos. Vários autores situam a média de F_0 para falantes masculinos em

torno de 120 Hz e para falantes femininos, em torno de 225 Hz (PITTAM, 1994), o que significa que as pregas vocais estão vibrando 120 e 225 vezes por segundo, respectivamente. Durante uma conversação normal, a faixa de extensão de *pitch* pode variar de 50 a 250 Hz para os homens e de 120 a 480 Hz para as mulheres (LAVER, 1994).

O termo *pitch habitual* (“*habitual pitch*”) é freqüentemente empregado para indicar a freqüência fundamental de fala (“*speaking fundamental frequency*” – *SFF*) habitual de um indivíduo. COLEMAN & MARKHAM (1991) afirmam que, relativamente, pouca atenção tem sido dada a variação de *SFF*, ou seja, ao *pitch* habitual do falante. Os autores realizaram uma pesquisa com falantes normais divididos em três grupos. O primeiro foi constituído por 11 estudantes do sexo feminino, cujas vozes foram gravadas durante um mês a cada três dias. O segundo grupo reunia indivíduos que, presumidamente, possuíam um padrão vocal modelo, por serem apresentadores de emissoras de televisão. Foram gravadas as vozes de 2 sujeitos do sexo feminino e 4 do sexo masculino, no mínimo 6 vezes, num período de 2 semanas. O terceiro grupo foi formado por sujeitos, cujas vozes foram acompanhadas por vários anos – um pastor de uma emissora de televisão local, gravado semanalmente por um período de 7 anos e uma contadora de histórias que foi acompanhada por 10 anos. Os resultados foram padronizados em semitons para uma comparação da variação de *SFF* nos três grupos. Os resultados sugerem que uma variação de 18% na freqüência fundamental de fala de um dia para outro pode ser considerada dentro dos limites normais. Os autores, contudo, alertam que esse dado não pode ser generalizado para toda a população.

COLEMAN & MARKHAM consideram que seu estudo teriam, pelo menos duas aplicações **clínicas** (grifo meu). A primeira é que a medida do *pitch* habitual em adultos pode ser considerada estável, pois, a variação encontrada foi pequena independentemente do período estudado, um mês, um ano ou uma década. Esta faixa de variação pode ser atribuída ao estado emocional e/ou de saúde do sujeito. A segunda aplicação relaciona-se ao processo terapêutico da voz de um sujeito. Os autores alertam que o terapeuta deve ser bastante criterioso em relação a mudar ou não o *pitch* habitual do paciente. Sugerem que esta é uma decisão que deve ser tomada levando-se em consideração o real desejo do sujeito de mudar um atributo vocal que lhe pertence, e deve-lhe ser esclarecido que a mudança pode afetar em alguns aspectos o seu comportamento. O conceito de o que é uma voz normal e a tarefa de manipular o *pitch* do falante devem ser, portanto, cuidadosamente analisados.

Um fato interessante a assinalar em relação a esse estudo é que, embora, os autores tenham analisado a variação normal do *pitch*, o interesse deles, na verdade, estava na sua aplicação clínica. Entretanto, ainda que não fosse objetivo da pesquisa, seus resultados trazem importante contribuição, não apenas para a clínica, mas também para os estudos da voz normal.

PITTAM (1994) indica outras medidas que podem ser realizadas baseadas na frequência fundamental, quais sejam, a faixa de extensão de F_0 , a variabilidade de F_0 e a média de F_0 em um enunciado de fala. Neste estudo serão medidas as variações tonais utilizadas pelos sujeitos como meio de sinalizar as ênfases no seu discurso oral.

Como já foi assinalado antes, tais proeminências ocorrem quando o falante realiza suas intenções de enfatizar uma palavra, ou qualquer outra unidade lingüística e, assim, interrompem o fluxo contínuo fala. Como se trata de fala, a interrupção é audível. Este fenômeno da interrupção, como assinalam SOUZA e SILVA & CRESCITELLI (1998), é constitutivo da natureza oral e os enunciados interrompidos podem ser considerados como uma marca de elaboração da própria oralidade. A ênfase é conseguida, tornando-se uma sílaba na palavra ou uma palavra no grupo ou um grupo de palavras em um contexto maior e assim por diante, mais proeminente do que os outros (CRYSTAL, 1969).

Não se pode esquecer que proeminência envolve uma questão de figura e fundo, como menciona BOLINGER (1985). Assim, será proeminente uma palavra produzida em voz sussurrada se o “*fundo*” for realizado com um registro de voz modal, por exemplo. Ao contrário, será proeminente uma palavra produzida em registro modal sobre um “*fundo*” de voz sussurrada. Há, portanto, várias maneiras de enfatizar uma unidade no fluxo de fala.

Os padrões de variação da proeminência silábica resultam da interação de quatro elementos: *pitch*, *loudness*, duração e qualidade articulatória (LAVER, 1994). Do ponto de vista perceptual, somente o aspecto de *loudness* não será considerado na análise da qualidade e dinâmica de voz dos sujeitos estudados, pois, como aponta ABERCROMBIE (1967) é um parâmetro que não tem tanta significância como o *pitch*, além de ser de difícil aferição. Além disso, a intensidade é um parâmetro acústico sensível às condições do meio como a presença de ruído ambiental e o nível de absorção acústica do ambiente, o que dificulta a realização de sua medida.

A melodia de uma elocução é comunicada, primordialmente, pelo movimento do *pitch* da voz do falante numa perspectiva temporal. LAYER (1994) faz uma distinção entre *pitch range* (faixa de extensão de *pitch*) e *paralinguistic range*. A primeira expressão corresponde à faixa máxima de extensão de *pitch* que o falante é fisicamente capaz de produzir delimitado pelas suas condições biológicas. A segunda forma corresponde aos ajustes que um falante pode realizar, dentro dos limites de sua faixa de extensão de *pitch* e em função de seus propósitos paralingüísticos, para marcar uma atitude particular, ajustes estes, que também estão submetidos às distinções culturais das diferentes línguas. A outra forma descrita por LAYER é o *pitch span*. Para o autor é a mais importante do ponto de vista fonológico, pois, é o local da faixa de extensão em que o falante organiza os valores de *pitch* com finalidades prosódicas nos limites, de parte ou do todo, de um enunciado de fala.

Nesta pesquisa, além de examinar a variação do *pitch span* em unidades prosódicas demarcadas pelas pausas nos enunciados produzidos pelos sujeitos, investigarei também o chamado ***pitch accent***, ou seja, como esclarece LAYER (1994), qualquer configuração de *pitch* que torne uma sílaba proeminente.

Outras alterações na qualidade de voz como a ***velocidade de fala*** e a ***duração***, serão parâmetros também analisados, pois, combinados ou não, podem ser utilizados pelos falantes para marcar as ênfases de tal forma que as unidades afetadas no fluxo de fala serão ouvidas como proeminentes.

A duração representa a diferença de tempo entre dois eventos. Como menciona BARBOSA (1999), é preciso que dois eventos acústicos singulares ocorram no tempo para que se possa experimentar a sensação perceptiva da

duração. O autor acrescenta que: “A sensação de duração percebida é obtida, portanto, pelo concurso dos parâmetros prosódicos como um todo e não apenas pela duração mensurável por instrumentos de medida de tempo.” (BARBOSA, 1999, p.23)

A duração de segmentos de fala apresentam uma faixa de variação em torno de 30 a 300 milissegundos aproximadamente. A duração será medida nas tônicas das palavras enfatizadas e na extensão total dos trechos de fala analisados. A duração total dos enunciados é uma medida que possibilitará a verificação da velocidade de fala (número total de sílabas produzido durante um turno de fala completo).

LAVER (1994) distingue três tempos de fala: lento, médio e rápido. O autor afirma que na língua inglesa a velocidade de fala média situa-se em torno de 5,3 sílabas por segundo. Denomina essa medida de “*velocidade de articulação*” (“*articulation rate*”), enquanto que, para ele, a velocidade de fala (“*speaking rate*”) corresponde, na verdade, ao número de “*palavras por minuto*”, o que no inglês apresenta um valor médio em torno de 200 palavras por minuto. No entanto, essa distinção não é usual em nosso meio e, por isso, mantereí o emprego do termo ***velocidade de fala*** como expressão do número de ***sílabas por segundo***.

ROCHA FILHO (1997), em seu estudo: “*Som e a Ação na Narração de Futebol do Brasil*”, assinala que a narração de um jogo de futebol na televisão requer um ritmo mais cadenciado, enquanto que a narração no rádio impõe um ritmo mais acelerado. O locutor de uma emissora de televisão apresentou uma *velocidade de fala* de 4,4 sílabas por segundo enquanto que o locutor de uma estação de rádio apresentou 6,9 sílabas por segundo. Ambos realizavam a

narração de um mesmo episódio – um lance de ataque de um jogo decisivo pelas finais do Campeonato Brasileiro de Futebol.

Estes valores comparados à média do inglês, indicada por LAYER (1994), permitem classificar os três tempos em lento (4,4 sílabas/segundo), médio (5,3 sílabas por segundo) e rápido (6,9 sílabas por segundo), e servem como uma estimativa ilustrativa do parâmetro que estou considerando como velocidade de fala.

Além desses parâmetros, um outro aspecto a ser examinado é a ocorrência de palavras enfatizadas. Do ponto de vista perceptual, são palavras articuladas mais claramente, que apresentam sons alongados, e soam mais fortes (*loudness*) e altas (*pitch*) do que as que estão na sua vizinhança. Além disso, vêm, geralmente, seguidas de pausa.

Como aponta BOLINGER (1985) é mais fácil perceber um aumento em intensidade quando o *pitch* aumenta. Assim, a intensidade maior é reforçada pelo *pitch* alto. Como mencionei antes, quando a proeminência restringe-se à variação de *pitch* da sílaba tônica corresponde à ocorrência de *pitch accent*, se ela, contudo, co-ocorre com a sensação auditiva de alongamento, intensidade e altura extras, que caracterizam as palavras enfatizadas, corresponde a múltiplos eventos que sinalizam a ênfase realizada pelo falante.

O estudo desses parâmetros físico-acústicos permite investigar a manifestação da materialidade fônica expressa por meio dos aspectos paralingüísticos, os quais configuram a qualidade e a dinâmica vocal do falante engendrada na situação social, o *locus* de pesquisa negligenciado, segundo GOFFMAN (1964), e o tópico do próximo capítulo.

CAPÍTULO II

“It is true, then, that the frameworks in which words are spoken pass far beyond ordinary conversation. But it is just as true that these frameworks are brought back into conversation, acted out in a setting which they initially transcended. What nature divides, talk frivolously embeds, insets, transcended. As dramatists can put any world on their stage, so we can enact any participation framework and production format in our conversation.”

(Erving GOFFMAN, 1974)

ANÁLISE DA VOZ: AS IMPLICAÇÕES DE NATUREZA SOCIAL

*“Frame analysis would then become the study of everything
but ordinary behavior.”*

(Erving GOFFMAN, 1974)

Neste capítulo, trato da situação social engendrada na comunicação face a face como *locus* de pesquisa – um lugar negligenciado, como afirmou GOFFMAN em 1964, em seu artigo *“The Neglected Situation”*, no qual o autor convida os pesquisadores – lingüistas, sociolingüistas, antropólogos, sociólogos – a examinarem este fenômeno – a situação social como *locus* de pesquisa, convite válido na atualidade, para alguns campos de pesquisa, como no caso dos estudos em voz.

GOFFMAN foi um sociólogo que despontou nos anos que se seguiram ao final da Segunda Guerra Mundial e seus escritos tiveram grande impacto dentro e fora da Sociologia.

Em seus trabalhos, o autor descreve a complexidade das variáveis sociolingüísticas envolvidas na interação e, em GOFFMAN (1964), define *situação social*, como aquela que emerge a qualquer momento em que dois ou mais sujeitos se encontrem na presença imediata um do outro e que dura até que o penúltimo saia da situação. Sugere que o ato de fala deve ser sempre remetido ao estado de conversa, sustentado, particularmente, através do turno de fala, e envolvendo um círculo de outros participantes, ratificados como co-participantes. A conversa é socialmente organizada em princípios que ratificam, governam e regulam esta situação, um encontro social. Nas palavras do autor:

“ A interação face a face tem seus próprios regulamentos; tem seus próprios processos e sua própria estrutura, e estes não parecem ser de natureza intrinsecamente lingüística, mesmo que freqüentemente expressos por um meio lingüístico.”

(GOFFMAN, 1964)²⁷

Os elementos de natureza não presentes em sua obra, os aspectos paralingüísticos, dentre outros que serão explicitados ao longo deste capítulo e, em especial, o seu interesse pela vida cotidiana, despertaram minha atenção, uma vez que nos remetem à questão da realidade social. Sua forma de investigação da interação face a face no dia a dia despontava, para mim, como uma forte possibilidade analítica para compreender as estruturas emergentes nas situações da vida cotidiana, em particular, as variações das vozes de falantes em diferentes situações interacionais.

A partir de 1971, a obra de GOFFMAN voltou-se decisivamente na direção sociolingüística, a partir de seus trabalhos mais sistemáticos sobre análise da conversação (MALUFE, 1992), e sua contribuição para o estudo da materialidade fônica sob uma ótica social pode ser localizada em seus escritos, no que ele denominou *análise de esquemas de referência (frame analysis)*, em sua proposta de examinar a organização da experiência social, considerando o ato de fala e, particularmente importante para este estudo, também a produção vocal, como uma experiência social.

²⁷ Esta citação foi extraída da versão do artigo de GOFFMAN, traduzido por Pedro M. Garcez, em RIBEIRO & GARCEZ, 1998, p.15.

GOFFMAN define *frame* da seguinte forma:

“...I assume that definitions of situations are built up in accordance with principles of organization which govern events – at least social ones – and our subjective involvement in them; frame is the word I use to refer to such of these basic elements as I am able to identify. That is my definition of frame.”

(GOFFMAN, 1974, p. 10 -11)

A partir desta definição, o autor esclarece o termo *frame analysis*: “My phrase ‘frame analysis’ is a slogan to refer to the examination in these terms of the organization of experience”. (Ibidem, p. 11)

Essas premissas organizacionais, portanto, ancorarão a proposta de investigação desta pesquisa do ponto de vista social. Cabe acrescentar que, para o autor, tais premissas são algo que a cognição alcança, de alguma forma, e não algo que a cognição cria ou gera. Frente à compreensão do que está acontecendo, os indivíduos ajustam suas ações e, comumente, descobrem que o mundo sustenta esses ajustes.

Diante destas considerações preliminares, discuto em seguida, alguns dos conceitos centrais na obra do autor, que serão utilizados neste trabalho, as noções de “*frame*” (doravante considerado como *enquadre*), “*footing*” (*alinhamento* ou *posicionamento*), “*keying*” (*transposição*) e “*fabrication*” (*fabricação*) e “*frame dispute*” (*disputa de enquadre*).

No final deste capítulo, discuto sobre **voz na situação social**, um lugar negligenciado.

1. A SITUAÇÃO SOCIAL COMO *LOCUS* DE PESQUISA – O LUGAR NEGLIGENCIADO

A partir das pesquisas de Gregory BATESON em 1954, com a introdução da noção de *enquadre*, diversos estudos da linguagem foram desenvolvidos sob a perspectiva da interação, entre os quais destaca-se a produção de GOFFMAN, que, na mesma época, já expressava suas preocupações sobre a compreensão das situações face a face em artigo intitulado “*On Face-Work: an Analysis of Ritual Elements in Social Interaction*” (1955) ²⁸.

Nele, GOFFMAN discute aspectos da interação face a face e procura mostrar como os participantes exploram diversas estratégias para anunciar sua identidade social, estabelecendo uma *face* particular, e como as ações individuais necessitam ser consistentes em relação às expectativas dos demais componentes dos grupo, podendo resultar em alienação ou integração grupal conforme correspondam, ou não, a essas expectativas.

Em seguida, GOFFMAN publicou “*The Neglected Situation*” (1964) e, a partir da sua proposta de análise da situação social, o estudo da língua passou a ser visto de uma forma diferente (grifos meus):

*“...o estudo da relação língua e sociedade passa a ser visto
a partir do uso da fala em contextos sociais específicos,*

²⁸ In: LAVER, J. & HUTCHESON, S., 1974 – **Communication in Face to Face Interaction**. Esse artigo foi publicado em 1967 em **Interaction Ritual. Essays on face to face behavior**, uma coletânea de artigos do autor escritos nos anos 50. Há, também, uma versão em espanhol - **Ritual de la Interacción**, Argentina: Editorial Tiempo Contemporáneo, 1970.

*assumindo um arcabouço teórico bem mais complexo. Já não se trata mais de realizarmos estudos correlacionais entre variáveis lingüísticas e variáveis sociais; nem de aprofundarmos estudos indicativos do comportamento lingüístico por si só. GOFFMAN nos convida a examinarmos **a situação social como o locus da pesquisa – o lugar que tem sido negligenciado** – e deixa um alerta aos interessados: uma vez cruzada a ponte entre os estudos da fala e da conduta social, nos tornaremos por demais ocupados para olharmos para trás.”²⁹*

Posteriormente, em *Frame Analysis - An Essay on the Organization of Experience* (1974), GOFFMAN aprofunda essa discussão, propiciando um extenso estudo sobre a estrutura de organização, conceitos e situações, apresentando e analisando exemplos que elucidam diversos tipos e níveis de *enquadres* que constituem a interação social no dia a dia.

Embora, não se tenha ocupado, particularmente, do fenômeno vocal, mas sim da **estrutura organizacional da situação de fala** como um todo, as pesquisas de GOFFMAN trazem uma contribuição substancial para a análise do discurso oral e da situação de interação e são, portanto, essenciais para estudos como o presente. Além disso, o lugar que o autor considerou negligenciado, continua sendo-o, principalmente no que se refere à voz.

Diversos trabalhos sofreram sua influência, entre os quais, alguns voltados para análise do discurso como os de TANNEN & WALLAT (1987 e 1993), que analisaram o discurso de uma consulta/exame pediátrico, mostrando a

²⁹ GOFFMAN, 1964, citação extraída da versão em português de 1998, p.11.

complexidade do jogo discursivo mesmo num contexto ritualizado como o da consulta médica; e o de RIBEIRO (1988 e 1991), que faz a análise de uma entrevista psiquiátrica de um paciente, demonstrando que a situação interacional é construída a partir dos papéis sociais projetados e vivenciados pela paciente, somado aos diferentes alinhamentos em relação à médica. TANNEN & WALLAT buscam aprofundar os conceitos introduzidos por GOFFMAN na análise do discurso.

Outra pesquisa que se utiliza dos fundamentos desse autor foi a de MADUREIRA (1992), que enfoca os efeitos dos sentidos produzidos pelos recursos fônicos, defendendo a posição de que o papel do som na construção do sentido, no discurso oral, pode ser descrito com base na relação entre sua manifestação no discurso e sua manifestação física, ou seja, a maneira como os fenômenos fônicos se apresentam no discurso pode refletir aspectos da natureza dos fenômenos físicos que ocorrem na produção, transmissão e recepção dos sons vocais. A pesquisadora recorre a GOFFMAN para analisar a situação de fala em que desenvolveu sua pesquisa – uma palestra, apontando como os recursos fônicos são utilizados pelo falante (palestrante) para veicular efeitos de sentido, sinalizando mudanças de *alinhamento* em relação ao público ou ao discurso.

Outro estudo também voltado aos aspectos fônicos foi o realizado por GROSJEAN (1993). Citando GOFFMAN e BAKHTIN como referenciais teóricos para sua pesquisa, ela investigou as funções interacionais da voz numa situação específica — de parteiras durante a condução de partos — a partir da

consideração de que as variações vocais ocorridas no desenrolar da situação eram um índice essencial das mudanças de posição do falante (parteiras).

Observa-se nestes trabalhos uma preocupação em investigar o discurso na organização social, ou seja, um movimento analítico em direção ao *locus* de pesquisa negligenciado para GOFFMAN – **a situação social**. Como escrevem RIBEIRO & GARCEZ (1998), em estudos desse tipo, ressalta-se:

*“... a natureza dialógica da comunicação humana e o intenso trabalho social e lingüístico implícito na co-construção do significado. Nesta abordagem do discurso, tanto o falante como o ouvinte têm papéis ativos na elaboração da mensagem e na definição ‘do que está acontecendo aqui e agora’. Não há, portanto, significado que não seja situado. A noção de contexto ganha relevância passando a ser entendida como criação conjunta de todos os participantes presentes ao encontro emergente a cada novo instante interacional. Os interagentes levam em consideração não somente os dados contextuais relativamente mais estáveis, sobre participantes (quem fala para quem), referência (sobre o quê) espaço (em que lugar) e tempo (em que momento), mas consideram sobretudo a maneira como **cada um dos presentes sinaliza e sustenta o contexto interacional em curso.**”*

(op.cit., p. 9, grifos meus)

Focalizar as pesquisas em uma perspectiva social, contudo, é ainda, um situação pouco privilegiada, e que merece a atenção de um maior número de pesquisadores.

Os grifos, que assinalei no trecho acima, ilustram a presença dos aspectos **paralingüísticos** em sua obra. O autor afirma que a forma como cada um dos participantes *sinaliza e sustenta* seu *posicionamento*, numa dada situação, pode ser analisada por meio de aspectos não lingüísticos, sinalizados por movimentos corporais, olhares e também pelo “*tone of voice*”. Desta forma, seus estudos tem uma contribuição decisiva para que a **situação social** comece a ser menos negligenciada nas pesquisas em voz.

1. 1. O Enquadre na Interação Face a Face – a Noção de “Frame”

O *enquadre* - “*frame*”- é um conceito introduzido por BATESON nas ciências sociais, em seu artigo “*A Theory of Play and Fantasy*”(1972)³⁰, no qual o autor discute a natureza da comunicação, assinalando que nenhum enunciado do discurso poderia ser compreendido sem referência a um *enquadre*. Esse conceito foi posteriormente aprofundado por GOFFMAN, que afirma que, em qualquer situação face a face, os participantes permanentemente introduzem, ou mantêm, enquadres que organizam a situação e ajudam os sujeitos a se orientar em relação a ela.

BATESON relata que o primeiro passo definido para a formulação de sua hipótese ocorreu em 1952, quando foi ao Zoológico em São Francisco e se deparou com um fenômeno bastante conhecido de todos: dois macacos *brincando* (grifo do autor), porém utilizando ações ou sinais que, individualmente, eram semelhantes (mas, não idênticos) aos usados em situação de combate.

³⁰ Este artigo foi escrito em 1954, re-publicado em 1972 no livro **Steps to na ecology of mind** (pp. 177-193) e traduzido por Lúcia Quental, em RIBEIRO & GARCEZ, 1998.

Qualquer observador conseguiria perceber que aquela situação não representava um combate. Deveria, portanto, haver algum tipo de metacomunicação veiculando a mensagem de que aquela situação específica era uma *brincadeira*. Essa mesma análise poderia ser estendida para situações de *ameaça* ou de *fantasia*. A questão que se colocava era “*Será isto uma brincadeira?*”. Seria um *enquadre* de brincadeira? Para que pudessem interpretar uma dada situação, bem como contribuir para construção da comunicação numa situação face a face, os participantes estariam constantemente atentos aos sinais que delimitariam ou contextualizariam os *enquadres* – “*Isto é uma brincadeira?*” ou “*Isto é uma ameaça?*”.

BATESON utiliza dois tipos de analogia para discutir essa noção - uma, física, análoga a uma moldura de um quadro (imagem concreta). Segundo ele, assim, como a moldura é indicativa da direção para a qual um observador deve dirigir seu olhar quando observa um quadro, o *enquadre* conteria uma série de informações para que o ouvinte pudesse entender uma dada mensagem. O *enquadre* delimitaria figura e fundo, ruído e sinal e, para BATESON seria a forma como as pessoas delimitam os eventos. A outra analogia utilizada pelo autor, mais abstrata, seria a de uma linha imaginária (imagem abstrata) de um conjunto matemático, de inclusão e exclusão de premissas comuns.

O primeiro passo para a definição de um *enquadre psicológico* seria dizer o que é ou como se delimita uma classe ou um conjunto de mensagens ou de ações significativas. Em termos “matemáticos”, essas mensagens poderiam ser representadas por pontos e o “conjunto” (aspas do autor) separado por uma linha imaginária de mensagens sobre o que não seria brincadeira. BATESON considera

essa analogia matemática falha, porque o enquadre psicológico não seria satisfatoriamente representado por essa linha imaginária. Em alguns casos, podem haver dados de realidade, como referências verbais explícitas dos enquadres e, em outros casos, os sujeitos podem não ter consciência deles.

O autor assinala que, ainda que a analogia do conjunto matemático possa ser considerada abstrata demais, a outra, análoga a uma moldura de um quadro, seria muito concreta, não se tratando de uma conceituação nem física, nem lógica. A mensagem *“Isto é uma brincadeira”* estabeleceria um enquadre que poderia precipitar um paradoxo, ou seja, uma tentativa de diferenciar ou de traçar uma linha entre categorias de tipos lógicos diferentes. BATESON conclui o artigo ressaltando que sua tese central pode ser resumida como a afirmação da necessidade dos paradoxos de abstração:

“... acreditamos que os paradoxos de abstração devem estar presentes em todos os tipos de comunicação mais complexos do que sinais de humor, e que sem esses paradoxos a evolução da comunicação estaria estagnada. A vida se tornaria então, uma troca infindável de mensagens estilizadas, um jogo com regras rígidas, sem o alívio da variação ou do humor.”

(BATESON, 1972)³¹

Suas idéias sobre o *enquadre* influenciaram pesquisas na área da comunicação e na da psicologia. Um dos mais importantes trabalhos sobre o

³¹ Este trecho foi extraído do artigo traduzido por Lúcia Quental, em RIBEIRO & GARCEZ, 1998, p.69)

tema, entretanto, foi desenvolvido no campo da sociologia, por GOFFMAN, em *Frame Analysis - An Essay on the Organization of Experience* (1974). Seu objetivo era tentar isolar alguns enquadres básicos de compreensão da sociedade que fizessem sentido fora dos eventos e analisar as vulnerabilidades especiais, às quais estes enquadres estariam sujeitos, propiciando, assim, uma extensa análise do discurso produzido na interação.

No início de seu trabalho, GOFFMAN (1974) informa que, embora faça uso do termo *frame*, como BATESON, baseou-se também em estudos da Sociologia e da Filosofia da Linguagem, para traçar seus fundamentos teóricos. Assume que definições de uma situação são construídas de acordo com os ***princípios de organização que governam os eventos*** (grifos meus), pelo menos os sociais, e pelo envolvimento subjetivo neles, explicando que *frame* é a palavra que ele utilizará para se referir a esses elementos básicos. Esclarece, ainda, que “*frame analysis*” é a expressão que usará para se referir ao exame desses elementos na organização de uma experiência. Segundo GOFFMAN, seu livro trata da organização de experiências - algo que um ator individual pode tirar de sua mente e não da organização da sociedade.

