



Universidade do Porto

---

Faculdade de Engenharia

**FEUP**

**A UTILIZAÇÃO DA RESSONÂNCIA MAGNÉTICA  
NA CARACTERIZAÇÃO FUNCIONAL DA FALA**

Sandra Moreira Rua Ventura

Janeiro de 2012



## **A UTILIZAÇÃO DA RESSONÂNCIA MAGNÉTICA NA CARACTERIZAÇÃO FUNCIONAL DA FALA**

Dissertação submetida ao Departamento de Engenharia Electrotécnica e  
de Computadores para satisfação parcial dos requisitos do Programa  
Doutoral em Engenharia Biomédica

### **Sandra Moreira Rua Ventura**

Licenciada em Radiologia pela Escola Superior de Tecnologia da  
Saúde do Porto (2002)

Mestre em Engenharia Biomédica pela Faculdade de Engenharia  
da Universidade do Porto (2007)

### **Orientador**

Diamantino Rui da Silva Freitas

Professor Associado do Departamento de Engenharia Electrotécnica e de  
Computadores da Faculdade de Engenharia da Universidade do Porto

### **Coorientador**

João Manuel R. S. Tavares

Professor Associado do Departamento de Engenharia Mecânica da  
Faculdade de Engenharia da Universidade do Porto

Janeiro de 2012



*Ao meu filho Martim*

*És a minha vida!  
O meu maior tesouro!  
Tu que tantas alegrias me dás,  
Tu que me fazes sorrir nos momentos em que estou triste...  
Tu que me fazes refletir sobre as mais pequenas coisas, coisas simples...  
Que possas ser soberano sem jamais perder a simplicidade,  
Que possas sorrir sem perderes a capacidade de chorar,  
Que possas ser Homem, sem perder a criança que há em ti,  
E principalmente, meu filho,  
Sorri e sê FELIZ!*



## AGRADECIMENTOS

Ao meu orientador, Prof. Dr. Diamantino Freitas, pelos momentos de reflexão, debate e sugestões prestadas ao longo de todo o meu percurso na FEUP, desde o Mestrado em Engenharia Biomédica até à data.

Ao meu coorientador, Prof. Dr. João Manuel Tavares, por toda a colaboração, disponibilidade prestadas, pelo incentivo, nomeadamente no apoio à produção científica.

À Professora Doutora Isabel Ramos e toda a sua equipa da Unidade de Ressonância Magnética do Serviço de Radiologia do Hospital de S. João, EPE, pela disponibilidade de recursos materiais, interesse no estudo e apoio prestados.

Ao meu colega e amigo, Vítor Silva, pela paciência e disponibilidade incansável durante a aquisição e recolha das imagens, e contribuição no estudo.

À Maria João Vasconcelos, companheira neste percurso de investigação e colaboração na aplicação dos modelos deformáveis ao estudo do trato vocal.

À Dra. Carla Moura, Médica Otorrinolaringologista do Hospital S. João, EPE, pela colaboração na aplicação dos questionários e palavras de incentivo.

Aos docentes e estudantes do Curso de Terapia da Fala da Escola Superior de Tecnologia da Saúde do Porto (ESTSP) do Instituto Politécnico do Porto (IPP), pela contribuição e apoio no que respeita ao domínio dos conceitos da Fala. Em especial, ao Terapeuta Pedro André Araújo, pela simpatia, disponibilidade, colaboração no estudo de casos e sugestões prestadas.

Aos estudantes do Curso de Radiologia da ESTSP-IPP, nomeadamente ao César Silva, Pedro Nuno Silva, Gisela Romano, Diana Atayde, Fernanda Silva, Beatriz Santos e Márcia Filipa Lima, pela sua colaboração durante a fase experimental de otimização das recolhas imagiológicas.

Às colegas, Susana Fafiães e Bebiana Carvalho pela ajuda prestada e simpatia.

À Mareike Sinderman, ao Rui Oliveira, ao Rodrigo Viterbo e à Inês Sofia pela participação e disponibilidade!

À Dra. Augusta Marques, pela disponibilização de materiais odontológicos e ajuda prestada nos ensaios para identificação dos dentes nas imagens de ressonância magnética.

À minha amiga Liliana por acreditar comigo, por me ouvir e estar presente.

À minha amiga Clarisse pela ajuda e tempo despendido.

À Dra. Ruth Sá pela colaboração prestada ao nível do Inglês.

À minha prima Manuela Moreira pelas sugestões ao nível do Francês.

À Escola Superior de Tecnologia da Saúde do Porto e ao Instituto Politécnico do Porto pela atribuição da Bolsa de Doutoramento no âmbito do Programa de Formação Avançada de Docentes do IPP e apoios concedidos.

