



UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
CENTRO DE CIÊNCIAS E TECNOLOGIA
COORDENAÇÃO DOS CURSOS DE PÓS-GRADUAÇÃO EM
ENGENHARIA ELÉTRICA

Tese de Doutorado

Estimulação Tátil Aplicada ao Ensino/Aprendizagem da Fala

Aléssio Trindade de Barros

Campina Grande – PB

Fevereiro - 2004

Estimulação Tátil Aplicada ao Ensino/Aprendizagem da Fala

Aléssio Trindade de Barros

Tese de Doutorado submetida à Coordenação dos Cursos de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Campina Grande, como parte dos requisitos necessários para obtenção do grau de Doutor em Ciências no domínio da Engenharia Elétrica.

Área de Concentração: Processamento da Informação

Raimundo Carlos Silvério Freire – Dr.

Gurdip Singh Deep – Ph.D

Orientadores

Campina Grande, Paraíba, Brasil



**B277e
2004**

Barros, Aléssio Trindade

**Estimulação tátil aplicada ao ensino/aprendizagem da fala /
Aléssio Trindade de Barros - Campina Grande - UFPG, 2004
169 p.: il.**

Inclui Bibliografia

Tese (Doutorado Engenharia Elétrica)UFPG/CCT/DEE.

**1.Engenharia de Reabilitação 2.Estimulação Tátil
3. Ensino da Fala**

CDU: 615.84:612-7

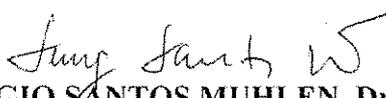
ESTIMULAÇÃO TÁTIL APLICADA AO ENSINO/APRENDIZAGEM DA FALA

ALÉSSIO TRINDADE BARROS

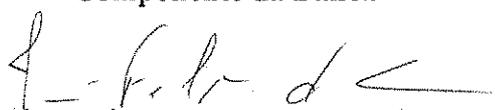
Tese Aprovada em 19.02.2004


RAIMUNDO CARLOS SILVÉRIO FREIRE, Dr., UFCG
Orientador


GURDIP SINGH DEEP, Ph.D., UFCG
Orientador


SÉRGIO SANTOS MUHLEN, Dr., UNICAMP
Componente da Banca


ALBERTO CLIQUET JUNIOR, Dr., USP-SCarlos / UNICAMP
Componente da Banca


JOSE FELÍCIO DA SILVA, Dr., UFPB
Componente da Banca


BENEDITO GUIMARÃES AGUIAR NETO, Dr.-Ing., UFCG
Componente da Banca

CAMPINA GRANDE – PB
Fevereiro - 2004

A Deus.

*A memória de meu pai, Pedro Matias de Barros, e
de meu avô, Zabulom Maribondo da Trindade.*

*A minha esposa, Maria José, e minha filha, Beatriz,
que está para nascer.*

*A minha mãe, Maria Violeta, minha avó, Augusta
Teixeira, e minha tia, Maria Augusta.*

Dedico humildemente este trabalho

Agradecimentos

Aos meus orientadores, professores Raimundo Carlos Silvério Freire e Gurdip Singh Deep, que desde os tempos de iniciação científica vêm contribuindo para minha formação, e pelo apoio, contribuições e paciência.

Aos professores Sérgio Mühlen, José Felício, Benedito Aguiar e Alberto Cliquet, pelas considerações e contribuições no Exame de Qualificação;

À André Trindade, Maria José e João Negrão por contribuírem gravando vozes e em testes de percepção.

À Eriádnes, fonoaudióloga da FUNAD, pelas discussões e contribuições.

Ao professor Hiran de Melo, pelo convívio agradável, pelas discussões e trabalhos conjuntos realizados neste período em Campina Grande.

Aos amigos do LIMC e da UFCG, Rômulo(Ana Cardoso), Antenor, Franklin, Cida, Guilherme, Lígia, Ivan, Brunelli, Hermano, Ronny, Ricardo, Lívia, Maxwell, Wendell, Lucilene, Will, Lucélio, Thayrone, Moacir, Zurita, Marcelo, Joelle, Denis, Figueiredo, Francisco, Lucélio, Márcia e os professores Gutemberg e Benedito Luciano pelo apoio, amizade e descontração.

À Ângela, Eleonôra e Pedro, pelo apoio e dedicação de sempre.

Aos amigos Ilton, Sabiniano, Fechine e Walmeran, da área de elétrica do CEFET-PB, pelo constante apoio e incentivo durante este curso.

A todos os professores da área de eletrônica do CEFET-PB, em especial ao prof. Dilço, na época coordenador da área, que permitiram e apoiaram meu afastamento para a realização deste curso.

À Universidade Federal de Campina Grande, ao Centro Federal de Educação Tecnológica da Paraíba e ao CNPq.

A todos que, direta ou indiretamente, me incentivaram no decorrer deste trabalho.

Muito Obrigado!

Resumo

A informação acústica pode ser transmitida aos surdos por amplificadores acústicos, implantes cocleares, ou através de sistemas de substituição sensorial. A escolha depende das características de perda auditiva e das necessidades do portador da deficiência. Os sistemas táteis de substituição sensorial vêm sendo desenvolvidos desde o início do século XX, e podem basear-se em estimulação vibrotátil, mas não existem relatos na literatura técnica de sistemas táteis desenvolvidos especificamente para emprego no auxílio ao ensino/aprendizagem da fala. Neste trabalho é apresentada uma investigação sobre a codificação de parâmetros de voz em estimuladores vibrotáteis, no contexto específico do ensino da produção da fala. Foram propostas e analisadas várias formas de codificação (transformações), para as quais foram desenvolvidos o respectivo modelo matemático. Com base no modelo matemático apresentado foram feitas diversas simulações, para subsidiar uma análise comparativa entre as transformações propostas. Ao final, as transformações propostas foram analisadas através dos resultados de experimentos de percepção de estímulos vibrotáteis. Também foi analisada a escolha de parâmetros de voz adequada para modulação direta de estimuladores monocanais, no contexto do português.

Abstract

Acoustic information can be transmitted to the deaf, using acoustic amplifiers, cochlear implants and sensory substitution systems. The choice depends on characteristics and degree of hearing impairment. Tactile stimulation devices can be used as sensory substitute and have been developed since the beginning of the 20th century. In this thesis the emphasis is on the investigation of the transformation of voice parameters into frequency of the pulsed tactile stimulus, in the context of the speech hearing on the part of the deaf for the Portuguese language. Different transformation schemes have been formulated and analyzed by simulation studies. The proposed transformations were also analyzed by vibrotactile perception sensitivity experiments.

ÍNDICE

1	INTRODUÇÃO.....
1.1	MOTIVAÇÃO	4
1.2	OBJETIVO	9
1.3	ORGANIZAÇÃO DO DOCUMENTO.....	10
2	COMUNICAÇÃO ORAL E DEFICIÊNCIA AUDITIVA	12
2.1	COMUNICAÇÃO E LINGUAGEM	13
2.2	VOZ, FALA E AUDIÇÃO	14
2.3	A PRODUÇÃO DA FALA.....	15
2.3.1	Vogais	19
2.3.2	Consoantes	20
2.3.3	Fatores que Interferem no Processo Articulatório	22
2.4	O APRENDIZADO DA FALA	24
2.5	DEFICIÊNCIA AUDITIVA.....	25
2.6	EDUCAÇÃO DE DEFICIENTES AUDITIVOS	27
2.6.1	Desenvolvimento da Fala da Criança Surda	29
2.6.1.1	Fatores Pedagógicos Envolvendo o Ensino da Fala	30
2.6.1.2	Dispositivos para Auxílio ao Ensino da Fala.....	31
3	ESTÍMULOS VIBROTÁTEIS	33
3.1	PERCEPÇÃO DE ESTÍMULOS TÁTEIS	34
3.2	CONSIDERAÇÕES SOBRE ESTÍMULOS	37
3.3	COMUNICAÇÃO DE INFORMAÇÕES ACÚSTICAS	37
3.3.1	Percepção de Informações Acústicas	38
3.3.2	Estimulação acústica do ambiente (background)	39
3.3.3	Detecção.....	39
3.3.4	Discriminação e Identificação	40
3.3.4.1	Envoltória do sinal	40
3.3.4.2	Parâmetros Espectrais	41
3.3.5	Localização	41
3.3.6	Processamento de Alto Nível.....	42
3.4	DISPOSITIVOS TÁTEIS DE UM CANAL.....	43
3.5	DISPOSITIVOS TÁTEIS MULTICANAL.....	44
3.5.1	Número de canais.....	45
3.6	ESTIMULAÇÃO VIBROTÁTIL: QUESTÕES DO ENSINO DA FALA	45
3.6.1	A Questão da Intensidade da Fala.....	46
3.6.2	A Questão da Respiração	47
3.6.3	A Questão da Entonação	49

3.6.4	A Questão da Articulação.....	51
3.6.4.1	As Vogais.....	52
3.6.4.2	As Consoantes.....	55
4	CODIFICAÇÃO DE PARÂMETROS EM ESTÍMULOS TÁTEIS	60
4.1	INTRODUÇÃO.....	61
4.2	TRANSFORMAÇÕES DE UM PARÂMETRO	63
4.2.1.1	Grau de Dificuldade.....	64
4.2.1.2	Cálculo de S_{max} e p_1 para a Transformação Linear.....	65
4.2.1.3	Cálculo de S_{max} e p_1 para a Transformação Logarítmica.....	65
4.2.1.4	Cálculo de S_{max} e p_1 para a Transformação Cúbica	66
4.2.1.5	Cálculo de S_{max} e p_1 para a Transformação Quadrática	66
4.2.1.6	Dependência de p_1 em função de p_2	66
4.2.1.7	Cálculo de S_{1max} e S_{2max}	67
4.2.1.8	Simulações.....	67
4.3	TRANSFORMAÇÕES DE DOIS PARÂMETROS	72
4.3.1	Transformação Quadrática de Dois Parâmetros	73
4.3.1.1	Simulações.....	74
4.3.1.2	Codificação Quadrática com Coeficientes de Ajuste	78
4.3.1.2.1	Simulação.....	79
4.3.2	Transformações Linear e Cúbica de Dois Parâmetros	81
4.3.2.1	Cálculo de S_{max} , p_{11} e p_{12} para a Transformação Linear	82
4.3.2.2	Cálculo de S_{max} , p_{11} e p_{12} para a Transformação Cúbica....	82
4.3.2.3	Simulações.....	82
4.4	TRANSFORMAÇÕES DE TRÊS PARÂMETROS.....	85
4.4.1	Cálculo de S_{max}	86
4.5	DISCUSSÕES.....	87
5	CODIFICAÇÃO E EXPERIMENTOS DE PERCEPÇÃO	88
5.1	INTRODUÇÃO.....	89
5.2	ESTIMULADOR VIBROTÁTIL	90
5.3	PARÂMETROS DE VOZ	92
5.3.1	Amplitude Média.....	92
5.3.2	Taxa de Cruzamentos por Zero.....	93
5.4	SINAL DE VOZ	93
5.4.1	Metodologia.....	95
5.4.1.1	Histogramas	97
5.4.2	Codificação Mista dos Parâmetros AVG e TCZ	101
5.4.3	Experimentos de Percepção	104
5.4.3.1	Metodologia	105
5.4.3.2	Codificação do Sinal de Estimulação Vibrotátil	106
5.4.3.3	Resultados Obtidos.....	107

5.4.4	Discussões.....	109
5.5	TRANSFORMAÇÕES DE UM PARÂMETRO.....	110
5.5.1	Frequência Fundamental.....	110
5.5.2	Experimentos de Percepção.....	110
5.5.2.1	Metodologia.....	111
5.5.2.2	Sinal de Estimulação Vibrotátil.....	113
5.5.2.3	Resultados Obtidos.....	114
5.5.3	Discussões.....	122
5.6	TRANSFORMAÇÕES DE DOIS PARÂMETROS.....	124
5.6.1	Vogais.....	124
5.6.2	Parâmetros Propostos.....	127
5.6.2.1	Parâmetro 1.....	127
5.6.2.2	Parâmetro 2.....	128
5.6.3	Coleta de Dados e Extração de Parâmetros.....	129
5.6.4	Análise dos Parâmetros.....	129
5.6.4.1	Parâmetro 1.....	129
5.6.4.2	Parâmetro 2.....	130
5.6.5	Experimentos de Percepção.....	133
5.6.5.1	Metodologia.....	133
5.6.5.2	Sinal de Estimulação Vibrotátil.....	136
5.6.5.3	Resultados Obtidos.....	136
5.7	DISCUSSÕES.....	139
6	CONCLUSÕES E SUGESTÕES.....	141
6.1	INTRODUÇÃO.....	142
6.2	SINAL DE VOZ.....	143
6.3	TRANSFORMAÇÕES DE UM PARÂMETRO.....	144
6.4	TRANSFORMAÇÕES DE MAIS DE UM PARÂMETRO.....	145
6.5	CONTRIBUIÇÕES.....	145
6.6	SUGESTÃO PARA TRABALHOS FUTUROS.....	146
6.6.1	Integração Estimulação Tátil e Visual.....	146
6.6.2	O desenvolvimento do Sistema para Uso com Surdos: Parametrização e Definição de Alvos.....	146
6.6.3	O Desenvolvimento de Sistemas Específicos para Surdocegos e Multideficientes.....	147
	BIBLIOGRAFIA.....	148
	ANEXOS.....	148

LISTA DE FIGURAS

FIGURA 1 - DIAGRAMA ESQUEMÁTICO DOS ÓRGÃOS ENVOLVIDOS NA PRODUÇÃO DA FALA	16
FIGURA 2 - DIAGRAMA DE BLOCOS DO APARELHO FONADOR.....	17
FIGURA 3 - PROCESSO DE PRODUÇÃO DA FALA (MARTINS, 1988)	18
FIGURA 4 - POSICIONAMENTO DOS ÓRGÃOS ARTICULADORES (LÍNGUA E DA MANDÍBULA) DE ACORDO COM O TIPO DE VOGAL (RUSSO & BEHLAU, 1993)	19
FIGURA 5 - REPRESENTAÇÃO ESQUEMÁTICA DOS PONTOS DE ARTICULAÇÃO DAS CONSOANTES DO PORTUGUÊS: 1. BILABIAL, 2. LÁBIODENTAL, 3. LINGUODENTAL, 4. LINGUOPALATAL, 5. LINGUOVELAR (RUSSO&BEHLAU, 1993)	21
FIGURA 6 - SINAL DE ESTIMULAÇÃO VIBROTÁTIL TÍPICO	37
FIGURA 7 – ESTIMULADOR VIBROTÁTIL MULTICANAL TÍPICO	44
FIGURA 8 - REPRESENTAÇÃO DE SINAL NA SAÍDA DE UMA MATRIZ DE ESTIMULADORES, PARA O SEGMENTO DE VOZ /BI/, DA LÍNGUA INGLESA. CADA PONTO REPRESENTA UM ESTIMULADOR. CADA PONTO PRETO REPRESENTA UM ESTIMULADOR ATIVO.	53
FIGURA 9 - MÉDIA DAS FREQUÊNCIAS DOS TRÊS PRIMEIROS FORMANTES, DAS VOGAIS ORAIS DO PORTUGUÊS BRASILEIRO (RUSSO&BEHLAU, 1993).....	55
FIGURA 10 - TRANSFORMAÇÕES DISCRETAS COM INDICAÇÃO DE ACERTO DE ALVO	62
FIGURA 11 - (A) – (H); SÃO APRESENTADOS RESULTADOS DE SIMULAÇÃO DA DEPENDÊNCIA DE $F_T(T)$ EM FUNÇÃO DE $P_E(T)$ PARA AS TRANSFORMAÇÕES PROPOSTAS, OBTIDOS A PARTIR DAS EQS. (3), (4), (5) E (6). FORAM PLOTADAS CURVAS PARA 10 VALORES DIFERENTES DE S. (A) TRANSFORMAÇÃO LINEAR, COM VALORES DE P_1 OBTIDOS A PARTIR DA EQ. (14); (B) TRANSFORMAÇÃO LINEAR, COM VALORES S_{1MAX} E S_{2MAX} OBTIDOS NA TABELA 3; (C) TRANSFORMAÇÃO LOGARÍTMICA, COM VALORES DE P_1 OBTIDOS A PARTIR DA EQ. (16); (D) TRANSFORMAÇÃO LOGARÍTMICA, COM VALORES S_{1MAX} E S_{2MAX} OBTIDOS NA TABELA 3; (E) TRANSFORMAÇÃO CÚBICA, COM VALORES DE P_1 OBTIDOS A PARTIR DA EQ. (18); (F) TRANSFORMAÇÃO CÚBICA, COM VALORES S_{1MAX} E S_{2MAX} OBTIDOS NA TABELA 3; (G) TRANSFORMAÇÃO QUADRÁTICA, COM VALORES DE P_1 OBTIDOS A PARTIR DA EQ. (20); (H) TRANSFORMAÇÃO QUADRÁTICA, COM VALORES S_{1MAX} E S_{2MAX} OBTIDOS NA TABELA 3.	71
FIGURA 12 - DEPENDÊNCIA DE $F_T(T)$ EM FUNÇÃO DE $P_{E1}(T)$ E $P_{E2}(T)$ PARA A TRANSFORMAÇÃO QUADRÁTICA	75
FIGURA 13 - $F_T(T)$ EM FUNÇÃO DE $P_{E1}(T)$ E $P_{E2}(T)$ PARA A TRANSFORMAÇÃO QUADRÁTICA(2-D).....	76
FIGURA 14 - DEPENDÊNCIA DE $F_T(T)$ EM FUNÇÃO DE $P_{E1}(T)$ E $P_{E2}(T)$ PARA A TRANSFORMAÇÃO QUADRÁTICA COM $(P_{22}-P_{12}) > (P_{21}-P_{11})$	76

FIGURA 15 - $F_T(T)$ EM FUNÇÃO DE $P_{E1}(T)$ E $P_{E2}(T)$ PARA A TRANSFORMAÇÃO QUADRÁTICA(2-D) COM $(P_{22}-P_{12}) > (P_{21}-P_{11})$	77
FIGURA 16 - $F_T(T)$ EM FUNÇÃO DE $P_{E1}(T)$ E $P_{E2}(T)$ PARA A TRANSFORMAÇÃO QUADRÁTICA(2-D) COM B DADO POR EQ. (31).....	79
FIGURA 17 - $F_T(T)$ EM FUNÇÃO DE $P_{E1}(T)$ E $P_{E2}(T)$ PARA A TRANSFORMAÇÃO QUADRÁTICA(2-D) COM B DADO POR EQ. (32).....	80
FIGURA 18 - $F_T(T)$ EM FUNÇÃO DE $P_{E1}(T)$ E $P_{E2}(T)$ PARA A TRANSFORMAÇÃO QUADRÁTICA(2-D) COM B DADO POR EQ. (33).....	80
FIGURA 19 - DEPENDÊNCIA DE $F_T(T)$ EM FUNÇÃO DE $P_{E1}(T)$ E $P_{E2}(T)$ PARA A TRANSFORMAÇÃO LINEAR.....	83
FIGURA 20 - $F_T(T)$ EM FUNÇÃO DE $P_{E1}(T)$ E $P_{E2}(T)$ PARA A TRANSFORMAÇÃO LINEAR(2-D) COM B DADO POR EQ. (33).....	84
FIGURA 21 - DEPENDÊNCIA DE $F_T(T)$ EM FUNÇÃO DE $P_{E1}(T)$ E $P_{E2}(T)$ PARA A TRANSFORMAÇÃO CÚBICA.....	84
FIGURA 22 - $F_T(T)$ EM FUNÇÃO DE $P_{E1}(T)$ E $P_{E2}(T)$ PARA A TRANSFORMAÇÃO CÚBICA(2-D) COM B DADO POR EQ. (33).....	85
FIGURA 23 - CONFIGURAÇÕES DE ESTIMULADORES VIBROTÁTEIS.....	90
FIGURA 24 - TRANSDUTOR INERCIAL COMERCIAL (VBW32, DA AUDIOLOGICAL ENGINEERING).....	91
FIGURA 25 - HISTOGRAMAS DA DISTRIBUIÇÃO DOS PARÂMETROS AVG E TCZ (/T, D, N/).....	97
FIGURA 26 - HISTOGRAMAS DA DISTRIBUIÇÃO DOS PARÂMETROS AVG E TCZ (/P, B, M/).....	98
FIGURA 27 - HISTOGRAMAS DA DISTRIBUIÇÃO DOS PARÂMETROS AVG E TCZ (/K, G, LH/).....	98
FIGURA 28 - HISTOGRAMAS DA DISTRIBUIÇÃO DOS PARÂMETROS AVG E TCZ (/S, Z/).....	99
FIGURA 29 - HISTOGRAMAS DA DISTRIBUIÇÃO DOS PARÂMETROS AVG E TCZ (/CH, J/)....	99
FIGURA 30 - HISTOGRAMAS DA DISTRIBUIÇÃO DOS PARÂMETROS AVG E TCZ (/F, V/)....	100
FIGURA 31 - HISTOGRAMAS DA DISTRIBUIÇÃO DOS PARÂMETROS AVG E TCZ (/L, R/)....	100
FIGURA 32 - DIAGRAMA DE BLOCOS DO SISTEMA DE CODIFICAÇÃO DE SINAL DE ESTIMULAÇÃO VIBROTÁTIL PROPOSTO.....	102
FIGURA 33 - - HISTOGRAMA GERAL PARA OS FONEMAS FORMADOS PELAS ARTICULAÇÕES /S/, /Z/ E /CH/, /J/, DO PARÂMETRO AVG.....	103
FIGURA 34 - - HISTOGRAMA GERAL PARA OS FONEMAS FORMADOS PELAS ARTICULAÇÕES /S/, /Z/ E /CH/, /J/, DA ARQUITETURA PROPOSTA, PARA UM CHAVEAMENTO COM TRX = 100.....	104
FIGURA 35 - RESULTADOS DE TESTES DE PERCEPÇÃO DE DISCRIMINAÇÃO ENTRE FONEMAS /P/, /B/ E /M/.....	107

FIGURA 36 - RESULTADOS DE TESTES DE PERCEPÇÃO DE DISCRIMINAÇÃO ENTRE FONEMAS /K/, /G/ E /LH/	108
FIGURA 37 - RESULTADOS DE TESTES DE PERCEPÇÃO DE DISCRIMINAÇÃO ENTRE FONEMAS /S/-/Z/, /CH/-/J/, /L/-/R/ E /F/-/V/.....	108
FIGURA 38 - RESULTADOS DE TESTES DE PERCEPÇÃO DE DISCRIMINAÇÃO ENTRE FONEMAS /S/-/Z/ E /CH/-/J/, COMPARANDO-SE OS PARÂMETROS AVG E PRP	109
FIGURA 39- SISTEMA TÍPICO PARA TREINAMENTO/APRENDIZAGEM DA FALA	111
FIGURA 40 – FLUXOGRAMA DO PROCEDIMENTO EXPERIMENTAL.....	113
FIGURA 41 - MÉDIA E DESVIO PADRÃO DA FREQUÊNCIA DO SINAL DE ESTIMULAÇÃO AUDITIVA PERCEBIDA COMO ACERTO DE ALVO PARA AS TRANSFORMAÇÕES PROPOSTAS.....	115
FIGURA 42 - MÉDIA E DESVIO PADRÃO DA FREQUÊNCIA FUNDAMENTAL CORRESPONDENTE À FREQUÊNCIA DO SINAL DE ESTIMULAÇÃO AUDITIVA PERCEBIDA COMO ACERTO DE ALVO PARA AS TRANSFORMAÇÕES PROPOSTAS	115
FIGURA 43 - MÉDIA E DESVIO PADRÃO DA FREQUÊNCIA DO SINAL DE ESTIMULAÇÃO VIBROTÁTIL PERCEBIDA COMO ACERTO DE ALVO PARA 8 DIFICULDADES DA TRANSFORMAÇÃO LINEAR, ALÉM DE RESULTADO OBTIDO PARA ESTIMULAÇÃO AUDITIVA	116
FIGURA 44 - MÉDIA E DESVIO PADRÃO DA FREQUÊNCIA FUNDAMENTAL CORRESPONDENTE À FREQUÊNCIA DO SINAL DE ESTIMULAÇÃO VIBROTÁTIL PERCEBIDA COMO ACERTO DE ALVO PARA 8 DIFICULDADES DA TRANSFORMAÇÃO LINEAR, ALÉM DE RESULTADO OBTIDO PARA ESTIMULAÇÃO AUDITIVA	116
FIGURA 45 - MÉDIA E DESVIO PADRÃO DO NÚMERO DE PASSOS EXCEDENTES REALIZADOS PARA INDICAÇÃO DE ACERTO DE ALVO PARA 8 DIFICULDADES DA TRANSFORMAÇÃO LINEAR, ALÉM DE RESULTADO OBTIDO PARA ESTIMULAÇÃO AUDITIVA	117
FIGURA 46 - MÉDIA E DESVIO PADRÃO DA FREQUÊNCIA DO SINAL DE ESTIMULAÇÃO VIBROTÁTIL PERCEBIDA COMO ACERTO DE ALVO PARA 8 DIFICULDADES DA TRANSFORMAÇÃO LOGARÍTMICA, ALÉM DE RESULTADO OBTIDO PARA ESTIMULAÇÃO AUDITIVA	117
FIGURA 47 - MÉDIA E DESVIO PADRÃO DA FREQUÊNCIA FUNDAMENTAL CORRESPONDENTE À FREQUÊNCIA DO SINAL DE ESTIMULAÇÃO VIBROTÁTIL PERCEBIDA COMO ACERTO DE ALVO PARA 8 DIFICULDADES DA TRANSFORMAÇÃO LOGARÍTMICA, ALÉM DE RESULTADO OBTIDO PARA ESTIMULAÇÃO AUDITIVA.....	118
FIGURA 48 - MÉDIA E DESVIO PADRÃO DO NÚMERO DE PASSOS EXCEDENTES REALIZADOS PARA INDICAÇÃO DE ACERTO DE ALVO PARA 8 DIFICULDADES DA TRANSFORMAÇÃO LOGARÍTMICA, ALÉM DE RESULTADO OBTIDO PARA ESTIMULAÇÃO AUDITIVA	118
FIGURA 49 - MÉDIA E DESVIO PADRÃO DA FREQUÊNCIA DO SINAL DE ESTIMULAÇÃO VIBROTÁTIL PERCEBIDA COMO ACERTO DE ALVO PARA 8 DIFICULDADES DA TRANSFORMAÇÃO CÚBICA, ALÉM DE RESULTADO OBTIDO PARA ESTIMULAÇÃO AUDITIVA	119

FIGURA 50 - MÉDIA E DESVIO PADRÃO DA FREQUÊNCIA FUNDAMENTAL CORRESPONDENTE À FREQUÊNCIA DO SINAL DE ESTIMULAÇÃO VIBROTÁTIL PERCEBIDA COMO ACERTO DE ALVO PARA 8 DIFICULDADES DA TRANSFORMAÇÃO CÚBICA, ALÉM DE RESULTADO OBTIDO PARA ESTIMULAÇÃO AUDITIVA.....	119
FIGURA 51 - MÉDIA E DESVIO PADRÃO DO NÚMERO DE PASSOS EXCEDENTES REALIZADOS PARA INDICAÇÃO DE ACERTO DE ALVO PARA 8 DIFICULDADES DA TRANSFORMAÇÃO CÚBICA, ALÉM DE RESULTADO OBTIDO PARA ESTIMULAÇÃO AUDITIVA.....	120
FIGURA 52 - MÉDIA E DESVIO PADRÃO DA FREQUÊNCIA DO SINAL DE ESTIMULAÇÃO VIBROTÁTIL PERCEBIDA COMO ACERTO DE ALVO PARA 8 DIFICULDADES DA TRANSFORMAÇÃO QUADRÁTICA, ALÉM DE RESULTADO OBTIDO PARA ESTIMULAÇÃO AUDITIVA.....	120
FIGURA 53 - MÉDIA E DESVIO PADRÃO DA FREQUÊNCIA FUNDAMENTAL CORRESPONDENTE À FREQUÊNCIA DO SINAL DE ESTIMULAÇÃO VIBROTÁTIL PERCEBIDA COMO ACERTO DE ALVO PARA 8 DIFICULDADES DA TRANSFORMAÇÃO QUADRÁTICA, ALÉM DE RESULTADO OBTIDO PARA ESTIMULAÇÃO AUDITIVA.....	121
FIGURA 54 - MÉDIA E DESVIO PADRÃO DO NÚMERO DE PASSOS EXCEDENTES REALIZADOS PARA INDICAÇÃO DE ACERTO DE ALVO PARA 8 DIFICULDADES DA TRANSFORMAÇÃO QUADRÁTICA, ALÉM DE RESULTADO OBTIDO PARA ESTIMULAÇÃO AUDITIVA.....	121
FIGURA 55 - DIAGRAMA DE BLOCOS DO ESQUEMA DO 1º PARÂMETRO PROPOSTO PARA AUXILIO AO ENSINO DE VOGAIS.....	128
FIGURA 56 - SINAL DE VOZ DE UMA VOGAL TÍPICA E SUA CORRESPONDENTE DISTRIBUIÇÃO ESPECTRAL.....	128
FIGURA 57 - GRÁFICO GERAL, DO PARÂMETRO I, PARA TODAS AS VOGAIS ESTUDADAS.....	130
FIGURA 58 - VARIAÇÃO DA TCZ DAS VOGAIS POSTERIORES.....	131
FIGURA 59 - VARIAÇÃO DA TCZ DAS VOGAIS ANTERIORES.....	131
FIGURA 60 - VARIAÇÃO DE TCZ DAS VOGAIS ANTERIORES.....	132
FIGURA 61 - VARIAÇÃO DE TCZ DAS VOGAIS POSTERIORES.....	133

LISTA DE TABELAS

TABELA 1 - QUADRO DAS CONSOANTES DO PORTUGUÊS, CLASSIFICADAS DE ACORDO COM O PONTO E MODO DE ARTICULAÇÃO.....	22
TABELA 2 - CARACTERÍSTICAS BÁSICAS DA PERCEÇÃO VIBROTÁTIL.....	36
TABELA 3 - EQUAÇÕES DE S_{1MAX} E S_{2MAX} PARA CADA TRANSFORMAÇÃO.....	68
TABELA 4 - CONDIÇÕES DE INFLUÊNCIA DE CADA PARÂMETRO NAS TRANSFORMAÇÕES DADAS PELAS EQUAÇÕES (45), (46) E (47).....	86
TABELA 5 - GRUPOS DE FONEMAS DA LÍNGUA PORTUGUESA CUJA DISTINÇÃO VIA LEITURA LABIAL APRESENTA DIFICULDADES.....	95
TABELA 6 - GRUPOS DE FONEMAS DA TABELA 4 CLASSIFICADOS QUANTO AO MODO E VOZEAMENTO.....	95
TABELA 7 - CONJUNTO DE PALAVRAS UTILIZADAS.....	96
TABELA 8 - CONJUNTO DE VISEMES.....	105
TABELA 9 - CONJUNTO DE OPÇÕES DE ESTÍMULOS DOS VISEMES COM TRÊS FONEMAS.....	106
TABELA 10 - CONJUNTO DE OPÇÕES DE ESTÍMULOS DOS VISEMES COM DOIS FONEMAS.....	106
TABELA 11 - MÉDIAS DAS FREQUÊNCIAS DAS TRÊS PRIMEIRAS FORMANTES: F1, F2 E F3, DAS VOGAIS ORAIS DO PORTUGUÊS BRASILEIRO, PARA HOMENS, MULHERES E CRIANÇAS (DADOS EXTRAÍDOS DE RUSSO & BEHLAU, 1993).....	125
TABELA 12 - MATRIZ DE CONFUSÃO DO PARÂMETRO 1, TRANSFORMAÇÃO QUADRÁTICA, PARA AS VOGAIS POSTERIORES, ANTERIORES E MEDIAL.....	136
TABELA 13 - MATRIZ DE CONFUSÃO DO PARÂMETRO 2 (TCZ), TRANSFORMAÇÃO QUADRÁTICA, PARA AS VOGAIS É, I E Ê.....	137
TABELA 14 - MATRIZ DE CONFUSÃO DO PARÂMETRO 2 (TCZ), TRANSFORMAÇÃO QUADRÁTICA, PARA AS VOGAIS Ó, Ô E U.....	137
TABELA 15 - PARÂMETROS UTILIZADOS NA CODIFICAÇÃO DAS TRANSFORMAÇÕES LINEAR, CÚBICA E QUADRÁTICA DE DOIS PARÂMETROS.....	137
TABELA 16- MATRIZ DE CONFUSÃO DOS PARÂMETROS 1 E 2 (TCZ), TRANSFORMAÇÃO LINEAR, PARA AS VOGAIS É, I, Ó, U Ê, Ô.....	138
TABELA 17- MATRIZ DE CONFUSÃO DOS PARÂMETROS 1 E 2 (TCZ), TRANSFORMAÇÃO QUADRÁTICA, PARA AS VOGAIS É, I, Ó, U Ê, Ô.....	138
TABELA 18 - MATRIZ DE CONFUSÃO DOS PARÂMETROS 1 E 2 (TCZ), TRANSFORMAÇÃO CÚBICA, PARA AS VOGAIS É, I, Ó, U Ê, Ô.....	138

GLOSSÁRIO

Alfabeto fonético: conjunto de símbolos que se destinam a representar graficamente os sons da linguagem.

Articulações: Configuração própria das cavidades supraglotais que assumem funções de filtro, atuando como cavidades de ressonância

Formantes: Zonas de frequência intensificadas pelas cavidades de ressonância. Os formantes são específicos para cada vogal e definem a sua qualidade. A qualidade vocal é essencialmente determinada pelos dois primeiros formantes, que permitem o reconhecimento da vogal.

Frequência Fundamental: (F_0) é a frequência de repetição de uma forma de onda complexa. Corresponde a taxa de vibração das cordas vocais. Note-se que os sons surdos não possuem F_0 definida.

Sons Surdos: Sons produzidos sem a vibração das cordas vocais.

Tadoma: método para comunicação com surdocegos no qual o "ouvinte" compreende a fala do locutor colocando sua mão na face do mesmo. A compreensão da fala ocorre pela percepção tátil dos movimentos das articulações, vibrações nasais e no pescoço e entrada e saída de ar na boca do locutor.

Taxa de Cruzamentos por Zero: Parâmetro bastante utilizado em aplicações de processamento digital de sinais de voz. Indica o número de vezes que as amostras de um sinal, em um determinado segmento, cruzam o zero.

Tom (pitch): O tom é a frequência fundamental subjetiva, isto é, "ouvida".

Viseme: é uma imagem facial genérica que pode ser usada para descrever um som particular. O viseme é o equivalente visual do fonema. Usando os visemes o surdo pode ver os sons, podendo fazer leitura labial.

1

INTRODUÇÃO

O ser humano vive em um processo de interação contínua entre os acontecimentos exteriores e sua vivência interior, que resulta nos seus atos cotidianos, e em uma escala maior, no desenvolvimento da sociedade. Este processo permite a vida em comunidade e a consciência do mundo como ele é, sendo essencial para a existência do homem e de sua sociedade.

O deficiente sensorial tem uma percepção dos acontecimentos exteriores alterada com relação aos outros indivíduos, o que influi em sua capacidade de comunicação. Um ser humano com deficiência sensorial possui dificuldades de desenvolvimento pessoal e de adaptação social, a depender do tipo e grau de sua deficiência.

Dentre as deficiências sensoriais, as da visão e da audição são as que mais influem na capacidade de integração social de um indivíduo, sendo que a deficiência auditiva, especificamente, influi na sua capacidade de comunicação oral. Um indivíduo com deficiência auditiva, além de não ser capaz de receber a informação acústica, tem dificuldade para aprender a falar naturalmente, pois o aprendizado da fala passa pelo automonitoramento da própria voz (Araújo & Lima, 1999; Szeto & Christensen, 1988).

A dificuldade de comunicação oral entre surdos e ouvintes é um problema que ao longo dos tempos relegou indivíduos com deficiência auditiva a condições inferiorizadas na sociedade. No entanto, a partir do século XVII começaram a ser desenvolvidos métodos para o ensino da fala a surdos, denominados métodos orais, e treinamentos para recepção da comunicação verbal através da leitura labial. Paralelamente também foram desenvolvidas linguagens gestuais, mais naturais para a comunicação entre os surdos (Skliar, 2001). Atualmente existem filosofias educacionais que se diferenciam quanto à aplicação das linguagens oral e gestual no ensino do surdo. Embora a aquisição da língua oral seja reconhecidamente difícil e demorada, a mesma é recomendada por quase todos os estudiosos, de quase todas as metodologias propostas para a educação de deficientes auditivos (Araújo, 2000).

Visando diminuir o problema do surdo quanto à recepção da informação acústica, pode-se transmitir a informação sonora através de várias formas: 1) com amplificadores acústicos; 2) com implantes cocleares; e 3) com sistemas de substituição sensorial.

Os amplificadores acústicos são dispositivos de ajuda à escuta de uso mais comum, estando restritos aos deficientes que possuem audição residual. Já os implantes cocleares e sistemas de substituição sensorial são indicados para indivíduos com perdas auditivas profundas.

A definição do dispositivo de auxílio à recepção da informação acústica adequado para determinado surdo envolve a análise de diversos aspectos, como a faixa etária, os custos envolvidos na implantação do dispositivo, a existência de resíduos auditivos, se a surdez é congênita ou se foi adquirida após a aquisição da linguagem, etc... Avaliando-se a literatura, constata-se que nenhum dos métodos de realimentação acústica pode ser considerado o mais adequado para todos os surdos, pois a escolha depende das características de perda auditiva e das necessidades do portador da deficiência (Chaba, 1991; Blamey & Cowan, 1992).

Os dispositivos de ajuda que se baseiam na substituição sensorial, em vez de utilizarem o canal auditivo, quer seja por estimulação sonora (amplificadores) ou estimulação elétrica (implante coclear), simplesmente não utilizam o mecanismo da audição, e a informação de voz é transmitida via outro canal sensorial, como o tato ou a visão. Alguns métodos naturais de comunicação utilizados por portadores de deficiência auditiva baseiam-se nos sentidos do tato ou da visão para efetuar a transmissão de informações acústicas. A leitura labial, que é um método visual, e o método de Tadoma¹, exclusivamente tátil, são exemplos que comprovam a possibilidade de compreender-se a fala através dos sentidos visual ou tátil, fato que motiva a realização de pesquisas para desenvolvimento de sistemas táteis ou visuais aplicados ao auxílio à comunicação oral de deficientes auditivos.

¹ O Método de Tadoma é um método para comunicação com surdocegos no qual o "ouvinte" compreende a fala do locutor colocando sua mão na face do mesmo. A compreensão da fala ocorre pela percepção tátil dos movimentos das articulações, vibrações nasais e no pescoço e entrada e saída de ar na boca do locutor.

Os sistemas visuais de substituição sensorial em geral são utilizados para o treinamento e aprendizado da fala (Watanabe, 1995). Já existem sistemas comerciais para ensino de aspectos da fala, como o Speechviewer, da IBM (Novaes & Mendes, 1998). Em termos do português, Araújo (2000) desenvolveu um sistema visual para auxílio ao aprimoramento da voz de deficientes auditivos, e Oliveira (1998) desenvolveu um sistema de estimulação visual para auxílio à oralização de surdos.

Os sistemas táteis de substituição sensorial podem basear-se em estimulação vibrotátil ou estimulação eletrocutânea (Kaczmarek, Webster & Bach-Y-Rita, 1991), sendo que a maior parte dos dispositivos desenvolvidos, aplicados ao auxílio auditivo, utiliza estimulação vibrotátil (Chaba, 1991). Estes dispositivos vêm sendo desenvolvidos desde o início do século XX (Gault, 1924), com o objetivo de permitir a recepção da fala, e não com ênfase no ensino da produção da mesma. Já existem diversos dispositivos disponíveis comercialmente (Thornton & Phillips, 1992; Goldstein & Proctor, 1985; Szeto & Christensen, 1988), que vêm sendo desenvolvidos e usados em países como a Austrália (Blamey & Cowan, 1992), Suécia (Söderlund, 1992), Inglaterra (Summers & Gratton, 1995) e Estados Unidos (Chaba, 1991). Na literatura pesquisada encontram-se poucos trabalhos científicos no Brasil relacionados ao desenvolvimento de estimuladores táteis, quer sejam aplicados no auxílio auditivo (Barros, 1996; Barros et al, 2001; Barros et al, 2002), ou aplicados à propriocepção (Nohama, 1997) e à deficiência visual (Antonino, 1993).

Neste trabalho investigou-se a codificação de parâmetros de voz em estímulos vibrotáteis, para aplicação no auxílio ao ensino/aprendizagem da fala.

1.1 Motivação

O aprendizado da fala inicia em uma criança com audição normal quando a mesma é estimulada por uma pessoa falando, e a criança pode observar o som da fala na sua forma acústica (som) e visual (articulação). Na seqüência de

diversas observações a criança emite um som (emissão da fala) e o avalia através da escuta, comparando-o com o som da fala anteriormente ouvido, em um processo repetitivo, de realimentação e ajuste do aparelho fonador, até que ocorra a produção do som ou segmento da fala da forma desejada.