Em algumas áreas, como na psicologia, os termos *enquadre* e *esquema* passaram a ser utilizados, indiscriminadamente, para descrever diferentes tipos de noções. Em uma revisão de literatura realizada por TANNEN (1979, apud TANNEN & WALLAT, 1987), visando a uma uniformização desses e de outros conceitos afins, a autora verificou que havia uma variedade de conceitos para *enquadre* e *esquema*, bem como pouco uso dos tipos de enquadres que GOFFMAN exaustivamente analisou em 1974. Preocupada com a confusão que

tal situação certamente criaria entre pesquisadores do tema, TANNEN, em parceria com a psicóloga Cynthia WALLAT, em seu artigo de 1987, propõe um melhor discernimento para essas noções, sugerindo a utilização do termo *enquadre* sempre que houver referência aos enquadres interativos de comunicação e, a de *esquema*³², quando a referência for a esquemas de conhecimento, conforme abalizado na psicologia e na inteligência artificial. Propuseram também ampliar a discussão sobre *enquadre*, para que o sentido antropológico e sociológico do termo pudesse ser abrangido e integrado.

As autoras consideram que os vários usos do termo *enquadre*, e outros afins, podem ser subdivididos em duas categorias: uma, de natureza interacional, refere-se aos “*enquadres de interpretação*”, da forma utilizada por BATESON, GOFFMAN e GUMPERZ (1982), e a outra, de estruturas de conhecimento, que as autoras preferem chamar de *esquemas*, distinguindo-a, assim, de *enquadres*, embora ambas guardem forte relação.

Para TANNEN & WALLAT (1987), a noção interativa de *enquadre* refere-se ao que acontece em uma dada interação, na qual os interlocutores têm a tarefa interpretativa de reconhecer em que tipo de *enquadre* um enunciado foi composto, como exemplo - “*É uma piada?*” ou “*É uma discussão?*”. Essa noção, segundo elas, diz respeito “... à percepção de qual atividade está sendo encenada, de qual sentido os falantes dão ao que dizem” (op.cit., p. 124). Acrescentam ainda que,

³² Não constitui objetivo desse estudo retomar nem aprofundar a noção de *esquema*, uma vez que não é um conceito necessário para a análise aqui proposta. RIBEIRO & GARCEZ (op.cit., p.70) esclarecem que Erving GOFFMAN, como sociólogo, optou pelo termo *frame* e não *schemas*, tendo em vista salientar os aspectos sociais descritos em seus estudos e não marcar a natureza cognitiva implícita na noção de *esquema*. Da mesma forma, os autores optaram pela utilização do termo “*enquadre*” para a tradução de *frame* em sua coletânea de textos, posição esta também seguida por mim.

em qualquer situação interacional, todos os participantes contribuem na negociação de todos os enquadres desenvolvidos dentro da interação.

Através da análise que fazem da entrevista pediátrica, TANNEN & WALLAT mostram como pistas lingüísticas, ou maneiras de falar, evidenciam e sinalizam a mudança de enquadres e de esquemas. Destacam, ainda, que a importância desse estudo vai além dos limites do cenário médico, uma vez que, acreditam, enquadres e esquemas operam de maneira semelhante em qualquer interação face a face. Apontam também a necessidade de investigações mais profundas, que busquem as diferenças individuais e sociais de enquadres e esquemas, lingüísticas e/ou paralingüísticas, através das quais eles podem ser construídos e identificados. Esse interesse das autoras vem corroborar com os interesses deste estudo, particularmente nos aspectos paralingüísticos.

Ao afirmarem que o termo *enquadre* diz respeito ao sentido que os participantes constróem acerca do que está ocorrendo na situação interacional, TANNEN & WALLAT estão, de certa maneira, refletindo a noção de *footing* desenvolvida por GOFFMAN, ou seja, do *alinhamento* que os participantes estabelecem para si e para os outros numa dada situação. O conceito de *footing* foi desenvolvido por GOFFMAN a partir da ampliação e do desdobramento de seus estudos sobre os *frames* utilizados no dia-a-dia, nos quais focalizava, mais especificamente, os usos da linguagem na análise lingüística do discurso. Sua primeira referência a *footing* ocorreu em sua coletânea de cinco artigos, escritos entre 1974 e 1980: "*Forms of Talk*", publicada em 1981. Para o autor, *footing* representa o alinhamento, a postura, a posição do participante de uma interação face a face na sua relação com o outro, consigo próprio e com o discurso em

construção, sinalizando aspectos pessoais e papéis sociais e/ou discursivos. Esta noção é abordada a seguir.

1.2. O Posicionamento na Interação Face a Face – a Noção de “Footing” e as “Transformações” nos Eventos de Fala

“This paper is largely concerned with pointing out that participants over the course of their speaking constantly change their footing, these changes being a persistent feature of a nature of talk.”

(GOFFMAN, 1981)

O próprio GOFFMAN explica a importância do termo *footing*, como mudanças que são características persistentes, próprias da natureza da conversação. Para introduzir o conceito de “*footing*”, retomo **um segmento de interação** utilizado pelo autor (1979 e 1981)³³. Trata-se de um *release* divulgado nos jornais em 1973 sobre atividades presidenciais — um episódio que envolvia o Presidente Nixon e uma jornalista. A cena se passa no Salão Oval da Casa Branca e os participantes eram um grupo de funcionários do governo e representantes da imprensa reunidos por razões profissionais para testemunhar um ritual político, a assinatura de um documento:

“WASHINGTON (UPI) - O Presidente Nixon, um cavalheiro à moda antiga, resolveu caçoar de uma jornalista que usava calças

³³ O artigo “*Footing*” foi publicado originalmente em 1979 e traduzido por Beatriz Fontana em RIBEIRO & GARCEZ, 1998. Foi re-publicado em **Forms of Talk**, 1981. Como utilizei os dois textos, cito as referências, conforme o emprego que faço dos mesmos.

compridas na Casa Branca, deixando bem clara sua preferência por vestidos.

Após a cerimônia de assinatura de um documento no Salão Oval, o Presidente se levantou de sua mesa de trabalho e, em tom de gracejo, dirigiu-se a Helen Thomas, da UPI: ‘Helen, você continua usando calças compridas. Você prefere mesmo? Quando eu vejo moças usando calças compridas, me lembro da China.’

Helen Thomas, um tanto desconcertada, disse ao Presidente que as chinesas estavam passando a usar roupas cada vez mais ocidentais.

‘Isto não é uma crítica, mas calças, às vezes, ficam bem para algumas pessoas e para outras não.’ Apressou-se em acrescentar, ‘mas acho que lhe caem bem. Dê uma voltinha.’

Enquanto Nixon, o Ministro da Justiça, Elliott L. Richardson, o diretor do FBI, Clarence Kelley, e outros altos funcionários jurídicos sorriam, Helen Thomas fez uma pirueta para o Presidente. Ela usava calças brancas, camisa de malha azul marinho, colar de contas brancas e sapatos debruados de vermelho.

Nixon perguntou à Helen Thomas o que seu marido, Douglas Cornell, achava de seu hábito de usar calças compridas.

- Ele não se importa. - ela respondeu.

- Calças são mais baratas do que vestidos?

- Não - disse Helen Thomas.

- Então mude - determinou o Presidente com um sorriso largo e malicioso, enquanto os outros repórteres e cinegrafistas caíam na gargalhada.”

(*The Evening Bulletin*, Filadélfia, 1973)³⁴

Esse incidente evidencia uma mudança nos papéis desempenhados pelos participantes do episódio: o presidente assumiu o papel de homem “galanteador” e a jornalista o de mulher, o que levou a um (re)ajuste dos participantes, que abandonam a postura e a linguagem formais, pertinentes à uma relação de trabalho, e adotam um tom mais informal de participação. Para GOFFMAN (1981), por trás desse fato há o pressuposto contemporâneo segundo o qual, as mulheres deveriam estar sempre preparadas para receber comentários sobre sua “aparência”. Isto, implicitamente, significaria que a mulher deve estar sempre pronta para mudar de posição e, quando é levada a mudar, transforma-se, ainda que, momentaneamente, no objeto de atenção e aprovação, ou não, perdendo seu papel de participante usual da situação.

Essa passagem mostra o poder de uma pessoa num cargo hierarquicamente superior – o de presidente, para forçar um indivíduo, que é do sexo feminino, do culturalmente chamado “*sexo frágil*”, a sair de seu lugar/papel profissional, para ocupar um lugar/papel “*sexual e doméstico*”, durante uma ocasião na qual, segundo GOFFMAN, ela (e as muitas mulheres que lhe atribuísem o papel simbólico de representante da classe feminina)³⁵ estaria, provavelmente, preocupada em ser respeitada como profissional e, somente como tal. O autor esclarece que esse segmento evidencia um momento da

³⁴ GOFFMAN (1979), extraído da tradução de Beatriz Fontana, em RIBEIRO & GARCEZ 1998, p.71.

³⁵ Trata-se de adendo do próprio GOFFMAN, traduzido por mim.

política de “*gênero*”(no sentido de identidade sexual)³⁶, no qual o presidente poderia, impensadamente, ter exercido esse poder.

A alterações como essas, que ocorreriam em qualquer situação face a face e seriam necessárias ao jogo interacional, GOFFMAN chamou de *footing*.

O autor afirma sua preocupação em mostrar que os participantes, ao longo de suas falas, mudam constantemente os *footings*, considerando essas mudanças como características pertinentes à fala natural. Como explica GOFFMAN uma mudança de *footing* implica numa mudança de posicionamento/alinhamento, que o sujeito assume para si próprio e para os outros, expresso na maneira com que o sujeito gerencia e recebe o enunciado. É uma mudança que está comumente vinculada à linguagem e, mesmo que não, pelo menos os marcadores paralingüísticos estarão presentes. Os *footings* são introduzidos, negociados, sustentados e modificados na própria situação de interação e, portanto, a análise de uma situação interacional nesta perspectiva permite que se compreenda como as identidades sociais e lingüísticas emergem e se constituem no discurso, bem como entender como elas afetam a interação em curso (RIBEIRO & GARCEZ, op.cit.). Por essa razão, é de grande importância a obra de GOFFMAN nesta linha de investigação.

Em seus estudos, GOFFMAN retoma a análise tradicional da atividade conversacional, cujo paradigma tradicional é o de que dois, e apenas dois, indivíduos estão juntos, envolvidos nesse ato, um na qualidade de ouvinte e outro na de falante. Os turnos de fala se sucedem, alternadamente, entre aqueles que têm o direito à fala e que ocupam a cena. A qualquer momento, e de cada vez, um estará falando e o outro ouvindo. Para GOFFMAN, trata-se de um arranjo

³⁶ Esclarecimento dos editores na tradução deste artigo feita por Beatriz Fontana (op.cit., p. 70).

para o qual os termos “*falante*” e “*ouvinte*”, embora leigos, se ajustam perfeitamente e se mostram adequados para necessidades técnicas deste tipo. O autor aponta a necessidade de ampliação dessa visão, em que sejam incluídos outros elementos que constituem uma situação de interação, tais como participantes eventuais em uma situação interacional e componentes paralingüísticos, como o “*tom*” de voz, os gestos, os olhares e as mudanças fisionômicas.

Em relação aos aspectos paralingüísticos, o autor assinala que conversar não é apenas *ouvir* o outro, mas também *ver* como algo é dito. Muito do que se conversa se processa no âmbito do visual. A visão é, do ponto de vista organizacional, muito importante para o gerenciamento da tomada de turnos, para a avaliação da recepção das pistas visuais, das gestuais e da sincronia do olhar, e para a atenção na situação de conversa.

Segundo o autor, a análise tradicional do dizer e do que é dito parece tacitamente comprometido. A qualquer momento um dos dois participantes exporá seus pensamentos e o outro estará ouvindo. O discurso em si, seria a principal preocupação de ambos. Apenas, estes dois indivíduos sabem quem está falando, quem está ouvindo, o que está sendo dito, aspectos “*inacessíveis*” ao outro. No curso desta situação, haverá a troca de papéis de falante e ouvinte para manter o formato afirmação/resposta: “...*diz-se que o que se passa é conversa, fala, conversa.*” (GOFFMAN, 1979)³⁷.

GOFFMAN considera necessário aperfeiçoar esse paradigma tradicional da conversa. Para tanto propõe que a contribuição seria reconhecer que *qualquer*

³⁷ Citação extraída da versão em português, 1998, p.75.

momento da conversa possa ser parte desta **conversa**, isto inclui qualquer trecho da conversa, inclusive aqueles em que não se fala nada.

O autor quer dizer com isso que, mesmo em períodos em que não ocorra conversa, os participantes estão em “estado de conversa” e, não necessariamente, é necessário que se mantenha a “obrigação” de os participantes de um encontro social sustentarem o envolvimento no que está sendo dito, para se assegurar que não haja períodos sem que ninguém faça uso da palavra. O autor afirma que creditando-se autonomia a uma conversa “*surge um domínio sui generis para análise*”. Diante desta possibilidade, propõe que os momentos de fala (grifo do autor) sejam considerados como unidades de análise.

Dessa forma, como assinala RIBEIRO & GARCEZ (1998, p.70), GOFFMAN “...desconstrói as noções clássicas de falante e ouvinte e discute as complexidades das relações discursivas presentes na estrutura de produção (relativa ao falante) e na estrutura de participação (relativa ao ouvinte).”

Tomando, inicialmente, a noção de **ouvinte** (*interlocutor*), GOFFMAN (1981) aponta que tanto um interlocutor participante ratificado (*ouvinte oficial*) de um encontro social pode simplesmente não estar ouvindo o falante, quanto um participante não oficial (“*não ratificado*”) pode estar escutando por acaso (às escondidas ou à distância, por exemplo). A este participante não ratificado, o autor denominou de “*circunstante*”(versão de Beatriz Fontana, op.cit.), que se torna ouvinte “*por acaso*”. Assim, um participante ratificado (“*ratified*”) pode não estar ouvindo e um outro que esteja escutando pode não ser o participante “*oficial*”. O autor distingue o ouvinte ratificado (a quem a conversa é endereçada) do ouvinte *não ratificado* (a quem a conversa não é endereçada), questionando, assim, o conceito clássico de ouvinte. Rompe-se, dessa forma, os limites

diádicos da fala e os circunstantes são, então, admitidos à cena, o que implica em ajustes no jogo conversacional em curso: tem-se, assim, um *“jogo paralelo”*, ou, às vezes, *“um jogo cruzado”*.

Esses participantes eventuais seriam *“circunstantes”* (*“by-standers”*), cuja presença seria uma regra e não uma exceção. Seriam espectadores de uma conversação em andamento. Sua participação pode não ser ouvida pelo simples fato de não estarem falando nada. Os sinais de acompanhamento da conversa, porém, são visíveis. Ele aponta que, no decorrer desse encontro, ocorrem fases distintas: o início é marcado pela aproximação dos participantes, uns se movem em direção aos outros, passando a se mover conjuntamente, e o encerramento é sinalizado pelo afastar-se.

Ao lidar com a noção de *“circunstantes”*, GOFFMAN alerta que foi tacitamente efetivada uma alteração no ponto de referência: o *encontro social* passa ao *status* de *“situação social”* - definida como a arena física absoluta na qual as pessoas presentes estão ao alcance visual e auditivo umas das outras. Acrescenta, também, que é comum os falantes modificarem seu modo de falar, e também o que estão dizendo, pelo fato de conduzirem sua fala dentro do limite auditivo e visual dos não participantes. Para ele, a evidência mais clara do significado estrutural da situação social para a fala pode, talvez, ser encontrada na atitude verbal de um sujeito quando está a sós, mesmo com a presença ocasional de desconhecidos. Algumas regras impeditivas de comunicação obrigam o sujeito a desistir de usar a fala e as palavras (*“som articulado”*).

No entanto, existe uma larga gama de circunstâncias, nas quais o sujeito dirige declarações a si mesmo e acaba por deixar aparecer exclamações do tipo:

“Oops!” “Eek!” (exemplos extraídos do texto original em inglês) e similares. Para o autor, essas são, indubitavelmente, manifestações do sujeito, na tentativa de encontrar alguma resposta, não necessariamente específica, de alguém que possa estar escutando. A intenção é fornecer informação a quem estiver ao alcance, mas sem fazer uso de um turno conversacional. O sujeito não está a procura de ouvintes propriamente ditos, mas sim de ouvintes ocasionais.

Examinando-se um sujeito específico falando (um recorte momentâneo), essa possibilidade de análise permite descrever o papel, ou a função, de cada um dos componentes desse agrupamento social — independente de serem participantes ratificados, ou não — a partir desse ponto de referência. Nessa abordagem, o “status de participação” (“*participation status*”) corresponde à relação de qualquer um dos participantes com um determinado enunciado, e a relação das pessoas participantes desse agrupamento com um dado enunciado pode ser chamado de “estrutura de participação” (“*participation framework*”) para esse momento de fala. Esses dois termos são empregados mudando-se o ponto de referência, de **falante** para algo mais abrangente, quer seja, **toda a atividade na situação em si**. GOFFMAN ressalta que o aspecto fundamental nessa mudança é que o enunciado não divide o mundo além do falante precisamente em duas partes – interlocutores e não interlocutores, mas, ao contrário, abre um leque de possibilidades que estabelece **a estrutura de participação segundo a qual o falante conduzirá sua fala**.

Decomposta a noção global de *ouvinte* no paradigma conversacional, torna-se necessário que o autor examine a noção de falante. Em uma conversa padrão o falante é a “*máquina de falar*”, um indivíduo que tem a tarefa de produzir

elocuções, nos termos de GOFFMAN. Ele está funcionando como um *animador*. No entanto, às vezes, há um *autor* do que é dito e um *principal* (“responsável”, na versão de Beatriz Fontana, 1998, que seguiu MALUFE, 1992), a quem pertence a posição e a crença do que é dito.

As noções de *animador*, *autor* e *responsável*, tomadas em conjunto, podem contribuir para elucidar o formato de produção de um ato (momento) de fala. Ao se empregar o termo *falante*, implicitamente, espera-se que o falante ocupe os três papéis, ou seja, que ele *anime* seu *próprio* texto e seja *responsável* pela posição estabelecidas pelas palavras faladas, no entanto, essa superposição implícita de papéis pode não ocorrer, conforme a situação social em curso.

GOFFMAN recomenda que para se chegar à base estrutural do *footing* é necessário tomar as noções tradicionais de ouvinte e falante em partes diferenciadas, isto é, estrutura de participação (relacionada ao ouvinte) e formato de produção (relacionada ao falante). Cada movimento do jogo conversacional traz consigo uma mudança de *footing*, embora o autor alerte que esse movimento não é algo “*mecânico*”. Como ele explica:

“ Frequentemente parece que ao **mudarmos de voz** – seja para falar por outro aspecto de nós mesmos ou por outra pessoa, ou para deixar nosso discurso mais leve com uma apresentação mais rápida de algum arranjo de uma interação fora do contexto – não estamos exatamente encerrando o alinhamento anterior, mas suspendendo-o temporariamente, com o entendimento de que ele será quase imediatamente restabelecido. Da mesma forma também, quando cedemos a palavra numa conversação,

*assumindo assim o 'footing' de interlocutor (endereço ou outros)
fica-nos garantida a expectativa de reingressar no papel de
falante, no mesmo 'footing' em que a interrompemos.*

(GOFFMAN, 1979, p. 96, grifo meu)

Retomo o texto original para completar essa citação: *"In truth, in talk it seems routine that, while firmly standing on two feet, we jump up and down on another."* (GOFFMAN, 1981, p. 155).

As figuras de linguagem e exemplos do cotidiano são freqüentes no texto de GOFFMAN e, por isso, no trecho acima, optei pelo original. Ele afirma que essas mudanças de *footing*, este "*saltitar*" de um pé para outro, e manter-se firme sobre os dois pés deveria servir como ajuda, por exemplo, para o enfrentamento de nichos institucionais, nos quais o indivíduo rotineiramente é forçado a sustentar simultaneamente mais do que um estado de fala.

O autor afirma que a Lingüística fornece pistas e marcadores por meio dos quais os *footings* se tornam manifestos e ajudam a encontrar o caminho de uma base estrutural para analisá-los.

Compreender a estrutura organizacional de uma situação discursiva (oral) pode fornecer subsídios para o entendimento da variação do fenômeno vocal num dado ***momento de fala***. Assim como GOFFMAN, considero que a análise de um ato de fala vai muito além da mera produção dos *sons*:

*"For example, the terms 'speaker' and 'hearer' imply that
sound alone is at issue, when, in fact, it is obvious that sight is
organizationally very significant too, sometimes even touch. In the*

management of turn-talking, in the assessment of reception through visual back channel cues, in the paralinguistic function of gesticulation, in the synchrony of gaze shift, in the provision assessment of engrossment through evidence of side-involvement and facial expression – in all of these ways it is apparent that sight is crucial, both for the speaker and for the hearer. For the effective conduct of talk speaker and hearer had best be in a position to watch each other.”

(GOFFMAN, 1981, pp. 129-130, grifo meu)

Evidentemente, apesar do valor dado aos aspectos paralingüísticos, estes não constituíam preocupação central nos trabalhos de GOFFMAN. Entretanto, ao longo do seu texto há várias referências sobre a necessidade dos marcadores e/ou pistas paralingüísticas para que o jogo interacional ocorra. Dentre eles, neste estudo, me deterei, particularmente, nas variações da voz em decorrência desse movimento entre os participantes em um dado momento de fala. Em seu trabalho, o autor escreve que sem o acesso à orientação corporal e *ao tom de voz* poderia haver um erro de análise de um dado acontecimento, pois mudanças significativas de alinhamento entre falantes e ouvintes, poderiam não ser percebidas (GOFFMAN, 1979).

No exemplo, citado anteriormente, relativo às exclamações que um sujeito utiliza ao falar sozinho, o autor observa que, para alcançar sua finalidade, ele fará um ***ajuste do volume do som*** (grifo meu) de tal forma que os envolvidos na situação social possam ouvir seus comentários.

Analisar eventos de fala implica em analisar coisas ditas e que estão sujeitas a transformações como “*keyings*” (*transposições*) e “*fabrications*” (*fabricações*). *Transposição* é um conceito que designa as convenções mediante as quais um fragmento de atividade pode ser transformado e enquadrado em diferentes esquemas de referência.

A *transposição* propicia uma forma básica na qual uma parte da atividade pode ser transformada, isto é, serve como modelo item por item para algo mais. Colocado de outra forma, “*keying*” representa a maneira básica pela qual uma atividade é vulnerável. Uma segunda vulnerabilidade transformacional é a *fabricação*. O termo refere-se ao esforço intencional de um ou mais indivíduos em manejar uma atividade de tal forma que um ou mais participantes poderão ser induzidos a ter uma “*falsa crença*” sobre o que está acontecendo.

Enquanto a *transposição* leva os participantes a terem a mesma visão do que está acontecendo, a *fabricação* implica em diferentes visões. A fabricação distintamente da transposição está sujeita a um certo *descrédito*.

GOFFMAN distingue as fabricações em “*benigna*” (*benign fabrications*) “e “*aproveitadora*” (*exploitive fabrications*). A primeira é arquitetada em nome do interesse ou benefício do participante, pelo menos não vai contra seu próprio interesse, e a outra vai contra seu interesse. As fabricações introduzem a possibilidade de um diferente tipo de interrupção, na qual descobertas podem, pontualmente, alterar a capacidade daqueles envolvidos em, novamente, participar juntos naquele tipo de atividade (evento de fala).

Outra noção importante é a de “*frame dispute*” (disputa de enquadre), que ocorre quando as partes envolvidas podem claramente disputar um com outro para definir o que está acontecendo.

Como os *frames* incorporam tanto a resposta do participante quanto a do mundo em que ele está respondendo, um elemento reflexivo deve necessariamente estar presente em qualquer visão dos eventos dos participantes, uma correta visão da cena deve ser a de incluir a visão da mesma como parte dela.

GOFFMAN propõe uma maneira básica de conceber o fenômeno para a análise da conversação, fundando-a em uma sequência de movimentos de respostas. Numa situação conversacional, cada elemento contém elementos de perguntas e de respostas, que se combinam de diferentes modos, segundo o contexto em que ocorre a conversação. Em alguns momentos, podem ocorrer *transgressões*, como no episódio entre o Presidente Nixon e a jornalista, em que, circunstancialmente, saem de um enquadre “*profissional e formal*” e passam a ocupar um outro enquadre, mais “*informal*”. Em outros momentos, esse movimento pode ser “*fabricado*”, algo que não ocorreu de fato.

Para o autor é verdadeiro que o contexto ajuda a compreender significados não intencionais e a suprimir mal entendidos, mas o poder do ambiente não pode ser independente da competência cultural dos intérpretes. A interpretação pressupõe competência cultural e lingüística.

Evidenciam-se, assim, a utilidade e a aplicabilidade dos conceitos de GOFFMAN nesta pesquisa. Como indica o autor, ao se considerar momentos de fala como unidade de investigação, creditar-se-ia maior autonomia a “*uma*

conversa”, como unidade de análise. Além disso, características importantes da conversação não se perderiam; os limites diádicos de fala do paradigma tradicional romper-se-iam e os “*circunstantes*” e/ou mais de um interlocutor “*ratificado*” seriam admitidos na cena, sem se perder de “*vista*” os aspectos paralingüísticos fundamentais na situação interacional.

2. A VOZ NA INTERAÇÃO SOCIAL – UM ESPAÇO NEGLIGENCIADO

“A voz, como vimos durante o decorrer do trabalho, causa impacto no outro, envolve-o numa grande malha de significações mediando a construção do conhecimento. Amalgamado a este propósito, a voz contém emoções. Suas mais sutis variações revelam vibração, alegria, surpresa. Ela repreende e coíbe, mas logo convida e permite. Deste modo revela emoção e cognição, locadas no sujeito, mas produto de um longo caminho de relações sociais.”

(Emilse Aparecida Merlin SERVILHA, 2000)

A autora, acima citada, retrata a voz como emoção, vida, mediadora do conhecimento, vibração, expressão do dono que proíbe, mas também, acolhe e, principalmente, como produto de um longo caminho de relações sociais. É esse caminho que pretendo abordar – a voz na interação social, um espaço negligenciado.

Ao longo dos tempos, os aspectos sociais da produção vocal têm sido citados na literatura, como um dos fatores que exercem influência na qualidade de voz. Poucos são os autores, contudo, que têm, de fato, seguido este rumo de investigação. Alguns foram comentados antes, como TANNEN & WALLAT, RIBEIRO, MADUREIRA e GROSJEAN. Outros, como SAPIR, LAVER e PITTAM, foram destacados pela crucial importância em fornecer subsídios fundamentais para um tratamento da voz no contexto social.

Nesta parte do trabalho, gostaria de me deter em alguns autores nacionais, particularmente da área de Fonoaudiologia, que apresentaram contribuição

substancial para mudar esse estado de coisas: MARTZ, FERREIRA e SERVILHA. Faço a ressalva de que, ao destacar estes trabalhos, estou “correndo o risco” de deixar de lado, outros estudos que também caminharam nesta direção, porém, assinalo que estas foram as pesquisas que, no meu levantamento da literatura sobre o assunto, se revelaram mais significativas, como subsídio para o tipo de investigação a que me proponho realizar.