E ainda, ao meu marido, filho, pais e irmão pelo incentivo a cada instante e por acreditarem comigo!

Obrigada a todos!



## SUMÁRIO

A articulação da fala constitui, pela sua complexidade, um desafio do ponto de vista das técnicas de análise, nomeadamente da imagiologia, para a adequada observação dos eventos articulatórios durante a produção da fala. Apesar das várias técnicas atualmente disponíveis, a informação morfológica e dinâmica é ainda exígua. A utilização da ressonância magnética (RM) neste âmbito, tem vindo a crescer, não existindo porém qualquer protocolo de estudo dirigido ao trato vocal. Assim, o tema central desta Tese está centrado na caracterização funcional da fala através da RM, e justifica-se principalmente com motivos clínico-terapêuticos bem como de engenharia.

Para o desenvolvimento deste projeto, procurou-se por um lado, o reconhecimento das etapas necessárias para o desenho, implementação e otimização de um protocolo de RM dedicado, e por outro, a exploração de técnicas de processamento e análise de imagens para a caracterização objetiva e quantitativa da morfologia e dinâmica articulatória. Por fim, no sentido de demonstrar a aplicabilidade do protocolo de RM desenvolvido, procedeu-se a um estudo de casos para avaliar diferentes contextos de voz profissional.

Pela análise estatística realizada, ficou demonstrado com clareza o modo como o estudo da articulação da fala se adequa às demais modalidades de voz profissional por comparação à voz falada de cada sujeito. De facto, tanto no canto como no teatro, foram encontradas diferenças nas cavidades de ressonância pela análise dos volumes extraídos do estudo estático 3D, indicando para a maioria dos sons estudados um aumento do volume das cavidades de ressonância, principalmente em resultado do aumento do volume da cavidade bucal no caso das cantoras, e também da faringe no caso do ator.

Relativamente às diferenças entre as médias dos resultados obtidos para cada medida articulatória, no contexto de uma melodia e para as quatro vozes em estudo, foi possível constatar que 80% destas medidas são significativamente afetadas pela variação tonal, destacando-se a altura da laringe e a protrusão da mandíbula com maior poder estatístico de significância ( $p < .05$ ).

Por fim, pelas várias aplicações práticas realizadas, foi possível demonstrar a utilidade e adequação da técnica de RM, de forma não-invasiva, no estudo da articulação da fala.

**PALAVRAS-CHAVE:** TRATO VOCAL, PROTOCOLOS DE IMAGEM POR RESSONÂNCIA MAGNÉTICA, ESTUDO MORFOLÓGICO-DINÂMICO, DINÂMICA ARTICULATÓRIA, ARTICULAÇÃO DA FALA, VOZ PROFISSIONAL, MODELOS TRIDIMENSIONAIS, PROCESSAMENTO E ANÁLISE DE IMAGEM.



## SUMMARY

Due to its complexity, the speech articulation is a challenge for the analytic techniques, including the ones of medical imaging, to observe and characterize accurately the articulatory events that occur during speech production. Despite the various techniques available nowadays, the dynamic and morphological information about this human motor activity is still scarce. Although the growing use of magnetic resonance imaging (MRI) that has been occurring recently in the last years, there is no vocal tract-oriented protocol. Thus, the main purpose of this Thesis is focused on the functional characterization of the speech production using MRI, in order to provide new insights, mainly for clinical, therapeutic and engineering uses.

This project was developed, firstly, with the identification of the steps required for the design, implementation and optimization of an MRI protocol vocal tract-oriented, and secondly, by the exploration of computational techniques of image processing and analysis for the real and quantitative description of the morphology and articulatory dynamics occurring during speech production. Finally, to demonstrate the applicability of the MRI protocol developed, a study case was addressed regarding diverse contexts of professional voice.

Based on the statistical analysis performed, it was possible to demonstrate the manner how the speech articulation study outfits with other styles of professional voice in comparison to the spoken voice of each subject under evaluation. In both singing and theatre voices studied, for the most sounds studied, there were observed differences in the resonance cavities analysis of the volumes extracted from the 3D static data, leading to an increase of the resonance cavity volumes, mainly as a result of the increase of the singers' oral cavity volumes, and also of the actor's pharynx volume.

Regarding the differences between the averaged results for each articulatory measure studied, 80% of these measures were significantly affected by tones in the melody framework and for the four vocal registers in study, highlighting the larynx's height and the mandible protrusion with values statistically significant ( $p < .05$ ).

Finally, the various practical applications performed allowed to demonstrate the usefulness and suitability of MRI, to study the speech articulation in a noninvasively way.