Com a falta da audição normal, a impossibilidade de comparar o som produzido com algum padrão auditivo dificulta todo o processo de aprendizado da fala.

Em alguns casos, o aproveitamento de resíduos auditivos, por meio de amplificadores pode restabelecer a realimentação auditiva. Em outras situações o implante coclear ou o uso de estimuladores táteis ("tactile aids") também podem auxiliar no aprendizado da fala, principalmente, de algumas características suprasegmentais, como a entonação.

A aprendizagem e desenvolvimento da fala para a vasta maioria das crianças surdas requerem instrução da fala. Em geral não se espera que estas crianças aprendam a falar por si próprios, (ou com intervenções mínimas), mas são submetidas a treinamentos específicos da fala. Dispositivos de auxílio sensorial, tanto táteis como visuais, podem ser usados na terapia da fala, ou para suplementar a audição residual, ou como substituto para a mesma.

Existem em geral quatro etapas envolvidas no ensino da fala, que são: emissão, automação, generalização e facilitação do uso lingüístico (Ling, 1976, Risberg, 1968), e nem todos os dispositivos são igualmente capazes de facilitar a produção nestas quatro etapas. As duas primeiras etapas situam-se no nível segmental da fala, ou seja, tratam do aprendizado de segmentos isolados da fala, como fonemas e sílabas isoladas. Nas duas últimas etapas, os segmentos são conectados, e o aprendizado e controle das características suprasegmentais da fala passa a ser fundamental.

Um passo inicial para o ensino de qualquer fundamento da fala é a emissão de um padrão. Por exemplo, o ensino da produção da consoante nasal /m/ envolve a demonstração para a criança do padrão do som (através da audição,

tato ou visão) e depois a criança deve executar diversas emissões até produzir um padrão aceitável.

Após uma produção aceitável, o padrão é praticado de forma que pode ser executado sem que o locutor tenha que prestar atenção nos detalhes de sua produção. Isto é referenciado como automação (Ling, 1976; Schmitt, 1988) e é considerado necessário para que a fala seja produzida suavemente e sem esforço.

Uma vez que o padrão possa ser produzido facilmente, a criança é ensinada a generalizar na produção do mesmo em diferentes contextos e localização nas sílabas e palavras.

Finalmente, a criança aprende a produzir o padrão em palavras e frases com significado. A facilitação do uso lingüístico envolve a transferência de fundamentos de padrões imitados e praticados anteriormente para situações de comunicação verbal. A situação desejada nesta etapa de ensino é o uso espontâneo deste fundamento na comunicação verbal.

Os sistemas de substituição sensoriais visuais são adequados para o ensino da fala nas duas primeiras etapas, ou seja, a emissão e a automação. Estas etapas atuam no nível segmental da voz, ou seja, na produção de fonemas separados. Muitas vezes, no entanto, um deficiente que fala e entende a comunicação verbal no nível segmental não consegue ter uma fala com frases inteligíveis.

A aplicação de estímulos táteis como forma de transmissão da informação acústica para deficientes auditivos vem sendo investigada desde o trabalho pioneiro de Gault (1924), no início do século XX, ao longo do qual foram desenvolvidos vários sistemas táteis com o objetivo de substituir o sentido auditivo, permitindo a recepção de voz (Sherrick, 1984; Bernstein, 1992; Weisenberger, 1992). Pelos resultados obtidos, no entanto, observa-se que a capacidade de transmissão de informações por estímulos táteis quer sejam monocanais ou multicanais, vibrotáteis ou eletrotáteis, é pequena, limitando a aplicação da estimulação tátil principalmente ao suporte à leitura labial (Summers et al, 1994). Também é importante observar que Tan (1996), utilizando estímulos vibrotáteis e

cinestésicos, baseados em estudos feitos com o método natural de Tadoma, desenvolveu um estimulador tátil com capacidade de transmissão de informações similar ao do canal auditivo, embora ainda não existam relatos de resultados de pesquisas referente à aplicação deste estimulador na recepção de voz.

Observa-se na literatura que, apesar dos dispositivos para ajuda tátil terem sido desenvolvidos com foco na recepção de fala, os mesmos têm apresentado bons resultados no auxílio ao seu aprendizado (Goldstein & Proctor, 1985). Sehgal et al (1998) investigaram a produção de consoantes de um grupo de crianças que utilizaram um estimulador tátil multicanal por um ano e meio, concluindo que houve um aumento da capacidade de produção de consoantes. Thornton & Phillips (1992) realizaram uma avaliação de alguns dispositivos de ajuda vibrotáteis comerciais, constatando que os mesmos contribuem para um melhor controle da voz, principalmente nos aspectos dinâmicos da mesma, que se relacionam com a qualidade da produção vocal e a inteligibilidade da fala. Atualmente, do ponto de vista comercial, a Audiological Engineering (www.tactaid.com) está produzindo o TACTAID VI, um estimulador multicanal que transmite diversos parâmetros de voz por estímulos vibrotáteis, e indicando-o tanto como ferramenta para ajuda à leitura labial, como ferramenta para o ensino da fala, de surdos e surdocegos.

Goldstein & Proctor (1985) apontam para a importância de desenvolver-se dispositivos para ajuda à produção da fala, enfatizando que um sistema para treinamento na produção da fala, usando sinalizações táteis e visuais simultaneamente pode ser eficiente. Mashie (1995) reafirmando o que foi colocado por Goldstein & Proctor dez anos antes, citou que o sentido tátil transmite informações dinâmicas da fala, e a visão, informações médias. Além disso, citou que são necessárias pesquisas adicionais para examinar a adequação das modalidades sensoriais táteis e visuais nas diferentes etapas do ensino da fala, com o objetivo de utilizá-las de forma integrada.

Um fator adicional que motiva o desenvolvimento de um sistema tátil específico para o ensino da fala é a existência de indivíduos com múltipla

deficiência, nos quais se colocam os surdocegos ou os surdos com deficiência visual. Em 1988, Szeto & Christsen descreveram a situação crítica do surdocego pré-lingual e alertou para a necessidade de se desenvolver instrumentos/sistemas que permitam a comunicação dos surdocegos com os outros indivíduos. Como na área do ensino da fala os sistemas existentes adequados para as etapas iniciais de automação e emissão utilizam a visão (Mashie, 1995), há a necessidade de desenvolver sistemas táteis com esta finalidade.

No Brasil, Araújo (2000) e Oliveira (1988) desenvolveram sistemas para auxílio a oralização de surdos, utilizando a visão como canal de realimentação. No sistema desenvolvido por Araújo, realimenta-se a informação acústica através de jogos lúdicos, e em Oliveira, além de jogos, há uma realimentação visual do posicionamento das articulações da fala. Segundo Oliveira, comentando a possibilidade de trabalhos futuros: "O uso de outros tipos de realimentação, paralelamente à visual, através de amplificadores e vibradores, complementa a informação recebida pelo surdo, além de poder auxiliar no treinamento da utilização dos resquícios auditivos, inclusive no caso de adaptação a próteses".

Nos congressos da área de educação especial e de engenharia biomédica há poucos relatos sobre a utilização da estimulação tátil, e principalmente os aplicados à reabilitação auditiva (Barros et al, 2002).

Com base no contexto descrito acima pode-se concluir que:

- a) Há fortes indícios que o tato pode desempenhar um papel importante no ensino da fala, de forma complementar e integrada ao sentido da visão, mas não existem ainda sistemas táteis desenvolvidos especificamente para o ensino/treinamento da fala;
- b) Os surdocegos já utilizam os dispositivos táteis desenvolvidos para transmissão de informação acústica para surdos, mas conforme se vê na literatura, tais sistemas não têm bom desempenho nas etapas de emissão e automação na aprendizagem da fala. Há necessidade de desenvolver-se sistemas de comunicação da informação acústica pelo tato específicos para

o ensino/aprendizagem da fala, abordando também exercícios que facilitem a emissão e a automação, que poderão ser utilizados por surdocegos e por surdos.

- c) Existe pouco conhecimento no Brasil sobre o uso de "tactile aids", além disto, tais dispositivos não foram desenvolvidos de acordo com a realidade da língua portuguesa. Isto motiva a realização de investigações para a parametrização dos sinais de voz e codificação de parâmetros de voz em estimuladores táteis.

Com relação à escolha da modalidade de estimulação tátil adequada para transmissão de informações acústicas, Summers (1992) concluiu que: "a percepção eletrotátil a alterações na frequência é pior que a vibrotátil", e que "o fato de algumas pessoas acharem a estimulação elétrica desagradável é um ponto adicional a favor do uso da estimulação vibrotátil". No entanto, eles reconhecem que "os eletrodos são menores e mais aceitáveis cosmeticamente que a alternativa dos vibradores eletromecânicos". Este fator, juntamente com o melhor consumo de potência, foi alegado por Nohama (1997) para afirmar que a estimulação eletrotátil seria mais adequada. Weisenberger (1992), contudo, analisando diversos trabalhos, concluiu que ainda não há uma palavra final sobre qual a melhor modalidade, elétrica ou vibratória, para transmissão da informação acústica pelo tato.

1.2 Objetivo

Considerando os itens a, b e c, e a recomendação de Summers quanto ao uso da estimulação vibrotátil, apresentados ao final da seção anterior, se teve como objetivo realizar uma investigação sobre a codificação de parâmetros de voz em estimuladores vibrotáteis, no contexto do ensino/aprendizagem da produção da fala, além da escolha de parâmetros de fala para codificação de estimuladores vibrotáteis, no contexto da língua portuguesa.

1.3 Organização do Documento

Este documento está dividido em seis capítulos.

No Capítulo 2 é mostrada a importância da linguagem para o desenvolvimento do ser humano, e como a deficiência afeta o desenvolvimento dos natissurdos e surdos pré-linguais. É mostrado o funcionamento e características do mecanismo de produção da fala no ser humano, e demonstrado porque o deficiente auditivo não aprende a falar normalmente. Ao final, são enfatizados os problemas que o surdo apresenta na fala e os métodos e técnicas para o ensino da produção da fala, baseados no uso de sentidos adicionais como o tato e a visão, com ênfase na necessidade de realimentação das informações acústicas.

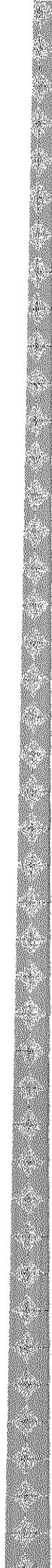
No Capítulo 3 são apresentadas as regras para transmissão de informação acústica por canais táteis, e mostrado os diversos tipos de estimuladores existentes e quais parâmetros acústicos podem ser transmitidos, a depender do parâmetro acústico que modula o estimulador. Será mostrado que os estimuladores vibrotáteis têm sido desenvolvidos com ênfase na aplicação de recepção de informações acústicas, mas diversos trabalhos citam as vantagens da utilização de estimuladores táteis no ensino da produção da fala. Ao final é feita uma revisão bibliográfica das características importantes do ensino da fala e resultados de pesquisas relacionados aos estimuladores táteis.

No Capítulo 4 são apresentadas propostas de transformações para codificação de parâmetros de voz no sinal de estimulação vibrotátil, além dos resultados de estudos por simulações.

No Capítulo 5 é apresentada inicialmente uma proposta de parametrização e codificação do sinal de voz em estímulos vibrotáteis, com ênfase no auxílio à discriminação de visemes da língua portuguesa. São apresentados estudos de caso para parametrização de características da voz e codificação de acordo com

as transformações apresentadas no Capítulo 4. São apresentados resultados de testes de percepção.

No capítulo 6 são apresentadas as conclusões e sugestões para continuidade desta investigação.



2

**COMUNICAÇÃO ORAL
E DEFICIÊNCIA
AUDITIVA**

2.1 Comunicação e Linguagem

Os seres humanos reúnem-se em sociedade, com cada indivíduo possuindo seus próprios interesses, que são associados aos de outras pessoas, juntamente com as leis, os costumes e tradições estabelecidas, em um processo de ajustamento social (Enciclopédia BARSÁ, 1995). A aglutinação de tais informações possui como etapa essencial a comunicação, que é um fator fundamental na integração e progresso de uma sociedade. A comunicação configura-se como forma de expressão dos interesses individuais e coletivos, sendo essencial para a formação do senso comum e a satisfação dos anseios individuais.

O ser humano comunica-se com o mundo externo através da emissão de estímulos que se tornam perceptíveis a outros seres humanos. Tal percepção se dá no correto funcionamento dos sentidos, na atividade mental normal e na codificação apropriada dos estímulos emitidos de acordo com os costumes dos seres humanos. Se um destes elos encontra-se defeituoso, há uma perda da capacidade de comunicação.

Desde o nascimento o ser humano encontra-se envolto em sons, e naturalmente aprende a decodificá-los, atribuindo-lhes um significado. Também apreende intuitivamente aspectos gramaticais da linguagem, coordenando palavras para formar frases que expressam seus anseios. De tanto escutar, aprende a coordenar seus órgãos fonoarticulatórios de modo a reproduzir os sons, aprendendo a falar.

A linguagem é o fenômeno da comunicação entre os seres humanos (Amorim, 1972). É a transmissão dos conceitos através de elementos simbólicos e convencionados. A voz, a fala, a escrita, a audição e os gestos são exemplos de tais elementos. A linguagem possibilita a organização, desenvolvimento e comunicação do pensamento, além de ser um permanente estímulo para que se formem conceitos, permitindo ao indivíduo a plena expansão de suas

potencialidades (Noronha & Rodrigues, 1974). Embora haja muitas formas de linguagem – como a escrita, por gestos, ou por sinais – a linguagem falada é o código mais usado na vida cotidiana.

A linguagem falada está profundamente arraigada nas estruturas de raciocínio do ser humano. A estrutura do pensamento termina por ser desenvolvida em função de estruturas da linguagem. O pensamento parece surgir mentalmente expressos em palavras e frases (Brooks, 1974). Para Helen Keller, cega e surda, que aprendeu a comunicar-se pelo tato, os problemas da surdez são mais profundos, mais complexos e mais importantes que os da cegueira (Almeida & Iorio, 1996). Para Helen Keller, “a surdez é o maior dos infortúnios, a perda do mais vital dos estímulos: o som da voz, que nos traz a linguagem, desencadeia os pensamentos e nos mantém na companhia intelectual dos homens”.

2.2 Voz, Fala e Audição

O som é a manifestação acústica percebida pelo sentido auditivo. Por isto é um fator primordial para a comunicação oral. Escutar e reproduzir sons, dentro de um código de sinais, resulta, para o ser humano, no ato de transmissão dos conteúdos da mente.

A voz, a fala e a audição são elementos fundamentais da linguagem (Martins, 1988). A voz é a produção de sons efetuada pelo ser humano através das cordas vocais. É o elemento sonoro da comunicação. Assim, apresenta características acústicas, tais como: a intensidade, que é a força ou volume em que o som é produzido; a altura, que determina o tom grave ou agudo do som, isto é, relaciona-se com sua frequência; e o timbre que, definido pelos componentes harmônicos do espectro do som é o que, como exemplo, caracteriza e diferencia sons de instrumentos distintos. É a detecção e análise de tais características que permitem o reconhecimento da informação contida na voz.

A fala é a produção dos fonemas com articulação na palavra. Articulação é o contato que os órgãos fonatórios têm, ora em uma região, ora em outra região

bucal, resultando em sons diferentes que se combinam, formando a comunicação verbal (Amorim, 1972). De fato, a fala diferencia-se da voz e da linguagem. A voz é simplesmente a sonoridade. A fala é a voz articulada; e a linguagem oral é a transmissão de conceitos através de voz e fala. Como exemplos: o cachorro tem voz; o papagaio tem voz e fala; e o homem, exclusivamente, é capaz de transformar a realidade em palavras e elaborar conceitos numa vivência interior.

A voz é controlada pela audição do locutor. Através dela o indivíduo determina a intensidade, a altura, o timbre e a duração de sua voz. A audição implica na aptidão para escutar, e é um elemento essencial para o aprendizado da fala. Assim, o deficiente auditivo, surdo ou profundamente surdo, é inapto naturalmente tanto ao reconhecimento como à produção da fala.

2.3 A Produção da Fala

A fala é uma onda de pressão acústica que se origina de movimentos voluntários das estruturas anatômicas que formam o aparelho de produção de voz.

As partes do corpo responsáveis pela produção dos sons da fala são os pulmões, a traquéia, a laringe, cavidade nasal, cavidade bucal, língua, lábios, véu palatino e dentes, que formam um tubo que vai dos pulmões aos lábios, conforme se vê na Figura 1. A parte superior deste tubo, que se inicia no final da laringe é chamada trato vocal, e consiste de parte da laringe, faringe, cavidade oral e nasal (Spinelli et al, 1999). A parte inferior, denominada sistema subglotal, funciona como uma fonte de energia para produção da voz.

A produção da voz depende dos órgãos e do sistema de respiração. O ar que respira-se permite produzir voz. Na fase de expiração, as ondas sonoras, modeladas pela laringe e as cavidades orais e nasais, dão as características da voz.

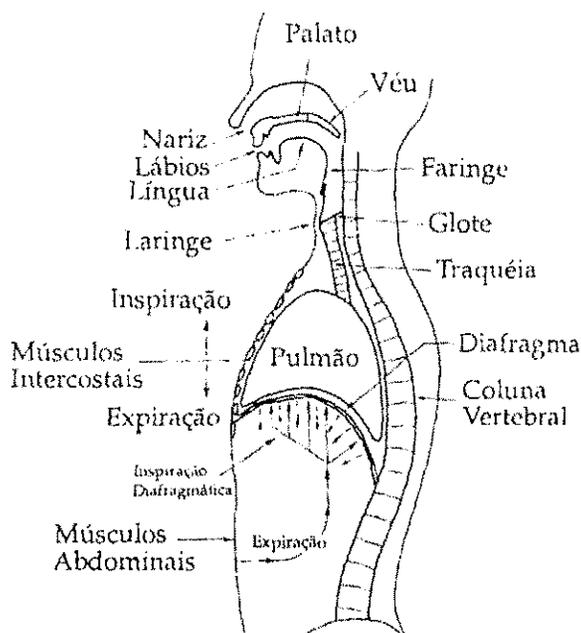


Figura 1 - DIAGRAMA ESQUEMÁTICO DOS ÓRGÃOS ENVOLVIDOS NA PRODUÇÃO DA FALA

Na Figura 2 é possível observar-se um diagrama de blocos do aparelho fonador. A corrente de ar sai dos pulmões durante a expiração e de acordo com a posição do véu palatino (velum) pode ser dirigida à cavidade oral ou nasal. A expiração, sempre nasal e silenciosa, torna-se audível por intervenção da laringe ou por constrições produzidas no trato vocal. A forma como os articuladores se posicionam para executar tais constrições é denominada de articulação, e as mesmas exercem uma filtragem seletiva sobre a excitação de ar produzida pelos pulmões.

Portanto, para que haja produção de fala correta, os pulmões devem ser capazes de produzir uma excitação adequada e os articuladores precisam adquirir a velocidade e exatidão suficientes para produção de sons inteligíveis. A excitação ainda pode sofrer ação vibratória das cordas vocais.

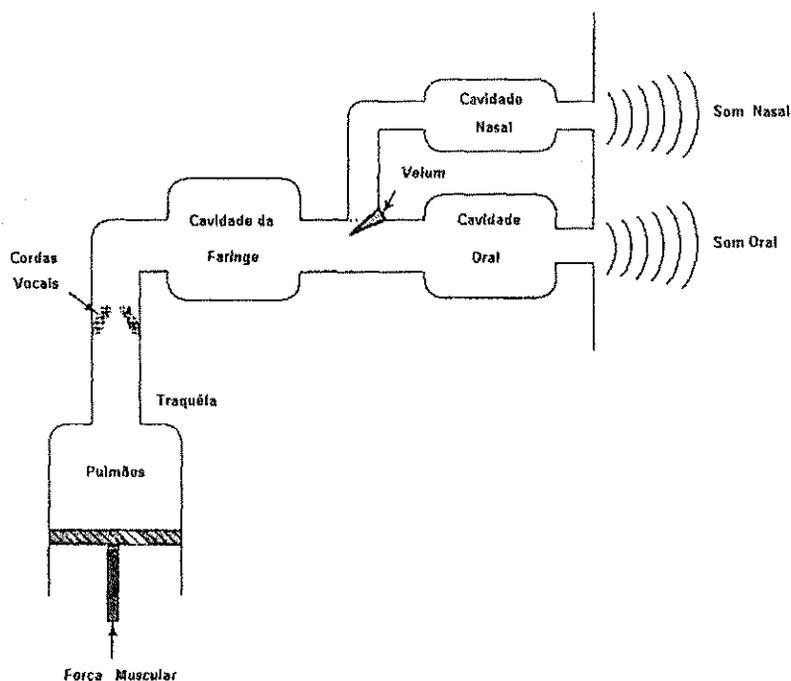


Figura 2 - DIAGRAMA DE BLOCOS DO APARELHO FONADOR

Ainda com relação à produção da fala, Martins (1988) menciona que a mesma passa por três etapas: a respiração, a fonação e a articulação. Martins expõe didaticamente que, como sistema acústico, a respiração é equivalente à fonte de energia acústica para a fonação, que se traduz em pressão de ar com determinado volume e velocidade. O som assim originado passa pelo trato vocal que atua como um filtro acústico, mudando as características do som que vai passando. Apresenta-se na Figura 3 um diagrama de blocos desse processo.

Os sons da voz podem ser agrupados em três classes distintas de acordo com o modo de excitação, que são os sons sonoros, surdos e explosivos (Fechine, 2000).

Nos sons sonoros, o fluxo de ar vindo dos pulmões é controlado pela abertura e fechamento das cordas vocais, que são semelhantes a dois lábios que podem ser tensionados e (ou) aproximados sob o controle do locutor. Quando estão completamente fechados, o fluxo de ar vindo dos pulmões é interrompido e a pressão aumenta até que as cordas vocais sejam separadas, liberando o ar

pressionado, gerando um pulso de ar de curta duração. Com o escoamento do ar, a pressão é reduzida, possibilitando uma nova aproximação das cordas vocais. O processo se repete de forma quase-periódica. Estas ondas de pressão estimulam o trato vocal, que atuando como ressonador modifica o estímulo, produzindo freqüências de ressonância, denominadas formantes, que caracterizarão os diferentes sons sonoros. As vibrações das cordas vocais são responsáveis pela variação do tom da voz, isto é, definem a freqüência fundamental de cada voz. Os fatores de tamanho da laringe, e conseqüentemente das cordas vocais, são responsáveis pelas diferenças de tom de voz entre os homens (freqüência fundamental entre 100 e 150 Hz), as mulheres (freqüência fundamental entre 200 e 250 Hz) e as crianças (entre 300 e 500 Hz).

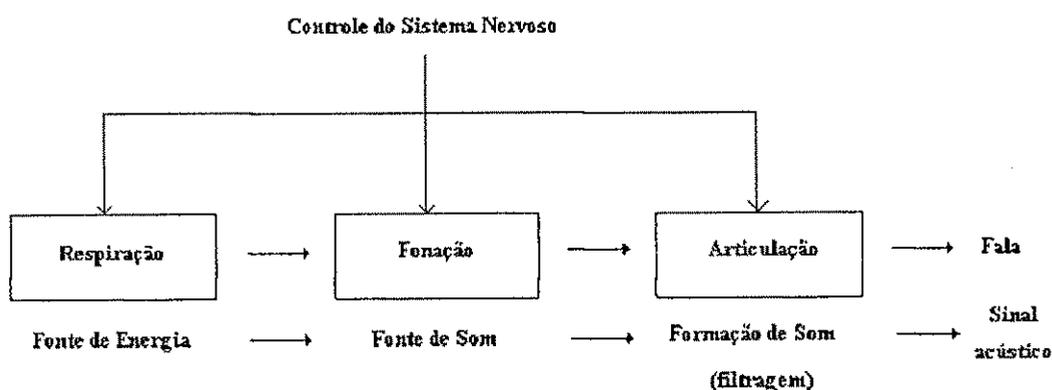


Figura 3 - PROCESSO DE PRODUÇÃO DA FALA (Martins, 1988)

Os sons surdos são gerados pela produção de uma constrição em algum ponto do trato vocal, assim o ar adquire velocidade suficientemente alta para produzir turbulência gerando um ruído de espectro largo para excitar o trato vocal. Na produção destes sons as cordas vocais permanecem abertas, não havendo vibração.

Nos sons explosivos, o ar é totalmente dirigido à boca, estando esta completamente fechada. Com o aumento da pressão, a oclusão é rompida bruscamente, gerando um pulso que estimula o aparelho fonador.

Os sons da fala podem ser divididos em vogais e consoantes, considerando-se segmentos que possuem propriedades acústicas semelhantes para curtos intervalos de tempo.

2.3.1 Vogais

As vogais são sons sonoros diferenciados principalmente pela posição da mandíbula e da língua. O diagrama da Figura 4 contém um mapeamento da posição da língua e da mandíbula para cada vogal, baseado em que pode-se classificar as vogais orais do português brasileiro.

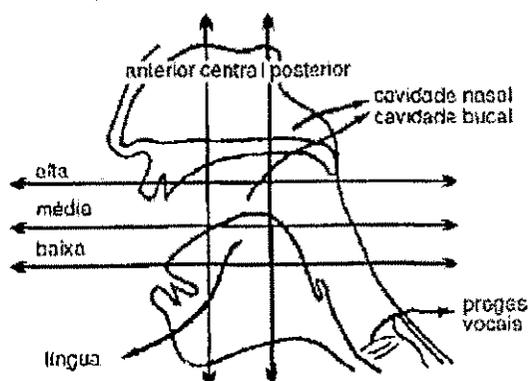


Figura 4 - Posicionamento dos órgãos articuladores (língua e da mandíbula) de acordo com o tipo de vogal (Russo & Behlau, 1993)

As vogais orais do português falado no Brasil são:

/a/, vogal central, baixa, conforme pronunciada em vaca;

/e/, vogal anterior, média, conforme pronunciada em pé;

- /e/*, vogal anterior, média, conforme pronunciada em *peso*;
- /i/*, vogal anterior, alta, conforme pronunciada em *tia*;
- /ɔ/*, vogal posterior, média, conforme pronunciada em *pote*;
- /o/*, vogal posterior, média, conforme pronunciada em *poço*;
- /u/*, vogal posterior, alta, conforme pronunciada em *uva*.

2.3.2 Consoantes

Segundo Spinelli et al (1999), as consoantes constituem diversas categorias. Elas podem ser surdas ou sonoras, orais ou nasais, oclusivas ou fricativas, etc., mas são mais bem descritas a partir da especificação de seus pontos de articulação e seu modo de articulação.

Russo & Behlau (1993) apresentaram sucintamente os pontos de articulação das consoantes da língua portuguesa, que são:

- *Os lábios entre si*: consoantes bilabiais - / p, b, m /
- *O lábio inferior e a arcada superior*: consoantes labiodentais - / f, v /
- *A língua e os dentes*: consoantes linguodentais ou alveolares - / t, d, n, l, r, s, z /
- *A língua e o palato*: consoantes linguopalatais ou palatais - / , j, λ /
- *A língua e o véu*: consoantes linguovelares ou velares - / k, g, ŋ /

Na Figura 5 é apresentada uma representação esquemática dos pontos de articulação das consoantes do português.

Ainda segundo Russo & Behlau (1993), a classificação quanto ao modo de articulação diz respeito à consoante ser o produto da passagem do ar através de um obstáculo que obstrui total ou parcialmente o trato vocal.

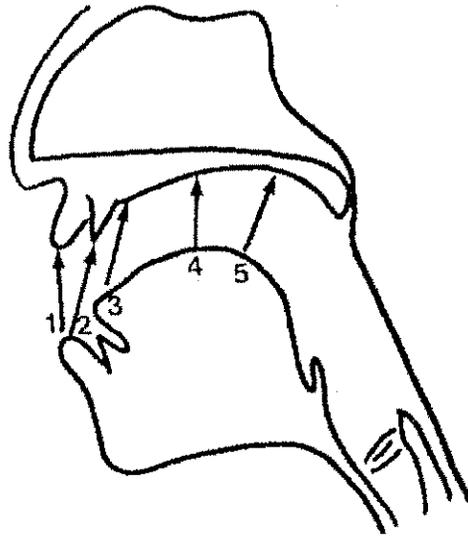


Figura 5 - Representação esquemática dos pontos de articulação das consoantes do português: 1. bilabial, 2. lábiodental, 3. linguodental, 4. linguopalatal, 5. linguovelar (Russo&Behlau, 1993)

Para o português, se tem os seguintes modos de articulação:

- Oclusivas: são produzidas pelo bloqueio da pressão do ar em algum ponto do trato vocal que a seguir é desfeito.
- Fricativas: são produzidas por uma constrição da corrente aérea em algum ponto do trato vocal, determinando uma grande turbulência e produção de sons de alta frequência.
- Líquidas: Compreendem as laterais e as vibrantes. As laterais se distinguem pela obstrução da corrente expiratória em um ponto central e esse fato permite sua passagem pelas regiões laterais. Nas vibrantes, a obstrução também ocorre em um ponto, mas a língua executa um ou mais movimentos rápidos.

As oclusivas também podem ser orais ou nasais, e surdas ou sonoras. As fricativas e líquidas podem ser surdas ou sonoras. Um resumo de todas estas classificações encontra-se na Tabela 1.

Tabela 1 - Quadro das consoantes do português, classificadas de acordo com o ponto e modo de articulação

Modo	oclusivas			fricativas		líquidas	
	orais		nasais	orais		surda	sonora
Ponto	surda	sonora	sonora	surda	sonora		
Bilabiais	p	b	m				
Labiodentais				f	v		
Linguodentais ou alveolares	t	d	n	s	z		l r
Linguopalatais ou palatais					j		λ
Linguovelares ou velares	k	g	ŋ				

2.3.3 Fatores que Interferem no Processo Articulatório: realimentação

Os fatores que interferem diretamente no processo articulatório da fala podem ser divididos em dois grupos de acordo com a função que exercem: os fatores aferentes, que levam informação ao sistema nervoso central, e eferentes, que trazem os estímulos do sistema nervoso central para a periferia (Spinelli et al, 1999).

Os fatores aferentes são constituídos por todas as informações que servem de modelo ou correção dos atos articulatórios, e apoiam-se em:

- 1) função auditiva – responsável pelo fornecimento do modelo acústico vindo do exterior e do próprio indivíduo (realimentação acústica);

- 2) função tátil – responsável pela informação sobre os pontos de contato durante a articulação (realimentação tátil);
- 3) função proprioceptiva – informa sobre as sensações dos músculos e tendões, fornecendo condições para análise do movimento articulatorio, da pressão no momento do contato e da tonicidade da musculatura envolvida (realimentação cinestésica);
- 4) função visual – responsável pela formação de modelos visuais das produções articulatorias.

Os fatores eferentes são constituídos principalmente pelas funções responsáveis pela execução dos movimentos articulatorios.

Segundo Zemlin (2000), a retroalimentação auditiva é uma via principal pela qual monitora-se nossa produção sonora, no entanto, uma vez que a fala esteja bem estabelecida, o papel da retroalimentação auditiva diminui, como demonstram os indivíduos que tiveram perdas de audição graves com mais idade, mas que conseguem manter a articulação adequada, principalmente através da realimentação cinestésica.

De acordo com Mysak (1988), o sensor auditivo, em circunstâncias normais, monitora o pensamento verbal e controla e monitora a efetiva produção de código verbo-acústico, tanto quanto suas dimensões de tempo e sonorização. Dependendo da idade do indivíduo, o sensor auditivo também faz uma importante contribuição ao controle e monitoração da articulação dos sons da fala. Teoricamente, quanto mais a idade do indivíduo se aproxima e ultrapassa os sete ou oito anos, menos é utilizado o sensor auditivo no controle e monitoração da articulação. A propriocepção coopera com o controle e monitoração da articulação efetiva da fala. Este mecanismo torna-se aparentemente mais importante nesta atividade, logo após a maturação (Araújo, 2000).

Segundo Zemlin (2000), os sensores proprioceptivos apresentam, em sua maior parte, adaptação muito rápida, enviando informações enquanto o movimento está ocorrendo. Muitas destas informações nunca atingem o nível da consciência.

2.4 O Aprendizado da Fala

A criança com audição normal aprende a falar a linguagem espontaneamente através da combinação entre a escuta e a leitura labial. O aprendizado da fala em uma criança se faz com base na existência de quatro tipos de informações: auditivas, táteis, visuais e cinestésicas. Há pesquisas que afirmam que a informação mais importante para recepção de voz no início do aprendizado da fala é a informação visual, depois a auditiva (Öster, 1995). A partir da informação visual a criança começa a tentar imitar as articulações e gerar sons, comparando a sua produção sonora com a de outras pessoas. Começa então a criar um conjunto de modelos cinestésicos, visuais, auditivos e táteis para a produção dos sons, que se auto-ajustam e se relacionam. Os passos seguintes são a seqüência mais provável envolvida no aprendizado da fala:

1. À medida que a criança escuta os sons da fala, armazena em sua memória um padrão auditivo, que servirá como guia para a produção motora (movimentação dos articuladores), e como referência de avaliação do aprendizado dos padrões motores;
2. A criança tenta produzir os sons da fala que ela escuta;
3. A consciência dos padrões da fala produzida não é concorrente com a produção, mas ocorre após a produção. A criança obtém informação sobre o resultado da tentativa de produção através da realimentação auditiva e então compara o padrão auditivo associado com sua imitação com os padrões de referência armazenados. Esta comparação funciona como um mecanismo de detecção de erro onde a exatidão da produção é avaliada.

4. Tentativas subseqüentes são adaptadas para reduzir o erro na produção, e as comparações são feitas novamente.

5. Depois de repetidas tentativas para produzir o padrão, a criança estabelece as transformações sensório-motoras requeridas para a produção do padrão.

6. O resultado é desenvolvido de transformações aprendidas entre os padrões auditivos e os padrões motores (associação sensório-motora).

A aquisição da fala, então, baseia-se e é mediada pela audição. Dentre as características motoras envolvidas na fala que a criança deve aprender através da audição estão:

1. Posicionamento dos articuladores;
2. Controle do fluxo de ar;
3. Coordenação dos articuladores;
4. Coordenação dos elementos respiratórios, articulatórios e fonatórios da fala;
5. Produção de fonação adequada (entonação e qualidade)
6. Padrões articulatórios da laringe precisos para a produção do contraste surdo/sonoro, junto com uma coordenação apropriada dos padrões;
7. Suporte da respiração.

2.5 Deficiência Auditiva

A deficiência auditiva ocorre em graus variados e é resultado da perda da capacidade de percepção normal dos sons. Além do grau de deficiência auditiva, o deficiente pode ser classificado pelo período em que a mesma surgiu (perdas congênitas ou adquiridas), pela associação com outras deficiências (surdocegueira, por exemplo), de acordo com a causa, etc. (Brasil, 1995). As

perdas adquiridas ocorrem por causas patológicas, como meningite, icterícia, outros processos infecciosos durante a infância, viroses, otites, intoxicações medicamentosas e outros; ou ainda por traumatismos, como exposição contínua a ruídos. As perdas congênitas podem ter origem hereditária ou durante o processo de gestação. No Brasil, cerca de 20% dos casos de deficiência auditiva grave são causados por rubéola durante a gestação.

Em termos de magnitude, pesquisas relatam que nos Estados Unidos, em 1985, 0,3% da população não conseguia ouvir e entender uma conversação normal, enquanto 3% escutavam e entendiam com dificuldades (Chaba, 1991). Outras pesquisas mostram que um ou dois bebês entre mil nascem com problemas auditivos sérios e de vinte a trinta irão apresentar alguma perda auditiva até a idade de cinco anos (Boothroyd, 1982; Northern & Downs, 1989). Estudos mais recentes estimam a existência de 3 a 12 crianças com perdas auditivas em cada 2000 nascimentos (Jcjh, 1995). No Brasil, segundo estimativas da Organização das Nações Unidas e da Organização Mundial de Saúde, os surdos atingem um total de quatro milhões de pessoas, enquanto estima-se que 2 % da população possui alguma deficiência auditiva.

Segundo definições da Sociedade Otológica Americana, *surdo* é o indivíduo cuja audição não é funcional na vida comum, e parcialmente surdo é aquele cuja audição, ainda que deficiente, é funcional com ou sem prótese auditiva. As perdas auditivas podem ser classificadas em grupos separados, segundo o valor da média dos limiares de audibilidade nas frequências de 500, 1000 e 2000 Hz. Um indivíduo com perda auditiva de grau severo é aquele cuja média dos limiares de audibilidade encontra-se entre 71 e 90 dB, enquanto que o indivíduo com perda profunda possui média dos limiares de audibilidade acima de 90 dB. Redondo & Lopes (1997), baseado na International Organization for Standardization, utilizam a seguinte classificação: normal para perdas até 25 dB; leve para perdas entre 26 e 40 dB; moderada para perdas entre 41 e 55 dB; moderadamente severa entre 56 e 70 dB; severa para perdas entre 71 e 90 dB; profunda para perdas acima de 91 dB.

2.6 Educação de Deficientes Auditivos

A dificuldade de comunicação oral entre surdos e ouvintes é um problema que ao longo dos tempos relegou indivíduos com deficiência auditiva a condições inferiorizadas na sociedade.

Segundo a filosofia de Aristóteles, todos os conteúdos da consciência deviam ser recolhidos primeiro por um órgão sensorial, tendo-se o ouvido como o órgão sensorial mais importante para a educação. No decorrer do tempo, isto foi interpretado como se Aristóteles tivesse negado ao surdo qualquer possibilidade de instrução. Este pensamento perdurou até o final da Idade Média. No entanto, a partir do Renascimento, os médicos e religiosos passaram a se dedicar ao estudo da fala dos surdos, bem como de suas possibilidades de aprendizagem.

No século XVI, o médico italiano Gerolamo Cardano (1501-1576) afirmou, baseado em seus estudos, que a mudez não se constituía em impedimento para que o surdo adquirisse conhecimento, e que a surdez, por si mesma, não modificava a inteligência da criança, de modo que a educação dos surdos deveria ser realizada pelo ensino da leitura e da escrita. Ainda no século XVI o frade espanhol Pedro Ponce de León conseguiu ensinar a linguagem articulada a surdos-mudos, filhos de membros da corte espanhola e, em 1620, outro espanhol, Juan Pablo Bonet, publicou o primeiro livro sobre o assunto, em que explica como exercitar o educando para a emissão dos sons. O reconhecimento de surdos filhos de famílias nobres como pessoas de lei (os mudos não eram reconhecidos), para que pudessem herdar títulos e a fortuna da família, foi um fator para o desenvolvimento de métodos educacionais especiais para os deficientes auditivos (Sacks, 1998).

Até 1750 o acesso à educação especial era um privilégio de poucos surdos, e a situação das pessoas com surdez pré-linguística era calamitosa, pois eram considerados incapazes de desenvolver a fala, e portanto "mudos", incapazes de comunicar-se livremente até mesmo com pais e familiares, restritos a

alguns sinais e gestos rudimentares. Isolados, exceto nas grandes cidades, até mesmo da comunidade de pessoas com o mesmo problema, privados de alfabetização e instrução, de todo o conhecimento do mundo, forçados a fazer trabalhos mais desprezíveis, vivendo sozinhos, muitas vezes à beira da miséria, eram considerados pela lei e pela sociedade como pouco mais do que imbecis (Sacks, 1998). Entretanto, em 1760 o abade francês Charles Michel de L'Épée fundou a primeira escola de surdos, em Paris e criou o chamado método silencioso, que permitia a realização de uma instrução rápida. Sem desprezar a importância da palavra oral, L'Épée deu relevo especial ao emprego de sinais manuais, estabelecendo uma linguagem convencional, como meio de instrução dos deficientes auditivos. Esse sistema foi amplamente difundido em toda a Europa até o aprimoramento do método oral, que emprega a palavra e a leitura labial na transmissão das idéias.

As instituições de educação de surdos disseminaram-se entre os principais países da Europa no século XIX, sendo que as escolas italianas alcançaram um notável desenvolvimento no que se refere ao ensino da linguagem oral. No primeiro Congresso Internacional de Surdos-mudos, que aconteceu em Paris em 1878, sustentou-se que o melhor método de ensino seria aquele que combinasse a articulação com a leitura das palavras nos lábios, mas conservando o uso do gesto, como medida de auxílio entre professores e alunos durante o período inicial. Porém em 1880, quando se realizou o segundo congresso, em Milão, o uso da fala e do gesto foi rechaçado e a utilização do método oral puro foi recomendada. A controvérsia entre defensores do método oral e os partidários da linguagem de sinais que se estabeleceu em princípios do século XIX, persiste ainda hoje (Enciclopédia BARSA, 1995).

A partir de 1960, a língua dos sinais tornou a ressurgir associada à forma oral, com o aparecimento de novas correntes, como a Comunicação Total e, mais recentemente, o Bilingüismo.

A Comunicação Total defende a utilização de todos os recursos lingüísticos, orais ou visuais, simultaneamente, privilegiando a comunicação, e não

apenas a língua. Já o Bilingüismo acredita que o surdo deve adquirir a língua dos sinais como língua materna, com a qual poderá desenvolver-se e comunicar-se com a comunidade de surdos, e a língua oficial de seu país como segunda língua.