Início pela autora citada no início deste item. SERVILHA (2000), em estudo recente, buscou resgatar o papel da voz na interação social e, como ela assinala, ressignificá-la. A autora afirma que, sem a amálgama que deve envolver a voz, a linguagem, a fala e a audição, estes atributos passam a ser vistos de forma fragmentada: a voz sob a perspectiva da *laringe*, a *linguagem* do *cérebro*, a *fala* da *boca* e a *audição* do *ouvido*. Recorte este que subtrai o essencial – o humano, o *sujeito* que se relaciona com o outro e que se constitui nesta relação na e pela linguagem. Uma busca que teve como meta superar a visão positivista da linguagem, e por consequência da voz, visão esta apoiada, tradicionalmente, nos aspectos biológicos. Seu olhar dirigiu-se à voz como atividade discursiva, momento em que há a constituição do ser humano como sujeito dessa voz.

Para essa busca, a autora ancorou-se em pesquisadores que consideram a linguagem como fundante das relações humanas. Fundamentou-se, assim, na perspectiva histórico-cultural de Vygotsky e na dialogia de Bakhtin, os quais lhe possibilitaram compreender a voz na sua “*legítima dimensão*”, nas palavras da autora, qual seja, “*social e interacional*”. Seu trabalho procurou ocupar este espaço negligenciado – a voz numa perspectiva social de análise.

O foco de seu estudo foi analisar *como as variações nas características da voz do professor*, no processo de interlocução da sala de aula, participavam da mediação pedagógica (grifo meu). Os resultados mostraram que é possível associar mudanças na qualidade de voz dos professores a diferentes objetivos que permeiam a relação professor/aluno em sala de aula, como o estabelecimento de acordos e negociações de sentidos:

“Neste espaço, a voz do locutor possibilita a reflexão do interlocutor, pois ele pode se distanciar e dialogar com seu próprio discurso através do outro. O discurso do interlocutor espelhado na voz do locutor permite negociações, interferências, acordos e desacordos e finalmente um re-arranjo de significações. ...É a voz do outro externo a nós, e a voz do outro que criamos dentro de nós (nos moldes do outro externo) que nos ajudam a resolver problemas, refletir o mundo e a nós próprios.”

(SERVILHA, 2000, p. 217-218)

Seu estudo pautou-se por criar categorias de análise que lhe propiciassem compreender o papel das variações vocais na díade professor e aluno. O diálogo e a negociação que envolve a relação professor/aluno são permeados pela voz, não como *“mero instrumento”*, mas pela flexibilidade de suas qualidades, que explicitam ou ocultam sentidos e emoções importantes na construção conjunta da relação ensino-aprendizagem.

Com relação às vozes dos alunos, a autora observa diferentes usos. Eles a usam de *“modo coloquial”*, quando questionam os professores. O aluno que

está fisicamente mais próximo do professor pergunta mais e não precisa elevar muito sua *loudness*. Por outro lado, o aluno realiza “intervenções bruscas por voz e interrupção solicitada por gesto” (grifos da autora), para esclarecimento ou complementação. Em uma das situações estudadas, as vozes dos alunos são utilizadas para tentar coibir o ruído produzido pelas vozes de outros colegas, que prejudicam a viabilização das aulas.

Os sujeitos de seu trabalho foram três professores (denominados de P1, P2 e P3) da primeira série do Curso de Pedagogia de uma Universidade do interior do estado de São Paulo. Apresento, como exemplo, dados da análise de um deles (P1). SERVILHA observa que ele diversifica sua voz com nuances que parecem explicitar repreensão, solicitação, negociação e inclusive, súplica.

Na análise perceptual, a autora constata que, em uma das situações examinadas, P1 inicia utilizando um *pitch* médio, *loudness* fraca, porém audível, velocidade controlada, lentificada e pausada, articulação precisa e entoação sem alterações de *pitch* bruscas. Conforme seu discurso vai se tornando inflamado, a investigadora constata que as características vocais de P1 vão se modificando: a *loudness* torna-se mais forte, o *pitch* agudiza-se um pouco, a articulação fica mais exagerada, a velocidade aumenta e a entoação assume um tom mais incisivo.

Diversas são as particularidades dialógicas que envolvem P1 e seus alunos - o sujeito que enuncia e que, por isso, estabelece relações e cria sentidos por meio da sua voz e do seu discurso. Quando esse sujeito está em um momento mais explicativo para a classe, sua voz apresenta-se mais aguda, forte, rápida, precisa em termos de articulação e mais afirmativa na entoação, resultando em

um discurso predominantemente encadeado, o que significa, para a pesquisadora, que P1 deve falar e o aluno, por sua vez, deve ouvir.

A autora cita outro episódio em que um aluno faz uma *intervenção brusca por voz para esclarecimento* (categoria de análise adotada pela autora, por isso coloco em itálico), e isso obriga P1, a reorganizar suas características vocais: ele diminui a velocidade de fala e articula mais, como forma de dirimir as dúvidas do seu aluno.

Observa-se, por meio deste exemplo, que a interação leva a modificações na voz do sujeito. Em função da sua relação com outro, da negociação e da construção de sentidos, P1 realiza ajustes em sua voz, que representam e manifestam a situação vivenciada, ocorrendo o mesmo com os demais sujeitos estudados (P2 e P3). Nas palavras da autora:

“Cada um tem um modo singular de relacionar-se com os mesmos alunos nos mesmos dias. Este modo de relacionamento tem história e sua gênese nas relações intersubjetivas que foram construídas ao longo da vida de cada um, professores e alunos. Quando em relação, os atores se ajustam, negociam limites e significações e a linguagem/voz é que permite esta construção e colaboração já que permeia as relações e constitui os atores como sujeitos interativos.”

(SERVILHA, 2000, p.211)

A autora indica, como resultado de sua pesquisa, o papel da mediação da linguagem/voz na relação ensino-aprendizagem entre professor e aluno, que

precisam estar, necessariamente, envolvidos em um objetivo comum, em um comungar de algumas questões. Ao mesmo tempo, SERVILHA (Ibidem, pp. 216-217) problematiza se o foco dessa relação, tendo como fio condutor a voz, não criaria *“um esquema mágico no qual a relação dialógica se caracteriza, assume o status de resolutiva de qualquer conflito presente em sala de aula e forma primeira, privilegiada de construir o conhecimento.”*

Desta maneira, as nuances vocais possibilitariam o destaque, o reforço e a ampliação de questões, o oferecimento de turnos e as colocações dos alunos, o que geraria um espaço de mobilidade de conhecimento. Fica claro que, quando SERVILHA discute a voz, está implícito o sujeito dessa voz (grifos da autora) e não apenas, um som específico produzido pelo órgão laríngeo, ou seja, ela mostra que:

“... é a voz do/no sujeito que interfere no outro e que o torna parceiro, já que ambos se interdependem, ambos se relacionam e na escola, esta relação procura desvendar, criar, refletir o conhecimento.”

(Ibidem, p. 221)

Da mesma forma, sua pesquisa procura desvendar, criar e refletir outra perspectiva de estudo da voz – uma abordagem que contemple a história e a singularidade do sujeito, dono da voz. Houve um intuito claro de superar o eixo saúde-doença. Um claro movimento em direção ao espaço negligenciado nas pesquisas em voz. Uma análise do movimento das vozes na interação e seus efeitos de sentido e de impacto sobre o outro. Efeitos de sentidos já anunciados

por FERREIRA (1993), em estudo anterior ao de SERVILHA, um dos pioneiros no cenário brasileiro a ocupar esse espaço tradicionalmente negligenciado. Postura que exige que se (re)pense a natureza do processo lingüístico, uma discussão, contudo, ainda bastante incipiente em nosso meio.

A autora ressalta que, embora a voz seja um produto de fatores biológicos, psicológicos e sociais, estes últimos parecem ser muito pouco focalizados na literatura:

“ Quando os autores fazem referência ao social, ou associam este a manifestações orgânicas e psicológicas, dizendo que suores nas mãos, falta de contato de olho, excesso de gestos são comportamentos que denotam ansiedade e, portanto, inabilidade para manter relação interpessoal (BOONE, 1971), ou reduzem a importância do social a um registro de profissão e tempo de uso vocal (BACKER e col., 1981).”

(FERREIRA, 1993, p. 31)

Para FERREIRA, o que está em jogo é a necessidade de construção de uma nova postura individual e, principalmente, coletiva, frente a essa situação, por meio do reconhecimento de que tanto o sujeito como o fonoaudiólogo estão inseridos num mundo social, o que implica em captar, de modo particular, a inserção deste sujeito na realidade social.

Embora FERREIRA discorra sobre estas questões para analisar a *avaliação* de voz de sujeitos *disfônicos*, acredito que suas reflexões e indagações

podem ser estendidas a *todo* processo de produção vocal, normal o u patológico, e trazem substancial contribuição para uma maior compreensão do sujeito e sua voz, nos aspectos sociais, históricos e culturais. Trata-se, como a autora indica, de se considerar a gênese histórica do sujeito e não simplesmente a evolução dos problemas vocais, de forma isolada:

*“Em suma, dar oportunidade ao sujeito de produzir e compreender a sua própria História, é possibilitar-lhe a interpretação de sua disfonia (na verdade da própria voz, adendo meu) e de seus conflitos inerentes a ela e ao seu mundo social, a partir de sua própria **posição** e dos **papéis** que desempenha na sociedade. Em outras palavras, diríamos, é questionar o que lhe foi dado como inteligível, “**natural**”, ao fazer uso da sua voz, em seu cotidiano.*

*... Desta forma, ambos, paciente e terapeuta (qualquer sujeito, adendo meu) ao **compreenderem o contexto da situação social**, simultaneamente, de um lado explicitam as condições de produção do discurso e conflitos que ele (paciente) produziu; de outro, problematizam estes no processo dialógico de avaliação, relacionando-os com o **social**, isto é, com a História e a cultura de sua sociedade e de seu tempo.*

(Ibidem, pp. 37-38, grifos meus)

Depreende-se do que ela escreve, a importância de um processo conjunto de troca e de interação na produção de sentidos. Frequentemente, o sujeito não tem noção clara de sua inserção social e dos efeitos desta em sua voz, como

comenta SERVILHA (2000), em seu estudo sobre a voz do professor. Ele, comumente, não acredita que seja possível mudar sua voz, desconhece os potenciais e os limites vocais que permitem que se possa fazer um melhor uso da voz.

Como assinala FERREIRA, é como se fosse, *dada* ao sujeito, a voz que tem, algo considerado como *natural* e *invariável*. Questionamento que SAPIR (1927) já apontava em seus estudos. O sujeito desconhece o próprio papel na produção de sentidos e a influência que o seu *posicionamento* na situação interacional pode acarretar em sua voz.

Acompanho FERREIRA (1993, p.38) em suas considerações sobre a avaliação vocal do sujeito, quando afirma que, em última instância, o fonoaudiólogo não deverá chamar a si a responsabilidade de atribuir sentidos ao processo de produção de voz, mas sim, resgatar o “*processo totalizante de significados que estão em movimento, imbricados e ocultos nas relações sociais de produção.*”

NASCIMENTO & FERREIRA (2000) realizaram uma pesquisa sobre voz no contexto não terapêutico, em que afirmam que, como essa inserção é recente, a Fonoaudiologia precisa repensar seus diferentes trabalhos com a voz.

FERREIRA representa uma das contribuições mais importantes, dentro da perspectiva de análise da voz na situação social, em nosso país. A partir de seu artigo “*A Avaliação de Voz: O Sentido Poderia ser Outro*”, publicado em 1993, vem caminhando mais incisivamente nessa direção, como mostra em uma publicação recente (FERREIRA, 2000). Nesta publicação ela pontua aspectos importantes da voz desde a década de 60 até aos dias de hoje, e assinala que o

pouco que se caminhou na direção de como interpretar o que o sujeito social traz para a sua relação com o profissional e de que teoria pode respaldar o profissional na condução de seu barco, para que não fique à deriva, impõe, na verdade, o quanto se tem para avançar.

MARTZ (1990 e 1999) representa outra, entre os poucos autores brasileiros a seguir este percurso. Sua dissertação (1990) configura-se, segundo suas próprias palavras, na composição de uma *escritura dos dizeres sobre a voz* e não propriamente *da* voz. A partir dessa *escritura*, na qual, ao mesmo tempo em que confronta os dizeres sobre a voz, busca também dizê-la, a autora procura redimensionar a voz como “*princípio humano de linguagem*”, a seu ver, ainda basicamente centrado e restrito à esfera da realidade técnica. Um relato de uma conversa sua com um locutor de rádio demonstra a sua ótica sobre voz:

“Lembro-me de certo dia ter conversado com um locutor de rádio e ele então, por me saber fonoaudióloga, contava-me sobre sua voz - que tinha tido um nódulo de cordas vocais porque impostava errado, agravando sua sonoridade natural para atender aos padrões do que entendia como uma voz bonita; que achava importante o trabalho fonoaudiológico para conhecer seu organismo tão bem quanto alguém precisa conhecer uma máquina para trabalhar com ela. Mostrou-me com que precisão sabia imitar várias vozes, suas flexíveis possibilidades quanto à variação de timbres, intensidades; imitou, também, e de forma quase convincente, o som de alguns animais. De seu trabalho ainda me foi possível assisti-lo enquanto gravava aquilo que, nas

rádios, se costuma chamar por anunciar e desanunciar as músicas, ou seja, falar os nomes das músicas e dos intérpretes e autores; essas chamadas foram gravadas para o programa que iria ao ar no período da tarde, quando então o técnico de som tocava alternadamente, as fitas em que estavam gravados os nomes das músicas, os discos relativos a elas, bem como outras fitas contendo diversos e variados comerciais.”

(MARTZ, 1990, p.28)

Dependendo da concepção que se tenha sobre voz, não há, aparentemente, nada de “anormal” ou diferente nesse relato, comum e freqüente entre os fonoaudiólogos que trabalham com voz profissional e que poderia ser considerado **tecnicamente correto** (grifo meu). No entanto, o posicionamento da autora frente à voz a conduz ao questionamento dessas afirmações, conforme segue em seu relato:

*“Impressionada com a **frieza** e mesmo **mecanicidade** com que este trabalho era realizado, perguntei ao locutor se ele não fazia um programa ao vivo, que fosse dele - disse-me que sim, em outra rádio, justamente no horário da tarde, e que lá, ao contrário das gravações que acabava de fazer numa entonação lenta e calma, **sua voz era ágil, sua linguagem mais efusiva, mais rápida**, e, conseqüentemente, com **mais entonação e muito mais emoção.**”*

(MARTZ, 1990, pp. 28–29, grifos meus)

Deste relato, destaco alguns pontos fundamentais. Observa-se, primeiro, grandes variações na voz desse locutor. A voz que esse sujeito utilizava com flexibilidade técnica, nas gravações, e que se caracterizava pela “...frieza e mesmo mecanicidade...”, na qual ele, possivelmente, se “obrigava” a produzir um padrão que não lhe era adequado nem natural - “... impostava errado, agravando sua sonoridade natural para atender aos padrões do que entendia como uma voz bonita ...” (Ibidem, 1990, p.29)

Evidencia-se, aqui, a interferência da situação interacional e a aplicabilidade dos conceitos de GOFFMAN. Quem são os interlocutores deste sujeito? Em que contexto ocorre o seu momento de fala? “**O que está acontecendo aqui e agora?**”. Pergunta chave nos estudos de orientação sociolingüística, nos termos de GOFFMAN.

Outras indagações poderiam ainda, ser feitas: Como essa situação verbal transforma a produção vocal do sujeito a ponto de lhe causar danos? (“... tinha tido um nódulo de cordas vocais ...”). Que **posições** assume esse sujeito de forma a interferir na sua voz dessa maneira?

A outra modificação na voz, que se observa, ocorria quando o sujeito a utilizava em seu programa ao vivo - uma voz mais “*ágil*”, uma linguagem mais “*efusiva*”, com mais “*entonação*” e “*emoção*”. Alinhava-se diferentemente nesta situação, e produzia uma outra voz, mais “*natural*” e *menos nociva* para sua saúde vocal, o que demonstrava suas possibilidades vocais e indicava um caminho, seguramente, melhor a seguir. Verifica-se que a análise feita pela autora decorre da sua concepção de voz - uma forte preocupação com a

influência dos usos sociais da voz. Preocupação que continua sendo foco de sua atenção.

Embora este não fosse o foco da autora, a maneira como trata a questão mostra a utilidade e a aplicabilidade de uma investigação da voz que considere a estrutura organizacional da situação de fala, como proposto por GOFFMAN, isto é, considerá-la a partir das noções de *alinhamento* e de *enquadre*, para a compreensão do fenômeno vocal.

Em artigo, publicado em 1999, MARTZ (op.cit., p.205) aprofunda a reflexão sobre *“quais significados atribuímos à voz e, por extensão, ... os significados que atribuímos ao trabalho fonoaudiológico.”*, no qual levanta algumas questões bastante interessantes, que reiteram a posição que venho defendendo em meu trabalho:

“A queixa de voz está necessariamente comprometida com o uso da voz, ou seja, surge quando é necessário estabelecer uma comunicação com o outro. Surge, portanto, na interação verbal (Bakhtin, 1979) – a voz está implicada com a linguagem oral e é nela que ganha sentido, orgânico inclusive.”

MARTZ (1999, p.206) retoma o fato de que os órgãos envolvidos na produção da voz não tinham, a princípio, essa função no homem, porém,

afirma que a produção vocal, *“tal qual a conhecemos na oralidade, só se tornou possível mediante um desdobramento dessas funções vitais no âmbito da linguagem.”*, afirmando, que se os aspectos orgânicos são fundamentais na produção vocal a linguagem também o é.

A voz, neste estudo, é considerada como aspecto constitutivo da oralidade e conseqüentemente, do sujeito em interação, portanto, considero-a como parte do movimento incessante da linguagem na relação com o outro. Compreender seu funcionamento, seus efeitos no próprio dono da voz e no seu interlocutor assume, assim, grande importância na perspectiva teórica adotada nesta investigação.

Por fim, cabe assinalar que, apesar de haver uma coincidência do ponto de partida deste estudo, com os de FERREIRA, MARTZ e SERVILHA, ou seja, a investigação do aspecto social da voz - sigo um rumo analítico um pouco diferente do das autoras, por estar ancorada em referenciais teóricos distintos, oriundos da Lingüística, em especial da Fonética e da Sociologia.

Segue a descrição do método de investigação, que utilizei nesta pesquisa.

CAPÍTULO III

*“Our aim is to observe how people talk when they are not
being observed.”*

(William LABOV, 1972)

MÉTODO

“Há tempos venho escutando vozes. Todas apesar de singulares, carregam semelhanças nas suas propriedades analisáveis. Deslocando-se num andamento fixado sobre um pulso abstrato, ora marcam um trote contínuo, dado pela cadência de impulsos e repousos de uma fala acelerada, ora estendem-se sobre significativos silêncios de uma fala lenta. A melodia oscila agilmente entre patamares distantes e alterna com naturalidade níveis de energia que vão do mais suave sussurro ao mais forte e áspero grito ”.

(ROCHA FILHO, 1997)

Assim como o autor acima, há tempos venho escutando vozes. Porém, outras vozes, em outros contextos. Vozes de professores, vozes de atendentes de enfermagem, vozes de recepcionistas, vozes de fonoaudiólogos, vozes de atores, vozes de coralistas, vozes de estudantes, vozes de donas de casa, vozes de telefonistas, entre tantas outras vozes, inclusive a minha própria voz. Seus donos participaram de diferentes “*Grupos de Vivência de Voz*” conduzidos por fonoaudiólogos ou por estudantes de Fonoaudiologia com a finalidade de conhecer e aprimorar a própria voz, bem como prevenir alterações vocais. Esses grupos tinham por objetivo a promoção da saúde vocal e constituíam-se, em sua maioria, de sujeitos que apresentavam vozes consideradas *normais*, ou seja, sem a presença de alterações vocais.

ROCHA FILHO (1997), em seu estudo, dirigiu seu foco de atenção a vozes de locutores esportivos de rádio e televisão, na busca de propriedades invariantes na singularidade das vozes, para caracterizar o estilo de narração de futebol, enquanto minha escuta, por sua vez, guiou-se pelas variâncias observadas nas vozes dos participantes dos “*Grupos de Vivência de Voz*”. Em minha “escuta” atenta, um fenômeno deixou-me bastante intrigada – a grande variabilidade da voz de uma mesma pessoa em diferentes contextos/situações. Mudanças de alinhamento dos participantes e/ou dos enquadres da situação de interação revelavam-se na materialidade sonora por meio de alterações no *pitch*, no *loudness*, na dinâmica acentual e/ou nas pausas entre outros aspectos. Pistas fônicas que contribuíam para marcar, por exemplo, um comentário, uma discussão ou o próprio estado emocional do falante.

Minha experiência com Grupos de Vivência de Voz iniciou-se em uma Unidade Básica de Saúde da Secretaria Municipal de Saúde de São Paulo, onde coordenei-os, juntamente com uma outra profissional da área de Fonoaudiologia e, posteriormente, como supervisora de um Curso de Graduação em Fonoaudiologia³⁸, no qual os discentes, alunos do último ano, durante suas atividades de estágio, coordenavam grupos similares, supervisionados por mim e por outros docentes especializados em voz. Além disso, durante a formação acadêmica eram ministradas duas disciplinas, uma de caráter teórico e outra, teórico-prática. Nesta última, um dos objetivos principais era propiciar ao aluno o

³⁸ Refiro-me ao Curso de fonoaudiologia da UNIMEP, Piracicaba, São Paulo.

conhecimento de sua própria voz de forma a poder trabalhar com a voz do outro³⁹.

O fenômeno – ***variação das vozes de um mesmo sujeito***, era observado em todos esses Grupos de Vivência de Voz e também, nos grupos de alunos de Fonoaudiologia, em sala de aula. Vale dizer que frente a essas constatações comecei a observar mais atentamente outras vozes em diferentes situações cotidianas e profissionais e pude verificar que o fenômeno se repetia, o que contribuiu para aumentar o meu interesse em estudar e compreender a materialidade fônica sob este aspecto.

Os Grupos de Vivência de Voz tinham, em média, de 12 a 15 participantes, e eram realizados uma vez por semana, com a duração de uma hora e meia, em um período aproximado de 4 meses. Os Grupos de Vivência de Voz realizados na Unidade Básica de Saúde eram voltados exclusivamente a professores da rede municipal de ensino da área de abrangência dessa Unidade de Saúde, enquanto que os Grupos de Vivência de Voz do Curso de Fonoaudiologia eram abertos à comunidade, para quaisquer interessados em conhecer e aprimorar a própria voz. Nestes, participavam não só, profissionais que dependiam da voz como instrumento de trabalho, mas diversas outras pessoas interessadas, cujas profissões principais, não se relacionavam diretamente ao uso profissional da voz, mas que desenvolviam atividades

³⁹ As disciplinas teóricas e práticas (estágios supervisionados) em Voz Profissional constituem uma inovação curricular desse Curso de Fonoaudiologia, nos níveis de prevenção primária e secundária de atenção à saúde, em relação à maioria dos Curso de Graduação em Fonoaudiologia em nosso país, nos quais a atenção maior tem se voltado à prática clínica. Para maiores esclarecimentos sobre a experiência relativa a esses Grupos de Vivência de Voz e especial, sobre a formação acadêmica nesse campo, ver CHUN & NAVARRO (1998).

paralelas que envolviam o uso profissional da voz, tais como, a participação em corais ou grupos de teatro amador.

Além de breve exposição teórica sobre as bases da produção da voz, os encontros tinham por objetivo principal promover vivências de atividades vocais. Os temas, geralmente, abrangiam os seguintes tópicos: a voz, usos e contextos; qualidade e dinâmica da voz; a voz enquanto expressão; respiração, movimento e voz; ressonância, articulação e voz; e a voz no diálogo, na leitura, na poesia e no canto.

1. CONSTITUIÇÃO DO CORPUS

Para melhor explicitar o método utilizado nesta pesquisa e para tornar mais claras minhas opções metodológicas, retomo aqui algumas questões de ordem teórica que nortearam não apenas a coleta dos dados, mas também a linha de análise adotada.

Considerando-se que na polaridade som-sentido, o **som** é o polo possuidor de materialidade e exterioridade e que a materialidade pode ser decomposta em parâmetros físicos, mensuráveis por meio da análise computadorizada, cabe assinalar, que por essa mesma característica realista e objetiva, o estudo do som tem-se revelado desafiador no que tange em alçar dessa materialidade (física), o **simbólico**. Em meu trabalho a polaridade som-sentido coloca-se a medida que relacionarei manifestações fônicas com os alinhamentos exercidos pelo sujeito no discurso oral em decorrência dos enquadres relativos ao contexto de produção vocal.

O trabalho é norteado por uma abordagem naturalista/observacional discutida por PERRONI (1996), que levanta reflexões importantes e úteis para esta pesquisa ao comparar os métodos experimental, de um lado, e naturalista/observacional, de outro. Afirma que uma das vantagens potenciais nesta segunda abordagem, é que a *"qualidade vem substituir a quantidade, já que não se trata de provocar respostas, mas deixar falar a criança."* (op. cit. p.22, grifos meus), o que coincide com o foco deste estudo, ou seja, deixar **falar o sujeito** (adulto, no caso).

Importa recuperar a história do dado por meio da descrição das suas condições de produção, o que pode permitir a identificação da constituição histórica do sentido. O não controle de variáveis pode possibilitar uma maior aproximação da naturalidade e o reconhecimento de que o que é dado é **o que acontece** e não **o que deveria acontecer** (PERRONI, 1996, grifos meus).

Interessa estudar a língua em atividade, isto é, os dados e suas variações conforme a situação de interação, e não apenas descrever produtos estáticos, o que, também, não significa estabelecer uma polarização radical e estanque entre quantidade e qualidade, como se uma fosse perversão da outra (FERREIRA, 1993). Portanto, desde a coleta dos dados, tive uma preocupação em garantir tais aspectos.

Segundo GAMA-ROSSI (1999) a variabilidade da produção da fala é aferida a partir das transformações dos valores das medidas acústicas do *corpus* estudado em *parâmetros estatísticos*⁴⁰ (nos termos da autora). Daí, se depreende que os valores absolutos extraídos por meio das palavras medidas precisam ser interpretados em termos de significância e, portanto, têm de ser submetidos à análise acústica. Frequentemente, o método utilizado nos estudos acústicos compreende um controle do contexto fonético, a repetição do *corpus* pelos sujeitos e a utilização de frases padronizadas que, se por um lado, "**facilitam**" a comparação dos dados entre si, levariam, por outro lado, a uma "**perda**" qualitativa dos dados e não contemplariam as metas de análise estabelecidas (grifos meus).

⁴⁰ A autora faz um excelente trabalho de revisão da falhas metodológicas da literatura nos estudos sobre a duração na fala infantil.

Uma abordagem naturalista/observacional (desejável teoricamente) apresenta, também, dificuldades. Utilizo BENNETT-KASTOR (1988, citado por PERRONI, 1996) para abordá-las. Segundo ele, nas duas etapas em que um método de trabalho se desdobra - coleta e análise (interpretação), atuam inevitavelmente dois sistemas de filtros como processo seletivo: um, nas **decisões** do pesquisador sobre a seleção e coleta dos dados e, o outro, na **interpretação** (grifos meus).