**KEYWORDS:** VOCAL TRACT, MAGNETIC RESONANCE IMAGING PROTOCOLS, MORPHOLOGICAL DYNAMIC STUDY, ARTICULATORY DYNAMICS, SPEECH ARTICULATION, PROFESSIONAL VOICE, THREE-DIMENSIONAL MODELS, IMAGE PROCESSING AND ANALYSIS.



## RÉSUMÉ

L'articulation de la parole, malgré sa complexité, représente un défi vu des techniques analytiques, et plus principalement pour l'imagerie et pour l'observation des mouvements articulatoires lors de la production de la voix. Nonobstant les différentes techniques actuellement disponibles, les informations morphologiques et dynamiques, sont encore insuffisantes. L'utilisation de l'imagerie par résonance magnétique (IRM) dans ce domaine a grandi, cependant il n'existe pas de protocole d'étude dédié à l'appareil vocal. Ainsi le thème de cette Thèse est orienté vers la caractérisation fonctionnelle de la parole basée sur l'imagerie par IRM, afin d'offrir de nouvelles perspectives clinico-thérapeutiques et d'ingénierie.

Ce projet a suivi, dans un premier temps, la recherche des étapes nécessaires à la conception, la mise en œuvre et l'optimisation d'un protocole d'IRM, pour ensuite, explorer les techniques de processus et analyse d'images pour la caractérisation objective et quantitative de la morphologie et de la dynamique articulatoire. Finalement, une étude de cas a été menée pour démontrer l'applicabilité du protocole d'IRM résultant, et ainsi évaluer les différents modalités de la voix professionnelle.

En ce qui concerne l'analyse statistique effectuée, nous avons pu démontrer comment l'articulation de la parole s'adapte face aux différentes modalités de voix professionnelle par rapport à la voix parlée de chaque individu. Entre le chant et le théâtre, des différences ont été trouvées dans l'analyse de les cavités résonnantes des volumes mesurée à partir de l'étude 3D statique, la majorité des phonèmes en étude ayant conditionnée un agrandissement des cavités résonnantes, principalement résultant du volume de la cavité buccale pour les chanteurs, et aussi du pharynx chez l'acteur.

En ce qui concerne les différences entre les moyennes des résultats pour chaque mesure articulatoire, dans le contexte d'une mélodie et pour les quatre voix, il a été constaté que 80% de ces mesures sont sensiblement affectées par la variation tonale, notamment pour la hauteur du larynx et protrusion de la mandibule qui présentent une force supérieure statistiquement significative ( $p < .05$ ).

En conclusion, les différentes applications pratiques réalisées, ont permis de démontrer l'utilité et la pertinence de l'IRM pour l'étude non invasive de l'articulation de la parole.

**MOTS-CLES:** APPAREIL VOCAL, PROTOCOLES DE L'IMAGERIE PAR RESONANCE MAGNETIQUE, ÉTUDES MORPHOLOGIQUE ET DYNAMIQUE, L'ARTICULATION DE LA PAROLE, ARTICULATOIRE DYNAMIQUE, VOIX PROFESSIONNELLE, MODELES A TROIS DIMENSIONS, TRAITEMENT ET ANALYSE DES IMAGES.



## ÍNDICE

<b>CAPÍTULO 1. INTRODUÇÃO À DISSERTAÇÃO E À SUA ESTRUTURA.....</b>	<b>1</b>
1.1 Introdução.....	2
1.2 Objetivos e abordagem seguida.....	3
1.3 Estrutura da Dissertação.....	6
1.4 Contribuições alcançadas .....	9
<b>CAPÍTULO 2. TÉCNICAS DE AVALIAÇÃO DA ARTICULAÇÃO DA FALA .....</b>	<b>13</b>
2.1 Produção da voz .....	14
2.2 Articulação da fala.....	16
2.2.1 Trato vocal e cavidade nasal.....	18
2.2.2 Técnicas de análise da dinâmica articulatória.....	20
2.3 Estimação dos formantes e classificação fonética .....	21
2.4 Conclusão.....	24
<b>CAPÍTULO 3. ESTUDO DAS VIAS AÉREAS SUPERIORES POR RESSONÂNCIA MAGNÉTICA.....</b>	<b>27</b>
3.1 Princípios físicos da ressonância magnética.....	28
3.1.1 O núcleo de Hidrogénio e o fenómeno de ressonância.....	28
3.1.2 O sinal de ressonância magnética.....	29
3.2 Instrumentação e segurança em ressonância magnética .....	31
3.2.1 Linhas de orientação da Food and Drug Administration.....	34
3.2.2 Controlo de qualidade em ressonância magnética .....	36
3.3 Técnicas de obtenção de imagens.....	37
3.3.1 Formação de imagens e espaço K.....	40
3.3.2 Parâmetros de aquisição e qualidade da imagem.....	41
3.3.3 Artefactos de imagem .....	44
3.4 Sequências de pulsos e a sua aplicação no estudo das vias aéreas superiores	46
3.4.1 Técnicas de sincronização.....	48
3.4.2 Monitorização de parâmetros fisiológicos.....	51
3.4.3 Potenciais aplicações clínicas no estudo das vias aéreas superiores .....	52
3.5 Conclusão.....	56
<b>CAPÍTULO 4. UTILIZAÇÃO DA RESSONÂNCIA MAGNÉTICA NA CARACTERIZAÇÃO DA FALA .....</b>	<b>59</b>
4.1 Perspetiva histórica: desenvolvimento da técnica e principais resultados obtidos.....	60
4.1.1 Estudos estáticos .....	61
4.1.2 Estudos dinâmicos .....	67
4.1.3 Estudos do Português Europeu .....	72