No Brasil, a educação dos surdos teve início durante o segundo império, com a chegada do educador francês H Ernest Huet. Em 1857, foi fundado o Instituto Nacional de Surdos-Mudos, atual Instituto Nacional de Educação dos Surdos (INES), que inicialmente utilizava a língua dos sinais, mas que em 1911 passou a adotar o oralismo puro. Na década de 70, com a visita de Ivete Vasconcelos, educadora de surdos da Universidade Gallaudet (Estados Unidos), chegou ao Brasil a filosofia da Comunicação Total, e na década seguinte, a partir das pesquisas da professora lingüista Lucinda Ferreira Brita sobre a Língua Brasileira de Sinais e da professora Eulália Fernandes, sobre a educação dos surdos, o Bilingüismo passou a ser difundido. Atualmente, estas três correntes educacionais ainda persistem paralelamente no Brasil.

A aquisição da língua oral, embora reconhecidamente difícil e demorada, é recomendada por quase todos os estudiosos, de quase todas as metodologias propostas para a educação de deficientes auditivos (Araújo, 2000). Neste contexto, esta aquisição da língua oral só pode ocorrer criando-se uma retroalimentação para a fala do indivíduo, através da utilização de resíduos auditivos, por implante coclear ou por meio de um sentido alternativo, como o tato ou a visão. Existem diversos trabalhos na literatura relacionando o tato e a visão à transmissão alternativa de informação acústica para surdos. Neste trabalho, apresenta-se um estudo dos estímulos vibrotáteis aplicados ao auxílio à comunicação oral de surdos.

2.6.1 Desenvolvimento da Fala da Criança Surda

As abordagens tradicionais para facilitar o desenvolvimento da fala em crianças surdas geralmente tentam imitar o processo de aquisição da fala das crianças de audição normal, embora com auxílio de amplificação e instrução

especializada. Quando a criança tem capacidade de extrair informações úteis do que escuta, esta abordagem pode ser benéfica. Enquanto apropriado para alguns, nem todos os surdos desenvolverão a fala automaticamente apenas fornecendo-se amplificação. As limitações da audição resultam na dificuldade de detecção e na dificuldade de processamento da fala. Como resultado, apenas a audição residual não é suficiente para habilitar a maioria dos surdos no desenvolvimento da fala.

Para muitas crianças surdas, a utilização de informações sensoriais alternativas é necessária. Isto envolve fornecer para a criança informações auditivas alternativas na forma de implante coclear, ou informações alternativas e suplementares na forma de estimulação vibrotátil. Para estas crianças o objetivo é facilitar a produção natural da fala e o desenvolvimento sensorial via um dispositivo protético, que deve ser utilizado todo o tempo (Mashie, 1995).

2.6.1.1 Fatores Pedagógicos Envolvendo o Ensino da Fala

Conforme está descrito na introdução deste trabalho, existem em geral quatro etapas envolvidas no ensino da fala, que são: emissão, automação, generalização e facilitação do uso lingüístico.

Cada uma destas quatro etapas pode ser subdividida em outras. Por exemplo, a generalização da consoante nasal /m/ poderia iniciar com a produção isolada, seguida pela produção da consoante após a vogal /e/. Exercícios posteriores poderiam variar o contexto da vogal e o contexto da sílaba.

Existem muitos dispositivos desenvolvidos para auxiliar o ensino da fala, no entanto não existam dispositivos que permitam trabalho extensivo em ambos os níveis, segmentais e suprasegmentais, ou seja, nem todos os dispositivos são igualmente capazes de facilitar a produção nas quatro etapas descritas acima (Mashie, 1995).

2.6.1.2 Dispositivos para Auxílio ao Ensino da Fala

A realimentação fornecida por dispositivos de ajuda sensorial (sensory aids) pode ser obtida a partir de transdutores acústicos ou fisiológicos, como microfones (Watson & Keley Port, 1989) dispositivos de medidas aerodinâmicas, como o pneumotacógrafo (Mashie & Yadav, 1990), acelerômetros (Stevens, Kalikow & Willemain, 1975) e dispositivos especiais, como o eletropalatógrafo que monitora o contato entre a língua e o palato (Fletcher & Hasegawa, 1983).

O primeiro estágio do ensino da fala, emissão, é adequado para o uso de realimentação da informação fisiológica, que leva o surdo a adquirir consciência de como controlar as posições articulatórias de um segmento de voz particular (Mashie, 1995; Spinelli et al, 1999). O eletropalatógrafo, por exemplo, fornece informações visuais sobre o contato entre a língua e o palato durante a fala. Contudo, embora a realimentação de padrões articulatórios seja útil para a etapa de emissão de segmentos de voz, não é ideal para os outros estágios do aprendizado da fala.

A maioria dos dispositivos para auxílio ao ensino da fala existentes baseia-se em parâmetros acústicos, e a realimentação pode ser visual ou tátil. Os sistemas de substituição sensoriais visuais são adequados para o ensino da fala nas duas primeiras etapas, ou seja, a emissão e a automação. Estas etapas atuam no nível segmental da voz, ou seja, na produção de fonemas separados. Thornton & Phillips (1992) realizaram uma avaliação de alguns dispositivos de ajuda vibrotáteis comerciais e constataram que os mesmos contribuem para um melhor controle da voz, principalmente nos seus aspectos dinâmicos, que se relacionam com a qualidade da produção vocal e a inteligibilidade da fala. Ou seja, estão relacionados às etapas de generalização e facilitação do aprendizado da fala. Assim o uso da estimulação vibrotátil integrada com a estimulação visual pode trazer benefícios ao aprendizado da fala e uma ligação mais natural entre as etapas de emissão, automação e generalização do seu aprendizado.



3

ESTÍMULOS VIBROTÁTEIS: CANAL SENSORIAL ALTERNATIVO À AUDIÇÃO

A sensação tátil é uma alternativa para substituição da visão, audição ou propriocepção, e possui vantagens com relação aos outros sentidos, como a liberação dos sentidos intactos para tarefas mais importantes e a possibilidade de construção de estimuladores pequenos, leves e com boa apresentação estética. Além disso, o tato é o único sentido a combinar as dimensões espacial e temporal em um só canal sensorial, e raramente estar ocupado, ao contrário da visão e da audição.

Diversas técnicas podem ser implementadas para evocação da sensação tátil, pois tanto estímulos elétricos quanto mecânicos, térmicos ou químicos são percebidos pelos receptores, seus periféricos e as fibras nervosas aferentes. Devido às limitações de ordem prática, somente os dois primeiros têm sido empregados em sistemas de substituição ou ampliação sensorial (Nohama, 1997).

O conceito da utilização do tato como substituto de outros sentidos tem as primeiras referências registradas nos trabalhos de Dalgarno e Rosseau, no período do Renascimento, em 1680 e 1762, respectivamente (Sherrick, 1984). No entanto, foi no início do século XX que houve um crescimento das pesquisas nesta área, devido ao desenvolvimento da eletrônica e dos estudos da psicofísica.

3.1 Percepção de Estímulos Táteis

A hipótese de que a audição humana desenvolveu-se a partir do tato deve ter sido um estímulo para que Gault (1924) realizasse os primeiros experimentos para transmissão da informação acústica por este sentido (Kirman, 1982). Gault utilizou um tubo oco, no qual falava-se em uma extremidade enquanto a outra estava em contato com a pele de um indivíduo. Mesmo com este sistema simplificado, algumas palavras puderam ser identificadas, o que levou ao desenvolvimento de sistemas com um amplificador eletrônico e um vibrador. Apesar de terem sido obtidos alguns resultados positivos, não foi possível demonstrar que a voz poderia ser interpretada por meio do tato (Gault, 1926).

Atualmente, sabe-se que muitas das dificuldades obtidas naquela época derivaram do desconhecimento sobre os atributos sensoriais dos receptores vibrotáteis. Por exemplo, a resposta do tato à estimulação vibrotátil situa-se na faixa de 50 a 500 Hz, que é bastante reduzida com relação às frequências que compõem um sinal de voz. Desde os tempos de Gault até hoje foram realizados diversos trabalhos para definir os atributos sensoriais do tato, quanto às estimulações vibrotátil e eletrotátil. Existem revisões bibliográficas sobre o assunto em Verrilo & Gescheider (1992), Kaczmarek et al (1991) e Sherrick (1984). O entendimento das características e limitações sensoriais da estimulação vibrotátil já estabelecidos e consolidados na literatura é essencial para o desenvolvimento de estimuladores vibrotáteis eficazes.

Algumas características que afetam a resposta do sistema tátil aos estímulos vibrotáteis podem ser classificadas em fatores temporais, espaciais e fatores subjetivos da estimulação (Verrilo & Gescheider, 1992).

Quanto aos aspectos temporais, diversos estudos mostram que o tato tem resposta diferenciada com a variação da frequência (Verrilo & Gescheider, 1992). Verrilo (1963) mostrou que quando um contato grande (área maior que 0,02 cm²) é utilizado para produzir vibrações táteis, o limiar de sensação é independente da frequência em baixas frequências (25 a 40 Hz), depois passa a variar, atingindo-se uma maior sensibilidade em torno de 250 Hz. Até 1000 Hz, a sensibilidade decresce rapidamente, e limiares válidos não podem ser obtidos após esta frequência. Quando o contato do estimulador tem área pequena (menor que 0,02 cm²) o limiar de sensação é independente da frequência do sinal de estimulação. A discriminação em frequências, ou sensibilidade a variações em frequências, para níveis de intensidade acima do limiar de sensação é melhor em frequências baixas e diminui com o aumento da mesma (Verrilo et al, 1969). Por exemplo, a sensibilidade à variação em frequências é melhor em torno de 100 Hz do que na frequência de 250 Hz. Outro atributo importante é o tempo em que devemos manter um sinal para que o mesmo seja percebido, e é chamado duração do estímulo. Verrilo (1965) mostrou que, quando a área do contato é grande (2,9cm²), é mais difícil detectar-se um sinal de pequena duração, e a capacidade de

detecção aumenta com a duração do estímulo. Neste caso, quanto menor a duração do estímulo, maior deverá ser a intensidade do mesmo para que possa ser percebido.

Quanto aos aspectos espaciais as diferenças na percepção e sensibilidade dos limiares com relação a partes do corpo definem as possibilidades de desenvolvimento de dispositivos para partes distintas do corpo. Verrilo (1992) observou os limiares de sensação para sinais senoidais aplicados no antebraço, na palma da mão e na ponta do dedo médio, obtendo que a maior sensibilidade situa-se no dedo médio e a menor no antebraço. Outro fator importante é a resolução espacial do tato, que é usualmente definida como a mínima distancia entre dois pontos de um estimulador. A ponta dos dedos e palma das mãos apresentam a melhor resolução, em torno de 2 a 10 mm (Kaczmarek et al, 1991).

Tabela 2 - Características Básicas da Percepção Vibrotátil

Faixa de Frequências	50 a 500 Hz
Limiar de Sensação (intensidade)	Curva em forma de U, com vale em torno de 250 Hz
Discriminação em Frequência (sensibilidade para sinais acima do limiar de sensação)	Melhor sensibilidade abaixo de 100 Hz. A sensibilidade diminui com o aumento da frequência
Percepção Vibrotátil em Diferentes partes do corpo	Melhor nas extremidades (ponta dos dedos, lábios), diminuindo em direção oposta às extremidades do corpo (palma das mãos, antebraço, etc...)

A estimulação vibrotátil vem sendo utilizada como canal alternativo para transmissão de informações acústicas, e visando subsidiar o desenvolvimento de um estimulador vibrotátil de um canal. São apresentados no Anexo 1 resultados de diversas pesquisas psicofísicas para a estimulação vibrotátil.

3.2 Considerações sobre Estímulos

A estimulação vibrotátil da superfície da pele é geralmente empregada através da aplicação de estímulos com componentes de baixa intensidade e alta frequência. Estes estímulos possuem quatro dimensões básicas, que são frequência, amplitude, duração e local do corpo (abordados no Anexo 1). Um sinal de estimulação típico, contendo os três primeiros parâmetros, é mostrado na Figura 6.



Figura 6 - Sinal de Estimulação Vibrotátil Típico

3.3 Comunicação de Informações Acústicas Através de Estimuladores Vibrotáteis

Através da estimulação vibrotátil é possível transmitir a um deficiente auditivo informações acústicas do ambiente que o cerca, com grau de complexidade que pode variar desde a aquisição de informações como a localização do som até o reconhecimento de fonemas. Os estimuladores podem ser de um canal, dois canais ou múltiplos canais.

A habilidade de comunicação demonstrada por surdocegos através da percepção tátil de informações, junto com dados de estudos psicofísicos e de percepção das habilidades de resolução espacial e temporal do sistema tátil, indicam que o canal tátil é viável como opção para comunicações de informações do ambiente acústico.

Reed et al mencionam a habilidade de muitos surdocegos treinados no uso de métodos naturais para a percepção da voz pelo tato (Reed et al, 1985; Reed et al, 1989; Delhorne et al, 1988). No método de TADOMA, a mão do receptor é colocada na face e no pescoço do locutor (Weisenberg & Miller, 1987). A informação é obtida através da mão, que capta os movimentos faciais e do pescoço do locutor (movimentos dos lábios e queixo, fluxo de ar na boca, vibração da laringe). Estas informações são suficientes para uma boa recepção de voz. (Delhorne et al, 1988). A performance de usuários de métodos naturais atesta a capacidade do sentido do tato para o processamento dos sinais complexos provenientes do ambiente acústico, no entanto, a sua limitação é que requer contato físico entre o locutor e o receptor, dificultando a recepção de informações acústicas de fontes distantes. Estes métodos envolvem apenas os sentidos vibrotáteis e cinestésicos ou proprioceptivos e não requerem informações adicionais de outro canal sensorial, como a visão.

3.3.1 Percepção de Informações Acústicas

Segundo Russo & Behlau (1993) a percepção auditiva envolve a recepção e a interpretação de estímulos sonoros através do sentido da audição, em um processo altamente complexo e elaborado que se desenvolve a partir da representação interna do objeto ou acontecimento percebido nos centros auditivos do cérebro. A forma exata como tais estímulos são percebidos ainda não é totalmente compreendida, mas pode-se identificar alguns dos componentes da percepção auditiva, tais como a detecção, a sensação sonora, a discriminação, localização e reconhecimento.

Russo & Behlau (1993), citando Bothroyd & Hnath (1986), apresentam uma maneira didática de se compreender os componentes básicos da percepção acústica, através de perguntas sobre os sons e eventos que o produziram:

- *Detecção*: “Existiu som?”

- *Sensação*: “Como era o som?”
- *Discriminação*: “O som se modificou?” ou “Em que este som é diferente daquele outro?”
- *Localização*: “Onde está a fonte que produziu o som?”
- *Reconhecimento*: “Qual foi o evento que causou o som?”

As características de percepção acústica são abordadas nas seções seguintes, fazendo-se uma relação com os deficientes auditivos e a possibilidade de transmissão de tais componentes através de estimulação vibrotátil.

3.3.2 Estimulação acústica do ambiente (background)

Mesmo sem a atenção consciente do ouvinte, a estimulação auditiva dos ruídos ou sons ambientais serve para conectar o ouvinte normal com o mundo. A perda da sensação de estar conectado ao mundo, que é encontrada em pessoas com deficiência auditiva, pode levar a confusão e depressão, pois o estado de alerta inconsciente dos eventos ao redor, fornece um componente importante para o senso de segurança e bem-estar. Deficientes auditivos que não podem ouvir estas alterações no estímulo acústico do ambiente sempre reagem com desgosto às suas falhas, ao entender o que está acontecendo ao seu redor.

As alterações no nível ou natureza da estimulação acústica do ambiente servem como um sinal para que o ouvinte fique alerta a danos potenciais, permitindo que tome medidas preventivas ou defensivas. Este nível de alerta envolve não somente a identificação (detecção/sensação/discriminação) do sinal, mas também a habilidade para determinar a sua fonte (localização).

3.3.3 Detecção

Dispositivos de ajuda tátil empregando apenas um canal, e apresentando informações sobre a envoltória do sinal no tempo, referentes a eventos acústicos, podem fornecer estimulação suficiente para restaurar o sentimento de ligação com

o mundo. Enquanto este aspecto da habilidade auditiva pode ser fornecido por dispositivos de ajuda a escuta, como amplificadores, para aqueles com menores perdas auditivas, dispositivos táteis ou outros dispositivos sensoriais podem ter mais sucesso para aqueles profundamente surdos.

3.3.4 Discriminação e Identificação

A Discriminação é definida como a habilidade de determinar se um dado estímulo A é diferente de um estímulo B.

A Identificação é descrita como a habilidade de isolar características únicas do estímulo, juntamente com a capacidade de nomear o mesmo.

Nas situações do dia a dia há uma vasta gama de estímulos que podem ser discriminados e identificados, desde os sons ambientais, passando por acentuação de sílabas, fonemas isolados, palavras isoladas, frases até a discriminação de características da fala em velocidade normal.

3.3.4.1 Envoltória do sinal

A identificação e discriminação de alguns sons podem ser realizadas utilizando-se a envoltória da forma de onda do sinal acústico. O receptor ("ouvinte") pode extrair informações sobre o estímulo a partir das variações temporais e espectrais da envoltória da forma de onda. Este nível de análise pode ser apropriado quando se quer distinguir classes de som com grandes variações nas suas características de envoltória. Os sons ambientais satisfazem este critério, de forma que muitos sons ambientais podem ser identificados ou discriminados com base em características distintas da envoltória temporal. Mesmo um dispositivo tátil relativamente simples, de um canal, pode ser capaz de apresentar informações sobre a envoltória da forma de onda com fidelidade suficiente para a identificação de sons ambientais (WEISENBERG, 1992).

Através da envoltória do sinal de voz também é possível distinguir entre sons de voz ou características das mesmas, como o número de sílabas, sílaba tônica, e ritmo da fala. Estudos de codificação de informação auditiva têm

demonstrado que mesmo um simples sinal modulado em amplitude que reproduz a forma de onda temporal de voz pode fornecer informações importantes para o auxílio à leitura labial (CARNEY & BEACHER, 1986).

3.3.4.2 Parâmetros Espectrais

Para compreender a voz, é necessário analisar a distribuição espectral do estímulo acústico. No sistema auditivo normal esta análise é feita até com ambigüidade, de uma forma além da necessária para a identificação dos fonemas. O sistema tátil, porém, demonstra uma resolução pobre em freqüência, particularmente na faixa em que há uma melhor detecção absoluta do estímulo vibrotátil, em torno de 250 Hz.

Uma forma de parâmetro espectral que tem se mostrado efetivo em estudos auditivos é a freqüência fundamental, utilizada como um suplemento para a leitura labial (Carney & Beacher, 1986). Para a identificação e discriminação de fonemas, está claro que a informação espectral é de grande importância. Muitos sons podem ser distinguidos com base nas freqüências dos formantes (F1, F2 ou F3), ou com base na mudança da freqüência dos mesmos. Há discordâncias quanto a qual formante é mais informativo quanto à transmissão da unicidade de um fonema, ou se um contraste de diversos valores de formantes pode ser necessário.

3.3.5 Localização

A localização da fonte do sinal acústico constitui-se em uma tarefa mais complexa que a detecção. O dispositivo de localização ideal envolveria a comparação de duas ou mais informações, e requereria mais de um transdutor (ou ao menos mais de um modo de estimulação). O ouvinte normal localiza sinais de baixa freqüência principalmente com base na diferença de fase entre os dois ouvidos, e localiza sinais de alta freqüência com base na diferença de intensidade entre os sinais que chegam nos ouvidos (Whigtman & Kistler, 1989a; Whigtman &

Kistler, 1989b). Richardson (1982), no entanto, observou que o sistema tátil é insensível a diferenças de fase entre estímulos de vibração. Por outro lado, como o tato é sensível à diferença de intensidades, um dispositivo que fornece estímulos em dois ou mais locais separados espacialmente é adequado para fornecer informações para a localização da fonte do som.

Segundo GRANT et al (1985), a importância da informação sobre a localização da fonte sonora é duvidosa, em função da limitada capacidade de transmissão de informações do canal tátil. Assim, GRANT et al (1985) concluíram que pode melhor transmitir-se informações que auxiliem na realização de tarefas mais complexas do processamento das informações acústicas. Além disso, Weisenberger (1992) afirmou que uma avaliação da localização do som pode ser feita pela movimentação do usuário do estimulador, mesmo que o dispositivo seja de um canal; e que a detecção do estímulo permite que o indivíduo fique alerta, e use outros sentidos, como a visão, para localizar a fonte do som.

3.3.6 Processamento de Alto Nível

A habilidade para identificar fonemas individuais na fala pode ser necessária para o entendimento da conversação do dia-a-dia, mas a proficiência na identificação de fonemas em sílabas isoladas não garante a proficiência do entendimento da fala normal. Este processamento inclui a integração das diversas modalidades da informação da leitura labial com aquela fornecida pelo dispositivo tátil, e ainda operações cognitivas, como informações contextuais, de natureza semântica ou sintática. Na tentativa de integrar informações visuais e táteis no processamento de sinais de voz, a questão colocada é se as duas entradas devem ser redundantes ou complementares. O balanço ótimo da redundância versus a complementaridade da informação nos sinais provenientes das duas modalidades, tátil e visual-labial não foram ainda demonstrados empiricamente, contudo, informações sobre o ponto de articulação são prontamente disponíveis via leitura labial, quando informações sobre o modo de articulação ou excitação vocal em

consoantes são menos disponíveis (Weisenberger, 1992). A maioria dos dispositivos táteis multicanais não são bons transmissores de informações ou pistas referentes ao ponto de articulação, mas em muitos casos são capazes de fornecer informações sobre o modo e a excitação vocal. Esta complementaridade pode contribuir para altos níveis de performance na identificação de sinais de voz.

3.4 Dispositivos Táteis de um Canal

A simplicidade dos dispositivos de estimulação tátil de um canal facilita o desenvolvimento de dispositivos portáteis. Várias possibilidades podem ser consideradas no que se refere à informação acústica que pode ser transmitida em um dispositivo deste tipo. Na abordagem mais simples apresenta-se a informação acústica, de forma limitada, pela transmissão da envoltória do sinal. A utilidade desta informação foi estudada por vários pesquisadores. Carney & Beacher (1986) e Carney (1988) verificaram que um estimulador deste tipo é efetivo na transmissão de informações de número de sílabas e padrão de acentuação. Resultados similares foram obtidos por Weisenberger (1989), que concluiu que o estimulador apresentou bons resultados quanto à detecção de sons, identificação de sons ambientais, ritmo silábico e acentuação.

De uma forma geral, os parâmetros acústicos relevantes da forma de onda de voz utilizados na codificação de estimuladores vibrotáteis são:

1. Envoltória do sinal, que indica a sílaba ou palavra forte de uma sentença e também contém alguma informação dos fonemas (Van Tasell et al, 1987 Summers, 1992)
2. Freqüência fundamental, que indica acento e entonação e pode fornecer informações referentes aos fonemas e informações suprasegmentais, sinalizando a presença ou não da excitação sonora (Summers & Gratton, 1995)

3. Parâmetros espectrais gerais contêm informações referentes a fonemas, como excitação surda, sonora e mista (Summers & Gratton, 1995). A taxa de cruzamentos por zero é um parâmetro com estas características.

3.5 Dispositivos Táteis Multicanal

Os estimuladores táteis multicanais utilizam um conjunto de transdutores para transmissão da informação acústica. Na configuração mais comum o sinal de áudio é dividido em faixas de frequência, por um banco de filtros (Figura 7), cujo sinal na saída de cada filtro passa por um detector de envoltória e depois alimenta os transdutores. Desta forma, consegue-se transmitir a informação espectral do sinal de áudio, mapeada espacialmente na saída de cada transdutor.

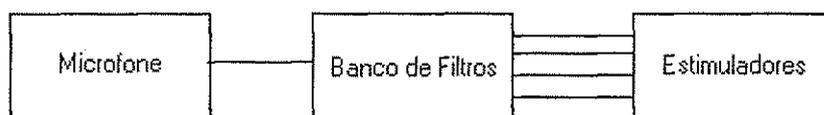


Figura 7 – Estimulador Vibrotátil Multicanal Típico

Brooks et al (1985) descrevem um experimento com um adulto utilizando o vocoder da Universidade de Queen, um dispositivo com 16 canais com um conjunto de transdutores dispostos linearmente, aplicados no ante-braço. O usuário adquiriu um vocabulário de 250 palavras utilizando apenas o vocoder, além de ter obtido resultados muito bons com a leitura labial. A performance deste indivíduo forneceu evidências que os dispositivos de ajuda tátil podem ser uma opção viável para a transmissão de informação acústica para surdos.

3.5.1 Número de canais

Existe uma relação custo-benefício entre o número de canais e a miniaturização e portabilidade do estimulador. Os transdutores também representam a maior parte do consumo de energia do dispositivo, e um dispositivo tátil comercial deve incorporar o menor número de canais transmitindo quantidade de informação suficiente para o usuário (Weisenberger, 1992).

É difícil avaliar o número de canais ideal para um estimulador multicanal, pois não existem trabalhos na literatura comparando estimuladores diferentes apenas quanto ao número de canais. Weisenberger (1989) comparou a performance do TACTAID V, um dispositivo de 5 canais com o Queen's Vocoder, de 16 canais. Como resultado, observou que os dois estimuladores tinham performance semelhante, embora a performance do estimulador de 16 canais foi melhor para situações mais complexas. Outro ponto observado na literatura é que a performance de dispositivos táteis em tarefas segmentadas, com conjunto fechado de vocabulário não é um indicativo dos resultados a serem observados na conversação fluente, ou fala conectada.

3.6 Estimulação Vibrotátil e as Questões do Ensino da Fala

A produção da fala envolve o controle de diversas ações do aparelho fonador, como o posicionamento dos articuladores, o controle do fluxo respiratório, a coordenação dos articuladores, a coordenação dos elementos respiratórios, articulatórios e fonatórios da fala, a produção de fonação adequada (entonação e qualidade), o controle dos padrões articulatórios da laringe para a produção do contraste surdo/sonoro, e o suporte respiratório. Para Behlau & Pontes (1995), por exemplo, a produção da fala exige uma série de atos coordenados pelo cérebro, que ocorrem na seqüência:

o cérebro dispara um comando para a laringe e articuladores; inicialmente precisamos inspirar o ar, o que requer um afastamento das pregas vocais; na seqüência as pregas vocais devem aproximar-se entre si, com tensão adequada, controlando e bloqueando a saída de ar dos pulmões; a pressão do ar coloca em vibração as pregas vocais; as caixas de ressonância devem estar ajustadas para facilitar e amplificar a saída do som pela boca; os articuladores devem ser posicionados adequadamente propiciando as necessárias freqüências de ressonância.

A realimentação dos sons produzidos permeia todas as etapas do aprendizado da fala, e o surdo precisará de realimentação externa até desenvolver uma realimentação própria, interna, que pode ser baseada no próprio sentido cinestésico envolvido no aparelho sensoriomotor ou auxiliada por dispositivos como amplificadores, implante coclear ou "tactile aids". Nas subseções seguintes é apresentada uma revisão de trabalhos referentes a aplicação da estimulação tátil nas questões essenciais para o aprendizado da fala.

3.6.1 A Questão da Intensidade da Fala

A intensidade vocal no deficiente auditivo é geralmente alterada e pode ser forte, por ser uma forma de se fazer ouvir ou pela excessiva tensão do trato vocal; ou ser fraca, quando existe a consciência de uma tendência em se falar forte (Rodrigues, 1997).

Wilson (1993), na terapia vocal para intensidade, sugere correlacionar vários graus de intensidade como muito forte, forte, correto e muito fraco, com movimentos musculares, por exemplo, batidas em um tambor, e ainda em situações concretas, como: chamar alguém do outro lado da rua ou falar com alguém que está próximo, sem que mais ninguém precise ouvir. Wilson (1993)

lembra que quando planejados cuidadosamente, os programas de modificação de intensidade ajudam o deficiente auditivo a obter uma comunicação mais eficaz.

A estimulação vibrotátil pode ser utilizada para realimentar o surdo sobre informações de intensidade da fala, sendo um componente auxiliar, por exemplo, às atividades descritas por Wilson (1993). Neste sentido, a envoltória do sinal é um parâmetro adequado para realimentar informações referentes à energia/altura da fala, dentro de um módulo desenvolvido especificamente para o ensino da fala.

3.6.2 A Questão da Respiração

O sistema respiratório comporta-se na fala um pouco diferente do que na respiração normal. Na fala se procura manter uma pressão constante subglotal durante a expiração, a fim de permitir uma fonte de energia constante para as pregas vocais. Isto requer uma inspiração mais rápida, que é seguida por uma expiração mais lenta, mais controlada (Almeida, 1997).

O padrão respiratório no deficiente auditivo apresenta-se geralmente adequado na função isolada, mas com problemas para coordenar fala-respiração. O deficiente auditivo tem dificuldade de sincronizar o início da fonação com início da respiração, tendendo a expelir ar antes de falar, existindo uso de ar de reserva e interrupção de frases, com pausas em momentos inadequados (Rodrigues, 1997).

Segundo Araújo (2000), em repouso a inspiração corresponde a 40% do ciclo e a expiração corresponde a 60% do ciclo. Na fala, a duração da expiração se amplia para 90% do ciclo e a inspiração se reduz para 10% do ciclo. Assim, o Tempo Máximo de Fonação (TMF) é uma das medidas fonatórias mais simples, cuja importância está no fato do falante, na fala encadeada, realizar recargas aéreas a cada um terço de seu tempo máximo de fonação. Citando Bloch (1963), Araújo (2000) menciona que o maior segredo da boa voz está na coordenação da respiração com a emissão sonora.

Araújo & Lima (1999) constataram que o TMF médio de deficientes auditivos pré-adolescentes estava 42,5% abaixo do padrão apresentado por locutores

ouvintes na mesma faixa etária, e citando diversos autores, Araújo (2000) mostrou que a dificuldade do deficiente auditivo na coordenação respiratória leva a uma menor produção de sílabas por ciclo respiratório e a um cansaço prematuro, que resultam na alteração do ritmo da fala e na perda da inteligibilidade da mesma.

Ao longo da primeira infância aprende-se que, durante a fala, deve-se inspirar rapidamente o ar e que este deve ser expirado lentamente durante a sua produção. Atividades para aprimoramento do controle respiratório para produção da fala visam estimular o aumento do tempo máximo de fonação, ou seja, aumentar o tempo máximo em que o indivíduo pode manter o fluxo respiratório, respeitando o ar de reserva (Araújo, 2000).

Um parâmetro simples para acompanhar o período de fonação, possibilitando o cálculo do TMF, é a intensidade ou energia do sinal acústico. Através do acompanhamento da energia do sinal de voz é possível detectar a presença de fonação (início e término), quando está ocorrendo a expiração. Assim, é possível, a partir deste parâmetro, gerar realimentações específicas que permitam ao surdo ou surdocego treinar e controlar os tempos de expiração e inspiração durante a fonação.

Segundo Weisenberger (1992), a detecção do sinal de voz é a tarefa mais simples que um ouvinte realiza. O grau de processamento requerido para a detecção de voz é mínimo, sendo suficiente apenas a presença ou ausência de energia do sinal.

Os estimuladores vibrotáctis de um canal, quando alimentados pelo sinal de voz ou modulados pela envoltória do sinal permitem a detecção do som e a identificação de sons ambientais (Carney & Beacher, 1986; Carney, 1988; Weisenberger, 1989), de forma que a envoltória do sinal é um parâmetro adequado para a realimentação de informações referentes às fases da inspiração e expiração durante a fala.

3.6.3 A Questão da Entonação

A entonação elevada e sua excessiva variabilidade são sintomas vocais comuns nos deficientes auditivos, e que os mesmos podem ser sanados através de uma orientação externa que substitua a realimentação auditiva.

Araújo (2000) citando vários autores, colocou que os deficientes auditivos apresentam frequência fundamental mais elevada que o normal, especialmente os do sexo masculino, sendo que este problema ocorre mais nos adolescentes do que em crianças surdas.

Bush (1981) observou uma forte relação entre a variação da frequência fundamental e a produção de diferentes vogais por surdos norte-americanos, o que afeta a inteligibilidade da fala. Os surdos que possuíam frequência fundamental média mais alta que a dos falantes normais apresentaram maiores variações na frequência fundamental na pronúncia das diferentes vogais. Este fato não ocorreu com os que possuíam frequência fundamental média no nível da dos falantes normais. A autora concluiu que há necessidade dos surdos adquirirem maior controle sobre as variações da frequência fundamental, através de treinamento. Para Monsen (1979), é importante ressaltar que não só a frequência fundamental deve estar em uma faixa apropriada, mas que a transição de um formante para outro esteja dentro dos padrões da normalidade, ou seja, ressalta as características dinâmicas da entonação.

Araújo & Lima (1999) mediram a frequência fundamental de deficientes auditivos pré-adolescentes e obtiveram valores 56% acima da média apresentada por ouvintes na mesma faixa etária. Este valor de frequência fundamental, inadequado para a realização de vogais fechadas como /i/ e /u/, resulta em um padrão harmônico restrito, que diminui a inteligibilidade da fala.

Os esforços para auxiliar na correção da frequência fundamental dos surdos por meio de instrumentos de realimentação visual têm registros desde 1935 (Hudgins, 1935; Sterne & Zimmerman, 1939), sendo que atualmente existem diversos softwares que, entre outras funções, possuem atividades para

treinamento do controle da frequência fundamental (Araújo, 2000; Oliveira, 1998; Novaes & Mendes, 1998; Mashie, 1995).

Em termos de estimulação tátil, não existem sistemas específicos para auxiliar no treinamento do controle da frequência fundamental e ajuste de seus valores. No entanto, existem sistemas desenvolvidos com o objetivo de auxiliar a leitura labial através da transmissão por estímulos táteis de informações sobre a frequência fundamental (Breeuwer & Plomp, 1985; Grant et al, 1985; Bernstein et al, 1989, 1993; Eberhardt et al, 1990).

Segundo Bernstein (1995), com base em diversos trabalhos, a transmissão da frequência fundamental pelo tato pode ser eficiente pelas seguintes razões: (1) A faixa de variação da frequência fundamental F_0 situa-se, ou então pode ser facilmente trasladada, na faixa de melhor sensibilidade vibrotátil (Rothenberg & Molitor, 1979); (2) As características da prosódia da fala transmitidas pela frequência fundamental variam lentamente, sendo adequadas a capacidade de processamento temporal do tato (Rothenberg et al, 1977).

Os resultados de várias pesquisas têm mostrado que é possível identificar-se características da prosódia, como a entonação e as sílabas acentuadas apenas com o uso de dispositivos táteis (Bernstein et al, 1989; Rothenberg & Molitor, 1979), embora com resultados mais modestos que aqueles obtidos pela transmissão acústica da frequência fundamental.

As comparações realizadas entre estimuladores multicanais e monocanais não apresentam vantagens de nenhum dos tipos de estimuladores com relação ao outro, quanto aos resultados na transmissão da frequência fundamental (Bothroyd & Hnath, 1986; Hanin et al, 1988). Os estimuladores monocanais, no entanto, têm vantagens com relação à portabilidade, consumo de potência, aceitação cosmiética, e a possibilidade de uso de canais adicionais para outras características da fala.

Eberhardt et al (1990) investigaram a transmissão da frequência fundamental da fala por um estimulador vibrotátil de um canal. Segundo os autores, a faixa ideal para a codificação da frequência do sinal de estimulação

vibrotátil varia de 10 a 100 pulsos por segundo, e os sistemas que trasladam a frequência fundamental com frequência central (média) em torno de 50 pulsos por segundo tendem a ter mais sucesso na transmissão da entonação. Bernstein et al (1989) também utilizaram um estimulador vibrotátil monocanal para transmissão da frequência fundamental aplicada como auxílio (suplemento) para a leitura labial, sendo que dos métodos de codificação que foram estudados, as transformações que apresentaram melhores resultados para a percepção da entonação foram:

$$P_t = \frac{1}{50 + (F_v - M) * S} , \text{ (codificação linear)} \quad (1)$$

$$P_t = \frac{1}{50 + (\log(F_v/M)) * S} , \text{ (codificação logarítmica)} \quad (2)$$

Nas quais, P_t é o período do pulso vibrotátil, F_v é a frequência fundamental detectada, M é a média das variações da frequência fundamental do locutor, e S é um fator de escala escolhido para manter ao menos 96% do valor de P_t na faixa de 0,01 a 0,1 segundos.

Assim, conclui-se que o parâmetro mais adequado para transmissão da entonação através de estímulos vibrotáteis é a própria frequência fundamental. O estimulador monocanal apresenta vantagens com relação ao multicanal, e o mesmo desempenho em termos de transmissão da informação.

3.6.4 A Questão da Articulação

A articulação está relacionada ao emprego dos órgãos fonoarticulatórios na produção da fala, ou seja, ao processo de ajustes motores dos órgãos fonoarticulatórios na produção e formação dos sons, que se dividem em vogais e consoantes.

3.6.4.1 As Vogais

Spinelli et al (1999) apresentam alguns recursos que podem ser utilizados para auxiliar o ensino da fala de vogais. Entre estes recursos estão algumas pistas táteis utilizadas para cada vogal, que são:

/a/ - vibração da laringe, colocando-se a mão no pescoço;

/i/ - mão na laringe, emitindo /a/ e /i/, para perceber a elevação da laringe neste último;

/e/ - mão embaixo da mandíbula, emitindo /a/ e /e/, para perceber a diferença;

/é/ - mão no pescoço, emitindo-se /e/ e /é/, para evidenciar o pequeno abaixamento da laringe;

/o/ - mão no pescoço, percebendo o abaixamento da laringe;

/ó/ - mão no pescoço, comparando o abaixamento da laringe entre /o/ e /ó/;

/u/ - mão no pescoço, percebendo que o abaixamento da laringe é maior do que na emissão das demais vogais.

Pode-se observar então que estas pistas estão relacionadas ao estágio inicial do aprendizado da fala, a emissão, e são características articulatórias gerais. Assim, como as informações acústicas podem ser geradas por configurações articulatórias diferentes, a depender da pessoa que fala, estas pistas constituem-se apenas de um estágio inicial do aprendizado. Através do resultado acústico, o indivíduo deve adquirir seu próprio modelo de posicionamento fonoarticulatório para a produção das vogais. Esta é uma das motivações para o desenvolvimento de ferramentas que possibilitem o treinamento de sons da fala e a automonitoração dos seus resultados. Também se pode observar que algumas vogais são aprendidas em comparação com outras, e a existência de uma realimentação tátil, direta, no momento do ensino destas posições articulatórias, pode auxiliar o surdo em seu aprendizado.

No Brasil, foram desenvolvidos sistemas de realimentação visual para auxílio ao ensino de vogais por Oliveira (1998) e Araújo (2000). Em Oliveira foi desenvolvido um módulo que realimenta visualmente o surdo com informações sobre o posicionamento dos articuladores, derivadas do sinal acústico da fala do surdo, e permite a comparação com uma figura esperada. Araújo desenvolveu alguns jogos, cujo controle é feito através das características da fala do surdo.

Em termos de estimulação tátil, Green et al (1983) estudaram a identificação de vogais através de estímulos vibrotáteis utilizando um estimulador vibrotátil multicanal, contendo uma matriz de estimuladores distribuídos em 24 linhas e 6 colunas, de acordo com o que se pode observar na Figura 8.

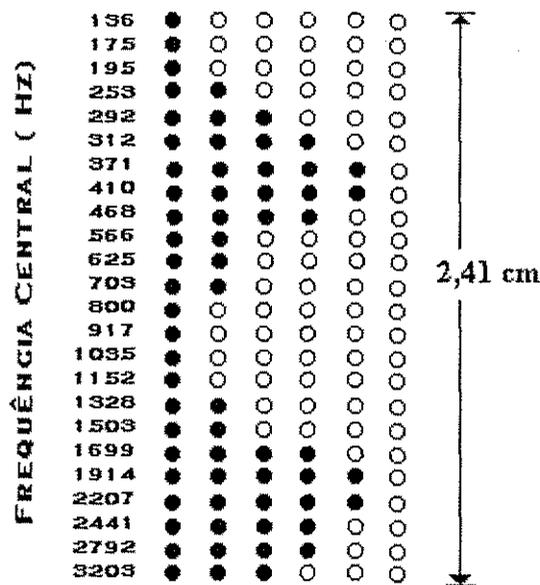


Figura 8 - Representação de sinal na saída de uma matriz de estimuladores, para o segmento de voz /bi/, da língua inglesa. Cada ponto representa um estimulador. Cada ponto preto representa um estimulador ativo.

O sinal da fala passa pelo conjunto de filtros, cujas frequências centrais podem ser vistas na Figura 8. A informação da envoltória do sinal, para cada banda de frequências selecionada, modula uma linha de estimuladores. Assim,

todos os estimuladores de uma linha vibrarão quando ocorrer um sinal de máxima amplitude na respectiva banda de freqüências.