Como eu optei por um "*filtro não experimental*" (no termos da autora), as limitações ficaram mais evidentes nessa fase da pesquisa, pois não havia um modelo pré-determinado a seguir, (indesejado do ponto de vista teórico). Essa "**falta**" (grifo meu), neste momento, implicou na necessidade do estabelecimento de critérios próprios para seleção das categorias a serem submetidas à análise acústica. Momento de tomada de decisões e de interpretações pessoais, embora fundadas nos autores utilizados. Cabe ressaltar a importância da análise perceptual como ponto norteador dessa seleção. Um intenso movimento de idas e vindas da análise perceptual à observação dos fenômenos fonético-acústicos, e vice-versa, bem como idas e vindas entre os referenciais teóricos e o trabalho de análise, e vice-versa, foi fundamental nessa empreitada, e permeou o desenvolvimento de todo trabalho.

O segundo ponto que norteou a constituição do *corpus*, foi o fato de que dados que vão ser submetidos a análise fonética-acústica requerem qualidade de gravação e cuidados técnicos para evitar distorções do material (LABOV, 1972-8 e PITTAM, 1994), uma das minhas primeiras preocupações do ponto de vista metodológico. Assim, a impossibilidade técnica de gravação dos participantes dos

Grupos de Voz, *in loco*, em condições técnicas favoráveis para esse tipo de análise, conduziu-me à opção pela gravação dos sujeitos em ambiente acusticamente tratado, e não, nas salas em que normalmente ocorriam os grupos, que eram acústica e tecnicamente mais desfavoráveis, tendo em vista garantir uma melhor qualidade técnica do material coletado. Por se tratar de uma análise de caráter **qualitativo** e não quantitativo, considerou-se que um número reduzido de sujeitos poderia ser utilizado, número este que foi fixado em **dois** (2), embora o número de falantes gravados tenha sido maior, sete (7).

Vale destacar que, a meu ver, o fato da coleta de dados ter sido realizada num estúdio de gravação não lhe conferiu um caráter experimental, pois não houve qualquer preocupação no sentido de controle de variáveis como, por exemplo, o conteúdo do que seria dito pelos participantes, o tempo de fala de cada um ou as próprias condições de produção, embora fosse claro que essa situação (gravação) poderia interferir nos dados obtidos.

Por outro lado, entre outros pontos, era importante que o material pudesse ser coletado em situações que favorecessem maior variabilidade da produção vocal e, por isso, os sujeitos foram inseridos dentro de grupos maiores.

Essa inserção teve, também, um propósito que remete ao que LABOV (1972-1978: 6) considerou como o *paradoxo do observador*: "...Our aim is to observe how people talk when they are not being observed.". O autor considera esta situação paradoxal como uma questão que não pode ser completamente solucionada — quando se grava um grupo em interação, limita-se a qualidade e a quantidade dos dados, e se, por outro lado, apenas se observa esse grupo, não haverá um registro gravado dos dados. Em razão disso, tanto na coleta de dados,

quanto na posterior seleção do material para constituição do corpus, procurei tomar alguns cuidados que minimizassem os efeitos da situação de gravação. Reafirmo, contudo, que não houve controle algum sobre a variável principal da pesquisa, isto é, a produção vocal dos falantes.

Os componentes dos grupos (de gravação), entre os quais, posteriormente, retiraria os sujeitos da pesquisa, foram escolhidos aleatoriamente entre aqueles que participavam dos *Grupos de Vivência de Voz*. Por se tratar de um trabalho realizado na Clínica-Escola do Curso de Fonoaudiologia, todos os participantes, no início desses *Grupos de Vivência de Voz*, caso houvesse concordância, autorizavam por escrito a realização de fotos, ou de gravações em áudio ou em vídeo, para fins didáticos e/ou de pesquisa. Desta forma, havia uma autorização prévia para registros gravados, bem como haviam sido feitas gravações em áudio e/ou vídeo nos *Grupos de Vivência de Voz*, para acompanhamento dos participantes no decorrer do semestre. A situação de gravação, portanto, não era um fato novo para essas pessoas, o que se constituiu em um elemento minimizador dos efeitos do paradoxo apontado por LABOV.

Os participantes foram convidados a participar de uma gravação em estúdio, com a proposta de se obter um registro vocal de melhor qualidade técnica, como se fosse uma continuidade do trabalho realizado nos *Grupos de Vivência de Voz*. A sugestão foi gravar situações vivenciadas⁴¹ durante o semestre, isto é, atividades vocais desenvolvidas nos *Grupos de Vivência de Voz*,

⁴¹ Como os objetivos principais dos *Grupos de Vivência de Voz* eram a promoção da Saúde Vocal, a prevenção de alterações de voz e a promoção do conhecimento da produção e usos da voz dos sujeitos, as atividades vivenciadas foram, por exemplo, a apresentação e discussão de vídeos sobre saúde vocal e atividades de produção de voz como, aquecimento e desaquecimento vocais, canto com variações de ritmo, leitura em voz alta, dramatização da voz no uso profissional ou de textos curtos, diálogo entre personagens com vozes distintas, entre diversas outras.

buscando-se mais uma vez, minorar os efeitos que uma situação nova poderia provocar.

A função dos discentes era a de coordenação desses grupos de gravação, a mesma da situação acadêmica, ou seja, nos estágios os alunos eram responsáveis, geralmente em duplas ou trios, pela condução dos *Grupos de Vivência de Voz* sob supervisão de docentes especializados na área de voz. Desta forma, todos ocupariam os mesmos papéis durante a situação de gravação. Além disso, essa coleta foi feita no final do semestre e, portanto, os participantes, se conheciam, e também, haviam realizado atividades vocais similares, conforme descrito anteriormente, às propostas para a gravação. Outro cuidado, que tomei, em relação à composição dos grupos de gravação, foi o de convidar duplas de estagiárias acostumadas a atuar juntas, bem como integrantes dos *Grupos de Vivência de Voz*, que fossem do mesmo grupo coordenado por elas, de forma a garantir que os participantes se entrosassem e favorecer que eles ficassem o mais à vontade possível, frente à situação de gravação.

Assim, foram constituídos 3 grupos de gravação.

O primeiro grupo (Grupo 1) compunha-se de 3 participantes, um ator amador e 2 discentes, estagiárias do último ano desse Curso de Fonoaudiologia. Os três pertenciam ao mesmo *Grupo de Vivência de Voz* e, portanto, o contexto do grupo de gravação era semelhante à experiência vivenciada por eles em seu cotidiano.

O segundo grupo (Grupo 2) foi constituído de forma similar ao primeiro, com 2 outros discentes, também estagiários do mesmo curso, e um terceiro participante, uma telefonista, componente do Grupo de Vivência de Voz, porém,

apenas de uma das estagiárias desse grupo. Pouco antes do início das gravações, a outra aluna convidada, por motivos particulares, não pôde comparecer, e pediu que uma colega a substituísse. Essa mudança levou à constituição de um grupo de gravação, diferente do habitual dos estágios. Embora o enquadre previsto, à semelhança do Grupo 1, fosse o de reprodução do *Grupo de Vivência de Voz*, isso não ocorreu, pois, nem todos os componentes deste grupo (Grupo 2) se conheciam.

O terceiro grupo (Grupo 3) foi composto por 5 participantes, os 4 discentes dos grupos anteriores, e uma professora desse Curso de Graduação em Fonoaudiologia, coincidentemente, participante de um dos *Grupos de Vivência de Voz*.

No grupo de gravação, no entanto, seria a docente, quem teria a função de coordená-lo. A proposição de coordenação foi feita apenas a ela, pouco antes de sua chegada ao estúdio, o que a levou a acreditar que as vozes a serem analisadas seriam as dos alunos e as dos participantes dos *Grupos de Vivência de Voz*, por eles terem participado dos outros grupos de gravação.

Os discentes, por sua vez, acreditavam que, como essa professora havia sido integrante de um dos *Grupos de Vivência de Voz*, ela participaria do grupo de gravação, na mesma condição que a dos outros componentes - o ator amador e a telefonista, ou seja, no papel de participante de um *Grupo de Vivência de Voz*. Havia, portanto, uma crença de que o contexto (enquadre) de gravação seria o mesmo dos demais grupos de gravação (Grupos 1 e 2) – uma reprodução dos *Grupos de Vivência de Voz*, e portanto, não sabiam que, na verdade, a professora iria assumir o papel exercido na vida institucional.

Desta forma, desconhecendo a proposta deste terceiro grupo de gravação, os alunos foram **surpreendidos** (grifo meu) com a nova proposição, ou seja, de mudança de papéis e de contexto. Cabe destacar, porém, que a proposta foi aceita **naturalmente**. Não houve, por parte de nenhum dos presentes no estúdio, qualquer demonstração de “constrangimento”, nem de “rejeição” diante do proposto. A professora foi bem recebida pelos alunos, que se **alinham** como tais, assim que ela chegou, e vale dizer, que ela também, se **alinhou como docente**, desde a sua entrada no estúdio. Houve, portanto, uma modificação dos enquadres e dos alinhamentos de **forma natural**, frente à nova situação proposta, estabelecendo-se assim, um contexto distinto dos Grupos de gravação 1 e 2, o que concorreu para variações espontâneas nas vozes dos presentes.

Considero que não houve qualquer manifestação de “estranhamento”, pois, o contexto proposto coincidia com a experiência institucional e cotidiana de cada um dos presentes, havia, portanto, um conhecimento prévio e mútuo, que favoreceu que a situação transcorresse sem quaisquer dificuldades. Houve um movimento, dos participantes e um re-arranjo natural ao jogo interacional, nos termos de GOFFMAN

Após o término das gravações, todos os participantes foram informados sobre os objetivos deste estudo, e foi-lhes então, solicitada uma autorização para a utilização do material gravado, nesta pesquisa. Todos concordaram.

1. 1. Delimitação do Corpus

Uma vez coletado o material, procedeu-se a delimitação do *corpus*. Como havia sido anteriormente planejado, foram selecionadas para este estudo as falas

de dois sujeitos, a de uma discente do grupo 1, doravante denominado de **Sujeito 1** e a fala da professora do grupo 3, doravante **Sujeito 2**. Diversas razões conduziram-me para essa escolha, conforme esclareço a seguir.

Inicialmente, o meu objetivo era analisar as vozes dos participantes dos *Grupos de Vivência de Voz*, isto é, a do ator amador e a da telefonista. Entretanto, como eles tinham a expectativa de que suas vozes seriam avaliadas, por já serem participantes dos *Grupos de Vivência de Voz*, demonstraram estar “preocupados” com a produção de suas vozes e assim, não selecionei suas produções vocais para análise. Optei por mudar o foco de análise para a fala de sujeitos, que não esperavam que suas vozes fossem estudadas, buscando assim, analisar produções de vozes mais “**espontâneas**”, ainda, que se considere a interferência da própria situação de gravação, um paradoxo de difícil solução, como discutiu LABOV.

As condições de constituição do Grupo 2 de gravação não se efetuaram da forma prevista, em função da ausência de uma das estagiárias da dupla convidada, o que gerou um certo “constrangimento” por parte do grupo, pois, eles não estavam acostumados a trabalhar juntos. Esse imprevisto interferiu na produção vocal dos seus integrantes e os deixou pouco à vontade nessa situação. Diante dessas considerações, os participantes desse grupo foram excluídos da pesquisa.

Frente às propostas iniciais para cada grupo de gravação, os sujeitos escolhidos para a constituição final do *corpus* acreditavam que as vozes dos outros interlocutores é que seriam estudadas. O Sujeito 1 desempenhava o papel de coordenador do Grupo 1 e, como julgava que o integrante do grupo de voz é

que seria analisado, ***não se preocupou*** em como produzir sua voz. O fato de este sujeito não se considerar como foco de estudo, favoreceu uma maior ***“espontaneidade”*** de sua parte na situação de gravação.

O Sujeito 2 tinha conhecimento de que os discentes haviam participado das gravações dos grupos anteriores e julgava que as demais vozes é que seriam analisadas. Esta pressuposição contribuiu, para sua maior ***“naturalidade”*** na situação de gravação.

Como interessava que os sujeitos a serem estudados apresentassem variações espontâneas em suas vozes, de acordo com o contexto de produção, os fatos acima expostos foram determinantes para a seleção desses dois sujeitos.

2. PROCEDIMENTOS DA COLETA E SELEÇÃO DOS DADOS

Os sujeitos foram gravados no Estúdio de Audio dos Laboratórios de Recursos Audiovisuais da Faculdade de Comunicação Social da mesma Universidade do Curso de Fonoaudiologia. Ficaram numa sala acusticamente tratada, sentados em torno de uma mesa redonda, com 3 microfones individuais. Eu, como pesquisadora, fiquei numa sala contígua junto com o técnico de som e acompanhei a gravação externamente, através de um painel de vidro.

No início dos *Grupos de Vivência de Voz*, eram feitas gravações em audio dos componentes, quando lhes era explicado que não deveriam considerar a gravação como uma situação de teste ou de avaliação, mas sim como um registro de suas vozes num dado momento. Desta forma, sugeria-se que procurassem ficar o mais à vontade possível durante a gravação. Antes de serem gravadas, as atividades propostas eram apresentadas e realizadas com os participantes dos grupos, buscando-se, assim, obter um registro um pouco mais “espontâneo” das pessoas. Essas orientações foram retomadas durante a coleta dos dados com a finalidade de minimizar os efeitos da situação de gravação e sua interferência na produção vocal.

Após conhecerem a sala em que seriam gravados, os participantes receberam a sugestão de conversarem livremente antes do início das gravações. Foi entregue uma proposta de atividades aos coordenadores dos grupos, à semelhança do que ocorria nos *Grupos de Vivência de Voz*, nos quais havia sempre um planejamento prévio das atividades, que eram discutidas com os

alunos durante as supervisões desse estágio, tendo em vista contemplar as necessidades/expectativas dos componentes dos *Grupos de Vivência de Voz*.

Foram propostas quatro situações, vivenciadas nos *Grupos de Vivência de Voz*, a saber: conversação livre, discussão a partir da leitura de um texto, dramatização de uma situação de trabalho e comentários sobre a voz nessas diferentes atividades.

A proposição de conversarem antes de gravar e as demais propostas foram feitas levando-se em conta, também, as considerações de LABOV (1972-1978) como: incentivar a conversação entre os componentes do grupo e registrar uma conversação que não fosse dirigida ao entrevistador (pesquisador), escolher tópicos de interesse dos falantes e eliciar narrativas de experiência pessoal de forma a poder alcançar as metas estabelecidas.

A partir da seleção dos sujeitos, realizei a transcrição ortográfica do material gravado em fita cassete dos Grupos 1 e 3 (Anexos 1 e 2). Em seguida, selecionei os trechos de fala a partir de análise perceptual. O critério principal foi a escolha de enunciados que ocorressem em contextos diferentes para análise da produção vocal, uma vez que a minha hipótese é que a situação de interação transforma a voz. Posteriormente, procedi a análise acústica. Nesta, foram realizadas medidas que na literatura fonética têm sido consideradas apropriadas para descrição da qualidade e dinâmica de voz, como a medida da frequência fundamental (Fo) e a duração do segmento de fala. O próprio processo de análise determinou a seleção de outros parâmetros que serão explicados mais detalhadamente à frente. Antes, faço a descrição dos recursos materiais utilizados.

3. RECURSOS MATERIAIS

O material foi gravado em Deck recorde (k-7) – Tascam ZOZ MK-III, Console (mixer) – Spirit studio e Microfones - Beta Shure (Beta 58) e Neumann V 87. A fita cassete utilizada foi *Basf Crhome Extra* – 60 minutos, por indicação do técnico de gravação do Estúdio da Universidade.

Os textos escolhidos pelos sujeitos foram: no grupo 1, matéria do jornal Folha de São Paulo sobre a luta de boxe entre Evander Holyfield e Mike Tyson (30 de junho de 1997, p. 6, Caderno de Esportes) e no grupo 3, um livro sobre Saúde Pública - VIEIRA, R.M. et al. (orgs.) (1995) - **Fonoaudiologia e Saúde Pública**, Carapicuíba: Prófono, p.27, assunto de especialidade do sujeito/docente.

O equipamento utilizado para análise acústica foi o software *CSRE 4.5*. (*Computerized Speech Research Environment*, AVAAZ Innovations Inc., 1995), instalado no Laboratório Integrado de Análise Acústica da Faculdade de Comunicação e Filosofia da Pontifícia Universidade Católica de São Paulo, composto de: *CSRE 4.5*, instalado num computador 486 e respectivos componentes, provido de placa de som *Sound Blaster 16*. Os dados foram digitalizados a partir de gravador analógico cassete, conectado ao computador, utilizando-se as gravações nas fitas K-7 BASF. Os dados digitalizados no computador foram utilizados para estudo dos diferentes parâmetros acústicos por meio do *CSRE 4.5*.

4. PROCEDIMENTOS DE ANÁLISE DOS DADOS

Foram selecionados cinco (5) trechos de fala para análise: três (3) do Sujeito 1, a aluna L., e dois (2) do Sujeito 2, a professora , conforme o contexto em que ocorreram. Em seguida os trechos foram analisados de acordo com:

- (a) O contexto em o que o trecho de fala foi produzido;
- (b) A variação da produção vocal percebida auditivamente, que norteou os parâmetros físicos e acústicos selecionados para análise e
- (c) Os parâmetros físicos e acústicos, buscando-se correlacioná-los com a análise perceptual e o contexto de produção do enunciado.

A **análise do contexto** em que os enunciados ocorreram, possibilitou categorizá-los inicialmente, da seguinte forma:

SUJEITO 1

- 1. Enunciado 1 - *coordenadora*.
- 3. Enunciado 2 - *personagem* (cliente)
- 4. Enunciado 3 - *aluna*

SUJEITO 2

- 5. Enunciado 4 – *professora*
- 6. Enunciado 5 - *mãe*

Na análise perceptual⁴² foram estudados os seguintes aspectos: proeminências, pausas e velocidade de fala. Como uma análise auditiva pressupõe que se escute o material atenta e repetidas vezes para maior e melhor percepção das nuances dos eventos fonéticos, utilizei o *CSRE 4.5*, nesta parte da análise, apenas como um recurso auxiliar e facilitador dessa escuta, e portanto, do estudo perceptual dos enunciados.

Esse *software* apresenta a possibilidade de marcar trechos dos enunciados, que podem ser visualizados na tela do computador e repetidos mais prontamente, e portanto, propicia um procedimento mais preciso do que a reprodução de áudio em fita cassete para a análise auditiva.

As palavras avaliadas como mais proeminentes foram anotadas e as pausas demarcadas após ouvir os enunciados diversas vezes para confirmação dos dados.

Para avaliar a velocidade de fala empregada, medi a duração total dos enunciados, marcando o início e o final da onda no espectograma. Obtive assim, o valor em milissegundos. Os resultados foram transformados em segundos para determinação da velocidade de fala.

Além disso, realizei a contagem do número total de sílabas. Estas foram contadas a partir das unidades silábicas produzidas oralmente pelo sujeito, como por exemplo em “... *depois um* ...”, o sujeito produziu “**de- poi- sum**”. O número

⁴² Vale lembrar, conforme expliquei na introdução do trabalho, que utilizo a expressão *análise perceptual*, na forma proposta por LAYER, ou seja, utilizo análise perceptual, em substituição a denominada análise perceptiva auditiva.

total resultante foi de 92 sílabas. Segue, a título de ilustração (vide QUADRO 2)⁴³, a marcação das sílabas do Enunciado 1.

QUADRO 2 - Contagem de sílabas do Enunciado 1

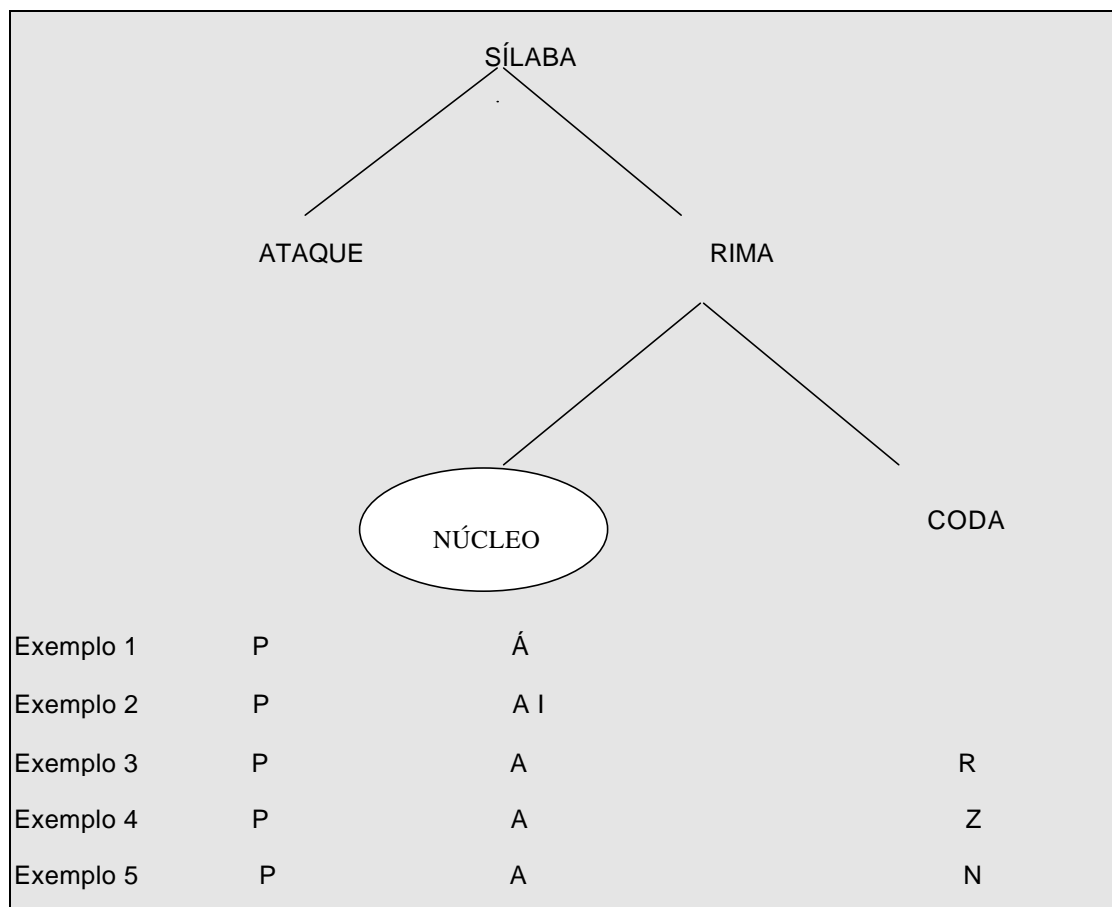
EN	TÃO	A	SSIM	A	GO	RA	UM	POU	CO	XX	XX	A GEN	TE
1	2	3	4	5	6	7	8	9	10	11	12	13	14
VAI A	SSIM	TEN	TAR	RE	PRO	DU	ZIR	NÉ	NAO	SEI	SE	A	IN
15	16	17	18	19	20	21	22	23	24	25	26	27	28
DA	QUÊ	CO	MEN	TAR	MAIS	AL	GU	MA	COI	SA	DA	DA	LU
29	30	31	32	33	34	35	36	37	38	39	40	41	42
TA	UM	POU	CO	DA	SI	TU	A	ÇÃO	DE	TRA	BA	LHO	SUA
43	44	45	46	47	48	49	50	51	52	53	54	55	56
MES	MO	NÉ	OU	NA	PRE	FEI	TU	RA	E	DE	POI	SUM	POU
57	58	59	60	61	62	63	64	65	66	67	68	69	70
QUI	NHO	DO	TE	A	TRO	NÉ	PRA	CE	TEN	TAR	DRA	MA	TI
71	72	73	74	75	76	77	78	79	80	81	82	83	84
ZAR	PRA	GEN	TE	UM	POU	CO	NÉ						
85	86	87	88	89	90	91	92						

Para avaliar as proeminências, investiguei os seguintes parâmetros acústicos: variação frequência fundamental (Fo), a duração e o valor de Fo (frequência fundamental) mais alto do núcleo das sílabas tônicas das palavras enfatizadas e a variação de Fo (frequência fundamental) no núcleo vocálico das tônicas dos *pitch accents*. Os parâmetros para análise acústica foram selecionados após repetida análise perceptual, que me possibilitou identificar os recursos/aspectos que marcavam as variações vocais dos sujeitos nos trechos de fala estudados.

Considere a estrutura silábica segundo MIRA MATEUS (1996) e LAVER (1994), conforme demonstro no QUADRO 3.

⁴³ As sílabas 11 e 12 foram omitidas no quadro, para preservar a identidade do participante.

QUADRO 3 - Esquema de Estrutura Silábica



Em relação à variação de *pitch*, considerei como unidade de análise cada trecho de fala entre as pausas nos enunciados. Em cada unidade prosódica de cada enunciado foram medidos três valores de Fo: inicial, máximo e final. Foi utilizado o extrator de *pitch* do CSRE 4.5. Os parâmetros empregados para obtenção desses valores, na maioria dos casos, encontram-se discriminados no QUADRO abaixo:

QUADRO 4 - Parâmetros para extração do Fo

PARÂMETROS	ALGORÍTIMOS E VALORES
Procedure (método)	Cepstrum
Wind (janela)	256
Overlap (percentagem de sobreposição)	60%
Sample Frequency (taxa de amostragem)	12,5 kHz
High Frequency (limite de frequência mais alta)	400 Hz ou 500 Hz
Silence Threshold (limiar mínimo do sinal)	5 dB
Zero Crossing Threshold	1500 Hz

O *Cesptal* é um método de análise da frequência fundamental desenvolvido desde os anos 60. Os harmônicos (múltiplos da frequência fundamental) se repetem num intervalo regular, que é a própria *freqüência fundamental*. Essa análise *cepstral* é um meio de extrair esses valores de forma precisa e automaticamente. *Cepstrum* corresponde a *spectrum* com a primeira sílaba lida de trás para frente (KENT & READ, 1992).

A janela *cepstral* (*wind*) tem que durar o suficiente para conter pelo menos dois pulsos de *pitch* no sinal acústico. O comprimento da janela é controlado por este parâmetro. Para mulheres e crianças, cujas frequências são mais altas, é preferível a utilização de janela com valores mais baixos como a que foi utilizada (AVAAZ, 1997).

O valor de sobreposição que determina a sucessão das janelas é recomendado para processar mais suavemente as mudanças sucessivas de janela no decorrer do tempo. Um *overlap* de 30 %, conduz à mudanças mais rápidas. (AVAAZ, 1997) .

Sampling é a operação pela qual o sinal analógico é convertido em digital, denominado de *samples* (amostras) (KENT & READ, 1992). *Sample Frequency* corresponde à taxa de amostragem do sinal acústico.

O valor de *High Frequency* estabelece o limite de frequência mais alta. O valor mínimo é fixado em 75 Hz e não é passível de ser modificado neste programa. Considera-se que o valor máximo de 400 Hz (0,40 kHz) corresponderia a uma faixa de extensão adequada para vozes masculinas e que abrangeria a maior parte das vozes femininas e infantis (AVAAZ, 1997). Somente um dos valores extraídos (Sujeito 1, Enunciado 1) nesta pesquisa ultrapassou 400 Hz e, assim, quando os valores marcados automaticamente pelo programa provocavam qualquer dúvida, esse parâmetro era modificado para 500 Hz, o que possibilitava a visualização de valores de F_0 mais altos no espectograma.