<b>4.2 Condicionantes da ressonância magnética na avaliação da articulação da fala.....</b>	<b>73</b>
<b>4.3 Conclusão.....</b>	<b>76</b>
<b>CAPÍTULO 5. PROTOCOLO DE ESTUDO DO TRATO VOCAL POR RESSONÂNCIA MAGNÉTICA.....</b>	<b>79</b>
<b>5.1 Desenho do protocolo .....</b>	<b>80</b>
<b>5.2 Implementação e otimização do protocolo.....</b>	<b>85</b>
5.2.1 Estudo estático 2D e 3D.....	86
5.2.2 Estudo dinâmico 2D.....	92
<b>5.3 Utilização da ressonância magnética de alto campo: reflexão .....</b>	<b>94</b>
<b>5.4 Conclusão.....</b>	<b>96</b>
<b>CAPÍTULO 6. PROCESSAMENTO E ANÁLISE DE IMAGENS DO TRATO VOCAL.....</b>	<b>99</b>
<b>6.1 Processamento e análise de imagens de ressonância magnética .....</b>	<b>100</b>
<b>6.2 Morfometria 2D/3D do trato vocal.....</b>	<b>102</b>
6.2.1 Interesse Clínico.....	104
6.2.2 Interesse em Engenharia .....	111
<b>6.3 Medidas articulatórias a partir de imagens sequenciais .....</b>	<b>113</b>
<b>6.4 Identificação dos dentes em imagens de RM.....</b>	<b>116</b>
<b>6.5 Conclusão.....</b>	<b>119</b>
<b>CAPÍTULO 7. ESTUDO DE CASOS: VOZ FALADA E VOZ CANTADA .....</b>	<b>121</b>
<b>7.1 Introdução.....</b>	<b>122</b>
7.1.1 Diferenças entre os contextos de voz profissional .....	123
<b>7.2 Metodologia .....</b>	<b>125</b>
7.2.1 Participantes e corpus da fala.....	125
7.2.2 Equipamento e protocolo de RM.....	128
7.2.3 Processamento e análise das imagens.....	130
7.2.4 Incerteza das medições.....	132
<b>7.3 Resultados e discussão .....</b>	<b>134</b>
7.3.1 Caracterização morfológica do trato vocal.....	134
7.3.2 Caracterização da dinâmica articulatória .....	143
<b>7.4 Conclusão.....</b>	<b>151</b>
<b>CAPÍTULO 8. CONCLUSÕES FINAIS E PERSPETIVAS FUTURAS.....</b>	<b>155</b>
<b>8.1 Conclusões finais.....</b>	<b>156</b>
<b>8.2 Perspetivas futuras .....</b>	<b>161</b>



<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>165</b>
<b>ANEXOS.....</b>	<b>177</b>
ANEXO 1. QUESTIONÁRIO	172
ANEXO 2. GUIA BÁSICO – INSTRUÇÕES DE PREPARAÇÃO E COLABORAÇÃO NO ESTUDO	174
ANEXO 3. QUESTIONÁRIO DE SEGURANÇA EM RM	175
ANEXO 4. FORMULÁRIO DE CONSENTIMENTO INFORMADO	177



