



**SUSANA MARIA
CAPITÃO DA SILVA**

**Traços Acústicos e Perceptivos de Sons Não
Verbais e da Fala**



**SUSANA MARIA
CAPITÃO DA SILVA**

**Traços Acústicos e Perceptivos de Sons Não
Verbais e da Fala**

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Ciências da Fala e da Audição, realizada sob a orientação científica do Prof. Dr. Luís Miguel Teixeira de Jesus, Professor Adjunto da Escola Superior de Saúde da Universidade de Aveiro.

o júri

Presidente

Prof. Dr. Nelson Fernando Pacheco da Rocha
Professor Catedrático da Universidade de Aveiro

Vogais

Prof. Dr. João Manuel Pires da Silva e Almeida Veloso
Professor Auxiliar da Universidade do Porto

Prof. Dr. Luís Miguel Teixeira de Jesus
Professor Adjunto da Escola Superior de Saúde da Universidade de Aveiro

agradecimentos

a todos os docentes do Mestrado em Ciências da Fala e da Audição que me permitiram adquirir conhecimentos em áreas diversas, e contribuíram para me despertar a curiosidade pela investigação,

ao Prof. Dr. Luís Jesus, pelo seu dinamismo, dedicação e entusiasmo, por assumir o seu compromisso na orientação deste projecto muito para além da realização da Dissertação de Mestrado,

a todas as pessoas que colaboraram neste projecto como informantes, pela sua disponibilidade e altruísmo,

a todos os meus alunos surdos, que me inspiraram para realizar este projecto,

ao André, à Helena e à Marisa, pelo companheirismo e pelas gargalhadas mesmo nos momentos de maior cansaço,

à Isabel, por continuar a ser mais do que professora, pela sua amizade e compreensão,

aos meus Pais e ao meu irmão Ricardo, por estarem sempre presentes e por todo o apoio incondicional,

ao Mário, pelo carinho, pela dedicação, e por percorrer este caminho sempre ao meu lado.

palavras-chave

Sons Não Verbais, Análise Acústica, Percepção, Produção de Onomatopeias.

resumo

Foram analisados parâmetros acústicos extraídos da forma de onda, espectrograma e espectro de estímulos de sons não verbais. Realizou-se uma experiência de percepção em que 8 falantes do Português Europeu produziram onomatopeias para descrever os sons não verbais, analisando-se as características acústicas dos fones utilizados e dos respectivos sons não verbais. Foram também efectuadas duas experiências de percepção em que os sujeitos procuraram identificar os sons não verbais manipulados com filtros passa-banda e com ruído em diferentes gamas de frequências. Os resultados obtidos indicam que as características acústicas dos sons não verbais estão relacionadas com o material que os produz e com o tipo de interacção entre os materiais. Foram também observadas correlações entre as características acústicas dos sons não verbais e as dos sons da fala utilizados para os descrever.

keywords

Environmental Sounds, Acoustic Analysis, Perception, Onomatopoea Production.

abstract

In this study the features of environmental sounds were analysed using parameters extracted from the acoustic signal waveforms, spectrograms and spectra. We also ran a perceptual experiment involving 8 Portuguese European speakers that produced onomatopoeas to describe the environmental sounds. Speech sounds used in onomatopoeas and the environmental sounds were then analysed. Two perception experiments were also developed, in which subjects had to identify band-pass filtered and noisy environmental sound stimuli. The results suggest that the acoustic features of environmental sounds are related to the materials and interactions involved in the sound-producing event. There is also evidence showing that speech and environmental sounds have related acoustic features.

ÍNDICE

LISTA DE FIGURAS	5
LISTA DE TABELAS	11
CAPÍTULO 1: INTRODUÇÃO	13
1.1 MOTIVAÇÕES E OBJECTIVOS DO ESTUDO	13
1.2 SONS NÃO VERBAIS.....	14
1.3 A PERCEPÇÃO AUDITIVA	15
1.3.1 <i>O Processamento Auditivo Central</i>	15
1.3.2 <i>Competências Auditivas: Aquisição e Desenvolvimento</i>	16
1.4 ORGANIZAÇÃO DA DISSERTAÇÃO	18
CAPÍTULO 2: SELECÇÃO E CLASSIFICAÇÃO DOS ESTÍMULOS	19
2.1 INTRODUÇÃO.....	19
2.2 REVISÃO DA LITERATURA	19
2.2.1 <i>Sons Não Verbais</i>	19
2.2.2 <i>A Classificação de Gaver (1993) dos Sons Ambientais</i>	21
2.3 SELECÇÃO DOS ESTÍMULOS.....	24
2.4 CLASSIFICAÇÃO DOS ESTÍMULOS	26
2.4.1 <i>Sólidos</i>	29
2.4.2 <i>Líquidos</i>	30
2.4.3 <i>Gases</i>	31
2.4.4 <i>Sons de Animais</i>	31
2.5 SUMÁRIO	32
CAPÍTULO 3: ANÁLISE ACÚSTICA DOS ESTÍMULOS	33
3.1 INTRODUÇÃO.....	33
3.2 REVISÃO DA LITERATURA	33
3.3 PROCEDIMENTOS.....	41
3.4 RESULTADOS	44
3.4.1 <i>Análise Temporal</i>	44

3.4.2 <i>Análise da Forma de Onda e Análise Espectral</i>	46
3.4.2.1 Sólidos.....	46
3.4.2.2 Líquidos.....	48
3.4.2.3 Gases.....	50
3.4.2.4 Animais.....	54
3.5 SUMÁRIO	56
CAPÍTULO 4: PERCEÇÃO DE SONS NÃO VERBAIS E PRODUÇÃO DE ONOMATOPEIAS ...	59
4.1 INTRODUÇÃO.....	59
4.2 SONS DA FALA.....	59
4.3 SONS DA FALA E SONS NÃO VERBAIS NA AVALIAÇÃO AUDITIVA	65
4.4 ONOMATOPEIAS	67
4.5 PROCEDIMENTOS.....	69
4.6 RESULTADOS	70
4.6.1 <i>Sólidos</i>	71
4.6.2 <i>Líquidos</i>	74
4.6.3 <i>Gases</i>	78
4.6.4 <i>Vocalizações de Animais</i>	81
4.6.5 <i>Comparação de Categorias</i>	83
4.7 SUMÁRIO	84
CAPÍTULO 5: PERCEÇÃO DE SONS NÃO VERBAIS FILTRADOS E COM RUÍDO	87
5.1 INTRODUÇÃO.....	87
5.2 REVISÃO DA LITERATURA	87
5.3 MÉTODO	93
5.3.1 <i>Percepção de Sons Não Verbais Filtrados</i>	93
5.3.2 <i>Percepção de Sons Não Verbais com Ruído</i>	96
5.4 RESULTADOS	99
5.4.1 <i>Sólidos</i>	99
5.4.2 <i>Líquidos</i>	101
5.4.3 <i>Gases</i>	103
5.4.4 <i>Sons de Animais</i>	106
5.5 SUMÁRIO	108
CAPÍTULO 6: CONCLUSÕES E TRABALHO FUTURO	111
6.1 INTRODUÇÃO.....	111
6.2 CONCLUSÕES.....	112
6.3 TRABALHO FUTURO	114
BIBLIOGRAFIA.....	115

PUBLICAÇÕES NO ÂMBITO DO MESTRADO 117

APRESENTAÇÕES ORAIS E EM CARTAZ NO ÂMBITO DO MESTRADO..... 121

ANEXOS 122

LISTA DE FIGURAS

Figura 1: O contínuo do contexto para a experiência. A fonte de um evento gera ondas sonoras: algumas são irradiadas directamente para um ponto de observação; outras são modificadas pelo contexto antes de chegarem ao ouvinte. O sinal à entrada do sistema auditivo, fornece informação sobre a experiência da “ everyday listening” . De Gaver (1993).....	21
Figura 2: Descrição hierárquica de eventos sonoros simples. De Gaver (1993).....	22
Figura 3: Um “ mapa” de sons não verbais adaptado de Gaver (1993, p. 24). Podem observar-se três fontes fundamentais (sólidos em vibração, líquidos e gases), sobrepondo-se nalgumas zonas. Dentro de cada secção, vemos os eventos de nível básico a negrito, e os seus atributos relevantes aparecem por baixo deles em itálico. A complexidade aumenta no sentido do centro da figura, desde eventos de nível básico, eventos que cariam no padrão temporal, eventos compostos e eventos e híbridos.	24
Figura 4: Ilustração dos eventos mecânicos: a) garrafa a balouçar, b) garrafa a partir. De Warren e Verbrugge (1984).....	34
Figura 5: Espectrogramas de: a) garrafa a balouçar, b) garrafa a partir. De Warren e Verbrugge (1984).	34
Figura 6: Espectro para cada palma dos 20 sujeitos. Cada espectro é a média aritmética dos espectros de cada palma isolada (das 10 <i>palmas</i>) obtida para os primeiros 10 ms após o início da palma. De Repp (1987).....	35
Figura 7: Forma de onda de um passo de um (a) homem e (b)mulher. De Xiaofeng et al. (1991). 36	
Figura 8: Espectros médios para os 16 andantes. As letras M e F representam o género dos sujeitos. O número de cada espectro representa a ordem de percepção do género. No eixo das abcissas está representado o intervalo linear das frequências até 5 kHz, e no eixo das ordenadas o intervalo linear da intensidade. De Xiaofeng et al. (1991).....	36
Figura 9: Forma de onda idealizada em que estão representados os períodos transitórios e o período de estabilidade. De Henrique (1997).	37

Figura 10: Representação tridimensional de um som onde são definidos os planos dinâmico, harmónico e melódico, e se pode observar a evolução temporal de cada parcial. De Meyer (1978), adaptado para o Português por Henrique (1997).....	38
Figura 11: Em cima som da caixa de rufo: A) representação temporal, B) representação espectral. Em baixo som do bombo: C) representação temporal, D) representação espectral. De Henrique (1997).....	38
Figura 12: Som de pratos <i>Zildjar</i> . A) representação temporal, B) representação espectral. De Henrique (1997).....	39
Figura 13: Forma de onda e o respectivo espectrograma de uma chamada de contacto feminina mostrando a frequência fundamental (F0) e harmónicos, e as primeiras quatro ressonâncias (F1 a F4). De Gargstang (2004).....	39
Figura 14: Grunhido emitido por uma fêmea adulta durante a perseguição (a). Fêmea adulta a berrar com um sub-adulto macho (b). De Range e Fischer (2004).....	40
Figura 15: Espectrograma da produção vocal dum mocho-real (A) e parâmetros medidos para descrever a produção vocal (B). De Lengagne (2001).....	40
Figura 16: Procedimentos utilizados no programa <i>Praat</i> para obter os valores de duração, F0 médio, F0 mínimo e F0 máximo dos estímulos.....	41
Figura 17: Procedimentos utilizados no programa <i>Praat</i> para obter a forma de onda dos estímulos.	42
Figura 18: Procedimentos utilizados no programa <i>Praat</i> para obter o espectrograma dos estímulos.....	43
Figura 19: Procedimentos utilizados no programa <i>Praat</i> para obter os espectros da densidade de potência dos estímulos.....	44
Figura 20: Figura da duração dos estímulos do corpus de sons não verbais. A vermelho estão os estímulos da categoria Sólido- <i>Impacto</i> (I), a laranja os estímulos Sólidos- <i>Deformação</i> (D), a amarelo os estímulos Sólidos- <i>Raspar</i> (R), a azul os estímulos do grupo dos Líquidos- <i>Pingar</i> (P), a azul turquesa os estímulos de Líquidos-” splash” (S), a branco o estímulo Líquido- <i>Ondular</i> (O) e a cor de rosa os estímulos do grupo dos Animais. Os números no eixo dos XXs correspondem ao número do estímulo.....	45
Figura 21: Forma de onda, espectrograma (à esquerda) e espectro (à direita), do som <i>martelar</i> (nº 6), pertencente ao grupo <i>sólidos-impacto</i>	46
Figura 22: Forma de onda, espectrograma (à esquerda) e espectro (à direita), do som <i>tambor</i> (nº 10), pertencente ao grupo <i>sólidos-impacto</i>	47

Figura 23: Forma de onda, espectrograma (à esquerda) e espectro (à direita), do som <i>serrar</i> (nº 16), pertencente ao grupo <i>sólidos-raspar</i> .	47
Figura 24: Forma de onda, espectrograma e espectro do som <i>encher um copo de água</i> (nº 18), pertencente ao grupo <i>líquidos-pingar</i> .	48
Figura 25: Forma de onda, espectrograma e espectro do som <i>mar</i> (nº 22), do grupo <i>líquidos- " splash"</i> .	49
Figura 26: Forma de onda, espectrograma e espectro do som <i>chuva</i> (nº 19), pertencente ao grupo <i>líquidos- pingar e " splash"</i> .	50
Figura 27: Forma de onda, espectrograma e espectro do som <i>rio</i> (nº 24), pertencente ao grupo <i>líquidos-ondular</i> .	50
Figura 28: Forma de onda, espectrograma e espectro do som <i>trovão</i> (nº 25), pertencente ao grupo <i>gases-explosão</i> .	51
Figura 29: Forma de onda, espectrograma e espectro do som <i>encher um pneu</i> (nº 26), pertencente ao grupo <i>gases- " gust"</i> .	52
Figura 30: Forma de onda, espectrograma e espectro do som <i>avião</i> (nº 27), pertencente ao grupo <i>gases- " gust"</i> .	52
Figura 31: Forma de onda, espectrograma e espectro do som <i>vento</i> (nº 29), pertencente ao grupo <i>gases-vento</i> .	53
Figura 32: Forma de onda e espectrograma do som <i>assobiar</i> (nº28), pertencente ao grupo <i>gases-vento</i> .	53
Figura 33: Forma de onda e espectrograma do som <i>vaca</i> (nº 30), pertencente ao grupo animais.	54
Figura 34: Forma de onda e espectrograma do som <i>pássaro</i> (nº 36), pertencente ao grupo animais.	55
Figura 35: Forma de onda e espectrograma do som <i>grilo</i> (nº 35), pertencente ao grupo animais.	55
Figura 36: Valores de F0 médio, mínimo e máximo dos estímulos produzidos por animais.	56
Figura 37: Espaço de algumas vogais do Inglês Americano com representação acústica tradicional. De Borden et al. (2003).	60
Figura 38: Espectrogramas das vogais /i, I, ε , æ, α , ɔ , ʊ , u/ de um falante de Inglês Americano. As setas indicam o valor das primeiras três formantes. De Ladefoged (2006, p.186).	61
Figura 39: Espaço das vogais do PE. De Cruz-Ferreira (1999).	62
Figura 40: Espectrogramas e padrões de F1-F2 obtidos para as palavras [bai], [bo i] e [bau]. De Kent e Read (2002).	62

Figura 41: Forma de onda e espectrogramas para as produções isoladas das fricativas estridentes. De Kent e Read (2002).	65
Figura 42: Formas de onda e espectrogramas para produções isoladas de fricativas não estridentes. De Kent e Read (2002).	65
Figura 43: Representação de sons não verbais num audiograma: do lado esquerdo de Monfort (2001), do lado direito de Roeser (2004).	67
Figura 44: Espectrograma do estímulo “ pardal” incluído no Grupo 1. De Tanaka et al. (Tanaka et al. 2006).	68
Figura 45: Distribuição dos parâmetros fonéticos pelo tipo de interacção entre sólidos.	71
Figura 46: Forma de onda, espectrograma e espectro da onomatopeia [tũ tũ] produzida pelo informante MS ao descrever o estímulo “ martelar” (nº 6).	72
Figura 47: Forma de onda, espectrograma e espectro da onomatopeia [pápã] produzida pelo informante RS ao descrever o estímulo <i>tambor</i> (nº 10).	73
Figura 48: Forma de onda, espectrograma e espectro da onomatopeia [tʃ iktʃ ktʃ i] produzida pelo informante BR ao descrever o estímulo <i>esfregar as mãos</i> (nº 15).	74
Figura 49: Distribuição dos parâmetros fonéticos pelo tipo de interacção entre líquidos.	75
Figura 50: Forma de onda, espectrograma e espectro da onomatopeia [tuf t tuf t tuf t] produzida pelo informante MR ao descrever o estímulo <i>encher um copo de água</i> (nº 18).	75
Figura 51: Forma de onda, espectrograma e espectro da onomatopeia [tʃ :] produzida pelo informante MR ao descrever o estímulo <i>chuva</i> (nº 19).	76
Figura 52: Forma de onda e espectrograma da onomatopeia [ua:] produzida pelo informante RS ao descrever o estímulo <i>mar</i> (nº 22).	77
Figura 53: Forma de onda, espectrograma e espectro da onomatopeia [fliflifi] produzida pelo informante TM ao descrever o estímulo <i>rio</i> (nº 24).	77
Figura 54: Distribuição dos parâmetros fonéticos pelo tipo de interacção entre gases.	78
Figura 55: Forma de onda, espectrograma e espectro da onomatopeia [ɤ : õ :] produzida pelo informante RS ao descrever o estímulo <i>água a ferver</i> (nº 23).	79
Figura 56: Forma de onda, espectrograma e espectro da onomatopeia [ʃ tʃ tʃ t] produzida pelo informante RS ao descrever o estímulo <i>encher um pneu</i> (nº 26).	80
Figura 57: Forma de onda e espectrograma da onomatopeia [u:] produzida pelo informante AM ao descrever o estímulo <i>vento</i> (nº 29).	80
Figura 58: Distribuição dos parâmetros fonéticos pela classe animal.	81

Figura 59: Forma de onda e espectrograma da onomatopeia [mu:] produzida pelo informante TM ao descrever o estímulo <i>vaca</i> (nº 30).	82
Figura 60: Forma de onda e espectrograma da onomatopeia [piupiupiu] produzida pelo informante MS ao descrever o estímulo <i>pássaro</i> (nº 36).	82
Figura 61: Forma de onda e espectrograma da onomatopeia [gr igr igr i] produzida pelo informante AM ao descrever o estímulo <i>grilo</i> (nº 35).	83
Figura 62: Distribuição dos parâmetros fonéticos pela categoria de sons não verbais.	84
Figura 63: Limiares obtidos na experiência de Myers et al. (1996).	89
Figura 64: Valores de detecção e identificação no ruído de 25 sons não verbais (onde Ev/N significa Event to Noise, traduzindo relação Sinal / Ruído). De Gygi et al. (1997).	89
Figura 65: Resultados da identificação de sons não verbais com filtros passa-alto. De Gygi (2001).	90
Figura 66: Resultados da identificação de sons não verbais com filtros passa-baixo. De Gygi (2001).	90
Figura 67: Os oito grupos obtidos com base nos resultados de identificação de 70 sons não verbais. De Gygi (2001).	92
Figura 68: Exemplo de espectrogramas de um som não verbal filtrado, neste caso o estímulo <i>avião</i> . Em cima o estímulo original, e em baixo os estímulos obtidos após a aplicação dos filtros passa-banda.	95
Figura 69: Imagens do programa <i>Praat</i> , na realização da experiência de percepção de sons não verbais filtrados, para o grupo líquidos. Observa-se em 1.1 a primeira imagem que surge no ecrã, para que se inicie a experiência; a imagem 1.2 corresponde ao grupo dos líquidos, com as opções de resposta; a imagem 1.3 é a que surge a cada 20 estímulos, para que o sujeito faça uma pausa se necessitar; e a imagem 1.4 a correspondente ao final da experiência.	96
Figura 70: Exemplo de espectrogramas de um som não verbal, neste caso do estímulo <i>avião</i> , com ruído sobreposto a -15 dB de relação S/N.	98
Figura 71: Imagem do Programa <i>Praat</i> , na realização da experiência 3.	99
Figura 72: Distribuição das respostas de identificação dos sons filtrados de <i>esfregar as mãos</i> . .	100
Figura 73: Distribuição das respostas de identificação dos sons filtrados de <i>maracas</i>	100
Figura 74: Distribuição das respostas de identificação dos sons filtrados de <i>tambor</i>	100
Figura 75: Distribuição das respostas de identificação dos sons de <i>tambor</i> com ruído.	101

Figura 76: Distribuição das respostas de identificação dos sons filtrados de <i>encher um copo de água</i>	101
Figura 77: Distribuição das respostas de identificação dos sons filtrados de <i>chuva</i>	102
Figura 78: Distribuição das respostas de identificação dos sons de <i>chuva</i> com ruído	102
Figura 79: Distribuição das respostas de identificação dos sons filtrados de <i>mar</i>	102
Figura 80: Distribuição das respostas de identificação dos sons filtrados de <i>rio</i>	103
Figura 81: Distribuição das respostas de identificação dos sons de <i>rio</i> com ruído	103
Figura 82: Distribuição das respostas de identificação dos sons filtrados de <i>avião</i>	104
Figura 83: Distribuição das respostas de identificação dos sons de <i>avião</i> com ruído	104
Figura 84: Distribuição das respostas de identificação dos sons filtrados de <i>vento</i>	105
Figura 85: Distribuição das respostas de identificação dos sons de <i>vento</i> com ruído	105
Figura 86: Distribuição das respostas de identificação dos sons filtrados de <i>trovoada</i>	105
Figura 87: Distribuição das respostas de identificação dos sons filtrados de <i>grilo</i>	106
Figura 88: Distribuição das respostas de identificação dos sons <i>grilo</i> com ruído	106
Figura 89: Distribuição das respostas de identificação dos sons filtrados de <i>pássaro</i>	107
Figura 90: Distribuição das respostas de identificação dos sons de <i>pássaro</i> com ruído	107
Figura 91: Distribuição das respostas de identificação dos sons filtrados de <i>vaca</i>	107
Figura 92: Distribuição das respostas de identificação dos sons de <i>vaca</i> com ruído	108

LISTA DE TABELAS

Tabela 1: Efeitos acústicos dos atributos das fontes sonoras. De Gaver (1993).	22
Tabela 2: Efeitos frequenciais e temporais. De Gaver (1993).	23
Tabela 3: Estímulos incluídos no estudo. Os sons com * correspondem a sons de instrumentos musicais. O código da primeira coluna indica a origem do estímulo: “ DQF” – corresponde ao jogo “ Digo o que Faço, Faço o que Digo” , “ GS1” – corresponde ao jogo “ Ginásio dos Sons 1” , “ GS2” , pertence ao jogo “ Ginásio dos Sons 2”). Os últimos 3 algarismos referem-se à ordem de aparecimento do estímulo no respectivo jogo.	28
Tabela 4: Exemplos de palavras com as vogais orais e nasais do PE. De Cruz-Ferreira (1999)...	61
Tabela 5: Exemplos de palavras com os ditongos orais e nasais do PE. De Cruz-Ferreira (1999).	63
Tabela 6: Consoantes do Português Europeu e exemplos de palavras onde ocorrem. De Cruz-Ferreira (1999).	63
Tabela 7: Classificação das consoantes do Português Europeu. De Cruz-Ferreira (1999).	63
Tabela 8: Valores das frequências de corte dos filtros passa-banda. De Gygi (2001).	91
Tabela 9: Corpus de sons não verbais utilizados na experiência de percepção em que foram utilizados filtros.	93
Tabela 10: Valores das frequências de corte dos filtros passa-banda utilizados na experiência de percepção de sons não verbais filtrados.	94
Tabela 11: Corpus de sons não verbais utilizados na experiência de percepção em que foram utilizados ruídos sobrepostos.	97
Tabela 12: Valores das frequências de corte dos ruídos aplicados na experiência de percepção de sons não verbais com ruído.	97

CAPÍTULO 1: INTRODUÇÃO

1.1 MOTIVAÇÕES E OBJECTIVOS DO ESTUDO

Perceber como funciona a audição humana tem sido o interesse de muitos investigadores de áreas variadas, como a medicina, a engenharia, a linguística e a fonética. São inúmeros os estudos na área da percepção de fala, procuram saber quais as características dos estímulos sonoros que transmitem informações ao nosso sistema auditivo e lhe possibilitam a compreensão do significado da fala, e que respostas desencadeiam no ouvido externo e interno.

No exercício profissional como Terapeuta da Fala a autora intervém junto de crianças e jovens com deficiência auditiva, que utilizam próteses auditivas ou implantes cocleares. Assim, compreender como se desenvolve a audição e como podem ser melhoradas as competências auditivas desta população são do extremo interesse para os terapeutas da fala. Mas, enquanto que a fala tem sido já amplamente estudada e caracterizada em termos da sua percepção, o mesmo não se verifica com os sons não verbais.

O Terapeuta da Fala intervém junto de populações com dificuldades de comunicação, linguagem e fala. A audição e a fala estão intimamente relacionadas, pelo que é pertinente tentar perceber e conhecer as competências auditivas das populações, e como as mesmas influenciam a sua aquisição e desenvolvimento da fala. Existem diferentes materiais didácticos que possibilitam a utilização de vários tipos de estímulos sonoros, desde sons não verbais a sons da fala. Procurou-se pois perceber que traços acústicos os caracterizam, e que propriedades em comum possuem a fala e os restantes tipos de sons, mais frequentes na actividade clínica dos terapeutas da fala.

A capacidade de ouvir e perceber aquilo que a nossa audição nos transmite é essencial para o desenvolvimento do ser humano. Ainda durante a gestação, o feto reage a estímulos sonoros, começando precocemente a interagir com a mãe. A audição será bastante importante para a socialização e comunicação do bebé com os seus pais, familiares e amigos. Os primeiros sons a que o bebé consegue atribuir significado são os sons do seu meio ambiente, sendo mais acessíveis os não verbais, do que os sons da fala que possuem uma semântica mais complexa e menos icónica. Então, para conseguirmos avaliar e estimular a audição de uma criança pequena,

devemos utilizar não apenas estímulos de fala, mas também estímulos não verbais (Monfort 2001).

Este projecto teve como principal objectivo a caracterização acústica de sons não verbais, tendo por base uma análise acústica e a realização de experiências de percepção. Outro objectivo foi comparar os sons não verbais com os sons da fala, através do estudo das características acústicas de ambos e da realização de uma experiência de produção de onomatopeias para descrever os sons não verbais incluídos na amostra. Quando o Terapeuta da Fala tem como objectivo de intervenção o desenvolvimento de competências auditivas, o facto de os graus e tipos de surdez poderem ser bastante variados de indivíduo para indivíduo, faz com que seja importante a utilização de estímulos sonoros adequados às características de cada um, para que não sejam exigidas tarefas impossíveis, tais como ouvir sons que não possuem parâmetros acústicos dentro da gama de frequências e intensidade que o indivíduo consegue detectar.

1.2 SONS NÃO VERBAIS

Tal como acontece com o normo-ouvinte é essencial que o ouvido de crianças e adultos deficientes auditivos seja estimulado desde cedo com estímulos de fala e com sons naturais que existem no meio ambiente. Os sons não verbais são utilizados desde fases muito precoces do treino auditivo, nomeadamente na identificação e discriminação de sons (Monfort 2001). A nível muito precoce é aconselhado aos pais que chamem a atenção da criança para as fontes sonoras que a rodeiam, de uma forma informal (Barker 1997). São posteriormente utilizados jogos sonoros didácticos específicos que possuem uma maior variedade de sons e que se utilizam nas sessões de terapia da fala para promover as competências do sujeito na realização de tarefas auditivas (Dumont 1995).

No nosso quotidiano estamos imersos num ambiente sonoro muito variado. A fala é o tipo de estímulo sonoro a que estamos mais atentos, seguida da música, sendo estes os sinais sonoros que têm recebido maior atenção da comunidade científica. Os restantes estímulos são muitas vezes referidos como som ambiente. Neste som ambiente os humanos são capazes de identificar e discriminar pequenas diferenças, úteis no seu dia-a-dia, tais como saber a que distância está um carro, ou se o metro que chega é o da sua linha ou não (Gygi 2001).

A fala resulta da produção de sons por um sistema complexo, os quais são articulados e combinados segundo um determinado código linguístico. Por sua vez a música é um som estruturado, organizado, que possui um conteúdo e transmite uma mensagem. Nem na fala, nem na música, a preocupação principal do ouvinte é identificar a fonte sonora. No entanto, quando ouvimos sons não verbais, que não possuem uma estrutura linguística ou estética de base, o nosso objectivo principal é identificar a fonte deste som (Gygi 2001).

A investigação na área dos sons não verbais iniciou-se relativamente recentemente, o que é espelhado na falta de consenso em relação à terminologia utilizada, ao contrário do que acontece com a fala e a música. São variados os termos que se usam para referir sons familiares que ocorrem naturalmente, e que não são de fala nem de música. O termo “environmental sounds” (traduzindo: *sons ambientais*) parece ser o mais comum (Gygi 2001), mas possui algumas limitações, nomeadamente o facto de poder ser muito geral (é difícil definir o ambiente e todos os sons ocorrem num ambiente) e muito específico (associado a locais exteriores, como por exemplo uma floresta). Neste trabalho, os sons acima referidos são designados por *sons não verbais*, porque serão objecto de estudo os sons que partilham a característica de não fazerem parte de palavras, não terem uma estrutura linguística própria, e de poderem ser produzidos por fontes diversas e em contextos ambientais também variados.

Os sons não verbais partilham atributos com a fala: são complexos, possuem formas de onda que variam no tempo, são muito familiares para o ouvinte, e transmitem significados diferentes consoante as propriedades acústicas do som (Gygi 2001).

1.3 A PERCEPÇÃO AUDITIVA

1.3.1 O PROCESSAMENTO AUDITIVO CENTRAL

O Processamento Auditivo Central refere-se à eficácia com que o Sistema Nervoso Central (SNC) utiliza a informação auditiva. De uma maneira geral, tem a ver com o processamento de informação auditiva no SNC e com a actividade neurobiológica envolvida nesse processamento e que origina os potenciais auditivos neurofisiológicos (ASHA 2005). Depois de o sistema nervoso periférico receber o estímulo sonoro, este é enviado ao sistema nervoso central, para as áreas responsáveis pelo processamento auditivo. É pois importante referir as competências de processamento auditivo envolvidas no processo de aprendizagem e desenvolvimento de competências auditivas. Segundo Keith (1997), podem ser descritas da seguinte forma:

- Discriminação: diferenciar sons de frequências, intensidades e durações distintas;
- Localização: localizar a fonte sonora;
- Atenção auditiva: dirigir a atenção para sinais acústicos relevantes e manter a atenção durante um período de tempo adequado;
- Separação da figura do fundo auditivo: identificar a fonte sonora principal num fundo de ruído;
- Discriminação auditiva: distinguir diferentes elementos fonémicos da fala que são acusticamente semelhantes;

- Encerramento auditivo: compreender a palavra ou mensagem total quando falta uma parte;
- Síntese auditiva: juntar fonemas isolados para formar palavras;
- Análise auditiva: identificar os fonemas constituintes de uma palavra ou frase;
- Associação auditiva: relacionar as palavras com o seu significado;
- Memória auditiva: armazenar e evocar estímulos na ordem ou sequência correcta.

Estes mecanismos podem ocorrer tanto para sons não verbais como sons verbais (Musiek 2001). Pereira (1997) sugere que os comportamentos auditivos referidos sejam agrupados segundo o tipo de tarefa auditiva: descodificação, codificação e organização. A descodificação ou gnosia acústica, refere-se aos processos envolvidos na aquisição de conhecimentos pela integração auditiva de eventos sonoros, ou seja, com os processos de transformação dos “inputs” numa forma de memória sensorial e transferência de alguma parte para a memória primária. Podem haver dificuldades ao nível da atenção selectiva para sílabas e de localização sonora. A codificação ou gnosia auditiva integrativa, é constituída pelos processos envolvidos na aquisição de conhecimentos pela integração de informações sensoriais auditivas e das auditivas com outras informações sensoriais. Quando há um défice a este nível pode estar afectada a atenção selectiva para sílabas, palavras e frases. A *organização* ou *gnosia auditiva sequencial temporal*, está relacionada com os processos de ordenação dos eventos sonoros no tempo. Havendo uma alteração nestes processos, pode estar comprometida a memória sequencial, a escuta direccionada ao ouvido direito ou esquerdo e/ou o fechamento (fala filtrada).

1.3.2 COMPETÊNCIAS AUDITIVAS: AQUISIÇÃO E DESENVOLVIMENTO

Os sons não verbais são importantes para o desenvolvimento da audição, já que preparam o ouvido para perceber diferentes frequências, intensidades e durações do sinal acústico. O sistema auditivo deverá ter cada vez maior acuidade para distinguir pequenas diferenças e discriminar sons cada vez mais semelhantes, como são os da fala.

O desenvolvimento da audição passa por diferentes etapas, que são a base dos programas de educação/treino auditivo desenvolvidos junto de crianças com défice auditivo (Dumont 1995, Monfort 2001). A educação auditiva, também denominada por treino auditivo, é uma educação perceptiva que tem como objectivos: despertar o interesse pelo mundo sonoro, desenvolver a vigilância e atenção auditivas, sensibilizar a criança para os ritmos e características dos sons, perceber as semelhanças e as diferenças, melhorar as discriminações, desenvolver as capacidades de reconhecimento, identificar e compreender os sons da língua. Estas competências são adquiridas ao longo de várias etapas (Dumont 1995):

1. Descobrir os sons, as fontes sonoras e os efeitos do ruído;
2. Distinguir som/silêncio;

3. Orientação auditiva;
4. Distinguir som longo/curto;
5. Percepção do ritmo;
6. Distinguir som forte/fraco;
7. Distinguir som grave/agudo;
8. Identificação dos sons;
9. Educação auditiva linguística.

Considerando as fases identificadas por diversos autores (Thompson 1995, Bell 1996, Barker 1997, Rodriguez 2002) é possível agrupá-las em cinco fases principais: *detecção* (1,2), *discriminação* (3,4,5), *identificação* (6,7,8), *reconhecimento* (9) e *compreensão* (9).

A *detecção* sonora é o primeiro passo para aprender a ouvir. Está relacionada com um comportamento de alerta para o som e com a produção de vocalizações. Os sons iniciais a que a criança irá estar alerta são os sons não verbais, evoluindo posteriormente para os sons da fala (Barker 1997). Esta capacidade permite-lhe responder à presença ou ausência de som. Assim, a criança aprende a responder ao som, a prestar atenção ao som, e a não responder quando não há som (Bell 1996). A progressão irá incidir sobre a redução da intensidade e duração do estímulo (Rodriguez 2002).

A fase de *discriminação* consiste em desenvolver a capacidade de perceber semelhanças e diferenças entre dois ou mais estímulos sonoros. A criança aprende a prestar atenção às diferenças entre os sons, ou a responder de forma diferenciada a sons distintos (Bell 1996). A discriminação grosseira é ensinada primeiro à criança porque é mais fácil de aprender. Refere-se à discriminação entre sons não verbais, tais como sons de instrumentos ou sons de transportes. A *discriminação fina* refere-se aos sons da fala, e requer que a criança seja capaz de usar informações apresentadas ao longo de todo o espectro de frequências usadas na fala (Barker 1997). Para que as crianças possam mostrar as suas capacidades de discriminação fina elas deverão possuir algumas competências linguísticas, ou seja relacionar as palavras com os objectos ou pessoas (Thompson 1995). Nesta fase as tarefas de discriminação serão cada vez mais complexas no sentido de passar de oposições máximas para oposições cada vez menores (Rodriguez 2002).

Durante a fase de *identificação*, treina-se a capacidade de atribuir o significado correcto a um estímulo auditivo, ou seja, ser capaz de relacionar o que ouviu com o objecto, animal ou pessoa correspondente (Bell 1996). Esta tarefa será realizada para um conjunto fechado de hipóteses, que podem ser apresentadas na forma de objectos, fotografias, listas de palavras ou frases ou categorias semânticas (Rodriguez 2002).

O *reconhecimento* é realizado apenas com a ajuda de uma pista contextual, ou seja, o interlocutor proporciona alguma informação sobre o tipo de mensagem que vai ser ouvida (uma

situação semi-aberta). A resposta pedida é a repetição da frase ouvida ou a resposta à mesma (Rodriguez 2002).

Por último, a *compreensão* é a capacidade para perceber o significado do discurso através da resposta a questões, cumprimento de ordens, reformulação do que ouviu, participação numa conversa (Bell 1996). Nesta etapa, pressupõe-se que a criança seja capaz de reconhecer um item verbal sem ajuda prévia e em situações acusticamente mais complexas, e.g., conversar em locais ruidosos, falar ao telefone, ouvir televisão, ouvir rádio, ir ao cinema ou ouvir música (Rodriguez 2002).

Sabendo as várias fases que constituem o treino auditivo é de igual modo importante realizar a adequada selecção dos estímulos sonoros a utilizar em cada uma delas. Para tal, é necessário conhecer as características acústicas dos sons não verbais e dos sons da fala utilizados no treino auditivo.

1.4 ORGANIZAÇÃO DA DISSERTAÇÃO

Incluem-se seis capítulos nesta dissertação. No presente capítulo, Introdução, procurou-se referir a pertinência do estudo e enquadrá-lo no âmbito do Mestrado em Ciências da Fala e da Audição e da prática profissional de terapia da fala, através de uma breve revisão teórica. No segundo capítulo é descrita a amostra e o método de selecção dos estímulos, à luz da revisão da literatura relativa ao estudo de sons não verbais. A análise acústica dos sons não verbais é apresentada no Capítulo 3. No quarto capítulo é apresentada a experiência de produção de onomatopeias para descrever a amostra de sons não verbais, analisando-se os resultados com base em estudos de produção de fala e da caracterização acústica dos sons do Português Europeu. No quinto capítulo são descritas duas experiências de percepção, em que numa delas os estímulos da amostra foram filtrados e na outra lhes foi aplicado ruído. No último capítulo os resultados descritos nos capítulos anteriores são relacionados e apresentam-se as conclusões, bem como propostas de trabalho futuro.

CAPÍTULO 2:

SELECÇÃO E CLASSIFICAÇÃO DOS ESTÍMULOS

2.1 INTRODUÇÃO

Neste capítulo faz-se a descrição do corpus seleccionado para a realização das experiências. Na primeira parte é descrita a pertinência do uso de estímulos de sons não verbais na investigação, através da revisão de alguns estudos nesta área. É dado realce ao estudo de Gaver (1993) no qual é proposta uma classificação de sons não verbais, que foi a adoptada nesta dissertação. Na segunda parte é descrita a selecção dos estímulos, e é apresentado o corpus e sua classificação.