Green et al (1983) concluíram que o display espectral utilizado na ponta do dedo fornece informações suficientes para a identificação de vogais estacionárias, no contexto de sílabas CV (consoante-vogal). Isto pode ser explicado pelo display espectral apresentar os picos de intensidade nas freqüências dos primeiro e segundo formantes das vogais, como pode ser observado na Figura 8, para o caso da vogal /i/ no segmento /bi/. Apesar do resultado positivo com relação à possibilidade de discriminação entre vogais estacionárias, ou sustentadas, não se encontra na literatura a continuidade destas pesquisas. Isto se explica pelo interesse dos pesquisadores no desenvolvimento de sistemas para transmissão de informações da fala, e não pelo desenvolvimento de etapas parciais como esta. No entanto, pensando-se em um sistema aplicado apenas para o auxílio ao ensino de vogais, tais resultados, tanto da análise do sinal de voz por divisão espectral, como pela aplicação de estimuladores multicanais, podem ser promissores.

O estudo da parametrização de sinais de fala para o reconhecimento de vogais é uma etapa importante no desenvolvimento de um sistema de auxílio ao ensino da fala. E esta parametrização pode variar com a língua em questão.

As vogais orais do português brasileiro foram estudadas por Russo & Behlau (1993), e de acordo com seus resultados os valores médios dos três primeiros formantes, encontram-se representados na Figura 9.

Na análise dos sinais de fala, as duas primeiras formantes (F1 e F2) são de maior importância, e concentram a maior parte da energia do sinal de voz. O reconhecimento de vogais em geral passa pela análise e obtenção dos valores das primeiras formantes, no entanto, diversos fatores limitam as técnicas para extração de formantes, entre as quais se destacam as diferenças físicas entre locutores, a diferença entre diversas locuções de um mesmo locutor, as diferenças de idade e sexo (Araújo, 2000).

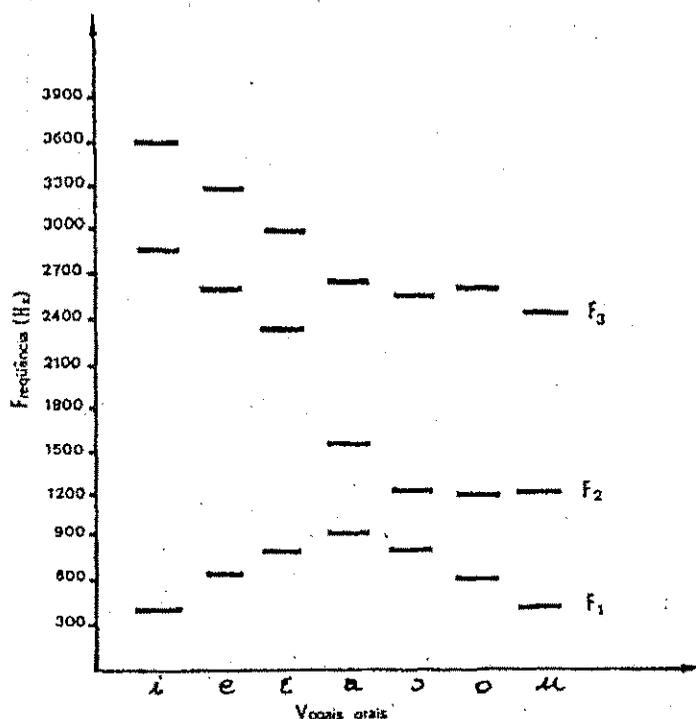


Figura 9 - Média das frequências dos três primeiros formantes, das vogais orais do português brasileiro (Russo&Behlau, 1993)

Para as vogais brasileiras, os valores de F2 são praticamente indistintos de uma vogal para outra, com pares de formantes próximas. No entanto, valores elevados de F2 estão associados às vogais anteriores, enquanto baixos valores de F2 estão associados às vogais posteriores. Assim, Araújo (2000) propôs um algoritmo que consiste em determinar a altura da vogal pela determinação de F1, e obter a posição articulatória a partir da distribuição espectral, ou seja, permitindo classificar a vogal em anterior ou posterior.

3.6.4.2 As Consoantes

Spinelli et al (1999) apresentam alguns recursos que podem ser utilizados para auxiliar o ensino da fala de consoantes. Observando as estratégias adotadas, se nota a divisão em dois grandes grupos majoritários: oclusivas e fricativas, com técnicas de ensino semelhantes, do ponto de vista de que primeiro são ensinados

os fonemas de articulação mais externa, e na seqüência, por analogia, as articulações mais internas. Este processo é resumido a seguir, com o objetivo de mostrar que os dois grupos são aprendidos separadamente, de forma que é possível parametrizar as consoantes considerando-se apenas as características de variação da fala restritas ao grupo que está sendo ensinado.

Grupo Oclusivas /p, b, t, d, k, g, m, n, h/:

- /p/ é aprendido inicialmente;
- /b/ é aprendido por analogia com o /p/;
- /t/ é aprendido isoladamente, podendo-se utilizar analogia com o /p/;
- /d/ é aprendido por analogia com o /t/;
- /k/ pode ser ensinado por analogia com /p/ e /t/;
- /k/ pode ser ensinado por analogia com /b/ e /d/;
- /m/, /n/ e /h/ : /m/ pode ser aprendido inicialmente, e /n/ e /h/ são ensinados por analogia; Faz-se comparações com /d/ e /b/ para desenvolver o ponto de articulação e a nasalidade dos fonemas.

Grupo Fricativas /f, v, s, z, ch, j/:

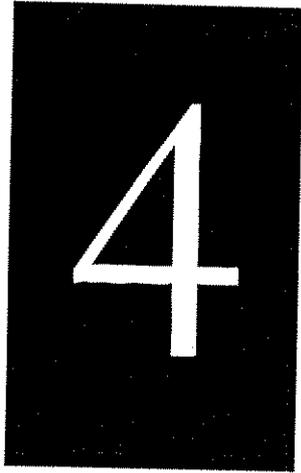
- /f/ é ensinado inicialmente, pois o ponto de articulação é anterior;
- /v/ tem o mesmo ponto de articulação do /f/ e estes dois fonemas são considerados visemes. O /v/ é ensinado comparando-se com o /f/;
- /s/ ensinado em comparação com o /f/ quanto à constrição da corrente de ar;
- /z/ forma, junto com o /s/, um viseme. O /z/ é ensinado por analogia com o /s/.
- /ch/ é ensinado fazendo-se analogia com o /s/;
- /j/ é ensinado por analogia com o /ch/

Além destes grupos, Spinelli et al fornecem pistas táteis, visuais e cinestésicas para o ensino e treinamento da pronúncia das consoantes vibrantes e laterais.

Segundo Russo & Behlau (1993), os fricativos são as consoantes mais fracas e mais agudas do português, representando uma tarefa de reconhecimento difícil para os deficientes auditivos. O par / f, v/, extremamente débil, é porém auxiliado pela leitura orofacial, o que não ocorre com / ch, j /. Além disso, as fricativas exigem, como no caso do /s/, a participação máxima de todos os parâmetros articulatórios, possibilitando o exercício de uma ampla gama de músculos necessários à produção da fala, como também ao aprendizado do sistema tátil-cinestésico (Araújo, 2000).

A existência de componentes significativas em altas frequências, sugere a utilização de taxas de cruzamentos por zero para modelar e localizar quadros contendo fricativas (Rabiner & Schafer, 1978; Vieira, 1989). O centro de gravidade espectral definido como sendo a frequência abaixo da qual estão concentradas 50 % da energia também pode ser utilizado para a caracterização das fricativas. Araújo (2000) utilizou dois parâmetros para, junto com o traço surdo/sonoro, diferenciar as fricativas. O primeiro foi a Frequência de Pico Suavizada, que sugerem a frequência para a qual a amplitude do espectro é máxima, a frequência de pico, como um parâmetro para diferenciar /s/ de /z/. O segundo foi uma medida do espalhamento espectral, denominada coeficiente de espalhamento espectral.

Em termos de estimulação tátil, Barros et al (2002) fizeram um estudo sobre a parametrização de sinais de voz para codificação de um estimulador vibrotátil monocanal, com ênfase no auxílio à leitura labial, e observaram que a taxa de cruzamentos por zero trouxe vantagens no auxílio à discriminação entre os visemes compostos por fricativos.



**Codificação de Parâmetros
em Estímulos Vibrotáteis**

4.1 Introdução

Os procedimentos para o ensino da fala podem ocorrer com a utilização de práticas pedagógicas diretas ou indiretas. Na abordagem direta o surdo ou aprendiz deve estar consciente do processo educativo, e deve fazer voluntariamente os exercícios fonoarticulatórios (Spinelli et al, 1999), que são intensos e repetitivos. Na abordagem indireta os objetivos dos exercícios não estão declarados nas atividades propostas. Eles são alcançados através da atuação do surdo/aprendiz em jogos e brincadeiras. Esta abordagem é utilizada principalmente com crianças pequenas e com aquelas que não se dispõem a participar diretamente do processo de treinamento fonoarticulatório (Spinelli et al, 1999).

Os sistemas de auxílio ao ensino da fala por realimentação visual apresentam geralmente a informação da fala por meio de gráficos e jogos, ou seja, de uma forma indireta.

Os sistemas táteis descritos na literatura pesquisada foram desenvolvidos para auxílio à percepção da fala ou como auxílio à leitura labial, e quando utilizados no auxílio ao aprendizado da fala enquadram-se na classe de abordagem direta do ensino da fala. Nestes sistemas um número limitado de parâmetros é extraído do sinal de voz e são utilizados para modular o sinal de estimulação tátil.

Este capítulo trata sobre a codificação de parâmetros em sinais de estimulação tátil, de forma a permitir o uso de sistemas táteis em abordagens indiretas do ensino da fala, baseada em jogos e em indicações de acerto de alvo.

A faixa de frequências perceptíveis pela estimulação vibrotátil não é a mesma da voz humana pelo sistema auditivo, de modo que é preciso que seja feito um mapeamento do parâmetro extraído da voz em um valor contido na faixa de frequências adequada para estimulação vibrotátil. Este é um procedimento tradicional utilizado no desenvolvimento de sistemas táteis para envio/recepção de informações de voz.

Propõe-se neste trabalho um mapeamento de parâmetros de voz em estímulos vibrotáteis baseado nas necessidades de ensino/aprendizagem da fala. Assim, propomos enviar por estímulos táteis, informações referentes a comparações com um padrão de referência, além de informações relacionadas à variação da fala, modulando o estímulo tátil com a variação do erro entre o valor do parâmetro extraído e um valor pré-definido (alvo).

Na Figura 10 podemos ver os gráficos representativos de duas transformações discretas para mapeamento da variação do parâmetro de voz extraído nas frequências de estímulos pulsados vibrotáteis. Na transformação com dois níveis o indivíduo recebendo estímulos vibrotáteis tem duas informações: acertou o alvo ou não acertou o alvo. Na transformação com três níveis, além das duas informações anteriores, o indivíduo sob estímulo vibrotátil pode saber se está falando em um nível acima ou abaixo do valor alvo. Aumentando-se os níveis discretos aumenta-se então a quantidade de informação transmitida, embora haja um aumento na complexidade da percepção vibrotátil. Considerando-se que o tato possui vantagem na transmissão de informações dinâmicas da voz, quando comparado à realimentação visual, o aumento do número de níveis, torna possível a transmissão da variação dinâmica da voz. No caso extremo, estão as transformações contínuas.

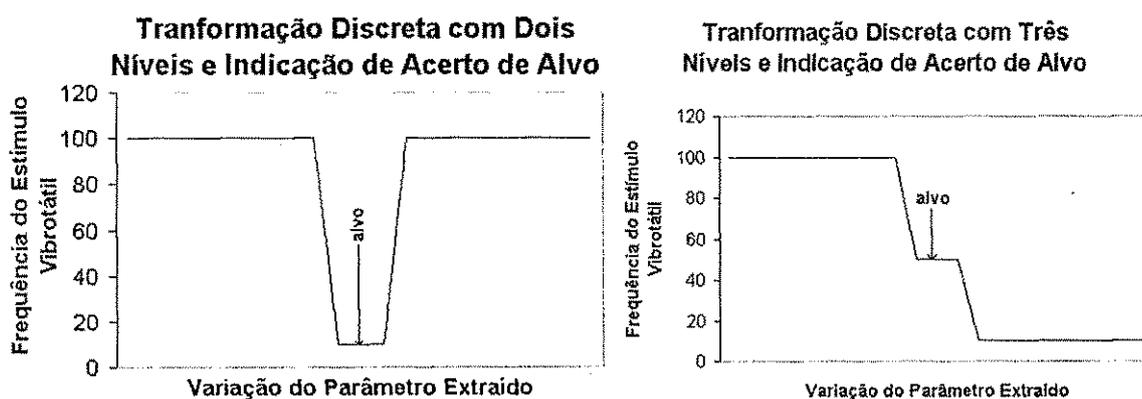


Figura 10 - Transformações Discretas com Indicação de Acerto de Alvo

$F_t(t)$ deve variar entre f_{min} e f_{max} , $P_{e1}(t)$ deve variar entre os valores p_{11} e p_{21} , e $P_{e2}(t)$ deve variar entre os valores p_{12} e p_{22} , definidos pelo usuário do sistema. Pode-se observar na Eq. (23) que quando $S = 0$, o grau de dificuldade é máximo, pois a realimentação independe de $P_{e1}(t)$, $P_{e2}(t)$, P_{a1} e P_{a2} . Quando $S = S_{max}$, deve-se ter a máxima excursão de $F_t(t)$, caracterizando o grau de menor dificuldade do exercício.

Assim, dado uma frequência F_c , os parâmetros alvo P_{a1} e P_{a2} , e os limites superiores das faixas dos parâmetros detectadas, p_{21} e p_{22} , poderemos obter a expressão de S_{max} , p_{11} e p_{12} para a transformação da equação 23.

Consideramos as seguintes relações:

$$P_{e1}(t) = p_{11}; P_{e2}(t) = p_{12} \Rightarrow F_t(t) = f_{max}, \quad (24)$$

$$P_{e1}(t) = p_{21}; P_{e2}(t) = p_{22} \Rightarrow F_t(t) = f_{max} \text{ e} \quad (25)$$

$$F_c = f_{min} \quad (26)$$

Substituindo-se as igualdades (24), (25) e (26) em (23), obtém-se:

$$S_{max} = \frac{f_{max} - f_{min}}{(p_{21} - P_{a1} + p_{22} - p_{a2})^2} = \frac{f_{max} - f_{min}}{(P_{a1} - p_{11} + P_{a2} - p_{12})^2} \quad (27)$$

De onde obtém-se:

$$p_{11} = -p_{21} + 2P_{a1}, \quad (28)$$

$$p_{12} = -p_{22} + 2P_{a2}, \quad (29)$$

4.3.1.1 Simulações

Com o objetivo de realizar uma avaliação do comportamento da transformação, através da visualização gráfica da equação (23), calculou-se a dependência de $F_t(t)$ em função de parâmetros genéricos $P_{e1}(t)$ e $P_{e2}(t)$ a partir da

Eq. (23). Foram escolhidos $f_{min} = 10\text{Hz}$, e $f_{max} = 100\text{Hz}$, que implicam $F_c = 10\text{Hz}$. As simulações foram realizadas no ambiente do MATLAB. Os valores de p_{11} e p_{12} utilizados para obtenção dos resultados foram calculados a partir das Eqs. (28) e (29), respectivamente, e os valores de S_{max} foram calculados a partir da Eq. (27).

Pode-se observar nas Figuras 12 e 13 a variação de $F_t(t)$ em função de $P_{e1}(t)$ e $P_{e2}(t)$ considerando-se os parâmetros limite $p_{11} = p_{12} = 1$ e os alvos $p_{a1} = p_{a2} = 0,5$. Na Figura 12 apresenta-se o diagrama em três dimensões, e na Figura 13 o diagrama em duas dimensões.

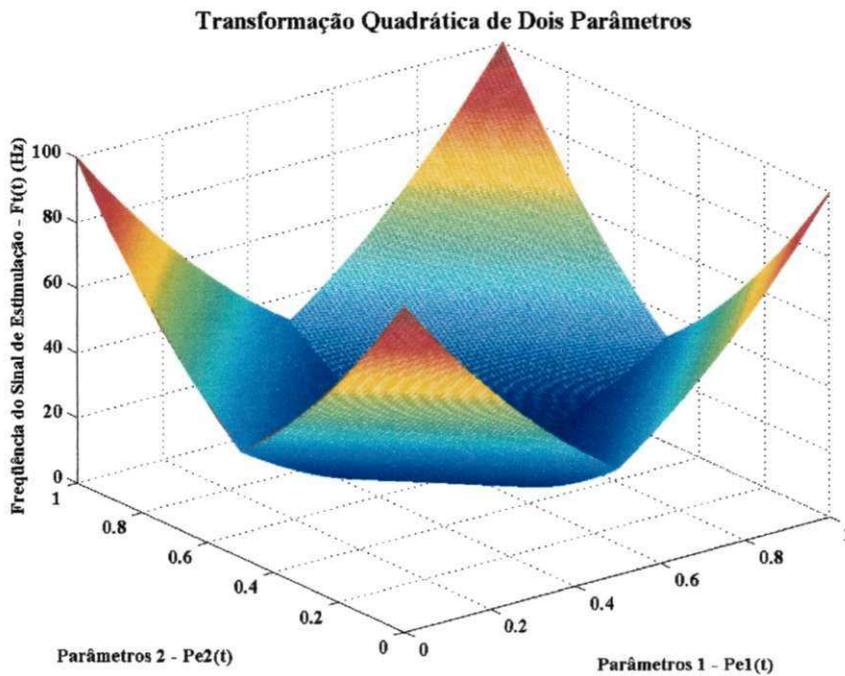


Figura 12 - Dependência de $F_t(t)$ em função de $P_{e1}(t)$ e $P_{e2}(t)$ para a transformação quadrática

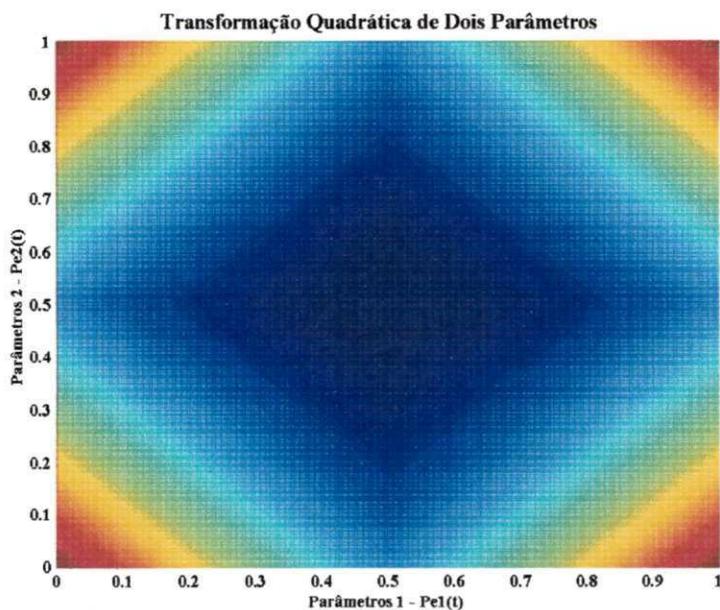


Figura 13 - $F_i(t)$ em função de $P_{e1}(t)$ e $P_{e2}(t)$ para a transformação quadrática(2-D)

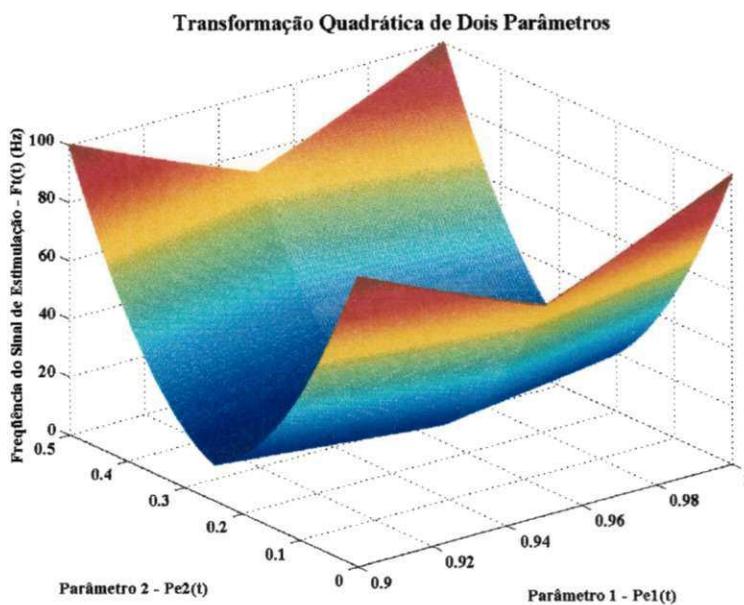


Figura 14 - Dependência de $F_i(t)$ em função de $P_{e1}(t)$ e $P_{e2}(t)$ para a transformação quadrática com $(p_{22}-p_{12}) > (p_{21}-p_{11})$

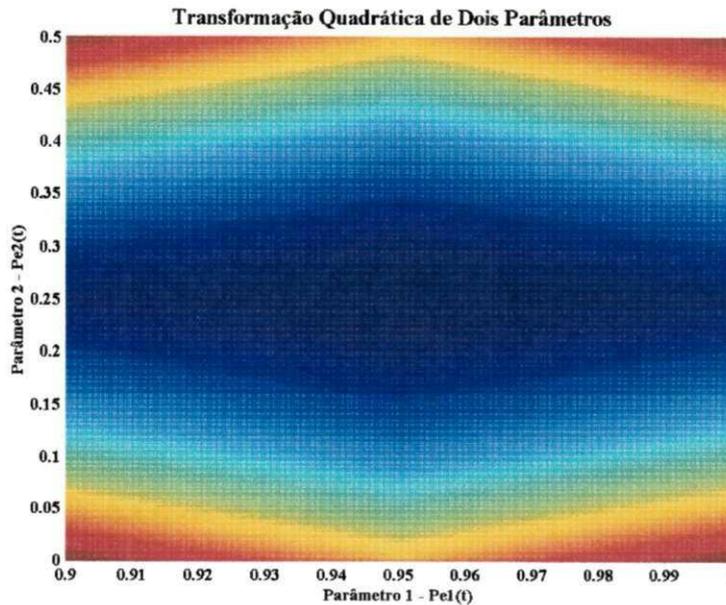


Figura 15 - $F_t(t)$ em função de $P_{e1}(t)$ e $P_{e2}(t)$ para a transformação quadrática (2-D) com $(p_{22} - p_{12}) > (p_{21} - p_{11})$

Nas Figuras 14 e 15 pode-se observar a variação de $F_t(t)$ em função de $P_{e1}(t)$ e $P_{e2}(t)$ considerando-se os parâmetros limite $p_{11} = 1$, $p_{12} = 0,5$ e os alvos $p_{a1} = 0,95$ e $p_{a2} = 0,25$, ou seja, as variações em $P_{e1}(t)$ são, em termos absolutos, bem maiores do que em $P_{e2}(t)$. Observa-se que neste caso é possível acertar-se alvo quase que apenas em função do parâmetro P_{e2} , e isto significa que neste caso a variação de $P_{e1}(t)$ é desprezível.

De fato, como o termo $(p_{21} - p_{a1})$ é muito menor que $(p_{22} - p_{a2})$ no exemplo da Figura 14, o cálculo de S_{max} , conforme a equação (27), é influenciado majoritariamente por p_{22} e p_{a2} . Assim, a transformação de dois parâmetros torna-se semelhante a uma transformação de um parâmetro, pois apenas $P_{e2}(t)$ influencia a frequência do estímulo pulsado.

4.3.1.2.1 Simulação

O comportamento da transformação de $F_t(t)$ em função de parâmetros genéricos $P_{e1}(t)$ e $P_{e2}(t)$ foi avaliado a partir de simulações baseadas nas equações (32- 38). Foram escolhidos $f_{min} = 10\text{Hz}$, e $f_{max} = 100\text{Hz}$, que implicam $F_c = 10\text{Hz}$.

Pode-se observar nas Figuras 16, 17 e 18 o diagrama em duas dimensões da variação de $F_t(t)$ em função de $P_{e1}(t)$ e $P_{e2}(t)$, para as condições dadas pelas equações (33), (34) e (35), respectivamente. Na Figura 16, tem-se $b = b_1 = b_{\text{equilíbrio}}$. Na Figura 17, $b = 2*b_1$, e na Figura 18, $b = 0,5*b_1$. Consideraram-se os parâmetros limite $p_{11} = 1$, $p_{12} = 0,5$; os alvos $p_{a1} = 0,95$ e $p_{a2} = 0,25$, e o coeficiente $a = 1$.

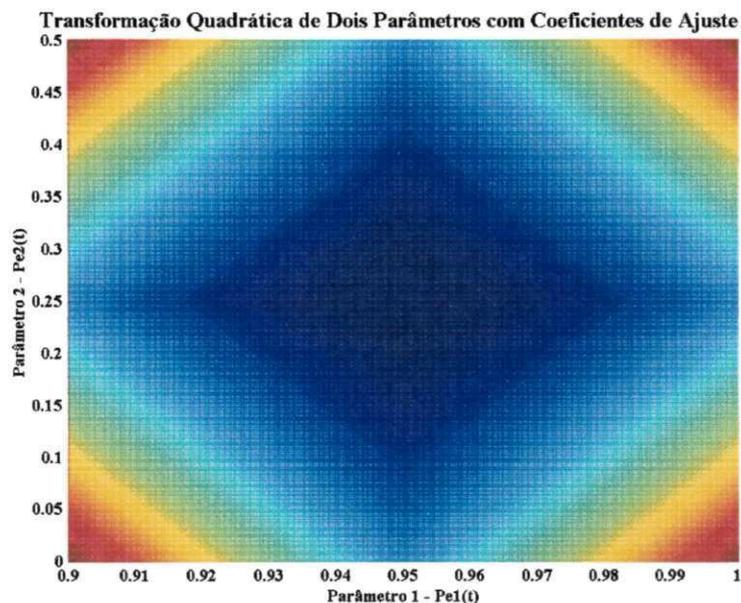


Figura 16 - $F_t(t)$ em função de $P_{e1}(t)$ e $P_{e2}(t)$ para a transformação quadrática(2-D) com b dado por Eq. (31)

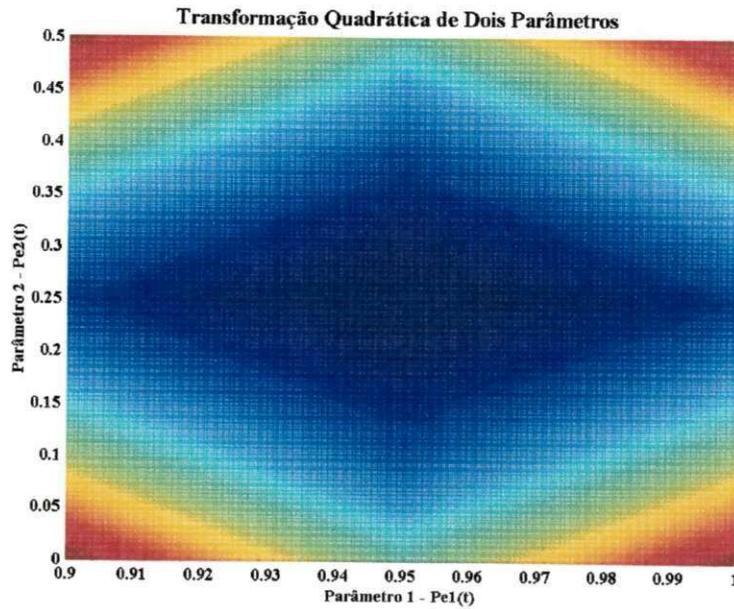


Figura 17 - $F_t(t)$ em função de $P_{e1}(t)$ e $P_{e2}(t)$ para a transformação quadrática(2-D) com b dado por Eq. (32)

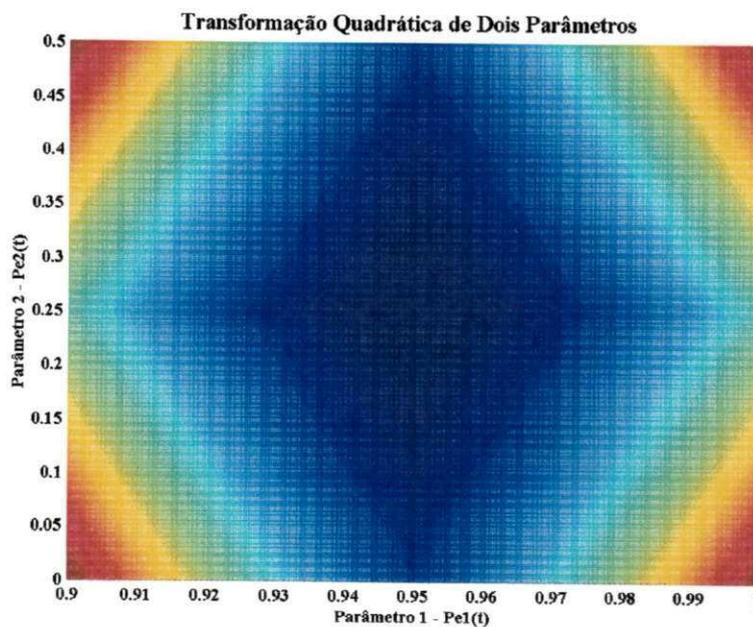


Figura 18 - $F_t(t)$ em função de $P_{e1}(t)$ e $P_{e2}(t)$ para a transformação quadrática(2-D) com b dado por Eq. (33)

Portanto, foi possível observar que a condição de equilíbrio empregada na equação da transformação é adequada, como também, que se pode adequar a influência de cada parâmetro a depender dos coeficientes a e b , havendo então a possibilidade de desenvolver-se sistemas inteligentes, com possibilidade de treinamento e aprendizado dos coeficientes mais adequados para a discriminação via estímulos táteis de características da fala.

4.3.2 Transformações Linear e Cúbica de Dois Parâmetros

As transformações linear e cúbica também podem ser utilizadas com dois parâmetros. As equações (37) e (38) definem estas transformações, já incorporando os coeficientes de ajuste. Da mesma forma que na transformação quadrática, para evitar que os erros positivos e negativos dos parâmetros P_{e1} e P_{e2} se anulem, calculou-se o módulo dos mesmos nas equações (37) e (38). $P_{e1}(t)$ e $P_{e2}(t)$ são parâmetros extraídos do sinal de voz, P_{a1} e P_{a2} são os respectivos parâmetros-alvo, F_c é a frequência do pulso vibrotátil obtida quando o usuário atinge os parâmetros alvo simultaneamente, e S é um fator de escala que define o grau de dificuldade do exercício.

$$F_t(t) = F_c + (a * |P_{e1}(t) - P_{a1}| + b * |P_{e2}(t) - P_{a2}|) * S \quad (37)$$

$$F_t(t) = F_c + (a * |P_{e1}(t) - P_{a1}| + b * |P_{e2}(t) - P_{a2}|)^3 * S \quad (38)$$

$F_t(t)$ deve variar entre f_{min} e f_{max} , $P_{e1}(t)$ deve variar entre os valores p_{11} e p_{21} , e $P_{e2}(t)$ deve variar entre os valores p_{12} e p_{22} , definidos pelo usuário do sistema. Assim, dado uma frequência F_c , os parâmetros alvo P_{a1} e P_{a2} , e os limites superiores das faixas dos parâmetros detectadas, p_{21} e p_{22} , pode-se obter a expressão de S_{max} , p_{11} e p_{12} para as transformações apresentadas.

4.3.2.1 Cálculo de S_{\max} , p_{11} e p_{12} para a Transformação Linear

Substituindo-se as igualdades (24), (25) e (26) em (37), obtém-se:

$$S_{\max} = \frac{f_{\max} - f_{\min}}{[a(p_{21} - P_{a1}) + b(p_{22} - p_{a2})]} = \frac{f_{\max} - f_{\min}}{[a(P_{a1} - p_{11}) + b(P_{a2} - p_{12})]} \quad (39)$$

De onde se obtém:

$$p_{11} = -p_{21} + 2P_{a1}, \quad (40)$$

$$p_{12} = -p_{22} + 2P_{a2}, \quad (41)$$

4.3.2.2 Cálculo de S_{\max} , p_{11} e p_{12} para a Transformação Cúbica

Substituindo-se as igualdades (24), (25) e (26) em (38), obtém-se:

$$S_{\max} = \frac{f_{\max} - f_{\min}}{[a(p_{21} - P_{a1}) + b(p_{22} - p_{a2})]^3} = \frac{f_{\max} - f_{\min}}{[a(P_{a1} - p_{11}) + b(P_{a2} - p_{12})]^3} \quad (42)$$

De onde se obtém:

$$p_{11} = -p_{21} + 2P_{a1}, \quad (43)$$

$$p_{12} = -p_{22} + 2P_{a2}, \quad (44)$$

4.3.2.3 Simulações

Com o objetivo de realizar uma avaliação do comportamento das transformações linear e cúbica, calculamos a dependência de $F_i(t)$ em função de parâmetros genéricos $P_{e1}(t)$ e $P_{e2}(t)$ a partir das equações (37) e (38). Foram escolhidos $f_{\min} = 10\text{Hz}$, e $f_{\max} = 100\text{Hz}$, que implicam $F_c = 10\text{Hz}$. As simulações foram realizadas no ambiente do MATLAB. Os valores de p_{11} e p_{12} utilizados para

obtenção dos resultados foram calculados a partir das Eqs. (43) e (44), respectivamente, e os valores de S_{max} foram calculados a partir das Equações (39) e (42).

Pode-se observar nas Figuras 19 e 20 a variação da transformação linear de $F_t(t)$ em função de $P_{e1}(t)$ e $P_{e2}(t)$ considerando-se os parâmetros limite $p_{11} = p_{12} = 1$ e os alvos $p_{a1} = p_{a2} = 0,5$. Na Figura 19 apresenta-se o diagrama em três dimensões, e na Figura 20 o diagrama em duas dimensões. Nas Figuras 21 e 22 são apresentados os mesmos resultados para a transformação cúbica.

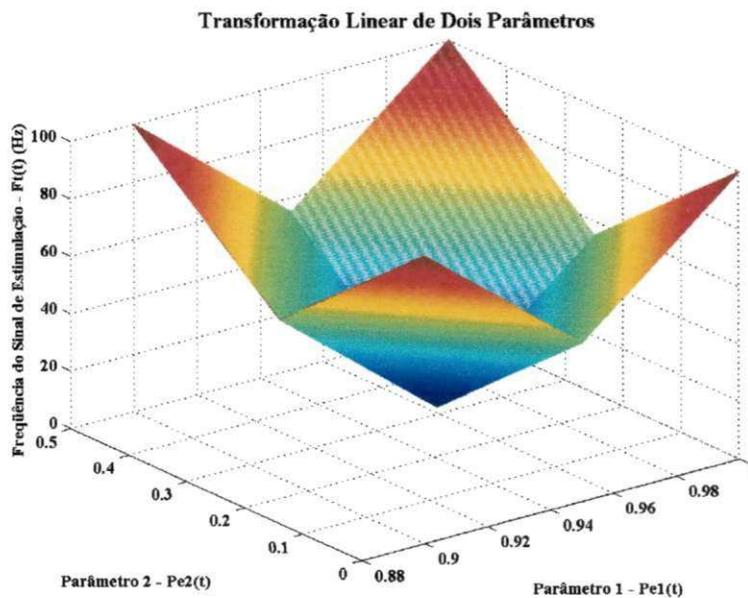


Figura 19 - Dependência de $F_t(t)$ em função de $P_{e1}(t)$ e $P_{e2}(t)$ para a transformação linear

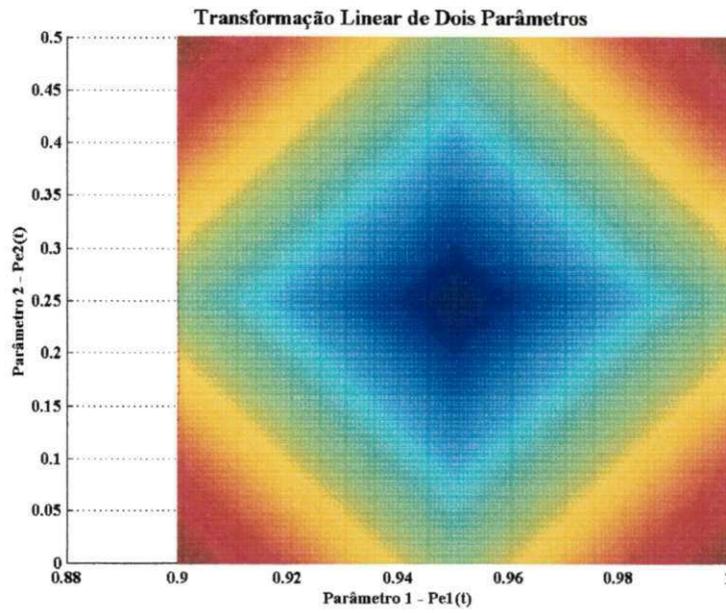


Figura 20 - $F_t(t)$ em função de $P_{e1}(t)$ e $P_{e2}(t)$ para a transformação linear(2-D) com b dado por Eq. (33)

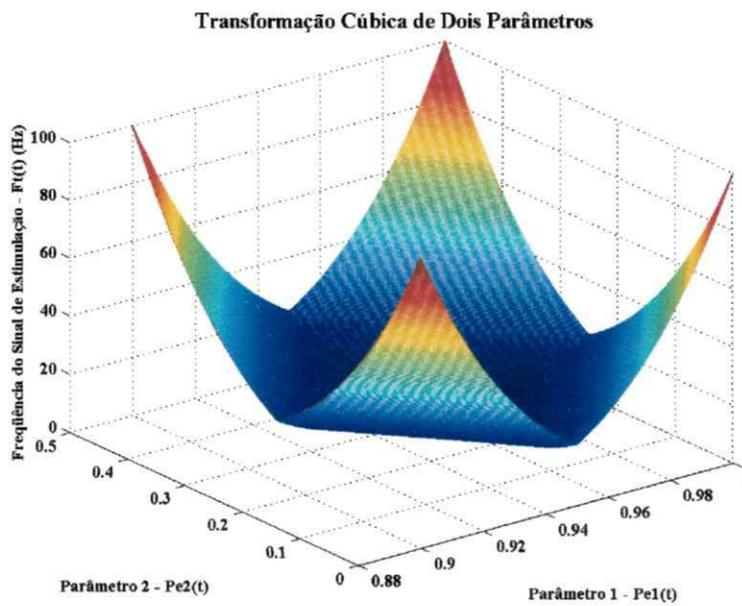


Figura 21 - Dependência de $F_t(t)$ em função de $P_{e1}(t)$ e $P_{e2}(t)$ para a transformação Cúbica

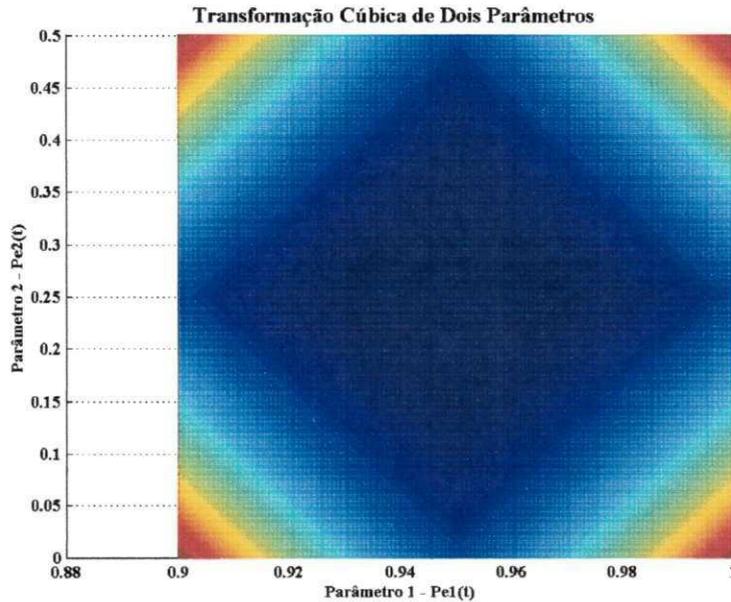


Figura 22 - $F_t(t)$ em função de $P_{e1}(t)$ e $P_{e2}(t)$ para a transformação Cúbica(2-D) com b dado por Eq. (33)

4.4 Transformações de Três Parâmetros

No caso de uma característica da fala ser mais bem definida por três parâmetros, apresenta-se a equação de transformação abaixo:

$$F_t(t) = F_c + (a * |P_{e1}(t) - P_{a1}| + b * |P_{e2}(t) - P_{a2}| + c * |P_{e3}(t) - P_{a3}|) * S \quad (45)$$

$$F_t(t) = F_c + (a * |P_{e1}(t) - P_{a1}| + b * |P_{e2}(t) - P_{a2}| + c * |P_{e3}(t) - P_{a3}|)^2 * S \quad (46)$$

$$F_t(t) = F_c + (a * |P_{e1}(t) - P_{a1}| + b * |P_{e2}(t) - P_{a2}| + c * |P_{e3}(t) - P_{a3}|)^3 * S \quad (47)$$

Assim, considerando-se

$$b_1 = \left(\frac{P_{21} - P_{11}}{P_{22} - P_{12}} \right) * a, \quad (48)$$

$$c1 = \left(\frac{p_{21} - p_{11}}{p_{23} - p_{13}} \right) * a \quad (49)$$

Pode-se verificar pela Tabela 4 as condições de influência de cada parâmetro nas transformações dadas pelas equações (45), (46) e (47).