## ÍNDICE DE FIGURAS

Figura 2.1 Órgãos do aparelho fonador num corte sagital de RM e respetiva segmentação e identificação do trato vocal e respetivos articuladores (VENTURA et al., 2011b).	14
Figura 2.2 Contornos sagitais da deformação da língua e respetiva posição na cavidade bucal extraídos a partir das imagens de RM, para a produção das vogais abertas do Português (VENTURA et al., 2008).	17
Figura 2.3 Regiões do trato vocal propriamente dito e suas relações anatómicas numa imagem sagital de RM.	18
Figura 2.4 Diferenças anatómicas do trato vocal num sujeito do sexo masculino (A) e feminino (B), visualizadas numa imagem sagital de RM.	19
Figura 2.5 Configuração do trato vocal durante a produção de uma vogal (A) e de uma consoante (B) com base em imagens sagitais de RM. Em (A) o trato vocal não apresenta qualquer obstrução/ constrição ao contrário da imagem (B).	19
Figura 3.1 Representação das diferentes velocidades de magnetização e desmagnetização de uma pequena área do trato vocal.	31
Figura 3.2 Localização do corte ao longo do gradiente e da frequência central do pulso de RF.	38
Figura 3.3 Localização espacial do sinal de RM num estudo do trato vocal e sua correlação com os planos anatómicos.	39
Figura 3.4 Imagens sagitais do trato vocal ponderadas em T1 e em T2. A gordura apresenta maior intensidade de sinal na imagem ponderada em T1. Na imagem ponderada em T2, observa-se maior intensidade de sinal no canal medular pela presença do líquido céfalo-raquidiano (seta).	43
Figura 3.5 Artefactos encontrados nos estudos do trato vocal por RM. A imagem (a) ilustra o artefacto de aliasing encontrando-se a parte posterior da cabeça (seta) à frente da parte anterior. Em (b) as setas ilustram dois artefactos metálicos causados por material chumbíneo existente nos dentes. E, a imagem (c) obtida na sequência do estudo dinâmico exemplifica o artefacto de movimento pela falta de definição (distorção) dos contornos de todo o trato vocal.	45
Figura 5.1 Posicionamento e dispositivos básicos para estudos de RM da cabeça e pescoço. Imagem de fundo: sistema de RM Siemens Magnetom® Trio 3.0 T.	83
Figura 5.2 Conjunto de etapas básicas para a instrumentação de um protocolo de imagem por RM.	84
Figura 5.3 Representação esquemática para programação das técnicas de aquisição de imagens usando multi-stacks (A) e um slab (B).	89
Figura 6.1 Segmentação dos contornos do trato vocal ao longo de várias imagens de RM, pelo método de Level-set e respetivas regiões de interesse obtidas.	105

- Figura 6.2 Análise comparativa da morfometria do trato vocal durante a produção de 5 sons vocálicos extraída a partir da técnica de aquisição volumétrica (VOL) e multi-stacks (MS). 106
- Figura 6.3 Sequência de tarefas de processamento de imagem realizadas para modelação 3D do trato vocal. As quatro janelas de visualização permitem executar a segmentação simultânea nos três planos ortogonais e visualizar o respetivo modelo 3D reconstruído. 107
- Figura 6.4 Modelos 3D do trato vocal reconstruídos a partir da aquisição volumétrica durante a produção de consoantes fricativas não vozeadas (A) e vozeadas (B). 109
- Figura 6.5 Módulos constituintes de um sistema de síntese articulatória. 112
- Figura 6.6 Gráfico com a distribuição das áreas seccionais extraídas desde a glote até aos lábios e respetiva localização dos cortes perpendiculares ao trato vocal na imagem médio sagital de RM. 113
- Figura 6.7 Medidas articulatórias realizadas imagens sequenciais de RM durante a produção da fala (VENTURA et al., 2011d). 115
- Figura 6.8 Análise comparativa das distâncias medidas para a vogal [a] quando produzida em contexto articulatório vogal-vogal (VV) ou consoante-vogal (CV), para cada sujeito. 116
- Figura 6.9 (A) Fotografia da moldeira. (B) Imagem sagital 2D de RM obtida durante a produção da vogal /a/, com utilização da moldeira impregnada com vaselina (indicada pelas setas) e marcas de óleo de fígado de bacalhau (sinalizadas por asteriscos). 118
- Figura 6.10 Tarefas de processamento e análise de imagem realizadas para a extração da morfologia 3D dos dentes da arcada superior a partir de imagens de RM, usando uma moldeira impregnada com vaselina. 119
- Figura 7.1 Parâmetros articulatórias definidos para caracterização da dinâmica articulatória a partir das imagens sequenciais de RM. 131
- Figura 7.2 Imagens sagitais 2D de RM obtidas durante a sustentação de três vogais, no contexto de voz falada de cada um dos sujeitos. 135
- Figura 7.3 Estudo estático 2D do padrão patológico “escape nasal” por RM, na palavra /faca/, observando-se a abertura do véu do palato (indicado pela seta representada) provocando a saída dos sons pela cavidade bucal e nasal, e ainda a modificação da morfologia global do trato vocal em resultado dessa alteração. 136
- Figura 7.4 Imagens sagitais 2D de cada um dos sujeitos obtidas durante a produção da vogal /a/ nos contextos de voz falada versus cantada/teatralizada e no padrão patológico de sigmatismo e respetivos modelos 3D. 142
- Figura 7.5 Notação musical da melodia “Parabéns a Você” e respetivo agrupamento das três variações tonais para o conjunto de sons musicais recolhidos. 148