2.2 REVISÃO DA LITERATURA

2.2.1 SONS NÃO VERBAIS

Gaver (1993) propõe a distinção entre dois tipos de escuta: *escuta musical* e a *escuta quotidiana*. Esta distinção prende-se com a experiência auditiva em si, não com os estímulos envolvidos. Ao ouvir um quarteto de cordas podemos estar atentos às sensações que estes estímulos provocam (*escuta musical*), ou à identidade dos instrumentos (*escuta quotidiana*), na qual procuramos identificar as fontes sonoras. Na *escuta musical*, as dimensões perceptivas e os atributos a que prestamos atenção estão relacionados com o som em si. Na *escuta quotidiana* estamos atentos a eventos, não a sons, ouvimos as coisas que nos rodeiam, e as dimensões perceptivas e os atributos em questão correspondem ao evento que produz o som, e não ao som em si. Gygi (2001) também defende esta ideia ao dizer que o facto de para os sons não verbais se dar ênfase à identificação da fonte sonora, enquanto que na percepção de fala e de música há uma maior atenção ao valor semântico ou valor expressivo do som em si, resultará em diferentes formas de escutar cada um destes tipos de sons. Resultados experimentais, como os de Remez et al. (1981), mostram que a forma como nós ouvimos os sons depende do que nós esperamos que esse som seja. Se o objectivo é identificar a fonte sonora, o ouvinte poderá ter de centrar-se nas

propriedades espectrais e temporais num curto intervalo de tempo (para o identificar o mais rapidamente possível) ao contrário de recolher informações durante um longo período de tempo, o que é necessário para saber qual a mensagem transmitida pela fala ou pela música (Gygi 2001).

Numa perspectiva antropológica, a capacidade para reconhecer fontes sonoras terá sido desenvolvida antes da percepção de fala e da capacidade para apreciar música. Studdert-Kennedy e House (1972) e Pribram (1976) sugerem que a capacidade de reconhecimento de sons não verbais originou a capacidade de compreender a fala, e que houve uma progressão de “sons” para “sons com significado” para “símbolos”.

Na área da psicolinguística, são utilizados principalmente dois tipos de estímulos: estímulos simples (sons gerados em laboratório, tais como tons puros, cliques ou ruído transitório), cujas propriedades acústicas podem ser facilmente manipuladas e controladas, e estímulos complexos, tais como a fala e a música. Os sons não verbais não têm sido muito utilizados pois existe a ideia generalizada que os resultados de experiências com estes sons não são aplicáveis à percepção da fala e os investigadores encontram dificuldades em caracterizá-los acusticamente (Gygi 2001, Gygi e Shafiro 2006). O estudo de sons não verbais tem, no entanto, aplicações práticas em áreas diversas, nomeadamente medicina, inteligência artificial, controlo de ruído e em projectos de design de contextos sonoros artificiais (Gygi e Shafiro 2006).

É através dos sons não verbais que o ouvinte consegue uma maior consciência situacional. Quando se tratam de indivíduos com défice auditivo estas capacidades estão comprometidas. A incapacidade para identificar as fontes de sons específicos num contexto sonoro pode não só impedir o sujeito de reagir correctamente a um determinado som, muitos deles de aviso de perigo, mas também prejudicar a orientação no contexto geral.

Os sons não verbais são usados em etapas precoces da reabilitação/habilitação auditiva após implante coclear, sendo que este tipo de sons tem sido aplicado de uma maneira geral para o treino auditivo (Dumont 1995, Maxon 1992, Barker 1997, Pereira e Schochat 1997). No entanto, a fundamentação destas práticas pela investigação tem sido escassa. Acresce ainda o facto de maioria dos testes com sons não verbais não estar padronizada, variando muito no número e tipo de sons, na qualidade das gravações, nas características acústicas das amostras, nos procedimentos de teste e nos formatos de resposta (Gygi e Shafiro 2006).

O facto de o processamento auditivo de sons não verbais não implicar competências linguísticas complexas faz com que sejam indicados para testar a audição de sujeitos com limitados conhecimentos da língua tais como crianças, ou pacientes com défices cognitivos e linguísticos (Finitzo-Hieber et al. 1980, Schnider et al. 1994, Fabiani et al. 1996). O uso de sons não verbais com gamas de frequências filtradas tem sido defendido como uma alternativa aos tons puros dos testes audiológicos feitos a crianças. A atenção das crianças é mantida mais facilmente e durante mais tempo do que com os tons puros, mantendo a capacidade de avaliar limiares de audição em gamas de frequências específicas (Myers et al. 1996). Este tipo de sons também tem

vido utilizado para diagnosticar lesões do sistema nervoso central, em que se procuram encontrar diferenças no processamento da fala, música e sons não verbais.

2.2.2 A CLASSIFICAÇÃO DE GAVER (1993) DOS SONS AMBIENTAIS

Gaver (1993) aborda diferentes questões relacionadas com a audição do dia a dia e propõe uma classificação para os sons que ocorrem quando há interacção entre dois materiais. O autor define “everyday listening” como a experiência de ouvir/escutar acontecimentos (ao contrário de sons). O ouvinte presta atenção a sons que o rodeiam, de forma a saber o que se passa à sua volta, i.e., o que produz o som e em que contexto. Por sua vez, aquilo que denomina de “musical listening” reporta-nos para uma procura do significado do som, das sensações que ele provoca, enquanto pela “everyday listening” tentamos identificar os instrumentos que originaram o som.

Gaver (1993) realça a importância da abordagem ecológica da percepção auditiva para os sons não verbais, segundo a qual a percepção normalmente se relaciona com acontecimentos/eventos e entidades complexas do mundo quotidiano. Gaver (1993) defende que esta percepção é directa, e não é mediada pela inferência nem pela memória. Na Figura 1 encontra-se uma representação da “everyday listening”, mostrando vários acontecimentos que influenciam a percepção do som desde que é produzido até que é percebido pelo ouvinte. Perceber o percurso desde que o evento acontece até que é experienciado engloba áreas como a mecânica, dinâmica de fluidos, acústica, anatomia e psicoacústica. Uma observação breve sobre este contínuo sugere que um dado som fornece informação sobre a *interacção* de *materiais* num determinado *local* do *meio/contexto*.

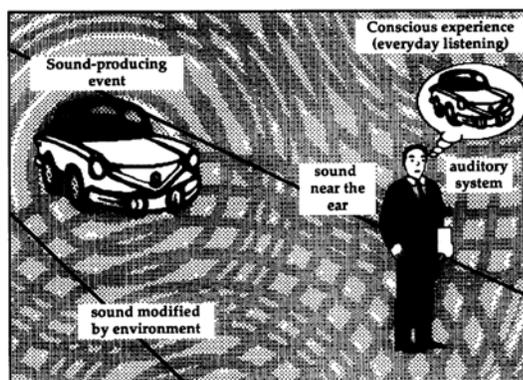


Figura 1: O contínuo do contexto para a experiência. A fonte de um evento gera ondas sonoras: algumas são irradiadas directamente para um ponto de observação; outras são modificadas pelo contexto antes de chegarem ao ouvinte. O sinal à entrada do sistema auditivo, fornece informação sobre a experiência da “everyday listening”. De Gaver (1993).

Em relação ao evento que produz o som, ou fonte de som, é importante identificar os atributos físicos da fonte e de que forma estes influenciam o som acusticamente. Todas as fontes de som envolvem a interacção de materiais, sendo importante perceber como é esta interacção. O contexto em que o som ocorre vai influenciá-lo: as ondas sonoras perdem energia, especialmente energia de alta-frequência, à medida que percorrem o ar, fornecendo informação quanto à distância da fonte. As mudanças no volume provocadas por alterações na distância da fonte fornecem informação quanto ao “tempo até ao contacto”.

Segundo a classificação de Gaver (1993) os materiais envolvidos na produção de sons são de extrema importância na análise dos sons obtidos pela interacção entre os materiais. Tal como se pode observar na Figura 2, Gaver (1993) considera que existem três tipos principais de sons, consoante a interacção dos materiais: sons de objectos em vibração, sons aerodinâmicos, e sons líquidos.

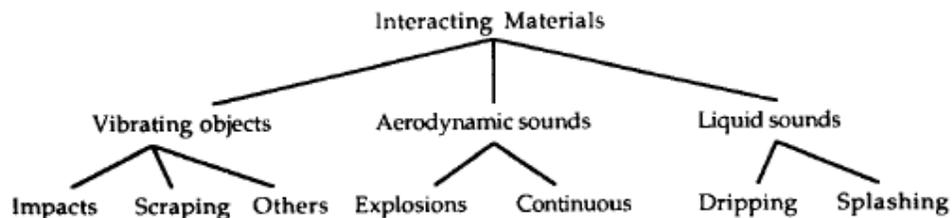


Figura 2: Descrição hierárquica de eventos sonoros simples. De Gaver (1993).

O padrão de vibração e seus efeitos acústicos num dado objecto é afectado por vários atributos físicos (ver Tabela 1). O *tipo de interacção* (*impacto*, *raspar*, ou *rolar*) afecta a amplitude ao longo do tempo e o espectro da vibração. A *força da interacção* determina a amplitude de vibração geral (quanto maior a força maior a amplitude); a força também afecta a largura de banda da vibração deformando os objectos mais bruscamente e aumento as componentes de alta-frequência (Gaver 1993).

Tabela 1: Efeitos acústicos dos atributos das fontes sonoras. De Gaver (1993).

Acoustic Effects of Source Attributes	
Source	Effects on the Sound Wave
Interaction	
Type	Amplitude function, spectrum
Force	Amplitude, bandwidth
Material	
Restoring force	Frequency
Density	Frequency
Damping	Amplitude functions; also frequency
Homogeneity	Complex effects on amplitude; also frequency
Configuration	
Shape	Frequency, spectral pattern
Size	Frequency, bandwidth
Resonating cavities	Spectral pattern
Support	Amplitude functions, frequency, spectrum

As características do *material* determinam a frequência das vibrações. Por exemplo, a *estrutura interna do material* pode ter efeitos complexos na vibração, principalmente no domínio temporal. A madeira parece amortecer de uma forma mais irregular do que os materiais como o metal.

A *forma* do objecto e o *tamanho* determinam a frequência de vibração (objectos grandes tendem a produzir sons mais graves do que os pequenos). Estas diferenças no espectro são também influenciadas pelo facto de haverem *cavidades de ressonância* ou não. Vibrações no ar dessas cavidades reforçam ou amortecem as vibrações do material circundante correspondente. Aquilo em que o objecto está apoiado pode alterar o *espectro* e a *frequência de amortecimento*.

Podemos agrupar os parâmetros descritos pelos efeitos que produzem: domínio frequencial ou temporal (ver Tabela 2). Atributos do objecto (forças de recuperação, tamanho e densidade) tendem a influenciar os sons na frequência, enquanto que as características da interacção (o seu tipo e força) influenciam o domínio temporal.

Tabela 2: Efeitos frequenciais e temporais. De Gaver (1993).

Frequency and Temporal Effects	
<i>Frequency Domain</i>	<i>Temporal Domain</i>
Restoring force	Interaction type
Density	Damping
Size	Internal structure
Shape	Support
Support	

Os sons de eventos aerodinâmicos resultam de diferenças de pressão atmosférica. A maior parte da informação acústica das *explosões* parece relacionar-se com o tamanho ou força da explosão. As componentes de alta-frequência tendem a indicar a brusquidão da mudança de pressão perto da fonte (e logo a duração da libertação de pressão inicial). Sons produzidos pela passagem de ar através de objectos ou superfícies resultam de *variações na pressão* (turbulência).

Um som de um líquido pode ocorrer ao *pingar* ou ao *salpicar*. São semelhantes aos sons de objectos em vibração no sentido em que dependem de uma *deformação* inicial que é contraposta pelas forças de recuperação de um material. Os sons resultantes desta interacção são determinados pela formação e alteração de cavidades de ressonância à superfície do líquido (Gaver 1993).

Na Figura 3 pode observar-se um quadro proposto por Gaver (1993) para classificar os sons não verbais, onde estão representados os vários tipos de interacção possíveis dentro de cada grupo, e exemplos de sons que se incluem em cada uma das categorias.

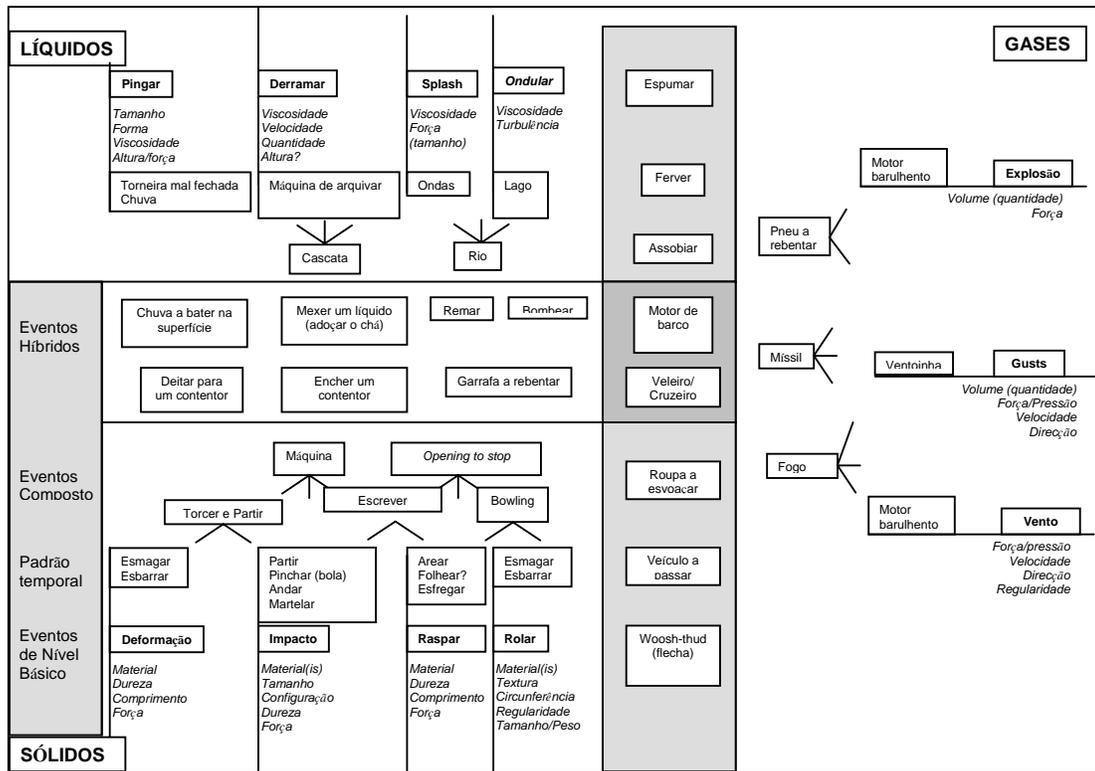


Figura 3: Um “mapa” de sons não verbais adaptado de Gaver (1993, p. 24). Podem observar-se três fontes fundamentais (sólidos em vibração, líquidos e gases), sobrepondo-se nalgumas zonas. Dentro de cada secção, vemos os eventos de nível básico a negrito, e os seus atributos relevantes aparecem por baixo deles em itálico. A complexidade aumenta no sentido do centro da figura, desde eventos de nível básico, eventos que cariam no padrão temporal, eventos compostos e eventos e híbridos.

2.3 SELECÇÃO DOS ESTÍMULOS

A variedade de sons não verbais, tal como já referimos, poderá ser virtualmente infinita. A falta de uma classificação ou de uma descrição de um conjunto de estímulos significativos para serem tomados como referência, faz com que seja difícil saber quais são os estímulos mais apropriados a incluir num teste. Além disso, há uma grande variabilidade nos métodos e estímulos utilizados em estudos anteriores, não havendo concordância quanto aos critérios de avaliação da qualidade e aplicabilidade dos estímulos a determinadas tarefas experimentais. Daí que, quando se inicia um trabalho nesta área, seja necessário pesquisar bases de dados de sons não verbais, definindo novas linhas de orientação.

Dependendo dos objectivos de cada estudo em particular, pode ser adequado usar sons produzidos naturalmente (e.g., por um Terapeuta da Fala) ou gravar sons especificamente para um projecto. No entanto, produzir sons em tempo real ou gravá-los nem sempre é possível, principalmente quando o objectivo do estudo é relativo à percepção de um grande número de sons

não verbais, já que irá implicar um grande investimento de tempo e de materiais. Neste tipo de situações, os investigadores tendem a utilizar compilações de gravações de sons não verbais já disponíveis no mercado (Shafiro e Gygi 2004). Segundo Shafiro (2004) o uso gravações já existentes fornece uma qualidade consistente dos estímulos e permite que se investigue um grande número de sons não verbais num só estudo. Além disso, também torna possível a replicação dos resultados e a comparação entre estudos.

Considerando os estudos efectuados com sons não verbais previamente gravados, podemos dividi-los em quatro categorias (Shafiro e Gygi 2004):

- Estímulos sonoros isolados múltiplos de uma mesma fonte física, ou do mesmo tipo de fonte sonora, em que se investiga a capacidade dos sujeitos em perceber uma propriedade particular do mesmo objecto ou evento que produz o som;
- Inventários relativamente vastos com sons não verbais isolados para investigar a exactidão na identificação das fontes sonoras respectivas, ou a concordância na nomeação entre populações diferentes de ouvintes, ou a relação entre parâmetros acústicos e factores cognitivos envolvidos na percepção;
- Sequências temporais de sons não verbais produzidos por fontes distintas que não se sobrepõem no tempo para que seja possível analisar os efeitos do contexto nos sons antecedentes e procedentes e a facilidade de aprendizagem de determinadas sequências sonoras;
- Misturas de sons não verbais, em que há uma sobreposição dos vários sons em tempo e frequência, para avaliar a identificação feita por humanos e por máquinas.

Os objectivos deste estudo relacionam-se com a análise e identificação de uma vasta gama de sons não verbais, em experiências distintas, as quais descreveremos nos Capítulos 3, 4 e 5. Estes objectivos foram traçados tendo por base a futura aplicação clínica na terapia da fala dos sons não verbais. Na terapia da fala são já utilizados vários materiais pedagógicos com este tipo de sons, mas de uma forma aleatória baseada nos materiais usados (e.g. representações de sons da floresta, de sons de acções) e não segundo uma caracterização acústica, tal como propomos ao longo deste trabalho. Este estudo estará enquadrado na segunda categoria da proposta de Gygi (2006) (relacionar a identificação de sons não verbais com tarefas cognitivas), uma vez que será realizada a identificação de sons não verbais, a sua nomeação através de onomatopéias, assim como a análise da relação entre parâmetros acústicos e os grupos de sons incluídos em determinadas categorias.

O âmbito deste projecto não permitiu grandes recursos materiais e humanos, o que leva a optar por sons já existentes em bases de dados. Os sons não verbais foram obtidos das ferramentas de trabalho do Terapeuta da Fala para actividades de treino auditivo, sendo seleccionados os seguintes jogos didácticos:

- “Soundtracks”, Living & Learning (cassete áudio);
- “Sons de Exteriores, Sons de Interiores”, Color Cards (cassete áudio);
- “Digo o que faço, Faço o que Digo”, Areal Editores (CD);
- “Ginásio dos Sons – 1: Sons isolados”, Edições Convite à Música (CD);
- “Ginásio dos Sons – 2”, Edições Convite à Música (CD).

Procedeu-se à catalogação dos sons não verbais existentes nas quatro bases de dados, aplicando-se um código a cada estímulo. Observou-se que entre as várias bases de dados haviam estímulos que eram comuns. Assim, considerou-se ser mais adequado optar pelas bases de dados em formato digital, no sentido de facilitar o processamento do sinal, já que não haveria a perda da variedade de estímulos. De seguida, tendo por base critérios de classificação dos estímulos (os quais serão abordados na secção 2.4 deste Capítulo) e o requisito de haver um conjunto representativo de sons não verbais para cada categoria, procedeu-se à selecção dos estímulos.

Os dados dos CDs originais dos jogos didácticos foram transferidos digitalmente para ficheiros *.wav* mono de 16 bits, com uma frequência de amostragem de 44 kHz. Foram codificados de acordo com o CD de origem e a ordem de aparecimento no mesmo e seleccionados os sons adequados à constituição da base de dados (ver na Tabela 3).

De seguida realizou-se a segmentação manual do Corpus de Sons Não Verbais utilizando o programa *Adobe Audition* onde se observou o sinal acústico de cada estímulo. Uma vez que alguns ficheiros possuíam o mesmo som repetido várias vezes (e.g., sino, andar, martelar, grasnar ou cacarejar) e outros tinham uma duração muito extensa (e.g., chuva, serrar, vento, cascata), foi necessário editar alguns dos estímulos.

2.4 CLASSIFICAÇÃO DOS ESTÍMULOS

Um dos objectivos deste estudo é tentar encontrar parâmetros acústicos que justifiquem a construção de grupos taxionómicos relativamente aos sons não verbais. A classificação pode basear-se no contexto sonoro em que os sons são ouvidos (e.g., sons da casa, sons da rua e sons da selva) tal como acontece em alguns dos jogos sonoros cujos estímulos foram utilizados (Ginásio dos Sons 1 e Ginásio dos Sons 2). Pode optar-se pelo critério da relevância dos sons para o bem-estar do ouvinte (e.g. sons de alarme, sons relaxantes e sons distractivos). No entanto, na taxionomia proposta por Gaver (1993) as categorias não são mutuamente exclusivas e não fornecem informações específicas para a identificação dos sons. Gaver (1993) apresenta uma classificação que se baseia precisamente em parâmetros acústicos que resultam das interacções dos objectos ou eventos que são as fontes sonoras. A percepção precisa das características da

fonte sonora de um determinado estímulo poderá ser mais útil para a descrição de um contexto sonoro ou para a avaliação da sua importância para o bem-estar da pessoa (Shafiro 2004).

Assim, procurou-se classificar os estímulos segundo Gaver (1993), de forma a posteriormente verificar se os resultados obtidos revelariam ou não a existência de características comuns entre os sons não verbais de cada categoria. Verificamos que na Classificação de Gaver (1993) os sons de animais não foram estudados. No entanto, devido ao facto de este tipo de estímulos sonoros ser bastante utilizado no treino auditivo, levou-nos a incluí-los na base de dados construída (ver Tabela 3).

Como se pode observar na Tabela 3 a base de dados é constituída por 17 sons produzidos pela interacção de sólidos, 7 sons cujas fontes sonoras são líquidos, 6 estímulos sonoros resultantes de fontes aerodinâmicas e 8 sons de animais. É de realçar que da categoria de sólidos fazem parte 5 sons de instrumentos musicais. No agrupamento dos estímulos da nossa base de dados seguimos o proposto por Gaver (1993), especificamente no mapa de sons não verbais (Gaver 1993, p. 24, ver Figura 3).

Tabela 3: Estímulos incluídos no estudo. Os sons com * correspondem a sons de instrumentos musicais. O código da primeira coluna indica a origem do estímulo: “DQF” – corresponde ao jogo “Digo o que Faço, Faço o que Digo”, “GS1” – corresponde ao jogo “Ginásio dos Sons 1”, “GS2”, pertence ao jogo “Ginásio dos Sons 2”). Os últimos 3 algarismos referem-se à ordem de aparecimento do estímulo no respectivo jogo.

Nº	Código	Descrição	Classificação de Gaver (1993)																			
			Sólidos				Líquidos				Gases											
			Impacto	Deformação	Raspar	Rolar	Pingar	Derramar	Splash	Ondular	Explosão	Gust	Vento									
1	GS1035	sino *																				
2	GS1020	estalos dedos																				
3	GS1016	andar																				
4	GS2029	talheres e prato																				
5	GS2052	machado a rachar lenha																				
6	GS1026	martelar																				
7	DQF017	sineta *																				
8	DQF021	palmas																				
9	DQF032	descer escadas																				
10	DQF030	tambor *																				
11	GS2051	pá a enterrar na areia																				
12	DQF014	porta a fechar																				
13	DQF041	maracas *																				
14	DQF043	pandeireta *																				
15	DQF024	esfregar as mãos																				
16	DQF051	serrar																				
17	DQF050	varrer																				
18	DQF034	encher um copo de água																				
19	GS2013	chuva																				
20	GS2008	cascata																				
21	GS2035	chuveiro																				
22	GS1042	mar																				
23	GS2027	água a ferver																				
24	GS2009	rio																				
25	GS1037	trovão/relâmpago																				
23	GS2027	água a ferver																				
26	DQF053	encher pneu																				
27	GS1031	avião																				
28	DQF025	assobiar																				
29	GS1039	vento																				
			Sons de Animais																			
30	GS1002	vaca	mamífero																			
31	GS1004	porco	mamífero																			
32	GS1005	galo	ave																			
33	GS1006	cão	mamífero																			
34	GS1008	pato	ave																			
35	GS1009	grilo	insecto																			
36	GS1010	pássaro	ave																			
37	GS1012	gato	mamífero																			

2.4.1 SÓLIDOS

Um som é o resultado de um *impacto* quando ocorre uma deformação rápida e discreta do(s) sólido(s) envolvido(s) (Gaver 1993). De seguida ir-se-á apresentar a lista de estímulos desta categoria e a respectiva justificação da classificação:

- Sino e sineta: o som produz-se porque há um impacto entre o pêndulo e a parte interior da cavidade do instrumento, são ambos incluídos no grupo dos sons de percussão de tipo idiofones em que há uma casca que vibra após o impacto com uma vara (Henrique 1997);
- Estalos de dedos: o som produz-se porque há um impacto entre os dedos;
- Andar: o som produz-se porque há um impacto entre a sola do sapato e o chão;
- Talheres e prato: o som produz-se porque há um impacto entre os talheres e o prato;
- Machado a rachar lenha: o som produz-se porque há um impacto entre o machado e a lenha;
- Martelar: o som produz-se porque há um impacto entre o martelo e a cabeça do prego.

No mapa de sons não verbais proposto por Gaver (1993) e apresentado na Figura 3, são exemplos de sons resultantes de uma *deformação* os sons *esmagar* e *esbarrar*, em que para além do impacto há uma deformação dos sólidos envolvidos. Assim, considerou-se o seguinte:

- *Palmas*: o som produz-se porque há um impacto entre as palmas das mãos e também uma deformação das mesmas;
- *Descer escadas*: o som produz-se porque há um impacto entre a sola do sapato e os degraus das escadas, e há uma deformação da madeira dos degraus das escadas;
- *Tambor*: o som produz-se porque há um impacto entre a baqueta do tambor e a pele que cobre a cavidade de ar, e há uma deformação da pele (Henrique 1997).

Raspar implica uma interacção mais contínua entre os sólidos, ao contrário de um impacto. Na base de dados encontramos os seguintes sons incluídos nesta categoria:

- *Esfregar as mãos*: o som produz-se porque as palmas das mãos raspam entre si,
- *Serrar*: o som produz-se porque o serrote raspa na madeira,
- *Varrer*: o som produz-se porque a vassoura raspa no chão.

Existem ainda alguns estímulos constituídos por duas partes, em que numa delas há um *impacto* e na seguinte há o *raspar*:

- *Pá a enterrar na areia*: o som produz-se porque há um impacto entre a pá e a areia e a pá raspa na areia,
- *Porta a fechar*: são produzidos dois sons, o primeiro é resultado da dobradiça a girar quando a porta começa a fechar, e os elementos da dobradiça raspam entre si; o segundo som é resultado do impacto entre a porta e a “moldura” de madeira,

- *Maracas*: o som produz-se porque há um impacto entre as areias da maraca e as suas paredes interiores, e também porque as mesmas areias raspam nas paredes internas da maraca (Henrique 1997),
- *Pandeireta*: o som produz-se porque há um impacto e fricção entre os pratos que estão à volta da pandeireta (Henrique 1997).

2.4.2 LÍQUIDOS

Os sons gerados por líquidos resultam da formação e alteração de cavidades de ressonância à superfície do líquido (Gaver 1993). Ao *pingar* o líquido entra em contacto com outro líquido ou com uma superfície sólida. No caso de *pingar* num líquido é formada uma cavidade de ressonância e posteriormente uma coroa e esguicho, à medida que a gota é misturada com o restante líquido. Assim sendo, considerámos a seguinte classificação:

- *Encher um copo de água*: o som é o resultado do pingar da água para o copo, onde há interacção entre a água e o vidro do copo e entre a água e a água que vai ficando no interior do copo,

No mapa de sons não verbais apresentado na Figura 3, Gaver (1993) propõe o tipo de interacção “*splash*” apresentando os exemplos das ondas do mar, da cascata e do rio. Por associação, considerámos na nossa base de dados que:

- *Mar*: o som é o resultado das ondas do mar a “desfazerem-se” na areia, gerando um som do tipo “*splash*”;
- *Água a ferver*: o som tipo “*splash*” resulta do líquido das bolhas de água que ao “rebentarem” fazem o líquido “bater” na restante água; este som também possui características de *explosão*, de que falaremos mais à frente.

Também é referido por Gaver (1993) a interacção *ondular* associada a águas mais paradas como o lago e o rio, daí que na base de dados consideramos:

- *Rio*: o som resulta do barulho da água a *ondular* e também terá características de “*splash*” ao ir de encontro às margens do rio.

Para além do que já referimos, existem ainda sons que resultam da co-ocorrência de dois tipos de interacção, nomeadamente:

- *Chuva, Cascata e Chuveiro*: considerou-se que incluem não apenas a interacção de *pingar*, mas também componentes de “*splash*” uma vez que se trata de um som contínuo e o pingar é bastante intenso e rápido, o que se traduz num som semelhante a “*splash*”. Na chuva e no chuveiro a interacção é entre o líquido e maioritariamente um sólido (chão ou banheira), enquanto que na cascata a interacção é entre líquidos,

o que gera maiores cavidades de ressonância, sendo a quantidade de líquido envolvida também maior.

2.4.3 GASES

Os sons que são incluídos nesta categoria resultam da ocorrência de modificações na pressão. Podem haver três tipos de interacção: *explosão*, *“gust”* e *vento*.

Numa *explosão* há uma quantidade de gás que é libertado na atmosfera, entrando em contacto dois gases com grande diferença de pressão, gerando repentinamente um som. Incluímos nesta categoria os seguintes sons não verbais:

- *Trovão*: consideramos tratar-se de um som do tipo explosão porque resulta do eco da onda de choque nas altas camadas da atmosfera e da geografia envolvente.
- *Água a ferver*: também incluímos na categoria de *explosão* porque para além do *“splash”* também ocorrem pequenas explosões de bolhas de vapor de água.

Sons classificados por *“gust”* são exemplificados por Gaver (1993) como uma ventoinha ou um míssil, em que há uma passagem rápida do ar, podendo ser mais ou menos contínua. Consideramos haver dois estímulos da base de dados que podem ser classificados nesta categoria:

- *Encher um pneu*: há uma entrada de ar rápida da atmosfera e esse ar entra em contacto com um sólido gerando o som;
- *Avião*: resulta um som tipo *“gust”* porque a superfície sólida do avião ao entrar em contacto com a atmosfera rapidamente gera o som.

Por último, os sons do tipo *vento* são aqueles em que há uma passagem do ar de forma mais contínua. Incluímos os seguintes sons nesta categoria:

- *Assobiar*: o som é o resultado da passagem do ar pelas cavidades orais;
- *Vento*.

2.4.4 SONS DE ANIMAIS

O último grupo de sons incluído na base de dados não é classificado por Gaver (1993), mas o facto de serem sons utilizados regularmente no treino auditivo e nos materiais lúdicos das crianças faz com que sejam elementos da base de dados desta dissertação. Assim, os estímulos incluem animais de três classes principais: mamíferos, aves e insectos. Tendo em conta que as estruturas anatómicas de cada uma das classes possuem elementos comuns nas suas produções

sonoras, considera-se pertinente analisar os resultados por categorias relativas à classe animal. Assim sendo:

- *Vaca*: pertence à classe animal dos mamíferos, onde o som se produz pela vibração de tecidos moles, nomeadamente pregas vocais;
- *Porco*: pertence à classe animal dos mamíferos, em que o som se produz pela vibração de tecidos moles, nomeadamente pregas vocais;
- *Cão*: pertence à classe animal dos mamíferos, na qual o som se produz pela vibração de tecidos moles, nomeadamente pregas vocais;
- *Gato*: pertence à classe animal dos mamíferos, onde o som vocal se produz pela vibração de tecidos moles, nomeadamente pregas vocais;
- *Galo*: pertence à classe animal das aves;
- *Pato*: pertence à classe animal das aves;
- *Pássaro*: pertence à classe animal das aves;
- *Grilo*: pertence à classe animal dos insectos; os grilos possuem uma série de pelos nas bordas de suas asas, alinhados como pentes, e produzem os sons roçando uma asa contra a outra.

2.5 SUMÁRIO

Neste capítulo fez-se uma descrição dos processos envolvidos na audição de sons não verbais, distintos da percepção de estímulos de fala. Na escuta quotidiana de sons não verbais o ouvinte procura perceber a fonte que gera esse som, enquanto que na escuta de música e de fala o objectivo é perceber uma mensagem transmitida ao longo de um período de tempo mais longo.

Os estudos de sons não verbais podem ser aplicados na avaliação, diagnóstico e intervenção do funcionamento do sistema auditivo. Neste capítulo foi abordada a Classificação de Sons Não Verbais proposta por Gaver (1993), que agrupa os sons de acordo com os eventos físicos que os geram, defendendo o pressuposto que as propriedades acústicas dependem do material (*sólido*, *líquido* ou *gasoso*) e da interacção dos próprios materiais (e.g., *impacto*, *pingar* ou *explosão*).

Na segunda parte deste capítulo foi descrita a selecção, segmentação e anotação dos estímulos, e apresentada a sua classificação com base no enquadramento teórico proposto por Gaver (1993).

CAPÍTULO 3: ANÁLISE ACÚSTICA DOS ESTÍMULOS

3.1 INTRODUÇÃO

Este capítulo descreve a análise espectral e temporal dos estímulos. No início faz-se referência a estudos onde é feita este tipo de análise. Posteriormente descrevem-se os procedimentos utilizados, e por último são apresentados e analisados os resultados obtidos para cada categoria de sons não verbais.

3.2 REVISÃO DA LITERATURA

O estudo dos sons não verbais, como se afirmou no capítulo anterior, tem merecido menos atenção do que outros estímulos sonoros. A fala e a música têm sido objectos de estudo muito frequentes, e as metodologias de análise acústica dos mesmos encontram-se muito desenvolvidas. Assim sendo, os procedimentos de análise acústica utilizados com a fala e a música têm sido usados por alguns autores no estudo de sons não verbais, daí que seja pertinente apresentar uma breve abordagem do estado da arte nestas áreas.

Os procedimentos utilizados para analisar os estímulos sonoros de fala podem basear-se na caracterização da frequência fundamental, na obtenção dos valores de formantes, da amplitude, da duração, do tempo máximo de fonação para determinadas tarefas fonatórias, do “jitter” e do “shimmer”, sendo obtidos muitas vezes os espectrogramas e formas de onda dos sinais (Colton 1996, Urrutia e Marco 2000, Behlau 2004).

O estudo dos sons não verbais poderá ser feito em diferentes âmbitos, e Gygi (2001) refere que podem ter como objectivo:

- obter valores qualitativos e diferenciais semânticos,
- relacionar a identificação de sons não verbais com tarefas cognitivas,
- estudar a acústica ecológica e a percepção de eventos sonoros,
- sintetizar sons.

Para este capítulo da dissertação os procedimentos utilizados estão relacionados com os estudos do primeiro e terceiro tipo, nomeadamente a obtenção de valores qualitativos e o estudo da acústica ecológica pela obtenção de valores quantitativos.

Alguns estudos de sons não verbais investigaram apenas um tipo de objecto ou evento, tais como garrafas a balouçar ou a partir (Warren e Verbrugge 1984), palmas (Repp 1987) e passos (Xiaofeng et al. 1991). Warren e Verbrugge (1984) investigaram a capacidade dos ouvintes identificarem correctamente fenómenos de garrafas a partir ou a balouçar através da informação acústica do sinal. Referem que estes sinais são constituídos por padrões de impulsos quase-periódicos isolados (garrafa a balouçar) ou múltiplos com um “burst” inicial de ruído (garrafa a partir) (ver Figuras 4 e 5).

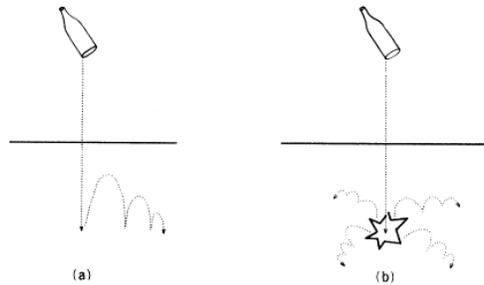


Figura 4: Ilustração dos eventos mecânicos: a) garrafa a balouçar, b) garrafa a partir. De Warren e Verbrugge (1984).

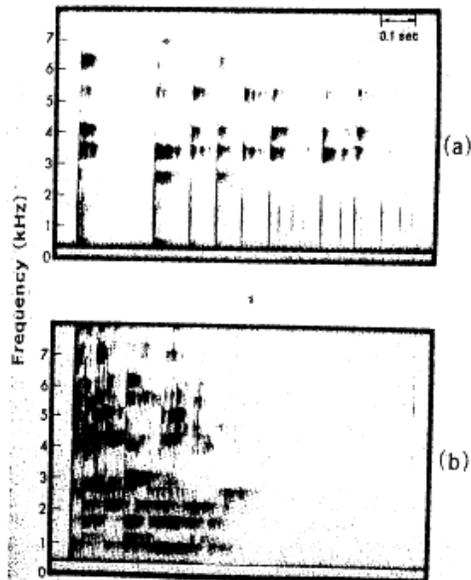


Figura 5: Espectrogramas de: a) garrafa a balouçar, b) garrafa a partir. De Warren e Verbrugge (1984).

A influência da configuração manual da produção de *palmas* na informação acústica do sinal gerado foi investigada por Repp (1987). Para todos os sujeitos foi gravada uma série de 10 *palmas* consecutivas. Para proceder à análise acústica dos sons foram medidos os intervalos de “onset-to-onset” (OOI), cuja média e desvio padrão forneceu a velocidade e ritmo das *palmas*, respectivamente. Foram também obtidos espectros de cada palma (ver Figura 6), foi calculada a amplitude média, e a variabilidade temporal. Esta análise serviu de base à discussão dos resultados da experiência de percepção em que os indivíduos tiveram de relacionar o estímulo sonoro com a configuração manual e o género da pessoa que produzia a palma. O ritmo médio foi de 4 s (média de OOI=250ms), sem haver diferenças entre os géneros. A variabilidade temporal foi em média de 6.8 ms, e o desvio padrão das amplitudes variou entre os 0.7 e os 5.2 dB, sem diferenças entre os géneros dos sujeitos.