Tabela 4 - Condições de influência de cada parâmetro nas transformações dadas pelas equações (45), (46) e (47).

Maior Influência	Condições		
Pe1	$b < b1$	$c < c1$	$b = \left(\frac{p_{23} - p_{13}}{p_{22} - p_{12}} \right) * c$
Pe2	$b > b1$	$c = c1$	-
Pe3	$b = b1$	$c > c1$	-
Pe1 e Pe2	$b = b1$	$c < c1$	-
Pe1 e Pe3	$b < b1$	$c = c1$	-
Pe2 e Pe3	$b > b1$	$c > c1$	$b = \left(\frac{p_{23} - p_{13}}{p_{22} - p_{12}} \right) * c$
Equilíbrio	$b = b1$	$c = c1$	

4.4.1 Cálculo de S_{\max}

Substituindo-se as igualdades (24), (25) e (26) em (45), (46) e (47), obtém-se:

$$S_{\max} = \frac{f_{\max} - f_{\min}}{[a(p_{21} - p_{a1}) + b(p_{22} - p_{a2}) + c(p_{23} - p_{a3})]^l} \quad (50)$$

$$S_{\max} = \frac{f_{\max} - f_{\min}}{[a(P_{a1} - p_{11}) + b(P_{a2} - p_{12}) + c(P_{a3} - p_{13})]^i} \quad (51)$$

com $i=1$ para a transformação linear, $i=2$ para a transformação quadrática e $i=3$ para a transformação cúbica.

Para $i = 1, i=2$ ou $i=3$, igualando-se as equações (52) e (53) obtém-se:

$$p_{11} = -p_{21} + 2P_{a1}, \quad (52)$$

$$p_{12} = -p_{22} + 2P_{a2}, \quad (53)$$

$$p_{13} = -p_{32} + 2P_{a3}, \quad (54)$$

4.5 Discussões

As transformações apresentadas podem ser utilizadas no desenvolvimento de sistemas de auxílio ao ensino da fala por estímulos vibrotáteis.

Dentre as transformações de um parâmetro, pôde-se observar pelos resultados de simulações que a transformação quadrática apresenta a melhor performance para indicação de acerto de alvo, pois a indicação do acerto não depende apenas do aprendizado da frequência de acerto, mas também da observação da inflexão da curva no momento do acerto.

As transformações de dois parâmetros atendem a uma necessidade de se codificar mais de um parâmetro em apenas um canal de estimulação. Com a utilização da função módulo com as transformações linear e cúbica, as mesmas passaram a apresentar a mesma característica de inflexão da função quadrática. Os resultados das simulações indicam que a função linear apresenta uma indicação de acerto de alvo mais exata que as demais.

5

CODIFICAÇÃO DE PARÂMETROS DE VOZ EM ESTÍMULOS VIBROTÁTEIS: EXPERIMENTOS DE PERCEPÇÃO

5.1 Introdução

Conforme foi citado nos capítulos anteriores, o aprendizado da fala está relacionado ao controle da respiração, da intensidade da voz, da entonação e o aprendizado da articulação de fonemas. A envoltória ou a energia do sinal de voz são parâmetros adequados para transmissão de informações sobre respiração e intensidade da fala. Para o caso da entonação, a frequência fundamental é o parâmetro mais adequado. Quanto à articulação de vogais e consoantes, mais de um parâmetro é necessário, tais como as duas primeiras formantes para identificação de vogais.

Atualmente, pela codificação direta do sinal de estimulação vibrotátil, é possível transmitir apenas um parâmetro por estimulador, no entanto, utilizando-se as transformações propostas no capítulo 4 deste trabalho é possível desenvolver sistemas vibrotáteis para auxílio ao ensino da fala usando apenas um estimulador, em uma configuração de indicação de acerto de alvo, podendo-se codificar o estimulador com informações de um ou mais parâmetros.

Assim, com o objetivo de comparar o desempenho das diversas transformações propostas, aplicadas à questão do aprendizado da fala, apresenta-se neste capítulo dois estudos-de-caso. O primeiro baseia-se na codificação da Frequência Fundamental, utilizando-se transformações de um parâmetro, e o segundo trata da identificação de vogais, com codificação de dois parâmetros (Amplitude Média e Taxa de Cruzamentos por Zero), utilizando-se as transformações de dois parâmetros.

Devido às vantagens da estimulação tátil com relação à transmissão de informações dinâmicas da fala, também estudou-se a parametrização do sinal de voz para transmissão direta via um estimulador vibrotátil monocanal, considerando-se o contexto do português. Um estimulador deste tipo pode ser usado paralelamente com estimuladores usados especificamente para o ensino da fala, como os baseados nas transformações propostas nesta tese.

5.2 Estimulador Vibrotátil

Os transdutores táteis são dispositivos que estimulam a pele e fornecem informações para indivíduos que possuem limitações, quer seja devido a uma deficiência sensorial ou por estar em uma situação na qual os outros sentidos estão sobrecarregados. Pode-se observar na Figura 23 alguns transdutores vibrotáteis típicos. Na Figura 23(a) observa-se um transdutor tipo bobina-móvel. Em 23(b) e 23(c) são apresentadas configurações de transdutores inerciais, similares aos utilizados nos estimuladores da Audiological Engineering. Em 23(d) pode-se observar a configuração de um display piezocerâmico.

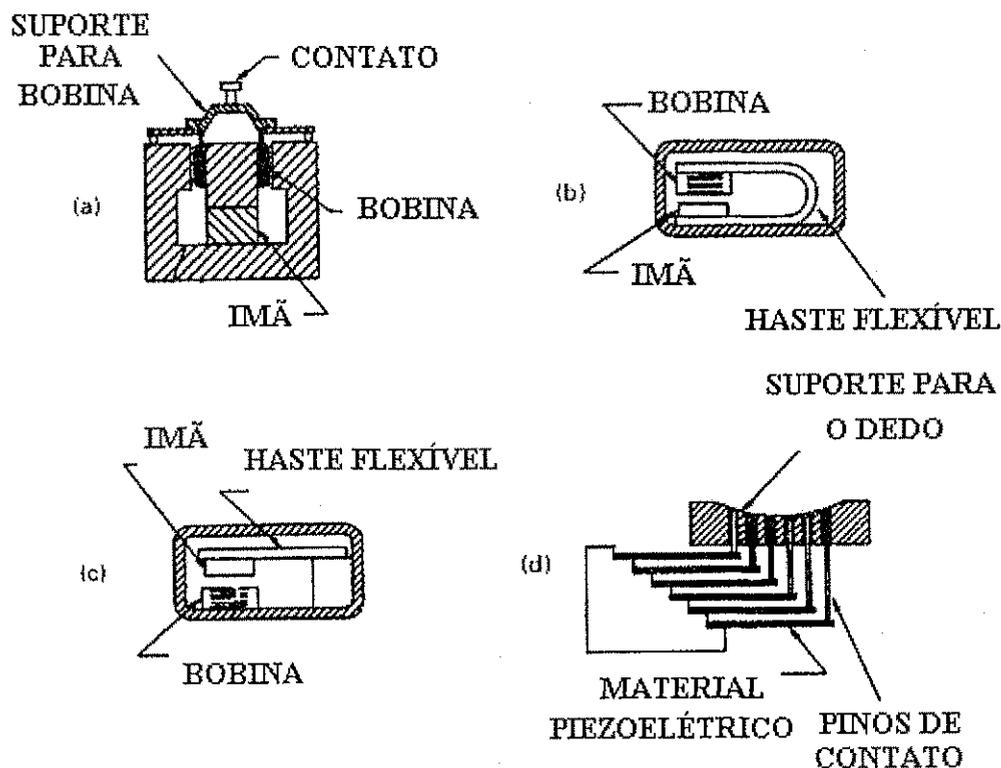


Figura 23 - Configurações de Estimuladores Vibrotáteis

A maioria dos dispositivos de ajuda vibrotátil portáteis utiliza transdutores inerciais. Neste tipo de transdutor, uma massa é suspensa por uma mola, ou haste flexível, fixada na base ou caixa que envolve todo o estimulador, conforme se pode observar nas Figuras 23(a) e 23(b). Uma força alternada é gerada entre a massa e a caixa externa, em geral através da passagem de corrente elétrica por uma bobina. Como a caixa que envolve o sistema não está fixada rigidamente, a mesma irá vibrar. Na Figura 24, pode-se observar um exemplo de transdutor inercial comercial.

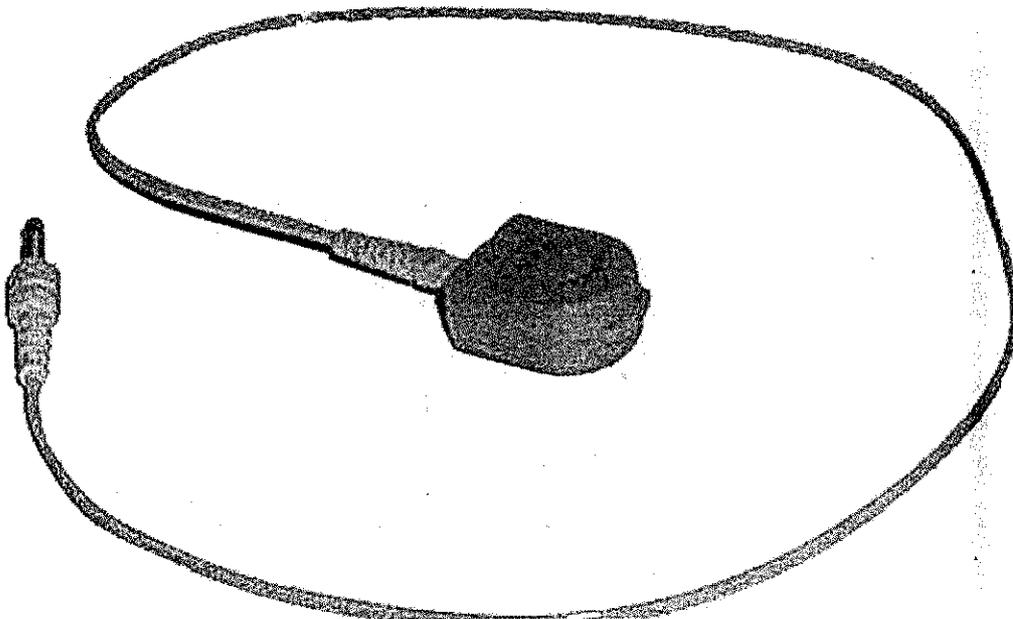


Figura 24 - Transdutor Inercial Comercial (VBW32, da Audiological Engineering)

Neste trabalho foi utilizado um transdutor comercial idêntico ao da Figura 24. Devido a sua baixa impedância o mesmo foi acionado pela saída de áudio de um microcomputador comum. Os estímulos foram aplicados na ponta do dedo indicador.

5.3 Parâmetros de Voz

O gráfico amplitude-x-tempo de um sinal permite a avaliação de muitas características importantes para uma completa descrição do mesmo. A partir do uso de parâmetros temporais torna-se possível identificar os sons básicos da fala. Dentre estes parâmetros destacam-se a amplitude média do sinal e a taxa de cruzamentos por zero.

O sinal de voz pode ser considerado invariante no tempo, para curtos intervalos, até 32 ms, sendo um valor típico 16 ms (Rabiner & Schafer, 1978). Assim, para se obter os parâmetros temporais do sinal de voz, é necessário particioná-lo em segmentos (ou blocos de amostras), visando trabalhar com o sinal dentro dos seus limites de estacionariedade (Fechine, 2000). Neste trabalho os parâmetros de voz foram extraídos a cada 5 ms, utilizando-se uma janela retangular de 25 ms (Rabiner & Schafer, 1978).

5.3.1 Amplitude Média

O parâmetro Amplitude Média (AVG) é definido por Rabiner & Schafer (1989).

$$AVG = \sum_{n=0}^{N_A-1} |x(n)| \quad (55)$$

em que $x(n)$ é o sinal de voz e N_A o tamanho da janela (bloco de amostras do sinal) em análise. O parâmetro AVG é obtido, portanto, simplesmente somando-se as amplitudes das N_A amostras do sinal contido na janela em análise.

5.3.2 Taxa de Cruzamentos por Zero

A taxa de cruzamentos por zero, TCZ, é outro parâmetro bastante utilizado em aplicações de processamento digital de sinais de voz, que utilizam métodos de análise no domínio do tempo. Esse parâmetro indica o número de vezes que as amostras de um sinal, em um determinado segmento, cruzam o zero (limiar) tomado como referência.

Este parâmetro é, geralmente, definido por:

$$TCZ = \sum_{n=0}^{N-1} |sinal[x(n)] - sinal[x(n-1)]| \quad (56)$$

$$sinal[x(n)] = \begin{cases} 1 & \text{se } x(n) \geq 0 \\ -1 & \text{se } x(n) < 0 \end{cases} \quad (57)$$

Isto significa que a contagem é efetuada sempre que sucessivas amostras do sinal tiverem polaridades contrárias.

Essa medida pode ser interpretada como uma forma simples de se determinar o conteúdo de frequência de um sinal

5.4 Parametrização do Sinal de Voz e Codificação do Sinal de Estimulação Vibrotátil

A estimulação vibrotátil monocanal é utilizada no auxílio à leitura labial e no aprendizado da fala (Proctor, 1995). Alguns sons da fala têm movimentos

articulatórios quase idênticos e outros têm posições articulatórias invisíveis. Qualquer grupo de fonemas que tem a mesma articulação visual é chamado "viseme", e há muita dificuldade para o surdo, tanto de aprender a oralizar tais sons, dada a falta de realimentação visual, como de percebê-los através da leitura labial (Öster, 1995).

Foi feito um estudo neste trabalho sobre a escolha e codificação de parâmetros de voz, para uso em estimuladores vibrotáteis monocanais aplicados ao auxílio na discriminação entre fonemas que compõem visemes da língua portuguesa, os quais encontram-se exclusivamente entre as consoantes, conforme é mostrado na Tabela 5.

Nos estimuladores monocanais, em função da limitada capacidade de transmissão de informações pelo tato, apenas um parâmetro de voz modula o sinal de estimulação (Summers, 1992). Os parâmetros de voz "Envoltória do Sinal", "Frequência Fundamental" e "Taxa de Cruzamentos por Zero" apresentam os melhores resultados na diferenciação entre consoantes, sendo que os parâmetros "Envoltória do Sinal" e "Taxa de Cruzamentos por Zero" apresentam boas respostas para sinais ambientais, e não apenas para a voz (Summers & Gratton, 1995).

Os parâmetros "Amplitude Média" e "Taxa de Cruzamentos por Zero" foram estudados quanto à capacidade de discriminação entre os fonemas da Tabela 5. Na seqüência, foi proposto e investigado um esquema misto utilizando os dois parâmetros, permitindo uma maior capacidade de discriminação quando comparado à aplicação de cada parâmetro em separado.

Na Tabela 6 pode-se observar que o problema do auxílio à leitura labial do português consiste também na ajuda a discriminação entre as características de modo e excitação das consoantes discriminadas.

Tabela 5 - Grupos de fonemas da língua portuguesa cuja distinção via leitura labial apresenta dificuldades

Grupos de Visemes para o Português			
1	s, z	5	g, k, lh
2	ch, j	6	l, r
3	t, d, n	7	f, v
4	p, b, m		

Tabela 6 - Grupos de fonemas da Tabela 4 classificados quanto ao modo e vozeamento.

Consoantes Oclusivas Orais:	
Surdas [p] [t] [k]	Sonoras [b] [d] [g]
Consoantes Oclusivas Nasais:	
	Sonoras [m] [n] [ɲ]
Consoantes Fricativas:	
Surdas [f] [s] [ʃ]	Sonoras [v] [z] [ʒ]
Consoante lateral:	
	[l]
Consoante Vibrante:	
	[r]

5.4.1 Metodologia

Summers & Gratton (1995) analisaram vários parâmetros de voz, através de experimentos de percepção, com o objetivo de determinar o parâmetro de voz mais adequado para modular um estimulador vibrotátil monocanal, tendo em vista o auxílio à compreensão da fala, concluindo que os parâmetros envoltória do sinal e taxa de cruzamentos por zero apresentam o mesmo desempenho.

Considerando que os estimuladores vibrotáteis monocanal são utilizados no auxílio à leitura labial, procurou-se identificar qual dos parâmetros, envoltória do sinal, ou TCZ, são mais adequados para discriminação entre os visemes do

português. Neste estudo, foram realizadas coletas de sons do tipo [VcV]², ou seja, consoante entre vogais, para cada consoante da Tabela 5, de 3 indivíduos, sendo um deles do sexo feminino.

A gravação dos dados foi feita diretamente em um microcomputador, em um ambiente tipo escritório. Os sinais foram adquiridos com 16 bits por amostra, e com frequência de amostragem de 44.100 Hz, assegurando uma qualidade de áudio em padrão CD.

A inserção da articulação consonantal entre vogais se justifica pela necessidade de isolar consoantes oclusivas surdas, por exemplo, que seriam difíceis de serem separadas no caso de estar iniciando um segmento [cV]. Além disto, foram escolhidas vogais distintas com relação ao posicionamento dos primeiros e segundos formantes, abrangendo a maioria das situações possíveis de ocorrências dos segmentos [VcV] nas palavras da língua portuguesa. O vocabulário escolhido encontra-se na Tabela 7.

Tabela 7 - Conjunto de Palavras Utilizadas

Conjunto de palavras utilizadas no experimento	
[i]c[u]	[i] em livro e [u] em pular
[u]c[a]	[u] em pular e [a] em pato
[a]c[i]	[a] em pato e [i] em livro
[o]c[a]	[o] em poço e [a] em pato
[ɔ]c[ɛ]	[ɔ] em gola e [ɛ] em terra

Foram extraídos os parâmetros amplitude média, que representa a envoltória do sinal, e taxa de cruzamentos por zero. Posteriormente, analisadas as forma de ondas dos sinais obtidos para cada parâmetro e extraídos a parte do sinal referente à articulação consonantal. Foi construído, para cada sinal um histograma representando a distribuição das amplitudes do sinal para cada

² [VcV] – Vogal-consoante-vogal, por exemplo [itu], [idu], [inu]

consoante. Os resultados foram comparados entre articulações do mesmo subgrupo mostrado na Tabela 5.

5.4.1.1 Histogramas

Os histogramas obtidos para os parâmetros amplitude média e taxa de cruzamentos por zero são mostrados a seguir, para um indivíduo. Os resultados obtidos foram similares entre todos os indivíduos:

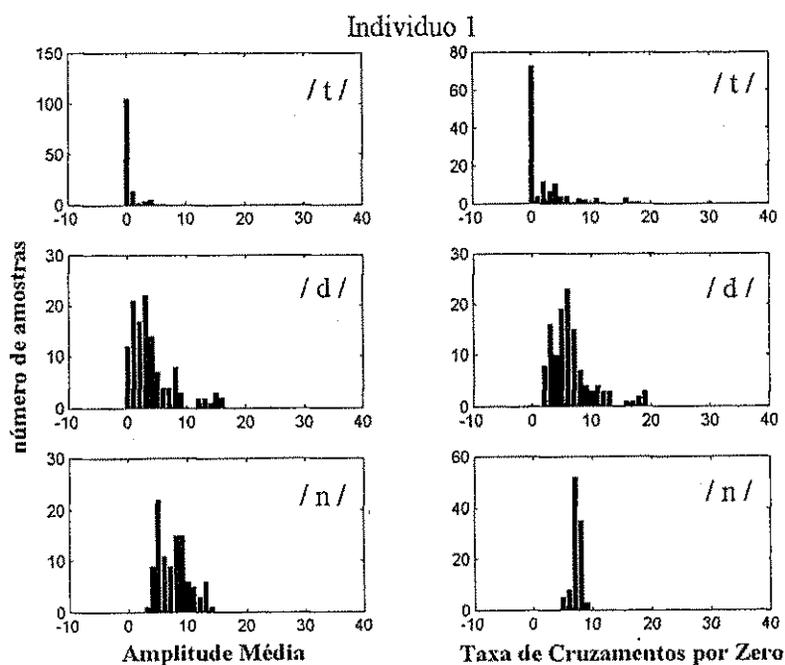


Figura 25 - Histogramas da distribuição dos parâmetros AVG e TCZ (/t, d, n/)

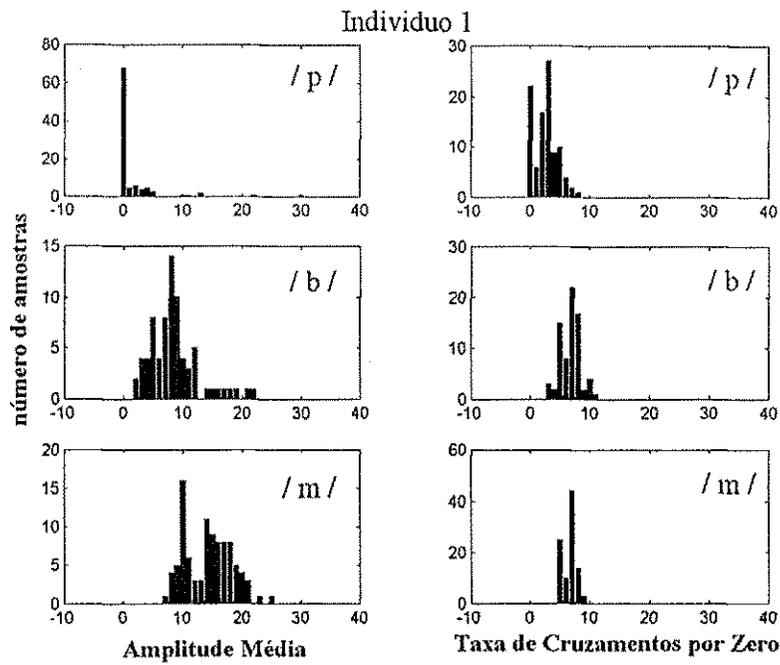


Figura 26 - Histogramas da distribuição dos parâmetros AVG e TCZ (/p, b, m/)

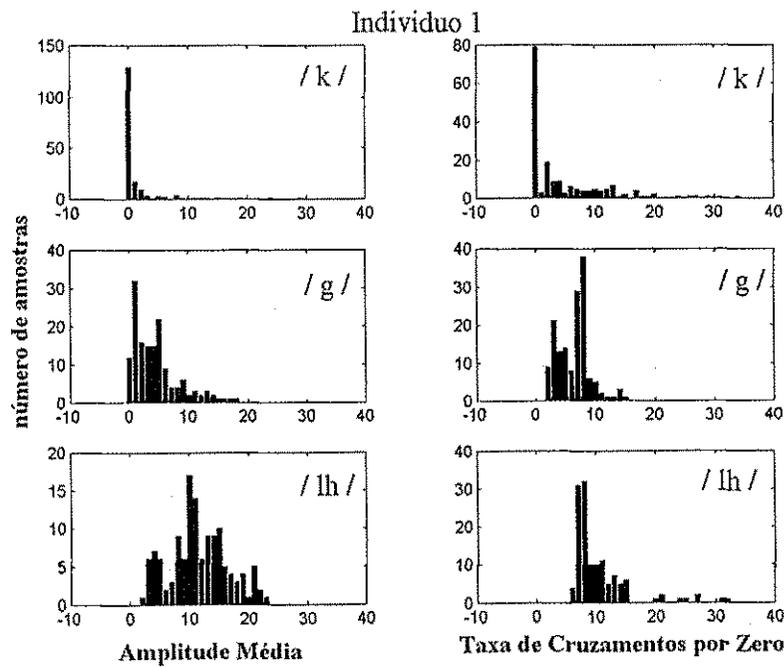


Figura 27 - Histogramas da distribuição dos parâmetros AVG e TCZ (/k, g, lh/)

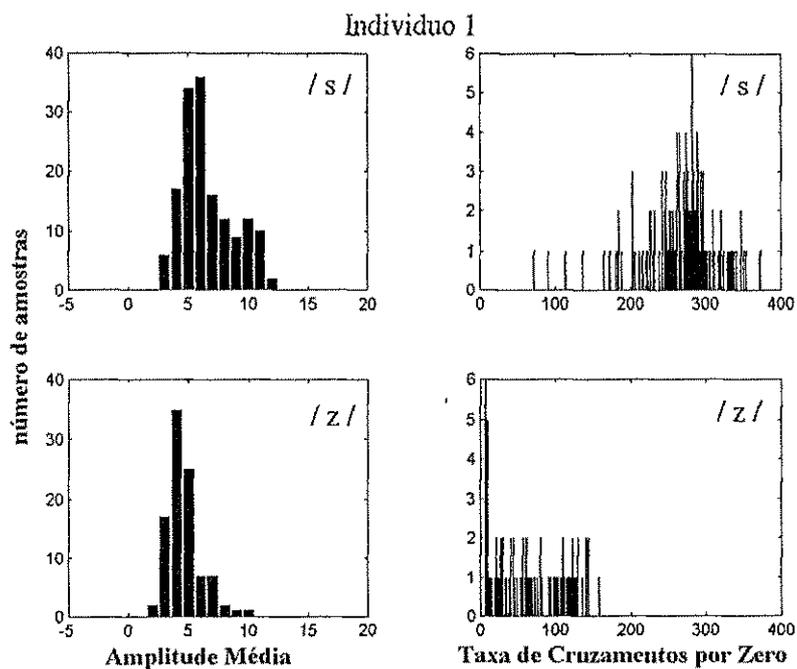


Figura 28 - Histogramas da distribuição dos parâmetros AVG e TCZ (/s, z/)

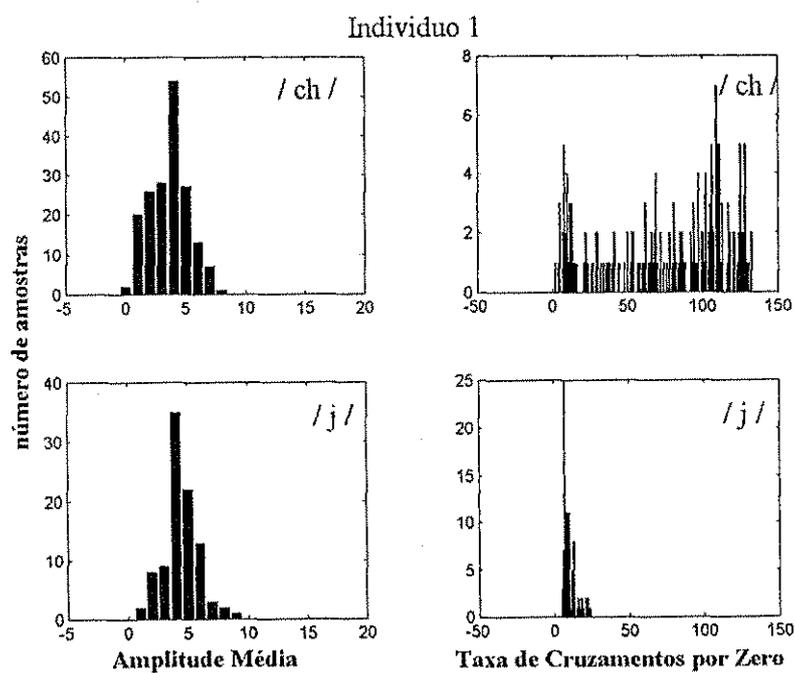


Figura 29 - Histogramas da distribuição dos parâmetros AVG e TCZ (/ch, j/)

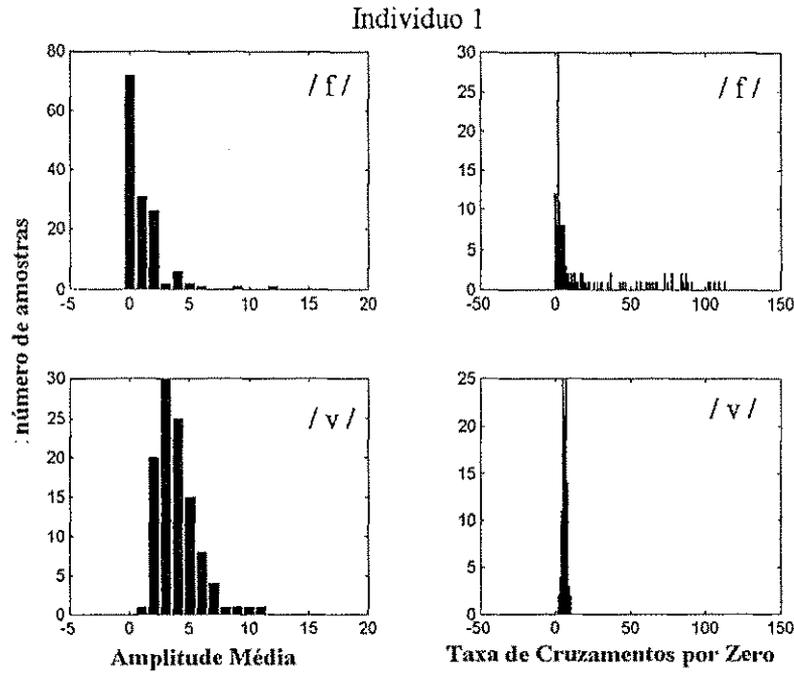


Figura 30 - Histogramas da distribuição dos parâmetros AVG e TCZ (/ f, v /)

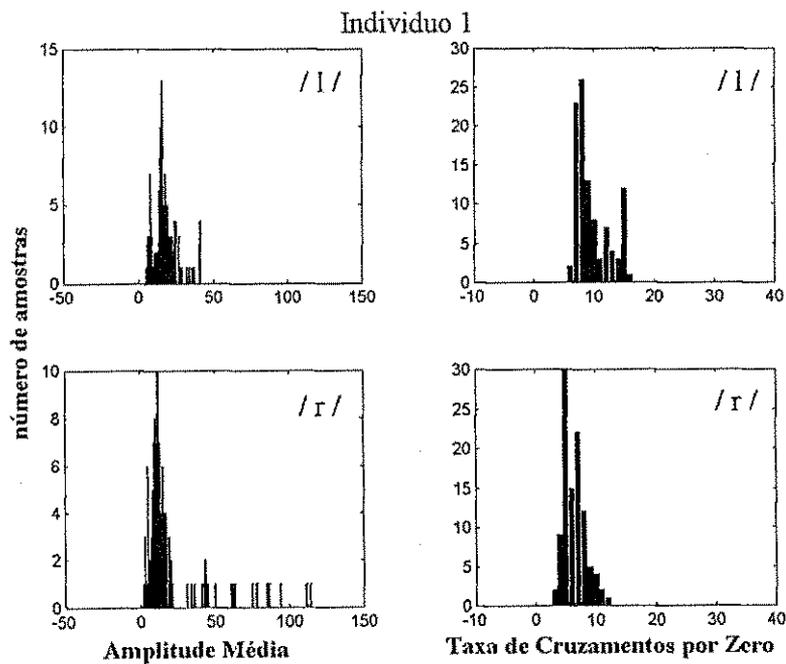


Figura 31 - Histogramas da distribuição dos parâmetros AVG e TCZ (/ l, r /)

A partir da observação dos resultados para a distribuição dos valores dos parâmetros AVG e TCZ para cada articulação consonantal estudada, apresentados nas Figuras (25 – 31) foi possível concluir que:

- 1) Nenhum dos dois parâmetros oferece bons resultados de discriminação para todos os grupos, para uma análise feita a partir da observação das faixas de valores nos histogramas.
- 2) A partir da análise da taxa de cruzamentos por zero é possível discriminar os fonemas dos grupos [s] e [z]; [ch] e [j]; [t] e [d]; [p] e [b]; [g] e [k]. Com maior incerteza, é possível tomar decisões quanto aos parâmetros [l] e [r].
- 3) A partir da amplitude média é possível discriminar [t], [d] e [n], [p], [b] e [m], [g], [k] e [ɲ]. Com maior incerteza, é possível tomar decisões quanto aos parâmetros [f] e [v].

Como os parâmetros possibilitam a discriminação de forma complementar, através de uma arquitetura que utilize os dois parâmetros poderá ser possível aumentar a capacidade de discriminação do sistema. Uma arquitetura deste tipo foi proposta na seção seguinte.

5.4.2 Codificação Mista dos Parâmetros AVG e TCZ

A partir das considerações da seção anterior foi proposta uma forma de codificação do sinal de estimulação vibrotátil (PRP), a qual é mostrada em diagrama de blocos na Figura 32. Esta configuração deve aumentar a capacidade de discriminação entre consoantes fricativas pertencentes aos visemes do português.

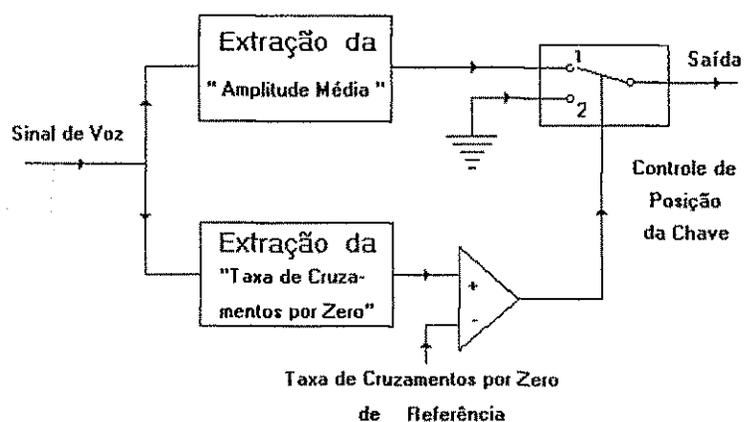


Figura 32 - Diagrama de Blocos do Sistema de Codificação de Sinal de Estimulação Vibrotátil Proposto.

O sinal de voz aplicado alimenta um módulo de extração da Amplitude Média e, paralelamente, um módulo para extração da taxa de cruzamentos por zero. A saída do extrator da taxa de cruzamentos por zero é aplicada na entrada de um comparador, sendo comparada com uma Taxa de referência pré-fixada (TRX). Quando a saída do extrator for maior que a taxa de cruzamentos por zero de referência (TRX), a chave muda da sua posição natural (posição 1), para a posição 2, de forma que o sinal de saída será aterrado neste instante. Para cada indivíduo foi levantado um histograma com o sinal de saída obtido para o sinal de estimulação, para as taxas de referência para a abertura da chave de 60, 80, 100, 120, 140 e 160 cruzamentos por zero.

Para o indivíduo 1, com taxa de abertura da chave em 60 cruzamentos por zero, teve-se uma melhora da discriminação em amplitude dos fonemas [ch] e [j], e uma alteração tendendo a zerar a amplitude tanto do fonema [s] como [z], embora tenha havido uma melhora na diferenciação entre os histogramas de amplitude destes fonemas. Com abertura de 80 cruzamentos por zero, melhorou a discriminação entre os histogramas dos fonemas [s] e [z] e [ch] e [j]. Ao aumentar este valor, há uma melhor discriminação entre [s] e [z], e diminui a capacidade de discriminação entre os fonemas [ch] e [j].

Para o indivíduo 2, com taxa de abertura da chave em 60 cruzamentos por zero, há uma sensível melhora da discriminação em amplitude dos fonemas [ch] e

[j], e [s] e [z]. Apenas quando a taxa passa de 140, diminui a capacidade de discriminação entre os fonemas [ch] e [j].

Para o indivíduo 3, com taxa de abertura da chave em 60 cruzamentos por zero, há uma sensível melhora da discriminação em amplitude dos fonemas [s] e [z]. Não há melhoras entre [ch] e [j], pois para este indivíduo, os valores dos parâmetros TCZ ocuparam a mesma faixa de valores.

Os resultados obtidos para PRP indicaram que se pode implementar um estimulador vibrotátil monocanal com boa discriminação entre fricativos. Nas Figuras 33 e 34 vê-se os histogramas gerais (dados dos três indivíduos) de saída para o parâmetro AVG e para PRP (TRX=100), respectivamente, através das quais é possível verificar que o esquema PRP permite melhor discriminação entre os grupos fricativos (/ch/ e /j/; /s/ e /z/) do que o parâmetro AVG.

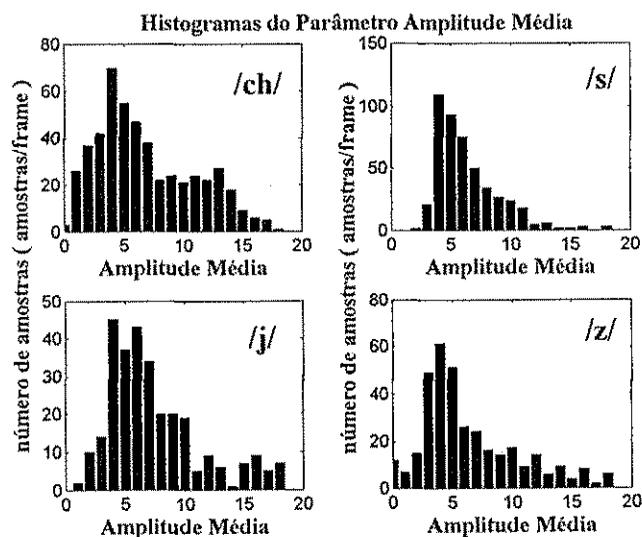


Figura 33 - - Histograma Geral para os fonemas formados pelas articulações /s/, /z/ e /ch/, /j/, do parâmetro AVG

Histogramas dos Sinais na Saída do Esquema Misto proposto - TRX = 100

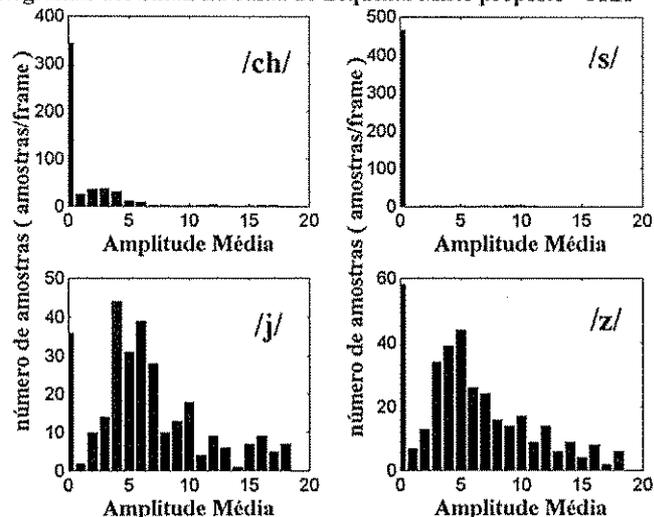


Figura 34 - - Histograma Geral para os fonemas formados pelas articulações /s/, /z/ e /ch/, /j/, da Arquitetura proposta, para um chaveamento com TRX = 100

5.4.3 Experimentos de Percepção

A comparação de resultados obtidos em testes com estimuladores vibrotáteis para transmissão de informações acústicas, efetuados em laboratórios diferentes é difícil, pois existem variáveis dos experimentos que são difíceis de controlar e de se repetir, como o nível de experiência do voluntário em percepção de estimulação vibrotátil, a forma de treinamento, a língua utilizada, e a integração nos testes de percepção de diversos fatores, como o parâmetro de voz utilizado, modulação do sinal de estimulação vibrotátil, tipo de estimulador, local do corpo, etc...

Com o objetivo de fazer uma análise por etapas, optou-se por realizar uma comparação da escolha de parâmetros de voz para uso com estimuladores vibrotáteis monocanal baseada na observação dos histogramas de distribuição de valores, conforme exposto nas seções anteriores. Posteriormente foram realizados experimentos de percepção com o objetivo de efetuar uma comparação com os resultados obtidos através da análise de histogramas.

5.4.3.1 Metodologia

Para cada conjunto de fonemas que compõem um viseme foi realizada uma experiência de percepção, com o objetivo de verificar a capacidade de discriminação entre os fonemas, considerando-se o parâmetro AVG. Apenas para os fricativos (/s/ , /z/ e /ch/, /j/) foi considerada a influência do parâmetro PRP.

Para os visemes formados por três fonemas, o indivíduo sob teste foi estimulado por um sinal composto por três palavras codificadas, no formato /VcV/, em uma seqüência aleatória. Para os visemes com dois fonemas, o estímulo foi composto por duas palavras codificadas no formato /VcV/. Pode-se observar na Tabela 8 o conjunto de consoantes para cada viseme, com ordem identificada por c1, c2 e c3. Todas as combinações para visemes com três e dois visemes são apresentadas nas tabelas 9 e 10, respectivamente. Para cada viseme foram geradas 100 estimulações de acordo com as opções das tabelas 9 ou 10, com a mesma probabilidade de ocorrer qualquer das opções. Um indivíduo normal recebeu os estímulos e por meio de uma interface gráfica, na qual com um clique no mouse foi possível identificar o estímulo percebido para cada estimulação, foi armazenado o valor do estímulo percebido, em conjunto com a informação sobre a opção de estímulo real.