Silence Threshold tem a finalidade de controlar o limiar mínimo do sinal. Se um sinal tem um amplitude muito baixa dificulta a distinção entre um ruído de fundo e o silêncio. Ajustando esse valor é possível estabelecer um critério do sinal que será considerado como silêncio (AVAAZ, 1997). Neste caso, segui o padrão estabelecido pelo *software* de 5,0 dB.

Zero Crossing Threshold é um parâmetro que permite que no *CEPSTRUM* sejam analisados somente valores de *pitch* abaixo de 1500 Hz - valor válido para todos falantes (AVAAZ, 1997).

Os valores utilizados em cada parâmetro mostraram-se adequados para a extração dos valores de F_0 . Quando havia alguma dúvida em relação as medidas por meio deste procedimento - *CEPSTRUM*, utilizei o *COMB*, um método alternativo do programa *CSRE* para análise do *pitch*. O sinal de fala neste outro método é dividido em blocos e filtrado de forma a oferecer a possibilidade de que as estimativas dos valores de *pitch* sejam examinadas individualmente, de forma mais detalhada, e por meio da edição da marcação dos valores de F_0 , conforme o formato da onda (repetição dos ciclos). Os pontos marcados automaticamente poderiam ser desmarcados e substituídos por outros assinalados por mim.

O número total de valores de F_0 (frequência fundamental) extraídos nos cinco enunciados analisados foi 255. Cada um dos enunciados foi dividido em unidades prosódicas (delimitadas a partir da ocorrência de pausas) e em cada unidade prosódica foram extraídos três valores de F_0 (inicial, máximo e final). O número de unidades prosódicas variou de acordo com o enunciado. Foram identificadas 12 unidades prosódicas no Enunciado 1, 09 no Enunciado 2, 19 no Enunciado 3, 27 no Enunciado 4 e 17 no Enunciado 5. Esses valores foram anotados numa tabela conforme exemplo abaixo, em que duas unidades prosódicas são apresentadas, e depois foi construída uma tabela no Excel com todos os valores, para posterior análise estatística da faixa de variação de F_0 dos enunciados.

TABELA 1 – Exemplo dos valores de *Fo* de algumas unidades prosódicas do Enunciado 1

UNIDADE PROSÓDICA	Fo INICIAL		Fo MÁXIMO		Fo FINAL	
1 – <i>Então, assim</i>	Vocábulo	<i>então</i>	Vocábulo	<i>assim</i>	Vocábulo	<i>assim</i>
	Fo	284 Hz	Fo	377 Hz	Fo	196 Hz
	Tempo	129 ms	Tempo	459 ms	Tempo	730ms
2 – <i>agora um pouco</i>	Vocábulo	<i>agora</i>	Vocábulo	<i>agora</i>	Vocábulo	<i>pouco</i>
	Fo	260 Hz	Fo	403,2 Hz	Fo	235,9 Hz
	Tempo	1429 ms	Tempo	1675 ms	Tempo	2621ms

Depois, procedi ao estudo das palavras enfatizadas marcando a duração (ms) e o valor do *Fo* mais alto do núcleo das sílabas tônicas. A duração foi extraída no *CSRE 4.5* na tela da análise espectral do enunciado. Essa tela de análise apresenta o espectograma e a forma da onda sincronizados, além de mostrar em zoom a porção da onda correspondente ao local onde se posiciona o cursor no sinal acústico. Mostra, também, o envelope espectral.

Dessa maneira, é possível que as variações de sinal sejam identificadas na forma da onda e no espectograma, ao mesmo tempo, possibilitando uma melhor delimitação do início e do término do núcleo vocálico da sílaba tônica e, portanto, um valor mais preciso da duração do segmento.

Para FANT (1973, versão para o português de GAMA-ROSSI, 1999, p.7) as fronteiras dos segmentos acústicos (*sounds segments*) não deveriam ser confundidas com as fronteiras de fonemas. De acordo com esse autor, a análise espectrográfica da fala mostra fronteiras distintas entre as partes sucessivas no eixo do tempo e variam em função da fonte primária do som e segundo o tipo de

abertura ou fechamento das cavidades. Em consequência, o número de segmentos acústicos num enunciado pode ser maior que o número de fonemas. O autor exemplifica com o caso de uma oclusiva surda, que apresenta pelo menos dois segmentos acústicos - a oclusão e o estouro, e seria uma questão de convenção estipular se o segmento acústico é atribuído à vogal ou à consoante precedente. Um outro aspecto importante, apontado pelo autor, é a relação entre segmentos acústicos e traços, como no caso da nasalização.

Tendo em vista tais considerações, esclareço que, nas medidas de duração das sílabas tônicas das palavras enfatizadas, considerei como segmento acústico de análise o *núcleo vocálico*, isto é, somente a própria vogal, pois, um monotongo e um ditongo (ver os exemplos 1 e 2 do Quadro 3 - Estrutura silábica) não poderiam ser considerados da mesma forma quanto à duração. Além disso, nas fronteiras dos segmentos acústicos, desconsiderei partes dos segmentos que poderiam ser atribuídos às consoantes que antecederiam ou sucediam o núcleo silábico como, por exemplo, o vozeamento, os estouros e/ou traços de nasalidade.

O valor da frequência fundamental (F_0 mais alto) foi extraído segundo o mesmo método utilizado para medida dos valores da variação de *pitch* das unidades prosódicas, ou seja, utilizei o extrator de *pitch* do CSRE 4.5 pelo método *CEPSTRUM*, na maioria dos casos. Quando os valores eram duvidosos, novamente a marcação das frequências fundamentais era realizada por mim, por meio do *COMB*.

Por último, foram analisados os *pitch accents* dos enunciados para verificação da variação tonal utilizada pelos sujeitos. Realizei as medidas de F_0

considerando o núcleo silábico das tônicas das palavras enfatizadas, extraindo valores de F_0 inicial, medial, final, máximo e mínimo. O valor da variação tonal do *pitch accent* foi obtido pela diferença entre os valores máximos e mínimos de F_0 . Os procedimentos para a delimitação das fronteiras dos segmentos acústicos desses núcleos foram os mesmos das demais medições.

Os dados, obtidos em termos de valores, foram submetidos a estudo estatístico, com a aplicação do Teste de Comparação de médias para amostras independentes. Para o estudo do conjunto de todas as variáveis dos enunciados foi utilizada a técnica multidimensional de Análise Fatorial dos Componentes Principais, para identificação dos fatores, estatisticamente mais significativos, que permitiram distinguir um enunciado do outro.

Os resultados são apresentados no capítulo que se segue, no qual serão discutidos, juntamente com a análise contextual e as demais categorias de análise, à luz de referenciais teóricos que enfocam uma perspectiva social, particularmente aqueles postulados em GOFFMAN.

CAPÍTULO IV

“When we use voice to communicate, we are engaging in social interaction. To say that voice is an important part of our communicative behavior, however does not adequately describe the way we use it.

... This is to say that vocal function in social interaction is dynamic and complex, requiring constant monitoring and shifts in usage and interpretation on the part of all participants.”

(Jeffery PITTAM, 1994)

RESULTADOS E DISCUSSÃO

Apresento aqui os resultados, dividindo-os nas categorias de análise que assinalei no capítulo anterior, a saber: a) **caracterização do contexto** em que o trecho de fala analisado é produzido; b) dados resultantes da **análise perceptual** e c) resultados da **análise de parâmetros físico-acústicos**.

A apresentação é feita por sujeitos, Sujeito 1 e Sujeito 2, e por trechos de fala, denominados, respectivamente, Enunciados 1, 2, 3, 4 e 5. Os três primeiros pertencem ao Sujeito 1 e os dois últimos ao Sujeito 2. Em seguida, faço uma comparação dos resultados dos enunciados e a discussão dos mesmos.

Embora os sujeitos tenham sido gravados em quatro atividades dialógicas diferentes, a saber — conversação livre, discussão a partir da leitura de um texto, dramatização de uma situação de trabalho e comentários sobre a voz nessas atividades, os trechos de fala escolhidos foram extraídos somente das situações de **discussão de texto** e **dramatização de uma situação de trabalho**, por corresponderem a momentos em que os sujeitos não tinham a atenção voltada para a própria voz, mas ao tópicos do discurso.

Nenhum dos enunciados dos períodos de conversação livre e de comentários sobre a própria voz foi selecionado, por opção metodológica. A primeira situação, conversação livre, era um período introdutório, de aquecimento, na qual os sujeitos, possivelmente, encontravam-se, ainda, pouco à vontade para falar e muito atentos à própria voz, em função da situação de gravação. A situação de comentários, por sua vez, além de corresponder ao final da gravação,

era uma situação em que os sujeitos estavam mais atentos às suas vozes por as estarem comentando.

1. ANÁLISE DOS ENUNCIADOS DO SUJEITO 1

1.1. Análise do ENUNCIADO 1

L - ENTÃO, ASSIM, // AGORA UM POUCO //..., J., // A GENTE VAI, ASSIM, // TENTAR **REPRODUZIR**, // NÉ? // NÃO SEI SE AINDA QUÉ COMENTAR MAIS ALGUMA COISA DA, // DA LUTA, UM POUCO DA SITUAÇÃO DE **TRABALHO**//, SUA MESMO, // NÉ? OU NA **PREFEITURA**, // E DEPOIS UM POUQUINHO DO **TEATRO**, NÉ? // PRA CÊ TENTAR DRAMATIZAR PRA GENTE UM POUCO //, NÉ? //

Legenda:

// - pausa

negrito – palavra enfatizada

sublinhado – *pitch accent*

a) Caracterização do Contexto

Este trecho foi extraído do Grupo1 (Anexo 1, p. 4), no qual o Sujeito 1 (L.), em conjunto com outra estagiária de Fonoaudiologia (M.), é responsável pela coordenação das atividades desenvolvidas no grupo. Há a participação de apenas um integrante (J.) do *Grupo de Vivência de Voz* - ator amador e funcionário público, que, durante o estágio, havia sido coordenado pela mesma dupla de alunas, conforme expliquei no capítulo anterior. Reitero que, há, portanto, um conhecimento prévio entre os interlocutores e, também, a reprodução de um

contexto vivenciado por eles no cotidiano. Inicialmente, o grupo conversa livremente Voz (vide Anexo 1, Grupo 1, turnos de fala iniciais).

L. e M. resgatam as atividades do *Grupo de Vivência Voz*, lembrando J., a facilidade que ele tinha em memorizar e reproduzir os textos de suas peças de teatro durante as atividades desenvolvidas nos *Grupos de Vivência de Voz*. Observa-se que L. e M. procuram incentivar a produção oral de J., retomando aspectos de experiências, anteriormente, vividas por eles.

O trecho de fala selecionado para análise corresponde ao momento em que o Sujeito 1 (L.) propõe a J. a dramatização de uma situação de sua vida profissional, como ator ou como funcionário público. Nos termos de GOFFMAN, o contexto do Sujeito, neste enunciado, pode ser brevemente caracterizado, da seguinte forma:

Interlocutores: participante dos *Grupos de Voz* (J.), duas estagiárias de Fonoaudiologia (L. e M.)

Enquadre: orientação de atividade de voz

Alinhamento do sujeito: coordenadora

b) Dados da Análise Perceptual

Pela análise auditiva desta situação observei:

- Ênfase em poucas palavras, considerando-se a extensão do trecho de fala. O sujeito enfatizou apenas quatro palavras, marcadas em negrito no segmento

transcrito anteriormente, quais sejam: REPRODUZIR, TRABALHO, PREFEITURA e TEATRO;

- Fala mais articulada, percebida pela maior clareza em sua produção oral e por meio da velocidade de fala do segmento, mais lenta que nos demais enunciados;
- Uso de poucos “*pitch accents*”, ou seja, uma considerável variação tonal, conferindo proeminência à sílaba tônica, foi percebida nas palavras AGORA e POUCO (sublinhadas no texto);
- Impressão auditiva, inicial, de utilização, pelo sujeito, de um *pitch* mais agudo nesta situação do que nos demais contextos; uma escuta mais atenta⁴⁴, entretanto, não confirma essa impressão.

c) Resultados da Análise de Parâmetros Físico-acústicos

Além dos resultados referentes aos parâmetros acústicos, incluo, neste item, aqueles relativos ao número de sílabas e pausas. Esclareço que estes últimos não estão sendo considerados como medidas acústicas, pois não medi a duração dos segmentos fônicos e nem, por exemplo, o intervalo de pausas silenciosas. Apenas verifiquei a quantidade de sílabas e pausas, por meio dos recursos que o *software CRSE 45* oferece, para que fosse realizada uma contagem mais precisa.

⁴⁴ Esclareço, atendendo à sugestão da banca de qualificação, que o uso da expressão “*escuta mais atenta*” relaciona-se ao fato de que qualquer análise auditiva pressupõe que o material investigado seja ouvido, várias vezes para estudo, conforme explico no capítulo anterior, e, portanto, não significa que a atenção seja menor nem diferente, nas primeiras vezes, em que se ouve um material, mas sim, que quanto mais se escuta, mais se capta nuances, às vezes, pouco ou nem percebidas antes.

A duração total deste enunciado foi de 19,48 segundos e o número de sílabas produzidas pelo sujeito foi 92 sílabas, resultando numa velocidade de fala de 4,72 sílabas/segundo. A duração do enunciado foi extraída no *CSRE 45*, demarcando-se o enunciado na forma da onda e utilizando-se o espectograma sincronizado na tela para auxiliar a demarcação. Identifiquei onze (11) pausas, neste segmento de fala.

Para a análise acústica, selecionei os parâmetros de duração e de frequência fundamental (Fo). Examinei a variação de Fo nas sílabas tornadas proeminentes por *pitch accent* e a variação de Fo nas unidades prosódicas, para investigação da variação de *pitch* dos sujeitos nos diferentes contextos.

Os valores de Fo foram extraídos no *CSRE 45*, por meio do extrator de *pitch*, segundo os procedimentos e parâmetros explicados no capítulo anterior. Foram medidos os valores de Fo inicial, máximo e final de 12 unidades prosódicas (36 valores de Fo medidos). O valor de Fo mínimo encontrado neste enunciado foi 196 Hz e o valor máximo, 403,2 Hz. Os valores médios de Fo encontrados foram: inicial - 252,0 Hz, máximo - 303,6 Hz e final da unidade prosódica - 234,3 Hz.

Em relação às palavras enfatizadas (04), extraí medidas relativas aos valores máximos de frequência fundamental (Fo) e medidas de duração dos núcleos vocálicos das tônicas de cada um dos vocábulos. O Fo mais alto variou de 277,7 a 331,4 Hz e a duração de 187 ms a 244 ms (vide TABELA 2).

TABELA 2 - Valor de Fo mais alto e duração dos *núcleos vocálicos* das tônicas das palavras enfatizadas do Enunciado 1

NÚCLEO VOCÁLICO	DURAÇÃO	VALOR DE Fo MAIS ALTO
REPRODUZIR	213 ms	331,4 Hz
TRABALHO	187 ms	297,9 Hz
PREFEITURA	213,6 ms	323,8 Hz
TEATRO	244 ms	277,7 Hz

Em relação aos valores de Fo no núcleo silábico das tônicas dos *pitch accents*, evidentes em AGORA e POUCO, os resultados encontrados (vide TABELA 3) expressam uma variação significativa, principalmente no vocábulo AGORA, cuja variação foi de 105,6 Hz. Estes dados confirmam os resultados da análise perceptual.

TABELA 3 – Variação de Fo no *núcleo silábico* das tônicas dos *pitch accents* do Enunciado 1

NÚCLEO SILÁBICO	AGORA	POUCO
Fo inicial	297,6 Hz	254,9 Hz
Fo medial	347,2 Hz	215,5 Hz
Fo final	403,2 Hz	235,9 Hz
Fo máximo	403,2 Hz	254,9 Hz
Fo mínimo	297,6 Hz	211,7 Hz
Variação entre Fo máximo e Fo mínimo	105,6 Hz	43,2 Hz

No outro vocábulo, em que o *pitch accent* foi percebido auditivamente, ou seja, em POUCO, verifica-se que a variação do valor de Fo no núcleo silábico da tônica foi menor - 42 Hz. Esta é, porém, uma diferença auditivamente significativa, pois variações acima de 20 Hz são perceptíveis ao ouvinte.

1.2. Análise do ENUNCIADO 2

L. - ENTÃO, ASSIM, // É, // EU TÔ COM MEU CARNEZINHO DE IPTU AQUI //
POR FAVOR // J., // EU GOSTARIA DE SABER // COMO EU POSSO
TRANSFERIR PARA O MEU **NOME** // ESSE CARNÊ? // TEM
POSSIBILIDADE?//

Legenda:

// - pausa

negrito – palavra enfatizada

sublinhado – *pitch accent*

a) Caracterização do Contexto

Este trecho de fala (Anexo 1, p. 6) corresponde ao momento em que o Sujeito 1 (L.) realiza uma atividade de dramatização de uma situação de trabalho, junto com o outro integrante do grupo (J.). A proposta, conforme expliquei no ENUNCIADO 1, foi feita para J., o qual escolheu dramatizar um momento da sua atividade profissional como funcionário público. **Espontaneamente**, (L.) introduz

sua participação nessa atividade, como se fosse uma cliente solicitando informações a J., posicionamento que denominei como *personagem*, considerando-se que se trata de uma situação de dramatização.

Interlocutores: participante dos grupos de voz (J), duas estagiárias de fonoaudiologia (L e M)

Enquadre: dramatização de atividade de trabalho

Alinhamento: personagem (cliente pedindo informação)

b) Dados da Análise Perceptual

Na análise auditiva desta situação observei:

- Ênfase em três palavras (marcadas em negrito no trecho de fala transcrito anteriormente): **AQUI**, **NOME** e **POSSIBILIDADE**;
- Velocidade de fala aumentada em comparação à situação circunscrita no enunciado 1 descrito anteriormente;
- Marcação da ênfase por meio da utilização do *pitch accent*, auditivamente percebida nas tônicas de duas palavras enfatizadas pelo sujeito - **AQUI** e **POSSIBILIDADE**.

c) Resultados da Análise de Parâmetros Físico-acústicos

A duração total deste enunciado foi de 8,3 segundos, e o número de sílabas produzidas pelo sujeito foi 53 sílabas, resultando numa velocidade de fala de 6,38 sílabas/segundo, o que confirma a análise perceptual, ou seja, a de que

há a utilização de uma velocidade aumentada em relação ao primeiro enunciado. A duração do segmento foi extraído no *CSRE 45*, e procedi da mesma forma que no segmento de fala anterior. Para o cômputo das pausas, segui também o mesmo procedimento descrito antes. Foram computadas 7 pausas.

Os valores relativos à variação de *Fo* deste enunciado foram também extraídos no *CSRE 45*. Da mesma forma, considerei como unidade de análise cada trecho de fala entre as pausas, e, em cada unidade prosódica, foram medidos os mesmos valores de *Fo*: início, ponto máximo e final da unidade. O *Fo* mínimo encontrado neste enunciado foi 199 Hz e o máximo, 347 Hz. Os valores médios de *Fo* foram: inicial - 266, 4 Hz, máximo - 298, 8 Hz e final - 263,5 Hz.

Em relação às palavras enfatizadas (03), extraí os valores máximos da frequência fundamental (*Fo*) e a medida da duração do núcleo vocálico das tônicas de cada uma, utilizando o procedimento já descrito. O *Fo* mais alto variou de 261,4 a 304,1 Hz e a duração de 202 ms a 257 ms. (vide TABELA 4).

TABELA 4 - Valor de *Fo* mais alto e duração do núcleo vocálico das tônicas das palavras enfatizadas do Enunciado 2

NÚCLEO VOCÁLICO	DURAÇÃO	VALOR DE <i>Fo</i> MAIS ALTO
AQUI	216 ms	304,1 Hz
NOME	202 ms	283,4 Hz
POSSIBILIDADE	257 ms	261,4 Hz

Duas dessas palavras enfatizadas apresentaram tônicas caracterizadas pela ocorrência de *pitch accent*, de forma mais evidente. No vocábulo AQUI, a variação tonal foi de 33,0 Hz e em POSSIBILIDADE, 42,0 Hz (vide TABELA 5).

TABELA 5 – Variação de Fo no núcleo silábico das tônicas dos *pitch accents* do Enunciado 2

NÚCLEO SILÁBICO	AQUI	POSSIBILIDADE
Fo inicial	297,3 Hz	219,7 Hz
Fo medial	273,8 Hz	255,4 Hz
Fo final	292,6 Hz	260,0 Hz
Fo máximo	304,0 Hz	261,4 Hz
Fo mínimo	271,0 Hz	219,3 Hz
Variação entre Fo máximo e Fo mínimo	33,0 Hz	42,0 Hz

Como auditivamente era bastante evidente um alongamento do segmento fônico ► / na outra palavra enfatizada - NOME, além da medida de Fo, cuja variação foi de apenas 16,9 Hz, realizei a medida de duração. O resultado obtido foi um tempo de 236 ms. Entretanto, como essa palavra - NOME, não foi produzida nenhuma outra vez nos trechos de fala analisados, comparei essa medida com a duração de um segmento fônico similar – ► /, produzido em contexto (antes de uma pausa e com uma certa ênfase) e vocábulo (dissílabo) semelhantes – FONO (extraído do ENUNCIADO 3). A duração do ► / neste

segundo vocábulo foi menor, 120 ms. Por se tratar de uma medida isolada, e pelo fato de os vocábulos serem diferentes, essa comparação não é suficiente para confirmar o alongamento como um recurso fônico utilizado para conferir proeminência a palavra NOME. Permitiu, entretanto, constatar que, de fato, este segmento apresentou uma duração maior.

1.3. Análise do ENUNCIADO 3

L - ASSIM, ACHO QUE NENHUM POSTO, // NENHUM DESCENTRALIZOU. // EM CADA **POSTO** // TERIA QUE TER ASSIM, // TODO UM ATENDIMENTO INTEGRADO // PRA TÁ SUPRINDO AS NECESSIDADES DAQUELE INDIVÍDUO, NÉ? // E A GENTE VÊ QUE NÃO É BEM ASSIM. // POR EXEMPLO, NÉ? // NA ESALQ, A GENTE VIU // NÃO TINHA FONO, // POR EXEMPLO PELO **SUS** // NÃO TINHA OFTALMOLOGIA, // ELE TINHA QUE PEGAR UMA // GUIA PRA E, TÁ NAQUELE POSTO, // NAQUELE NÃO TINHA AQUELE ATENDIMENTO.// ENTÃO, // NÃO SEI ESTAVA ASSIM INTEGRADO.// VAMOS PENSAR ASSIM, VENDO O INDIVÍDUO COMO UM TODO.// ELE TINHA AQUELA NECESSIDADE, // NÃO TINHA COMO BUSCAR ALI. //

Legenda:

// - pausa

negrito – palavra enfatizada

sublinhado – *pitch accent*

a) Caracterização do Contexto

Este segmento de fala (Anexo 2, p.8) foi extraído do terceiro grupo de gravação, constituído pelo Sujeito 1, três outras discentes do Curso de Fonoaudiologia e uma docente (Sujeito 2) do mesmo curso. Neste contexto, o Sujeito 1 ocupa o papel de aluno diante do Sujeito 2 (professora), papéis que realmente exercem na Universidade. Após a leitura de um texto específico sobre Saúde Pública, área de especialidade da docente que coordenava o grupo, seguiu-se uma discussão do tema entre a professora e os demais componentes do grupo. Neste trecho, o Sujeito 1 expressa sua opinião sobre uma questão colocada pela professora (Sujeito 2).

O grupo causou a impressão de que estava bastante envolvido com a discussão, como se de fato estivessem numa situação de aula: a professora atuava de forma a facilitar o debate e as alunas demonstravam interesse em participar da discussão. Na transcrição ortográfica (Anexo 2) é possível constatar que, no final dos turnos da fala da professora, ela, freqüentemente, introduzia uma pergunta, de forma a incentivar a discussão entre o grupo. Por outro lado, observam-se entradas e interrupções de turnos entre as alunas, o que parece ter sido a forma de elas garantirem suas participações na situação interacional. Por este motivo, considerei o posicionamento deste sujeito, neste contexto, como *aluna*.

Interlocutores: professora (R), quatro discentes de Fonoaudiologia (L., M. AP. e A.)

Enquadre: discussão de tema específico

Alinhamento: aluna

b) Dados da Análise Perceptual

Na análise auditiva desta situação observei:

- Número reduzido de palavras enfatizadas, considerando-se a maior extensão deste trecho de fala em relação aos demais segmentos. O sujeito apresentou ênfase em apenas dois vocábulos (marcados em negrito) - POSTO e SUS;
- Fala menos articulada e velocidade de fala aumentada em relação ao primeiro segmento e um pouco maior do que a do segundo trecho de fala analisado;
- Utilização de *pitch accents*, de uma maneira geral, não perceptível, auditivamente, neste segmento, exceto no vocábulo CADA, que não foi considerado como palavra enfatizada na análise perceptual;
- Impressão auditiva, inicial, de variação na faixa de *pitch*, isto é, um agravamento do *pitch* neste contexto, principalmente em relação ao primeiro enunciado; essa impressão, contudo, não se confirma numa escuta mais atenta.

c) Resultados da Análise de Parâmetros Físico-acústicos

A duração total deste enunciado foi de 26 segundos e o número de sílabas produzidas pelo sujeito foi 182 sílabas, resultando numa velocidade de fala de 7,0 sílabas/segundo, o que coincide com a análise perceptual de que ,

neste enunciado, o Sujeito 1 aumenta a velocidade de fala. Os procedimentos para o cômputo da duração e da contagem das pausas foram os mesmos dos outros enunciados, aproveitando-se os dados digitalizados no *CSRE 45*. Em relação às pausas, o sujeito apresentou 20 pausas neste segmento.

Os valores relativos à variação de *Fo* deste enunciado foram extraídos no *CSRE 45*, da mesma forma que nos demais enunciados, ou seja, considere cada unidade prosódica como unidade de análise e extraí três valores de *Fo*: inicial, máximo e final da unidade. O *Fo* mínimo obtido foi 147 Hz e o máximo 377 Hz. Os valores médios de *Fo* foram: inicial- 250,1 Hz, o máximo de 297,8 Hz e o final, 244,7 Hz.

Em relação às palavras enfatizadas (02), extraí os valores de *Fo* máximo e a duração de cada *núcleo vocálico* das tônicas. O *Fo* mais alto variou de 312,0 Hz a 336,0 Hz e a duração de 66 ms a 134 ms, conforme demonstro na TABELA 6.