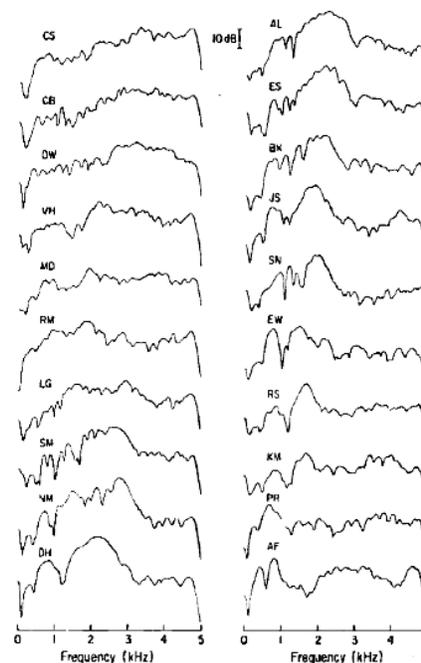


Figura 6: Espectro para cada palma dos 20 sujeitos. Cada espectro é a média aritmética dos espectros de cada palma isolada (das 10 *palmas*) obtida para os primeiros 10 ms após o início da palma. De Repp (1987).

Xiaofeng et al. (1991) analisaram a capacidade de sujeitos perceberem características da fonte sonora, nomeadamente o género da pessoa, na realização de um evento sonoro natural: um humano a andar. Os estímulos sonoros utilizados na experiência de percepção foram analisados acusticamente, com base nas formas de onda (Figura 7) em que o início foi identificado pela observação de um aumento súbito de amplitude seguido de oscilações rápidas da mesma (Figura 7), e o fim pela observação de oscilações lentas e baixas da amplitude, representando o ruído entre dois passos adjacentes. O início e o fim de cada passo eram identificados nos valores de amplitude mais próximos de zero. Estes dados eram sempre confirmados pela audição de cada passo. Para além da amplitude foi analisado o ritmo, nomeadamente o número de passos por

segundo. A observação dos espectros obtidos (Figura 8) evidenciou diferenças entre os gêneros, nomeadamente que para os homens a moda espectral tem frequências mais baixas, o espectro tem um pico mais largo, é mais inclinado ao longo das frequências baixas e que tem um aumento nas frequências baixas e uma queda nas frequências altas mais rápidos.

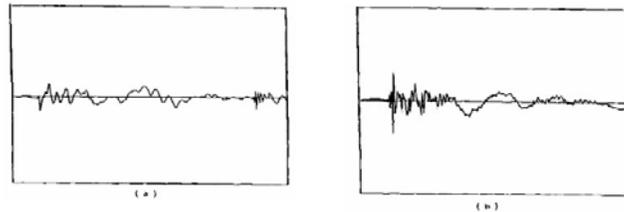


Figura 7: Forma de onda de um passo de um (a) homem e (b)mulher. De Xiaofeng et al. (1991).

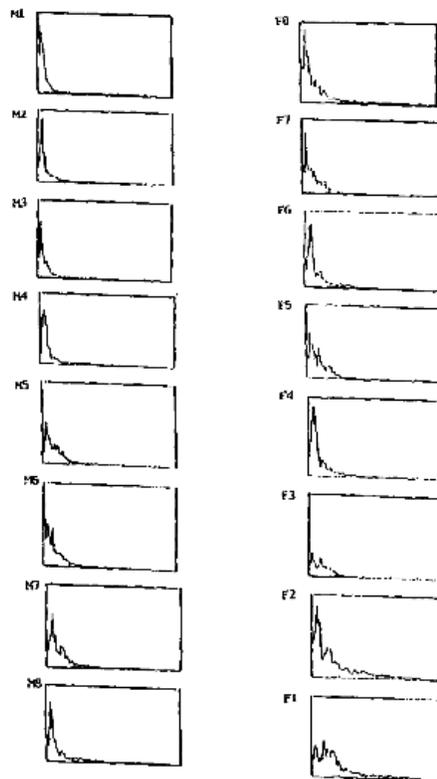


Figura 8: Espectros médios para os 16 andantes. As letras M e F representam o género dos sujeitos. O número de cada espectro representa a ordem de percepção do género. No eixo das abcissas está representado o intervalo linear das frequências até 5 kHz, e no eixo das ordenadas o intervalo linear da intensidade. De Xiaofeng et al. (1991).

Relativamente aos sons de instrumentos musicais, as medidas acústicas de duração, intensidade e frequência das ressonâncias são as mais utilizadas na análise acústica destes estímulos (Henrique 1997). Para analisar um som musical podem considerar-se três períodos (Figura 9):

- Transitório de ataque: corresponde à passagem do silêncio ao som, dura na maior parte dos instrumentos musicais entre alguns milissegundos a alguns centésimos de segundo; é muito importante para o reconhecimento do timbre, e se for cortado, o instrumento passa a ser auditivamente inclassificável na maior parte dos casos;
- Período de estabilidade: tempo entre o anterior e o seguinte, que dura normalmente alguns décimos de segundo a alguns segundos; é extremamente importante para certas características sonoras como a altura e a intensidade;
- Transitório de extinção (decaimento): período em que o som se extingue; é muito importante para o reconhecimento do timbre e se for cortado, o instrumento passa a ser auditivamente inclassificável na maior parte dos casos.

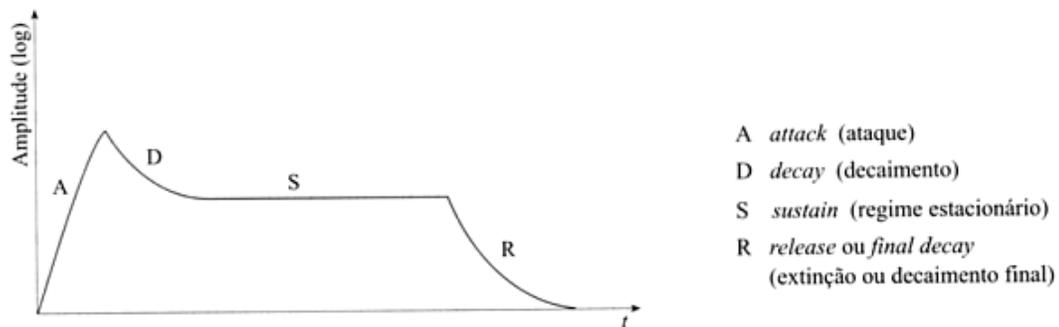


Figura 9: Forma de onda idealizada em que estão representados os períodos transitórios e o período de estabilidade. De Henrique (1997).

Se o instrumento musical for analisado como sistema, pode afirmar-se que este pode oscilar em regime transitório ou em regime permanente, dependendo do tipo de excitação (Henrique 1997). O regime permanente de um som é o estado estacionário em que as características do som se mantêm invariáveis no tempo. Por sua vez, o regime transitório é aquele que transmite a sensação de relevo acústico e dá uma estimativa da distância da fonte sonora. Henrique (1997) refere que um som musical contém informações relativas a altura, sensação de intensidade e timbre, se o som é liso ou com vibrato. Além disso, também transmite como é a sua evolução no tempo (alterações de intensidade e timbre), o tipo de ataque, e permite a apreciação do som como um todo musical e o reconhecimento do instrumento nesse todo, o tamanho e características do local onde o som é produzido (Figura 10).

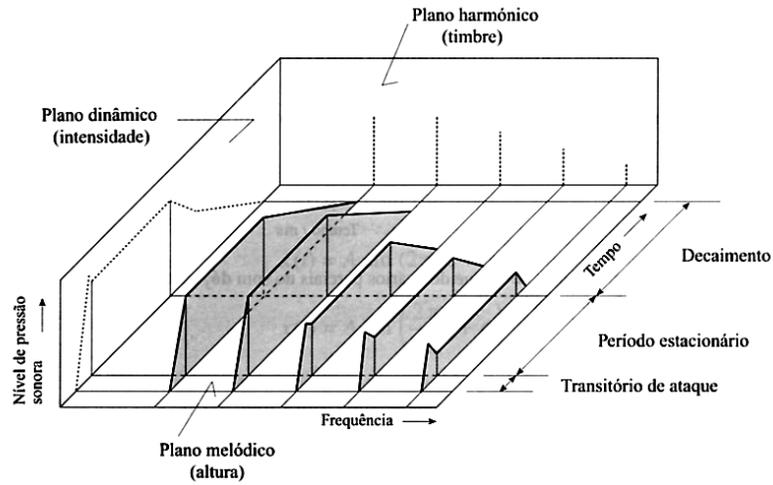


Figura 10: Representação tridimensional de um som onde são definidos os planos dinâmico, harmônico e melódico, e se pode observar a evolução temporal de cada parcial. De Meyer (1978), adaptado para o Português por Henrique (1997).

O uso da representação temporal e espectral na análise dos estímulos também é realizado no estudo de sons musicais (Henrique 1997). Nomeadamente na descrição do som produzido por membranofones (Figura 11) e idiofones (Figura 12).

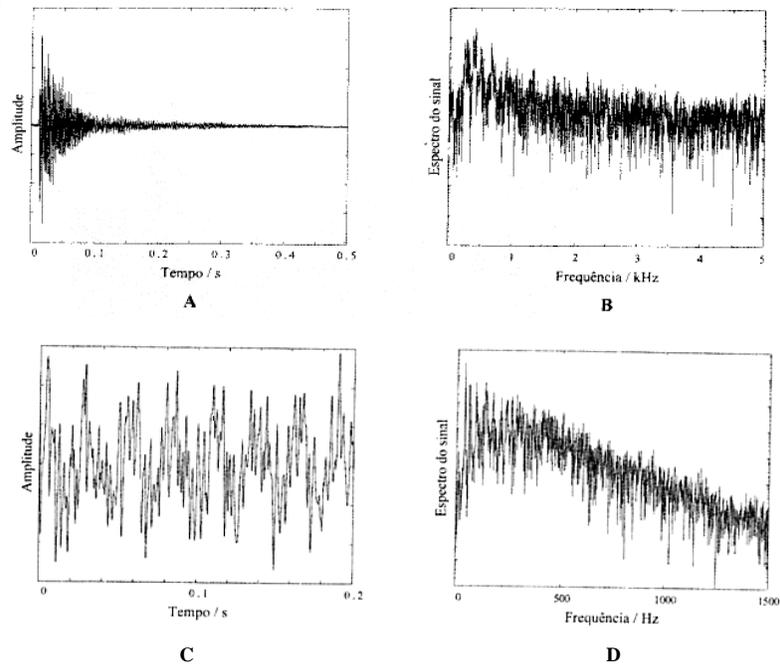


Figura 11: Em cima som da caixa de rufo: A) representação temporal, B) representação espectral. Em baixo som do bombo: C) representação temporal, D) representação espectral. De Henrique (1997).

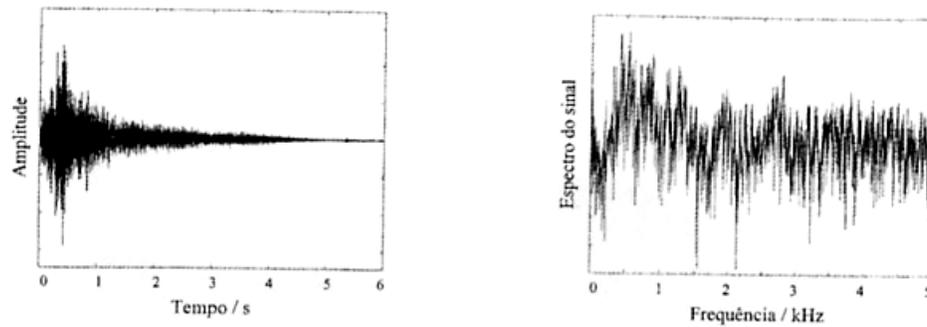


Figura 12: Som de pratos *Zildjan*: A) representação temporal, B) representação espectral. De Henrique (1997).

Os estudos com estímulos de sons de animais são normalmente realizados no âmbito do comportamento e comunicação animal, analisando-se o uso do som para transmitir mensagens diferentes, que podem estar relacionadas por exemplo com a luta ou o acasalamento (Riede et al. 2001, Brown et al. 2006). Para além da observação do comportamento, é feita a gravação dos estímulos, sendo obtidas medidas acústicas que servem de base à análise e discussão dos referidos estudos. Outro âmbito de aplicação dos sons de animais é a área da investigação veterinária, em que se analisam as patologias vocais destes animais (Ikeda et al. 2003, Rousseau et al. 2006), ou então em que se utilizam animais para testar hipóteses quanto a determinadas opções de tratamento para a voz humana e animal (Katada et al. 2004).

Gargstang (2004) estudou elefantes ao nível da produção de sons de baixa-frequência, analisando separadamente as características da fonte e dos filtros (onde se incluem os pulmões, laringe e tracto vocal, bem como a composição do ar expirado e da temperatura atmosférica). Gargstang (2004) usou uma representação em espectrograma dos sons registados, tal como é ilustrado na Figura 13 onde se consegue observar a F0, bem como as primeiras quatro ressonâncias (F1, F2, F3 e F4).

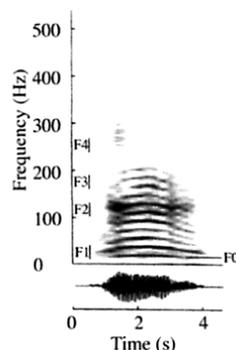


Figura 13: Forma de onda e o respectivo espectrograma de uma chamada de contacto feminina mostrando a frequência fundamental (F0) e harmónicos, e as primeiras quatro ressonâncias (F1 a F4). De Gargstang (2004).

Range e Fischer (2004) analisaram as vocalizações de primatas de acordo com determinados comportamentos. Para cada tipo de produção vocal foi calculada a duração. Os tipos de vocalização foram analisados a partir de espectrogramas e diferentes parâmetros foram utilizados para descrever a estrutura acústica de cada sinal. Na Figura 14 são apresentados espectrogramas de alguns tipos de vocalizações estudadas.

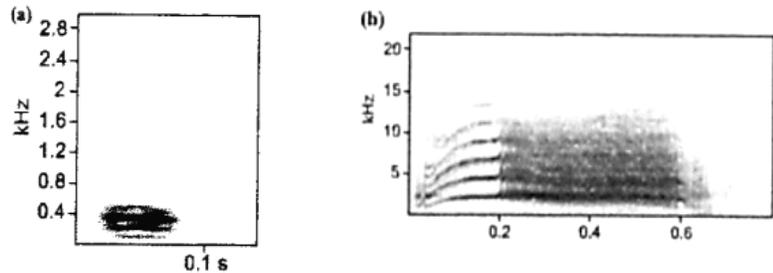


Figura 14: Grunhido emitido por uma fêmea adulta durante a perseguição (a). Fêmea adulta a berrar com um sub-adulto macho (b). De Range e Fischer (2004).

Os estudos de Riede et al. (2001) e de Hartwig (2005) tiveram ambos como objecto de estudo vocalizações caninas. No primeiro é afirmado que os latidos são tipicamente uma mistura de componentes acústicas periódicas e aperiódicas (ruído), em que a parte regular do sinal é fornecida por uma série de harmônicos devida a vibrações periódicas das pregas vocais, enquanto que o ruído resulta de vibrações inarmônicas do tecido das pregas vocais ou à turbulência do fluxo. Riede et al. (2001) e de Hartwig (2005) quantificaram a relação da energia harmónica/inarmónica (HNR), concluindo que esta medida é válida para classificar as produções vocais.

Lengagne (2001) gerou espectrogramas de sinais acústicos, descrevendo a produção de vocalizações de mochos-reais em termos de duração total, duração da parte ascendente, frequência mínima e frequência máxima (Figura 15).

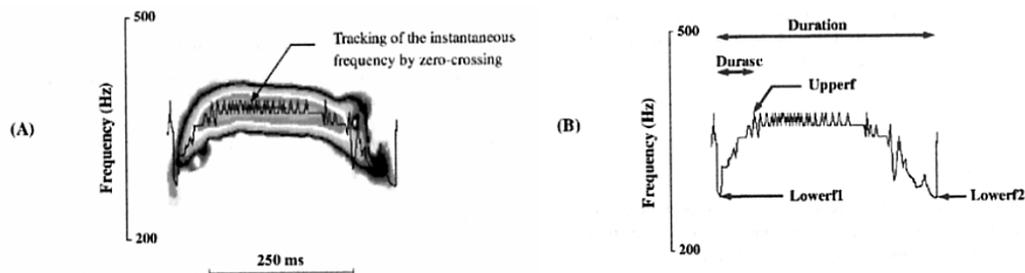


Figura 15: Espectrograma da produção vocal dum mocho-real (A) e parâmetros medidos para descrever a produção vocal (B). De Lengagne (2001).

3.3 PROCEDIMENTOS

Extraíu-se um conjunto de parâmetros acústicos do corpus de sons não verbais (Tabela 3), no sentido de encontrar características acústicas comuns aos elementos das quatro categorias, já que Gaver (1993) sustentou a sua classificação na existência de fenômenos físicos comuns que originam parâmetros acústicos comuns. Através do programa *Praat versão 4.3.11* (Boersma e Weenink 2005) procurou-se obter os seguintes dados:

- Duração (Figura 16);
- Frequência Fundamental (F0) média (Figura 16);
- F0 mínima (Figura 16) e F0 máxima (Figura 16);
- Formas de onda (Figura 17);
- Espectrogramas (Figura 18);
- Espectros da densidade de potência (Figura 19).

As formas de onda e os espectrogramas foram agrupados por categoria, de forma a facilitar a visualização para a análise e comparação entre sólidos, líquidos, gases e vocalizações de animais. Os valores das medidas acústicas foram introduzidos em tabelas de *Excel* para uma análise posterior.

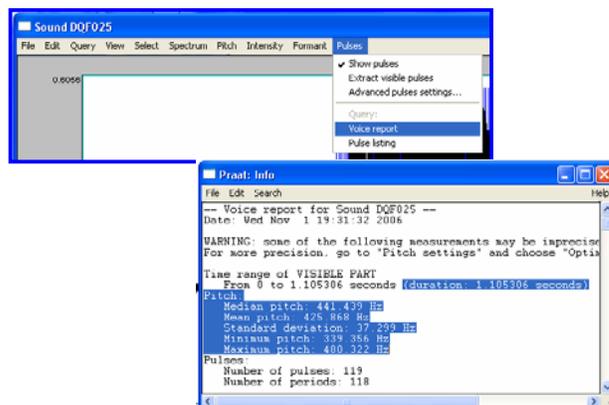


Figura 16: Procedimentos utilizados no programa *Praat* para obter os valores de duração, F0 médio, F0 mínimo e F0 máximo dos estímulos.

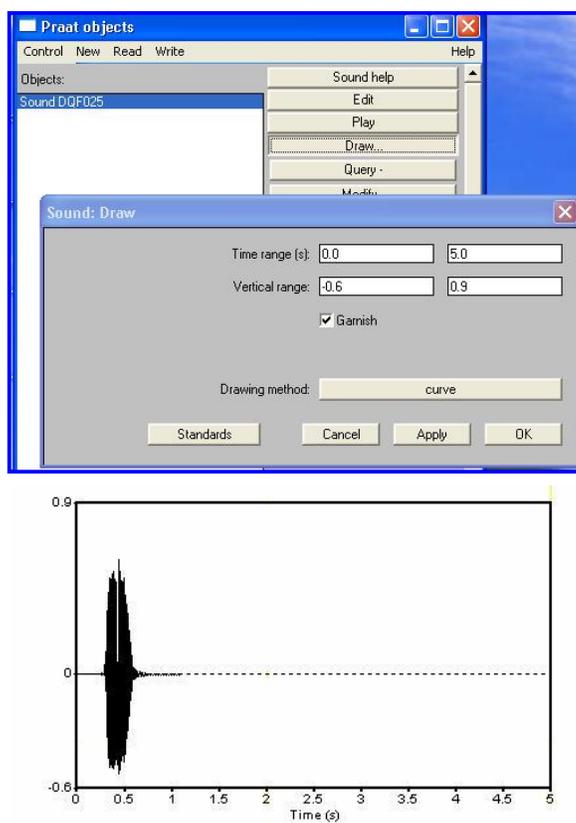


Figura 17: Procedimentos utilizados no programa *Praat* para obter a forma de onda dos estímulos.

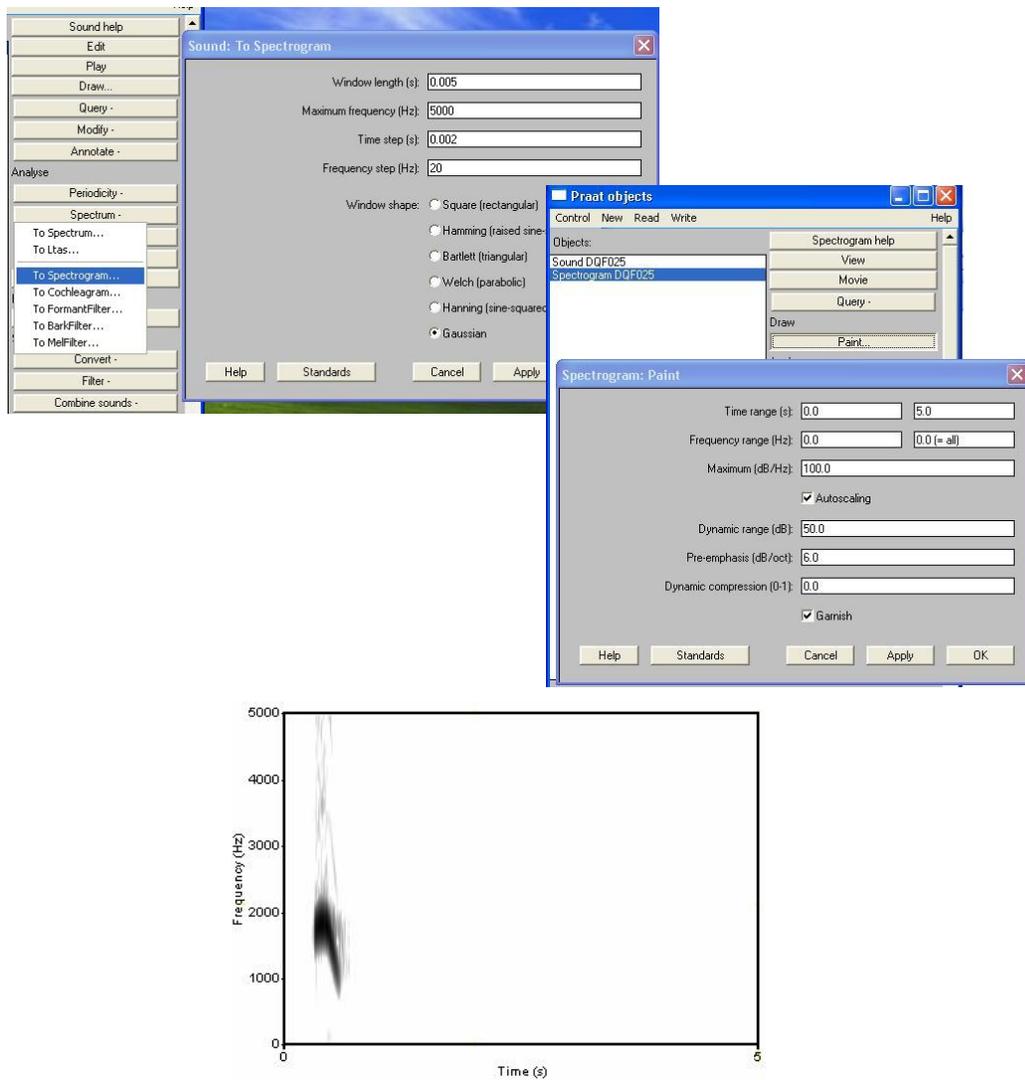


Figura 18: Procedimentos utilizados no programa *Praat* para obter o espectrograma dos estímulos.

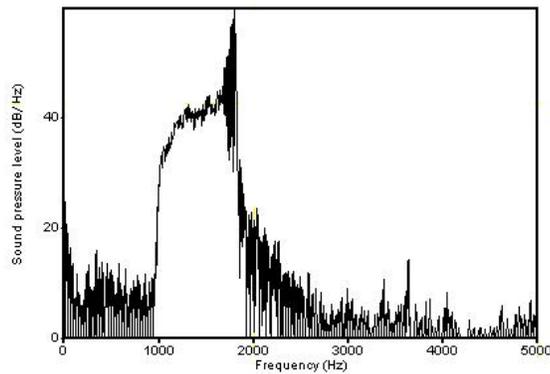
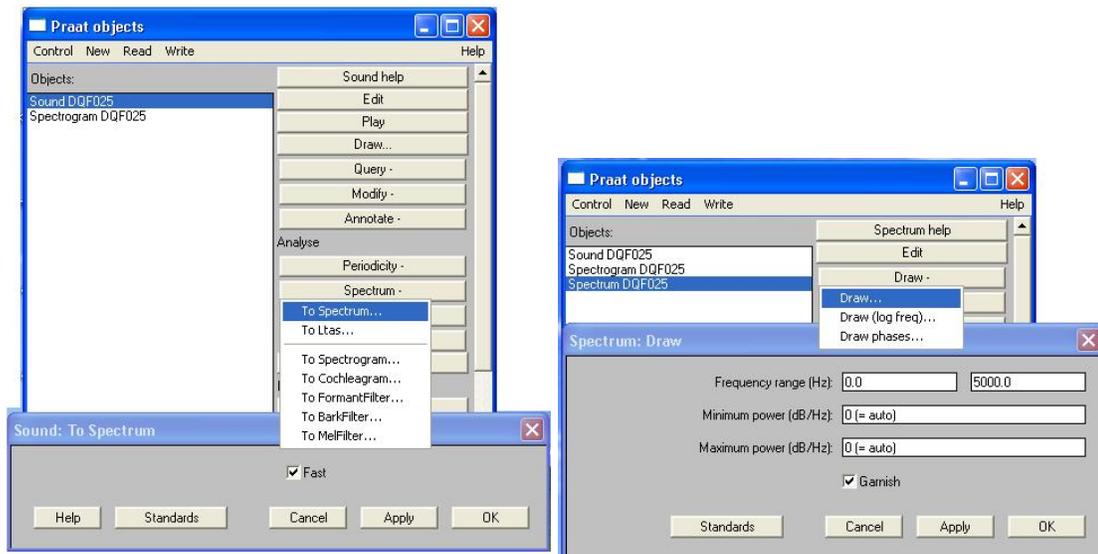


Figura 19: Procedimentos utilizados no programa *Praat* para obter os espectros da densidade de potência dos estímulos.

3.4 RESULTADOS

3.4.1 ANÁLISE TEMPORAL

Os dados obtidos foram agrupados por categoria de som não verbal e tipo de interação. Na Figura 20 observam-se diferenças na duração entre as categorias de sons não verbais. Os sons produzidos por sólidos têm uma menor duração (média=1.352 s e desvio padrão=1.129 s), principalmente os que são causados por uma *deformação*. O grupo de sons a que corresponde a interação *impacto* apresenta dois estímulos com uma duração muito superior à média (estímulo 4 – talheres e prato, e estímulo 11 – pá a enterrar na areia), que para além de um *impacto* envolvem

a interacção *raspar*, o que resulta numa maior duração. Na fala, os sons de consoantes com um modo de articulação do tipo oclusiva são os de menor duração (Kent and Read 2002, Stevens 1998).

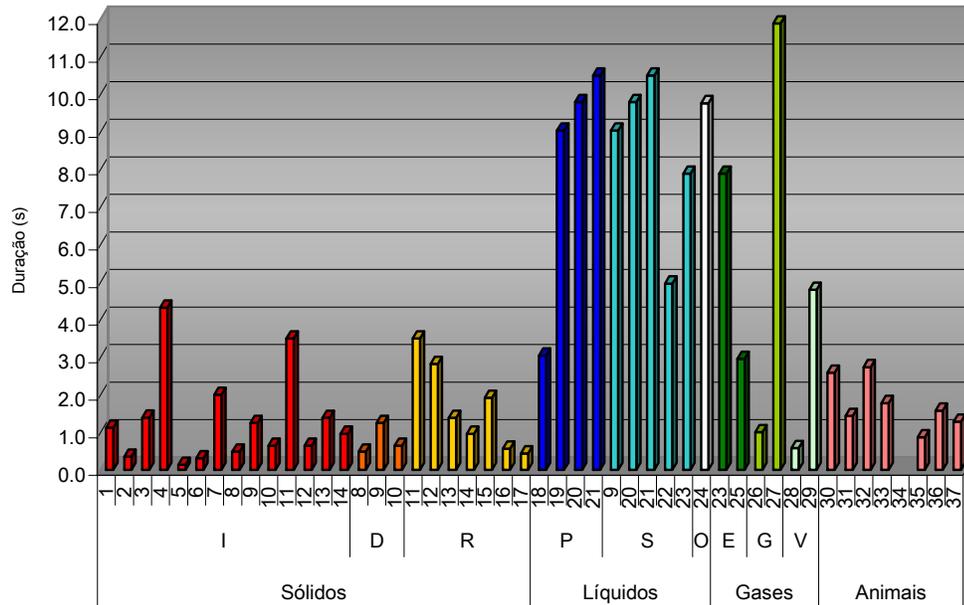


Figura 20: Figura da duração dos estímulos do corpus de sons não verbais. A vermelho estão os estímulos da categoria Sólido- *Impacto* (I), a laranja os estímulos Sólidos-*Deformação* (D), a amarelo os estímulos Sólidos-*Raspar* (R), a azul os estímulos do grupo dos Líquidos-*Pingar* (P), a azul turquesa os estímulos de Líquidos-*splash* (S), a branco o estímulo Líquido-*Ondular* (O) e a cor de rosa os estímulos do grupo dos Animais. Os números no eixo dos XXs correspondem ao número do estímulo.

Por outro lado os sons produzidos por líquidos e por gases são os que possuem estímulos com maior duração, sendo o de duração máxima o som do avião (nº 27). Relacionando com a fala, sabe-se que o modo de articulação do tipo fricativa é o que apresenta uma maior duração nas consoantes, e que as vogais são sons da fala com valores ainda superiores (Stevens 1998, Kent e Read 2002). Apesar de o estímulo com maior duração se incluir no grupo dos gases, a média da duração dos estímulos do grupo dos líquidos apresenta um valor superior à dos gases assim como um desvio padrão menor ao dos gases (média líquidos=8.284s, desvio padrão líquidos=2.499s; média gases=4.852s, desvio padrão gases=4.368 s).

Os sons do grupo dos animais apresentam uma média de 1.752s, a segunda maior dos quatro grupos e um desvio padrão bastante reduzido (0.684s). É de referir que no estudo de Range e Fischer (2004) a duração média dos gritos dos primatas foi de 0.8 s, e que a mesma medida para os mochos-reais do estudo de Lengagne (2001) teve o valor de 0.25 s, ambos inferiores à duração média obtida para os estímulos do presente estudo.

3.4.2 ANÁLISE DA FORMA DE ONDA E ANÁLISE ESPECTRAL

Foram analisados os espectrogramas e espectros de densidade de potência de todos os estímulos. De uma maneira geral, os dados obtidos possuem características semelhantes de acordo com o grupo de sons não verbais.

3.4.2.1 SÓLIDOS

Na Figura 21 pode observar-se o exemplo da forma de onda, espectrograma e espectro de sons resultantes da interacção de materiais sólidos num *impacto*. Observa-se que os dois sons têm um início abrupto, com uma elevada amplitude. O espectrograma e a forma de onda são característicos de um sinal de ruído, sendo um som de reduzida duração. No espectro observa-se que existe energia em toda a gama de frequências. Xiaofeng et al. (1991) tomaram como referência para identificar o início do sinal correspondente a *andar* o aumento súbito da amplitude, com uma maior concentração de energia nas baixas frequências, algo que não se observou neste estudo.

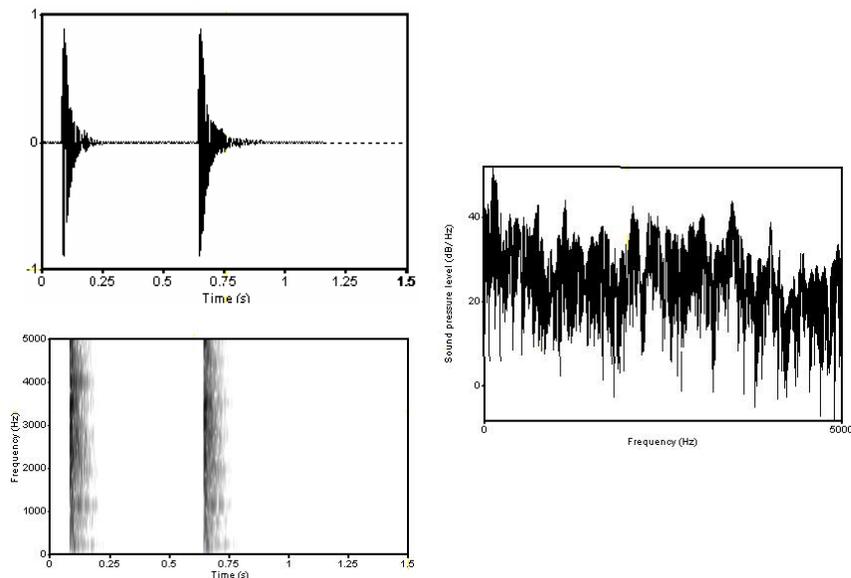


Figura 21: Forma de onda, espectrograma (à esquerda) e espectro (à direita), do som *martelar* (nº 6), pertencente ao grupo *sólidos-impacto*.

Para ilustrar um exemplo de forma de onda e espectrograma de um som produzido por material sólido em interacção numa *deformação*, foi seleccionado o som *tambor*, apresentado na Figura 22. O som *tambor* é constituído por quatro batidas consecutivas. Para cada um destes sons no início há distribuição de energia por toda a gama de frequências. Ao contrário do estímulo *martelar* não se observa ausência de som entre as batidas, o que poderá ser explicado pelo facto

de haver uma membrana em vibração (a pele do tambor) o que gera um sinal periódico em que a amplitude decai exponencialmente (ver forma de onda na Figura 22). Henrique (1997) obteve para a caixa de rufo (ver Figura 11) uma forma de onda semelhante à do estímulo *tambor*, e um espectro com decaimento da amplitude em ambos os instrumentos musicais (caixa de rufo e bombo). Os mesmos resultados podem ser observados no nosso estímulo (ver Figura 22).

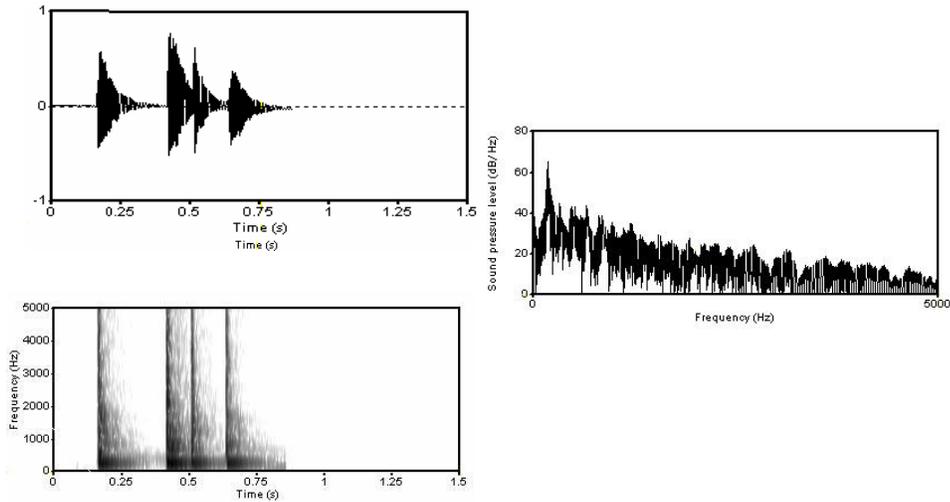


Figura 22: Forma de onda, espectrograma (à esquerda) e espectro (à direita), do som *tambor* (nº 10), pertencente ao grupo *sólidos-impacto*.

Um estímulo classificado como resultante de uma interacção do tipo *raspar* entre sólidos apresenta uma forma de onda e espectrograma exemplificados na Figura 23. O estímulo de serrar possui três repetições de duas fases distintas, que correspondem ao movimento do serrote para a frente e para trás. Nos dois primeiros pares, os valores de intensidade são mais altos no movimento para a frente e parece haver pouca energia nas frequências mais baixas. Existe sempre som entre cada movimento e a energia cobre toda a gama de frequências.

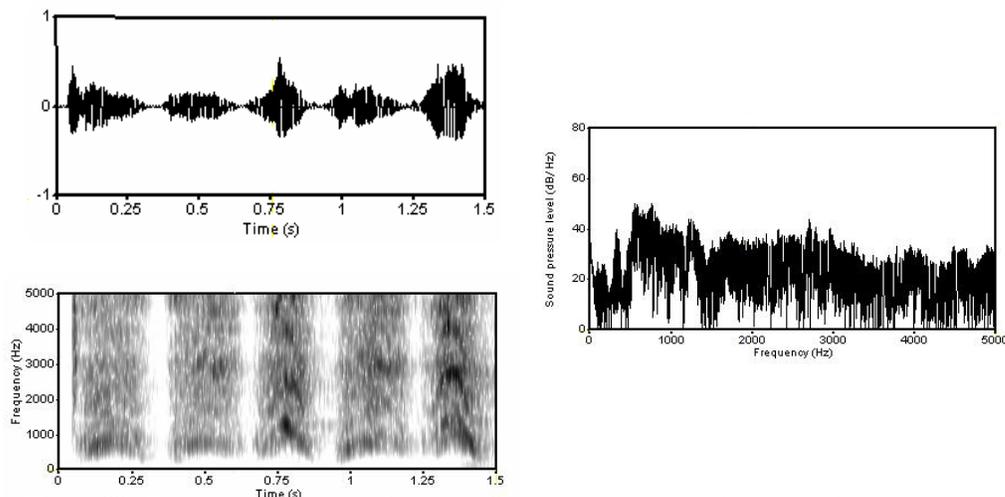


Figura 23: Forma de onda, espectrograma (à esquerda) e espectro (à direita), do som *serrar* (nº 16), pertencente ao grupo *sólidos-raspar*.