## ÍNDICE DE TABELAS

Tabela 2.1 Breve descrição das técnicas de medição funcional da produção da voz.	15
Tabela 2.2 Descrição sumária das principais técnicas de análise da articulação da fala.	21
Tabela 2.3 Características distintivas das classes fonéticas do Português Europeu (adaptado de RIBEIRO, 2003).	22
Tabela 2.4 Descrição das classes fonéticas do Português Europeu de acordo com os alfabetos IPA e SAMPA (adaptado de RIBEIRO, 2003).	23
Tabela 3.1 Resumo do processo de formação de imagem por RM.	31
Tabela 3.2 Limites de segurança em ressonância magnética (Food and Drug Administration, 2003).	35
Tabela 3.3 Análise comparativa das diferentes técnicas de obtenção de imagens por RM e principais características (adaptado de LUFKIN, 1999).	37
Tabela 3.4 Eixos dos gradientes para cada plano de corte.	39
Tabela 3.5 Parâmetros de aquisição em RM determinantes do tipo de ponderação e da qualidade de imagem.	41
Tabela 3.6 Fatores que aumentam a SNR e respectivas consequências.	44
Tabela 3.7 Descrição sumária das sequências de pulsos básicas usadas em ressonância magnética (adaptado de LUFKIN, 1999).	47
Tabela 5.1 Comparação dos diferentes ensaios e tipos de estudo 2D realizados e respetivo efeito no tempo de aquisição <i>versus</i> resolução das imagens.	87
Tabela 5.2. Comparação dos diferentes ensaios e tipos de estudo 3D realizados e respetivo efeito no tempo de aquisição <i>versus</i> resolução das imagens.	88
Tabela 5.3 Vantagens e limitações das várias técnicas de obtenção de imagens por RM: 3D <i>multi-stacks</i> e 3D <i>Slab</i> .	90
Tabela 5.4 Comparação dos diferentes ensaios e tipos de estudo 3D realizados e respetivo efeito no tempo de aquisição <i>versus</i> resolução das imagens.	91
Tabela 5.5 Comparação dos diferentes ensaios e tipos de estudo realizados para a implementação e otimização do estudo dinâmico durante a produção da fala.	93
Tabela 7.1 Características distintivas dos contextos de voz falada <i>versus</i> cantada/teatralizada (adaptado de BEHLAU & REHDER, 1997).	123
Tabela 7.2 Avaliação perceptiva e acústica dos três participantes para caracterização do registo de voz.	127
Tabela 7.3 Tipos de estudo realizados para avaliação das diferentes modalidades de voz profissional.	128
Tabela 7.4 Protocolo de RM para estudo da voz falada <i>versus</i> cantada.	130

Tabela 7.5 Parâmetros articulatórios definidos para avaliação da dinâmica articulatória e respetiva medição pretendida.	131
Tabela 7.6 Determinação da confiabilidade do avaliador pelo cálculo do ICC, quanto à consistência das medições e concordância absoluta dos avaliadores.	133
Tabela 7.7 Volumes 3D do trato vocal e das principais cavidades de ressonância durante a produção das três vogais no contexto da voz falada de cada um dos sujeitos.	137
Tabela 7.8 Comparação dos volumes 3D do trato vocal e das principais cavidades de ressonância durante a produção das três vogais nos contextos da voz falada <i>versus</i> cantada.	138
Tabela 7.9 Comparação dos volumes 3D do trato vocal e das principais cavidades de ressonância durante a produção das três vogais no contexto de voz teatralizada.	139
Tabela 7.10 Comparação dos volumes 3D do trato vocal e das principais cavidades de ressonância durante a produção das três vogais em dois padrões patológicos.	141
Tabela 7.11 Análise quantitativa das medidas articulatórias (independentemente da variação tonal) em arpejo para as duas cantoras a partir das 30 imagens de RM obtidas.	144
Tabela 7.12 Análise quantitativa da variação tonal em arpejo para as duas cantoras a partir do estudo dinâmico de RM.	145
Tabela 7.13 Medidas articulatórias obtidas para cada modalidade de voz dos sujeitos agrupadas pelas variações tonais: grave, intermédio e agudo, no contexto de uma melodia.	149
Tabela 7.14 Testes de Kruskal-Wallis para avaliar a diferença entre as médias dos resultados obtidos para cada medida articulatória no conjunto dos três grupos de tons produzidos pelas modalidades de voz no contexto da melodia.	150