TABELA 6 - Valor de *Fo* mais alto e duração do núcleo vocálico das tônicas das palavras enfatizadas do ENUNCIADO 3

NÚCLEO VOCÁLICO	DURAÇÃO	VALOR DE <i>Fo</i> MAIS ALTO
SUS	66 ms	336,0 Hz
POSTO	134 ms	312,0 Hz

Como evidenciado na análise perceptual, o sujeito utilizou o *pitch accent* apenas no vocábulo CADA. Os valores de *Fo* encontram-se na TABELA 7:

**TABELA 7 – Variação de Fo no núcleo silábico da tônica
do *pitch accent* do Enunciado 3**

NÚCLEO SILÁBICO	CADA
Fo inicial	254,0 Hz
Fo medial	265,0 Hz
Fo final	284,0 Hz
Fo máximo	284,0 Hz
Fo mínimo	254,0 Hz
Variação entre Fo máximo e Fo mínimo	30,0 Hz

2. ANÁLISE DOS ENUNCIADOS DO SUJEITO 2

2.1. Análise do ENUNCIADO 4

P - A IDÉIA DO **SUS** // É, // NA HORA EM QUE VOCÊ FALA EM DESCENTRALIZAÇÃO, // EM DIREÇÃO ÚNICA. // A IDÉIA É QUE VOCÊ PEGUE TODOS OS EQUIPAMENTOS QUE ANTES ERAM, POR EXEMPLO, DO MINISTÉRIO DA **SAÚDE**, // DO GOVERNO DO **ESTADO** // E PRÓPRIA DA **PREFEITURA**, // VOCÊ JUNTE SOBRE UMA MESMA DIREÇÃO, // TÁ? // É A DIREÇÃO // **ÚNICA** QUE A GENTE FALA. // QUE SERIA QUEM? // O MUNICÍPIO GERENCIANDO, // TÁ? // É, // MAS ISSO, ISSO AINDA, ESTÁ EM TRANSIÇÃO. // NÃO É TODO LUGAR DO BRASIL QUE CONSEGUIU, // É, JUNTAR TODOS EQUIPAMENTOS DESSA FORMA, // TÁ? // POR ISSO QUE A GENTE VÊ ESSA COISA. // POR EXEMPLO, A **ESALQ**, // UMA POPULAÇÃO... // O CORRETO QUE QUE SERIA? // **INCORPORAR** A UBS DA ESALQ // PRO SISTEMA DE SAÚDE DE PIRACICABA // E **ABRIR** // A ESALQ // PRO PÚBLICO GERAL. //

Legenda:

// - pausa

negrito - palavra enfatizada

sublinhado - *pitch accent*

a) Caracterização do Contexto

Este trecho foi extraído do Grupo 3 (Anexo 2, p.7), composto pela professora (Sujeito 2) e pelas 4 discentes do Curso de Fonoaudiologia. Após leitura de um texto sobre assunto da área de especialidade da docente - Saúde Pública, o grupo iniciou uma discussão sobre alguns princípios básicos do Sistema Único de Saúde (SUS), sob a coordenação do Sujeito 2. Este trecho de fala corresponde ao momento em que ele explica um desses princípios aos outros participantes do grupo. A impressão causada por ele é a de estar bastante envolvido com a explicação e com o assunto, como se estivesse ministrando uma aula às alunas, o que me levou a caracterizar o contexto deste enunciado como se segue:

Interlocutores: docente e quatro (04) discentes de Fonoaudiologia

Enquadre: discussão de texto sobre tema específico

Alinhamento: professor

b) Dados da Análise Perceptual

Na análise auditiva observei:

- Pequeno número de palavras enfatizadas, em relação à extensão total do enunciado, apesar de terem sido enfatizadas oito palavras (assinaladas em negrito): SUS, SAÚDE, ESTADO, PREFEITURA, ÚNICA, ESALQ, INCORPORAR e ABRIR;
- Velocidade de fala média em relação aos Enunciados 2 e 3 do Sujeito 1;

- Uso de poucos *pitch accents*. Uma variação tonal considerável no *núcleo silábico* das tônicas foi percebido em três vocábulos, que também foram enfatizados: PREFEITURA, ESALQ e INCORPORAR.
- Impressão auditiva de que o sujeito utiliza um *pitch* grave.

c) Resultados da Análise de Parâmetros Físico-acústicos

Este enunciado foi o que apresentou maior duração total - 44 segundos e, conseqüentemente, um número maior de sílabas: 228 sílabas, o que resultou em uma velocidade de fala de 5,19 sílabas/segundo. A duração do segmento foi extraída no *CSRE 45*. Segui os mesmos procedimentos detalhados na apresentação dos resultados do Sujeito 1, ou seja, marquei todo o segmento no espectograma e obtive o valor correspondente ao tempo do trecho marcado.

A contagem de sílabas foi computada da mesma forma que a dos enunciados do Sujeito 1, ou seja, utilizei os dados digitalizados no *CSRE 45* e ouvi trecho a trecho várias vezes. Usei esse mesmo recurso para marcar as pausas. Os valores obtidos foram revistos e confirmados. O sujeito apresentou vinte e oito (28) pausas neste segmento de fala.

Os valores de F_0 das unidades prosódicas foram extraídos por meio do *CSRE 45*, da maneira explicada nos enunciados antecedentes. Foram extraídos 84 valores ao todo, sendo que o F_0 mínimo encontrado neste contexto foi de 128 Hz e o máximo de 348 Hz.

Os valores médios de F_0 foram: inicial - 209,1 Hz, máximo - 246,9 Hz e final - 178,0 Hz. Os valores de F_0 mais alto e duração do *núcleo vocálico* das

sílabas tônicas das palavras enfatizadas estão discriminados na TABELA 8. Esclareço que na medida da duração do núcleo silábico da palavra ESALQ não considerei a semivogal /w/, pois, como ditongo, sua duração não poderia ser comparada aos demais núcleos.

TABELA 8 - Valor de Fo mais alto e duração do núcleo vocálico das tônicas das palavras enfatizadas do Enunciado 4

NÚCLEO VOCÁLICO	DURAÇÃO	VALOR DE Fo MAIS ALTO
SUS	177,51 ms	304,0 Hz
SAÚDE	274,0 ms	204,9 Hz
ESTADO	231,0 ms	186,9 Hz
PREFEITURA	80,56 ms	260,4 Hz
ÚNICA	157,0 ms	266,0 Hz
ESALQ	103,71 ms	162,5 Hz
INCORPORAR	187,37 ms	313,8 Hz
ABRIR	105,0 ms	304,0 Hz

Tendo em vista as considerações de FANT (op.cit.) e a noção de estrutura silábica (MIRA MATEUS, 1996 e LAVER, 1994), esclareço que me restringi à medição dos núcleos silábicos das tônicas dos vocábulos ABRIR e INCORPORAR, eliminando a *coda* / R /.

A variação tonal dos núcleos silábicos das tônicas dos *pitch accents* foi medida na palavras PREFEITURA, ESALQ e INCORPORAR, vocábulos nos quais essa variação era auditivamente mais proeminente (vide TABELA 9).

TABELA 9 - Variação de Fo no núcleo silábico dos *pitch accents* do Enunciado 4

NÚCLEO SILÁBICO	PREFEITURA	ESALQ	INCORPORAR
Fo inicial	254,8 Hz	158,6 Hz	211,7 Hz
Fo medial	255,8 Hz	154,3 Hz	240,0 Hz
Fo final	227,2 Hz	198,0 Hz	313,8 Hz
Fo máximo	260,4 Hz	162,5 Hz	313,8 Hz
Fo mínimo	227,2 Hz	154,3 Hz	211,7 Hz
Variação de Fo na própria tônica	33,2 Hz	43,7 Hz	102,1 Hz

Os resultados relativos à variação de Fo no núcleo silábico dos *pitch accents* confirmam a análise auditiva, principalmente, no caso do vocábulo INCORPORAR, em que o valor de mudança tonal foi alto - 102,1 Hz.

2.2. Análise do ENUNCIADO 5

P - É, // EU TÔ INTERESSADA SIM. // MINHA FILHA, // ELA TEM TRÊS ANOS //
E ELA, // PARECE QUE ELA // TROPEÇA NAS PALAVRAS. // ELA, // É, // TÁ
GAGUEJANDO JÁ, // SABE? EU CORRIJO ELA. // FALO PRA ELA RESPIRAR
// ANTES DE FALAR. // EU FALO PRA ELA OLHAR // PRA GENTE. // MAS,
NÃO TEM JEITO, // ELA TÁ MESMO GAGUEJANDO. //

Legenda:

// - pausa

negrito - palavra enfatizada

sublinhado - *pitch accent*

a) Caracterização do Contexto

Este trecho foi extraído do Grupo 3 (Anexo 2, p.24), em que o Sujeito 2, a professora, exercia a função de coordenação das atividades. Neste contexto, após propor aos demais componentes (discentes) que dramatizassem um atividade de trabalho, o sujeito, ***espontaneamente***, participa da situação quando uma das discentes (M.) se dispõe a desenvolver a proposta feita.

M. apresenta-se como estagiária de Fonoaudiologia, como se estivesse num Posto de Saúde orientando os usuários, função que exercia, na realidade, como parte de suas atividades acadêmicas. Ela (M.) inicia esta situação de dramatização perguntando ao interlocutor (P.) se estava interessado em alguma informação relativa à alimentação, chupeta ou mamadeira, no desenvolvimento infantil. O Sujeito 2 responde à pergunta e atua como um personagem, ou seja, alinha-se como ***mãe*** de uma criança com problemas de fala. Desta forma, circunstancialmente, sai da sua posição de professora.

Interlocutores: docente (P.) e quatro (04) discentes de Fonoaudiologia

Enquadre: dramatização de atividade profissional (orientação fonoaudiológica)

Alinhamento: mãe (de uma criança com problemas de fala)

b) Dados da Análise Perceptual

Observei que Sujeito 2, neste contexto, apresentou mudança significativa na qualidade e dinâmica de voz, em relação ao enunciado anterior expresso por:

- Maior quantidade de palavras enfatizadas (22), assinaladas em negrito no trecho de fala transcrito acima;
- Velocidade de fala reduzida em relação ao enunciado anterior;
- Uso de um grande número de *pitch accents* (13). Auditivamente, percebe-se variação tonal em uma parcela significativa das palavras enfatizadas pelo sujeito e
- Impressão auditiva de utilização de um *pitch* mais agudo em relação ao Enunciado 4.

c) Resultados da Análise de Parâmetros Físico – Acústicos

O enunciado teve uma duração total de 19,0 segundos e 86 sílabas, resultando numa velocidade de fala de 4,53 sílabas/segundo. Sua duração foi também extraída por meio do *CSRE 45*, segundo os procedimentos já descritos.

Como nos demais enunciados, utilizei os dados digitalizados no *CSRE 45*, para a contagem das sílabas e das pausas (17).

Os valores de *Fo* foram também extraídos por meio do *CSRE 45*. Medi os valores de *Fo* inicial, máximo e final. Foram obtidos, neste segmento de fala, 51 valores, o *Fo* mínimo foi de 115,0 Hz e o máximo 390,7Hz. Os valores médios foram: inicial - 225,3 Hz, máximo - 286,1 Hz e o final - 217,6 Hz.

Os valores de F_0 mais alto e duração do núcleo vocálico das sílabas tônicas das palavras enfatizadas encontram-se na TABELA 10.

TABELA 10 - Valor de F_0 mais alto e duração do *núcleo vocálico* das tônicas das palavras enfatizadas do Enunciado 5

NÚCLEO VOCÁLICO	DURAÇÃO	F_0 MAIS ALTO
INTERESSADA	92,6 ms	271,1 Hz
FILHA	198,56 ms	319,9 Hz
TRÊS	54,2 ms	304,9 Hz
ANOS	172,16 ms	291,1 Hz
ELA (1)	77,92 ms	390,7 Hz
PARECE	100,13 ms	260,2 Hz
ELA (2)	64,95 ms	277,1 Hz
TROPEÇA	214,04 ms	304,4 Hz
PALAVRAS	226,9 ms	226,6 Hz
ELA (3)	52,38 ms	208,4 Hz
GAGUEJANDO	172,17 ms	219,1 Hz
CORRIJO	128,73 ms	255,0 Hz
ELA (4)	187,96 ms	201,2 Hz
RESPIRAR	365,96 ms	208,3 Hz
FALAR	131,29 ms	195,4 Hz
OLHAR	117,1 ms	204,8 Hz
GENTE	86,49 ms	219,9 Hz
NÃO	56,33 ms	348,4 Hz
JEITO	125,65 ms	297,6 Hz
TÁ	155,83 ms	297,6 Hz
MESMO	122,29 ms	348,8 Hz

A variação tonal nos núcleos silábicos das tônicas das palavras enfatizadas, mais evidente em 13 vocábulos, ou seja, os *pitch accents* do Enunciado encontram-se na TABELA 11.

TABELA 11 - Variação de Fo nos núcleos silábicos dos *pitch accents* do Enunciado 5

NÚCLEO SILÁBICO	Fo INICIAL	Fo MEDIAL	Fo FINAL	Fo MÁXIMO	Fo MÍNIMO	VARIAÇÃO DE Fo NA TÔNICA
INTERESSADA	271,1 Hz	244,8 Hz	235,0 Hz	271,1 Hz	235,7 Hz	35,4 Hz
FILHA	290,2 Hz	305,2 Hz	250,2 Hz	319,9 Hz	250,2 Hz	69,7 Hz
ANOS	235,6 Hz	277,1 Hz	278,5 Hz	291,1 Hz	235,6 Hz	55,5 Hz
TROPEÇA	181,0 Hz	205,3 Hz	304,4 Hz	304,4 Hz	181,0 Hz	123,4 Hz
PALAVRAS	158,6 Hz	192,0 Hz	223,7 Hz	226,6 Hz	158,4 Hz	68,2 Hz
GAGUEJANDO	220,0 Hz	194,0 Hz	187,0 Hz	220,0 Hz	187,0 Hz	33,0 Hz
CORRIGO	219,4 Hz	255,0 Hz	245,2 Hz	255,0 Hz	219,4 Hz	35,6 Hz
ELA (4)	160,5 Hz	176,3 Hz	198,8 Hz	201,2 Hz	160,5 Hz	40,7 Hz
RESPIRAR	148,8 Hz	198,6 Hz	208,3 Hz	208,3 Hz	147,4 Hz	20,9 Hz
FALAR	150,6 Hz	189,1Hz	195,4 Hz	195,4 Hz	150,6 Hz	44,8 Hz
OLHAR	171,1 Hz	178,5 Hz	204,8 Hz	204,8 Hz	171,1 Hz	33,7 Hz
NÃO	305,8 Hz	345,8 Hz	377,8 Hz	377,8 Hz	305,8 Hz	72 Hz
MESMO	297,5 Hz	298,6 Hz	348,8 Hz	348,8 Hz	297,5 Hz	51,3 Hz

3. COMPARAÇÃO E DISCUSSÃO DOS RESULTADOS DOS ENUNCIADOS

Início pela análise do *pitch* utilizado pelos dois sujeitos nos diferentes contextos. Para a investigação deste aspecto, foram realizadas as medidas de F_0 inicial, máximo e final das unidades prosódicas de cada enunciado, cujos resultados foram apontados no item anterior. Para comparação foi realizado o Teste de Comparação de Amostras Independentes. Seguem os valores das médias e medianas do Sujeito 1 na TABELA abaixo:

TABELA 12 - Valores das médias e medianas de F_0 inicial, medial e final dos Enunciados do Sujeito 1

ENUNCIADO	VALORES	F_0 INICIAL (Hz)	F_0 MÁXIMO (Hz)	F_0 FINAL (Hz)
1. Coordenadora	Mediana	257,5	300,9	231,4
	Média	252,0	303,6	234,3
	N	12	12	12
	Desvio Padrão	19,2	48,8	23,7
	Mínimo	219,0	250,0	196,0
	Máximo	284,0	403,2	277,7
2. Personagem	Mediana	277,0	304,0	260,0
	Média	266,4	298,8	263,5
	N	9	9	9
	Desvio Padrão	30,3	27,8	23,0
	Mínimo	199,0	250,0	231,0
	Máximo	298,0	347,0	305,0
3. Aluna	Mediana	245,0	290,0	260,0
	Média	250,1	297,8	244,7
	N	19	19	19
	Desvio Padrão	46,0	35,7	47,5
	Mínimo	147,0	223,0	149,0
	Máximo	321,0	377,0	338,0

A análise dos resultados mostra que o Sujeito 1 apresenta uma faixa de variação de Fo (inicial, máximo e final) similar nos três contextos de produção analisados. O *pitch*, portanto, não se constitui como uma pista paralingüística, que se configure como marca distintiva de um enunciado em relação a outro, ou seja, no caso deste sujeito, não há uma variação de *pitch* significativa .

Um dado bastante interessante é que, comparando-se a média dos valores de Fo com os da mediana (correspondente à 50% dos valores extraídos), conforme demonstrado na TABELA 12, observo que o Sujeito 1 apresenta resultados próximos, o que indica que não há valores discrepantes, exceto por alguns valores isolados, como o Fo máximo de 403,2 Hz no Enunciado 1.

Assim, os resultados demonstram que a faixa de frequência usual do Sujeito 1 não se modifica muito, de um contexto para outro e, nos trechos de fala analisados, pode ser situado em torno dos 230 a 300 Hz.

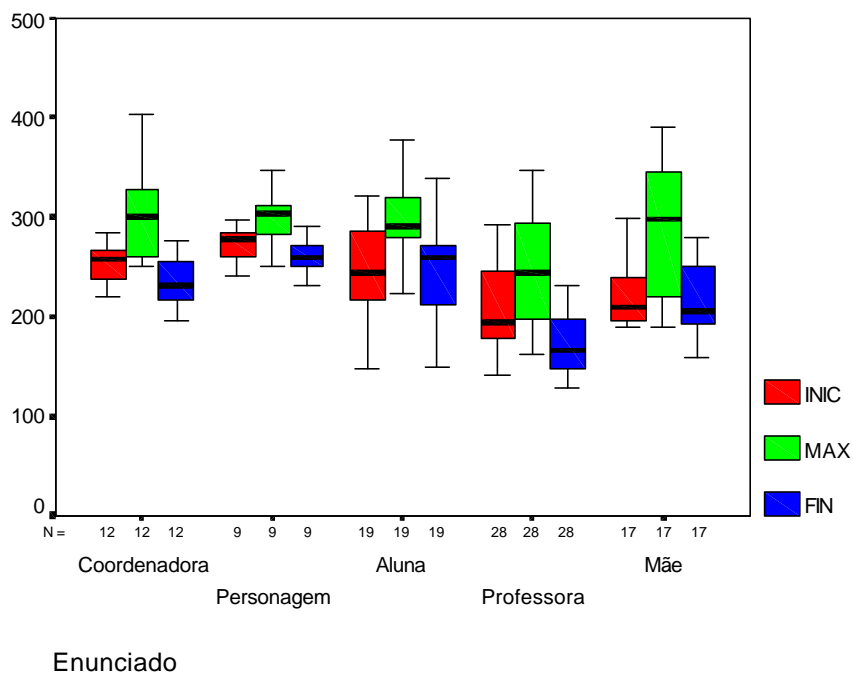
Embora, na análise perceptual, o sujeito tenha causado a impressão auditiva de utilizar um *pitch* mais agudo no primeiro contexto (coordenadora), a análise acústica não confirma essa impressão. A comparação dos demais resultados, que realizo adiante, mostra que os recursos fônicos utilizados pelo sujeito, e que caracterizam sua qualidade e dinâmica de voz, foram outros.

No entanto, retomando essa variável (*pitch*), verifica-se que esta se mostrou relevante no Sujeito 2, distintamente do Sujeito 1. No segundo sujeito, a variação de frequência, se constitui como uma das pistas paralingüísticas, que permitem caracterizar sua voz, de acordo com o contexto de fala produzido. Na análise auditiva observei clara diferença entre um contexto de produção de fala (professora) e outro (mãe) no tocante a esse aspecto. Como professora, o Sujeito

2 utiliza um *pitch* mais grave, que se modifica, quando se alinha como mãe, para um *pitch* mais agudo, além de apresentar maior variação tonal expressa nas palavras enfatizadas e nos *pitch accents* (vide resultados anteriores).

De forma a possibilitar uma melhor comparação e visualização desses resultados, apresento as medianas e demais quartis dos valores de Fo inicial, máximo e final dos enunciados dos dois sujeitos em um único gráfico (vide GRÁFICO 1), dispostos em ordem seqüencial por sujeito, Sujeito 1 (coordenadora, personagem e aluna) e Sujeito 2 (professora e mãe), respectivamente. Os resultados evidenciam a diferença de freqüência entre os dois sujeitos e a variação dessa variável no caso do Sujeito 2, nos dois enunciados analisados (vide professora e mãe, no GRÁFICO 1).

GRÁFICO 1 - Mediana e quartil dos valores de Fo dos Enunciados



GOFFMAN (1981) assinala que, freqüentemente, quando o sujeito muda a voz, não está, necessariamente, encerrando o *alinhamento* anterior, mas, suspendendo-o temporariamente, para, depois, reingressar no *footing* interrompido. Assim, o *pitch* revelou ser uma pista paralingüística, que marca essa mudança de *posicionamento* do Sujeito 2. A interrupção se manifesta na variação de um *pitch* mais agravado para um *pitch* mais agudo. O autor afirma que o sujeito pode manter o *footing* através de vários dos turnos de fala em curso. Trata-se da habilidade do falante em assumir um outro papel no fluxo imediato da situação conversacional, uma habilidade de ir e vir, mantendo em ação diferentes círculos. Desta forma, observa-se que o sujeito sai da sua posição de professora, manifesto fonicamente pela variação do *pitch*, para a de um personagem – “mãe” e retoma seu alinhamento, predominante na situação estudada, o de professora.

Os valores médios de Fo (inicial, máximo e final) mostram que o Sujeito 2 , ao se posicionar como professora, apresenta uma variação de 178,0 Hz a 246, 9 Hz mas, ao se *alinhar* como mãe, essa faixa situa-se em torno de tons mais agudos, e a variação é maior, de 217,6 Hz a 286,1 Hz (vide TABELA 13), constituindo-se em um dos aspectos que caracterizam a mudança circunstancial de *footing*, em consequência dos diferentes enquadres.

Interessante notar que, no caso deste sujeito, os valores médios e das medianas de Fo (inicial, máximo e final) dos trechos de fala analisados apresentam-se de forma diferente. A média de todos os valores de Fo inicial e final é maior que as respectivas medianas, indicando que há uma maior heterogeneidade dos valores, diferentemente do Sujeito 1.

A variação de *pitch*, expressa acusticamente nas medidas de frequência, portanto, constitui em um dos aspectos paralingüísticos que caracterizam a fala do sujeito como **personagem** (mãe), e expressa seu *alinhamento* (“*footing*”) nesta situação, ou seja, a posição que o falante assume para si e para os interlocutores, além da forma como o sujeito conduz ou recebe essa situação de interlocução.

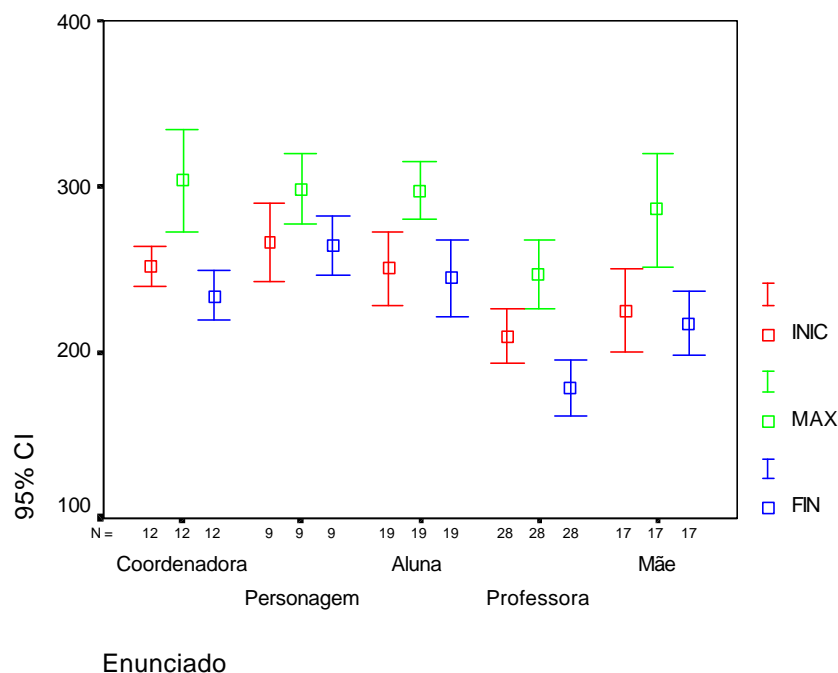
Como professora (Enunciado 4), os resultados das médias e medianas estão próximos, indicando que não há valores discrepantes. Nessa posição, portanto, a faixa de extensão de frequência utilizada é mais homogênea, ou seja, não sofre muitas variações. Outras pistas paralingüísticas marcam esse posicionamento, como demonstrarei adiante.

TABELA 13 – Valores das médias e medianas de Fo inicial, medial e final dos Enunciados do Sujeito 2

ENUNCIADO	VALORES	Fo INICIAL (Hz)	Fo MAXIMO (Hz)	Fo FINAL (Hz)
4. Professora	Mediana	195,0	245,0	165,7
	Média	209,6	246,9	178,0
	N	28	28	28
	Desvio Padrão	41,9	53,1	43,8
	Mínimo	142,0	162,5	128,0
	Máximo	292,0	348,0	304,0
5. Mãe	Mediana	225,3	286,1	217,6
	Média	266,4	298,8	263,5
	N	17	17	17
	Desvio Padrão	49,1	66,3	38,6
	Mínimo	115,0	189,0	158,0
	Máximo	299,0	390,7	278,5

Segue gráfico das médias de Fo inicial, máximo e final dos enunciados:

GRÁFICO 2 - Média dos valores de Fo Inicial, Máximo e Final dos Enunciados



Os resultados demonstram, portanto, que o Sujeito 1 não apresenta uma variação de *pitch* (expressa nas medidas de frequência) significativa de um contexto para outro, diferentemente do Sujeito 2, que apresenta uma variação considerável do *pitch* em decorrência da situação.

Faço a seguir, uma comparação dos demais resultados da análise dos parâmetros físico-acústicos dos enunciados, procurando mostrar como eles expressam o posicionamento e o enquadre dos sujeitos, conforme o contexto de produção de fala. Para tanto, apresento-os, inicialmente, discriminados em forma

de tabela (vide TABELA 14) para facilitar a visualização e a comparação dos mesmos.

TABELA 14 - Resultados das variáveis por enunciado

ENUNCIADO/ ALINHAMENTO	DURAÇÃO TOTAL (segundos)	Nº TOTAL DE SÍLABAS	VELOCIDADE (sílabas/ segundo)	PAUSAS	Nº de PALAVRAS ENFATIZADAS	Nº DE <i>PITCH</i> <i>ACCENTS</i>
1. Coordenadora	19,48	92	4,72	11	4	2
2. Personagem	8,30	53	6,39	7	3	2
3. Aluna	26,00	182	7,00	20	2	1
4. Professora	44,00	228	5,18	28	8	3
5. Mãe	19,00	86	4,53	17	22	13

Considerando-se que os enunciados apresentam durações distintas, os valores relativos às variáveis discriminadas na TABELA 14 foram transformados em índices, à semelhança do cálculo de velocidade (número de sílabas/segundo), a fim de possibilitar uma comparação entre os resultados. Obtive assim, os seguintes índices: pausas/segundo, palavras enfatizadas/segundo, *pitch accents*/segundo. Os índices foram padronizados e encontram-se na TABELA 15.