3.4.2.2 LÍQUIDOS

Os sons produzidos por líquidos são resultado de interações mais contínuas do que os sons de sólidos. O som de *encher um copo de água* (nº 18) é o único apenas classificado com uma interação do tipo *pingar*, daí que sejam apresentadas na Figura 24 a sua forma de onda, espectrograma e espectro para ilustrar este tipo de interação. Observa-se um som contínuo, mas com amplitudes que crescem e diminuem rapidamente ao longo do estímulo. A energia sonora cobre todas as frequências.

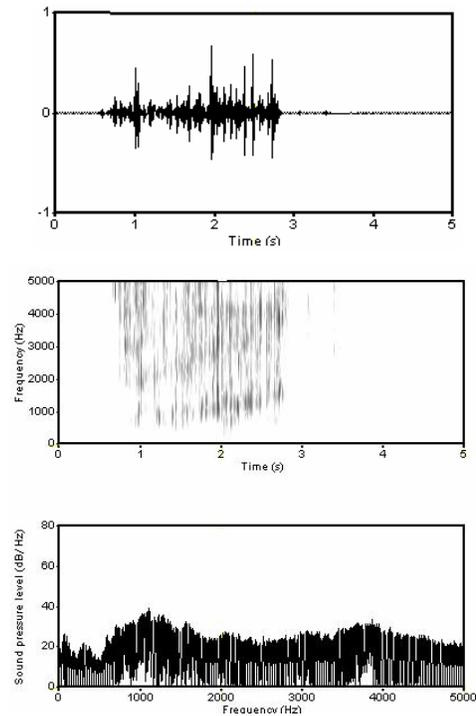


Figura 24: Forma de onda, espectrograma e espectro do som *encher um copo de água* (nº 18), pertencente ao grupo *líquidos-pingar*.

O estímulo *mar* inclui-se no grupo dos sons produzidos por líquidos numa interação “*splash*”, sendo o único estímulo que envolve apenas esta interação. É possível observar na Figura 25 um som contínuo, com amplitude crescente e decrescente, associada ao movimento da onda na direcção da praia, a “desfazer-se” na areia e a recuar para o mar, sendo as variações de amplitude graduais. A potência do sinal tem valores entre 20 e 50 dB para as frequências até 3500 Hz.

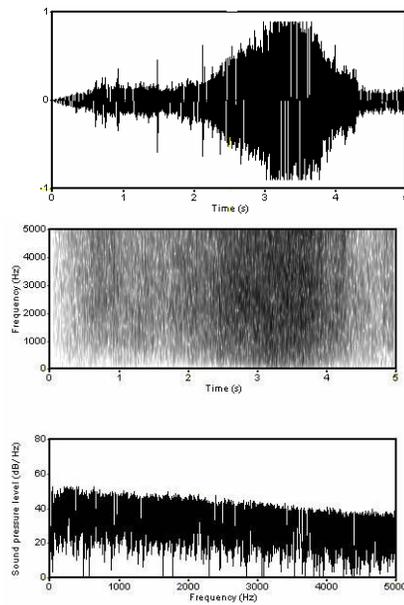


Figura 25: Forma de onda, espectrograma e espectro do som *mar* (nº 22), do grupo *líquidos-“splash”*.

Como no corpus de líquidos existem três estímulos que envolvem a interação de *pingar* e *“splash”*, será apresentado um exemplo destes estímulos na Figura 26. Os sons que resultam de *pingar* e *“splash”* também apresentam forma de onda e espectrograma característicos de um estímulo contínuo, mas com amplitude mais constante do que os estímulos de líquidos já referidos. Para o exemplo da *chuva*, apresentado na Figura 26, a energia sonora encontra-se distribuída por todas as frequências, a partir de 500-700 Hz.

Por último, o som *rio* representa uma interação *ondular*, e possui a forma de onda e espectrograma apresentados da Figura 27. Verifica-se que também o estímulo resultante de *ondular* possui características de um som contínuo. A amplitude apresenta variações ao longo de todo o estímulo, tal como se observou na interação *pingar*. A energia sonora cobre com a mesma amplitude todas as frequências a partir dos 500 Hz, tal como se observou com o estímulo *chuva*.

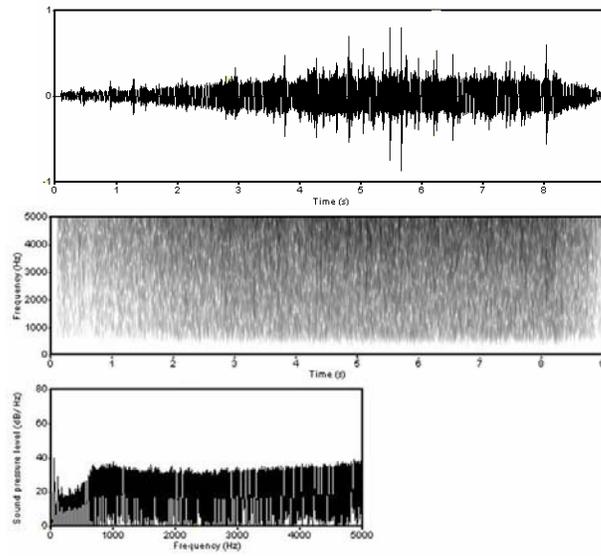


Figura 26: Forma de onda, espectrograma e espectro do som *chuva* (nº 19), pertencente ao grupo *líquidos-pingar* e “*splash*”.

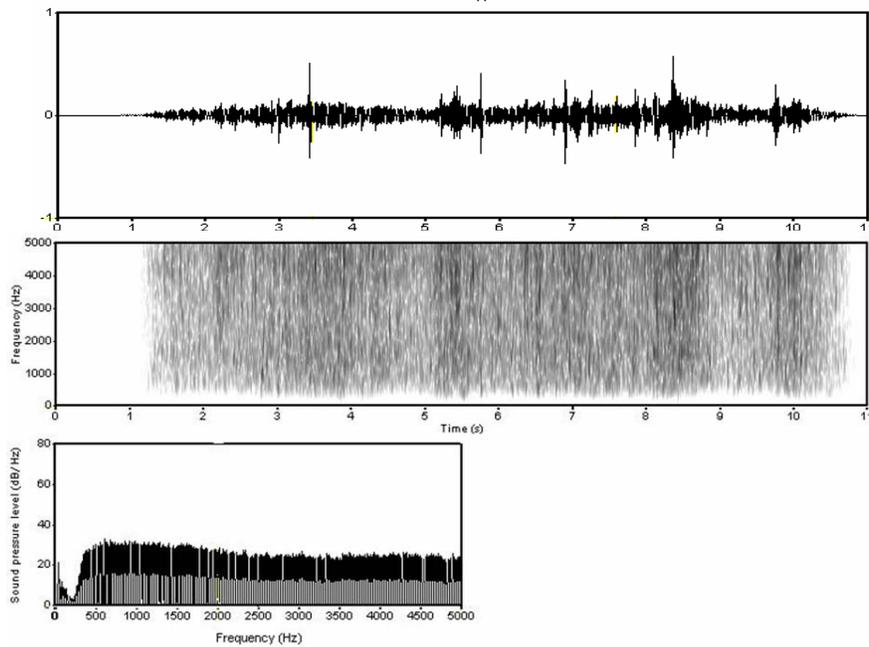


Figura 27: Forma de onda, espectrograma e espectro do som *rio* (nº 24), pertencente ao grupo *líquidos-ondular*.

3.4.2.3 GASES

Os sons que resultam da interacção de materiais gasosos podem acontecer por uma *explosão*, um “*gust*”, ou *vento*. No corpus seleccionado, foi classificado como som gerado numa *explosão* o som *trovão*. Observe-se na Figura 28 a forma de onda, espectrograma e espectro

obtidos. O som *trovão* apresenta um sinal característico de ruído, com variações de amplitude ao longo da produção. O ruído encontra-se distribuído por toda a gama de frequências, com valores mais elevados nas frequências inferiores a 1000 Hz.

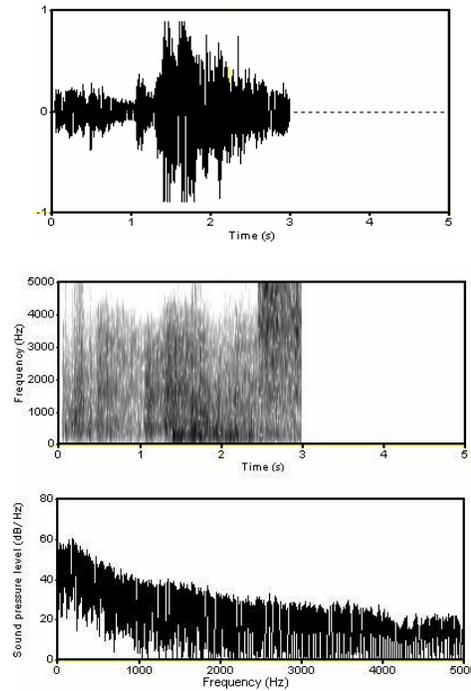


Figura 28: Forma de onda, espectrograma e espectro do som *trovão* (nº 25), pertencente ao grupo *gases-explosão*.

Os sons do corpus classificados como sendo produzidos através de um “*gust*” são *encher um pneu* e *avião*. Como se poderá observar nas Figuras 29 e 30 possuem forma de onda e espectrograma muito distintos entre si, daí que sejam ambos apresentados para ilustrar a interação “*gust*”. O estímulo *encher um pneu* é composto por quatro sons de curta duração, em que a amplitude ao longo da produção aumenta e decresce rapidamente. O som *avião* traduz-se num estímulo contínuo, onde a amplitude apresenta pequenas variações e cresce gradualmente a partir dos 5.8 s, altura em que o avião se aproxima do local onde está a ser recolhido o sinal acústico, diminuindo com o afastamento do avião. Em ambos os estímulos a energia sonora cobre toda a gama de frequências. No som *encher um pneu* há um crescimento exponencial, sendo os valores mais levados acima dos 3000 Hz, enquanto que no *avião* observa-se que é nas frequências inferiores a 1000 Hz que há maior concentração de energia.

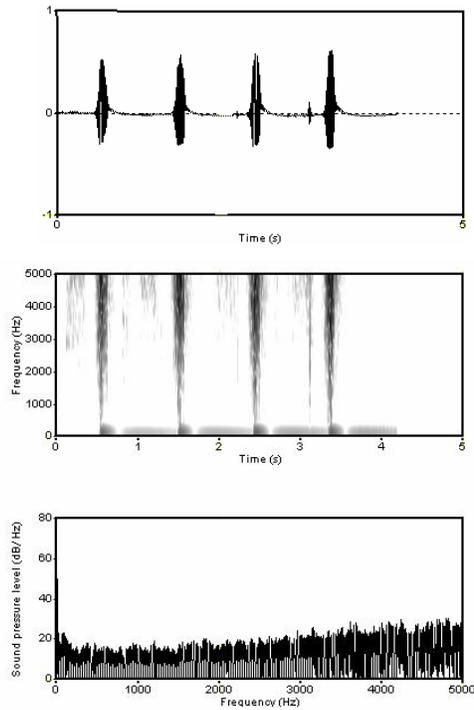


Figura 29: Forma de onda, espectrograma e espectro do som *encher um pneu* (nº 26), pertencente ao grupo *gases-“gust”*.

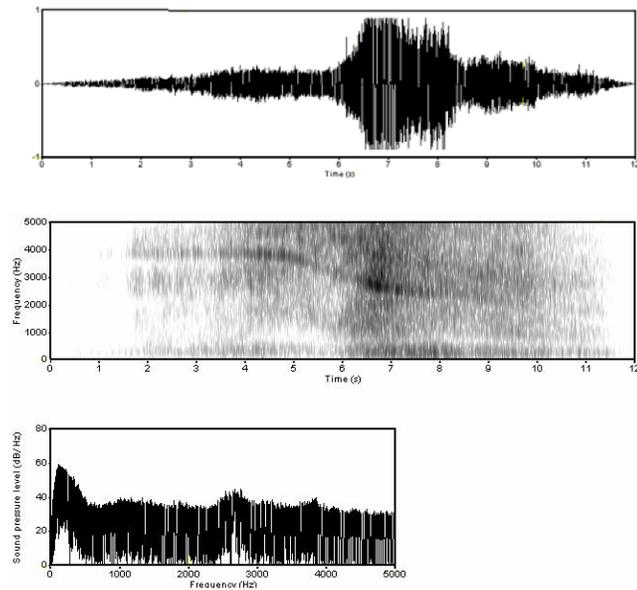


Figura 30: Forma de onda, espectrograma e espectro do som *avião* (nº 27), pertencente ao grupo *gases-“gust”*.

Foram incluídos na categoria *gases-vento* os estímulos *assobiar* e *vento*. Na Figura 31 pode observar-se que o estímulo *vento* apresenta um sinal contínuo, cuja amplitude aumenta

gradualmente, com pequenas variações, e se mantém com valores elevados durante a sua produção. A energia está concentrada na gama de frequências inferior a 2000 Hz, com valores mais elevados até aos 500 Hz.

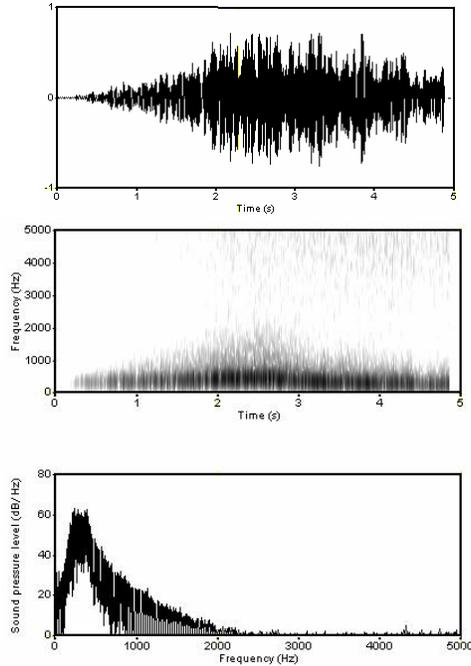


Figura 31: Forma de onda, espectrograma e espectro do som *vento* (nº 29), pertencente ao grupo gases-*vento*.

No espectrograma e forma de onda do estímulo *assobiar* (Figura 32) também se observou a existência de uma gama de frequências onde a energia está concentrada (entre os 1000 Hz e os 2000 Hz).

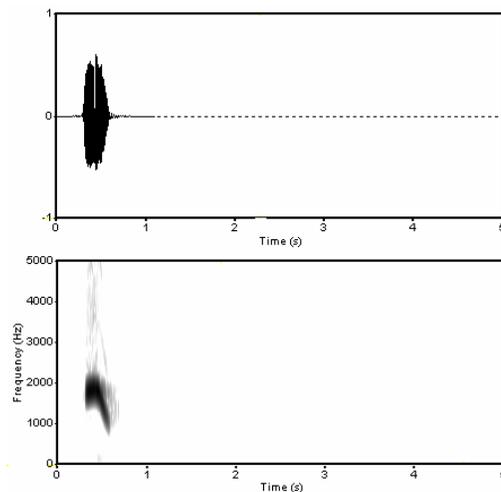


Figura 32: Forma de onda e espectrograma do som *assobiar* (nº 28), pertencente ao grupo gases-*vento*.

3.4.2.4 ANIMAIS

Nesta secção serão apresentados a forma de onda, espectrograma e espectro de sons produzidos por animais de cada um dos três grupos: mamíferos, aves e insectos.

Na Figura 33 observa-se a forma de onda e espectrograma do estímulo *vaca*. O som apresenta um sinal contínuo e com uma amplitude variável. Os estímulos correspondentes ao *cão* e ao *gato*, apresentam características em comum ao som da *vaca*, nomeadamente o sinal contínuo (de maior duração para o *gato* e *vaca*, do que para o *cão*) e a amplitude média elevada. Quanto à gama de frequências com maior energia, observou-se que no som *gato* corresponde aos valores superiores a 2000 Hz, enquanto que no som *cão* aos valores entre os 100 e os 3000 Hz. Em vários estudos em que foram analisadas vocalizações de mamíferos (Riede et al. 2001, Ikeda et al. 2003, Garstang 2004, Katada et al. 2004, Range e Fischer 2004, Rousseau et al. 2006) é referido que estes sons apresentam características periódicas e aperiódicas, e para alguns animais é possível observar a existência de harmónicos.

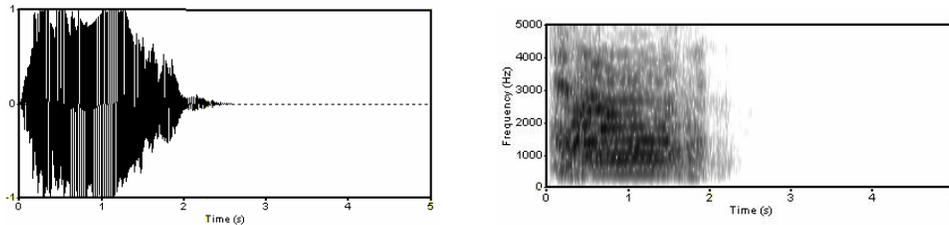


Figura 33: Forma de onda e espectrograma do som *vaca* (nº 30), pertencente ao grupo animais.

No corpus foram incluídos três estímulos produzidos por aves: galo, pássaro e pato. Na Figura 34 é apresentada a forma de onda, espectrograma e espectro do som *pássaro*. Os três estímulos produzidos por aves partilham uma amplitude média elevada durante a produção do estímulo, sendo que no *pato* e no *pássaro* ocorrem interrupções entre os sons de curta duração, enquanto que no *galo* tal só acontece antes da produção de maior duração. Os três sons deste grupo apresentam uma reduzida energia na gama de frequências mais baixas, sendo que para o *pássaro* a energia está concentrada nos valores acima de 2000 Hz, para o *pato* na gama de frequências acima dos 1000 Hz e no *galo* nos valores superiores a 500 Hz. Também no estudo de Lengagne (2001) as produções das corujas foram de curta duração e com uma amplitude que ascende e decresce, com valores distintos de frequência mínima (inicial e final) e frequência máxima.

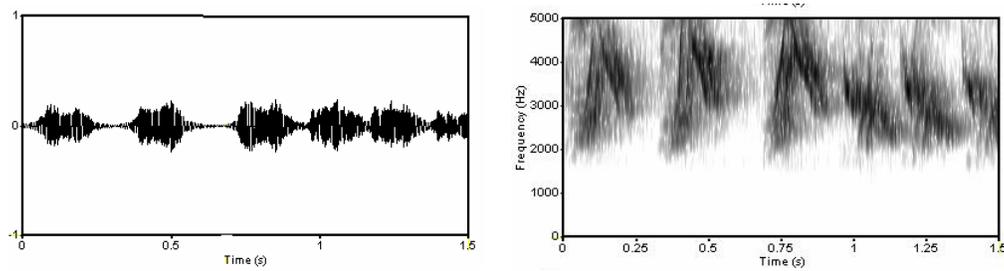


Figura 34: Forma de onda e espectrograma do som *pássaro* (nº 36), pertencente ao grupo animais.

O *grilo* é o único som produzido por um insecto incluído no corpus. Na Figura 35 observa-se que o som do grilo é constituído por 6 a 7 estímulos de curta duração e consecutivos, nos quais a amplitude se mantém constante, o que também foi observado por Brown et al. (2006) que descreve o som do grilo como muito repetitivos. É clara a existência de uma frequência de ressonância de cerca de 3000 Hz que decresce até aos 2250 Hz ao longo de 30 ms.

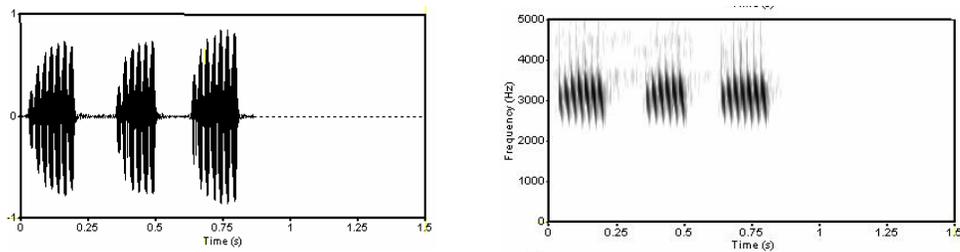


Figura 35: Forma de onda e espectrograma do som *grilo* (nº 35), pertencente ao grupo animais.

No grupo de estímulos denominado *animais*, observa-se que a maior parte dos estímulos possui valores de F0 médio entre os 250-450 Hz (ver Figura 36). No entanto, é de realçar que para os estímulos nº 32, nº 33 e nº 34 (sons do animal galo, cão e pato, respectivamente), há uma grande diferença entre o F0 mínimo e o F0 máximo.

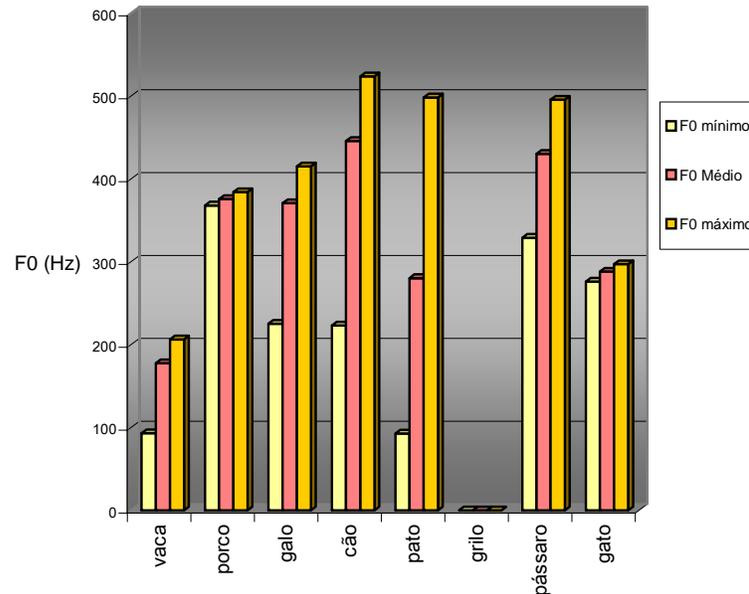


Figura 36: Valores de F0 médio, mínimo e máximo dos estímulos produzidos por animais.

3.5 SUMÁRIO

Neste capítulo verificou-se ser prática em estudos com sons não verbais realizar a análise acústica dos estímulos. A duração e a amplitude são medidas obtidas na maior parte dos estudos, independentemente do tipo de som não verbal analisado. Os valores das ressonâncias são obtidas principalmente em estudos com sons de animais. Além disso, também se verificou nos vários estudos a obtenção da forma de onda, espectro e espectrograma para análise dos estímulos.

Assim, depois de descrever os procedimentos utilizados para a análise da base de dados foram apresentados os resultados da análise temporal e espectral, procurando-se verificar a existência de características acústicas comuns entre grupos de sons não verbais.

Em relação à análise temporal, observou-se que os sons originados por sólidos apresentam a duração média menor, principalmente os do tipo impacto e deformação, verificando-se uma maior duração nos do tipo raspar. Os líquidos e gases são os que possuem durações médias maiores, o que estará relacionado com o facto de a maior parte das interações deste grupo resultarem em sons mais contínuos do que no caso dos sólidos. Por sua vez, os sons produzidos por animais também têm durações elevadas.

Quanto à análise espectral, foram obtidos o espectrograma e espectro dos sons não verbais e apresentados exemplos para cada tipo de interação dos diferentes grupos. Foram

observadas características comuns aos elementos de cada grupo e os dados obtidos foram concordantes com a bibliografia consultada. O início dos estímulos produzidos por *sólidos* em *impacto* ou *deformação* apresenta uma maior amplitude que decai exponencialmente e têm uma curta duração, tal como observaram diversos estudos (Warren e Verbrugge 1984, Repp 1987, Xiaofeng et al. 1991, Henrique 1997). Por sua vez, os sons resultantes de *raspar* apresentaram um sinal contínuo. Todos os estímulos gerados por líquidos apresentaram características de um sinal contínuo com poucas variações da amplitude ao longo do estímulo e com energia sonora em toda a gama de frequências, mas com maior amplitude acima dos 500 Hz para as interações do tipo *pingar* e *ondular*. Quanto ao grupo dos *gases*, observou-se que na maioria das interações a amplitude do espectro era maior nas frequências inferiores a 1000 Hz e que também apresentavam características de um sinal de ruído. Os *sons de animais* apresentaram uma estrutura periódica, excepto o grilo, sendo que nos mamíferos também se observou a existência de sinal aperiódico. Nos mamíferos *vaca* e *cão* observou-se uma maior acumulação de energia em frequências inferiores a 3000 Hz, enquanto que nas aves a energia sonora predominou acima dos 2000 Hz. No insecto *grilo* verificou-se serem produzidos sons de curta duração consecutivos e com uma gama de frequências restrita aos valores entre 2500 e 3500 Hz.

CAPÍTULO 4: PERCEPÇÃO DE SONS NÃO VERBAIS E PRODUÇÃO DE ONOMATOPEIAS

4.1 INTRODUÇÃO

Neste capítulo são apresentados os resultados da primeira experiência de percepção, na qual foi pedido aos sujeitos para utilizarem onomatopeias para descrever os sons não verbais. No início é apresentada uma breve descrição da fala, e como os sons que a constituem se caracterizam e são classificados. De seguida são apresentados estudos que relacionam a fala e os sons não verbais, e é abordada a especificidade das onomatopeias como sons da fala usados para referir sons não verbais. Por último é apresentada a experiência em si, descrevendo-se os procedimentos e analisando-se os resultados.

4.2 SONS DA FALA

As vogais são produzidas quando há vibração das pregas vocais, sem que ocorram constrictões significativas no tracto vocal, o que lhes confere ressonâncias visíveis na representação em espectrograma. Modificações nas configurações do tracto vocal produzem as diferentes vogais (Velooso 1999, Mateus et al. 2005). A principal dificuldade na descrição das vogais prende-se com o facto de não haverem fronteiras distintas entre um tipo de vogais e o outro, sendo possível formar um contínuo entre os sons de vogais (Ladefoged 2006). Quando a língua muda de posição entre as vogais, há uma alteração na qualidade auditiva da vogal. No entanto, é difícil saber precisamente a configuração do sistema de produção de fala, daí que sejam simplesmente aplicados “rótulos” (alta, baixa, anterior ou posterior) às vogais para distingui-las quanto às suas qualidades auditivas, tais como os que podemos observar no espaço de vogais representado na Figura 37. É de realçar que os “rótulos” utilizados não devem ser associados a posições da língua, aos articuladores, mas a qualidades auditivas de percepção.

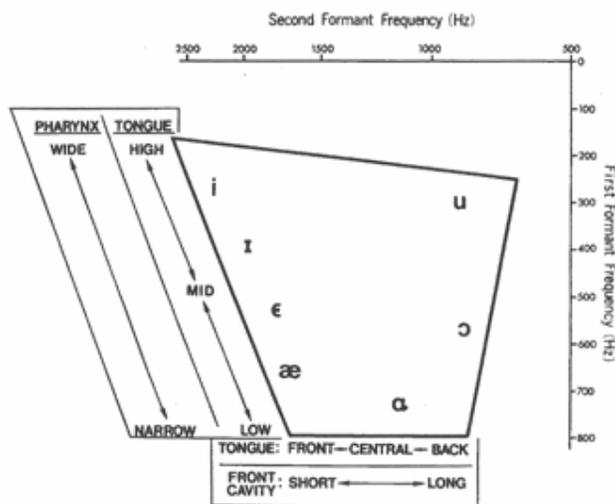


Figura 37: Espaço de algumas vogais do Inglês Americano com representação acústica tradicional. De Borden et al. (2003).

As principais características acústicas a ter em atenção ao estudar as vogais são: a frequência das formantes, a largura de banda das formantes, a amplitude das formantes, a duração e a F0. Caracterizar uma vogal como sendo “alta” está mais relacionado com a frequência da primeira formante (F1) do que com a altura da língua (Ladefoged 2006). A dimensão antero-posterior é complexa de associar a determinados valores de formantes. A segunda formante (F2) é afectada quer pela posteriorização do corpo da língua quer pelo arredondamento dos lábios. O grau de posteriorização da língua está relacionado com a diferença entre a frequência da primeira e da segunda formante: quanto menor é a diferença mais recuada soa a vogal. A dimensão (perceptual) antero-posterior relaciona-se com a segunda e terceira formantes; a dimensão relativa à altura, influencia a F1; e a dimensão relativa à tensão, correlaciona-se com a duração. De uma forma geral, vogais recuadas possuem uma frequência de F2 baixa (ver Figuras 37 e 38) e normalmente uma diferença pequena entre F1 e F2, enquanto que vogais anteriores têm uma F2 relativamente alta e uma grande diferença entre F1 e F2. Vogais baixas, possuem uma frequência de F1 alta e vogais altas uma F1 baixa (Ladefoged 2006).

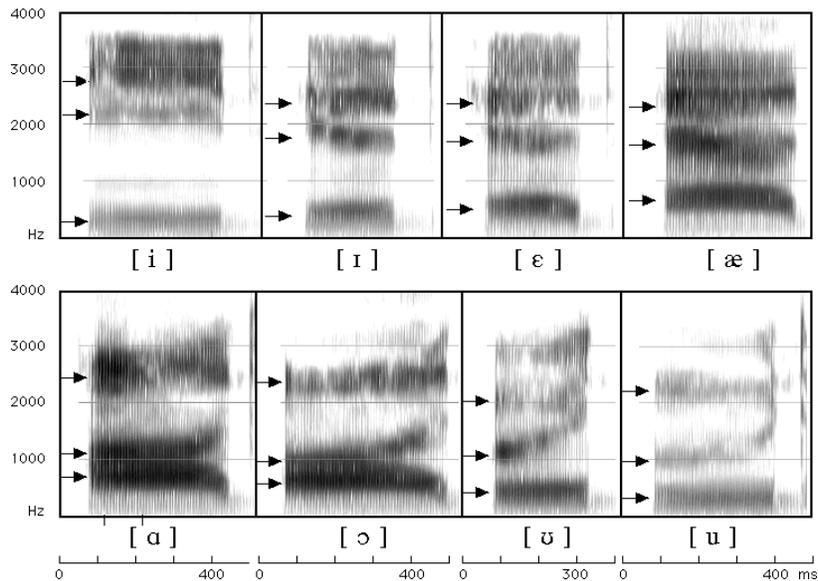


Figura 38: Espectrogramas das vogais /i, ɪ, ɛ, æ, ɑ, ɔ, ʊ, u/ de um falante de Inglês Americano. As setas indicam o valor das primeiras três formantes. De Ladefoged (2006, p.186).

Para as vogais do Português Europeu (PE) não há consenso na literatura quanto aos valores típicos de F1 e F2, pelo que se apresentam na Figura 38 os valores típicos propostos por Ladefoged (2006). Na Tabela 4 podem observar-se exemplos de palavras que contêm as vogais orais e nasais do PE, e na Figura 39 o espaço das vogais (Cruz-Ferreira 1999).

Tabela 4: Exemplos de palavras com as vogais orais e nasais do PE. De Cruz-Ferreira (1999).

Vogais Oraís			Vogais Nasais		
/i/	[vi]	<vi>	/ĩ/	[vĩ]	<vim>
/e/	[ve]	<vê>	/ẽ/	[vẽ]	<vem>
/ɛ/	[sɛ]	<sé>	/ã/	[ãtru]	<antro>
/a/	[va]	<vá>	/õ/	[sõ]	<som>
/ɔ/	[sɔ]	<só>	/ũ/	[mũdu]	<mundo>
/o/	[ˈsopɐ]	<sopa>			
/u/	[ˈmudu]	<mudo>			
/e/	[pɐˈgar]	<pagar>			
/u/	[puˈgar]	<pegar>			

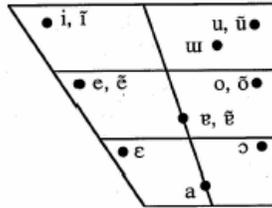


Figura 39: Espaço das vogais do PE. De Cruz-Ferreira (1999).

Os ditongos são sons dinâmicos cuja configuração (e padrão de formantes) varia lentamente durante a sua produção (Kent e Read 2002), tal como pode ser observado na Figura 40. Ditongos que terminam em /i/, tais como /ai/, /ɛi/, /ɔi/, são acompanhados por um movimento anterior de elevação da língua, acima das posições do /a/, /ɛ/ e /ɔ/. Por sua vez os que terminam em /u/, tais como /au/ e /ou/, envolvem um movimento de retracção e elevação da língua, em simultâneo com uma protrusão labial (Borden et al. 2003). No PE existem 14 ditongos, 10 orais e 4 nasais, representados na Tabela 5.

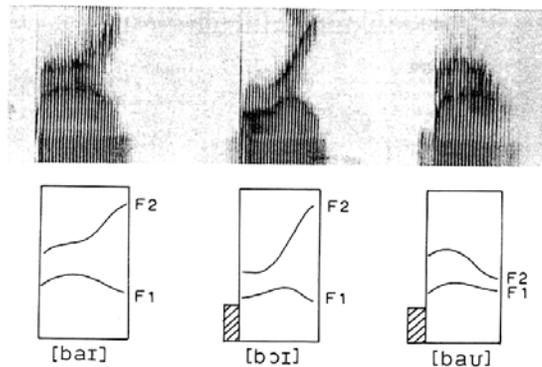


Figura 40: Espectrogramas e padrões de F1-F2 obtidos para as palavras [bai], [bɔi] e [bau]. De Kent e Read (2002).

Tabela 5: Exemplos de palavras com os ditongos orais e nasais do PE. De Cruz-Ferreira (1999).

Ditongos Orais			Ditongos Nasais		
ɛi	[ɛ'nɛi]	<anéis>	ɛi	[sɛi]	<cem>
ai	[sai]	<sai>	õi	[ɐnõi]	<anões>
ɐi	[sɐi]	<sei>	ũi	[muĩte]	<muita>
ɔi	[mɔi]	<mói>	ãu	[mãu]	<mão>
oi	[moite]	<moita>			
ui	[ɐ'nui]	<anuis>			
iu	[viu]	<viu>			
eu	[meu]	<meu>			
ɛu	[sɛu]	<céu>			
au	[mau]	<mau>			

As consoantes são sons que se produzem com uma constrição no tracto vocal, a qual pode ser de vários graus e acontecer em pontos distintos (Mateus et al. 2005). A estrutura acústica das consoantes é mais complexa do que a das vogais. Na Tabela 6 são descritas as consoantes do PE e exemplos da sua utilização. A classificação das consoantes de acordo com o International Phonetic Alphabet (IPA) está representada na Tabela 7.

Tabela 6: Consoantes do Português Europeu e exemplos de palavras onde ocorrem. De Cruz-Ferreira (1999).

/p/	[ˈpatu]	<pato>	/r/	[ˈpire]	<pira>
/b/	[ˈbatu]	<bato>	/l/	[ˈliɲe]	<linha>
/m/	[ˈmatu]	<mato>	/k/	[ˈkatu]	<cacto>
/f/	[ˈfatu]	<fato>	/g/	[ˈgato]	<gato>
/v/	[ˈviɲe]	<vinha>	/ɲ/	[ˈpiɲe]	<pinha>
/t/	[ˈtatu]	<tacto>	/ʃ/	[ˈʃatu]	<chato>
/d/	[ˈdatu]	<dato>	/ʒ/	[ˈʒatu]	<zatu>
/n/	[ˈnatu]	<nato>	/ʁ/	[ˈʁatu]	<rato>
/s/	[ˈsaie]	<saia>	/ʎ/	[ˈpiʎe]	<pilha>
/z/	[ˈkazu]	<caso>			

Tabela 7: Classificação das consoantes do Português Europeu. De Cruz-Ferreira (1999).

	Bilabial	Lábio-dental	Dental	Alveolar	Palato-alveolar	Palatal	Velar	Uvular
Oclusiva	b p		d t				g k	
Nasal	m		n			ɲ		
Fricativa		v f		z s	ʒ ʃ			ʁ
Vibrante				r				
Aproximante Lateral			l			ʎ		

Uma oclusiva envolve um encerramento completo do tracto vocal e, dependendo do seu contexto fonético, uma libertação da oclusão e um movimento em direcção a outra configuração do tracto vocal. A oclusão está associada à aplicação de pressão intra oral, a qual é libertada quando termina a oclusão. A evidência acústica desta distensão é uma explosão ou transiente (Borden et al. 2003, Kent e Read 2002).

As consoantes nasais do português são o /m/, /n/ e /ɲ/. Para haver ressonância nasal é necessário que o véu palatino esteja descido e desencostado da faringe, para que o fluxo de ar passe pelas cavidades oral e nasal (vogais nasais), ou apenas pela cavidade nasal (consoantes nasais). Quando há uma abertura da cavidade nasal e obstrução oral observa-se uma característica acústica denominada de murmúrio nasal. As formantes da cavidade nasal dependem do comprimento da cavidade desde a úvula até às narinas (Borden et al. 2003, Kent e Read 2002).

Existem duas formas principais de produzir fricativas: o som é gerado quando o ar impinge num corpo rígido normal à direcção do fluxo; o som é gerado ao longo de uma parede rígida e paralela à passagem de ar. O primeiro tipo de fonte gera fricativas normalmente designadas por estridentes/sibilantes (e.g. do tipo alveolar /s, z/ e do tipo palato-alveolar /ʃ, ʒ/ na Figura 41) e o segundo tipo de fonte não estridentes/não sibilantes (e.g. /f, v/ na Figura 42).

Assim sendo, as fricativas resultam de uma constrição algures no tracto vocal e da formação de ruído de turbulência. As fricativas têm durações de ruído mais longas do que as consoantes oclusivas, sendo este longo intervalo de energia aperiódica que as distingue como uma classe de sons. A duração de ruído de uma fricativa varia com o ponto de articulação, sendo que em média a duração aumenta na seguinte ordem: lábio-dentais, alveolares e palato-alveolares. As fricativas sibilantes têm uma energia de ruído muito maior do que as não sibilantes, e a diferença de amplitude é considerada um factor perceptual importante.

As fricativas são ainda classificadas em vozeadas/não vozeadas. As vozeadas têm duas fontes de energia, a energia quasiperiódica da vibração das pregas vocais e a energia aperiódica do ruído de turbulência. As não vozeadas apresentam apenas uma fonte de ruído. As fricativas vozeadas possuem uma menor amplitude no intervalo de fricativação do que as correspondentes não vozeadas e durações de segmentos de ruído mais curtos do que as não vozeadas.