Tabela 8 - Conjunto de Visemes

c1	c2	c3
t	d	n
p	b	m
k	g	lh
f	v	-
l	r	-
ch	j	-
s	z	-

Tabela 9 - Conjunto de Opções de Estímulos dos Visemes com Três Fonemas

Opção	Palavras		
1	/Vc1V/	/Vc2V/	/Vc3V/
2	/Vc1V/	/Vc3V/	/Vc2V/
3	/Vc2V/	/Vc1V/	/Vc3V/
4	/Vc2V/	/Vc3V/	/Vc1V/
5	/Vc3V/	/Vc1V/	/Vc2V/
6	/Vc3V/	/Vc2V/	/Vc1V/

Tabela 10 - Conjunto de Opções de Estímulos dos Visemes com Dois Fonemas

Opção	Palavras	
1	/Vc1V/	/Vc2V/
2	/Vc2V/	/Vc1V/

5.4.3.2 Codificação do Sinal de Estimulação Vibrotátil

A frequência do sinal de estimulação vibrotátil, $f(n)$, foi obtida de acordo com as equações abaixo:

$$f(n) = k * AVG(n) + B \quad (58)$$

$$f(n) = k * PRP(n) + B \quad (59)$$

nas quais n define o segmento de voz e k e B são parâmetros que determinam o comportamento linear da codificação.

Neste trabalho, considerou-se $B=30$ nas equações (58) e (59), pois assim garante-se um valor mínimo de frequência, $f(n)$, igual a 30 Hz, que está na faixa inicial de sensação tátil perceptível. Para definição do parâmetro k foram realizados alguns testes de percepção de discriminação dos fonemas / t /, / d / e

/n/, com $k=8$, $k=4$ e $k=16$. Os resultados do percentual de acerto para cada fonema indicaram melhores resultados para $k=8$.

5.4.3.3 Resultados Obtidos

Os resultados obtidos para os parâmetros amplitude média para o restante dos grupos da Tabela 8 são mostrados a seguir, nas Figuras 35, 36 e 37. Na Figura 38 são apresentados os resultados obtidos com a percepção do parâmetro proposto PRP, em comparação com os resultados obtidos para AVG.

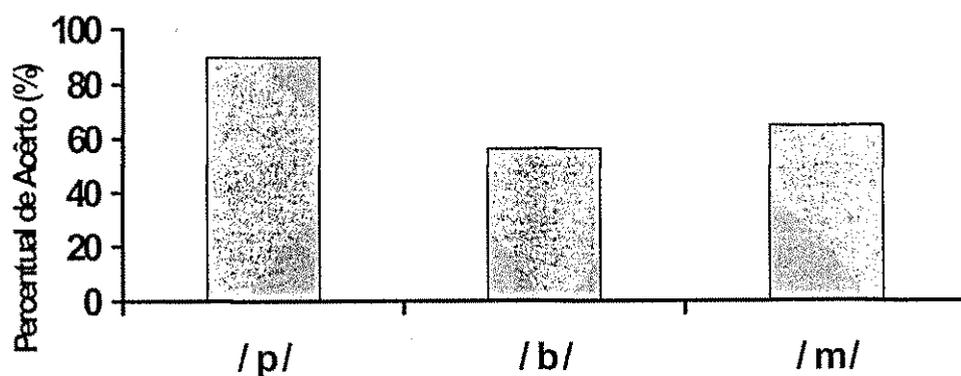


Figura 35 - Resultados de Testes de Percepção de Discriminação entre Fonemas /p/, /b/ e /m/

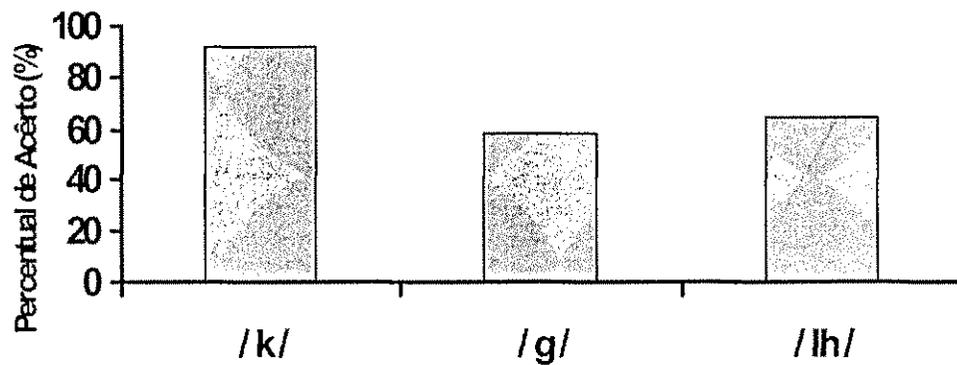


Figura 36 - Resultados de Testes de Percepção de Discriminação entre Fonemas /k/, /g/ e /lh/

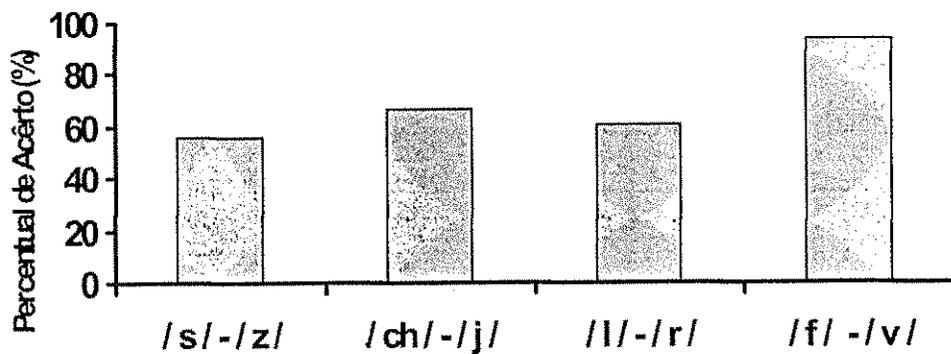


Figura 37 - Resultados de Testes de Percepção de Discriminação entre Fonemas /s/-/z/, /ch/-/j/, /l/-/r/ e /f/-/v/

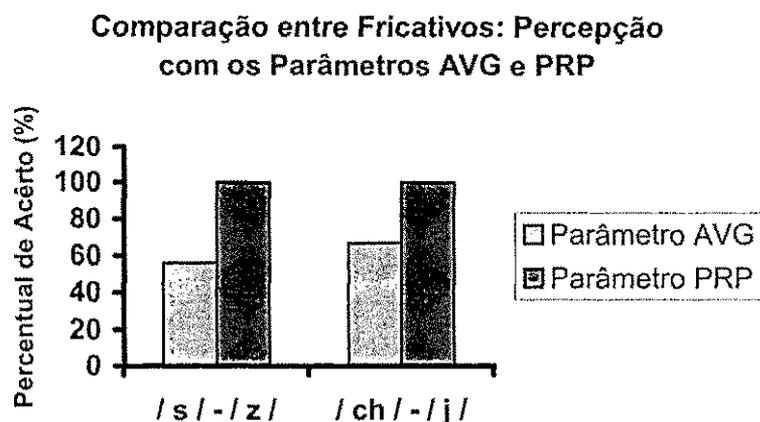


Figura 38 - Resultados de Testes de Percepção de Discriminação entre Fonemas /s/-/z/ e /ch/-/j/, comparando-se os parâmetros AVG e PRP

5.4.4 Discussões

Observou-se que os resultados obtidos com os testes de percepção da discriminação entre fonemas que compõem os visemes do português, utilizando-se o parâmetro AVG para modular o sinal de estimulação vibrotátil comprovaram as conclusões obtidas a partir da análise de histogramas, excetuando-se o fato que houve uma ótima discriminação entre os fonemas /f/ e /v/.

Para os visemes formados por fonemas oclusivos, houve uma melhor discriminação entre /d/ e /n/, do que entre /b/ e /m/, e /k/ e /lh/. Conclui-se que isto se deve ao fato da equação de codificação da frequência do sinal de estimulação ter sido ajustada para os fonemas /t/, /d/ e /n/.

Na Figura 38 pode-se observar que a utilização mista dos parâmetros AVG e TCZ de acordo com o esquema proposto resulta em uma melhor discriminação entre os fricativos, no caso de segmentos /VcV/.

5.5 Transformações de Um Parâmetro

As transformações de um parâmetro foram desenvolvidas no capítulo 4, podendo ser linear, cúbica logarítmica ou quadrática. Cada transformação pode ser configurada para trabalhar em níveis distintos de dificuldade. Em termos de aplicação no auxílio ao aprendizado da fala, pode ser utilizada com o parâmetro frequência fundamental, em sistemas para auxílio ao controle da entonação.

5.5.1 Frequência Fundamental

Os deficientes auditivos possuem uma entonação inadequada que afeta a qualidade e a inteligibilidade da fala, pois em geral as suas frequências fundamentais são mais elevadas do que as normais, tendo também uma excessiva variabilidade (Araújo, 2000; Monsen, 1979; Bush, 1981; Boone, 1994; Araújo et al, 1999). Em um sistema específico para o ensino da fala, dois fatores são importantes (Mahshie, 1995): a realimentação, que deve variar com o nível do aprendizado do deficiente auditivo; e a possibilidade do terapeuta configurar parâmetros do sistema, adequando-o às necessidades do usuário. Como já foi mencionado, geralmente os sistemas de realimentação visual contém módulos para treinamento da frequência fundamental, com indicativo de acerto de alvo. Também há a opção de ajuste da frequência fundamental de acerto pelo terapeuta.

O estudo-de-caso investigado nesta seção contempla a possibilidade de variação da quantidade de realimentação, em graus de dificuldade dependentes do aprendizado do usuário e a possibilidade de configuração da frequência fundamental alvo.

5.5.2 Experimentos de Percepção

Os experimentos de percepção foram realizados com o objetivo de avaliar comparativamente o desempenho entre as quatro transformações, em termos de

facilidade de acerto de alvo, e comparar, em cada transformação, o desempenho entre os vários níveis de dificuldade.

5.5.2.1 Metodologia

O processo de aprendizado do controle da frequência fundamental de voz utilizando as transformações propostas se dá da seguinte forma: o indivíduo que utiliza o sistema fala com determinada frequência fundamental que é transformada em determinado estímulo vibrotátil. A partir da percepção do estímulo, o indivíduo aumenta ou diminui o tom de voz, em função da realimentação vibrotátil que recebe, com o objetivo de acertar o alvo.

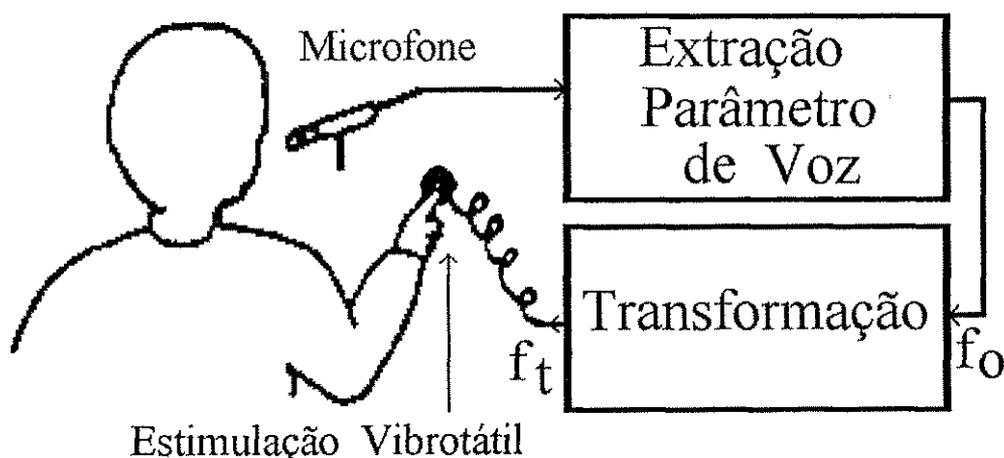


Figura 39- Sistema Típico para treinamento/Aprendizagem da fala

Pode-se observar na Figura 39 que o indivíduo sob teste gera o sinal de voz e percebe o sinal de estimulação vibrotátil, atuando tanto como gerador como sensor. Esta característica dificulta a avaliação sobre a escolha da forma de mapeamento de f_0 em f_t , pois a performance do sistema pode ser influenciada pela habilidade do indivíduo no controle e geração da frequência fundamental.

Com o objetivo de isolar a habilidade do indivíduo estimulado no controle da frequência fundamental no resultado da avaliação das transformações, o

processo de acerto de alvo com base no controle da freqüência fundamental foi simulado, de acordo com o fluxograma do procedimento experimental, mostrado na Figura 40. Inicialmente é gerado um valor aleatório de freqüência fundamental e um valor de incremento Δf_0 , e o indivíduo é estimulado no dedo por um sinal vibrotátil $s(t)$, gerado através da escolha aleatória de uma das transformações disponíveis. Em função da percepção do estímulo vibrotátil, o indivíduo estimulado tem a opção de aumentar a freqüência do estímulo ($k = 1$), diminuir a freqüência do estímulo ($k = -1$) ou definir que o alvo foi acertado (acerto). Na última opção, o sistema grava a quantidade de passos gasta para o acerto do alvo e a freqüência de estimulação percebida pelo voluntário como freqüência de acerto de alvo. Estes dois parâmetros são a figura de mérito para avaliação comparativa do desempenho das transformações.

Assim, para cada transformação, foram testados 8 graus de dificuldade ($s_1 - s_8$), sendo s_1 o grau de menor dificuldade, correspondente a S_{max} , e s_8 o grau de maior dificuldade, correspondente a $(S_{max}/8)$. Além disto, foram feitos testes de acerto de alvo a partir de realimentação auditiva, para o grau de menor dificuldade.

Os testes foram feitos com um voluntário e gerados 100 acertos de alvo para cada transformação, em cada grau de dificuldade. Os estímulos foram aplicados em sessões de 25 acertos de alvo, para cada transformação.

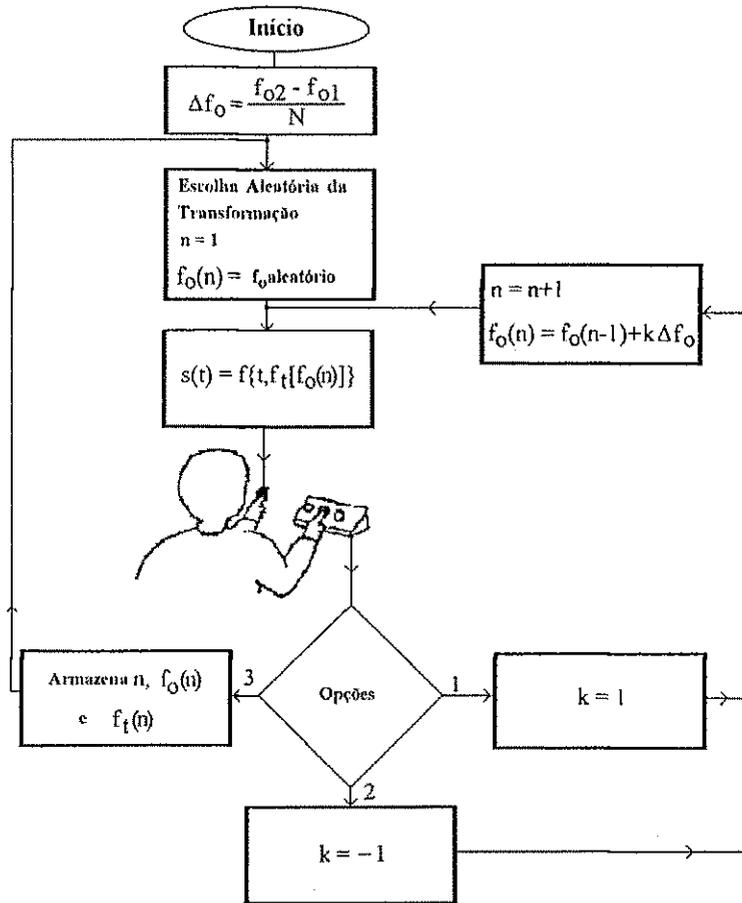


Figura 40 – Fluxograma do Procedimento Experimental

5.5.2.2 Sinal de Estimulação Vibrotátil

O sinal de estimulação vibrotátil foi gerado a partir das equações (60 - 63), para as transformações linear, logarítmica, cúbica e quadrática, respectivamente.

$$Ft(t) = Fc + (Fv(t) - Fa) * S \quad (60)$$

$$Ft(t) = Fc + \log\left(\frac{Fv(t)}{Fa}\right) * S \quad (61)$$

$$Ft(t) = Fc + (Fv(t) - Fa)^3 * S \quad (62)$$

$$F_t(t) = F_c + (F_v(t) - F_a)^2 * S \quad (63)$$

Nas quais, $F_v(t)$ é a frequência fundamental detectada do sinal de voz, F_a é a frequência fundamental alvo, F_c é a frequência do pulso vibrotátil obtida quando o usuário atinge a frequência fundamental alvo, e S é um fator de escala que define o grau de dificuldade do exercício.

$F_t(t)$ deve variar entre f_{min} e f_{max} , e $F_v(t)$ deve variar entre os valores f_1 e f_2 . Para as simulações realizadas neste trabalho escolheu-se: $f_2 = 400$ Hz e $F_a = 200$ Hz, representando uma situação típica.

Eberhardt et al (1990), afirmaram que a faixa ideal para a frequência do sinal de estimulação vibrotátil está entre de 10Hz e 100Hz, e que os sistemas de auxílio a leitura labial por realimentação vibrotátil, nos quais a frequência fundamental de voz média é transladada para a frequência de estimulação vibrotátil de 50Hz, tendem a ter mais sucesso na transmissão da entonação. Baseado nestas considerações foram escolhidos $f_{min} = 10$ Hz, e $f_{max} = 100$ Hz, que implicam $F_c = 10$ Hz para a transformação quadrática e $F_c = 55$ Hz para as demais transformações.

Os sinais de estimulação foram gerados no ambiente do MATLAB, e alimentaram o estimulador vibrotátil através da saída de áudio de um microcomputador.

5.5.2.3 Resultados Obtidos

Os resultados obtidos podem ser observados nas Figuras 41 a 54. Nas Figuras 41 e 42 pode-se observar os resultados para a frequência do sinal de estimulação auditiva percebida como acerto de alvo e a correspondente frequência fundamental. Nos gráficos 43 a 54 podem ser observados para cada transformação os resultados para a frequência do final de estimulação vibrotátil percebida como acerto de alvo, a correspondente frequência fundamental e o número de passos excedentes.

Estimulação Auditiva (S = Smax)

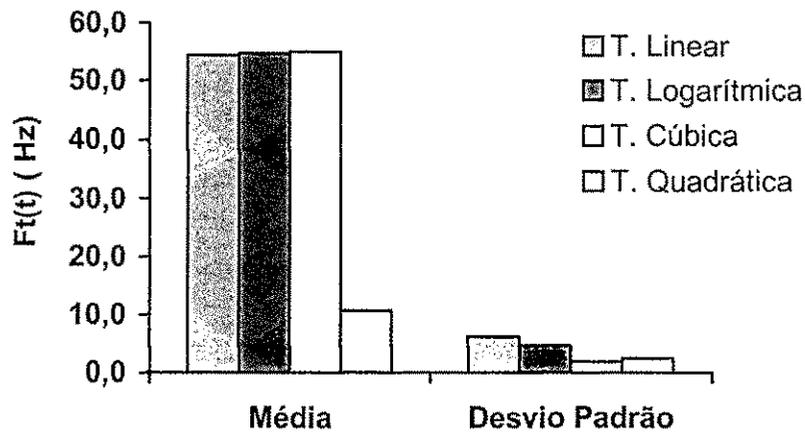


Figura 41 - Média e Desvio Padrão da Freqüência do Sinal de Estimulação Auditiva percebida como acerto de alvo para as Transformações Propostas

Estimulação Auditiva (S = Smax)

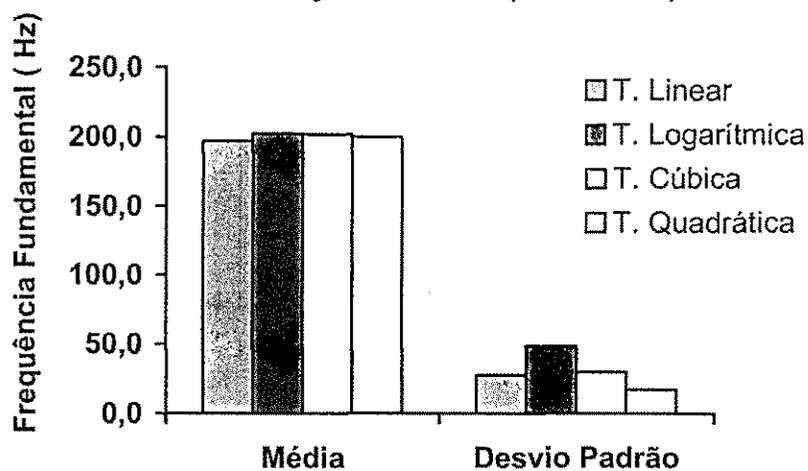


Figura 42 - Média e Desvio Padrão da Freqüência Fundamental Correspondente à Freqüência do Sinal de Estimulação Auditiva percebida como acerto de alvo para as Transformações Propostas

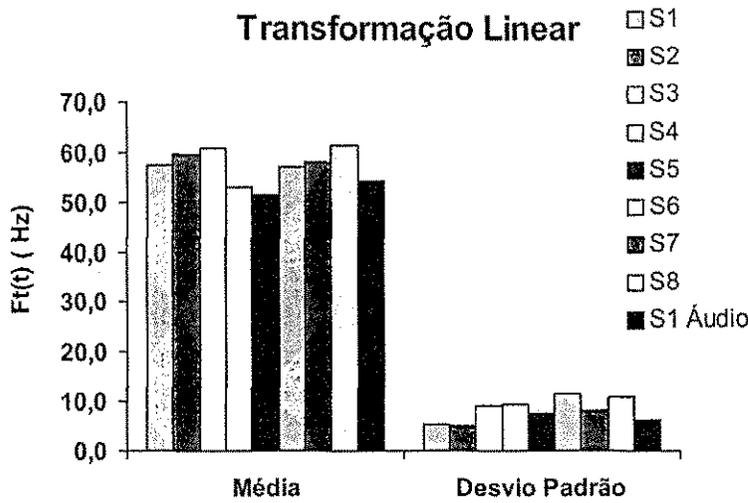


Figura 43 - Média e Desvio Padrão da Freqüência do Sinal de Estimulação Vibrotátil percebida como acerto de alvo para 8 dificuldades da Transformação Linear, além de resultado obtido para Estimulação Auditiva

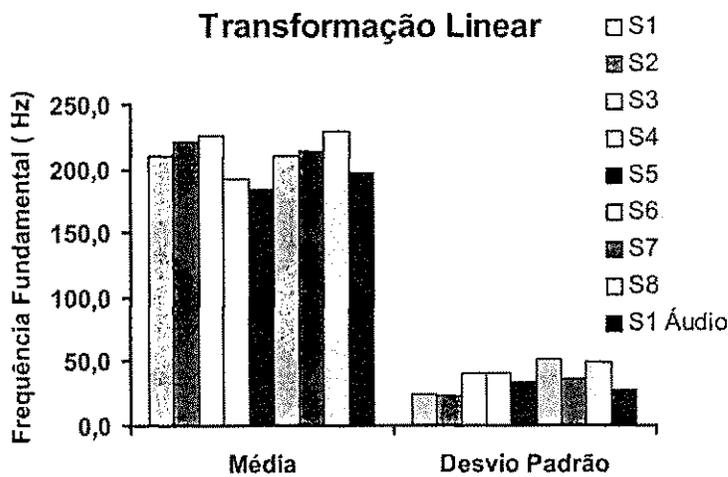


Figura 44 - Média e Desvio Padrão da Freqüência Fundamental Correspondente à Freqüência do Sinal de Estimulação Vibrotátil percebida como acerto de alvo para 8 dificuldades da Transformação Linear, além de resultado obtido para Estimulação Auditiva

Transformação Linear

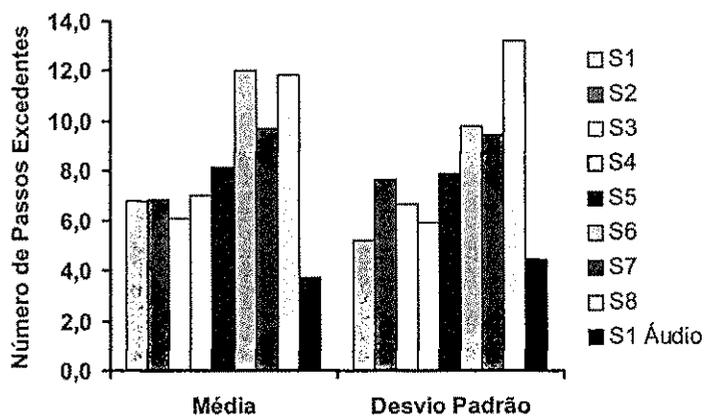


Figura 45 - Média e Desvio Padrão do Número de Passos Excedentes Realizados para Indicação de Acerto de Alvo para 8 dificuldades da Transformação Linear, além de resultado obtido para Estimulação Auditiva

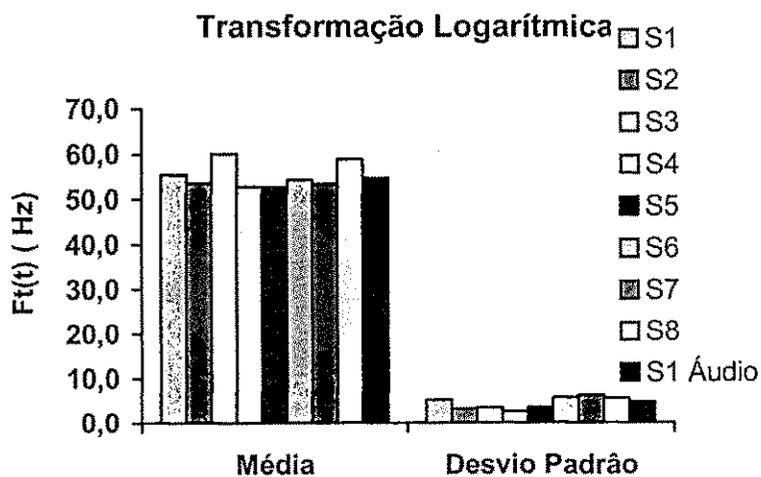


Figura 46 - Média e Desvio Padrão da Frequência do Sinal de Estimulação Vibrotátil percebida como acerto de alvo para 8 dificuldades da Transformação Logarítmica, além de resultado obtido para Estimulação Auditiva

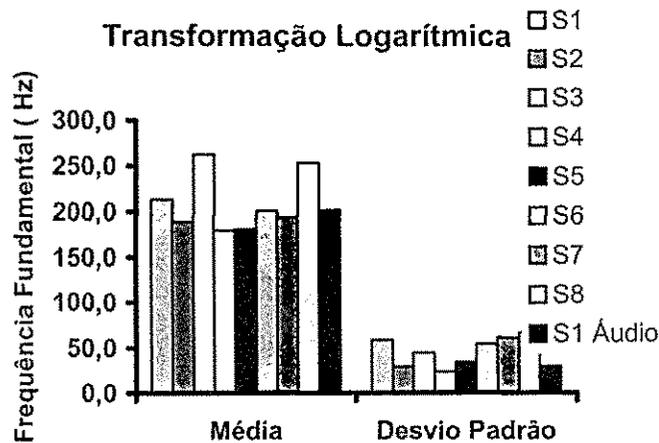


Figura 47 - Média e Desvio Padrão da Frequência Fundamental Correspondente à Frequência do Sinal de Estimulação Vibrotátil percebida como acerto de alvo para 8 dificuldades da Transformação Logarítmica, além de resultado obtido para Estimulação Auditiva

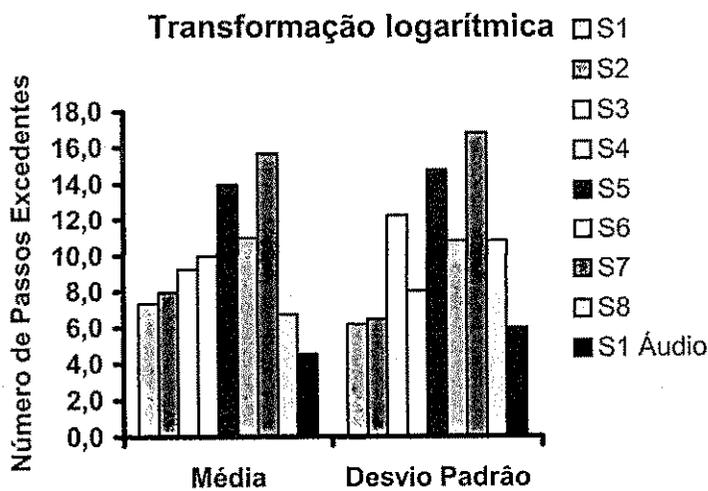


Figura 48 - Média e Desvio Padrão do Número de Passos Excedentes Realizados para Indicação de Acerto de Alvo para 8 dificuldades da Transformação Logarítmica, além de resultado obtido para Estimulação Auditiva

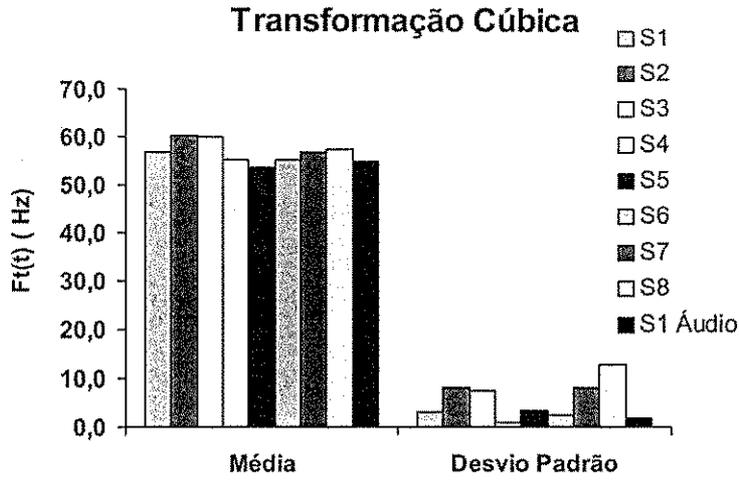


Figura 49 - Média e Desvio Padrão da Frequência do Sinal de Estimulação Vibrotátil percebida como acerto de alvo para 8 dificuldades da Transformação Cúbica, além de resultado obtido para Estimulação Auditiva

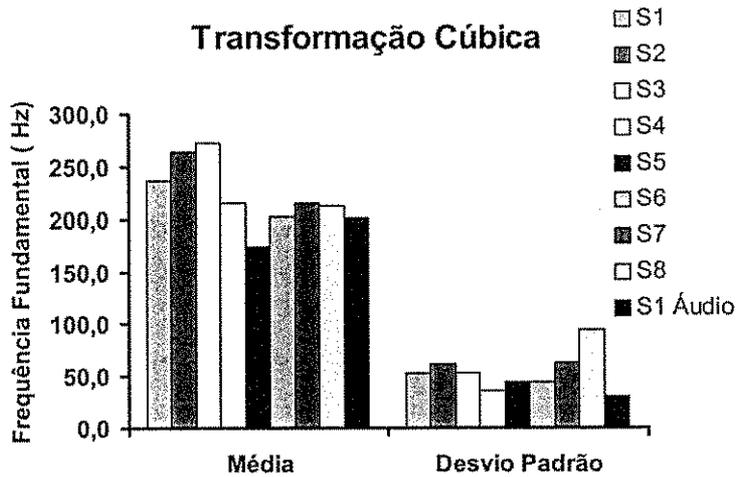


Figura 50 - Média e Desvio Padrão da Frequência Fundamental Correspondente à Frequência do Sinal de Estimulação Vibrotátil percebida como acerto de alvo para 8 dificuldades da Transformação Cúbica, além de resultado obtido para Estimulação Auditiva

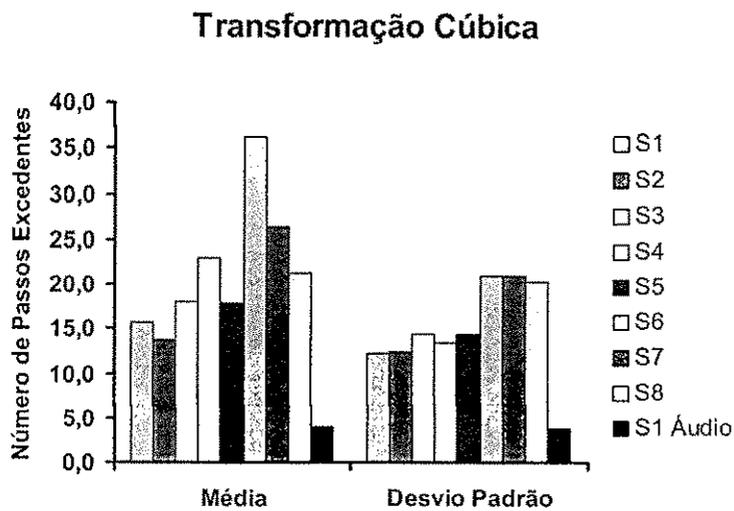


Figura 51 - Média e Desvio Padrão do Número de Passos Excedentes Realizados para Indicação de Acerto de Alvo para 8 dificuldades da Transformação Cúbica, além de resultado obtido para Estimulação Auditiva

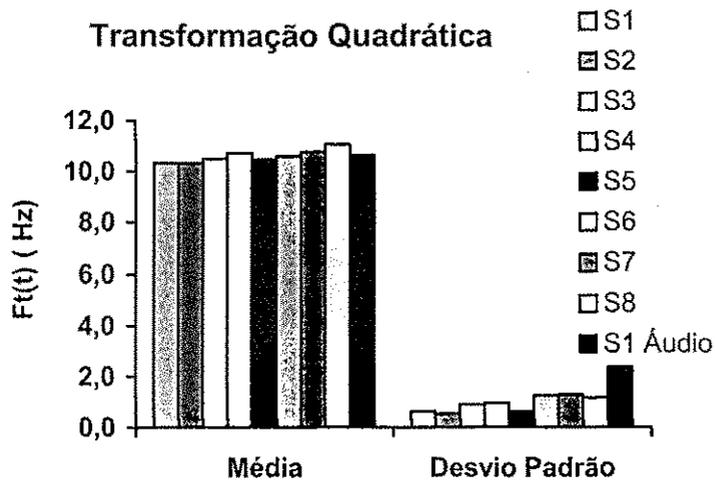


Figura 52 - Média e Desvio Padrão da Frequência do Sinal de Estimulação Vibrotátil percebida como acerto de alvo para 8 dificuldades da Transformação Quadrática, além de resultado obtido para Estimulação Auditiva

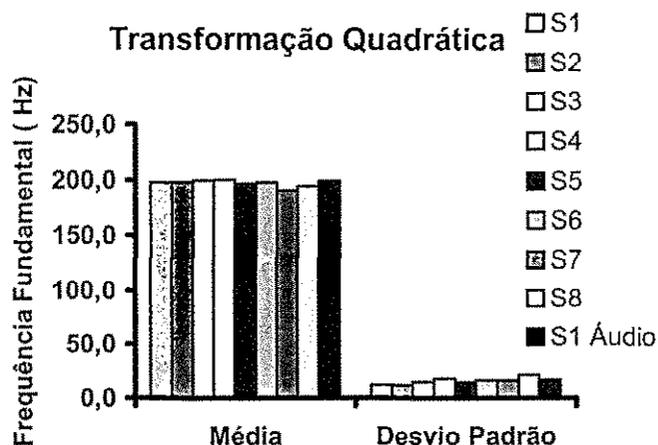


Figura 53 - Média e Desvio Padrão da Frequência Fundamental Correspondente à Frequência do Sinal de Estimulação Vibrotátil percebida como acerto de alvo para 8 dificuldades da Transformação Quadrática, além de resultado obtido para Estimulação Auditiva

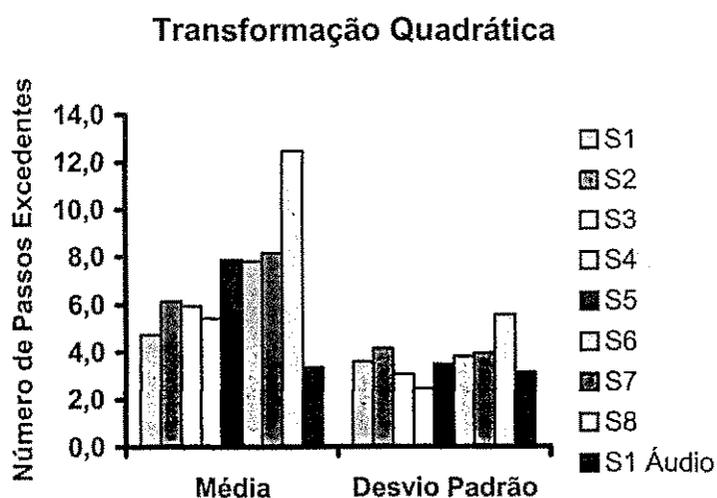


Figura 54 - Média e Desvio Padrão do Número de Passos Excedentes Realizados para Indicação de Acerto de Alvo para 8 dificuldades da Transformação Quadrática, além de resultado obtido para Estimulação Auditiva

5.5.3 Discussões

Os testes de percepção para acerto de alvo realizados com estimulação auditiva demonstram uma maior capacidade de identificação da frequência alvo através das transformações quadrática e cúbica (Figura 41), comprovando-se que as transformações linear e logarítmica são mais dependentes do aprendizado da identificação da frequência alvo por parte do indivíduo estimulado. No caso da transformação quadrática, a inflexão da curva na frequência alvo indica o acerto de alvo, e na transformação cúbica o achatamento da curva próxima a frequência alvo indica a frequência de acerto de alvo, pela sensação de mesma estimulação, mesmo com variações no parâmetro de entrada.

Na Figura 42, por sua vez, observa-se que a transformação cúbica apresenta o mesmo padrão de variação (desvio padrão) em torno da frequência alvo da transformação linear, enquanto a transformação logarítmica apresenta a maior variação e a transformação quadrática apresenta a menor variação, mantendo o melhor desempenho entre as quatro transformações de um parâmetro. Quanto ao número de passos excedentes para acerto de alvo, observa-se nas Figuras 45, 48, 51 e 54 que, na estimulação auditiva, acerta-se o alvo com menos passos nas transformações quadrática e cúbica.

Quanto à percepção vibrotátil da frequência do sinal de estimulação, observa-se nas Figuras 43, 46, 49 e 52 que no nível de menor dificuldade o desempenho das transformações é semelhante àquele obtido com estimulação auditiva. Observa-se que com o aumento do grau de dificuldade, o desvio padrão aumenta, indicando maior dificuldade no acerto do alvo. A transformação quadrática apresentou os resultados mais uniformes em todos os graus de dificuldade, em torno da frequência alvo, indicando maior facilidade no acerto do alvo. Quanto à observação da frequência fundamental correspondente à frequência de estimulação vibrotátil detectada, verifica-se nas Figuras 46 e 47 que os erros no acerto da frequência fundamental são bem maiores que os erros da percepção da frequência do sinal de estimulação vibrotátil, exatamente pelo

achatamento da curva em torno do parâmetro alvo, conforme se observou nas simulações do capítulo 4.

Quanto ao número de passos excedentes, pode-se observar nas Figuras 45 e 48 que nas transformações linear e logarítmica eles aumentam bastante com o grau de dificuldade do exercício, indicando o aumento das dúvidas e conseqüentemente da dificuldade no acerto do alvo. Na Figura 51, observa-se que até nos graus de menor dificuldade este número é bem maior que aquele obtido nas transformações linear e logarítmica, indicando que pelo achatamento da curva em torno do parâmetro alvo, há dificuldade e incerteza na identificação do acerto de alvo, embora haja facilidade de se identificar à faixa em torno do parâmetro alvo, conforme se concluiu nas simulações do capítulo 4. Já na transformação quadrática, o número de passos excedentes é pequeno em todos os graus de dificuldade, embora haja um pequeno aumento com o aumento dos graus de dificuldade.

Assim, concluiu-se que os testes de percepção baseados na simulação do controle da freqüência fundamental permitiram a realização de uma avaliação comparativa, entre as quatro transformações propostas, resultando na confirmação dos resultados esperados com base na análise das simulações efetuada no capítulo 4 deste trabalho. Observou-se que a transformação quadrática apresenta um desempenho bem melhor que as demais no acerto do alvo, mas que com o treinamento é possível melhorar o acerto do alvo com as demais. O domínio do controle da freqüência fundamental é um fator que pode auxiliar no acerto do alvo das transformações linear, cúbica e logarítmica, remetendo à possibilidade de implementação de jogos com diversas características e graus de dificuldade. Por fim observou-se maior confusão no uso da transformação logarítmica, pela não simetria de sua curva em torno do alvo.

5.6 Transformações de Dois Parâmetros

As transformações de dois parâmetros foram desenvolvidas no capítulo 4, podendo ser linear, cúbica ou quadrática. Cada transformação pode ser configurada para trabalhar em níveis distintos de dificuldade. Em termos de aplicação no auxílio ao aprendizado da fala, pode ser utilizada para o aprendizado da articulação de vogais.

5.6.1 Vogais

Segundo Russo & Behlau (1993), as vogais são definidas basicamente pela amplificação inserida na energia glótica. As faixas de frequência amplificadas, ou seja, os picos de energia variam de acordo com a vogal emitida e, no caso específico das vogais, representam grupos de harmônicos e recebem o nome de formante do som. Teoricamente tem-se um número infinito de formantes, mas os três primeiros são os mais importantes, suficientes para oferecer a identidade da vogal.