TABELA 15 - Índices padronizados das variáveis por enunciado

ENUNCIADO	ÍNDICE PADRONIZADO			
	VELOCIDADE	PAUSAS	PALAVRAS ENFATIZADAS	PITCH ACCENTS
1. Coordenadora	-0,671	-1,222	-0,435	-0,405
2. Personagem	0,447	0,883	-0,058	0,344
3. Aluna	1,565	0,131	-0,741	-0,817
4. Professora	-0,671	-0,814	-0,501	-0,715
5. Mãe	-0,671	1,023	1,734	1,593

A análise dos índices padronizados indicam que, no primeiro Enunciado, no qual o Sujeito 1 orienta uma atividade, alinhando-se como coordenadora, a velocidade de fala está abaixo da média (lenta) e o índice de uso de pausas encontra-se ainda mais baixo que a média dos demais enunciados.

A utilização das palavras enfatizadas e a variação tonal nas sílabas tônicas, ou seja, dos *pitch accents*, resultaram em índices próximos. Representam, porém, uma utilização desses recursos abaixo da média dos outros trechos de fala.

Quando as pessoas estão na presença uma das outras, suas atitudes verbais e não verbais são fontes potenciais de comunicação e suas ações e intenções de significados só podem ser entendidas se relacionadas ao contexto imediato (TANNEN & WALLAT, 1987). Vale reassegurar, portanto, que estes dados só fazem sentido se interpretados em relação ao contexto em que foram produzidos, como lugar de trabalho social, um lugar de produção de significados. Nesse sentido, os pressupostos de GOFFMAN são fundamentais para esta análise.

Os resultados indicam que o posicionamento do Sujeito 1 como **coordenadora** é expresso, marcadamente, por uma **velocidade de fala reduzida** e uma **articulação clara**. As outras pistas fônicas relacionam-se ao **pouco uso de palavras enfatizadas** e de **pitch accents**. Nas poucas vezes em que L. utiliza esses recursos, dá destaque às palavras chaves do tópico a ser discutido.

Segundo os princípios que regulam uma situação conversacional, analisados por GOFFMAN (1974), essas manifestações fônicas decorrem de como o sujeito se alinha na situação interacional, de forma a assegurar a *condução da atividade* proposta no enquadre. As manifestações da materialidade fônica do sujeito fornecem também indícios da representação⁴⁵ que este falante tem neste enquadre, ou seja, de que a voz de um coordenador seria uma **"voz de comando"** ou de "liderança", como ela própria expressa ao final da gravação, quando tece alguns comentários sobre este tópico:

"... Eu, consigo assim, voltar e impostar mais, dá uma coisa assim, mais de liderança, né ?..."

(L., Anexo 2, p.26, grifo meu)

O falante, consciente ou inconscientemente, escolhe recursos fônicos que, segundo sua própria interpretação, possam causar o impacto ou o efeito esperado no ouvinte. A estabilidade e a variabilidade do que ocorre na interação resulta da natureza emergente do discurso e é consequência do contexto social, como

⁴⁵ Cabe a ressalva de que a representação que o sujeito tem de sua voz, neste e em qualquer outro contexto interacional, envolve uma discussão mais abrangente de aspectos sociais, culturais e históricos, mas que fogem ao escopo desta pesquisa. Restrinjo-me, portanto, àqueles capturáveis por este trabalho.

afirmam TANNEN & WALLAT (1987). Numa perspectiva de análise que não reduza a produção vocal a um mero ato laríngeo, é possível ver a materialidade fônica como resultante/manifestação dos movimentos realizados pelos participantes numa dada situação interacional, como neste caso e no dos demais enunciados, cuja discussão segue-se a esta.

Na seqüência dos comentários de L., há passagens em que ela demonstra claramente esse movimento de idas e vindas, demonstrando sua habilidade de falante, uma habilidade de mudar o *footing* dentro de uma situação. (GOFFMAN, 1981). Ela refere-se ao fato, afirmando que ao ler "... *coloco um tom mais grave porque eu sei que assim, que o outro tá me escutando e tem que entender...*" (ANEXO 2, p.26), acrescentando que, quando a colega (M.) interagiu com ela, como se fosse J., pôde perceber que seu tom de voz ficou "*mais agudo*".

Esse movimento é governado por princípios de organização que repercutem nas manifestações fônicas decorrentes do *alinhamento* do sujeito em função do *enquadre*, nos termos de GOFFMAN. Nas palavras do autor, os participantes usam *esquemas de referências* ("*frame analysis*"), estruturados conforme a sociedade, mas que estão sujeitos a vulnerabilidades, como erros e disputas de enquadres. A *transposição* ("*keying*"), um dos conceitos centrais em sua obra, considera o processo como um fragmento de atividade que pode ser transformado e enquadrado diferentemente, como o que ocorre no segundo enunciado. Neste, o *enquadre* muda para uma dramatização de uma atividade profissional, levando o Sujeito 1 a *posicionar-se* como *personagem* (cliente pedindo informações) e "sair", ainda que brevemente, da posição de *coordenadora*.

A partir dessa *transposição*, o *posicionamento* do Sujeito 1 repercute em mudanças na produção vocal, conforme os resultados anteriormente descritos. A padronização dos índices confirma esses dados, indicando que uma das diferenças mais significativas em relação ao primeiro enunciado é a velocidade de fala aumentada do Enunciado 2, um pouco acima da média em comparação com os demais enunciados. Por se tratar de um trecho de fala com uma duração mais curta, o mesmo número de *pitch accents* (2) do Enunciado 1 resultou num índice indicativo de maior uso desse recurso, embora ainda pouco produtivo, principalmente se comparado ao Enunciado 5, no qual houve a maior utilização de *pitch accents*. Há um aumento significativo das pausas e o índice de palavras enfatizadas é um pouco maior do que no primeiro contexto (vide TABELA 15).

Nesse caso, a *transformação* da atividade corresponderia à *fabricação do tipo "benigna"* ("*benign fabrication*"), pois houve uma ***elocução fabricada***. Neste contexto, o Sujeito 1 representa uma identidade fictícia e, talvez por isso, parece ***ter uma menor preocupação com o seu desempenho de fala***, ou com a demonstração de suas habilidades como falante, em comparação com sua atuação nos demais enunciados. Os demais parâmetros apresentam-se em torno de valores médios, com exceção do índice de ***pausas um pouco maior*** (o segundo maior dos enunciados), o que pode ser interpretado como um recurso fônico para ***conferir a clareza e a inteligibilidade de fala*** que, na representação que o sujeito tem desse papel, seriam necessárias à um ***cliente***.

O sujeito causa a impressão de que, neste segmento de fala (Enunciado 2), sua preocupação é com uma ***produção vocal clara e inteligível***, por "estar pedindo informações". Nesse enquadre, ela não tem que se preocupar com uma

"voz de comando", uma vez que, naquele momento, não é a coordenadora quem fala, e sim uma cliente, em mútua colaboração com seu interlocutor (J.), que se coloca como funcionário, fornecendo as explicações solicitadas. Isso coincide com as afirmações de TANNEN & WALLAT (1987) de que, em acordo com os pressupostos de GOFFMAN, todo participante, em qualquer interação, colabora na negociação dos enquadres que operam numa dada situação.

Assim, as necessidades que se colocam no jogo interacional nos Enunciados 1 e 2 são distintas e geram ajustes de *alinhamento* do Sujeito 1, que se manifestam, lingüística e paralingüisticamente, de formas diferentes, como demonstram os resultados. É importante reafirmar que a participação de L., e conseqüentemente seu ***posicionamento no enquadre 2***, ocorreu ***espontaneamente*** e, portanto, suas manifestações vocais ocorreram em decorrência do novo contexto. Como esclareci antes, reconheço que outros fatores estão envolvidos neste *alinhamento*, como a representação que o sujeito teria da voz de um cliente, que perpassam pelos *esquemas de conhecimento*⁴⁶ discutidos por TANNEN & WALLAT (op.cit.). Esta discussão, porém, foge aos propósitos deste trabalho.

Em relação ao terceiro Enunciado, em que o sujeito se *alinha* como ***aluna***, a mudança mais evidente está no índice de ***velocidade de fala aumentado***, o maior índice de todos os enunciados, conforme demonstra a TABELA 15. Não há uma utilização produtiva de palavras enfatizadas ou de *pitch accents*. As pausas

⁴⁶ As autoras esclarecem que utilizam o termo "*esquema de conhecimento*" em referência às expectativas dos participantes acerca das pessoas, objetos, eventos e cenários do mundo. Acrescentam que os lingüistas, particularmente, aqueles dedicados ao campo da Semântica, têm interesse nesse fenômeno e observam que o significado literal de um enunciado somente pode ser entendido relacionado à um "modelo" de conhecimento anterior (aspas minhas, pois considero que não, necessariamente, relaciona-se a um "modelo" e sim, a vivências, experiências e/ou conhecimentos anteriores).

são expressas por um valor apenas um pouco acima da média geral dos demais enunciados, apesar de L. ter apresentado 20 pausas neste trecho de fala, em função da duração maior do mesmo.

Essas mudanças estão relacionadas ao seu ***posicionamento como aluna***, decorrente da transposição de enquadre. Esse aumento da velocidade de fala pode ser interpretado como uma forma de L. ***assegurar sua participação no contexto*** - discussão de um tema específico. Neste *enquadre*, a ***disputa*** (restringo-me ao *posicionamento* como falantes) entre os participantes é maior. Nos termos de GOFFMAN (1981), corresponde a uma resposta à *referência* que caracteriza esse tipo de enquadre. *Referência* é o termo que o autor propõe em substituição a *pergunta/declaração*, para designar o elemento iniciador da conversação, como esclarece MALUFE (1992). Neste caso, a *referência* é a proposta da professora (Sujeito 2) para que se discutisse sobre aspectos da Saúde Pública, à qual L. teria que responder. Sua "resposta" manifesta-se num ***aumento da velocidade de fala e num uso pouco produtivo de recursos fônicos***, como ênfases e variação tonal nas sílabas tônicas.

Além disso, para ocupar essa posição de aluna, L. teria a necessidade de possuir ou, pelo menos, aparentar, convincentemente, que possuía as qualificações pessoais (*role attributes*, segundo GOFFMAN, 1981) esperadas para este papel, seja pela auto-imagem que ela própria se atribui, seja pela imagem que ela julga que os outros participantes fazem dela e que poderia, ou não, estar sendo cobrada, tanto por ela própria, quanto pelos outros, frente a essa situação de interação face a face. Tais pressuposições, além do atravessamento

institucional e sociocultural da representação de L., como aluna, podem ter provocado essas mudanças vocais.

Distintamente do primeiro contexto, em que L. busca garantir a condução de uma atividade por meio de uma velocidade mais lenta e do uso de ênfase nas palavras-chaves do tópico, para, possivelmente, demonstrar maior clareza e habilidade de comando, nesse terceiro enunciado, o sujeito procura mostrar sua competência como aluna, manifesta, principalmente, por meio do aumento de velocidade de fala e pela utilização de um turno mais longo. Além disso, apresenta índices mais baixos de palavras enfatizadas e de *pitch accents*, refletindo uma utilização pouco produtiva desses recursos fônicos.

A padronização das variáveis estudadas mostra que os Sujeitos 1 e 2 apresentam resultados similares nos Enunciados 1 (***coordenadora***) e 4 (***professora***), conforme demonstram os resultados discriminados na TABELA 15. L. e P. apresentam, nestes dois contextos, uma produção vocal manifesta por ***velocidade lenta, poucas pausas e utilização pouco produtiva de palavras enfatizadas e pitch accents***. Como as duas situações são de condução, os dados sugerem que a preocupação de ambas estaria na clareza e na inteligibilidade de fala, o que pode ser interpretado como uma forma de os dois sujeitos garantirem a atenção e a compreensão por parte do interlocutor (es).

Cada enquadre pressupõe atitudes diferentes do sujeito, que se expressam na materialidade fônica, ou seja, na maneira de falar, ou ainda, na expressão e estabelecimento do *footing*. Assim, como ***personagem*** (cliente), a velocidade de fala é média, os índices de palavras enfatizadas e *pitch accents* são um pouco mais produtivos do que como aluna ou coordenadora, mas não se comparam, por

exemplo, aos altos índices desses itens no Enunciado 5 (Sujeito 2). Os ajustes fônicos ocorridos são inerentes à cada uma das situações descritas ou, como esclarece GOFFMAN (op.cit.), são necessários ao contexto, ao que ele denominou de “*jogo conversacional*”, e refletem o alinhamento do sujeito em decorrência do enquadre em que o trecho de fala ocorreu.

No Enunciado 5, o Sujeito 2, como coordenadora do grupo, faz a proposta da atividade de dramatização e, nessa situação, *alinha-se* espontaneamente como **mãe** de uma criança com dificuldades de fala, o que resulta em variações em sua voz, bastante distintas das dos demais trechos de fala analisados. O sujeito mantém a mesma velocidade que utiliza na situação como professora, uma **velocidade lenta**, porém, há uma **variação tonal significativa**, uma produção vocal **marcada por proeminências mais evidentes**, expressa em um uso mais produtivo de palavras enfatizadas, de *pitch accents* e de pausas.

O Sujeito 2 modifica sua voz neste contexto, de forma **não deliberada**, ainda, que seja uma situação de dramatização, portanto, uma **elocução fabricada**. Sua preocupação não era como produzir essa determinada voz: os ajustes fônicos ocorreram **naturalmente** em decorrência da representação que o sujeito faz da voz de uma mãe, na situação de interlocução com um “profissional” (o aluno no papel de terapeuta). Assinalo, novamente, que reconheço, que esta representação está atravessada por valores, vivências e expectativas pessoais, sociais, culturais e históricas, mas não é meu objetivo tratar, aqui, destes aspectos.

Para completar a análise estatística dos dados, foi realizado um estudo do conjunto de todas as variáveis, por meio da técnica multidimensional de *Análise Fatorial de Componentes Principais*.

Tal técnica possibilitou verificar quais parâmetros estudados eram, estatisticamente, mais significativos para distinguir um enunciado do outro. Essa análise indicou a existência de 2 fatores principais que explicam 65,78% e 32,31%, respectivamente, da variação total entre os enunciados.

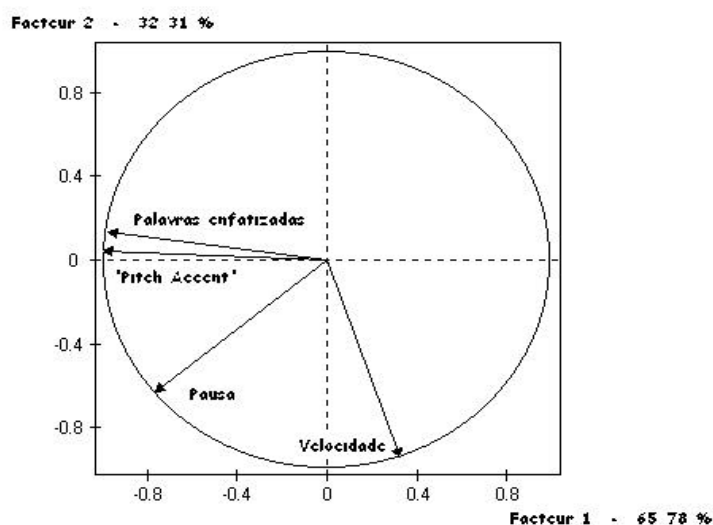
O Fator 1 (porcentagem maior) é definido pelas variáveis: ***pausa***, ***palavra enfatizada*** e ***pitch accent***, e o Fator 2 (porcentagem menor) pela ***velocidade***, conforme demonstra a TABELA 16 :

TABELA 16 - Correlação das variáveis ativas dos eixos fatoriais 1 e 2

VARIÁVEL ATIVA	CORRELAÇÕES	
	FATOR 1	FATOR 2
Velocidade	0,33	-0,94
Pausa	-0,76	-0,63
Palavra Enfatizada	-0,98	0,13
<i>Pitch Accent</i>	-0,99	0,04

O GRÁFICO 3 é uma outra forma de representação que possibilita demonstrar as relações entre as variáveis ativas (índices de velocidade de fala, de pausas, de palavras enfatizadas e de *pitch accente* as suas relações com os fatores indicados.

GRÁFICO 3 - Relações entre as variáveis ativas



A TABELA 17 mostra a posição dos enunciados nos eixos fatoriais 1 e 2:

TABELA 17 - Posição dos enunciados nos eixos fatoriais 1 e 2

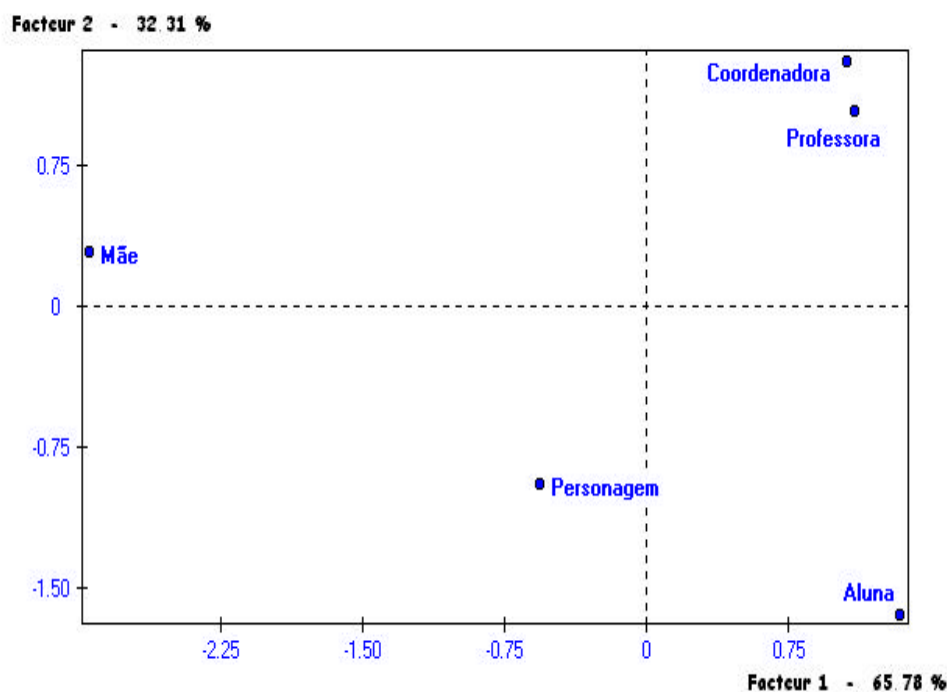
ENUNCIADOS	EIXOS FATORIAIS	
	1	2
1. Coordenadora	1,06	1,30
2. Personagem	-0,56	-0,95
3. Aluna	1,34	1,65
4. Professora	1,10	1,03
5. Mãe	-2,95	0,28

Estes valores demonstram resultados bastante interessantes, como a discrepância do Enunciado 5, nos dois eixos fatoriais, em relação aos demais e,

em menor grau, ao Enunciado 2. Estes dois enunciados podem ser considerados como um mesmo tipo de transformação da atividade, ou seja, uma *fabricação*. Por outro lado, os valores dos demais Enunciados (1, 3 e 4) guardam proximidade e correspondem a enquadres mais próximos da realidade dos Sujeitos 1 e 2 na relação acadêmica – aluna e professora.

A Análise Fatorial dos Enunciados resultou no GRÁFICO discriminado a seguir:

GRÁFICO 4 - Análise Fatorial dos enunciados



Este gráfico revela-se bem elucidativo quanto à manifestação fônica dos sujeitos em cada Enunciado analisado, e permite confirmar a hipótese inicial de que a voz muda conforme o contexto. Pode-se notar a proximidade entre os

Enunciados 1 (Sujeito 1, aluna no papel de coordenadora) e 4 (Sujeito 2, professora no seu papel institucional, ou seja, como professor). Essa relação se justifica em função dos dois Enunciados terem sido produzidos em enquadres semelhantes, ou seja, de "condução de atividade". Além disso, o gráfico demonstra que estes Enunciados são antagônicos ao Enunciado 5 (Sujeito 2, mãe), uma situação "*fabricada*" e de não condução.

Esse antagonismo expressa-se paralinguisticamente em função das variáveis do Eixo fatorial 1 composto por: pausa, palavra enfatizada e *pitch accent* (vide TABELA 16).

O Enunciado 3 (Sujeito 1, aluna) aproxima-se dos Enunciados 1 e 4 pelo Eixo fatorial 1, contudo, distancia-se dos mesmos pelo Eixo fatorial 2, cuja variável é a velocidade. Retomando os resultados da TABELA 15, verifica-se que, nesse Enunciado (3), ao alinhar-se como aluna, o Sujeito 1 apresenta velocidade de fala mais rápida em relação aos demais enunciados, o que leva a esse distanciamento. O aumento de velocidade justifica-se, por tratar-se de um enquadre em que o sujeito procura ocupar os turnos de fala, em detrimento dos demais interlocutores (colegas discentes), como forma de demonstração de "*conhecimento*", frente ao professor. Este resultado coincide com os estudos de SERVILHA (2000), que assinala que essa fala encadeada significa que um interlocutor deve falar (no caso do estudo da autora, o professor - P1), enquanto que o outro deve ouvir (o aluno, no estudo de SERVILHA). Sem dúvida, as forças atuantes eram a autoridade de professor e a competição, ainda que velada, entre os alunos.

O Enunciado 5 diferencia-se dos demais pelos altos índices de pausas, de palavras enfatizadas e de *pitch accents*, indicando um padrão vocal bem diferente dos demais. Guarda bastante semelhança com a descrição de fala de uma pediatra discutida em TANNEN & WALLAT (1987). As autoras assinalam que, durante uma consulta pediátrica, a médica utiliza um "*maternalês*"⁴⁷, ao dirigir-se à criança, sua cliente, caracterizado por mudanças exageradas no *pitch* e por uma prosódia marcada por pausas longas e surtos de vocalizações. Todos os participantes colaboram na negociação dos enquadres operantes numa dada situação conversacional. O Sujeito 2 atende à demanda associada à figura de *mãe* de uma criança com dificuldades de linguagem, fazendo uso de uma entoação bem marcada, o que coincide com o estudo das autoras.

O GRÁFICO 4 demonstra, de forma clara (visualmente), as diferenças e semelhanças dos sujeitos em cada contexto de fala analisado.

Finalizando, analiso os resultados sob duas formas: intrasujeitos e intersujeitos.

Na primeira forma, intrasujeitos, constato que o Sujeito 1 apresentou variações em sua voz nos três contextos analisados, quais sejam, como coordenadora, aluna e cliente. Suas produções vocais guardam semelhanças em dois contextos da sua vida acadêmica, como aluna e coordenadora, no aspecto de proeminência (pouco uso de palavras enfatizadas e de *pitch accents*), e se distanciam nos aspectos de velocidade de fala (maior como aluna) e índice de pausas (bastante baixo como coordenadora). Essas variáveis apresentam valores distintos do terceiro contexto analisado, ou seja, do sujeito como

⁴⁷ Termo utilizado por Parmênio Camurça Citó, na tradução do texto original das autoras, em RIBEIRO & GARCEZ, 1998.

personagem, o que confirma que a variação da voz do primeiro sujeito nas três situações estudadas. O Sujeito 2, por sua vez, apresentou, somente uma velocidade de fala idêntica nos dois contextos estudados. Os demais parâmetros resultaram em valores bem distintos, caracterizando uma variação significativa de voz, do segundo sujeito nas situações analisadas.

Interpretando-se os resultados intersujeitos, confirma-se que, também, há variação de voz, conforme o contexto de produção. A variável, estatisticamente mais significativa entre os sujeitos, foi a variação de frequência. Um aspecto importante nas mudanças vocais expressas pelo Sujeito 2, mas não, para o Sujeito 1. Na situação de condução, as vozes dos dois sujeitos mais se aproximam. Os valores dos vários parâmetros são muito parecidos, exceto no índice de pausas. Comparando-se os sujeitos nas duas elocuições fabricadas, ou seja, quando ambas atuam como personagens (cliente e mãe), constato que há coincidência somente em relação ao aumento do índice de pausas. Os demais parâmetros resultaram em valores bem distintos, indicando que os recursos fônicos utilizados não guardam semelhanças nesse enquadre (dramatização de um personagem). Cabe retomar, que os papéis representados pelos sujeitos eram diferentes e reafirmar que cada um possui representações próprias das vozes que produzem, de acordo com a experiência e a história própria de vida, aspectos que, certamente contribuíram para as diferentes manifestações fônicas e que constituem fértil campo de pesquisa.

Os dois sujeitos, comparados nos contextos semelhantes ao da vida institucional, ou seja, o primeiro como aluna, e o segundo, como professora, se distanciam nos aspectos de velocidade de fala e no índice de pausas e se

aproximam quanto à utilização pouco produtiva de proeminências em seus discursos orais, nestes contextos.

Assim, analisando-se as manifestações fônicas, os resultados mostram que há variação da materialidade fônica, conforme o contexto em que a voz é produzida tanto intrasujeitos como intersujeitos.

Por fim, teço, no último capítulo, as considerações finais deste trabalho.

CAPÍTULO V

“Os recursos sonoros são trabalhados para significar.”

(Sandra MADUREIRA, 1992, grifo da autora)

CONSIDERAÇÕES FINAIS

Na introdução desta tese, indiquei que pretendia tratar de questões relativas à voz, de um modo distinto. Fui motivada, como anunciei logo de início, por uma certa inquietude diante das minhas observações sobre as mudanças na qualidade vocal em situações de interação, que não seriam (e não eram) suficientemente explicadas, se analisadas somente sob o prisma predominante na literatura. Apontei a necessidade de um outro foco de análise para compreender este fenômeno como algo além de um mero ato laríngeo. Posso afirmar agora que essa crença se transformou em uma possibilidade científica de investigação da voz, à luz do reconhecido (teoricamente), porém, notadamente negligenciado, **aspecto social**. Quero dizer, com isso, que o conhecido chavão, de que a voz deveria ser compreendida em seus aspectos biopsicosociais, ganha, de fato, concretude, a partir da abordagem analítica que utilizei nesta pesquisa.

Apresenta-se, assim, uma outra instância de análise, que se distancia de uma visão estática e individual do fenômeno vocal, para outra, dinâmica e social. Este trabalho representa um passo adiante no estudo do **papel da materialidade fônica na interação** e uma possibilidade de demonstrar que os recursos sonoros são realmente utilizados para **gerar significado**.

Função primordial de qualquer ato lingüístico, a produção de sentidos se constrói na relação de um indivíduo com o outro, e se manifesta através de diferentes recursos, entre eles, a voz – foco desta pesquisa, e meus achados podem ser considerados como uma confirmação do que anunciei no início deste estudo: a produção vocal não pode ser vista somente em sua dimensão biológica.