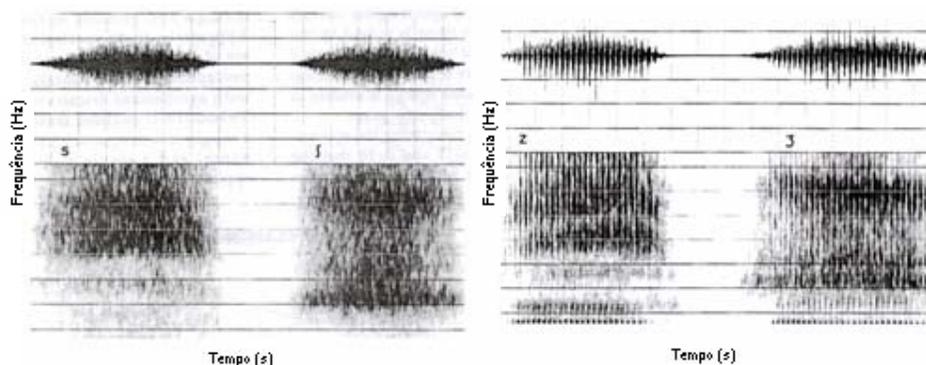


Figura 41: Forma de onda e espectrogramas para as produções isoladas das fricativas estridentes. De Kent e Read (2002).

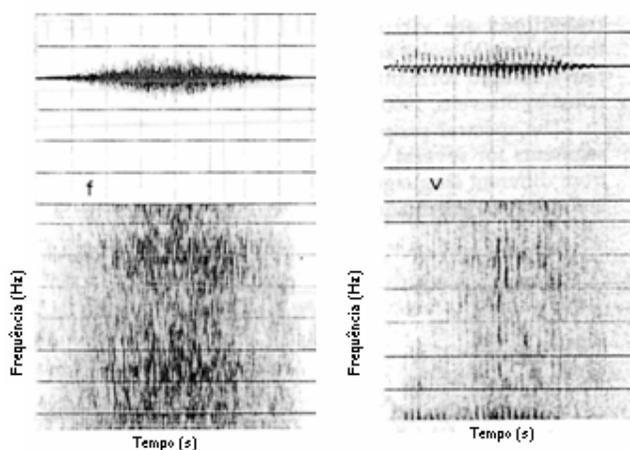


Figura 42: Formas de onda e espectrogramas para produções isoladas de fricativas não estridentes. De Kent e Read (2002).

As consoantes líquidas (/r, l, ʎ/ no Português Europeu) possuem propriedades semelhantes às vogais (padrões de formantes bem definidos) e partilham características das oclusivas (Kent and Read 2002). As líquidas /r/ e /l/ são produzidas em posição inicial de sílaba pela elevação da ponta da língua. As diferenças entre a configuração e posição da ponta da língua criam as distinções entre os dois fonemas.

4.3 SONS DA FALA E SONS NÃO VERBAIS NA AVALIAÇÃO AUDITIVA

Na intervenção ao nível da educação auditiva o Terapeuta da Fala deve considerar principalmente três características do som: *intensidade*, *frequência fundamental* e *duração*. Ao

longo das diferentes etapas do treino auditivo (Dumont 1995), os estímulos sonoros variam em complexidade pela combinação daqueles três parâmetros acústicos. Estas características estão presentes quer na fala quer nos sons ambientais. Será então de prever que hajam características comuns a ambos os tipos de estímulo (verbais e não verbais), e que o sistema cortical auditivo seja capaz de os processar e atribuir significado. Por exemplo, os sons vozeados partilham excitação periódica com certos instrumentos musicais e vocalizações de animais. Delgutte (1997) sugere que na percepção e interpretação de vocalizações específicas (e.g. alerta, chamamento, acasalamento) várias espécies animais utilizam mecanismos neuronais gerais comuns aos humanos.

Considerando três dimensões do som, a medição da audição deveria envolver a avaliação da capacidade de detectar diferenças na *intensidade*, *frequência*, e *duração* dos sons. No entanto, na audiometria convencional apenas se abrange duas dessas dimensões: intensidade e frequência (Katz 1992). Os audiogramas baseiam-se tipicamente em limiares de “estado estável” de tons puros (sons que não variam em intensidade nem em frequência). A fala é bem distinta deste tipo de sons. As características da fala são complexas e estão em constante mudança. Para além disso, os audiogramas não reflectem a experiência de aprendizagem auditiva da criança nem o potencial da mesma para aprender auditivamente. Assim sendo, é possível que duas ou mais crianças com deficiência auditiva tenham audiogramas idênticos, mas competências auditivas muito diferentes, i.e., níveis muito diferentes de perceber a fala através da audição (Ling 1989).

Monfort (2001) e Roeser (2004), utilizaram representações gráficas das características de sons não verbais, semelhantes à Figura 43, sobrepondo num audiograma, alguns sons não verbais, bem como sons verbais. Estes também foram representados em audiograma (Ling 1989, Maxon 1992) com o objectivo de permitir prever o impacto na audição de uma determinada perda ou défice auditivo, quer na percepção de sons da fala quer de sons ambientais. Existe pois uma grande necessidade de aferir a validade de tais representações.

As características acústicas dos sons da fala (vogais e consoantes) e os fenómenos de coarticulação, têm sido amplamente estudados ao longo dos últimos 50 anos (Stevens 1998). As relações entre os traços acústicos e perceptivos de sons não verbais e os traços acústicos e perceptivos dos sons da fala têm sido pouco estudados. No entanto, há evidência (Gygi et al. 2004) que os humanos têm capacidade de fácil e rapidamente identificar uma grande variedade de sons que ocorrem naturalmente e se associam a objectos e interacções no meio circundante. Devido à grande variedade de fontes sonoras, a variação espectral e temporal dos sons não verbais é superior à dos estímulos de fala (Gygi et al. 2004).

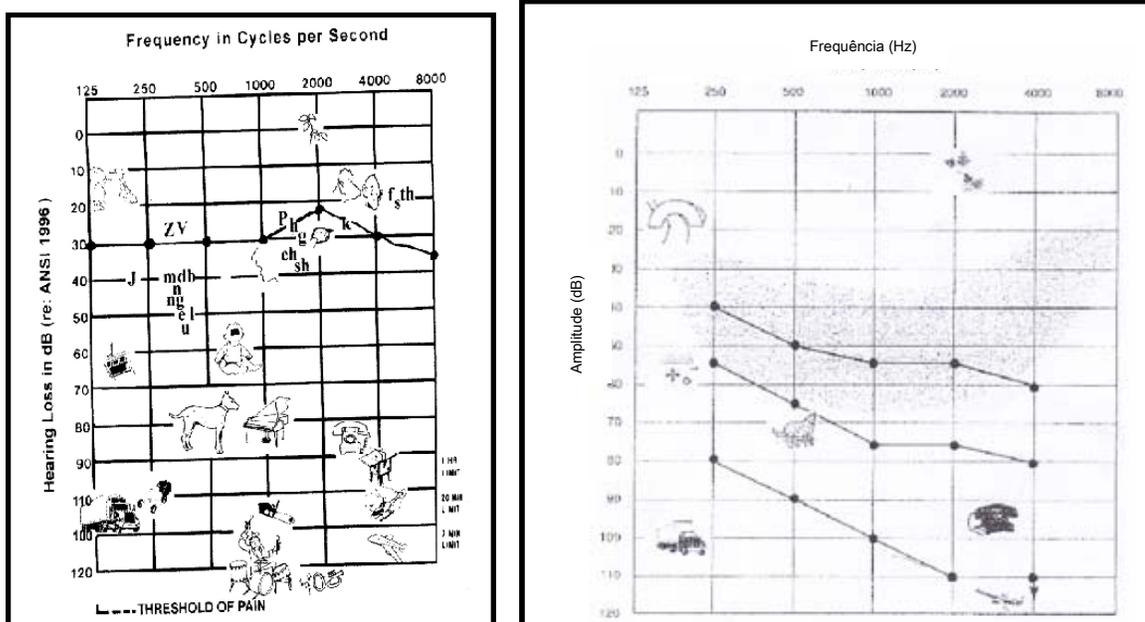


Figura 43: Representação de sons não verbais num audiograma: do lado esquerdo de Monfort (2001), do lado direito de Roeser (2004).

4.4 ONOMATOPEIAS

Uma onomatopeia é um tipo de palavra que existe em diversas línguas, e que é criada a partir da imitação de um som natural (Xavier e Mateus 1990). É diferente da imitação não linguística por ser constituída por segmentos fonéticos integrados no sistema fonológico de uma determinada língua. Nogueira (1950) propôs a classificação das onomatopeias em termos genéticos e morfológicos. A nível genético podem ser de dois tipos: onomatopeias puramente fonéticas, em que a formação tem por base a imitação fonética dos sons não verbais que representam (e.g.: trrrrim, tic-tac), e onomatopeias fonético-ideológicas que tentam imitar os sons através de vocábulos e expressões frásicas pronunciados de forma semelhante (e.g.: pouca terra pouca terra). Quanto à classificação morfológica, podem haver onomatopeias não vocabulizadas, que imitam os sons mas não se constituem como palavras da língua (e.g., pfff, ruf ruf) e onomatopeias vocabulizadas, que são palavras como outras quaisquer, cuja ortografia se deve adequar às regras da língua (e.g.: tic-tac, miau).

No estudo de Tanaka et al. (2006) foram utilizados dois tipos de metodologias de testes de percepção (Gygi 2001) - escalas de qualidade e diferencial semântico, tendo sido aplicadas para descrever o estímulo; acústica ecológica e percepção de eventos, procedendo-se à obtenção de medidas acústicas, bem como à aplicação de filtros aos estímulos para posterior realização de

novo estudo de percepção. Tal como noutros estudos, os estímulos de sons não verbais foram obtidos a partir de CDs comercializados, e a amostra incluiu sons variados comuns no dia-a-dia.

Na primeira experiência, os 8 participantes (6 homens e 2 mulheres) normo-ouvintes descreveram livremente através de onomatopeias os estímulos que ouviam. As onomatopeias foram codificadas usando-se 24 parâmetros fonéticos, que incluíam os 7 pontos de articulação, os 6 modos de articulação, as 5 vogais japonesas, consoantes vozeadas e não vozeadas, nasais silábicas, africadas, consoantes palatizadas, e vogais longas. Classificaram-se as onomatopeias pelas suas semelhanças e agruparam-se os resultados da seguinte forma:

Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5
Palma	Bebé a chorar	Telefone	Mocho	Trovão
<i>Palmas</i> de madeira	Sino	Cricrilar (som de grilo ou cigarra)	Veículo a engrenar	Comboio a passar
Fogo de artifício	Praia		Moinho de pedra	Local de obras
Chicote	Cascata		Chupar esparguete	Veículo a engrenar
Latido de cão	Buzina		Dactilografar	
Assobio	Trompete		Andar	
Pardal	Sirene		Pingar	
Microondas	Ruído branco		Gongo	
Espanta-espíritos	Vento		Bater à porta	
Tom complexo			Ribeiro	
			Câmara de filmar	

Na discussão dos resultados Tanaka et al. (2006) fizeram uma análise grupo a grupo de estímulos. Observaram que o Grupo 1 possuía a menor duração média (0.3 s). Neste grupo haviam 5 estímulos com valores elevados de energia nas frequências altas no espectro. Todos os sujeitos usaram vogal [i] na onomatopeia destes estímulos. No espectrograma da Figura 44, representativo dos estímulos do Grupo 1, observa-se uma região de elevada energia acima dos 2 kHz.

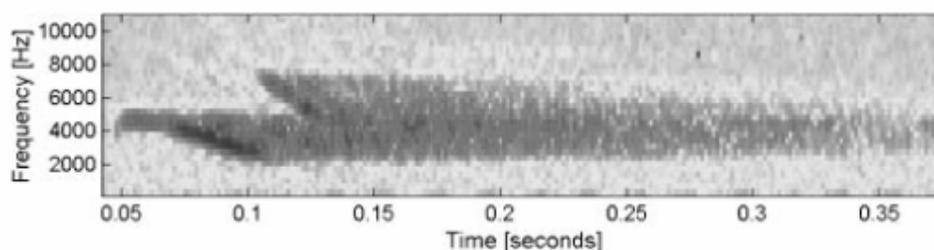


Figura 44: Espectrograma do estímulo “pardal” incluído no Grupo 1. De Tanaka et al. (Tanaka et al. 2006).

O Grupo 2 apresentou a maior duração média: 2.5 s. Estes estímulos revelaram característica temporais uniformes no seu espectrograma. As vogais sustentadas foram frequentemente utilizadas para representar estes estímulos. As vogais [a] e [u] foram mais utilizadas do que [i] ou [ɔ].

O Grupo 3 teve uma duração média de 1.2 s. Tanaka et al. (2006) referiram haver flutuações no envelope temporal, sendo os valores de modulação de 11 Hz (telefone) e 45 Hz

(grilo). As flutuações rápidas tendem a ser representadas onomatopeicamente por consoantes vibrantes múltiplas, associadas à vogal [i], no caso de haverem componentes de alta-frequência.

O Grupo 4 obteve a segunda menor duração média: 0.7 s. Os estímulos na sua maioria trataram-se de sons breves repetidos, representados nas onomatopeias por oclusivas não vozeadas.

O Grupo 5 também possuía sons breves repetidos, mas com a segunda maior duração média: 2.0 s. Foram utilizadas muitas oclusivas vozeadas para representar estes sons em onomatopeias. A energia dos estímulos estava distribuída pela gama de frequências entre os 50 Hz e os 8 kHz, mostrando um domínio de componentes de baixas frequências.

4.5 PROCEDIMENTOS

Realizou-se uma experiência de percepção, complementada com a produção de onomatopeias. Participaram 8 sujeitos, 4 homens e 4 mulheres. Efectuaram-se exames de rastreio auditivo por audiometria de tons puros, com a colaboração de um Técnico de Audiologia, antes da execução da experiência. Os resultados revelaram que todos apresentavam uma audição normal (ver audiogramas em Anexo). Foram utilizados todos os estímulos do corpus de sons não verbais (Tabela 3).

Os estímulos foram apresentados a cada ouvinte através de auscultadores Senheiser eh 1430 ligados a um computador. A ordem de apresentação dos estímulos foi aleatória relativamente aos grupos de sons. Foi pedido aos sujeitos que produzissem uma onomatopeia para descrever o som que ouviam, e forneceu-se o exemplo do som produzido por um cuco que pode ser representado pela onomatopeia [kuku].

Os sujeitos podiam ouvir o estímulo sonoro quantas vezes fosse necessário. Ao mesmo tempo, as onomatopeias produzidas oralmente pelos sujeitos eram gravadas no programa *Adobe Audition* utilizando um microfone unidireccional da marca Philips SBC ME 400 ligado a um PC através de uma placa de som Sound Blaster Live! 24 bit, com uma frequência de amostragem de 44 kHz para permitir comparar os dados de produção com os estímulos, também gravados a 44 kHz.

A realização desta experiência teve a duração média de 10 minutos, com uma duração mínima de 7 minutos e máxima de 15 minutos. Observou-se que o tipo de sons que os sujeitos normalmente ouviam repetidamente antes de produzir a onomatopeia foi o dos líquidos.

A segmentação de cada onomatopeia foi realizada pela observação do sinal acústico no *Adobe Audition* e respectiva audição dos estímulos. Realizou-se a transcrição fonética dos resultados, e procedeu-se à sua classificação segundo 15 parâmetros fonéticos que consistiram em:

- 7 pontos de articulação:
 - B: bilabial;
 - LD: lábio-dental;
 - D: dental;
 - A: alveolar;
 - P: palato-alveolar e palatal;
 - V: velar;
 - U: uvular.
- 4 modos de articulação:
 - O: oclusiva;
 - F: fricativa;
 - L: líquida (onde se incluem as vibrantes e as aproximantes laterais)
 - N: nasal.
- 2 modos de vozeamento:
 - VOZ: vozeada;
 - NVOZ: não vozeada.
- 3 grupos de vogais:
 - Grupo 1: [a, e, ε, u, e];
 - Grupo 2: [i];
 - Grupo 3: [u, o, ə].

Os parâmetros fonéticos foram definidos tendo por base o estudo de Takada et al.(2006), cujos resultados reflectiram serem estes os parâmetros mais importantes quando se caracterizam as onomatopeias usadas pelos sujeitos, em experiências de metodologia similar. Para além desta classificação fonética das onomatopeias foram analisadas formas de onda, espectrogramas e espectros de densidade de potência através dos mesmos procedimentos descritos no Capítulo 3 para o corpus de sons não verbais.

4.6 RESULTADOS

Os resultados obtidos foram analisados através dos programas *Excel* e *SPSS*. A análise dos resultados fez-se por categoria de sons não verbais, sendo construídas figuras representativas, onde o valor de 100% corresponde ao caso em que todos os sujeitos utilizaram fones caracterizados por esse parâmetro para descrever o estímulo (ver Figuras 45, 49, 54, 58 e 62).

4.6.1 SÓLIDOS

A análise dos sons deste grupo fez-se por tipo de interacção envolvida: *impacto*, *deformação* e *raspar* (Tabela 3). Pode observar-se na Figura 45 que os sons de *impacto* são descritos maioritariamente com onomatopeias que possuem consoantes oclusivas dentais não vozeadas (fone [t]), sem um tipo específico de vogais. Observe-se a Figura 46 com a forma de onda, espectrograma e espectro de uma das onomatopeias usadas pelos sujeitos para estes sons. Na análise acústica de estímulos deste grupo, que foi descrita no capítulo anterior, observaram-se características comuns com as onomatopeias obtidas nesta experiência, nomeadamente o facto de serem sons descontínuos de curta duração. As consoantes oclusivas e os sons que resultam de um *impacto* entre sólidos têm em comum a interacção entre duas estruturas em que é aplicada pressão, a qual termina quando termina essa interacção (Gaver 1993, Kent and Read 2002), sendo referido na literatura (Kent e Read 2002) que para as oclusivas a produção reduzida de energia, tal como é observado nos sons resultantes de impactos entre sólidos.

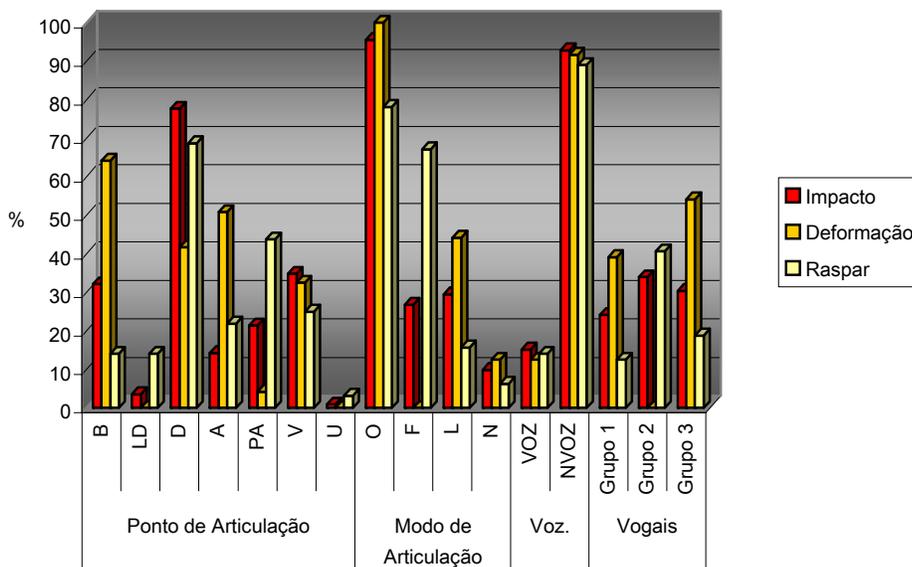


Figura 45: Distribuição dos parâmetros fonéticos pelo tipo de interacção entre sólidos.

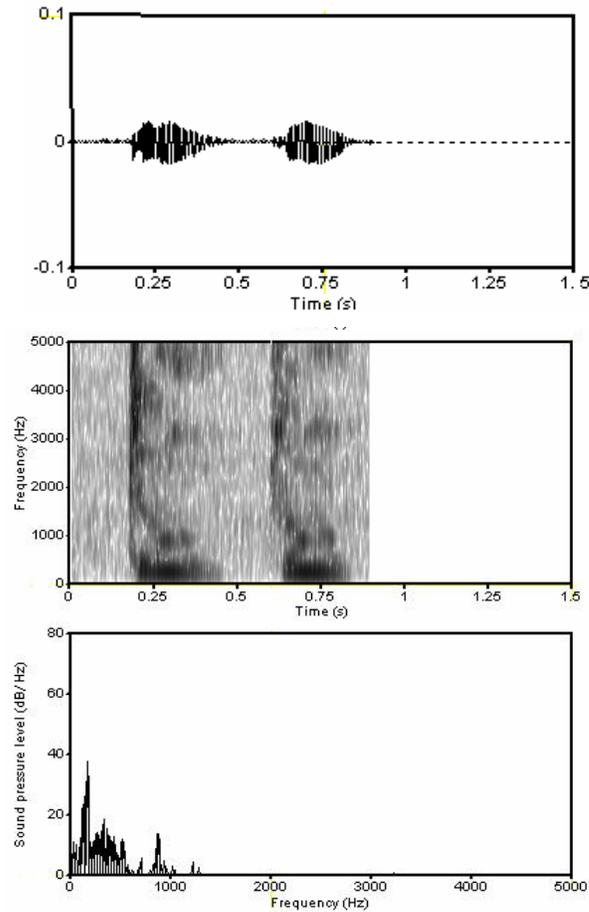


Figura 46: Forma de onda, espectrograma e espectro da onomatopeia [tütü] produzida pelo informante MS ao descrever o estímulo “martelar” (nº 6).

Aos sons do grupo *deformação* correspondem onomatopeias com consoantes oclusivas bilabiais não vozeadas (fone [p]) com vogais dos Grupos 3 e/ou Grupo 1. Na Figura 47 é visível a forma de onda, espectrograma e espectro de uma das onomatopeias usadas pelos sujeitos para descrever sons deste grupo. Mais uma vez, numa interacção entre sólidos são usadas consoantes oclusivas, os sons da fala com menor duração, e vogais de F2 baixo (Borden et al. 2003), nomeadamente valores inferiores a 1500 Hz, o que também se pode observar na Figura 47. Tal como na forma de onda e espectrograma do estímulo *tambor*, nas onomatopeias produzidas verifica-se um aumento da amplitude no início, que decresce exponencialmente, e ainda uma maior concentração de energia nas frequências inferiores a 1000 Hz.

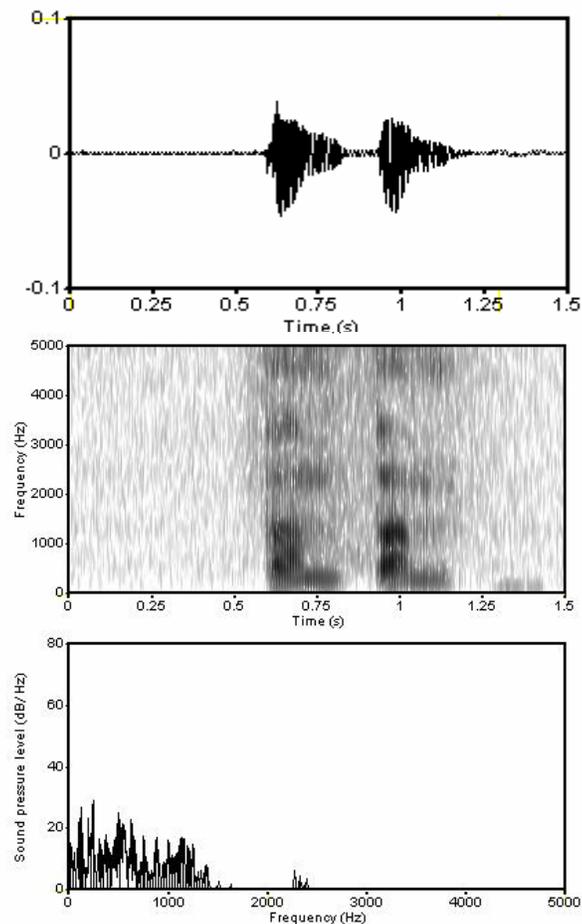


Figura 47: Forma de onda, espectrograma e espectro da onomatopeia [pãpã] produzida pelo informante RS ao descrever o estímulo *tambor* (nº 10).

Por sua vez aos estímulos que resultam da interação *raspar* associam-se onomatopeias com consoantes oclusivas dentais e fricativas palato-alveolares não vozeadas (fones [t] e [ʃ]) e vogal [i] (Figura 48). As fricativas têm durações mais longas do que as oclusivas, sendo as palato-alveolares as de maior duração (Kent e Read 2002) e as sibilantes como o [s], [z] e [ʃ] as que possuem ruído de maior energia. Na Figura 41 em que se apresentou a forma de onda e espectrograma da fricativa [ʃ], observa-se baixa energia nas frequências inferiores a 2 kHz, o que também se verifica na onomatopeia da Figura 48 e no espectrograma do estímulo *serrar* (Figura 23). Esta fricativa terá sido escolhida pelo facto de os sons que resultam de uma interação *raspar* apresentarem as maiores durações dos sons gerados por sólidos. Segundo Borden et al. (2003) a vogal /i/ é a que apresenta valores de F2 mais elevados, o que estará associado à percepção de um estímulo com energia concentrada nas frequências superiores.

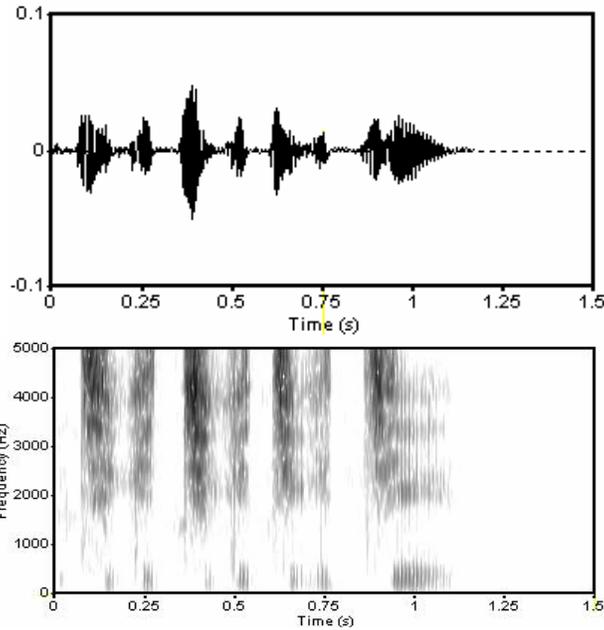


Figura 48: Forma de onda e espectrograma da onomatopeia [tʃiktʃktʃi] produzida pelo informante BR ao descrever o estímulo *esfregar as mãos* (nº 15).

4.6.2 LÍQUIDOS

De seguida serão descritas as onomatopéias usadas para descrever o grupo de sons gerados por líquidos. Na Figura 49 observa-se que os estímulos produzidos por um líquido a *pingar* são descritos com onomatopéias que possuem consoantes oclusivas dentais e fricativas palato-alveolares não vozeadas (fonos [t] e [ʃ]) com a vogal [i]. Na Figura 50 está representada uma das onomatopéias usadas pelos sujeitos na descrição de sons deste grupo. São usados sons consecutivos descontínuos, tal como se caracteriza o estímulo original (ver Figura 24), o que explica o uso de oclusivas. O facto de ser um estímulo com repetições consecutivas também resultará na percepção de um som com traços contínuos justificando a utilização das fricativas.

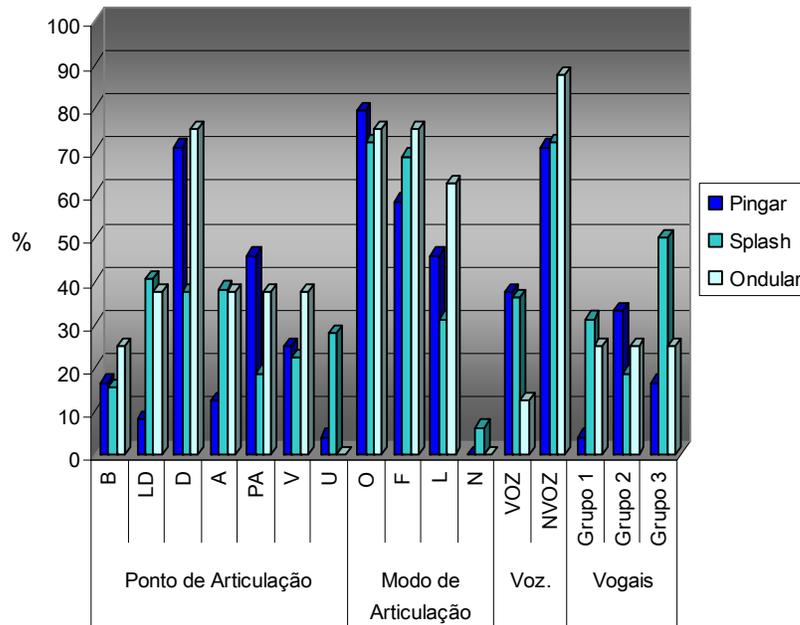


Figura 49: Distribuição dos parâmetros fonéticos pelo tipo de interacção entre líquidos.

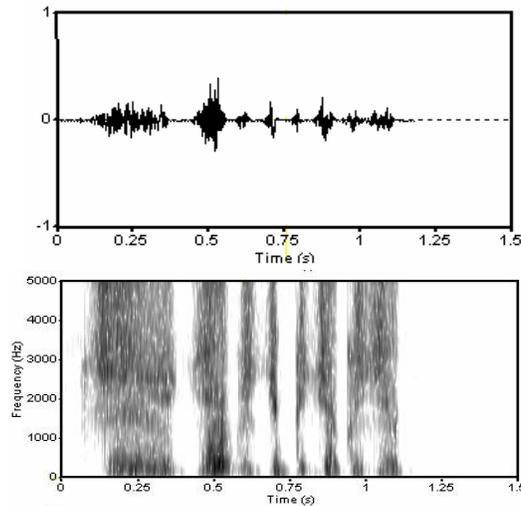


Figura 50: Forma de onda e espectrograma da onomatopoeia [tuʃt tuʃt tuʃt] produzida pelo informante MR ao descrever o estímulo *encher um copo de água* (nº 18).

Em relação às onomatopoeias usadas para descrever sons do tipo “*splash*” apresentam-se exemplos para a *chuva* e o *mar* (Figuras 51 e 52), sendo que a *chuva* é um som em que também ocorre uma interacção *pingar*. No primeiro som há concentração de energia nas frequências superiores a 2000 Hz, algo que no espectrograma e espectro do estímulo original se verifica a partir dos 1000 Hz (Figura 26). A percepção deste estímulo pelos sujeitos estará associada a informação contida nas frequências mais elevadas.

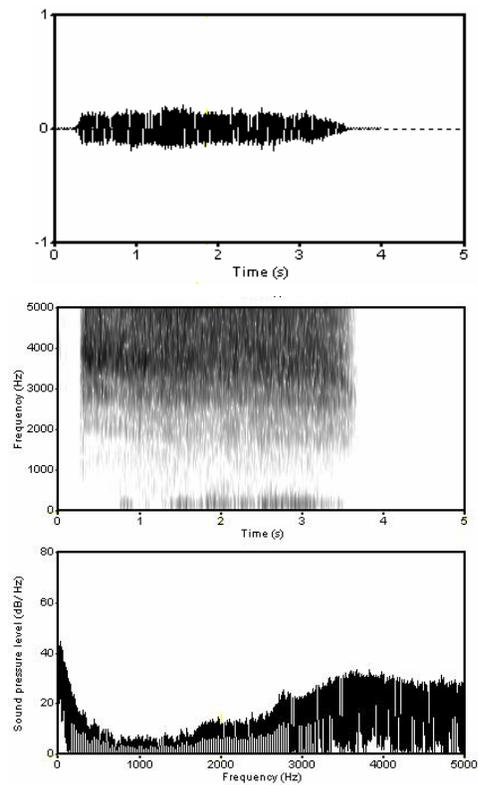


Figura 51: Forma de onda, espectrograma e espectro da onomatopeia [tʃ] produzida pelo informante MR ao descrever o estímulo *chuva* (nº 19).

Para os sons do grupo de interação “*splash*” foram usadas consoantes fricativas lábio-dentais e alveolares não vozeadas (fones [f] e [s]) e oclusivas dentais não vozeadas (fone [t]) com vogais dos grupos 3 e 1. Na forma de onda e espectrograma da Figura 52 observe-se um exemplo de onomatopeia em que foram usadas vogais de ambos os grupos.

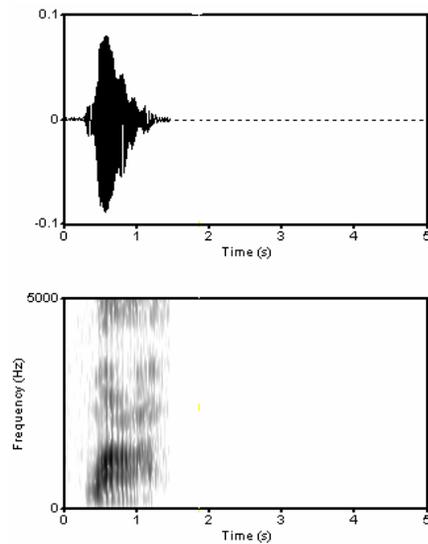


Figura 52: Forma de onda e espectrograma da onomatopeia [ua:] produzida pelo informante RS ao descrever o estímulo *mar* (nº 22).

Por último o som *rio* foi descrito com onomatopeias caracterizadas foneticamente pelo uso de consoantes não vozeadas, principalmente com o ponto de articulação dental de modo de articulação oclusiva e lateral (fones [t] e [l]) mas também com uso de fricativas lábio-dentais, alveolares e palato-alveolares e oclusivas velares (fones [f], [s], [ʃ] e [k]), respectivamente. Não se verificou o uso preferencial de nenhum grupo de vogais. Observe-se na Figura 53 características de um sinal de ruído, correspondente à fricativa não sibilante [f]. O uso da líquida [l] e da vogal [i] faz com que seja possível observar no espectrograma a presença de formantes (Kent e Read 2002), e explicará a maior energia nas frequências inferiores a 500 Hz.

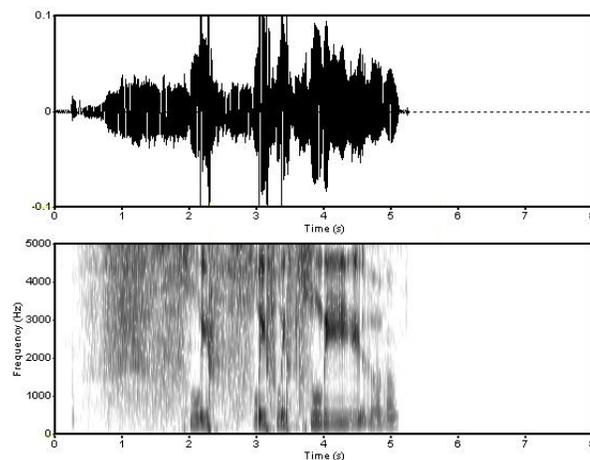


Figura 53: Forma de onda e espectrograma da onomatopeia [fliflifi] produzida pelo informante TM ao descrever o estímulo *rio* (nº 24).

4.6.3 GASES

Na Figura 54 podem observar-se os resultados obtidos para as onomatopeias que descreveram sons gerados por gases. Para os sons do grupo *gases-explosão* foram usadas onomatopeias que continham consoantes oclusivas bilabiais não vozeadas e vozeadas (fones [p] e [b]) e fricativas uvulares (fone [ʁ]) com vogais pertencentes ao grupo 3. Na Figura 55 observa-se uma das onomatopeias usadas na descrição dos estímulos deste grupo. A vogal [u] do terceiro grupo caracteriza-se por F1 e F2 de valores reduzidos, com maior acumulação de energia nas frequências mais baixas (Borden et al. 2003), o que se observa no espectro da Figura 55, e também no espectro do som *trovão* apresentado no capítulo anterior (Figura 28).

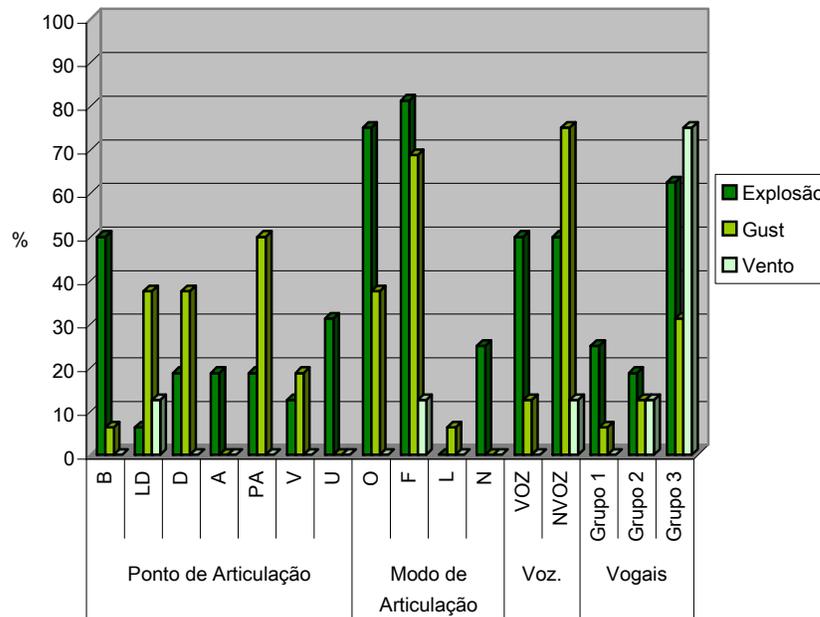


Figura 54: Distribuição dos parâmetros fonéticos pelo tipo de interação entre gases.