## LISTA DE SIGLAS, ABREVIATURAS E SÍMBOLOS

$\gamma$	Razão giromagnética
$\omega_0$	Frequência de precessão ou de Larmor
®	Marca registada
°	Grau
<sup>13</sup> C	Carbono
<sup>19</sup> F	Flúor
<sup>1</sup> H	Hidrogénio
2D	Bidimensional
<sup>31</sup> P	Fósforo
3D	Tridimensional
AL	Abertura labial
AM	Altura da mandíbula
ATD	Ator de Dobragens
B <sub>0</sub>	Campo magnético principal
BW	Banda de Receção
CFA	Codificação de Fase
CFR	Codificação de frequência
cm	Centímetros
CV	Consoante-vogal
dB	Decibel
DICOM	Digital Imaging and Communications in Medicine
ECG	Eletrocardiografia
EGG	Eletroglotografia
EMA	Articulografia eletromagnética
EPE	Entidade Pública Empresarial
EPG	Eletropalatografia
EPI	Localização da epiglote
ETL	Fator turbo (do Inglês, <i>Echo Train Length</i> )
EUA	Estados Unidos da América
F <sub>0</sub>	Frequência fundamental
F1	Primeira formante
F2	Segunda formante
FA	Flip angle
FDA	Food and Drug Administration
FID	Free Induction Decay
FISP	<i>Fast Imaging with Steady State Precession</i>
FLASH	<i>Fast Low Angle SHot</i>
FOV	Field Of View
FSE	<i>Fast Spin-eco</i>
GAP	Espaçamento entre cortes
GE	Gradiente-eco
GLO	Abertura glotal
GRAPPA	<i>Generalized Autocalibrating Partially Parallel Acquisition</i>
GRASS	<i>Gradient Recalled Acquisition in the Steady State</i>
GSC	Gradiente de Seleção de Corte

<b>Gx</b>	Gradiente aplicado ao longo do eixo x
<b>Gy</b>	Gradiente aplicado ao longo do eixo y
<b>Gz</b>	Gradiente aplicado ao longo do eixo z
<b>Hz</b>	Hertz
<b>ICC</b>	<i>Intraclass Correlation Coefficient</i>
<b>IPA</b>	International Phonetic Alphabet
<b>JAZ</b>	Jazz
<b>LAL</b>	Localização do ápex da língua
<b>LAR</b>	Altura da laringe
<b>LDL</b>	Localização do dorso da língua
<b>LEDs</b>	Luz de díodos foto emissores (do Inglês, <i>Light-emitting diode</i> )
<b>LIR</b>	Lírico
<b>MHz</b>	Megahertz
<b>MIPAV</b>	<i>Medical Image Processing, Analysis and Visualization</i>
<b>mm</b>	Milímetros
<b>ms</b>	Milissegundos
<b>MS</b>	<i>Multi-stacks</i>
<b>mT/m</b>	Militesla por metro
<b>n.º</b>	Número
<b>NEX</b>	Número de excitações
<b>NIFTI</b>	<i>Neuroimaging Informatics Technology Initiative</i>
<b>ORO</b>	Largura da orofaringe
<b>p</b>	Percent sampling
<b>PACS</b>	<i>Picture Archiving and Communication System</i>
<b>PCA</b>	<i>Principal Component Analysis</i>
<b>PE</b>	Português Europeu
<b>PM</b>	Protrusão da mandíbula
<b>px/mm</b>	<i>Pixel</i> por milímetro
<b>r</b>	coeficiente de correlação de Pearson
<b>RF</b>	Radiofrequência
<b>RM</b>	Ressonância Magnética
<b>ROI</b>	Região de interesse (do Inglês, <i>Region Of Interest</i> )
<b>r<sub>s</sub></b>	coeficiente de correlação de Spearman
<b>s</b>	Segundos
<b>SAMPA</b>	Speech Assessment Methods Phonetic Alphabet
<b>SAOS</b>	Síndrome da Apneia Obstrutiva do Sono
<b>SAR</b>	<i>Specific Absorption Ratio</i>
<b>SE</b>	Spin-eco
<b>SNR</b>	<i>Signal-to-Noise Ratio</i>
<b>T</b>	Tesla
<b>T/m/s</b>	Tesla por metro por segundo
<b>T1</b>	Relaxação longitudinal
<b>T2</b>	Relaxação transversal
<b>TE</b>	Tempo de eco
<b>TF</b>	Terapeuta da Fala
<b>TI</b>	Tempo de Inversão
<b>Tk</b>	Espessura de corte (do Inglês, <i>Thickness</i> )
<b>TR</b>	Tempo de repetição
<b>VEL</b>	Abertura véu do palato



<b>VME</b>	Vetor de magnetização efetivo
<b>VOL</b>	Volumétrica
<b>VTAF</b>	<i>Vocal tract area function</i>
<b>VV</b>	Vogal-vogal
<b>W/Kg</b>	Watts por Quilograma
<b><math>\rho</math></b>	Valor-p ou <i>nível descritivo</i>



## **CAPÍTULO 1. INTRODUÇÃO À DISSERTAÇÃO E À SUA ESTRUTURA**

## 1.1 Introdução

A Imagiologia médica teve um crescimento espantoso durante o século XX, consensualmente aceite, tornando-se imprescindível na prática clínica e prevendo-se uma continuidade no seu desenvolvimento no século XXI; nomeadamente a sua utilização em novos domínios de grande relevância, como a biologia molecular, a investigação biomédica e a intervenção terapêutica não-invasiva (CHAN, 2002).