Segui a linha defendida por PITTAM (1994), de uma abordagem multidimensional do estudo de voz, que elege, como objeto de estudo, o fenômeno vocal, como importante aspecto da comunicação humana, isto é, o **estudo da voz na interação social**. Esse tipo de abordagem implicou que eu não me restringisse a medir a voz acusticamente, nem a avaliá-la perceptualmente, mas, que essas modalidades de análise fossem estudadas de uma forma integrada, sob uma perspectiva social.

Para que uma tarefa desse tipo fosse possível, necessitei de um quadro teórico composto de várias dimensões complementares e integradas. A primeira implicou o uso de um modelo fonético centrado na voz do sujeito, como o proposto por LAVER, que possibilitou relacionar os ajustes articulatórios da produção vocal (analisados perceptualmente) às medidas acústicas da onda sonora.

Ao fazer esse percurso, **pude distanciar-me do fenômeno vocal como ato perceptível apenas individualmente** e tornar o fenômeno físico em algo **perceptível ao outro**, o que serviu de ponte para a outra dimensão: a **social**. Assim, a utilização de um **modelo fonético** tornou possível a análise da voz, **perceptual e acusticamente**, enquanto o uso de um **modelo social** possibilitou o estudo do cotidiano, ou seja, **dos atos de fala**. Esse estudo integrado proporcionou uma visão clara de como os **eventos fonéticos se processam e transcorrem dentro da interação**. Tentar compreendê-los nesta perspectiva constituiu o ponto de partida da minha pesquisa e, agora, essa compreensão tornou-se um ponto de chegada, cientificamente demonstrável. Viabiliza-se um caminho para o estudo da normalidade, deslocando-se, assim, o estudo de voz,

da dicotomia saúde/doença, tão preconizada na Fonoaudiologia e na Laringologia, deslocamento este que representa uma exigência para que se possa apreender a complexidade e a riqueza dos recursos fônicos das variações das vozes ouvidas e, desta forma, superar essa dicotomia e o atrelamento da produção vocal a seu aspecto essencialmente fisiológico.

O estudo, assim posto, possibilita demonstrar que o fenômeno vocal não significa, meramente, dois (ou mais) indivíduos se comunicando, e sim que eventos sociais podem ser, e geralmente o são, influenciados pelo institucional (neste caso, pela relação acadêmica dos sujeitos estudados) e pela natureza do arcabouço social e estrutural da sociedade em que eles se inserem, e, ainda, que os eventos sonoros tem um papel crucial no efeito do sentido.

Assim, iniciei esta proposta de investigação no indivíduo. Entretanto, em função das dimensões analíticas mencionadas, não me restringi a ele (de forma isolada), mas sim, pude estudá-lo, colocando-o em contato com os outros e com o mundo em que ele está inserido. Essas associações entre as diferentes dimensões proporcionaram um caminho para que esta abordagem analítica fosse feita não de forma fragmentada, e, sim, permitindo que o objeto de estudo fosse focalizado de modo integrado.

Considero que o fruto principal desta pesquisa foi a constatação da **variação das vozes dos sujeitos** (estudados) **na interação social** ou, em outras palavras, **como a interação transforma a voz**, o que veio confirmar minha hipótese inicial.

Assim, o estudo mostrou que a variação dos recursos fônicos utilizados pelos sujeitos analisados estava diretamente relacionada a seus *alinhamentos*,

em função do contexto de produção vocal. Os eventos de fala, como afirma MADUREIRA (1992, p.162), se constroem a partir da *“contrastividade entre recursos e da superposição de recursos, o que concorre para a implementação da variabilidade na fala.”*, e a consideração desse parâmetro reside, intrinsecamente, nos papéis assumidos pelo falante e em seus *posicionamentos*.

A análise da materialidade fônica, da forma realizada, evidencia a riqueza e a variabilidade da produção vocal desses falantes na interação social e demonstra **a relevância do recurso fônico como marca de constituição da própria oralidade**, posição sustentada tanto nos referenciais teóricos que adotei no trabalho, quanto nos resultados encontrados.

As contribuições deste estudo são de grande importância dentro do panorama atual das pesquisas em voz, uma vez que **revelam uma perspectiva de investigação do fenômeno sob um outro olhar: o social**. Olhá-lo sob este prisma implica um compromisso com o fator social, freqüentemente relegado, bem como um rompimento com os limites estritos do orgânico, circunscrito, primordialmente, ao patológico. Ao mesmo tempo, denuncia a necessidade de se compreender a voz de forma mais abrangente, fornecendo-lhe, assim, um outro tratamento.

É interessante sublinhar que esse enfoque permite que a produção vocal seja estudada como um elemento constitutivo da e na interação social: um espaço de negociação e colaboração mútua, expresso na materialidade fônica, como demonstram os resultados, que, por sua vez, coincidem com a afirmação de SOUZA e SILVA (1995, p.107) de que: *“... a interação verbal para ser bem sucedida depende da cooperação ativa dos interlocutores e de que a interpretação produzida se*

constrói durante a troca interativa.”. Os participantes abstraem elementos do diálogo em curso, e esses elementos determinarão a manifestação fônica de cada falante presente no contexto, ou seja, a escolha dos recursos fônicos está na dependência das práticas de linguagem ou, ainda, da interação verbal em andamento.

Reafirmo que esta tese representa, portanto, um modo diferente de compreensão da materialidade fônica: uma possibilidade de análise dos atos do cotidiano, isto é, dos momentos de fala produzidos pelos sujeitos na interação. Abrem-se, assim, novas direções.

Um modo novo de olhar a vida cotidiana e de examinar os eventos fônicos na situação social, esta pesquisa mostra, portanto, a aplicabilidade dos pressupostos de GOFFMAN, isto é, de análise da produção vocal à luz de noções, como *“frame”* e *“footing”*.

Por essa razão, ela passa a representar uma contribuição substancial tanto para a Fonoaudiologia quanto para a Lingüística, que atuou como um importante posto de observação da voz, distinto do tradicional, e conduziu a um deslocamento da abordagem analítica da materialidade fônica, marcadamente organicista e clínica, para uma abordagem ***social e fonética***.

A redundância dos eventos fônicos puderam ser estudados por meio da análise perceptual e acústica, com resultados algumas vezes coincidentes, e outras vezes não, o que demonstra que um recurso não exclui o outro, ao contrário se complementam e, portanto, há a necessidade de que ambos sejam utilizados. Assim, o papel da análise auditiva fica reiterado, ao mesmo tempo em que se reafirma a importância da análise acústica.

As nuances da voz marcaram o **posicionamento** dos sujeitos e puderam ser analisadas **perceptual e acusticamente**, o que demonstra a inserção na Fonética dos estudos sobre a voz. Quando esses estudos são feitos de forma **contextualizada**, o **contexto** passa a jogar importante papel aí, e o modelo de GOFFMAN, pelo seu poder explanatório, contribui de forma crucial para conferir-lhe concretude e sentido.

A voz causa impacto no outro e, também, sofre transformações pela relação com esse mesmo interlocutor. Um movimento de idas e vindas dos sujeitos, expresso em variações da **qualidade e dinâmica de voz**, podem ser resgatadas a partir da análise fonética e consideradas no estudo do uso e do papel dos recursos fônicos.

Assim, outro fruto importante, que posso citar, está diretamente ligado à **dimensão interdisciplinar**, presente através das vozes dos diversos autores que utilizei, como referencial teórico, para este estudo, e que permitiu que a voz fosse olhada sob esta outra perspectiva analítica. Essa interdisciplinariedade representa um avanço para a Fonoaudiologia e para a Laringologia, pelo compartilhamento de conhecimentos da Lingüística e da Sociologia nos estudos sobre voz.

A mudança de foco dessa perspectiva de análise, antes mais **estática e individualizada** e, agora, mais **dinâmica e intersubjetiva**, contribuiu para que as fronteiras do individual e do orgânico - a voz como mero ato laríngeo - fossem transcendidas, para a compreensão da **voz em sua legítima dimensão, a social** (SERVILHA, 2000), ou seja, a voz como produto social.

Os resultados mostram que a participação de cada sujeito na sua relação com o outro é permeada pela voz, isto é, é expressa pelos recursos fônicos que o indivíduo utiliza para demonstrar sua posição no mundo. Mostra, ainda, que tais escolhas são demonstráveis por meio das análises perceptual, acústica e social, ou seja, através de uma abordagem analítica multidimensional e integrada, como a que elegi aqui.

Também para a Fonoaudiologia, os achados deste trabalho representam importante contribuição para a chamada área de Voz Profissional, nos diferentes níveis de atenção à saúde (promoção da saúde vocal e prevenção de alterações de voz), por meio de propostas de atuação nesse campo através de assessorias, ou *Grupos de Vivência de Voz*, como os citados neste trabalho, junto à empresas, escolas, academias de ginásticas, grupos de teatro ou de corais, assim como em outros espaços em que a voz seja usada como instrumento de trabalho.

Eles repercutem também para a Clínica Fonoaudiológica, por proporcionarem uma reflexão e um redimensionamento do processo terapêutico em relação a conceitos como **qualidade e dinâmica de voz** e, principalmente, em relação a compreensão da **voz como fenômeno histórico, cultural e social**, o que poderá beneficiar e favorecer a qualidade de vida dos sujeitos envolvidos.

Embora haja, ainda, um longo caminho a percorrer, para que as fronteiras conceituais e disciplinares, consagradas, ao longo do tempo, na prática fonoaudiológica, sejam transcendidas, confirma-se, a partir deste estudo, que cada produção vocal é única e singular, composta de elementos individuais e sociais, influenciados pelo contexto de produção, e que se manifestam fonicamente. Confirma-se, ainda, que essa manifestação é cientificamente

analisável, a partir de uma abordagem multidimensional e integrada, indicando ser este um complexo e fértil campo para pesquisas interdisciplinares para áreas como a Lingüística, a Fonoaudiologia, a Sociologia e a Medicina (Laringologia), entre outras. É, portanto, uma interface necessária e produtiva, que abre horizontes para outras investigações dentro da abordagem social, como propuseram LAVER e PITTAM, e como anunciava SAPIR.

REFERÊNCIAS BIBLIOGRÁFICAS

ABERCROMBIE, D. (1967) – **Elements of General Phonetics**. Edinburgh: Edinburgh University Press.

_____ (1968) – “Paralanguage”. In: LAVER, J. & HUTCHESON, S.

(1972) – **Communication in Face to Face Interaction**. Great Britain: Penguin Books, pp. 64-70.

AVAAZ Innovations Inc. (1995) – **Computerized Speech Research Environment 4.5** - Canada: Sound Scientific Solutions.

AVAAZ Innovations Inc. (1997) – **Computerized Speech Research Environment 4.5 – User’s Guide**. Canada: Sound Scientific Solutions

BAKHTIN, M. (1990) - **Marxismo e filosofia da linguagem**. São Paulo: Hucitec.

BARBOSA, P. A (1999) – “Revelar a estrutura rítmica de uma língua construindo máquinas falantes: pela integração da ciência e da tecnologia de fala”. In: SCARPA, E. M. (org.) – **Estudos de Prosódia**. Campinas, SP: Editora da Unicamp, pp. 21-52.

BATESON, G. (1972) - “Uma Teoria sobre a Brincadeira e a Fantasia”, traduzido por Lúcia Quental. In: RIBEIRO, B. T. & GARCEZ, P. M. (orgs.) – **Sociolingüística Interacional**. Porto Alegre: Age, 1998, pp.57-69.

BEHLAU, M. & PONTES, P. (1995) – **Avaliação e Tratamento das Disfonias**. São Paulo: Lovise.

BOLINGER, D. (1985) – **Intonation and its Parts: Melody in Spoken English**. Edward Arnold Ltd.

CAMARGO, Z. (2000) – “Avaliação Objetiva da Voz”. In: CARRARA-DE ANGELIS, E.; FURIA, C.L.B.; MOURÃO, L.F. & KOWALSKI, L.P. (orgs.) – **Atuação da Fonoaudiologia no Câncer de Cabeça e Pescoço**. São Paulo: Lovise, pp.175 -194.

- CAVALCANTI, D. & BOMPET, R. (2000) – “A Voz do Advogado: Atuação Vocal do Advogado no Curso de Oratória da Escola Superior de Advocacia”. In: FERREIRA, L.P. & COSTA, H.O. – **Voz Ativa – Falando sobre o Profissional da Voz**. São Paulo: Roca, pp. 181-188.
- COLEMAN, R. F. & MARKHAM, I.W. (1991) - “Normal Variations in Habitual Pitch”. In: **Journal of Voice**, vol 5, nº .2, pp. 173 –177.
- CHUN, R.Y.S. & NAVARRO, C. (1998) – “Voz Profissional. A Formação Acadêmica.” In: LACERDA, C.B.F. de & PANHOCA, I. (org.) – **Tempo de Fonoaudiologia II** . Taubaté/São Paulo: Cabral Editora Universitária.
- CHUN, R.Y.S. et al. (2000) – “Voz Profissional: grupos de Voz na Comunidade de Piracicaba” . In: FERREIRA, L.P. & COSTA, H.O. – **Voz Ativa – Falando sobre o Profissional da Voz**. São Paulo: Roca, pp.79 – 90.
- CRYSTAL, D. (1969) – **Prosodic Systems and Intonation in English**. London: Cambridge University Press.
- FANT, G. (1973) – **Speech Sounds and Features**. USA: The MIT Press.
- FERREIRA, L.. [org.] (1988) – **Trabalhando a Voz: Enfoques em Fonoaudiologia** . São Paulo:Summus.
- _____ (1993) - “A Avaliação de Voz: O Sentido poderia ser outro?”. In: FERREIRA, L.P. [org.] (1993) - **Um pouco de nós sobre voz**. Barueri, São Paulo: Pró-Fono, pp.29-38.
- _____ (2000) - “Voz e Sistema Sensório Motor Oral: O problema dos Objetos”. In: FREIRE, R.M. (org.) – **Fonoaudiologia: Seminários de Debates: volume 3**. São Paulo: Roca (Série Interfaces), pp.128-130.

FERREIRA, L .P. et al. [orgs.] (1995) - **Voz Profissional: O Profissional da Voz.**

Carapicuíba: Pró-Fono Departamento Editorial.

FERREIRA, L.P. ; ALGODOAL, M.J. e ANDRADA e SILVA, M. (1998) – A

“Avaliação de Voz na Visão (e no Ouvido) do Fonoaudiólogo: Saber o que se Procura para se Entender o que se Acha”. In: MARCHESAN, I.Q. et al. (orgs.)

– **Tópicos em Fonoaudiologia 1997/1998.** São Paulo: Lovise, Volume IV, pp. 393 – 414.

FERREIRA, L.P. & COSTA, H.O. [orgs.] (2000) – **Voz Ativa – Falando sobre o Profissional da Voz.** São Paulo: Roca,

FIGUEREDO NETO, L.E. (1988) – **O Início da Prática Clínica Fonoaudiológica na Cidade de São Paulo.** Dissertação de Mestrado em Distúrbios da Comunicação. Pontifícia Universidade Católica de São Paulo, São Paulo .

GAMA-ROSSI, A. J. A. (1999) – **Relações entre Desenvolvimento Lingüístico e Neuromotor: a Aquisição da Duração no Português Brasileiro.** Tese de Doutorado em Lingüística. Universidade Estadual de Campinas, São Paulo.

GOFFMAN, E. (1964) - “A situação negligenciada.” Traduzido por Pedro M. Garcez. In: RIBEIRO, B. T. & GARCEZ, P. M. (orgs.) – **Sociolingüística Interacional.** Porto Alegre: Age, 1998, pp.11-30.

_____ (1974) - **Frame Analysis.** New York: Harper and Row.

_____ (1979) – “Footing”. Traduzido por Beatriz Fontana. In: RIBEIRO, B. T. & GARCEZ, P. M. (orgs.) – **Sociolingüística Interacional.** Porto Alegre: Age, 1998, pp.70-97.

_____ (1981) - "Footing". In: **Forms of Talk.** Philadelphia . The University Press. pp. 124 -157.

- GROSJEAN, M. (1993) – “Polyphonie et Positions De La Sage-Femme Dans La Conduite De L`Accouchement”. In: COSNIER, J., GROSJEAN, M. e LACOSTE, M. (org.). **Approaches Interactionnistes des relations de soins**. Collection Ethologie et Psychologie des Communications Lyon: Presses Universitaires de Lyon, pp. 121 – 158.
- GUMPERZ, J. J. (1982) – “Convenções de contextualização”. In: RIBEIRO, B. T. & GARCEZ, P. M. (orgs.) – **Sociolingüística Interacional**. Porto Alegre: Age, 1998, pp.98-119.
- HOLLIEN, H. (2000) – “ The Concept of Ideal Voice Quality”. In: KENT, R. D. & BALL, M. J. (2000) – **Voice Quality Measurement**. California: Singular Publishing Group, pp.13-24
- KENT, R. D. & READ, C. (1992) - **The Acoustic Analysis of Speech**. Singular Publishing Group, INC.
- KENT, R. D. & BALL, M. J. (2000) – **Voice Quality Measurement**. California: Singular Publishing Group.
- KREIMAN et al. (1993) – “ Perceptual Evaluation of Voice Quality: Review, Tutorial, and a Framework for Future Research” . **Journal of Speech and Hearing Research**, Vol. 36, 21- 40, February.
- KREIMAN & GERRAT (2000) – “Measuring Vocal Quality”. In: KENT, R. D. & BALL, M. J. (2000) – **Voice Quality Measurement**. California: Singular Publishing Group, pp. 73-101.
- LABOV, W. (1972-8) – **Field Methods Used by the Project on Linguistic Change & Variation**. Supported by National Science Foundation. Philadelphia, University of Pennsylvania.

- LAYER, J. (1968) – “Voice Quality and Indexical Information”. In: LAYER, J. – **The Gift of Speech** . Edinburgh: Edinburgh University Press, 1991, pp.145-161.
- _____ (1980) - **The Phonetic Description of Voice Quality** . Cambridge: Cambridge University Press.
- _____ (1981) – “The Analysis of Vocal Quality: from the Classical Period to the Twentieth Century”. In: LAYER, J. – **The Gift of Speech** . Edinburgh: Edinburgh University Press, 1991, pp. 350 - 371.
- _____ (1991) – **The Gift of Speech** . Edinburgh: Edinburgh University Press.
- _____ (1994) – **Principles of Phonetics** . Great Britain: Cambridge University Press.
- _____ (1995) – “The Phonetic Description of Paralinguistic Phenomena. **Anais do XIII International Congress of Phonetic Sciences**, Stockholm.
- _____ (2000) – “Phonetic Evaluation of Voice Quality. In: KENT, R. D. & BALL, M. J. (edit.) – **Voice Quality Measurement**. California: Singular Publishing Group, pp. 37 - 48.
- LAYER, J. & TRUDGILL, P. (1979) - "Phonetics and Linguistic Markers in Speech". In: SCHERER, K. & GILES, H. - **Social Markers in Speech**. London: Cambridge University Press, pp. 1-32.
- LEDEN, H. Von (1982) – “A Cultural History of the Larynx and Voice”. In: SATALOFF, R. (1996) - **Professional Voice. The Art and Science of Clinical Care**. San Diego, California: Singular Publishing Group, pp.7-86.

- _____ (1982) - "A Cultural History of the Human Voice". In: SATALOFF, R. (1998) – **Voice perspectives**. San Diego, California: Singular Publishing Group Inc. Reprinted with permission, pp. 15 – 86.
- LOPES FILHO, O et al. (orgs.) – **Tratado de Fonoaudiologia**. São Paulo/SP: Editora Roca.
- LOPES, V.A.R. (1997) – **A Oratória nos Caminhos da Fonoaudiologia Estética**. Dissertação de Mestrado em Distúrbios da Comunicação. Pontifícia Universidade Católica de São Paulo, São Paulo.
- MADUREIRA, S. (1992) - **O Sentido do Som**. Tese de Doutorado em Lingüística Aplicada ao Ensino de Línguas. Pontifícia Universidade Católica de São Paulo, São Paulo.
- _____ (1999) – "Entoação e Síntese de fala: modelos e parâmetros." In: SCARPA, E. M. (org.) – **Estudos de Prosódia**. Campinas, SP: Editora da Unicamp, pp. 53-65.
- MALUFE, J. R. (1992) – **A Retórica da Ciência: uma Leitura de Goffman**. São Paulo: Educ.
- MARTZ, M.L.W. (1990) – **Vozes da Voz**. Dissertação de Mestrado em Distúrbios da Comunicação. Pontifícia Universidade Católica de São Paulo, São Paulo.
- _____ (1999) – "Algumas reflexões sobre a terapia de voz". **Distúrbios da Comunicação**. São Paulo, vol.10, nº. 2, pp. 205-211, jun.
- MATOS, S. (org.) (1998) - **Cantando, Brincando, Sonhando**. Piracicaba (SP), Colégio Piracicabano.
- MIRA MATEUS, MH. (1996) – "Fonologia" . In: FARIA, H. I. et al. (orgs.) – **Introdução à Lingüística Geral e Portuguesa**. Coleção Universitária, Série

- Lingüística, dirigida por Maria Raquel Delgado Martins. Lisboa: Editorial Caminho, pp. 171-199.
- NASCIMENTO, MA. & FERREIRA, L.P. (2000) – “A Voz no Contexto Não Terapêutico: Visão do Fonoaudiólogo”. In: FERREIRA, L.P. & COSTA, H.O. – **Voz Ativa – Falando sobre o Profissional da Voz**. São Paulo: Roca, pp.163-180.
- NAVARRO, C.A. (1996) – “O efeito do trabalho de estética para estudantes de graduação de Fonoaudiologia.” In: **Anais do Congresso Brasileiro de Fonoaudiologia**. Goiânia. p.19.
- OLIVEIRA, I.B. de (1988) – “A Educação Vocal na Radiodifusão”. In: : FERREIRA, L.P. (org.) **Trabalhando a Voz**. São Paulo: Summus, pp. 28 - 39.
- ORLANDI, E.P. (1988) – **Discurso & Leitura**. São Paulo: Cortez.
- PERRONI, M.C. (1996) - “O Que é Dado em Aquisição de Linguagem?”. In: CASTRO, M.F. (org.) – **O Método e o dado no Estudo da Linguagem**. Campinas, SP: Editora da UNICAMP, pp.15 – 30.
- PINTO, A.M.M. & FURCK, M.A.E. (1988) – “Projeto Saúde Vocal do Professor”. In: FERREIRA, L.P. (org.) **Trabalhando a Voz**. São Paulo: Summus, pp. 11-27.
- PITTAM, J. (1994) - **Voice in Social Interaction: An Interdisciplinary Approach**. Language and Language Behaviors, vol. 5. California: Sage Publications Inc.
- PLATÃO (1973) - **Diálogos de Platão - Teeteto e Crátilo**. Tradução de Carlos Alberto Nunes. V. IX , Universidade Federal do Pará (Coleção Amazônica. Série Farias Brito).

- QUINTEIRO, E.A. (1989). **Estética da Voz: Uma Voz para o ator**. São Paulo: Summus.
- QUINTILIANO, M.F. (1944) – **Instituições Oratórias**. Traduzido por Jerônimo Soares Barbosa. São Paulo: Edições Cultura, Tomo I.
- RAMECK, M.F. (2000) – “ O Professor de Educação Física de Academia de Ginástica: Comportamento e Perfil Vocal. In: FERREIRA, L.P. & COSTA, H.O. – **Voz Ativa – Falando sobre o Profissional da Voz**. São Paulo: Roca, pp.189 – 206.
- RIBEIRO, B. T. (1988) - **Coherence in Psychotic Discourse: Frame and Topic**. PhD. Dissertation, Georgetown University, Washington, D.C.
- RIBEIRO, B.T. - (1991) - "Papéis e Alinhamento no Discurso Psicótico". In: **Cadernos de Estudos Lingüísticos**, Campinas , vol 20, pp. 133-138, jan/jun.
- RIBEIRO, B. T. & GARCEZ, P. M. [orgs.] (1998) – **Sociolingüística Interacional**. Porto Alegre: Age.
- ROCHA FILHO (1997) – **Som e Ação na Narração de Futebol do Brasil**. Tese de Doutorado em Ciências, Universidade Estadual de Campinas, São Paulo.
- RODRIGUES, S. et al. (1996) – “Considerações sobre Voz Profissional Falada.” In: MARCHESAN, I.Q. et al. (orgs.). **Tópicos em Fonoaudiologia 1996**. São Paulo: Lovise, volume III, pp.701-712
- SAPIR, E. (1927) – “Speech as a Personality Trait”. In: LAVER, J. & HUTCHESON, S. (orgs.). **Communication in Face to Face Interaction**. Great Britain: Penguin Books, 1972, pp.71-81.

SATALOFF, R. (1996) - **Professional Voice. The Science and Art of Clinical Care**. 2nd. ed. San Diego, California: Singular Publishing Group.

SCHERER, K. (1979) – “Personality Markers in Speech”. In: SCHERER, K. & GILES, H. (Eds.) – **Social Markers In Speech**. Great Britain: Cambridge University Press, pp. 147-209.

SIMÃO, A.L.F. & CHUN, R.Y.S. (1997) - "Do Movimento, a Voz Surge Naturalmente". In: LACERDA, C. & PANHOCA, I. (org.). **Tempo de Fonoaudiologia**. Taubaté, SP: Cabral Editora Universitária, pp. 61 - 83.

SCIENTIFIC PROGRAM (1999) – 2nd. World Voice Congress. 5th. International Symposium on Phonosurgery. São Paulo.

SERVILHA, E. A . M. (2000) – **A Voz do Professor: Indicador para Compreensão da Dialogia no Processo Ensino-Aprendizagem**. Tese de Doutorado Em Ciências (Psicologia). Pontifícia Universidade Católica de Campinas, São Paulo.

SOCIEDADE BRASILEIRA DE LARINGOLOGIA E VOZ (2000) – **Manual de Orientação da Campanha Nacional do Dia da Voz**, São Paulo.

SOUZA e SILVA, M.C.P. (1995) – “Os Efeitos da ‘Não Negociação’ nas relações de trabalho”. **The ESpecialist**, vol. 15, n^{os}. 1 e 2, pp. 107-111.

SOUZA e SILVA, M.C.P. & CRESCITELLI, M. F. (1998) – “Retomando a Interrupção (*Getting Back to Interruption...*)”. **D.E.L.T.A.**, vol.14, nº. Especial, pp. 243-254.

STEMPLE, J.C. & GLAZE (1995) - **Clinical Voice Pathology: Theory and Management**. San Diego, California: Singular Publishing Group

TANNEN, D. & WALLAT, C. (1987) – “Enquadres Interativos e Esquemas de Conhecimento em Interação: Exemplos de um Exame/Consulta Médica”, traduzido por Parmênio Camurça Citó. In: RIBEIRO, B. T. & GARCEZ, P. M. (orgs.) – **Sociolingüística Interacional**. Porto Alegre: Age, 1998, pp.120 – 141.

_____ (1993) – “Interactive Frames and Knowledge Schemas in Interaction: Examples from a Medical Examination/Interview.” In: TANNEN, D. – **Framing in Discourse**. New York, London: Oxford University Press, pp.57-76.

VIGOTSKY, L.S. (1979) - **Pensamento e Linguagem**. Lisboa: Antídoto.

ANEXOS