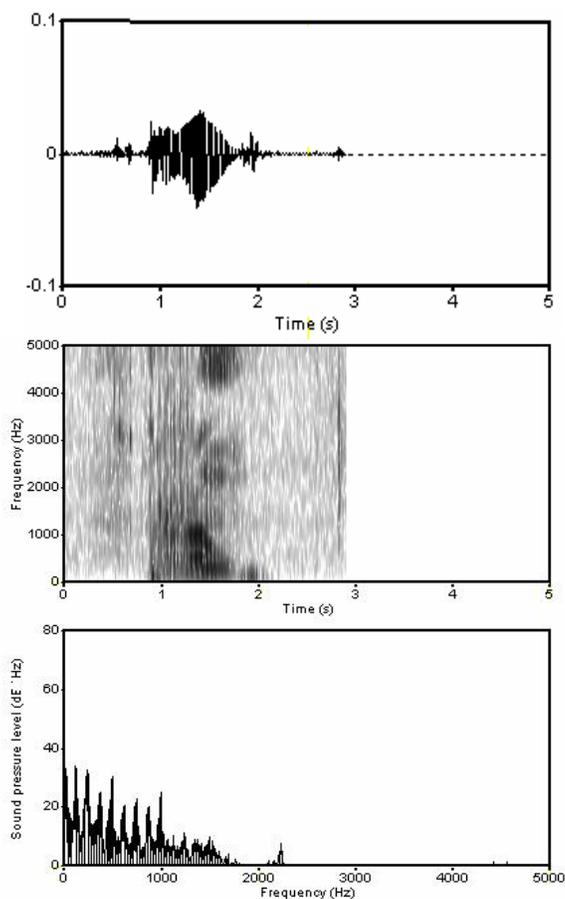


Figura 55: Forma de onda, espectrograma e espectro da onomatopeia [ʁ:ũ:] produzida pelo informante RS ao descrever o estímulo *água a ferver* (nº 23).

Relativamente aos estímulos gerados por *gases- "gust"*, as consoantes fricativas lábio-dentais e palato-alveolares não vozeadas são muito utilizadas (fones [f] e [ʃ]), assim como as oclusivas dentais não vozeadas (fone [t]), com vogais do terceiro grupo. Na Figura 56, tal como no estímulo original *encher um pneu* (Figura 29), pode observar-se na onomatopeia a produção de sons consecutivos de curta duração com maior concentração de energia nas frequências superiores a 2500 Hz.

Os sons do grupo *gases-vento* foram descritos com onomatopeias constituídas principalmente por vogais, ao contrário dos estímulos já descritos. Assim, as poucas consoantes utilizadas foram fricativas lábio-dentais não vozeadas (fone [f]), e as vogais pertencentes ao grupo 3. Na Figura 57 observe-se a forma de onda e espectrograma de uma onomatopeia para o estímulo *vento*. O uso da fricativa estará relacionado com a duração elevada deste tipo de estímulos, tal como se observou no Capítulo 3. O facto de a vogal [u] apresentar F1 e F2 baixos (Ladefoged 2006), justifica a sua utilização no sentido em que pela observação do espectrograma e espectro dos sons não verbais (Figura 31) se verifica que há uma maior concentração de energia nas frequências mais baixas.

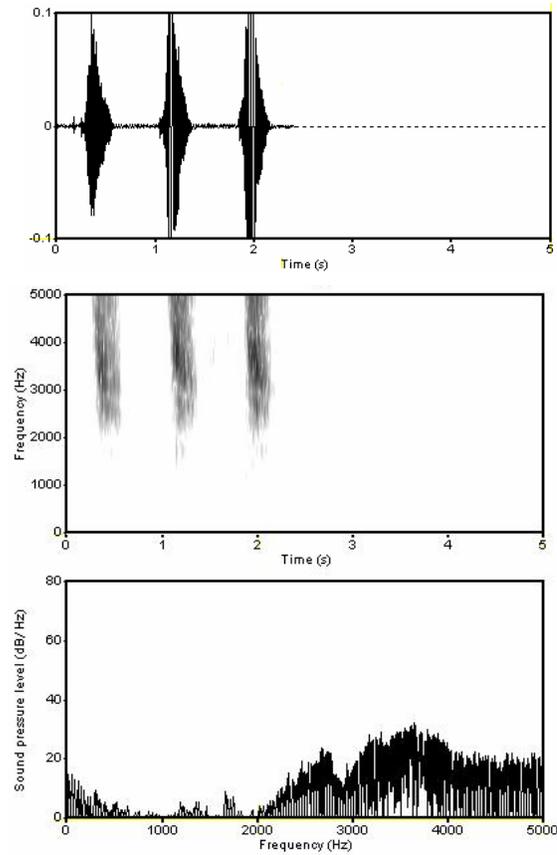


Figura 56: Forma de onda, espectrograma e espectro da onomatopeia [ft ft ft] produzida pelo informante RS ao descrever o estímulo *encher um pneu* (nº 26).

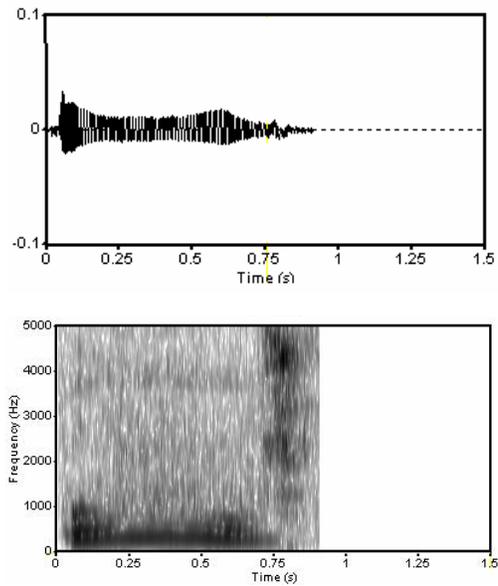


Figura 57: Forma de onda e espectrograma da onomatopeia [u:] produzida pelo informante AM ao descrever o estímulo *vento* (nº 29).

4.6.4 SONS DE ANIMAIS

Na Figura 58 apresentam-se os resultados obtidos para os sons gerados por animais. Para a classe animal *mamíferos* observa-se o uso de onomatopeias com fones de ponto de articulação bilabial e uvular, modo oclusiva e lateral, com traço de nasalidade e vozeamento (fones [m] e [ɰ]) onde são utilizadas maioritariamente vogais do terceiro grupo. Na Figura 59 podem observar-se a forma de onda e espectrograma da onomatopeia [mu:] usada por um dos informantes para descrever o som *vaca*. Em comparação com o estímulo original (Figura 33), verifica-se que também na onomatopeia há uma maior concentração de energia nas frequências mais baixas. Assim, justifica-se a escolha da oclusiva nasal [m] e da vogal [u] que apresentam ambos os valores da frequência de F1 e F2 reduzidos (Ladefoged 2006).

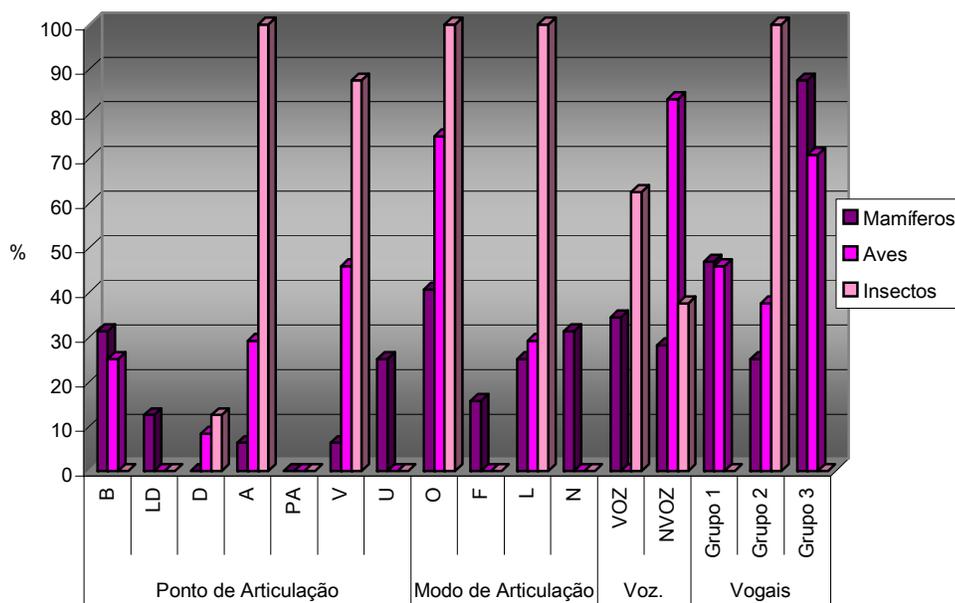


Figura 58: Distribuição dos parâmetros fonéticos pela classe animal.

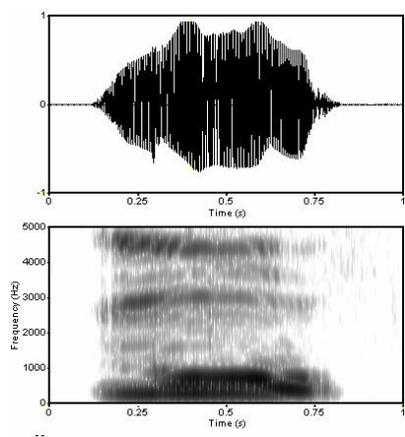


Figura 59: Forma de onda e espectrograma da onomatopeia [mu:] produzida pelo informante TM ao descrever o estímulo *vaca* (nº 30).

No grupo de aves foram usadas consoantes oclusivas velares e bilabiais não vozeadas e líquidas alveolares (fones [k], [p] e [r]). O grupo de vogais mais utilizado foi o terceiro, seguido do segundo grupo (ver Figura 58). No exemplo apresentado na Figura 60, foi utilizado o ditongo [iu], com uma frequência de F2 elevada no início e baixa no fim, observando-se no espectrograma essa transição. Também no espectrograma do estímulo original é observada uma transição, mas que é ascendente e descendente (Figura 34). Na onomatopeia são gerados sons consecutivos de curta duração, tal como nos estímulos originais.

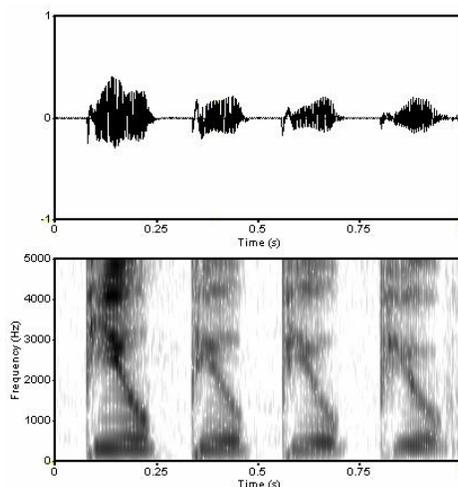


Figura 60: Forma de onda e espectrograma da onomatopeia [piupiupi] produzida pelo informante MS ao descrever o estímulo *pássaro* (nº 36).

Quanto ao estímulo *grilo*, representante da classe animal dos insectos, foram utilizadas consoantes oclusivas velares vozeadas e líquidas alveolares (fones [g] e [r]), com a vogal [i]. O uso de uma oclusiva vozeada estará relacionado com a percepção de uma estrutura harmónica

(Figura 35) observando-se no espectrograma da Figura 61 a barra de vozeamento característica (Mateus et al. 2005). Os informantes terão utilizado a vibrante [r] porque no som *grilo* há transições acústicas rápidas e consecutivas. O facto de este estímulo ter uma concentração de energia significativa entre os 2500 e os 3500 Hz, explica também o uso da vogal [i] cujos valores de F2 se situam nesta região (Ladefoged 2006).

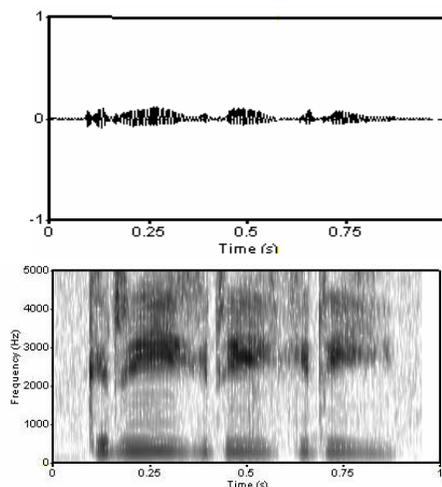


Figura 61: Forma de onda e espectrograma da onomatopeia [grigrigri] produzida pelo informante AM ao descrever o estímulo *grilo* (nº 35).

4.6.5 COMPARAÇÃO DE CATEGORIAS

Depois de ter sido realizada a análise das onomatopéias utilizadas em cada grupo de sons não verbais, comparando-se os resultados entre as interações envolvidas e a classe de animal, foi realizada uma análise dos resultados entre as categorias principais, de uma maneira mais geral.

Na Figura 62 observa-se que o ponto de articulação dental é o mais utilizado para sólidos e líquidos, enquanto que para os sons gerados por gases e por animais são utilizados vários pontos de articulação.

No que concerne o vozeamento todos os grupos foram descritos principalmente com consoantes não vozeadas, realçando-se de qualquer forma que os líquidos e os animais foram os que apresentaram mais consoantes vozeadas.

Por último, as vogais mais usadas para descrever sólidos pertencem ao Grupo 2, mas também se usaram em número não muito inferior vogais dos outros dois grupos, enquanto que para os sons gerados por líquidos, gases e animais as vogais do terceiro grupo se destacam por oposição às do primeiro e segundo grupo.

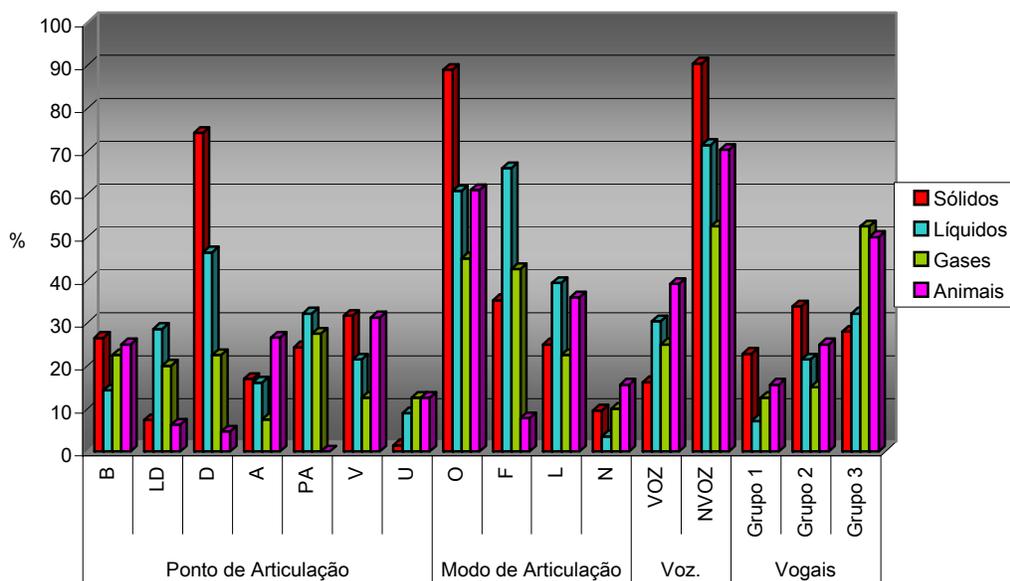


Figura 62: Distribuição dos parâmetros fonéticos pela categoria de sons não verbais.

4.7 SUMÁRIO

Neste capítulo foi descrita a experiência na qual os informantes descreveram os sons não verbais da amostra através de onomatopeias. Desta forma, pretendeu-se encontrar características comuns entre os sons não verbais e os sons da fala produzidos na caracterização da base de dados. Para conseguir analisar as onomatopeias foi necessário inicialmente realizar uma breve caracterização acústica dos fonemas do Português Europeu. Além disso, fez-se uma breve abordagem à especificidade das onomatopeias, no âmbito da sua relação próxima aos sons não verbais, e foi apresentado um estudo onde também foi realizada uma experiência semelhante à que é aqui apresentada.

Os resultados obtidos foram analisados por grupo de sons, comparando com os dados da análise acústica, e também tendo em conta 15 parâmetros fonéticos relacionados com o ponto de articulação, o modo de articulação e de vozeamento, e as vogais utilizadas.

Assim, observou-se que *sólidos-impacto* foram descritos com onomatopeias com oclusivas dentais não vozeadas, sons descontínuos de curta duração, tais como os estímulos deste grupo. Para os *sólidos-deformação* usaram oclusivas bilabiais não vozeadas e vogais como [a] e [u], que apresentam F2 baixo, tal como os estímulos deste grupo que apresentaram maior concentração de energia abaixo dos 1000 Hz. Quanto a *sólidos-raspar* os sujeitos usaram

oclusivas dentais e fricativas palato-alveolares não vozeadas, com a vogal [i], sendo de realçar que os sons deste tipo de interacção de sólidos foram os que se verificou terem maior duração, tal como acontece com as fricativas em relação às oclusivas.

Em relação aos *líquidos-pingar* foram usadas oclusivas dentais e fricativas palato-alveolares não vozeadas com a vogal [i], para os *líquidos-"splash"* fricativas lábio-dentais e alveolares e oclusivas dentais não vozeadas, e nos *líquidos-ondular* oclusivas dentais e velares e também fricativas não vozeadas. O uso de fricativas para descrever sons de líquidos independentemente da interacção estará relacionado com o facto de este grupo de sons não verbais apresentar maior duração tal como as fricativas. Tal como se observou nos sólidos, o traço de vozeamento está ausente na maioria das consoantes produzidas o que se relaciona com a ausência de uma estrutura periódica nos estímulos.

Para o grupo *gases-explosão* os informantes usaram oclusivas bilabiais vozeadas e não vozeadas e fricativas uvulares, com vogais do grupo 3, principalmente a vogal [u], o que estará relacionado com a maior energia nas frequências mais baixas que caracteriza os sons deste grupo. Nos *gases-"gust"* e *gases-vento* usaram-se fricativas lábio-dentais não vozeadas, e vogais do grupo 3, estando estes resultados relacionados com a duração dos sons não verbais destes grupos que é elevada e com a maior energia na gama de frequências mais baixas.

Para os sons de animais verificou-se que na descrição dos sons de *animais-mamíferos* usaram oclusivas bilabiais e vibrantes, com traço de nasalidade e vozeamento e vogais do terceiro grupo. O uso de oclusivas e vibrantes estará relacionado com a natureza repetitiva de dois dos sons não verbais (cão e porco). O facto de ser usado o traço de vozeamento, estará relacionado com a percepção de um sinal periódico. Nos sons de *animais-aves* usaram-se principalmente oclusivas bilabiais não vozeadas e líquidas alveolares, com vogais do grupo 2, nomeadamente o [i]. O uso das oclusivas estará relacionado com o facto destes estímulos se traduzirem em sons curtos consecutivos, usando-se as oclusivas para descrever as transições onde há ausência de som, e as líquidas para caracterizar a rapidez dessas transições. O som *animal-insecto* originou onomatopeias bastante similares ao estímulo, com a oclusiva velar vozeada, a líquida alveolar e a vogal do segundo grupo. Tal como nas aves o som do grilo é repetitivo, o que será percebido como um sinal periódico, daí provavelmente o uso de vozeamento e da vogal que se caracteriza por F1 e F2 dentro da gama de frequências do estímulo original.

CAPÍTULO 5: PERCEÇÃO DE SONS NÃO VERBAIS FILTRADOS E COM RUÍDO

5.1 INTRODUÇÃO

Neste capítulo serão discutidos os resultados de duas experiências de percepção em que alguns sons do corpus foram manipulados: numa os sons foram filtrados e na outra foi-lhes aplicado ruído. Na segunda secção são abordados estudos de percepção que utilizam metodologia de referência para esta dissertação. Na terceira secção é descrita a manipulação dos estímulos para a obtenção do corpus de cada experiência. No final os resultados são apresentados e analisados, comparando-se os dados obtidos entre as duas experiências.

5.2 REVISÃO DA LITERATURA

Um dos primeiros estudos em que a perspectiva ecológica foi aplicada à percepção auditiva foi o de Vanderveer (1979), que englobou três metodologias diferentes. Primeiro foi pedido aos sujeitos que ouvissem 30 sons não verbais diferentes, através de uma cassete áudio, e escrevessem o que ouviam. A descrição baseou-se mais em características do som, tais como as acções que causavam o som, do que em propriedades acústicas do mesmo. Na segunda parte do estudo foi pedido aos sujeitos para classificarem 20 dos sons num número de categorias diferentes. Concluiu-se que os sons eram agrupados mais pelo seu padrão temporal: sons rítmicos, repetitivos ou de percussão tendiam a ser agrupados. Na última experiência foi pedido aos sujeitos para avaliarem numa escala de 1 a 10 a semelhança entre dois grupos de 5 sons (10 sons diferentes). Concluiu-se que as características acústicas mais determinantes para os ouvintes considerarem os estímulos semelhantes eram: variáveis temporais (descontínuo vs contínuo, existência de padrão rítmico, ataque e decaimento), ressonância ou outras propriedades de objectos, superfícies e substâncias.

Têm sido realizadas várias experiências para analisar aquilo a que o ouvido humano recorre para identificar sons não verbais (Gygi et al. 2004). Muitos destes trabalhos têm

examinado a associação entre as propriedades acústicas e a identificação de sons específicos, e.g., o tamanho e a forma de um objecto. Outros centraram-se em decisões subjectivas sobre a acústica dos objectos (há uma caracterização semântica associada ao som, no sentido do contexto em que é previsível que ocorra ou do grupo ou categoria de sons a que pertence (e.g., sons de transportes). Um exemplo de uma taxonomia dos sons não verbais, baseada na dimensão do tipo de material em interacção (gás, líquido ou sólido) e o tipo de interacção (impacto, explosão ou “*splash*”) foi proposta por Gaver (1993) e encontra-se descrita no Capítulo 2.

Myers et al. (1996) analisaram a capacidade de detecção e identificação de 25 sons não verbais no silêncio e com ruído de banda larga de 60 dB. O objectivo era explorar a possibilidade de usar sons não verbais filtrados em substituição de tons puros na avaliação audiométrica de crianças e de sujeitos com um reduzido tempo de atenção, já que é difícil para estes grupos focarem-se em tons puros, mesmo por curtos períodos de tempo. Os estímulos foram obtidos a partir de CDs de efeitos sonoros, tal como a amostra desta dissertação. Cada som foi filtrado em oito bandas com frequências centrais que variaram de 250 a 4000 Hz, que são frequências padrão usadas na audiometria. Para além da identificação e detecção dos sons não verbais, os limiares de detecção de tons puros para as mesmas frequências, na condição de silêncio e de ruído também foram obtidos. Os resultados de detecção e identificação foram semelhantes, realçando-se que a percepção da frequência de 4000 Hz foi a menos afectada com a aplicação de ruído. Os resultados obtidos em relação aos valores de dB SPL a que os diversos sons não verbais foram identificados e detectados, no silêncio e no ruído, encontram-se representados na Figura 63.

Em relação aos sons não verbais filtrados, os resultados da detecção no silêncio e no ruído foram semelhantes, e foram concordantes com a avaliação audiométrica realizada com tons puros, onde os sons filtrados a 250 Hz corresponderam aos limiares mais elevados e os mais baixos à frequência de 1000 Hz (apesar de a diferença entre as duas frequências centrais ter sido de apenas 4 dB). Myers et al. (1996) concluíram que para cada frequência central houve pelo menos dois sons com limiares no silêncio e no ruído com apenas 2 dB de diferença em relação aos limiares de tons puros. Assim, concluiu-se que sons não verbais filtrados em bandas estreitas podem ser usados como um substituto dos tons puros na determinação de limiares auditivos.

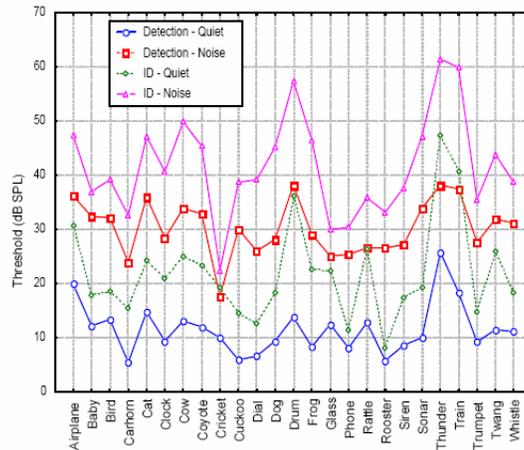


Figura 63: Limiares obtidos na experiência de Myers et al. (1996).

Gygi et al. (1997) analisaram a detecção e identificação de sons não verbais em ruído de banda larga. Usaram-se dois tipos de condições: incerteza mínima, em que os sujeitos sabiam qual o som que estava a ser testado, e incerteza máxima em que não sabiam. Os resultados apresentados na Figura 64 mostram que entre as duas condições de incerteza não há diferenças significativas e os valores variam entre -21dB Ev/N e -11.29dB Ev/N. Na identificação o intervalo é bastante maior, estando os valores compreendidos entre -29.28 dB Ev/N e -7.8 dB Ev/N, média = -19.04 dB Ev/N. Gygi et al. (1997) concluíram que os resultados de detecção são fracos preditores dos de identificação, tendo em conta as diferenças obtidas, que chegam a ser da ordem dos 8 dB para alguns estímulos.

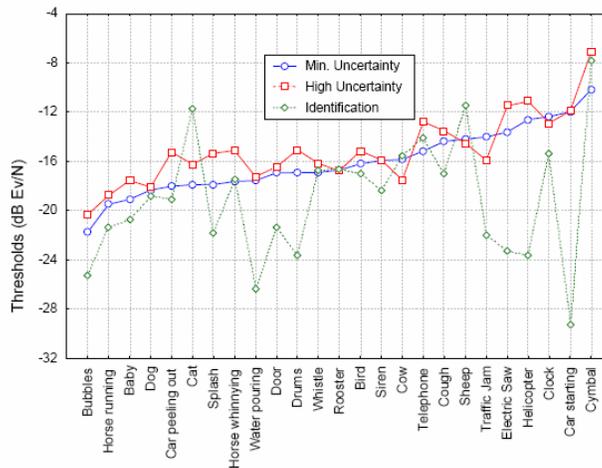


Figura 64: Valores de detecção e identificação no ruído de 25 sons não verbais (onde Ev/N significa Event to Noise, traduzindo relação Sinal / Ruído). De Gygi et al. (1997).

Gygi (2001) realizou outra experiência em que utilizou filtros passa-alto (Figura 65) e passa-baixo (Figura 66). Os valores das frequências de corte dos filtros passa-alto foram: 300, 600, 1200,

2400, 4800 e 8000 Hz; dos filtros passa-baixo foram: 300, 600, 1200, 2400 e 4800 Hz. Os sujeitos tinham de seleccionar o estímulo numa lista de 70 estímulos. De uma maneira geral os sujeitos obtiveram melhores resultados com filtros passa-alto, o que Gygi (2001) referiu indicar serem as regiões de alta frequência (Figura 65) as que transmitem maior informação ao ouvinte.

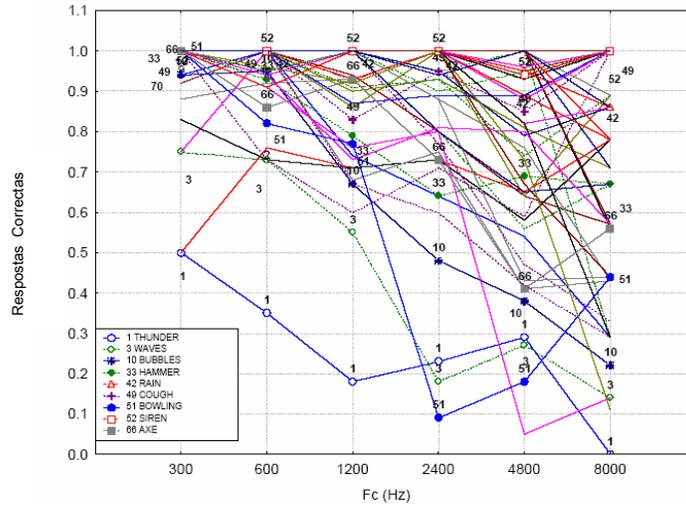


Figura 65: Resultados da identificação de sons não verbais com filtros passa-alto. De Gygi (2001).

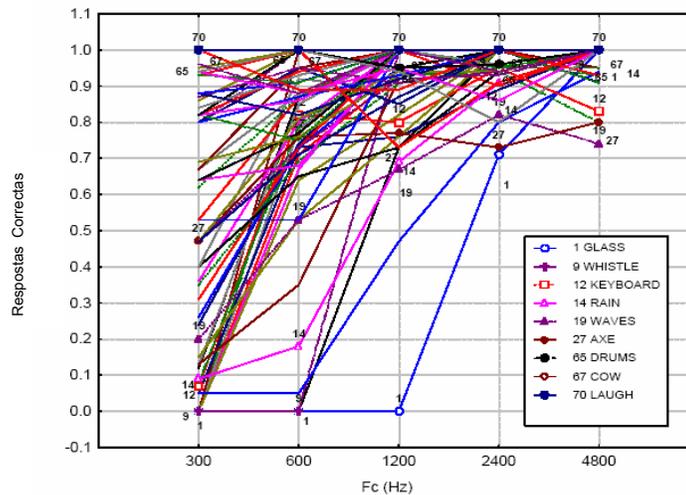


Figura 66: Resultados da identificação de sons não verbais com filtros passa-baixo. De Gygi (2001).

Gygi (2001) concluiu que haveriam sons em que as frequências baixas são as mais importantes, tais como a *trovoada* e as *ondas*, enquanto que para *vidro a partir* e *bowling* é necessária a informação acústica das altas-frequências para a identificação. Assim, Gygi (2001) procedeu a uma experiência para analisar a existência de regiões mais específicas, tendo aplicado filtros passa-banda. Foram utilizados 6 filtros, cujos valores das frequências de corte estão descritos na Tabela 8. Os sujeitos tinham de escolher o estímulo numa lista 70 opções.

Tabela 8: Valores das frequências de corte dos filtros passa-banda. De Gygi (2001).

	Fc Low (Hz)	Fc High (Hz)	Mid point (Hz)
BP1	150	300	212.13
BP2	300	600	424.26
BP3	600	1200	848.53
BP4	1200	2400	1697.06
BP5	2400	4800	3394.11
BP6	4800	9600	6788.23

Os resultados obtidos indicaram que quase todos os estímulos tiveram um filtro que impossibilitou a sua identificação, mas de um maneira geral os sons continuaram a ser identificados, sendo a maioria dos estímulos utilizados “resistente” aos efeitos da filtragem. Ao analisar os sons de acordo com a influência dos filtros na identificação Gygi (2001) obteve 8 grupos (Figura 67). No entanto, referiu ser difícil explicar os resultados já que entre os sons do mesmo grupo não encontrou uma relação lógica, mas considerou a hipótese de estarem principalmente relacionados com o tipo de evento que gera o estímulo, tal como Gaver (1993) argumenta.

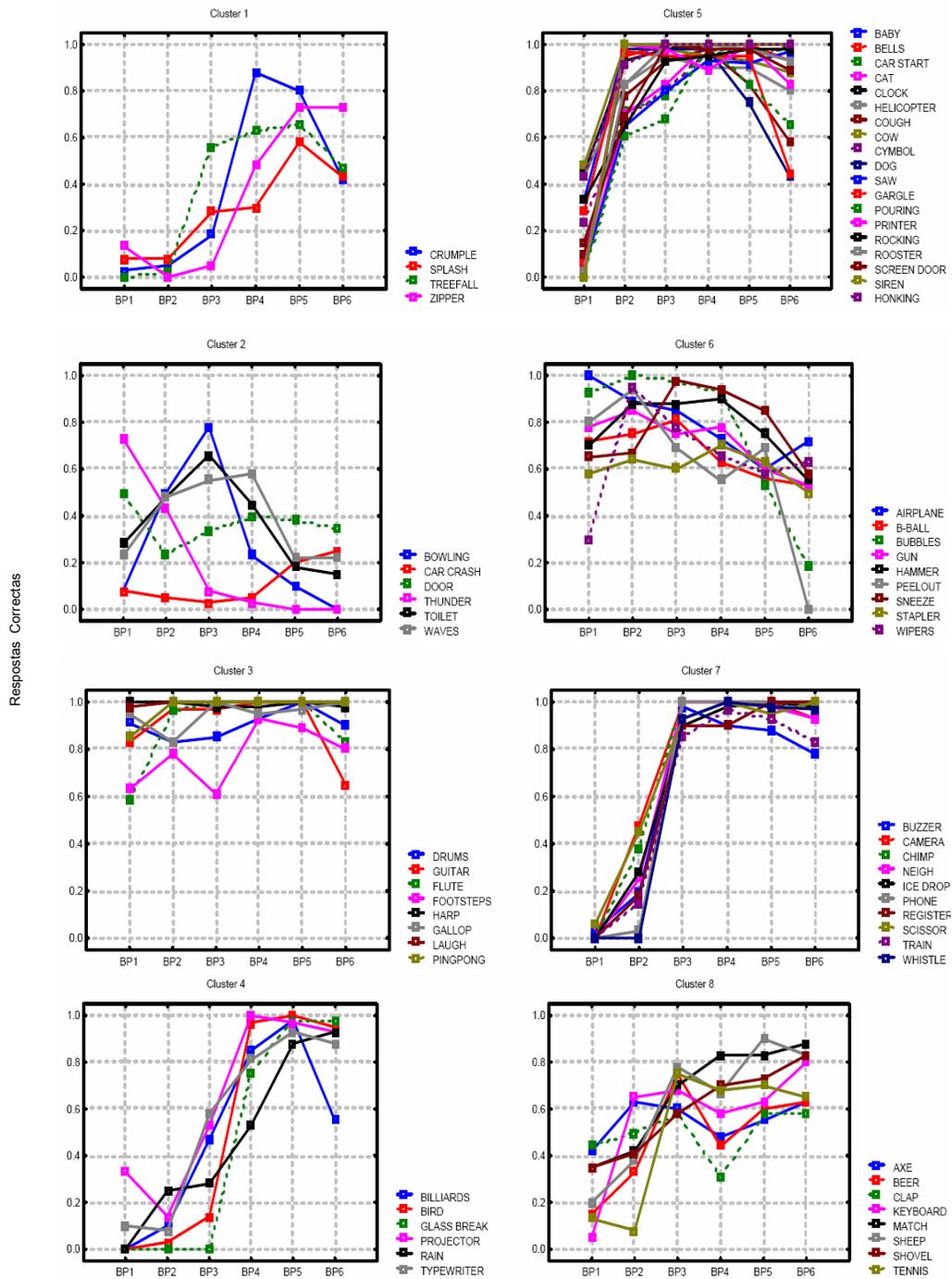


Figura 67: Os oito grupos obtidos com base nos resultados de identificação de 70 sons não verbais. De Gygi (2001).

5.3 MÉTODO

5.3.1 PERCEÇÃO DE SONS NÃO VERBAIS FILTRADOS

Para a experiência de percepção de sons não verbais filtrados, seleccionaram-se 6 sons de cada uma das categorias já referidas: sólidos, líquidos, gases e vocalizações de animais, procurando-se que nos seis estímulos estivessem incluídas as interações possíveis (e.g., *impacto* para os *sólidos*) dentro de cada um dos três primeiros grupos, e sons de mamíferos, aves e insectos. Os estímulos utilizados estão representados na Tabela 9. O estímulo *água a ferver* pertence a dois grupos: gases e líquidos o que faz com que haja um total de 23 estímulos.

Tabela 9: Corpus de sons não verbais utilizados na experiência de percepção em que foram utilizados filtros.

Código	Descrição	Classificação de Gaver										
		Sólidos				Líquidos				Gases		
		Impacto	Deformação	Raspar	Rolar	Pingar	Derramar	Splash	Ondular	Explosão	Gust	Vento
DQF024	esfregar as mãos											
DQF030	tambor											
DQF041	maracas											
DQF051	serrar											
GS1016	andar											
GS1035	sino											
DQF034	encher um copo de água											
GS1042	mar											
GS2008	cascata											
GS2009	rio											
GS2013	chuva											
GS2027	água a ferver											
DQF025	assobiar											
DQF053	encher pneu											
GS1031	avião											
GS1037	trovão/relâmpago											
GS1039	vento											
GS2027	água a ferver											

Sons de Animais		
GS1002	vaca	mamífero
GS1004	porco	mamífero
GS1006	cão	mamífero
GS1008	pato	ave
GS1009	grilo	insecto
GS1010	pássaro	ave

Depois de terem sido seleccionados os estímulos procedeu-se à manipulação dos mesmos com filtros passa-banda. Uma vez que as principais frequências utilizadas para avaliação da audição por meio de exames de audiologia são as de 500 Hz, 1000 Hz, 2000 Hz e 4000 Hz (Gelfand 2001), definiram-se os filtros passa-banda correspondentes a gamas de frequência que incluíssem as referidas frequências centrais (Tabela 10).

Tabela 10: Valores das frequências de corte dos filtros passa-banda utilizados na experiência de percepção de sons não verbais filtrados.

	Frequência inferior (Hz)	Frequência superior (Hz)	Frequência central (Hz)
Filtro 1	355	710	500
Filtro 2	710	1400	1000
Filtro 3	1400	2800	2000
Filtro 4	2800	5600	4000

Os estímulos foram filtrados através do programa *Speech Filling System (SFS) versão 4.6* (Huckvale 2000), no menu Tools → Speech → Process → Filtering → Band-pass filter, obtendo-se 4 estímulos filtrados para cada som (Figura 68). Uma vez que os estímulos gerados eram do tipo *.sfs* procedeu-se à conversão para ficheiros do tipo *.wav*. Desta forma foi obtido um total de 23 estímulos com 4 tipos de filtro mais o estímulo original, o que se traduziu em 115 estímulos, divididos em quatro categorias (30 estímulos por categoria, onde o estímulo *água a ferver* está incluído em dois grupos).