As vogais orais do português brasileiro podem ser classificadas com relação ao deslocamento da língua no plano horizontal (anterior, central, posterior) e pelo deslocamento da língua no plano vertical (alta, média, baixa). O primeiro formante está relacionado ao deslocamento da língua no plano vertical, e o segundo formante está relacionado ao deslocamento da língua no plano horizontal. São vogais do português brasileiro:

/a/, vogal central, baixa, conforme pronunciada em vaca;

/e/, vogal anterior, média, conforme pronunciada em pé;

/e/, vogal anterior, média, conforme pronunciada em peso;

/i/, vogal anterior, alta, conforme pronunciada em tia;

/ɔ/, vogal posterior, média, conforme pronunciada em pote;

/o/, vogal posterior, média, conforme pronunciada em poço;

/u/, vogal posterior, alta, conforme pronunciada em uva.

As vogais orais do português brasileiro foram estudadas por Russo & Behlau (1993), e de acordo com seus resultados, os valores dos formantes para as vogais orais de homens, mulheres e crianças, da cidade de São Paulo, encontram-se na Tabela 11. Os valores para os homens são os mais graves, por apresentarem os maiores tratos vocais, enquanto que para as crianças, os formantes das vogais são mais agudos, pela reduzida dimensão da caixa de ressonância.

Tabela 11 - Médias das frequências das três primeiras formantes: F1, F2 e F3, das vogais orais do português brasileiro, para homens, mulheres e crianças (Dados extraídos de Russo & Behlau, 1993)

		Vogais Analisadas						
Grupos	Formantes	/i/	/e/	/ɛ/	/a/	/ɔ/	/o/	/u/
Homens	F1 (Hz)	398	563	699	807	715	558	400
	F2 (Hz)	2456	2339	2045	1440	1201	1122	1182
	F3 (Hz)	3320	2995	2848	2524	2481	2520	2452
Mulheres	F1(Hz)	425	628	769	956	803	595	462
	F2(Hz)	2984	2712	2480	1634	1317	1250	1290
	F3(Hz)	3668	3349	3153	2721	2602	2668	2528
Crianças	F1(Hz)	465	698	902	1086	913	682	505
	F2(Hz)	3176	2825	2606	1721	1371	1295	1350
	F3(Hz)	3980	3637	3243	2873	2793	2823	2667
Média	F1(Hz)	429	630	790	950	810	612	456
	F2(Hz)	2872	2625	2377	1598	1296	1222	1274
	F3(Hz)	3656	3327	3081	2706	2625	2670	2549

De acordo com a freqüência dos formantes da Tabela 11, Russo & Behlau (1993) observaram:

1. Com relação às vogais anteriores, há um incremento em F1 na direção da vogal central /a/. As freqüências de F2 e F3 decrescem em direção à vogal central /a/.
2. Com relação às vogais posteriores, há um decremento em F1, partindo-se da vogal /a/ em direção às vogais posteriores. Não há um incremento significativo em F2 e F3, que tem valores praticamente iguais.

Analisando as informações da Tabela 11, Araújo (2000) observou que:

1. A utilização de F2 como parâmetro para classificação das vogais pode produzir confusões inerentes aos valores dos próprios parâmetros, dada a proximidade de valores, ou possíveis sobreposições de valores entre as vogais. A primeira formante, por outro lado, está intimamente relacionada à altura da língua, mas é insuficiente para propiciar a separação entre as sete vogais orais da língua portuguesa.
2. Os valores elevados de F2 estão associados a ressonâncias em freqüências altas, sendo característicos das vogais anteriores, enquanto baixos valores de F2 constituem uma característica das vogais posteriores. A presença de ressonâncias em freqüências baixas, nas vogais posteriores, faz com que a energia do sinal se concentre em baixas freqüências, enquanto que para as vogais anteriores, o espectro do sinal se mostra distribuído. Uma medida da distribuição de energia no espectro pode permitir avaliar-se se a vogal é posterior ou anterior.

5.6.2 Parâmetros Propostos

A partir das conclusões de Russo & Behlau (1993) e Araújo (2000), propõe-se dois parâmetros que podem fornecer pistas a um deficiente auditivo, via estimulação vibrotátil, sobre as vogais orais, funcionando como realimentação básica em um módulo de ensino da fala de vogais.

5.6.2.1 Parâmetro 1

O primeiro parâmetro proposto está representado no diagrama de blocos da Figura 55. Este parâmetro deve permitir a discriminação entre vogais anteriores, medial e posteriores, e baseia-se na análise da distribuição de energia no espectro da fala. O parâmetro Amplitude Média, que representa a envoltória do sinal de voz será tanto mais intenso quanto maior for a energia do sinal. Assim, inicialmente é extraído o parâmetro AVG do sinal de voz de entrada (AVG1). Paralelamente, o sinal de voz é aplicado em um filtro passa-altas e na seqüência é extraído o parâmetro AVG do sinal filtrado (AVG2). O filtro passa-altas tem uma faixa de passagem a partir da freqüência de 2.100 Hz. Observando-se a Tabela 7, pode-se verificar que as vogais posteriores possuem segunda formante abaixo da freqüência de 2.100 Hz, e as vogais anteriores, apresentam segunda formante acima desta freqüência. Dando continuidade à análise do diagrama de blocos da Figura 55, os sinais AVG1 e AVG2 são processados de forma a obter-se o valor percentual de AVG2 sobre AVG1. Desta forma, o parâmetro de saída fica independente das variações de intensidade do sinal. Espera-se então, que este valor percentual seja maior para as vogais anteriores, e menor para as vogais posteriores.

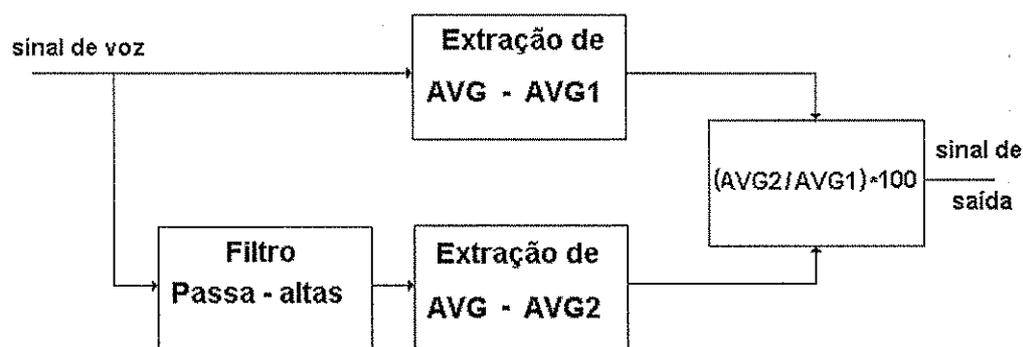


Figura 55 - Diagrama de Blocos do Esquema do 1º parâmetro proposto para auxílio ao ensino de vogais

5.6.2.2 Parâmetro 2

O segundo parâmetro proposto deve fornecer pistas sobre a variação do primeiro formante. Na Figura 56, pode-se observar um sinal de voz típico de uma vogal, com sua respectiva distribuição espectral. O primeiro formante corresponde ao primeiro pico de intensidade no espectro do sinal. Levando-se em consideração a maior influência do primeiro formante nas oscilações do espectro do sinal, propõe-se utilizar neste trabalho, pelo grau de relação com a frequência do primeiro formante o parâmetro taxa de cruzamentos por zero (TCZ), com o objetivo de comparar o desempenho das transformações de dois parâmetros propostas.

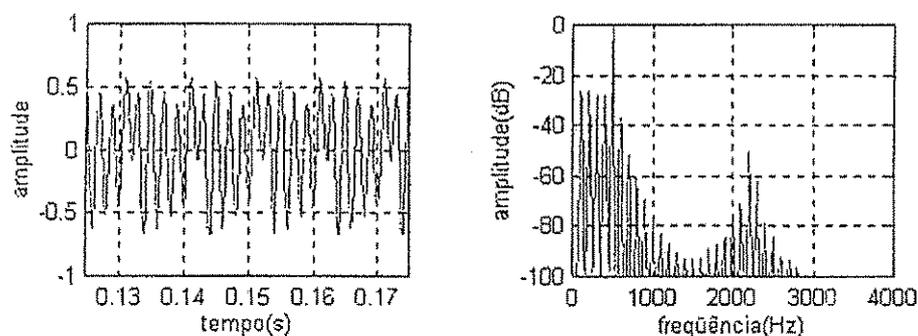


Figura 56 - Sinal de Voz de uma vogal típica e sua correspondente distribuição espectral

5.6.3 Coleta de Dados e Extração de Parâmetros

Foi montado um banco de dados com vozes sustentadas de dois voluntários adultos e do sexo masculino. A gravação dos dados foi feita diretamente em um microcomputador, em um ambiente tipo escritório. Os sinais foram adquiridos com 16 bits por amostra, e com frequência de amostragem de 44.100 Hz, assegurando uma qualidade de áudio em padrão CD. Os parâmetros AVG foram extraídos através de um programa em linguagem C. No ambiente MATLAB foi implementado o filtro e foram gerados os sinais dos parâmetros de saída, além de todo o tratamento para a apresentação gráfica.

Os parâmetros de voz foram extraídos a cada 5 ms, utilizando-se uma janela retangular de 25 ms (Rabiner & Schafer, 1978). O parâmetro amplitude média (AVG), que representa a "envoltória do sinal" foi obtido a partir da equação (55).

A Taxa de cruzamentos por zero foi obtida a partir das equações (56) e (57).

O filtro passa-altas do parâmetro 1, cujo diagrama de blocos pode ser visto na Figura 56, foi implementado no MATLAB, com as seguintes características: elíptico, de 7ª ordem, com faixa de passagem a partir de 2.100 Hz.

5.6.4 Análise dos Parâmetros

Nesta seção são apresentados os resultados típicos para os parâmetros 1 e 2.

5.6.4.1 Parâmetro 1

Os sinais dos parâmetros 1 obtidos para cada vogal foram plotados no mesmo gráfico, na Figura 57. A partir deste gráfico, é possível concluir que o

Parâmetro 1 ocupa faixa de valores distintos para as vogais anteriores, posteriores e média, constituindo-se em um indício que o mesmo pode ser uma pista acústica adequada para ser apresentada via estímulos.

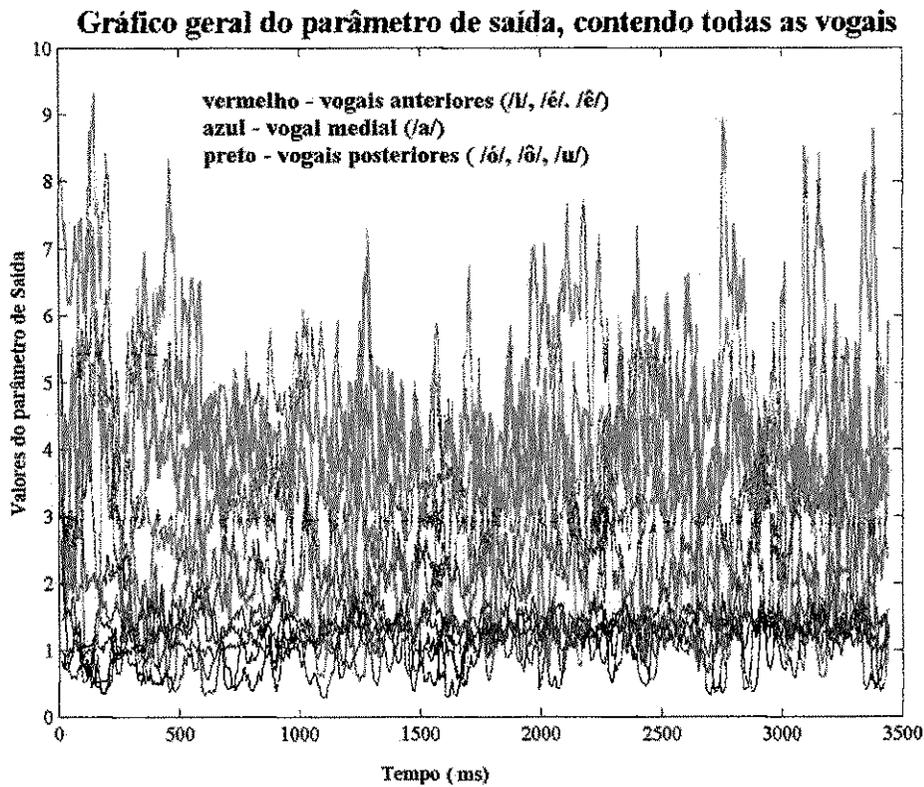


Figura 57 - Gráfico Geral, do parâmetro 1, para todas as vogais estudadas

5.6.4.2 Parâmetro 2

Nas Figuras 58 e 59 pode-se observar os gráficos de TCZ para o grupo de vogais posteriores e anteriores.

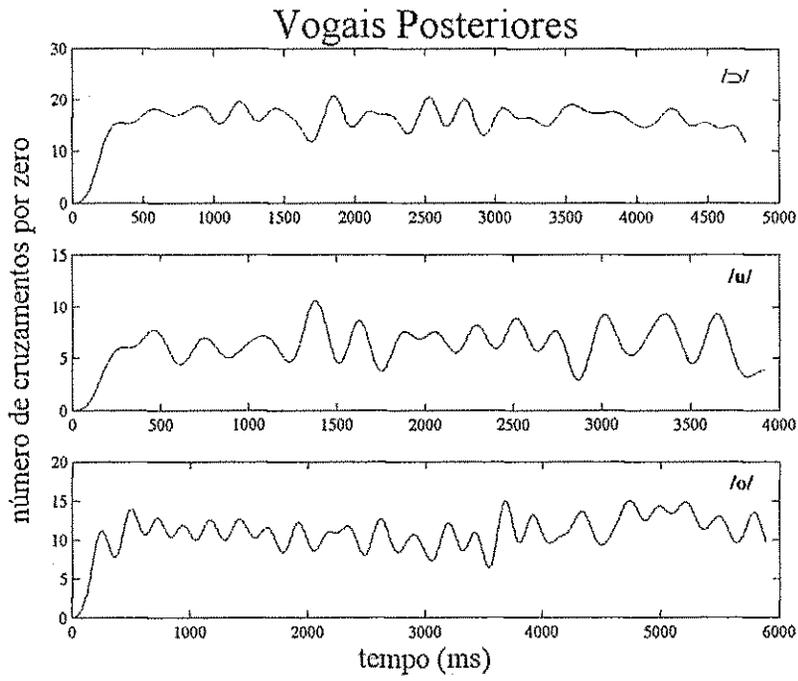


Figura 58 - Variação da TCZ das vogais posteriores

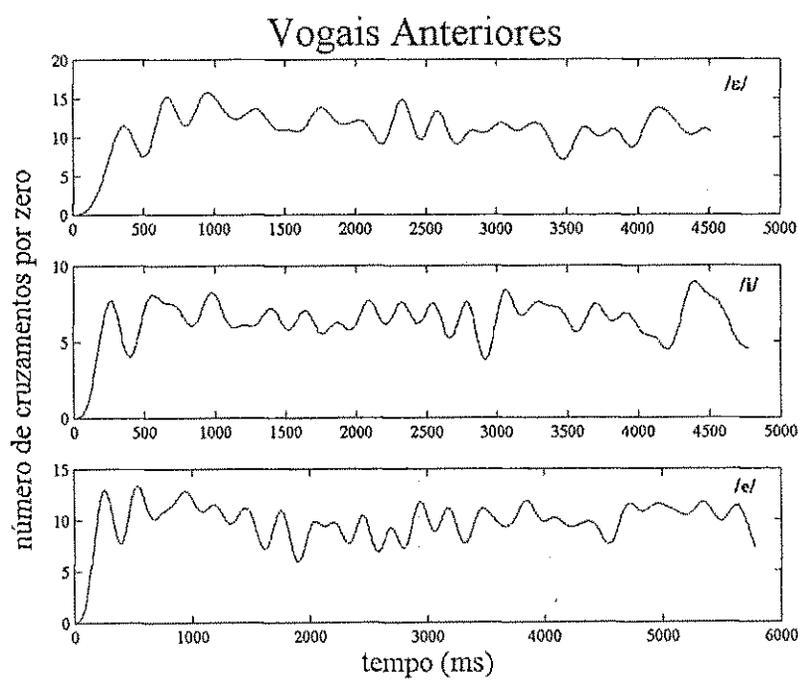


Figura 59 - Variação da TCZ das vogais anteriores

Observando-se os gráficos das Figuras 58 e 59, pode-se ver que tanto no grupo de vogais anteriores como no grupo de vogais posteriores, as vogais de primeira formante mais elevada apresentaram maiores valores de TCZ, e as de menores valores de primeira formante apresentaram menores valores de TCZ.

Assim, foi aplicada uma filtragem passa-baixas nos sinais de voz das vogais sustentadas, com frequência de corte em 1000 Hz, e calculada a TCZ para cada vogal. Os resultados encontram-se nas Figuras 60 e 61.

Pode-se observar na Figura 61 que a TCZ acompanhou o comportamento esperado da primeira formante. Na Figura 60, observa-se o mesmo para as vogais /i/ e /ê/, mas não para a vogal /é/.

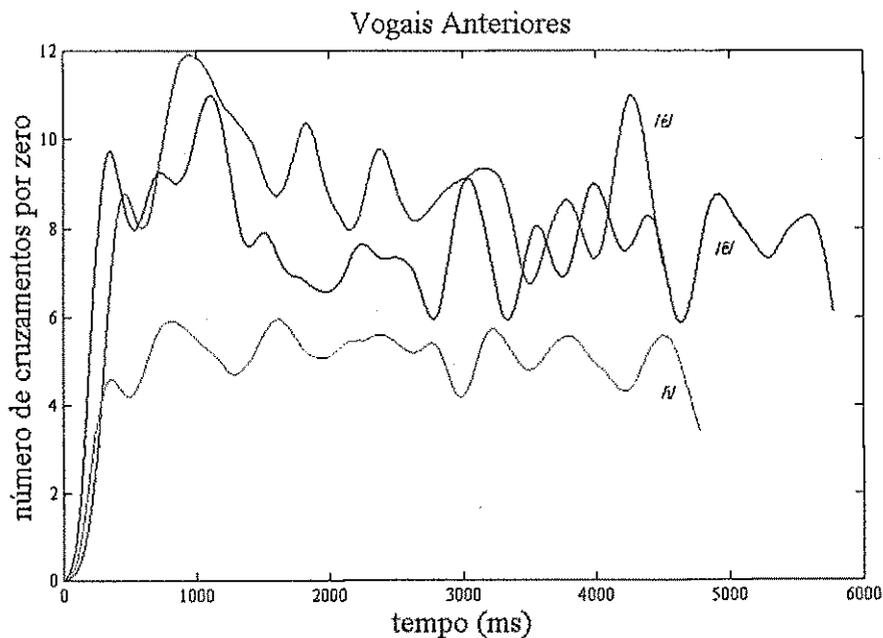


Figura 60 - Variação de TCZ das vogais anteriores

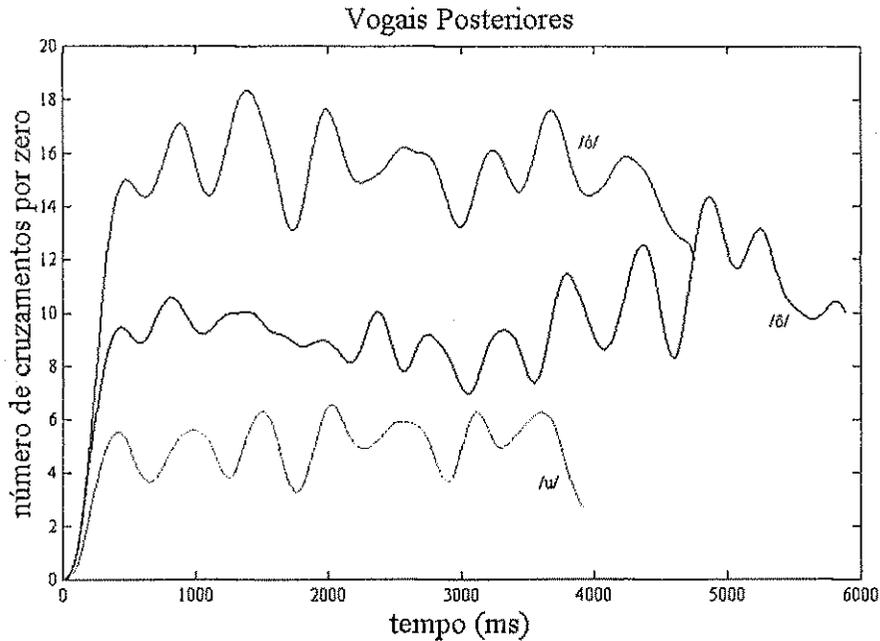


Figura 61 - Variação de TCZ das vogais posteriores

5.6.5 Experimentos de Percepção

Os experimentos de percepção foram realizados com o objetivo de avaliar comparativamente o desempenho entre as três transformações de dois parâmetros propostas no capítulo 4, em termos de facilidade de acerto de alvo.

5.6.5.1 Metodologia

Considerando-se a mesma base de dados de vogais descrita na seção 5.5.3, foi montada uma série de experimentos de acerto de alvo para ajuste dos parâmetros alvo e os parâmetros limites dos parâmetros 1 e 2 (TCZ) para cada vogal. De posse destes valores, foi montada uma nova série de experimentos para verificar a discriminação entre vogais para as transformações linear, cúbica e quadrática de dois parâmetros. Os experimentos são descritos a seguir:

Escolha de Pa, p1 e p2 para o parâmetro 1:

Observando-se os histogramas de distribuição de valores do parâmetro 1 para as vogais anteriores, foram feitos testes de identificação de vogais anteriores a partir da percepção da estimulação vibrotátil com o sinal de estimulação codificado através de uma transformação quadrática de um parâmetro. Na Tabela 12, pode-se verificar os valores dos parâmetros Pa, p1 e p2 utilizados. No experimento foram aplicados aleatoriamente sinais de estimulação vibrotátil modulados por parâmetros extraídos de qualquer uma das vogais. Cada experimento continuou até serem atingidos 25 percepções de identificação de alvo (estimulação de vogal anterior). A comunicação de acerto de alvo foi realizada através de um clique no mouse. Dentro de um experimento completo foram contadas o número de vogais anteriores que não foram identificadas, e a partir daí extraído o índice de aproveitamento (I_a), conforme se vê na equação (64).

$$I_a = \frac{N_a}{N_e} \quad (64)$$

na qual N_a é o número de identificações corretas e N_e é o número de estimulações corretas.

O procedimento feito com as vogais anteriores foi feito também com as vogais anteriores e medial, e pode-se verificar os resultados na Tabela 12. As estimulações foram feitas em um voluntário com experiência em estimulação vibrotátil.

Escolha de Pa, p1 e p2 para o parâmetro 2:

Observando-se os histogramas de distribuição de valores do parâmetro 2 para as vogais *é*, *í* e *ê*, foram feitos testes de identificação a partir da percepção da estimulação vibrotátil com o sinal de estimulação codificado através de uma transformação quadrática de um parâmetro. Na Tabela 13, pode-se verificar os

valores dos parâmetros P_a , p_1 e p_2 utilizados. No experimento foram aplicados aleatoriamente sinais de estimulação vibrotátil modulados por parâmetros extraídos de qualquer uma das vogais anteriores. Cada experimento continuou até serem atingidos 25 percepções de identificação de alvo (estimulação de vogal anterior). A comunicação de acerto de alvo foi realizada através de um clique no mouse. Dentro de um experimento completo foram contadas o número de vogais que não foram identificadas, e a partir daí extraído o índice de aproveitamento (I_a), conforme se vê na equação (67).

O mesmo procedimento foi feito para cada vogal do grupo das posteriores, e os resultados podem ser vistos na Tabela 14.

Identificação de Vogais:

Os experimentos de identificação de vogais foram feitos para as três transformações de dois parâmetros propostas, com base nos parâmetros alvo e parâmetros-limite definidos para a transformação quadrática, nos experimentos anteriores, conforme se pode ver na Tabela 15.

Para a vogal *é*, foram aplicados aleatoriamente sinais de estimulação vibrotátil modulados por parâmetros extraídos de qualquer uma das vogais. Cada experimento continuou até serem atingidos 25 percepções de identificação de alvo (estimulação de vogal *é*). A comunicação de acerto de alvo foi realizada através de um clique no mouse. Dentro de um experimento completo foram contados os números de vogais *é* que não foram identificadas, e a partir daí extraído o índice de aproveitamento (I_a), conforme se vê na equação (67).

O mesmo procedimento foi feito para cada vogal, excetuando-se a vogal *a*. A vogal *a* foi excluída pelo fato do parâmetro 2 não permitir a discriminação adequada desta vogal. Os resultados obtidos para as transformações quadrática, linear e cúbica encontram-se nas Tabelas 16, 17 e 18, respectivamente.

5.6.5.2 Sinal de Estimulação Vibrotátil

O sinal de estimulação vibrotátil foi codificado de acordo com as equações (30), (37) e (38) para as transformações quadrática, linear e cúbica, respectivamente, e o parâmetro b foi calculado de forma que as transformações sejam equilibradas entre os dois parâmetros. Os demais parâmetros foram escolhidos com base na Tabela 15, e foi utilizado o menor grau de dificuldade, isto é, $S = S_{max}$, conforme se obteve a expressão no capítulo 4.

5.6.5.3 Resultados Obtidos

Os resultados obtidos podem ser observados nas Tabelas 12 a 18.

Tabela 12 - Matriz de Confusão do Parâmetro 1, Transformação Quadrática, para as Vogais Posteriores, Anteriores e Medial

Vogal	p1(%)	Pa (%)	p2(%)	Anterior (%)	Medial (%)	Posterior (%)	Aproveitamento (%)
Anterior	0,0	10,0	20,0	56,0	12,0	32,0	96,5
Anterior	4,0	13,0	22,0	100,0	0,0	0,0	100,0
Anterior	5,0	14,0	23,0	100,0	0,0	0,0	100,0
Medial	1,0	4,0	7,0	0,0	84,0	16,0	100,0
Medial	0,0	4,0	8,0	0,0	100,0	0,0	80,6
Posterior	0,0	1,5	3,0	0,0	92,0	8,0	79,3
Posterior	0,0	2,0	4,0	0,0	100,0	0,0	75,7
Posterior	0,0	1,0	2,0	0,0	100,0	0,0	89,2

Tabela 13 - Matriz de Confusão do Parâmetro 2 (TCZ), Transformação Quadrática, para as Vogais é, i e ê.

Vogal	p1	Pa	p2	é (%)	i (%)	ê (%)	Aproveitamento (%)
é	10,0	12,5	15,0	92,0	0,0	8,0	88,5
i	4,0	6,0	8,0	0,0	100,0	0,0	96,2
ê	7,0	9,0	11,0	8,0	0,0	92,0	92,0

Tabela 14 - Matriz de Confusão do Parâmetro 2 (TCZ), Transformação Quadrática, para as Vogais ó, ô e u.

Vogal	p1	Pa	p2	ó (%)	ô (%)	u (%)	Aproveitamento (%)
ó	13,0	16,0	19,0	100,0	0,0	0,0	100,0
ô	9,0	11,0	15,0	0,0	100,0	0,0	100,0
u	7,0	9,0	11,0	0,0	0,0	100,0	100,0

Tabela 15 - Parâmetros Utilizados na Codificação das Transformações Linear, Cúbica e Quadrática de Dois parâmetros

Vogal	p11	pa1	p21	p12	pa2	p22
é	5	14	23	10	12,5	15
i	5	14	23	4	6	8
ó	0	1	2	13	16	19
u	0	1	2	4	7	10
ê	5	14	23	7	9	11
ô	0	1	2	9	12	15

Tabela 16- Matriz de Confusão dos Parâmetros 1 e 2 (TCZ),
Transformação Linear, para as Vogais é, i, ó, u ê, ô.

Vogal	a (%)	é (%)	i (%)	ó (%)	u (%)	ê (%)	ô (%)	Aproveitamento (%)
é	44,0	44,0	0,0	8,0	0,0	4,0	0,0	84,6
i	0,0	12,0	40,0	0,0	24,0	16,0	8,0	100,0
ó	8,0	0,0	0,0	36,0	28,0	0,0	28,0	100,0
u	0,0	0,0	0,0	4,0	28,0	0,0	68,0	46,7
ê	0,0	8,0	20,0	8,0	20,0	36,0	8,0	81,8
ô	12,0	0,0	0,0	20,0	4,0	4,0	60,0	60,0

Tabela 17- Matriz de Confusão dos Parâmetros 1 e 2 (TCZ),
Transformação Quadrática, para as Vogais é, i, ó, u ê, ô.

Vogal	a (%)	é (%)	i (%)	ó (%)	u (%)	ê (%)	ô (%)	Aproveitamento (%)
é	24,0	28,0	0,0	44,0	28,0	4,0	0,0	70,0
i	0,0	0,0	52,0	0,0	48,0	0,0	0,0	86,7
ó	4,0	0,0	4,0	92,0	0,0	0,0	0,0	95,8
u	0,0	0,0	0,0	0,0	96,0	0,0	4,0	100,0
ê	0,0	12,0	0,0	0,0	0,0	32,0	56,0	88,9
ô	0,0	0,0	0,0	0,0	4,0	0,0	96,0	96,0

Tabela 18 - Matriz de Confusão dos Parâmetros 1 e 2 (TCZ),
Transformação Cúbica, para as Vogais é, i, ó, u ê, ô.

Vogal	a (%)	é (%)	i (%)	ó (%)	u (%)	ê (%)	ô (%)	Aproveitamento (%)
é	0,0	92,0	0,0	0,0	0,0	8,0	0,0	82,1
i	0,0	0,0	100,0	0,0	0,0	0,0	0,0	100,0
ó	4,0	0,0	4,0	36,0	8,0	4,0	44,0	55,0
u	0,0	0,0	4,0	0,0	96,0	0,0	0,0	82,8
ê	0,0	4,0	12,0	0,0	0,0	84,0	0,0	95,5
ô	0,0	0,0	0,0	4,0	0,0	0,0	96,0	96,0

5.7 Discussões

Dos resultados obtidos na simulação, comparando-se as transformações linear, cúbica e quadrática de dois parâmetros, feito no capítulo 4, concluiu-se em linhas gerais que a transformação linear apresenta a melhor definição no acerto do alvo, e no outro extremo, a transformação cúbica apresentou a maior facilidade de obter-se percepção de acerto de alvo, embora, como no caso da transformação cúbica de um parâmetro, esta transformação pode mascarar erros no acerto real do parâmetro alvo. A transformação quadrática apresentou comportamento intermediário entre estes dois extremos, em função de que, para estes experimentos de percepção, foi utilizada para se definir os parâmetros alvo e limite para configuração das três transformações, com uma base comum para serem comparadas.

Na Tabela 12, se pode observar que foi possível configurar os parâmetros da transformação quadrática de um parâmetro de modo a obter discriminação entre o grupo de vogais anteriores, grupo de vogais posteriores e vogal medial. No entanto houve maior dificuldade de identificar com ótimo aproveitamento as vogais medial e posterior, indicando uma distribuição dos parâmetros em torno do parâmetro alvo, gerando, apesar de um erro menor que o das vogais de outros grupos, que permite a sua discriminação, um distanciamento da freqüência ideal de acerto de alvo.

Nas tabelas 13 e 14 é possível analisar a discriminação no grupo de vogais posteriores e anteriores, por exemplo, pode-se verificar se é possível, através do parâmetro 2 (TCZ), que configura uma transformação quadrática de um parâmetro, discriminar o fonema *ê* dos fonemas *é* e *i*. Assim, verifica-se que é possível, havendo menor aproveitamento entre as vogais anteriores e maior facilidade entre as vogais posteriores, de modo contrário ao que ocorreu com o parâmetro 1.

Portanto, com base nos resultados das tabelas 12, 13 e 14, configurou-se as transformações linear, quadrática e cúbica de dois parâmetros, tendo em vista o estudo comparativo do comportamento entre estas três transformações. Os parâmetros utilizados encontram-se sumarizados na Tabela 15.

Pode-se observar na Tabela 16 os resultados obtidos para a transformação linear quando se verifica que pelo fato dos parâmetros 1 e 2 ocuparem uma faixa em torno do parâmetro alvo, os sinais saíram fortemente da região de acerto do alvo, e o desempenho da transformação foi muito ruim. Assim, para casos em que há uma faixa de parâmetros alvo, esta configuração não é adequada. Abaixo se apresenta uma possibilidade de ajustar-se a transformação por meio da colocação de um termo DC (offset), conforme na equação (68), que pode ser pesquisado em trabalhos futuros.

$$Ft(t) = Fc + (a * |Pe1(t) - Pa1| + b * |Pe2(t) - Pa2| - offset) * S \quad (65)$$

Na Tabela 17 é possível verificar os resultados obtidos pela transformação quadrática, que também apresenta resultados ruins, embora melhores que o da transformação linear. Porque não se obteve bons resultados se os parâmetros foram ajustados com base nesta transformação? Concluiu-se então que isto ocorreu porque quando os erros obtidos individualmente, que eram satisfatórios para as transformações quadráticas de um parâmetro se somaram na equação da transformação quadrática de dois parâmetros, o resultado foi para um ponto mais distante da região do alvo, dificultando a discriminação entre as vogais. Logo, este é um problema que aumentará com o número de parâmetros da transformação e a solução pode ser a colocação de um fator de $1/n$, onde n é o número de parâmetros da transformação, multiplicando o erro da transformação de n parâmetros. Uma segunda opção é configurar os parâmetros individualmente com uma transformação linear de um parâmetro (utilizando o módulo). Outra opção é o uso de sistemas inteligentes com possibilidade de aprendizado, para definição dos parâmetros alvo, onde além da variação no parâmetro alvo e das faixas de parâmetros, pode-se variar os parâmetros a e b , e trabalhar com sistemas desequilibrados.

Quanto à transformação cúbica, verificou-se que o ajuste dos parâmetros com base na transformação quadrática é adequado, pois apesar dos erros individuais se

somarem, sendo menores que um, diminuem quando elevados ao cubo. Talvez se o sistema tivesse sido configurado por transformações cúbicas de um parâmetro (utilizando o módulo), tivéssemos observado as mesmas dificuldades verificadas na transformação quadrática.



CONCLUSÕES E SUGESTÕES

6.1 Introdução

A estimulação tátil vem sendo estudada desde o início do século XX como um canal alternativo para a transmissão da fala para surdos. No decorrer deste período muitas pesquisas foram feitas, constatando-se a partir de meados de 1970 que o tato possuía capacidade de transmissão de informações inferior ao canal auditivo, razão pela qual não se conseguiu transmitir a informação completa da fala por este canal. Apesar disto, pesquisas subseqüentes baseadas no método natural de Tadoma demonstraram que estimuladores baseados na estimulação conjunta dos sentidos vibrotátil e proprioceptivos tinham capacidade de transmissão de informações no mesmo nível da do canal auditivo. Ainda não se encontram relatos, entretanto, de testes destes sistemas na transmissão da fala.

Apesar de não se ter conseguido empregar os estimuladores táteis para a transmissão completa da fala, os estimuladores de um ou mais canais, vibrotáteis ou eletrotáteis desenvolvidos, apresentaram resultados positivos no auxílio ao aprendizado da fala, em termos de melhor controle da entonação e aumento da qualidade da vocalização. Quando comparados ao sentido visual, cita-se que o tato contribui na transmissão de variações instantâneas da fala, enquanto a visão contribui com variações médias. A possibilidade de integração de estímulos táteis e visuais em uma ferramenta de auxílio ao ensino da fala é um tema de pesquisa que está em aberto.

Os estimuladores táteis podem ter um ou mais canais, sendo que cada canal é acionado por um sinal de estimulação. Em termos de estimulação vibrotátil, o sinal de estimulação possui as características de intensidade e freqüência, que podem ser moduladas por parâmetros de voz. Os estimuladores vibrotáteis de um canal são mais simples e mais fáceis de serem usados, mas pesquisas concluíram que devem ser modulados por apenas um parâmetro de voz. Assim, quando há necessidade de transmissão de mais de um parâmetro, deve-se aumentar o número de estimuladores.

Considerando-se o contexto exposto nos parágrafos anteriores, apresentou-se neste trabalho uma investigação sobre o desenvolvimento de um sistema para auxílio ao ensino da fala utilizando-se estimulação vibrotátil, ou seja, as possibilidades de transmissão de informação acústica pelo tato não foram estudados com o objetivo de transmissão e recepção da fala, mas no contexto reduzido do auxílio ao aprendizado da fala.

Em termos de tecnologias para auxílio a oralização de surdos, os sistemas por realimentação visual existentes para auxílio ao ensino da fala utilizam jogos com indicação de acerto de alvo, e em termos de estimulação tátil, a transmissão da envoltória do sinal de voz é importante para melhorar a qualidade da vocalização e o controle da entonação.

Assim, para possibilitar o uso de estimuladores táteis na indicação de acerto de alvo foram propostas transformações de parâmetros genéricos na faixa de frequências de estimulação vibrotátil.

A modulação do sinal de estimulação vibrotátil por parâmetros de voz, com fim de transmitir a informação acústica foi estudada no contexto da língua portuguesa, através do estudo de caso da discriminação entre visemes do português.

Para avaliar comparativamente as transformações propostas foram feitos testes de percepção, com base no estudo de caso do controle da frequência fundamental (um parâmetro) e discriminação entre vogais (dois parâmetros).

6.2 Parametrização do Sinal de Voz e Codificação do Sinal de Estimulação Vibrotátil

Levando-se em consideração a indefinição dos trabalhos na literatura quanto ao parâmetro mais adequado para codificação de um estimulador vibrotátil monocanal, tendo em vista o auxílio à leitura labial e a não existência, na literatura pesquisada de análises por etapas, dado que as análises são feitas com base em resultados de percepção ao estímulo vibrotátil ou auditivo, estudou-se a

parametrização do sinal de voz no contexto do português, em duas etapas: análise de distribuição de valores dos parâmetros e testes de percepção. Para poder analisar a distribuição de valores do parâmetro, estudou-se o caso específico da discriminação entre fonemas que compõem visemes, que são um problema no auxílio à leitura labial. Desta investigação, concluiu-se:

- Um esquema misto considerando os dois parâmetros, AVG e TCZ, foi proposto, testado e mostrou melhores resultados do que cada parâmetro isoladamente;
- Os resultados são válidos apenas para segmentos de voz tipo /VcV/;
- Os testes de percepção confirmaram os resultados deduzidos pela análise dos histogramas de distribuição de valores dos parâmetros.

6.3 Transformações de Um Parâmetro

- As transformações propostas permitem a implementação de sistemas para indicação de acerto de alvo.
- Em termos da implementação de jogos, as transformações propostas apresentam, quanto à variação do parâmetro S, diversos níveis de dificuldade, e entre elas próprias também, sendo a transformação quadrática a de maior nível de facilidade;
- Os resultados obtidos de percepção de estímulos vibrotáteis, no nível de menor dificuldade, foram próximos dos obtidos utilizando-se o sentido da audição, principalmente quanto a transformação quadrática;
- A transformação quadrática de um parâmetro é a mais simples de se utilizar, pois a característica de inflexão da curva no momento do acerto

do alvo diminui a necessidade de se memorizar a frequência que está sendo estimulada.

6.4 Transformações de Mais de Um Parâmetro

- As transformações de dois parâmetros permitem a percepção de acerto de características da voz mais complexas, como as vogais, pela codificação de dois parâmetros em um só canal.
- Com o uso das transformações de dois ou mais parâmetros é possível transmitir-se mais de um parâmetro por estimuladores vibrotáteis monocanal, o que não é possível por outras formas de codificação citadas na literatura.
- Com a utilização da função módulo com as transformações linear e cúbica, as mesmas passaram a apresentar a mesma característica de inflexão da função quadrática.
- A definição de valores de parâmetros com base em testes isolados utilizando estimuladores de um parâmetro leva a um erro nas transformações de dois parâmetros, já que o erro fica com valor duplicado. Verificou-se que a definição dos valores com base na transformação quadrática levou a bons resultados na transformação cúbica.
- Utilizando-se a função módulo é possível implementar-se transformações linear e cúbica de um parâmetro com a característica de inflexão da transformação quadrática de um parâmetro;

6.5 Contribuições

- Escolha de parâmetros de voz não apenas por resultados de teste de percepção, mas através da análise das características do sinal de voz;
- Aplicação da estimulação vibrotátil no contexto da língua portuguesa;
- Abordagem para desenvolvimento de sistemas de estimulação vibrotátil com base nas necessidades e características do ensino ao aprendizado da fala;
- Proposta de transformações de um, dois, três e n parâmetros para transmissão de informações indicativas de acerto de alvo, por estimulação vibrotátil, com aplicação na transmissão de informações acústicas, como também se podendo expandir os resultados e propostas para outras áreas, como tele-operação, auxílio a outros tipos de deficiência, etc...

6.6 Sugestão para Trabalhos Futuros

6.6.1 Integração Estimulação Tátil e Visual

A questão que está em aberto é qual a contribuição do estímulo tátil e visual na transmissão de informações acústicas? São complementares? Como integrá-los? Neste contexto também pode ser estudada a possibilidade do uso misto de sistemas vibrotáteis e eletrotáteis. Além disso, pode-se verificar a adequação, resultados e adaptações das transformações propostas neste trabalho para a estimulação eletrotátil.