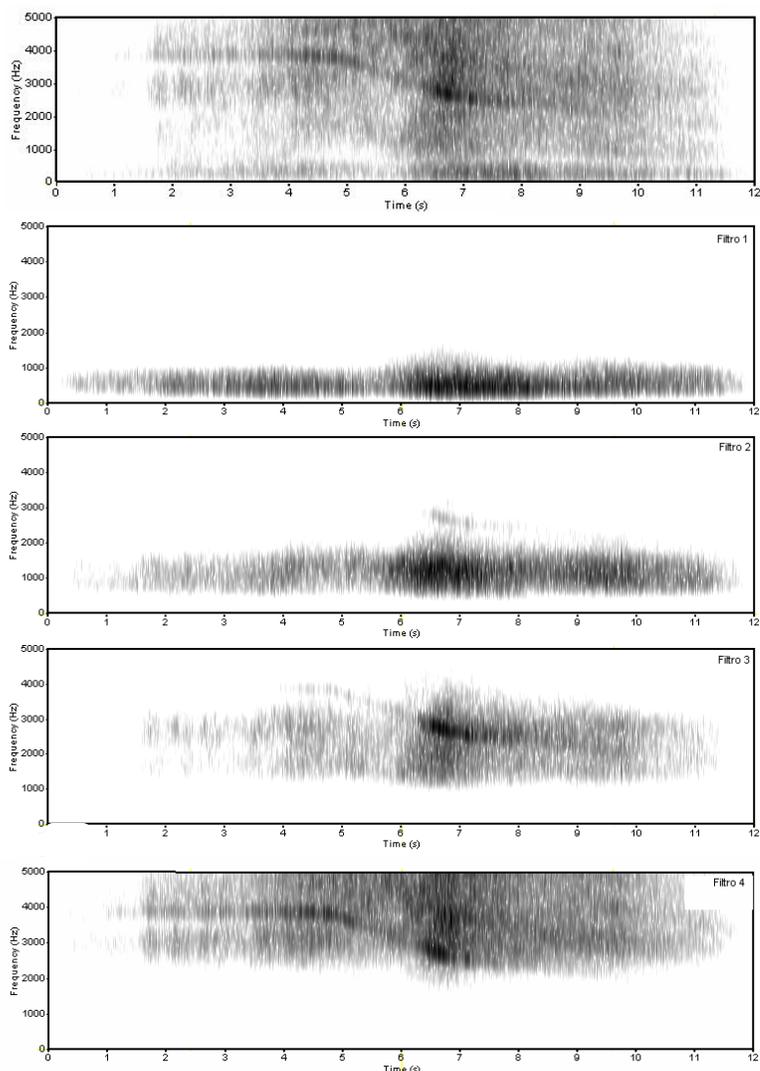


Figura 68: Exemplo de espectrogramas de um som não verbal filtrado, neste caso o estímulo *avião*. Em cima o estímulo original, e em baixo os estímulos obtidos após a aplicação dos filtros passa-banda.

A experiência foi dividida em quatro partes, correspondentes a cada uma das categorias de sons. A tarefa consistia em identificar o estímulo sonoro numa lista fechada de opções. Participaram os mesmos sujeitos que na experiência das onomatopeias. Os estímulos utilizados foram os do corpus de sons não verbais filtrados e foram apresentados a cada ouvinte com o programa *Praat versão 4.3.11* (Boersma e Weenink 2005) (ver Figura 69) através de auscultadores Senheiser eh 1430 ligados a um computador. A ordem de apresentação dos estímulos foi aleatória, sendo possível o sujeito realizar uma pausa ao fim de 20 estímulos (Figura 69: 1.3). Os sujeitos apenas ouviam uma vez cada som, e tinham de escolher uma das seis opções apresentadas no monitor (ver Figura 69).

No final de cada experiência, os resultados foram extraídos do *Praat* (Boersma e Weenink 2005) e transferidos para uma tabela de *Excel*, e posteriormente foram construídos gráficos representativos.

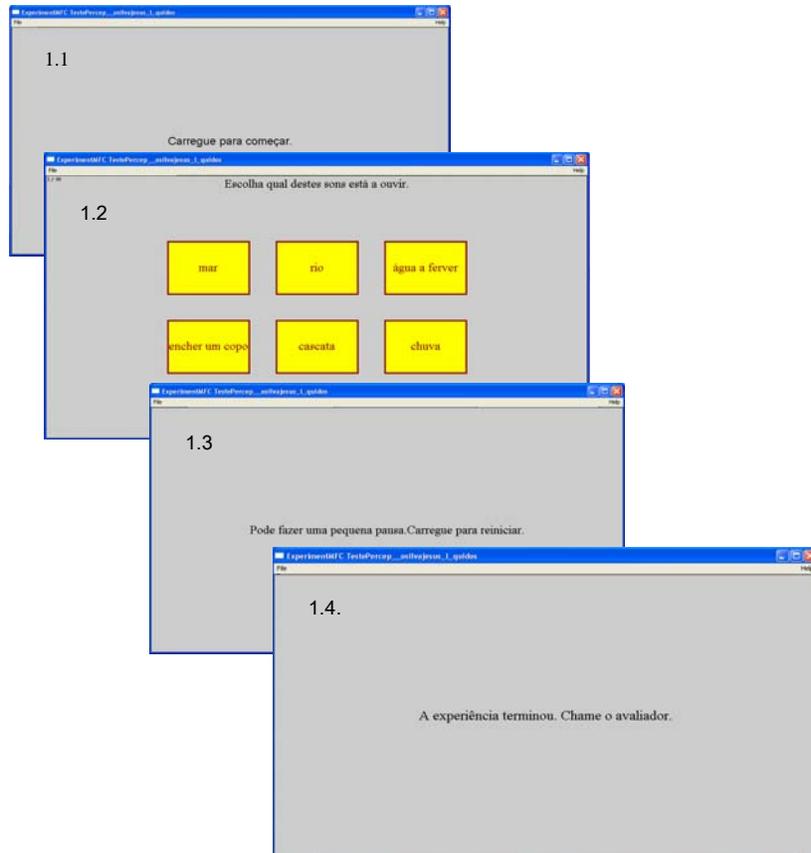


Figura 69: Imagens do programa *Praat*, na realização da experiência de percepção de sons não verbais filtrados, para o grupo líquidos. Observa-se em 1.1 a primeira imagem que surge no ecrã, para que se inicie a experiência; a imagem 1.2 corresponde ao grupo dos líquidos, com as opções de resposta; a imagem 1.3 é a que surge a cada 20 estímulos, para que o sujeito faça uma pausa se necessitar; e a imagem 1.4 a correspondente ao final da experiência.

5.3.2 PERCEÇÃO DE SONS NÃO VERBAIS COM RUÍDO

Para a experiência de percepção de sons não verbais com ruído sobreposto, seleccionaram-se 2 sons de cada uma das categorias já referidas, excepto para os sons de animais em que se utilizaram três estímulos (ver Tabela 11). Para os sons incluídos na classificação de Gaver (1993) seleccionaram-se estímulos que representassem os tipos de interacção possíveis. Para os sons de animais, incluíram-se três estímulos para haverem tipos diferentes de produção sonora.

Tabela 11: Corpus de sons não verbais utilizados na experiência de percepção em que foram utilizados ruídos sobrepostos.

		Classificação de Gaver										
		Sólidos				Líquidos				Gases		
Código	Descrição	Impacto	Deformação	Raspar	Rolar	Pingar	Derramar	Splash	Ondular	Explosão	Gust	Vento
DQF030	tambor											
DQF051	serrar											
GS2009	rio											
GS2013	chuva											
GS1031	avião											
GS1039	vento											

		Sons de Animais										
GS1002	vaca	mamífero										
GS1009	grilo	insecto										
GS1010	pássaro	ave										

Estando seleccionados os estímulos procedeu-se à sua manipulação com sobreposição de ruído. As gamas de frequência de ruído (Tabela 12) coincidiram com as frequências da experiência com filtros (Tabela 10) e para que a duração da experiência não fosse excessiva ao ponto de provocar o cansaço dos informantes e prejudicar os resultados obtidos, foram avaliadas apenas três gamas de frequência e ruído branco. Os estímulos de ruído foram gerados no programa *SFS versão 4.6*. (Huckvale 2000) Depois, através do programa *Colea* (Loizou 1999), os 6 tipos de sinais de ruído foram adicionados aos sons não verbais (Figura 70), através do menu *Tools-add noise from a file*, especificando-se nesta tarefa a relação sinal/ruído (S/N) pretendida. Utilizaram-se cinco relações S/R para cada um dos estímulos: -15 dB, -10 dB, -5 dB, 0 dB e 5 dB. Assim obteve-se um total de 9 estímulos com 4 tipos de ruído a 5 relações sinal/ruído diferentes, mais 9 estímulos originais, o que se traduziu em 189 estímulos.

Tabela 12: Valores das frequências de corte dos ruídos aplicados na experiência de percepção de sons não verbais com ruído.

	Frequência inferior (Hz)	Frequência superior (Hz)
Ruído 1	355	1400
Ruído 2	1400	2800
Ruído 3	2800	5600
Ruído branco	Cobre toda a gama de frequências	

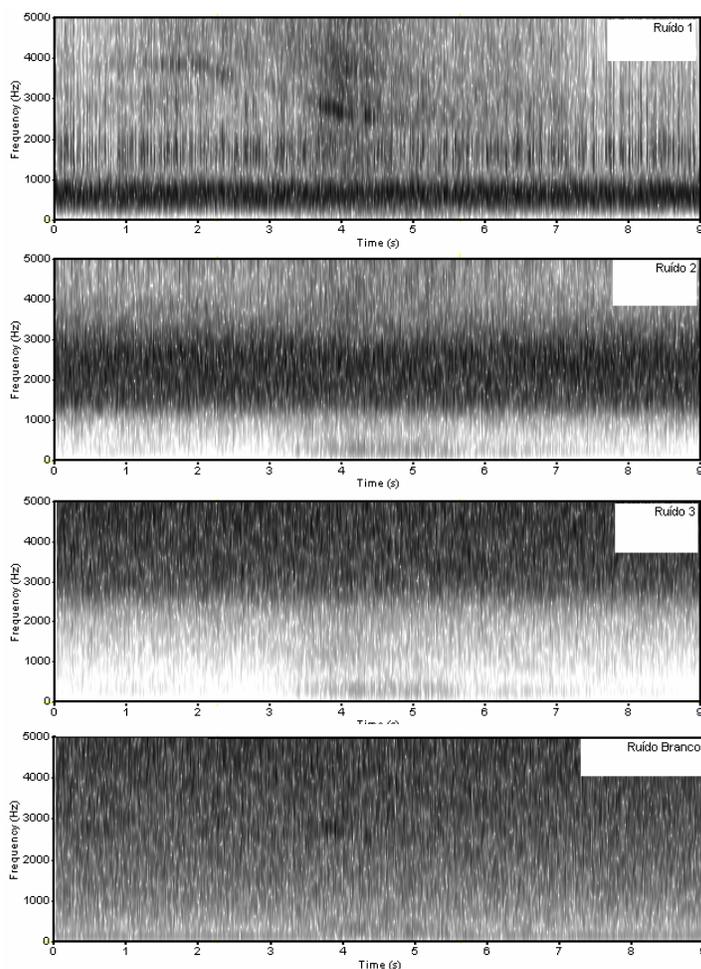


Figura 70: Exemplo de espectrogramas de um som não verbal, neste caso do estímulo *avião*, com ruído sobreposto a -15 dB de relação S/N.

Procedeu-se a uma experiência de percepção em que os sujeitos tinham de seleccionar o estímulo sonoro numa lista fechada de opções. Participaram os mesmos sujeitos que nas primeiras duas experiências de percepção. Os estímulos utilizados foram os do corpus de sons não verbais com ruído sobreposto (Tabela 11). Os estímulos foram apresentados a cada ouvinte com o programa *Praat versão 4.3.11* (Boersma e Weenink 2005) (ver Figura 71) através de auscultadores da Senheiser eh 1430 ligados a um computador.

A ordem de apresentação dos estímulos foi aleatória. Os sujeitos ouviam somente uma vez cada som, podiam fazer uma pausa a cada 20 estímulos, e tinham de escolher uma das opções apresentadas no monitor (ver Figura 71), as quais correspondiam aos nove estímulos incluídos na experiência. No final de cada experiência, os resultados foram extraídos do *Praat* (Boersma e Weenink 2005) e transferidos para uma tabela de *Excel*.

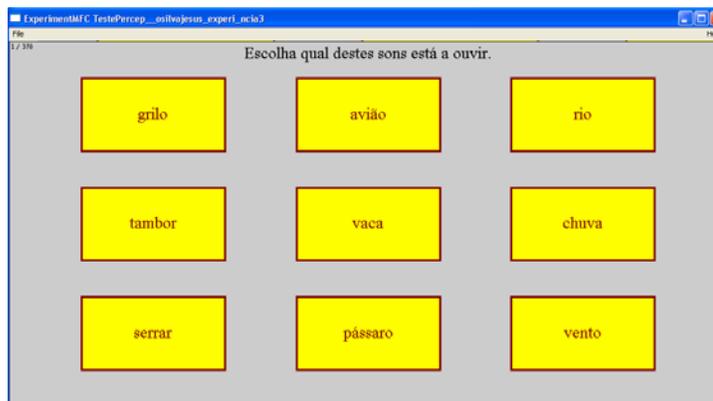


Figura 71: Imagem do Programa Praat, na realização da experiência 3.

5.4 RESULTADOS

De seguida serão apresentados os resultados obtidos nas experiências de percepção. É de referir que em relação à última experiência, a única relação sinal/ruído que interferiu com a percepção dos sujeitos foi a de -15 dB. Assim, serão apresentados os dados para todas as gamas de frequência e apenas para a relação sinal/ruído de -15 dB. Estes dados vão de encontro aos que Gygi et al. (1997) obtiveram, onde a maior parte dos sons era identificado em relações sinal/ruído superiores a -12 dB.

5.4.1 SÓLIDOS

Em relação aos sons gerados por sólidos, houve três estímulos que originaram resultados diferentes na identificação dos sons de acordo com o tipo de filtro utilizado (ver figuras 72, 73 e 74). Observa-se nas figuras 72 e 73 que as gamas de frequências em que os informantes erram mais na identificação dos sons *esfregar as mãos* e *maracas*, são correspondentes aos filtros 1 e 2, o que indicará que são necessárias as gamas de frequências entre os 1400 e os 5600 Hz, onde há maior percentagem de acerto, para serem identificados estes dois sons que envolvem a interacção *raspar*.

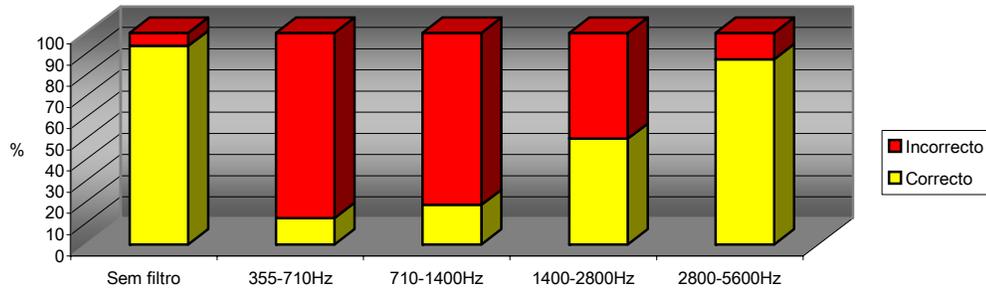


Figura 72: Distribuição das respostas de identificação dos sons filtrados de *esfregar as mãos*.

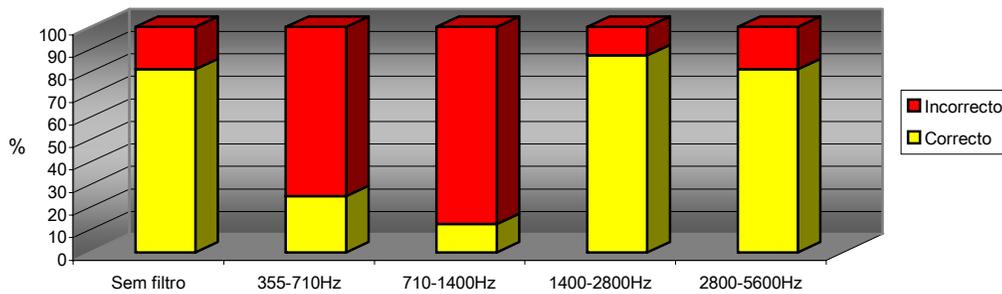


Figura 73: Distribuição das respostas de identificação dos sons filtrados de *maracas*.

Quanto ao *tambor* os resultados na Figura 74 indicam que os sujeitos têm dificuldades em identificá-lo tendo por base apenas a gama de frequências mais alta (2800 a 5600 Hz), e que com sobreposição de ruído a identificação do estímulo não é comprometida (Figura 75). Estes dados são concordantes com a análise espectral deste estímulo, onde se observou haver concentração de energia nas frequências inferiores a 2000 Hz. A aplicação de ruído não teve influência significativa na percepção do estímulo, tal como aconteceu no estudo de Gygi et al. (1997) onde o tambor foi identificado numa relação S/R de -24 dB.

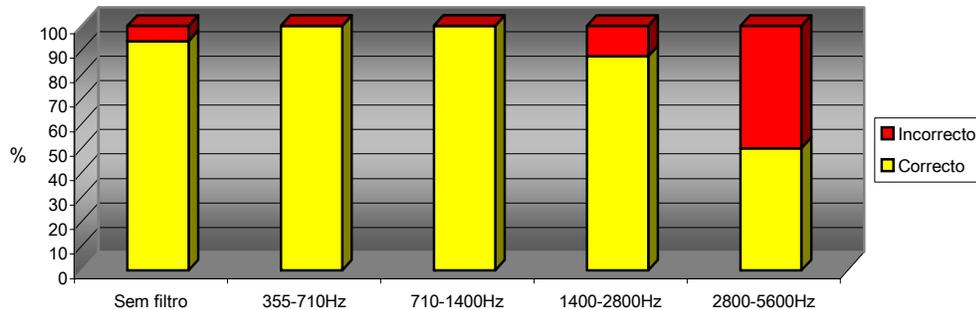


Figura 74: Distribuição das respostas de identificação dos sons filtrados de *tambor*.

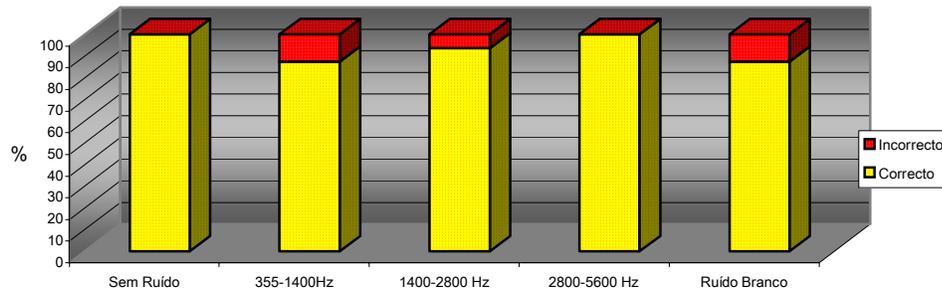


Figura 75: Distribuição das respostas de identificação dos sons de *tambor* com ruído.

5.4.2 LÍQUIDOS

Nos estímulos gerados por materiais líquidos os sujeitos tiveram dificuldades de identificação superiores a qualquer um dos outros grupos (ver Figuras 76 a 81). O estímulo gerado por uma interação do tipo *pingar*, foi identificado com dificuldade quando apenas foi ouvida a gama de frequências mais baixa, o que indicará que a maior informação acústica para a identificação do som *encher um copo de água*, estará presente entre os 710 e os 5600 Hz (ver Figura 76).

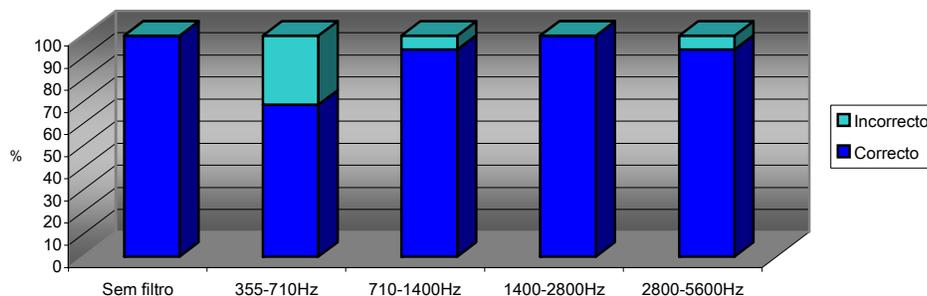


Figura 76: Distribuição das respostas de identificação dos sons filtrados de *encher um copo de água*.

No caso do som *chuva*, a identificação foi menos comprometida quando eram ouvidas as gamas de frequências mais altas (1400 a 2800 Hz e 2800 a 5600 Hz), e quase totalmente sem sucesso quando eram transmitidas as gamas de frequências inferiores, tal como pode ser observado na Figura 77. Na análise acústica deste estímulo observou-se uma maior amplitude acima dos 500 Hz, e nas onomatopeias utilizadas para o descrever foi clara a concentração de energia acima dos 2000 Hz. Também Gygi (2001) em experiência com filtros passa baixo, observou que a identificação do estímulo chuva aumentava com a frequência central do filtro passa baixo, ou seja, quando apenas eram ouvidas as frequências mais baixas o som não era identificado, atingindo-se os 70% de identificação a partir de 1200 Hz. A aplicação de ruído a

diferentes gamas de frequências não originou resultados diferentes entre as várias zonas do espectro (ver Figura 78).

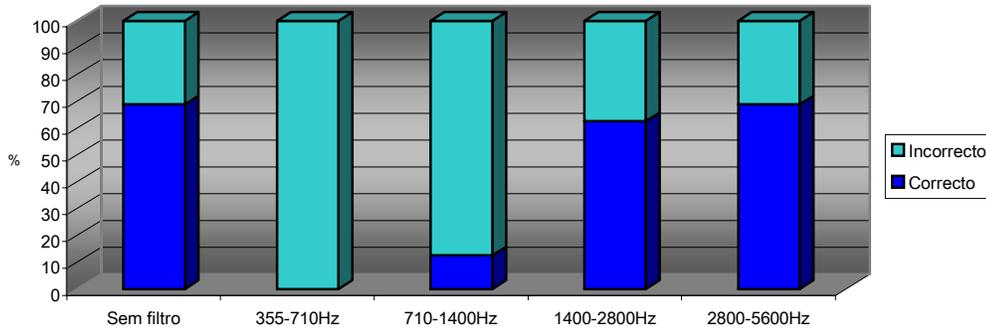


Figura 77: Distribuição das respostas de identificação dos sons filtrados de *chuva*.

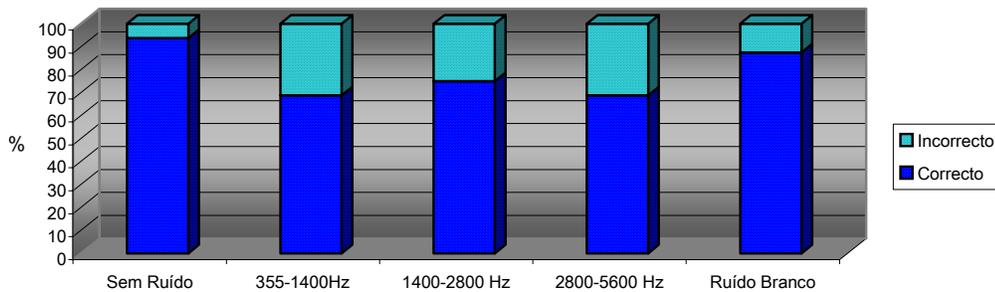


Figura 78: Distribuição das respostas de identificação dos sons de *chuva* com ruído.

O estímulo *mar*, decorrente de uma interação “*splash*”, foi difícil de identificar, mesmo sem filtro, sendo confundido com outros estímulos do grupo dos líquidos (ver Figura 79). As gamas de frequências mais baixas (entre os 355 e os 1400 Hz) parecem ser as que garantem a mais correcta identificação deste som, enquanto que as mais altas por si só não o possibilitam. Gygi (2001) obteve resultados semelhantes quando aplicou filtros passa-alto e passa-baixo, sendo a identificação das “ondas do mar” foi maior quando se incluíam as frequências até 1200 Hz.

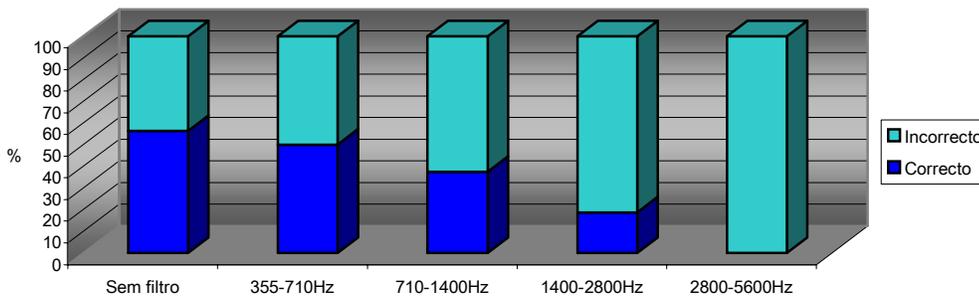


Figura 79: Distribuição das respostas de identificação dos sons filtrados de *mar*.

O som *rio* representa uma interacção do tipo *ondular*. Observa-se na Figura 80 que a gama de frequências de 710 a 1400 Hz é a que permite aos sujeitos uma melhor identificação deste estímulo, mesmo em relação ao estímulo sem filtro. No espectro deste estímulo também se observou haver pouca concentração de energia nas frequências inferiores a 500 Hz, assim como no espectro das onomatopéias usadas para o descrever. Ao ser aplicado ruído, este compromete a identificação quando cobre toda a gama de frequências, ou seja com ruído branco (ver Figura 81).

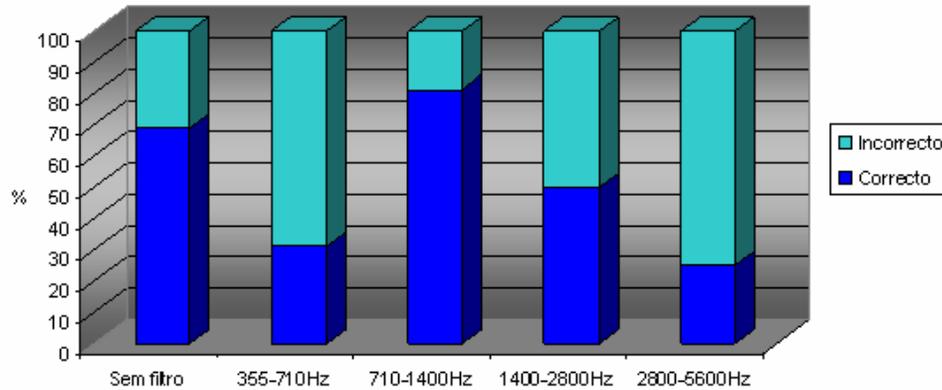


Figura 80: Distribuição das respostas de identificação dos sons filtrados de *rio*.

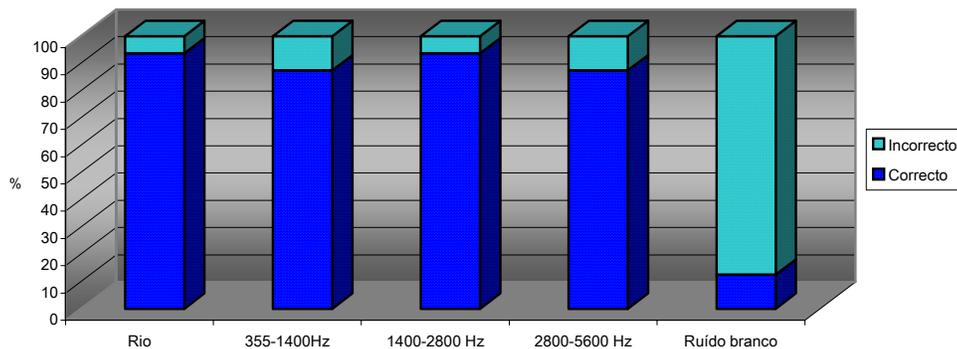


Figura 81: Distribuição das respostas de identificação dos sons de *rio* com ruído.

5.4.3 GASES

No que concerne os estímulos gerados por interacções entre gases, os resultados foram distintos para cada tipo de interacção. Em relação ao som do *avião* observa-se na Figura 82 que a gama de frequências que possui menos informação para os sujeitos realizarem a identificação é entre os 355 e os 710 Hz, o que contradiz os resultados de (Gygi 2001), em que o som *avião* foi melhor identificado com filtros passa-banda com frequências de corte mais baixas. Quanto à influência do ruído na identificação deste estímulo do grupo de interacção "*gust*", o ruído

sobreposto às gamas de frequência mais baixas (355 a 1400 Hz) foi o que prejudicou menos a identificação do *avião* (ver Figura 83), o que corrobora os resultados dos sons filtrados (Figura 82).

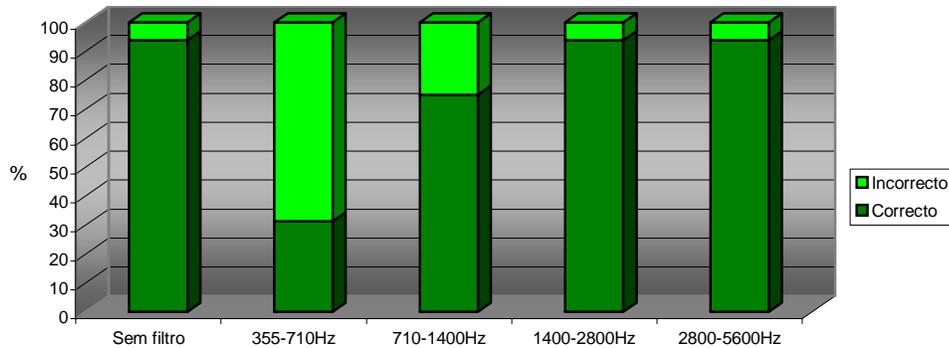


Figura 82: Distribuição das respostas de identificação dos sons filtrados de *avião*.

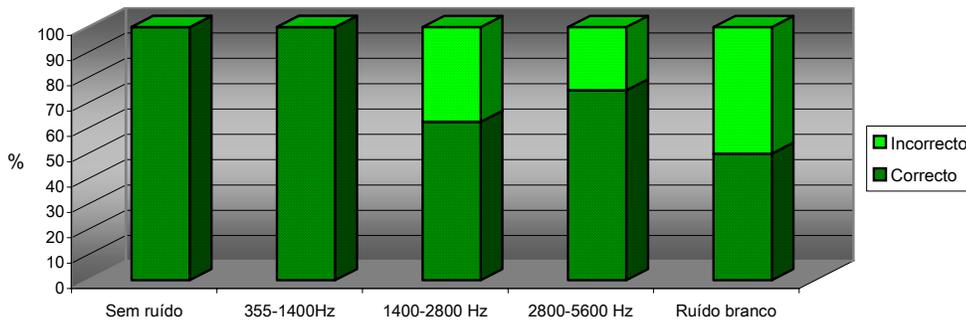


Figura 83: Distribuição das respostas de identificação dos sons de *avião* com ruído.

Quanto ao estímulo *vento* as gamas de frequências superiores (1400 a 2800 Hz e 2800 a 5600 Hz) parecem ser as que possuem menos informação acústica para os sujeitos identificarem o estímulo (Figura 84), daí que o ruído sobreposto às frequências mais baixas tenha prejudicado mais a identificação do que na gama entre os 2800 e os 5600 Hz (Figura 85). Na análise acústica deste som é clara a concentração de energia nas frequências inferiores a 1500 Hz, assim como nas onomatopeias utilizadas para o descrever onde houve uma maior utilização de fones com maior energia nas frequências mais baixas.

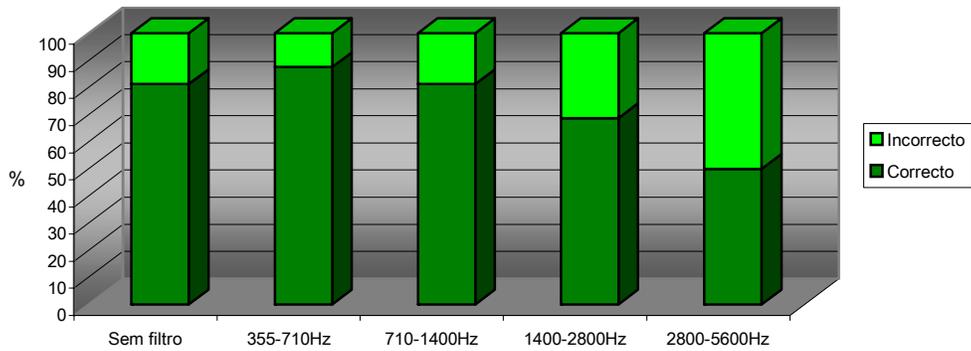


Figura 84: Distribuição das respostas de identificação dos sons filtrados de *vento*.

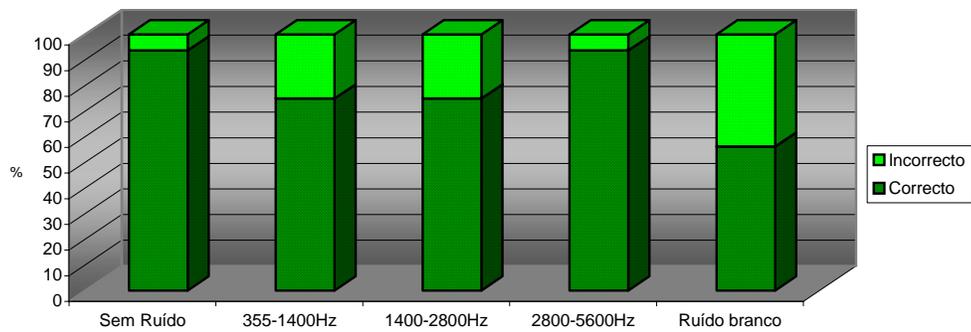


Figura 85: Distribuição das respostas de identificação dos sons de *vento* com ruído.

Por último, o estímulo *trovoada*, resultante de uma interação *explosão*, é dificilmente identificado quando apenas são transmitidas as gamas de frequências mais altas, o que indicará que a identificação depende fundamentalmente de informações acústicas contidas nos valores mais baixos (Figura 86). Realmente, o espectro deste estímulo evidencia uma maior concentração de energia abaixo dos 1000 Hz. Gygi (2001) observou o contrário, sendo na banda de frequências mais baixas que os seus sujeitos obtiveram maior sucesso de identificação.

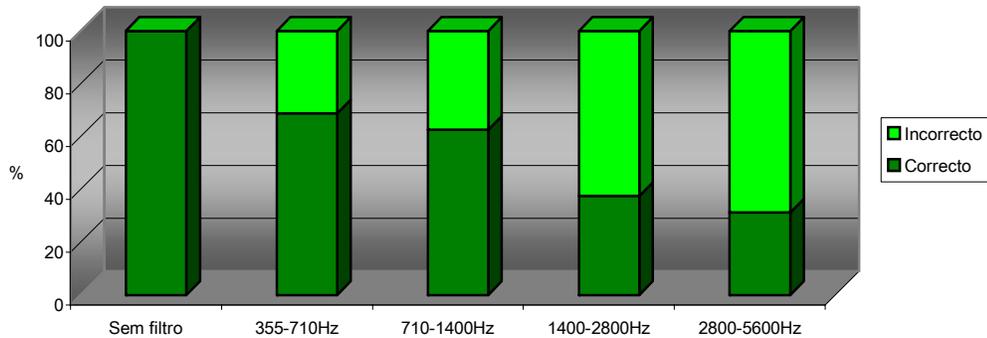


Figura 86: Distribuição das respostas de identificação dos sons filtrados de *trovoada*.

5.4.4 SONS DE ANIMAIS

O grupo dos sons gerados por animais originou resultados distintos de acordo com a classe animal. O som *grilo* foi identificado com maior sucesso para as gamas de frequência mais elevadas (1400 a 2800 Hz e 2800 a 5600 Hz), tal como se observa na Figura 87. Na experiência de percepção com ruído, os resultados foram coincidentes já que o ruído nas frequências mais altas foi aquele que prejudicou mais a identificação do estímulo. Na análise acústica do som *grilo* também se observou que a energia se concentrava entre 2000 e 4000 Hz.

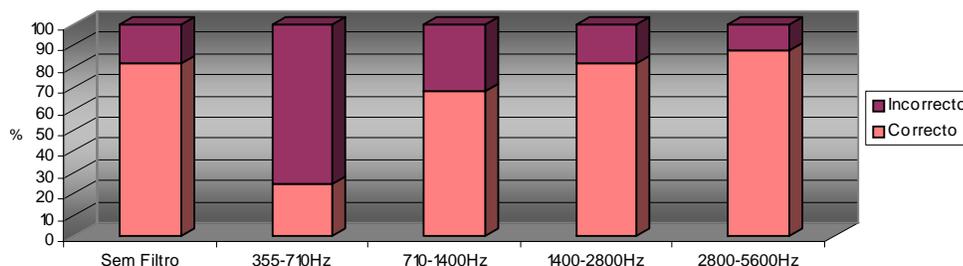


Figura 87: Distribuição das respostas de identificação dos sons filtrados de *grilo*.

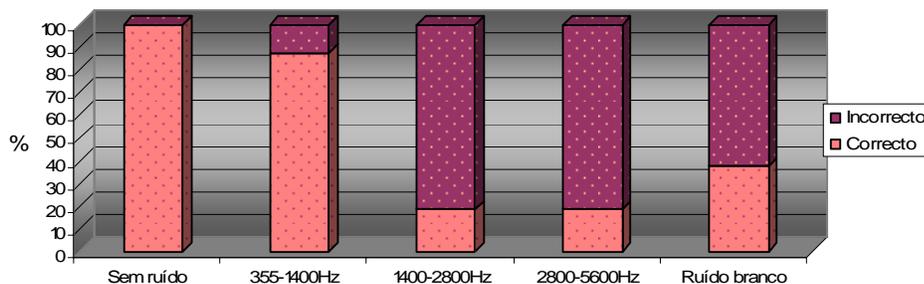


Figura 88: Distribuição das respostas de identificação dos sons *grilo* com ruído.

O estímulo gerado pela ave *pássaro* foi correctamente identificado com as componentes frequenciais entre 2800 e 5600 Hz (Figura 89) e foi também esta a gama de frequências que os sujeitos menos identificaram o *pássaro* na presença de ruído (Figura 90). O espectro deste estímulo também revelou uma maior energia nas frequências entre 2000 e 4000 Hz. Gygi (2001) observou que o som de outra ave, o *galo*, também originava uma melhor identificação com filtros passa-banda com frequências de corte superiores a 2400Hz.

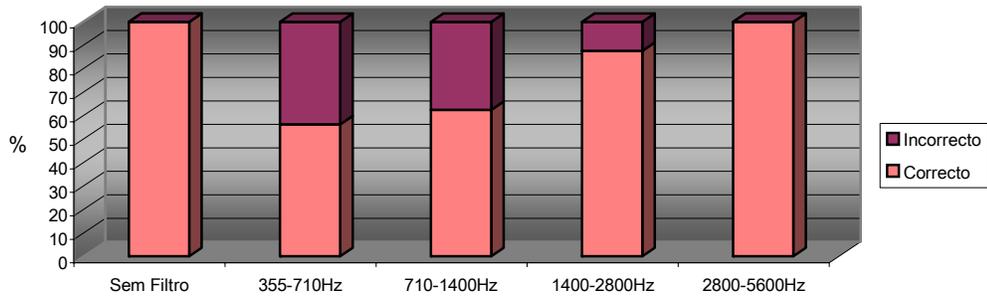


Figura 89: Distribuição das respostas de identificação dos sons filtrados de *pássaro*.

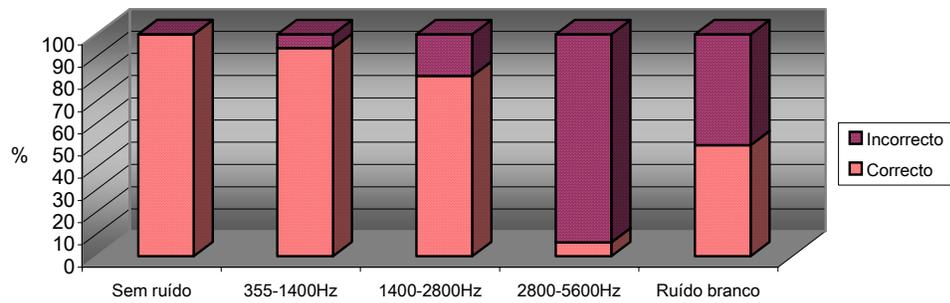


Figura 90: Distribuição das respostas de identificação dos sons de *pássaro* com ruído.

Por último, o som *vaca* foi sempre identificado independentemente do filtro (Figura 91), mas a sobreposição de ruído nas duas gamas de frequências mais baixas prejudicou significativamente a identificação deste som (Figura 92). No seu espectro, verificou-se haver mais energia nas frequências inferiores a 2000 Hz.

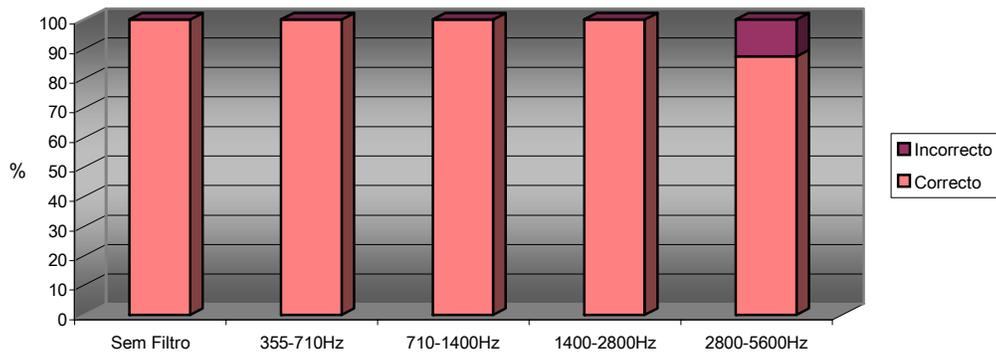


Figura 91: Distribuição das respostas de identificação dos sons filtrados de *vaca*.

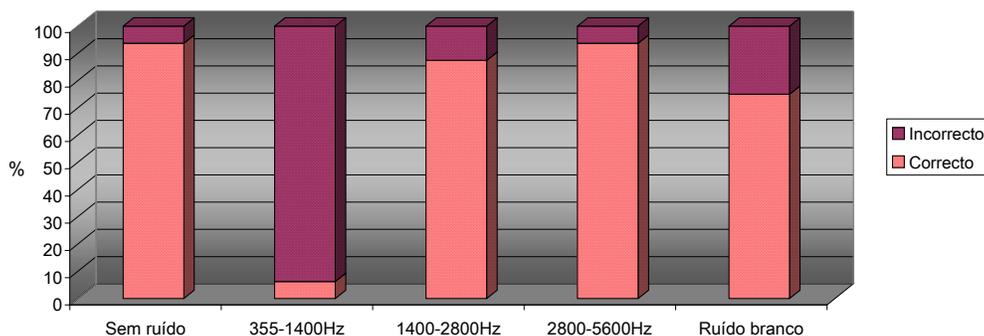


Figura 92: Distribuição das respostas de identificação dos sons de *vaca* com ruído.

5.5 SUMÁRIO

Neste capítulo foram apresentados estudos que envolveram testes de percepção de sons não verbais filtrados e onde houve aplicação de ruído a estes estímulos.

Na terceira parte foram apresentadas as duas experiências de percepção de sons complexos realizadas nesta dissertação, onde se procurou analisar quais as gamas de frequências que são necessárias à identificação de sons não verbais de diferentes categorias. A primeira experiência consistiu na selecção de 24 estímulos da base de dados e na aplicação de filtros passa-banda correspondentes às quatro principais frequências utilizadas na avaliação audiológica (em que os estímulos utilizados são tons puros). Foi pedido aos sujeitos para identificarem os sons em tarefas de escolha forçada, sendo apresentados como opção os estímulos da mesma categoria. Na segunda experiência foram utilizados 9 estímulos, aos quais se aplicou ruído branco e três ruídos com gama de frequências diferentes. Foram utilizados estímulos com uma relação sinal/ruído de 5 dB, 0 dB, -5 dB, -10 dB e -15 dB. Pediu-se aos participantes para identificarem o que ouviam em escolha forçada com os 9 estímulos possíveis.

Observou-se que para os sons gerados por *sólidos* num *impacto* ou numa *deformação* os filtros correspondentes às 3 frequências inferiores garantiam a identificação. Pelo contrário, as interações *raspar* apenas foram percebidas com filtros na gama de 1400 a 5600 Hz. O ruído não teve influência na percepção dos estímulos do grupo *sólidos*.

No grupo do *líquidos* a identificação foi mais afectada pelos filtros e ruído do que nos outros grupos. Assim, estímulos causados por *pingar* foram melhor percebidos com as frequências mais elevadas, enquanto que o evento simples de *“splash”* (*mar*) obteve valores superiores de identificação com as gamas de frequências mais baixas e total insucesso com a mais elevada. Por sua vez a *chuva*, evento complexo de *pingar* e *“splash”* foi incorrectamente identificada por todos os sujeitos quando apenas ouviam a gama de frequências de 355 a 710 Hz, tendo havido cerca de

10 % de respostas correctas na gama de 710 a 1400 Hz, tal como aconteceu com o estímulo *encher um copo de água (pingar)*, o que indica resultados para a *chuva* mais coincidentes com os resultados do evento simples com interacção *pingar* do que com o de *“splash”*. Quanto a *ondular*, os resultados indicaram que a gama de frequências entre 710 e 1400 Hz é de tal forma relevante para a percepção que a identificação foi melhor com este filtro do que na identificação do estímulo sem filtro. Apenas no caso do som *rio* (interacção *ondular*) houve influência do ruído na percepção, mas apenas na aplicação do ruído branco.

Para os *gases* observou-se que sons gerados por uma *explosão* foram melhor identificados nas gamas de frequências mais baixas. Estímulos do tipo *“gust”* obtiveram percentagens de sucesso inferiores a 30% quando era apenas ouvida a gama de frequências mais baixa, e quase 100% de acerto nas duas gamas mais elevadas, tendo-se observado que o ruído aplicado a estas duas gamas de frequências prejudicou a identificação dos estímulos, assim como o ruído branco. Quanto ao tipo *vento* houve melhor identificação com os dois filtros de gama inferior, e o ruído aplicado nestas gamas prejudicou mais a identificação do que aplicado na gama de frequências de 2800 a 5600 Hz.

Quanto aos *sons de animais mamíferos* a aplicação de filtros não prejudicou a sua identificação, mas ao ser aplicado ruído às duas gamas de frequências inferiores, a identificação do som foi bastante prejudicada (menos de 5% de respostas correctas). Em relação ao grupo *aves* a melhor identificação verificou-se com os dois filtros com frequências de corte mais elevadas, o que se confirmou com as respostas da experiência com ruído, em que a aplicação de ruído na gama mais elevada (2800 a 5600 Hz), prejudicou a quase totalidade das respostas (menos de 10 % de respostas correctas). O estímulo do grupo *insectos* foi dificilmente identificado com o filtro com frequências de corte de 1400 e 2800 Hz, e a aplicação de ruído na mesma gama de frequências prejudicou a identificação do estímulo resultando em menos de 20 % de respostas correctas.

É de realçar o facto de a aplicação de ruído ter influenciado mais a percepção dos sons gerados por *animais* e por *gases*, em que nos *animais* chegaram a haver percentagens de respostas correctas inferiores a 10%. Além disso, é de referir o facto de nos *sólidos* e nos *gases* os sons terem sido identificados com filtros de várias gamas de frequências, enquanto que nos *sons de animais* e de *líquidos* a gama de frequências necessária à identificação ter sido mais restrita e dependente normalmente de uma ou duas gamas de frequências específicas.

CAPÍTULO 6: CONCLUSÕES E TRABALHO FUTURO

6.1 INTRODUÇÃO

Neste último capítulo é apresentada uma síntese desta dissertação. É dada ênfase às principais conclusões relacionando-se a bibliografia consultada com os resultados obtidos na análise acústica e nas três experiências de percepção realizadas com sons não verbais. Além disso, é referido o contributo deste trabalho para a intervenção em terapia da fala, terminado com algumas propostas de continuidade desta temática num trabalho futuro.

Este estudo procurou dar resposta à necessidade dos terapeutas da fala perceberem de que forma os sons não verbais podem ser utilizados de forma objectiva na intervenção para o desenvolvimento de competências auditivas. A deficiência auditiva pode ser de diferentes graus (relativamente ao limiar de audição) e afectar de formas distintas a percepção da larga gama de frequências que o ouvido humano é capaz de ouvir (podem estar afectadas as gamas mais altas, as mais baixas, num intervalo restrito ou mais amplo). Assim, quando pedimos a uma criança para identificar o som de um pássaro é essencial saber se as características acústicas deste som são possíveis de identificar pela criança, tendo em conta o seu limiar de audição. Ou seja, se a criança tiver uma perda auditiva em que a gama de frequências mais altas está afectada e se o som do pássaro apresentar componentes acústicos na mesma gama de frequências, será muito difícil ou até impossível à criança identificar aquilo que o seu sistema auditivo não tem capacidade para ouvir.

A amostra deste estudo foi constituída por sons não verbais seleccionados de jogos didácticos utilizados na terapia da fala. A categorização da amostra teve por base a perspectiva de Gaver (1993) que defendeu a existência de uma relação entre o fenómeno que provoca o som e o resultado acústico originado, no que diz respeito aos materiais (sólidos, líquidos ou gases) e às interacções entre materiais. O facto de na intervenção em terapia da fala serem também utilizados sons de animais fez com que estes também fossem incluídos na amostra.

6.2 CONCLUSÕES

Os sons gerados por *sólidos* num *impacto* ou *deformação* são os que apresentam duração menor, e uma amplitude inicial maior que diminui rapidamente num curto espaço de tempo. Existe uma relação entre a percepção de um som não verbal e a produção de um som verbal como uma onomatopeia, quando se verifica que os sujeitos utilizaram oclusivas não vozeadas para descrever aqueles sons, com os quais partilham a curta duração, a reduzida energia e a ausência de periodicidade. O facto de na manipulação destes sons com filtros e com a aplicação de ruídos a sua identificação não ter sido muito prejudicada, leva a concluir que a sua identificação se baseia principalmente no seu padrão temporal, do que numa determinada gama de frequências, algo que também se verifica nas oclusivas. Uma interacção de *sólidos* a *raspar* apresentou características acústicas diferentes das outras interacções de sólidos. Neste caso, a duração é maior e o sinal é contínuo, e existe pouca concentração de energia nas frequências inferiores a 500 Hz. Assim, foi observável o uso de fricativas não vozeadas para descrever este grupo e o uso único da vogal [i], que apresenta valores de F2 elevados. Além disso, na percepção deste tipo de sons filtrados, as respostas correctas verificaram-se com filtros passa-banda nas gamas de frequências mais elevadas (entre 1400 e 5600 Hz), em cujo intervalo se encontra F2 da vogal [i]. Conclui-se que a percepção de *sólidos-raspar* depende das gamas de frequências mais elevadas.

No caso dos *líquidos*, observou-se que em todas as experiências os sujeitos mostraram mais dificuldade em descrever e identificar estes sons relativamente aos outros grupos. A nível temporal tratam-se de sons contínuos, com os maiores valores de duração de toda a amostra e amplitude relativamente constante. Nas interacções do tipo *pingar* e *ondular* observou-se pouca energia nas frequências inferiores a 500 Hz. Uma vez que as fricativas são os sons contínuos com menos energia abaixo dos 500 Hz, foram maioritariamente utilizadas pelos informantes na descrição dos estímulos gerados por *líquidos*. Verificou-se ainda o uso de fricativas não sibilantes para descrever “*splash*” e sibilantes para os outros dois tipos de interacção, o que indicará uma percepção de menor energia sonora nos sons gerados por “*splash*”. Apesar de na análise acústica não se terem observado diferenças espectrais nas várias interacções de líquidos, nas experiências de percepção houve resultados distintos de acordo com a gama de frequências ouvida. Os sujeitos basearam-se nas frequências mais elevadas para identificarem sons do tipo *pingar*, para os do tipo “*splash*” as mais baixas, e para os do tipo *ondular* a gama de frequências entre 710 e 1400 Hz.

Em estímulos gerados por *gases* a duração também é elevada e há uma maior concentração de energia nas frequências inferiores a 1000 Hz. Na interacção do tipo *explosão*, a qual resulta num som com um início abrupto foram usados sons da fala com esta característica, nomeadamente as oclusivas. Já nos outros dois tipos de interacção usaram-se fricativas sibilantes não vozeadas, traduzindo a percepção de um sinal contínuo aperiódico. Neste grupo de estímulos

foram usadas as vogais do grupo /u, o, ɔ/, que apresentam valores de F1 e F2 baixos, o que pensamos estar relacionado com a maior concentração de energia abaixo dos 1000 Hz que se observou na análise espectral dos estímulos não verbais. Através das experiências de percepção foi possível observar que quer a interacção de tipo *explosão* quer a de *vento* foram melhor identificadas com a audição das frequências mais baixas confirmando ser nestes valores que se encontram as informações acústicas para a percepção dos estímulos em causa. Pelo contrário, para identificar estímulos do tipo “*gust*” os sujeitos basearam-se nas gamas de frequências mais elevadas.

Relativamente aos *animais*, tratam-se de sons longos, onde se observam características periódicas e aperiódicas. Nos *mamíferos* há maior energia nas frequências menores que 3000 Hz, e foram usadas oclusivas vozeadas e nasais, fricativas uvulares e as vogais /u, o, ɔ/, estando estes dados relacionados com o início das produções, que é abrupto tal como nas oclusivas, com a existência de periodicidade traduzida no vozeamento e com a menor energia nas frequência elevadas como as vogais do Grupo 3 apresentam. Para além disso, o facto de ao ser aplicado ruído nas frequências inferiores (355 a 1400 Hz) a identificação dos estímulos ter sido bastante prejudicada, confirma a hipótese de que é nesta gama de frequências que se encontram as informações acústicas necessárias à percepção destes estímulos.

Quanto às *aves* observou-se haver uma maior concentração de energia acima dos 2000 Hz, e no espectrograma observaram-se transições rápidas de ressonâncias, tal como acontece na produção de ditongos e de líquidas. Assim, verificou-se que os sujeitos utilizaram oclusivas bilabiais e líquidas alveolares com a vogal /i/, que também apresentam maior energia nas frequências altas. Desta forma, os resultados das experiências de percepção mostraram que é na gama de frequências mais elevada, nomeadamente de 1400 a 5600 Hz que se encontram as informações acústicas necessária à percepção destes sons, já que os sujeitos tiveram melhor identificação com os filtros destas gamas e pior identificação quando era aplicado ruído nestas regiões do espectro.

Quanto ao estímulo do tipo *insecto*, na análise acústica não se observou a existência de um sinal periódico mas a gama de frequência onde a energia sonora está concentrada é bastante restrita e corresponde aos valores entre 2500 e 3500 Hz. Devido à repetição rápida de estímulos pela fricção de estruturas sólidas (asas do insecto), na percepção este som parece ser identificado como periódico, já que os sujeitos usaram oclusivas vozeadas para o descrever. Além disso, o uso da vogal com maior energia nas frequências altas é concordante com o espectro obtido deste som. Outros dados que confirmam a importância da gama de frequências mais elevada é a melhor identificação do estímulo com os dois filtros com frequências de corte mais altas e a pior identificação com o ruído aplicado a essa gama de frequências.

Assim, parece haver evidência para existência de características acústicas comuns aos sons não verbais e aos sons da fala, já que de uma maneira geral se observou que os sons da fala

usados para descrever os sons não verbais possuíam características acústicas comuns aos mesmos. Além disso, foi possível verificar, tal como Gaver (1993) sugere, que eventos físicos diferentes originam fenómenos acústicos também diferentes. É possível identificar características acústicas comuns a elementos de cada um dos grupos, mas as interações que ocorrem são determinantes para o sinal acústico que é gerado. Considera-se que os sons de menor duração, nomeadamente os que resultam de *sólidos* a interagir num *impacto* ou *deformação*, não serão percebidos com base nas características espectrais mas sim no seu padrão temporal, o que não foi possível confirmar neste estudo.

O facto de se tratar de uma temática pouco estudada, a nível internacional e ainda menos no âmbito nacional, dificultou a realização desta investigação, quer na selecção de bibliografia relacionada com o tema, quer na selecção da metodologia específica para analisar sons não verbais. De referir ainda, a limitação da generalização dos resultados devido ao tamanho reduzido da amostra e à complexidade de sons não verbais existentes no meio ambiente.

6.3 TRABALHO FUTURO

As aplicações do estudo de sons não verbais são variadas, tal como foi referido no segundo capítulo desta dissertação. No presente estudo há um contributo para a compreensão do funcionamento do ouvido humano, no sentido em que foi encontrada evidência para a relação entre as características acústicas de sons não verbais e sons de fala, estando percepção e produção directamente relacionadas.

A realização deste estudo sobre sons não verbais, pioneiro em Portugal, abre a possibilidade de haver uma maior contribuição em três áreas principais: o diagnóstico e avaliação da audição humana (normal e com patologia), a intervenção no desenvolvimento de competências auditivas e a investigação das características acústicas e da percepção de sons não verbais manipulados.

No diagnóstico e avaliação auditiva são actualmente utilizados tons puros e palavras faladas, que muitas vezes não se mostram eficazes quando a população em causa se constitui por crianças pequenas ou por pessoas com défice cognitivo. Através do conhecimento das características acústicas dos sons não verbais será possível que estes sejam utilizados na avaliação auditiva em complemento e/ou substituição aos tons puros e à fala. A nossa audição desenvolve-se pela exposição a sons não verbais e a sons de fala, e com este estudo foi possível concluir que a percepção de ambos apresenta semelhanças. Seria pertinente, então, realizar um estudo em que fossem utilizados na avaliação audiológica tons puros e sons não verbais, com as

mesmas características espectrais, de forma a aferir a possibilidade de usar os últimos na avaliação e diagnóstico da patologia auditiva.

Quanto à intervenção em terapia da fala propriamente dita seria útil analisar os benefícios da utilização de sons não verbais no treino auditivo. Num estudo longitudinal com crianças surdas, verificar se há diferenças significativas no tempo de aquisição de competências auditivas através da utilização de sons não verbais propostos neste estudo. Tomemos como exemplo os sons produzidos por *líquidos* que parecem apresentar semelhanças com as fricativas: seria interessante analisar se o treino auditivo para a melhor percepção de sons gerados por *líquidos* favoreceria a percepção dos sons da fala caracterizados pelo modo de articulação fricativa. A classificação de sons não verbais proposta neste estudo poderá também ser uma mais valia na monitorização dos progressos obtidos com o treino auditivo, possibilitando ao Terapeuta da Fala verificar de forma objectiva a evolução do utente na percepção de sons não verbais.

O que se pretendia neste estudo era a produção de onomatopeias puramente fonéticas, e verificou-se haver uma grande proximidade sonora entre as onomatopeias e os sons naturais que pretendiam imitar. Este estudo apresenta assim uma outra área de interesse, relativa à génese das onomatopeias. Seria pertinente aplicar o mesmo estudo a falantes de outras línguas e observar se as mesmas relações fonético-acústicas eram originadas. Em relação à manipulação de sons não verbais em experiências de percepção, seria de interesse aumentar a amostra deste estudo nas várias categorias utilizadas. A manipulação dos estímulos poderia ser feita de novas formas, tais como variar a relação sinal-ruído em valores inferiores a -15 dB e manipular o sinal no padrão temporal.

BIBLIOGRAFIA

- American Speech and Hearing Association 2005. *Central Auditory Processing Disorders: Current status of research and implications for clinical practice*. American Journal of Audiology 5: 41-54.
- Barker, E. J., Dettman, S. J. e Dowell, R. C. 1997. *Habilitation: Infants and Young Children*. Singular, San Diego.
- Behlau, M. 2004. *Voz: O Livro do Especialista*. Revinter, Rio de Janeiro.
- Bell, A. G. 1996. *Hierarchy of Listening Skills - Auditory-verbal Therapy for Parents and Professionals*. Estabrooks, Washington.
- Boersma, P. e D. Weenink. 2005. *Praat: doing phonetics by computer*. version 4.3.11.
- Borden, L., K. Harris, e L. Raphael. 2003. *Speech Science Primer*. 4ª Edição. Lippincott Williams & Wilkins, New York.
- Brown, W. D., A. T. Smith, B. Moskalik, e J. Gabriel. 2006. *Animal Behaviour* 72: 225-233.
- Colton, R. C. J. e Hirano, M. 1996. *Understanding Voice Problems: A Physiological Perspective for Diagnosis and Treatment*. Lippincott Williams & Wilkins, New York.
- Cruz-Ferreira, M. 1999. Portuguese (European). In *Handbook of the International Phonetic Association*. Cambridge University Press. pp. 126-130.
- Delgutte, B. 1997. *Auditory Neural Processing of Speech*. pp. 507-538 in Hardcastle, J. e Laver, J. ed. *The Handbook of Phonetic Sciences*. Blackwell, Oxford.
- Dumont, A. 1995. *L'orthophoniste et l'enfant sourd*. Masson, Paris.
- Fabiani, M., V. A. Kazmerski, e Y. M. Cycowicz. 1996. *Naming norms for brief environmental sounds: Effects of age and dementia*. *Psychophysiology* 33: 462-475.
- Finitzo-Hieber, T., I. J. Gerling, N. D. Matkin, e E. Cherow-Skalka. 1980. *A sound effects recognition test for the pediatric audiological evaluation*. *Ear Hear* 1: 271-276.
- Garstang, M. 2004. *Long-distance, low-frequency elephant communication*. *Journal of Comportamental Physiology* 190: 791-805.
- Gaver, W. W. 1993. *What in the World Do We Hear?: An Ecological Approach to Auditory Event Perception*. *Ecological Psychology* 5: 1-29.
- Gelfand, S. 2001. *Essentials of Audiology*. Thieme, New York.

- Gygi, B. 2001. *Factors in the Identification of Environmental Sounds*. Tese de Doutorado Department of Psychology. Indiana University, Indiana.
- Gygi, B., G. R. Kidd, e C. S. Watson. 2004. *Spectral-temporal factors in the identification of environmental sounds*. Journal of the Acoustic Society of America 115: 1252-1265.
- Gygi, B., e V. Shafiro. 2006. General Functions and Specific Applications of Environmental Sound Research. *Frontiers of Bioscience* (submetido).
- Gygi, B., C. S. Watson, e G. R. Kidd. 1997. *Factors in the detection of environmental sounds*. J. Am. Acad. Audiol. 101: 3149.
- Hartwig, S. 2005. *Individual acoustic identification as a non-invasive conservation tool: an approach to the conservation of the African wild dog Lycaon pictus*. *Bioacustics*. 15: 35-50.
- Henrique, L. 1997. *Acústica Musical*. Fundação Calouste Gulbenkian, Lisboa.
- Huckvale, M. 2000. *Speech Filing System User Manual*. University College of London.
- Ikedo, Y., Jahns, G., Nishizu, T., Sato, K. and Morio, Y. 2003. *Individual identification of dairy cows by their voice*. *Precision Livestock Farming*, Berlin: 81-86.
- Katada, A., S. Nonaka, M. Adachi, T. Arakawa, I. A. Masanobu, T. Hayashi, D. L. Zealer, e Y. Harabuchi. 2004. *Functional Electrical stimulation of laryngeal adductor muscle restores mobility vocal folds*. *Neuroscience Research* 50: 153-159.
- Katz, J. S., .A., Henderson, D. 1992. *Central Auditory Processing: A Transdisciplinary View*. Mosby, Chicago.
- Keith, R. W. 1997. *Auditory Processing Disorders* In C. e. D. Clark, R. et al., ed. *Cochlear Implantation for infants and children*. Singular, San Diego.
- Kent, R. D., e C. Read. 2002. *The Acoustic Analysis of Speech*. Singular, Australia.
- Ladefoged, P. 2006. *A Course in Phonetics*. 5ª Edição. Thomson Wadsworth, Boston.
- Lengagne, T. 2001. *Temporal stability in the individual features in the calls of eagle owls (bubo bubo)*. *Behaviour* 138: 1407-1419.
- Ling, D. 1989. *Foundations of Spoken Language for Hearing Impaired Children*. Alexander Graham Bell Association for the Deaf, Washington.
- Loizou, P. 1999. *COLEA: a MATLAB software tool for speech analysis*. University of Texas at Dallas.
- Mateus, M. H. M., I. Falé, e M. J. Freitas. 2005. *Fonética e Fonologia do Português*. Universidade Aberta, Lisboa.
- Maxon, A. e B., D. 1992. *The Hearing Impaired Child – infancy through High-School Years*. Andover Medical Publishers, London.
- Meyer, J. 1978. *Acoustics and the Performance of Music*. Verlag das Musikinstrument, Frankfurt am Main.
- Monfort, M. e J., Adoración 2001. *Algo que decir*. Entha Ediciones, Madrid.
- Musiek, F. E. e R., William 2001. *Perspectivas Actuais em Avaliação Auditiva*. Editora Manole, São Paulo.

- Myers, L. L., T. R. Letowski, K. S. Abouchacra, J. T. Kalb, e E. C. Haas. 1996. *Detection and Recognition of octave-band sound effects*. J. Am. Acad. Audiol. 7: 346-357.
- Nogueira, R. S. 1950. *Estudos sobre Onomatopeias*. Livraria Clássica Editora, Lisboa.
- Pereira, L. D., e E. Schochat. 1997. *Processamento Auditivo Central - manual de avaliação*. Editora Lovise, São Paulo.
- Pribram, K. H. 1976. *Language in a Sociobiological frame*. Annals of the New York Academy of Sciences: Origins and Evolution of Language 280: 708-717.
- Range, F., e J. Fischer. 2004. *Vocal Repertoire of Sooty Mangabeys (Cercocebus torquatus atys) in the Tai National Park*. Ethology 110: 301-321.
- Remez, R. E., P. E. Rubin, D. B. Pisoni, e T. D. Carrell. 1981. *Speech perception without traditional speech cues*. Science 212: 947-950.
- Repp, B. 1987. *The sound of two hands clapping: an exploratory study*. J. Acoust. Soc. Am. 81: 1100-1109.
- Riede, T., L. Brunberg, G. Tembrock, H. Herzel, e K. Hammerschmidt. 2001. *The harmonics-to-noise ratio applied to dog barks*. J. Acoust. Soc. Am. 110: 2191-2197.
- Rodriguez, M. M. e A. H., Ibañez 2002. *Implantes Cocleares*. Masson, Madrid.
- Roeser, R. e D., M. . 2004. *Auditory Disorders in School Children*. Thieme, New York.
- Rousseau, B., I. Tateya, X. H. Lim, Munoz-Del-Rio, e D. A. Bless. 2006. *Investigation of phytochemicals on in-vitro hyaluronan production*. Journal of Voice 20: 443-451.
- Schnider, A., D. F. Benson, D. N. Alexander, e A. Schnider-Klaus. 1994. *Non-Verbal Environmental Sound Recognition after Unilateral Hemispheric Stroke*. Brain 117: 281-287.
- Shafiro, V. 2004. *Perceiving the Sources of Environmental Sounds with a varying number of spectral channels*. Tese de Doutorado. University of New York, New York.
- Shafiro, V., e B. Gygi. 2004. *How to select stimuli for environmental sound research and where to find them*. Behaviour Research Methods, Instruments Computers 36: 590-598.
- Stevens, K. H. 1998. *Acoustic Phonetics*. MIT Press, Cambridge.
- Studdert-Kennedy, M., e A. S. House. 1972. *Speech perception*. Pages 1-62 in J. Tobias, ed. *Foundation of modern auditory theory*. Academic Press, New York.
- Tanaka, K., M. Takada, e S.- I. Iwamiya. 2006. *Relationships between auditory impressions and onomatopoeic features for environmental sounds*. Acoustic Science Technology 27: 67-79.
- Thompson, M. D., Atcheson, J., Pious, C. C. 1995. *Zero to Three – A Curriculum for Parents, Parent Trainers, and Teachers of Very Young Hearing-Impaired Children*. University of Washington Press, London.
- Urrutia, R., e I. Marco. 2000. *Diagnostico y Tratamiento de Los Transtornos de la Voz*. Ed. Garsi, Madrid.
- Vanderveer, N. J. 1979. *Ecological acoustics: human perception of environmental sounds*. Tese de Doutorado. Dissertation Abstracts International 40: 4543B.
- Veloso, J. 1999. *Na Ponta da Língua: Exercícios de Fonética do Português*. Granito. Porto.

- Warren, W. H., e R. R. Verbrugge. 1984. *Auditory Perception of Breaking and Bouncing Events: A case study in ecological acoustics*. *Journal of Experimental Psychology: Human Perception and Performance* 10: 704-712.
- Xavier, M., e M. Mateus. 1990. *Dicionário de Termos Linguísticos*. Edições Cosmos, Lisboa.
- Xiaofeng, L., R. J. Logan, e R. E. Pastore. 1991. *Perception of acoustic source characteristics: Walking sounds*. *Journal of the Acoustic Society of America* 90: 3036-3049.

PUBLICAÇÕES NO ÂMBITO DO MESTRADO

Capitão, S. e L. M. T. Jesus (2006). Acoustics of Speech and Environmental Sounds. In *Proceedings of ISCA Tutorial and Research Workshop on Experimental Linguistics*, Atenas, Grécia, pp. 221-224.

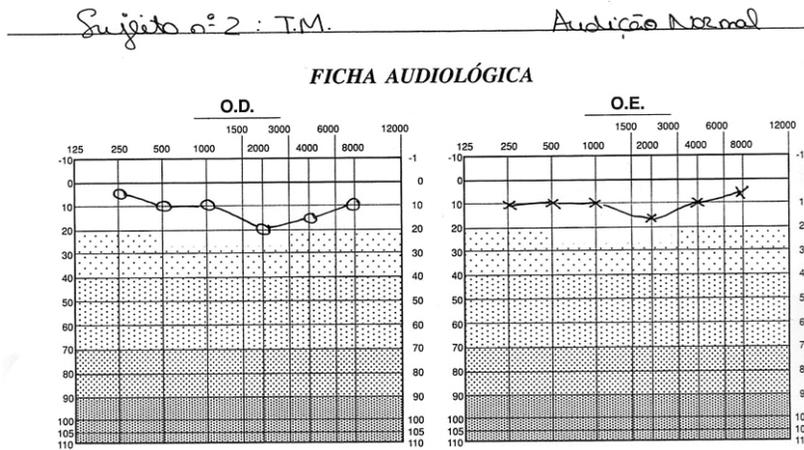
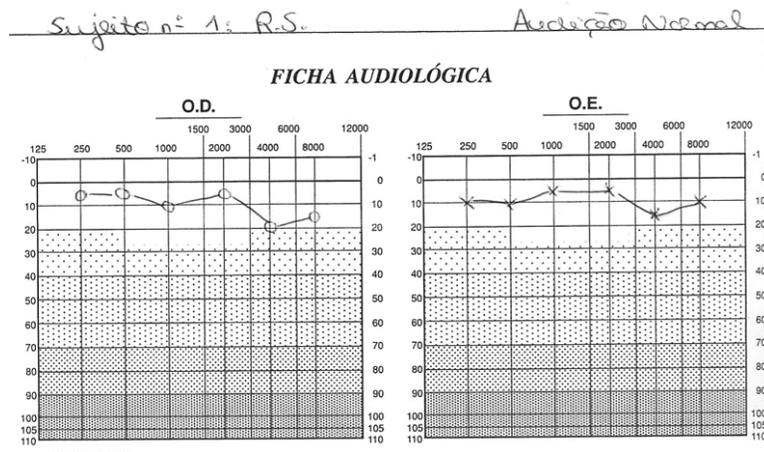
APRESENTAÇÕES ORAIS E EM CARTAZ NO ÂMBITO DO MESTRADO

Apresentação de Cartaz: “ Ressonância Magnética Funcional e o Processamento Central da Audição e do Discurso” , na 1ª Reunião do Hospital D. Estefânia: Audiologia Pediátrica e Workshops; Lisboa 2005.

Comunicação Oral em co-autoria: “ Contributo para a Padronização de Parâmetros Vocais na População Infantil falante de Português Europeu” , no IV Congresso Nacional da Associação Portuguesa de Terapeutas da Fala; Estoril 2005.

ANEXOS

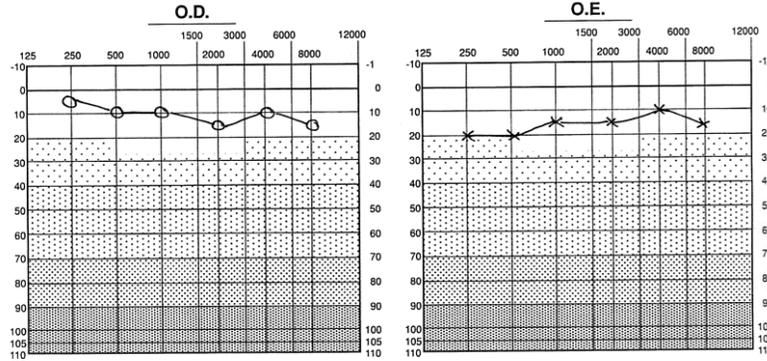
Avaliação Audiológica realizada pelo Dr. Mário Alves, Técnico de Audiologia, através de Audiometria Tonal Simples por via aérea.



Sujeito nº 3: B.R.

Audição Normal

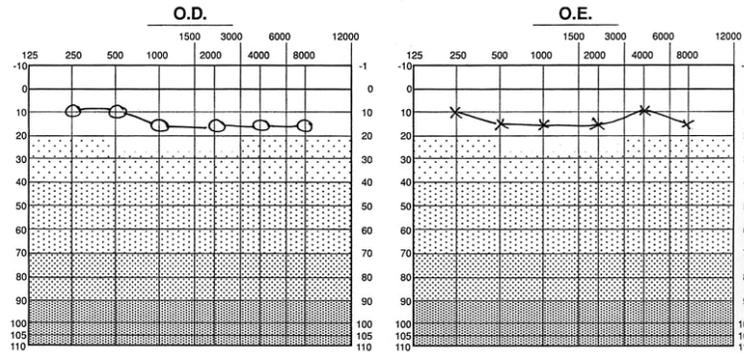
FICHA AUDIOLÓGICA



Sujeito nº 4: M.S.

Audição Normal

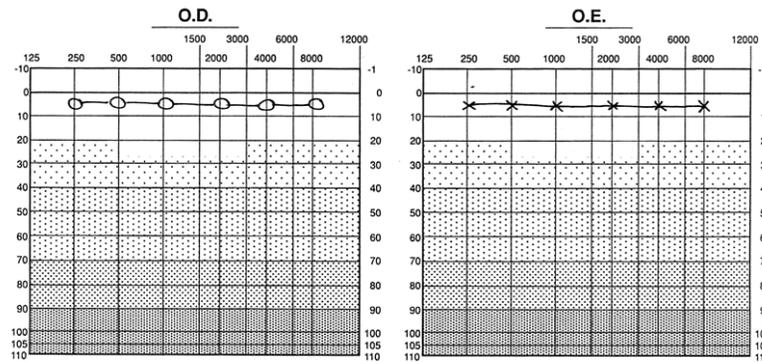
FICHA AUDIOLÓGICA



Sujeito nº 5: J.N.

Audição Normal

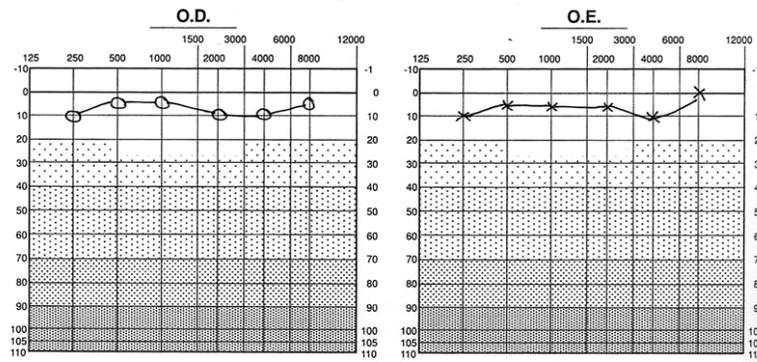
FICHA AUDIOLÓGICA



Sujeito nº 6: L.F.

Audição Normal

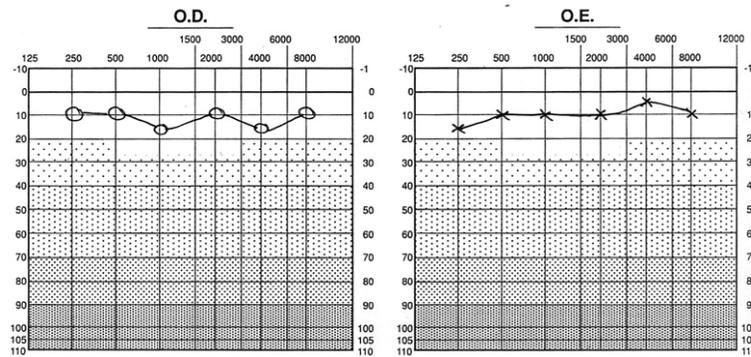
FICHA AUDIOLÓGICA



Sujeito nº 7: A.M.

Audição Normal

FICHA AUDIOLÓGICA



Sujeito nº 8: M.R.

Audição Normal

FICHA AUDIOLÓGICA