6.6.2 O desenvolvimento do Sistema para Uso com Surdos:

Parametrização e Definição de Alvos

A questão que está em aberto é: Pode-se parametrizar e definir exercícios/alvos para um sistema de auxílio ao aprendizado da fala com base no

que se conhece da fala correta, ou tal sistema deve ser bem aberto para que o terapeuta configure? Quais as vantagens e desvantagens do ponto de vista das transformações propostas neste trabalho? É possível desenvolver sistemas inteligentes, que variando os parâmetros de configuração das transformações possam ser treinados para ajustar a condição de acerto de alvo? Neste sentido pode-se aprofundar nos vários algoritmos para processamento de voz existentes tendo em vista a parametrização de sinais de voz no contexto da estimulação vibrotátil.

6.6.3 O Desenvolvimento de Sistemas Específicos para Surdocegos e Multideficientes

A classe de deficientes surdocegos e pessoas com plurideficiência tem dificuldades sérias de comunicação, e os profissionais da área não possuem muitas ferramentas para auxílio à comunicação. O tato, sem dúvidas é muito utilizado neste caso. Assim, quais as necessidades pedagógicas destes deficientes? Quais as carências de ferramentas? Como adaptar a estimulação vibrotátil para estas necessidades, podendo-se utilizar jogos lúdicos com base nas transformações aqui propostas?

BIBLIOGRAFIA

1. ALMEIDA, K. & IORIO, M. C. M. *Próteses auditivas – fundamentos teóricos & aplicações clínicas*. São Paulo: Lovise, 1996.
2. ALMEIDA, M. S. T. *Compreendendo a Síndrome Pseudobulbar*. Monografia de Especialização. Centro de Especialização em Fonoaudiologia Clínica (CEFAC), São Paulo, 1997.
3. AMORIM, A. *Fonoaudiologia Geral*. 1ª edição. 1972.
4. ANTONINO, P. H. D. *Estudo Preliminar para o Desenvolvimento de um Sistema de Substituição da Visão através da Sensação Tátil, utilizando estimulação Eletrocutânea*. Dissertação de mestrado – UNICAMP, 1993.
5. ARAÚJO, A. M. L., & LIMA, M. C. P. M. *Trabalhando a voz de deficientes auditivos com auxílio do computador*. In: Quevedo, A. A. F., Oliveira, J. R. & Mantoan, M. T. E. (org.) *Mobilidade e comunicação: Desafios à tecnologia e à inclusão social*, Campinas: Unicamp, 1999.
6. ARAÚJO, A. M. L. *Jogos Computacionais Fonoarticulatórios Para Crianças com Deficiência Auditiva*. Tese de Doutorado – UNICAMP, 2000.
7. BARROS, A. T. *Estudo de Sistemas de Estimulação Eletrotátil para Percepção de Informações Sonoras*. Dissertação de Mestrado – UFPB, 1996.
8. BARROS, A. T., FREIRE, R. C. S. & DEEP, G. S. *Estimulador Vibrotátil de um Canal*. Anais do II Congresso Internacional METROCAL, Associação Chilena de Metrologia, 2001.

9. **BARROS, A. T., DEEP, G. S. & FREIRE, R. C. S.** *Análise de Parâmetros Temporais de Voz para Uso em Dispositivos de Ajuda Tátil.* Anais do XVIII Congresso Brasileiro de Engenharia Biomédica, Subárea Engenharia de reabilitação, São José dos Campos, 2002.
10. **BARROS, A. T., FREIRE, R. C. S. & DEEP, G. S.** *Choice of Speech Features for Single-channel Vibrotactile Stimulators.* The Journal of the Acoustical Society of America, 112(5), 2002.
11. **BEHLAU, M. & PONTES, P.** *Avaliação e tratamento das disfonias.* São Paulo: Lovise, 1995.
12. **BERNSTEIN, L. E., EBERHARDT, S. P., & DEMOREST, M. E.** *Single-channel Vibrotactile Supplements to Visual perception of Intonation and Stress.* The Journal of Acoustical Society of America, 85(1), janeiro, 1989.
13. **BERNSTEIN, L. E., AUER, E. T., COULTER, D. C., TUCKER, P. E., & DEMOREST, M. E.** *Single versus Multichannel Vibrotactile Supplements to Intonation and Stress by Normal-hearing and Hearing-impaired Adults.* The Journal of the Acoustical Society of America, 93, 1993.
14. **BERNSTEIN, L. E.,** *The Evaluation of Tactile Aids (capítulo 8),* em Tactile Aids for The Hearing Impaired, Ians Summer (ed), Whurr Publishers, London, 1992.



15. **BERNSTEIN, L. E.**, *Toward Future tactile Aids*(capítulo 8). In: Plant, G & Spens K. (org.) *Profound deafness and speech communication*, Whurr Publishers, Londres, 1995.
16. **BLAMEY, P. J. & COWAN, R. S. C.** *The Potential Benefit and Cost-effectiveness of Tactile Devices in Comparison with Cochlear Implants*, capítulo 9, em *Tactile Aids for The Hearing Impaired*, Ians Summer (ed), Whurr Publishers, London, 1992.
17. **BLOCH, P.** *Problemas da voz e da fala*. Rio de Janeiro: Letras e Artes, 1963.
18. **BOONE, D.** *A voz e a Terapia Vocal*. Artes Médicas. Porto Alegre, 1994.
19. **BOOTHROYD, A.** *Hearing impairments in children*. New York: Prentice Hall, 1982.
20. **BOOTHROYD, A & HNATH, T.** *Lipreading with Tactile Supplements*. *Journal of Rehabilitation Research Devices*, 23, 1986.
21. **BRASIL, Ministério da Educação e do Desporto, Secretaria de Educação Especial**, *Subsídios para organização e funcionamento de serviço de educação especial: área de deficiência auditiva*. Brasília: MEC, Série Diretrizes, nº 6, 1995.
22. **BREEUWER, M. & PLOMP, R.** *Speechreading Supplemented with Formant-frequency Information from Voiced Speech*. *The Journal of Acoustical Society of America*, 77(1), janeiro, 1985.

23. **BROOKS, P. L. & FROST, B. J.** *Evaluation of a Tactile Vocoder for Word Recognition*. The Journal of Acoustical Society of America, 74(1), julho, 1983.
24. **BROOKS, P. L., FROST, B. J., MASON, J. L. & CHUNG, K.** *Acquisition of a 250-word Vocabulary through a Tactile Vocoder*. The Journal of Acoustical Society of America, 77(4), abril, 1985.
25. **BROOKS, W. D.** *Speech Communication*. Segunda edição, 1994.
26. **BUSH, M. A.** *Vowel Articulation and Laryngeal control in the Speech of the Deaf*. Ph.D. dissertation, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology-MIT, 1981.
27. **CARNEY, A. E. & BEACHER, C.R.** *Vibrotactile Perception of Suprasegmental Features of Speech: A Comparison of Single-channel and multichannel Instruments*, Journal of Acoustical Society of America. 79(1), janeiro, 1986.
28. **CARNEY, A. E.** *Vibrotactile Perception of Segmental Features of Speech: A Comparison of Single and Multichannel Aids*, Journal of Speech Hearing Research, 31, 1988.
29. **CHABA, Working Group on Communication Aids for the Hearing-Impaired, Committee on Hearing, Bioacoustics and Biomechanics.** *Speech-perception aids for hearing-impaired people: Current status and needed research*. Journal of Acoustical Society of America, 90 (2), agosto, 1991.

30. **DELHORNE, L.A.** *The Reception of Sign Language Through the Tactile Sense.* American Speech Language Hearing Association. 30, 136(A), 1988.
31. **EBERHARDT, S. P., BERNSTEIN, L. E., DEMOREST, M. E. & GOLDSTEIN JR, M. H.** *Speechreading Sentences with Single-channel Vibrotactile Presentation of Voice Fundamental Frequency.* Journal of Acoustical Society of America, vol 88, 3, setembro, 1990.
32. **EBERHARDT, S. P., BERNSTEIN, L. E., COULTER, D. C. & HUNCKLER, L. A.** *OMAR – A Haptic Display for Speech perception by Deaf and Deaf-blind Individuals.* Proceedings Virtual reality Annual International Symposium, The IEEE Neural Networks Council, Setembro, 1993.
33. **ENCICLOPÉDIA BARSA,** Rio de Janeiro, 1995
34. **FECHINE, J. M.** *Reconhecimento Automático de Identidade Vocal Utilizando Modelagem Híbrida: Paramétrica e Estatística.* Tese de Doutorado – UFPB, 2000.
35. **FLETCHER, S. & HASEGAWA, A.** *Speech Modification by a Deaf Adult through Dynamic Orometric modelling and feedback.* Journal of Speech and hearing Disorders 48, 1983.
36. **GAULT, R. H.** *Progress in Experiments on Tactual Interpretations of Oral Speech.* Journal of Abnormality Society Psychological. V19, 1924.
37. **GOLDSTEIN, M. H. & PROCTOR, A.** *Tactile Aids for Profoundly Deaf Children.* Journal of Acoustical Society of America. 77, 1985.

38. GRANT, K. W., ARDELL, L., KUHL, P. & SPARKS, D. *The Contribution of Fundamental Frequency, Amplitude Envelope, and Voicing-duration Cues to Speechreading in normal-hearing Subjects*. The Journal of the Acoustical Society of America, 77, 1985.
39. GRANT, K. W. *Encoding Voice Pitch for profoundly Hearing-impaired Listeners*. The Journal of Acoustical Society of America, 82, 1987.
40. GREEN, B. G., CRAIG, J. C., WILSON, A. M., PISONI, D. B. & RHODES, R. P. *Vibrotactile identification of Vowels*. The Journal of Acoustical Society of America, 73(5), 1983.
41. HANIN, J., BOOTHROYD, A. & HNATH-CHISOLM, T. *Tactile presentation of Voice Fundamental Frequency as an Aid to the Speechreading of Sentences*. Ear Hearing, 9, 1988.
42. HUDGINS, C. V. *Visual Aids in the Correction of Speech*, Volta Review 37, 637-643, 1935.
43. JCIH - JOINT COMMITTEE ON INFANT HEARING, *1994 Position Statement*, Pediatrics. 95(1), Janeiro, 1995.
44. KACZMAREK, K. A., WEBSTER, J. G. & BACH-Y-RITA, P. *Electrotactile and Vibrotactile Displays for Sensory Substitution Systems*. IEEE Transactions on Biomedical Engineering, vol 38, Jan, 1991.
45. KIRMAN, J. H. *Tactile perception of Computer-derived Formant Patterns from Voiced Speech*. The Journal of Acoustical Society of America, 55(1), janeiro, 1974.

46. **KISHON-RABIN, L., BOOTHROYD, A. & HANIN, L.** *Speechreading Enhancement: A Comparison of Spatial-Tactile Display of voice fundamental frequency (Fo) with Auditory Fo.* The journal of Acoustical Society of America, 100(1), julho, 1996.
47. **KNUDSEN, V. O.** *Hearing with the sense of touch.* Journal of General Psychology, 1, 320–352, 1928.
48. **LING, D.** *Speech and the Hearing Impaired Child: Theory and practice.* Washington: The Alexander Graham Bell Association for the Deaf, 1976.
49. **LING, D.** *Amplification to speech.* In CALVERT, D. R. & SILVERMAN, S. R. *Speech and deafness.* 2a ed. Washington: The Alexander Graham Bell Association for the Deaf, 1978.
50. **MARTINS, M. R.,** *Ouvir Falar,* Editora Caminho, Lisboa, 1988.
51. **MASHIE, J. & YADAV, P.** *The Gallaudet University Speech training and Evaluation System for Deaf Children.* American Speech Language Hearing Association 32:10, 1990.
52. **MASHIE, J.** *The use of sensory aids for teaching speech to children who are deaf.* In: Plant, G & Spens K (org.) *Profound deafness and speech communication,* Whurr Publishers, Londres, 1995.
53. **MONSEN, R. B.** *Acoustic qualities of phonation in young hearing impaired children.* Journal of Speech Hearing Research. 22: 270-288, 1979.

54. MOWBRAY, G. H., & GEBHARD, J. W. *Sensitivity of the skin to changes in the rate of intermittent mechanical stimuli*. *Science* , 125, 1.297–1.298, 1957.
55. MYSAK, E. D. *Patologia dos sistemas da fala*. 2ª ed. Rio de Janeiro: Atheneu, 1988.
56. NOHAMA, P. *Investigação em Propriocepção Artificial*. Tese de Doutorado, UNICAMP, 1997.
57. NOVAES, B. C. & MENDES, B. C. A. *O IBM Speechviewer II na Terapia Fonoaudiológica, Capítulo XII*, em *A Tecnologia Informática na fonoaudiologia*, Flávia Benevides Foz e outros, Ed. Plexus, São Paulo, 1998.
58. NORONHA, M. H. & RODRIGUES, M. H. *O Deficiente da Audição e a Educação Especial*. Ed José Olympio, 1974.
59. NORTHERN, J & DOWNS, M. P. *Audição em Crianças*. 3º edição. São Paulo. Manole, 1989.
60. NORTON, S. J. *Analytic Study of the Tadoma Method: Background and Preliminary Results*. *Journal of Speech Hearing Research*. 20, 574 – 595.
61. OLIVEIRA, P. M. T. *Auxílio Visual à Oralização de Surdos*. Dissertação de Mestrado, COPPE/UFRJ, 1998.
62. ÖSTER, A. *Principles for a complete description of the phonological system of deaf children as a basis for speech training*. In: Plant, G & Spens K.

- (org.) *Profound deafness and speech communication*, Whurr Publishers, Londres, 1995.
63. **PROCTOR, A.** *Tactile aid usage in young deaf children*. In: Plant, G & Spens K. (org.) *Profound deafness and speech communication*, Whurr Publishers, Londres, 1995.
64. **RABINER, L. R. & SCHAFER, R. W.** *Digital Processing of Speech Signals*, Prentice-Hall, 1978.
65. **REDONDO, M. C. & LOPES FILHO, O. C.** *Testes Básicos de Avaliação Auditiva*. In: LOPES FILHO, O.C. (editor.) *Tratado de Fonoaudiologia*. São Paulo: Editora Roca, 1997.
66. **REED, C.M.** *Analytic Study of the Tadoma Method: Discrimination Ability of untrained observers*. *Journal of Speech Hearing Research*, 21, 625 – 637, 1978.
67. **REED, C. M., RABINOWITZ, W. M., DURLACH, N. I., BRAIDA, L. D., CONWAY-FITHIAN, S. & SCHULTZ, M. C.** *Research on The Tadoma Method of Speech Communication*. *The Journal of the Acoustical Society of America*, 1985, 77(1), 247-257.
68. **REED, C. M., DURLACH, N. I., e DELHORNE, L. A., RABINOWITZ, W. M., & GRANT, K. W.** *Research on Tactual Communication of Speech: Ideas, Issues and Findings*. *Volta Review*. 91, 65-78, 1989 .
69. **RICHARDSON, B.**, *Using the Skin for the Purpose of Sound Localization*. In: Gatehouse, R.W. (ed.), *Localization of Sound: Theory and Applications*. Groton, CT: Amphora Press, pp. 155-168, 1982.

70. **RISBERG, A.** *Visual Aids for Speech correction.* American Annals of the Deaf 113(2), 1968.
71. **RODRIGUES, M. V.,** *Aspectos Vocais no Deficiente Auditivo.* Monografia de Especialização. Centro de Especialização em Fonoaudiologia Clínica (CEFAC), São Paulo, 1997.
72. **ROTHENBERG, M., VERRILLO, R. T., ZAHORIAN, S. A., BRACHMAN, M. L., & BOLANOWSKI, S. J. J.** *Vibrotactile frequency for encoding a speech parameter.* The Journal of the Acoustical Society of America, 62, 1003–1012.1977.
73. **ROTHENBERG, M. & MOLITOR, R. D.** *Encoding Voice Fundamental frequency into Vibrotactile Frequency.* The Journal of the Acoustical Society of America, vol 66, 4, outubro, 1979.
74. **RUSSO, I. C. P. & BEHLAU, M.** *Percepção da fala: análise acústica do português brasileiro.* São Paulo: Lovise, 1993.
75. **SACKS, O.** *Vendo vozes.* São Paulo, Companhia das Letras, 1998.
76. **SCHIFF, W. & FOULKE, E.,** *Tactual Perception: a sourcebook,* Cambridge University Press, 1982.
77. **SCHMIDT, R.** *Motor Control and learning: A behavioral Emphasis.* 2ª edição. Human Kinetics Publishers, Inc. Champaign, Il, 1988.
78. **SEHGAL, S. T., KIRK, K. I., SVIRSKY, M., ERTMER, D. J. & OSBERGER, M. J.** *Imitative Consonant Feature Production by Children with Multichannel Sensory Aids.* Ear and Hearing, vol 19, nº 1, 1998.

79. **SHERRICK, C. E.**, *Basic and Applied Research on Tactile Aids for Deaf people: progress and prospects*. Journal of Acoustics Society of America, 75, 1984.
80. **SKLIAR, C.** *Os Estudos Surdos em Educação: Problematizando a Normalidade (Capítulo 1)*. Em A Surdez, Um Olhar sobre as Diferenças. Carlos Skliar (editor), Editora Mediação, 2ª Ed., Porto Alegre, 2001.
81. **SÖDERLUND, G.** *Tactiling and tactile Aids: A User's Viewpoint*. In: Plant, G & Spens K. (org.) *Profound deafness and speech communication*, Whurr Publishers, Londres, 1995.
82. **SPINELLI, V. P., MASSARI, I. C. & TRENCH, M. C. B.** *Distúrbios Articulatorios*. In. *Temas de Fonoaudiologia*, Edições Loyola, São Paulo, 1999.
83. **STERNE, T. A. & ZIMMERMAN, H. J.** *A thyatron indicator for teaching the deaf*, Journal of Science instrumentation. 16, 334-336, 1939.
84. **STEVENS, K., KALIKOW & WILLEMANN, T.** *A Miniature Accelerometer for Detecting Glottal Wave forms and nasalization*. Journal of Speech and Hearing research 18:594, 1975.
85. **SUMMERS, I. R.** *Single Processing Strategies for Single Channel Systems (chapter 5)*, in *Tactile Aids for The Hearing Impaired*, Ians Summer (ed), Whurr Publishers, London, 1992.
86. **SUMMERS, I. R., DIXON, P. R., & COOPER, P. G.** *Vibrotactile and Electrotactile Perception of Time-varying Pulse Trains*, Journal of Acoustical Society of America, 95(3), 1994.

87. **SUMMERS, I. R. & GRATTON, D. A.** *Choice of Speech Features for Tactile Presentation to the Profoundly Deaf*, IEEE Transactions on Rehabilitation Engineering, vol 3, No 1, março, 1995.
88. **SZETO, A. Y. J. & CHRISTSEN, K. M.** *Technological Devices for Deafblind Children: Needs and Potential Impacts*, IEEE Engineering in Medical and Biology Magazine, setembro, 1988.
89. **TAN, H. Z.** *Analysis of Synthetic Tadoma System as a multidimensional Tactile display*. M.S. Thesis, Department of Electrical Engineering and Computer Science, MIT, 1988.
90. **TAN, H. Z.** *Information Transmission with a Multi-finger Tactual Display*, Ph.D. dissertation, Department of Electrical Engineering and Computer Science, MIT, 1996.
91. **THORNTON, A. R. D. & PHILLIPS, A. J.** *A Comparative Trial of Four Vibrotactile Aids (capítulo 11)*, em *Tactile Aids for The Hearing Impaired*, Ians Summer (ed), Whurr Publishers, London, 1992.
92. **VAN TASELL, D. J., SOLI, D. S., KIRBY, V. M. & WIDIN, G. P.** *Speech Waveform Envelope Cues for Consonant Recognition*. Journal of Acoustical Society of America. Vol 82, outubro, 1987.
93. **VERRILLO, R. T.** *Effect of contactor area on the vibrotactile threshold*. Proceedings of the Winter Annual Meeting of the American Society of Mechanical Engineers, 35, 1.962–1.966. 1963.
94. **VERRILLO, R. T., FRAIOLI, A. J., & SMITH, R. L.** *Sensation magnitude of vibrotactile stimuli*. Perception & Psychophysics, 6, 366–372. 1969.

95. VERRILLO, R. T., & GESCHEIDER, G. A. *Perception via the sense of touch*. In I. R. Summers (Ed.), *Tactile Aids for the Hearing Impaired* (p. 1–36). London: Whurr Publishers, 1992.
96. VIEIRA, M. N. *Módulo frontal para um sistema de reconhecimento automático de voz*. Dissertação de Mestrado, UNICAMP, 1989.
97. WATANABE, A. *Speech Visualization System as a Basis for Speech Training and Communication Aids*, capítulo 24, em *Profound Deafness and Speech Communication*, Plant, G. & Spens, K. (Eds), Whurr Publishers, London, 1995.
98. WATSON, C. & KEWLEY-PORT, D. *Advances in Computer-based Speech training: Aids for the profoundly hearing impaired*. *Volta Review* 4, 1989.
99. WEISENBERGER, J. M. *Sensitivity to Amplitude-modulated Vibrotactile Signals*. *The Journal of the Acoustical Society of America*, 86, 1989a.
100. WEISENBERGER, J. M. & MILLER, J. D., *The Role of Tactile Aids in Providing Information about Acoustic Stimuli*, *The Journal of the Acoustical Society of America*, 82, 902-916, 1987.
101. WEISENBERGER, J. M., *Evaluation of the Siemens Minifonator Vibrotactile Aids*. , *Journal of Speech Hearing Research*. 32, 24-32, 1989b.
102. WEISENBERGER, J. M. *Communication of the Acoustic Environment via Tactile Stimuli (chapter 4)*, in *Tactile Aids for The Hearing Impaired*, Ians Summer (ed), Whurr Publishers, London, 1992.

103. **WIGHTMAN, F. L. & KISTLER, D. J.**, *Headphone Simulation of Free-field listening. I: Stimulus Synthesis*. The Journal of the Acoustical Society of America, 85, 858-867, 1989a.
104. **WIGHTMAN, F. L. & KISTLER, D. J.**, *Headphone Simulation of Free-field listening. II: Psychophysical Validation*. The Journal of the Acoustical Society of America, 85, 868-878, 1989b.
105. **WILSON, D. K.** *Problemas de voz em deficientes auditivos – Problemas de voz em crianças*- Ed. Manole 301-327 - São Paulo, 1993.
106. **ZEMPLIN, W. R.** *Princípios de anatomia e fisiologia em fonoaudiologia*. 4ª ed., Porto Alegre: Artmed, 2000.

ANEXO A – RESULTADOS DA PSICOFÍSICA: ESTIMULAÇÃO TÁTIL

DISCRIMINAÇÃO DE INTENSIDADE

A faixa de variação dinâmica dos sistemas de estimulação vibrotáteis é limitada, indo do limiar de detecção a 55 dB acima deste limiar, após isto as vibrações tornam-se desconfortáveis. O parâmetro que mede o grau de discriminação de intensidade é o JND (Just Noticeable Difference). Basicamente, este parâmetro indica qual a mínima variação que é possível ser dada em um atributo de referência de modo que o indivíduo estimulado perceba uma variação de estímulo aplicado. Assim, foi verificado por KNUDSON (1928) que o JND de intensidade $[(A+DA)/A]$ (em que A é amplitude de vibração e DA é o incremento de amplitude) apresenta independência em frequência, ou seja, mantém-se constante com a variação da frequência. Por outro lado, o mesmo decresce com o aumento da intensidade. Este último aspecto indica que a discriminação de intensidade é maior quando o estímulo possui maior amplitude.

PARÂMETROS TEMPORAIS

FREQÜÊNCIA

Quanto à frequência do estímulo, diversos estudos mostram que o tato tem resposta diferenciada para a discriminação em frequência com a variação da mesma. De acordo com VERRILO&GESCHIEDER (1992), quando um dispositivo com área maior que 0,02 cm² é utilizado para produzir vibrações táteis aplicadas na ponta dos dedos, o limiar de sensação é independente da frequência em baixas frequências (25 a 40 Hz), depois passa a variar, atingindo uma maior sensibilidade em torno de 250 Hz. Até 1000 Hz, a sensibilidade decresce rapidamente, e limiares válidos não podem ser obtidos acima desta frequência. Este comportamento pode ser observado na Figura 58. O estimulador vibrotátil implementado no trabalho aqui apresentado possui uma área de contato em torno de 0,4 cm², ou seja, o limiar de sensação das vibrações percebidas é menor para frequências em torno de 250 Hz.

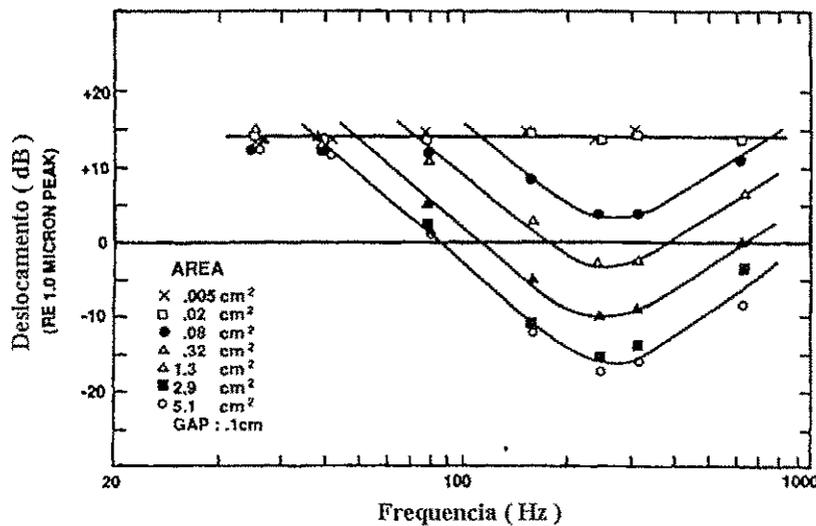


Figura 58 - Limiares de detecção de estímulos vibrotáteis (Verrillo, 1963)

DISCRIMINAÇÃO EM FREQUÊNCIA

A habilidade para detectar-se estímulos para diferentes frequências (limiar de sensação) é importante, mas representa apenas uma parte do problema. De igual importância é a habilidade para discriminar-se frequências entre frequências (discriminação em frequência – JND), quando o estímulo é apresentado com intensidade acima do limiar de sensação. MOWBRAW e GEBHARD (1957) mostraram que a discriminação é boa em baixas frequências, mas deteriora-se rapidamente com o aumento da frequência. Além disto, foi observado que existe uma melhor discriminação em frequência quando são utilizados pulsos na excitação do que quando são usados sinais senoidais. Através dos resultados mostrados na Figura 59 pode-se observar que os pulsos produzem uma melhor discriminação em frequências que os sinais senoidais, e que a discriminação em frequências é melhor em baixas que em altas frequências. As pesquisas mostram que o tato, quando comparado ao sentido auditivo, é pobre em discriminação em frequências, e a mesma diminui com o aumento da frequência.

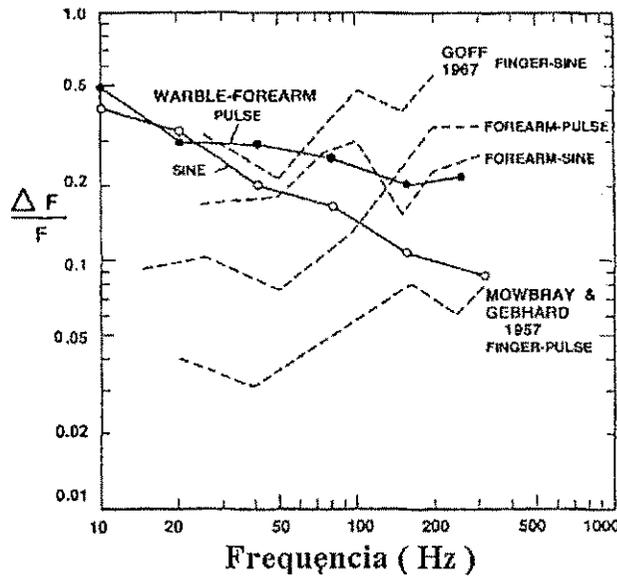


FIGURA 59 - LIMIARES DE DISCRIMINAÇÃO EM FREQUÊNCIA (ROTHENBERG ET ALII, 1977)

DURAÇÃO DO ESTÍMULO

Na figura 60 pode-se observar os limiares de percepção obtidos para duração do sinal, tendo-se como referência um estímulo aplicado com duração de 1000 ms. Quando a área do contato é grande (2,9cm²), um sinal de pequena duração é mais difícil de detectar-se, e o aumento da capacidade de detecção aumenta com a duração do estímulo.

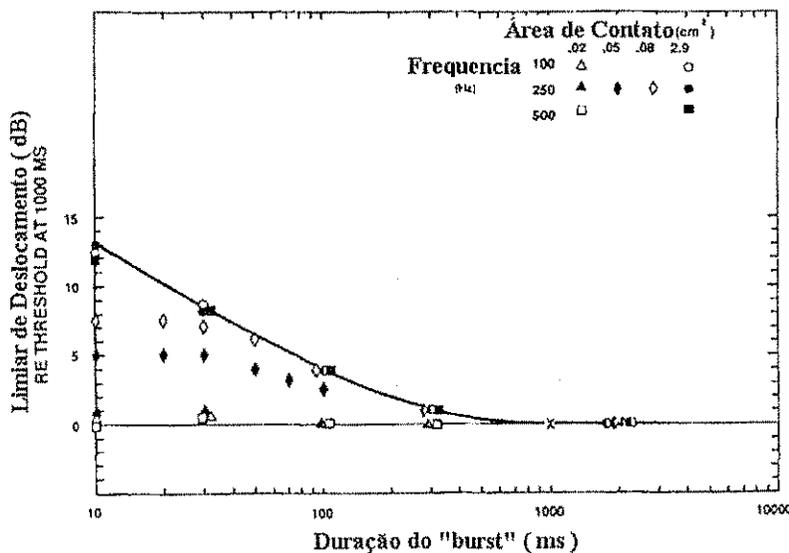


FIGURA 60 - LIMIARES DE DETECÇÃO VIBROTÁTIL PLOTADOS EM FUNÇÃO DA DURAÇÃO DO ESTÍMULO (VERRILO, 1965)

PARÂMETROS ESPACIAIS

DIMENSÃO DO CONTATO

De acordo com a Figura 61, em frequências altas (80 – 320 Hz), a sensibilidade aumenta diretamente com o aumento da superfície vibratória. Em baixas frequências (abaixo de 40 Hz), a dimensão do contato não tem efeito no limiar de detecção.

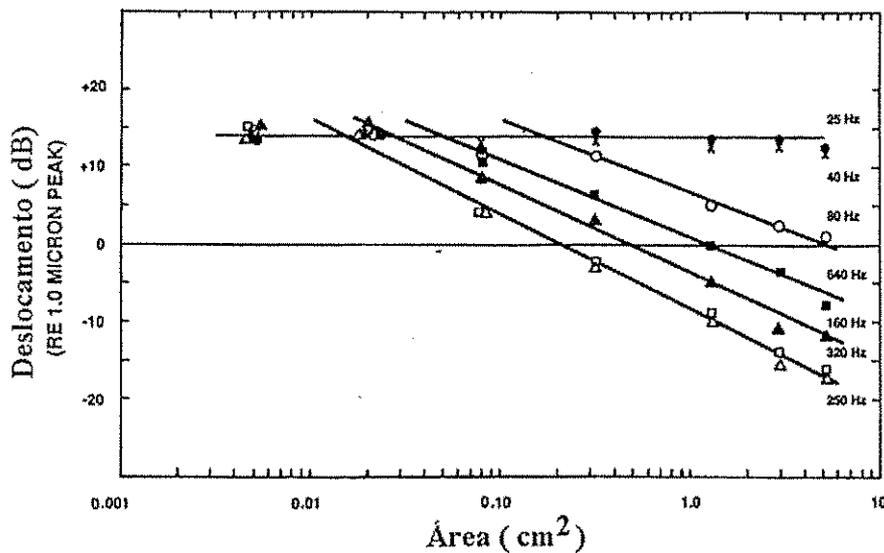


FIGURA 61 - : LIMIARES DE DETECÇÃO VIBROTÁTIL PARA DIFERENTES FREQUÊNCIAS, EM FUNÇÃO DO TAMANHO DA ÁREA DE CONTATO.

LOCAL DO CORPO

O estudo da sensação tátil em diferentes partes do corpo é importante porque as diferenças na percepção e sensibilidade dos limiares definem as possibilidades de desenvolvimento de dispositivos para partes distintas do corpo, de acordo com o uso do estimulador. Na figura 62, vemos um gráfico com uma comparação entre os resultados de limiares de sensação para sinais senoidais aplicados no ante-braço (▲), na palma da mão (●) e na ponta do dedo médio(O). A dimensão do contato é de 0,28 cm². A maior sensibilidade situa-se no dedo médio e a menor no ante braço.

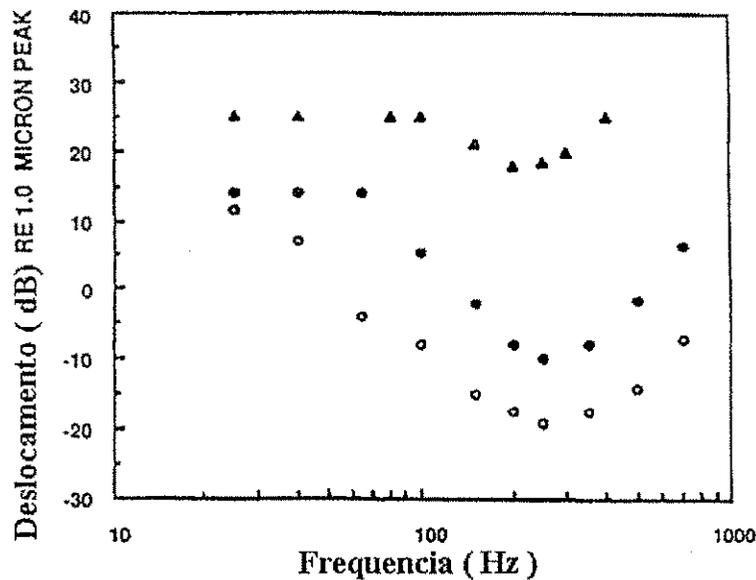


FIGURA 62 - COMPARAÇÃO DOS LIMIARES DE DETECÇÃO EM FUNÇÃO DA FREQUÊNCIA APLICADA NO ANTE-BRAÇO, NA PALMA DA MÃO E NA PONTA DO DEDO MÉDIO.

O projetista de um dispositivo para comunicação tátil que utiliza múltiplos pontos de estimulação deve estar alerta para a possibilidade de os estímulos aplicados por um estimulador poderem degradar a resposta ao estímulo aplicado em outro local do corpo.

PARÂMETRO DE INTENSIDADE

INTENSIDADE SUBJETIVA

Em função do estímulo ser aplicado normalmente acima dos limiares de sensação, a informação contida na Figura 63 mostra ao projetista de dispositivos vibrotáteis as diferenças na intensidade subjetiva produzidas por vibrações acima do limiar de detecção e de igual amplitude, que serão importantes na concepção de um projeto de estimulador vibrotátil. Estes dados estão dispostos nas curvas de sensação de magnitudes plotadas abaixo, em função da frequência. São curvas de intensidades subjetivas iguais. A partir destas curvas pode ser determinado qual a curva de intensidade de vibração necessária para manter a mesma sensação subjetiva de intensidade em qualquer frequência.

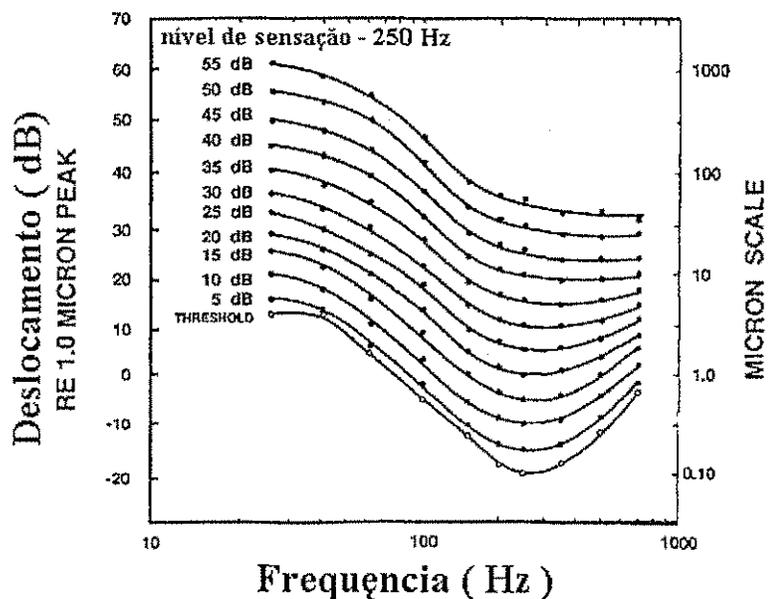


FIGURA 63 - CURVAS DE MESMA SENSÇÃO DE MAGNITUDE, EM FUNÇÃO DA FREQUÊNCIA DE ESTIMULAÇÃO (VERRILO ET ALI, 1969)

VARIÁVEIS SUBJETIVAS

GÊNERO

Na figura 64 pode-se observar os limiares vibrotáteis medidos em homens (●) e mulheres (○). A barra vertical representa o desvio padrão. Não há diferença aparente nos dados coletados.

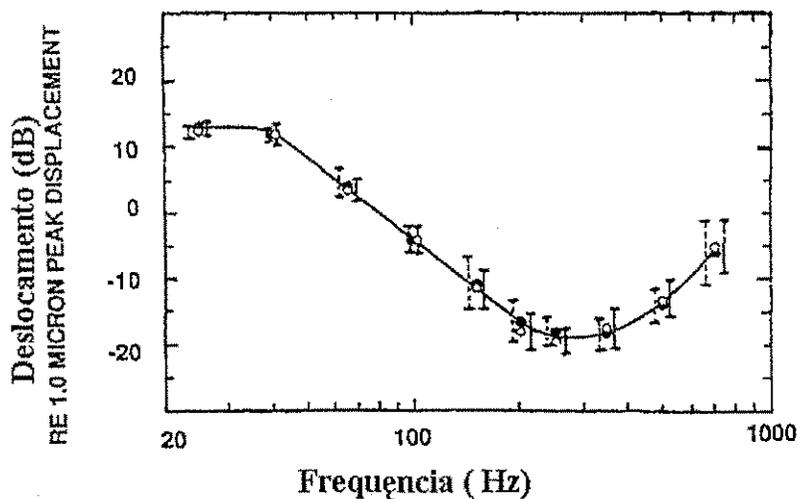


FIGURA 64 - LIMIARES VIBROTÁTEIS MEDIDOS EM HOMENS E MULHERES (VERRILO, 1979)

IDADE

Na figura 65 pode-se observar os limiares de sensação vibrotátil medidos para cinco grupos de idades: 10, 20, 35, 50, 65. No gráfico, as curvas de menor sensibilidade (mais alta) são das maiores idades, ou seja, a sensibilidade vibrotátil decresce com a idade.

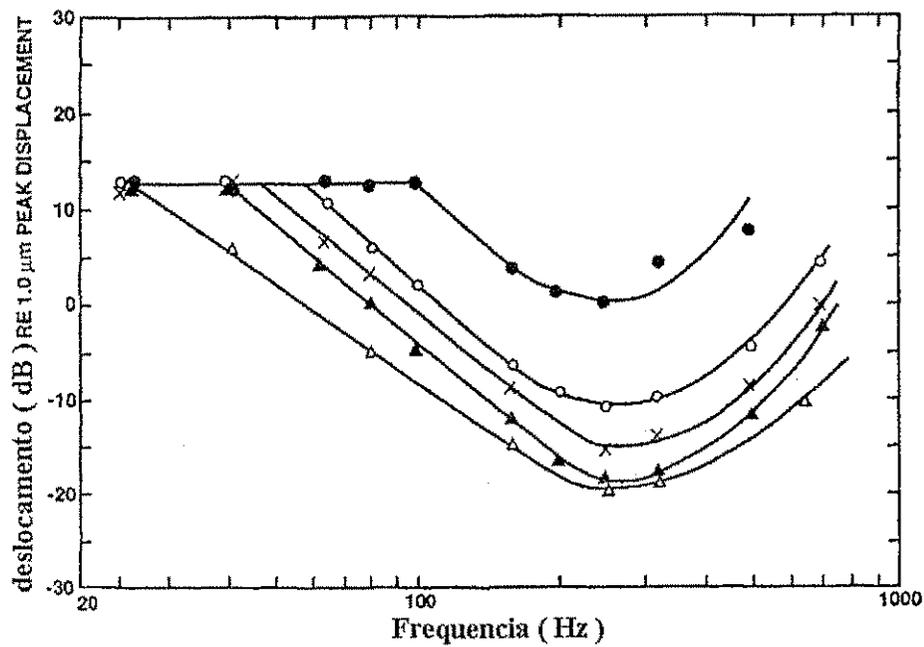


FIGURA 65 - LIMIARES DE DETECÇÃO VIBROTÁTEIS MEDIDOS EM CINCO GRUPOS (VERRILO, 1979)