Este crescimento advém essencialmente dos avanços que se verificaram nas modalidades de tomografia computadorizada e de ressonância magnética (RM), nomeadamente a decorrente da introdução das aquisições volumétricas – a Imagiologia médica tridimensional (3D). A ideia é aparentemente simples e consiste em aproveitar os dados resultantes de aquisições seccionais e reconstruir computacionalmente a região examinada também 3D.

Enquanto que a visão endoscópica é obtida através de um modelo simplificado de iluminação, a imagem de RM corresponde a uma visão gerada por computação gráfica. A utilização da ressonância magnética para fins médicos, emerge só a partir da década de 1970, ainda que este fenómeno quântico, que ocorre em escala nuclear, tenha sido descoberto anteriormente por Felix Bloch e Edward Purcell, logo depois da II Guerra Mundial, valendo-lhes o prémio Nobel de Física em 1952 (PISCO & SOUSA, 1999).

A possibilidade de se estudar, de forma não-invasiva, qualquer região do corpo humano e sem recurso a radiação ionizante, asseguraram um posicionamento da RM na prática clínica de grande valor em variadas especialidades médicas.

A aliança entre a Imagiologia médica e a Engenharia abriu novos horizontes para o diagnóstico médico com repercussões na abordagem clínica aumentando a precisão e exatidão da análise (MARGULIS & SUNSHINE, 2000). Para além disso, os benefícios estenderam-se a outra vertente médica – a terapêutica, como por exemplo, no planeamento e na simulação de cirurgias.

Com base em algoritmos computacionais variados é agora possível a visualização de dados de diferentes modos. Para além disso, esta manipulação de dados permite que sejam examinados na sua totalidade e de uma forma

interativa: é possível aplicar transformações geométricas (como, translação, rotação, escalamento, isolamento) de determinadas características do volume de dados para melhor observação e compreensão. Acresce que o processo é não invasivo, não requer sedação, com menor desconforto para o paciente e que produz vistas impossíveis de ser obtidas por meios convencionais, sem perda de orientação e ultrapassando obstruções. Os exemplos mais conhecidos na prática clínica incluem a broncoscopia virtual e a colonoscopia virtual.

No domínio da caracterização da fala humana, apesar de várias técnicas de observação e análise estarem atualmente disponíveis, a informação morfológica e dinâmica dos órgãos envolvidos neste mecanismo é consideravelmente escassa, principalmente no que concerne à articulação da fala. Acresce a inexistência de um protocolo de RM dirigido para estudo do trato vocal, de modo a potenciar não só o diagnóstico de problemas articulatorios como a terapêutica.

A qualidade de imagem de RM é afetada por vários fatores, iniciando-se no processo de aquisição, na instrumentação utilizada, e determina a forma como as imagens são visualizadas, e conseqüentemente o diagnóstico baseado nas mesmas. A ressonância magnética permite o estudo dos tecidos moles com elevada resolução de contraste; no entanto, do ponto de vista da instrumentação e técnicas de aquisição, é um método complexo, cuja interpretação das imagens obtidas é dificultada pela presença de ruído e de artefactos.

A RM aplicada ao estudo da produção da fala, tem vindo a crescer, não existindo porém linhas de orientação ou protocolos dirigidos que garantam a sua utilização e implementação rotineira na prática clínica. Neste sentido, o tema desta Dissertação está centrado na utilização da RM na caracterização funcional da fala. Justifica-se pela necessidade da adoção de protocolos de estudo dirigidos às especificidades anatómicas e funcionais dos órgãos que garantam a qualidade das imagens e permitam assim, estimular a utilização da RM por parte dos clínicos ou terapeutas da fala.

## **1.2 Objetivos e abordagem seguida**

Para o desenvolvimento deste projeto determinaram-se os seguintes objetivos principais:

- Definir e implementar um protocolo de aquisição de imagens de ressonância magnética para estudo do trato vocal, em situações estáticas e dinâmicas para a caracterização da fala;
- Desenvolver as potencialidades técnicas da ressonância magnética e propor linhas de orientação e respectivas aplicações para um protocolo específico para o estudo da articulação da fala.

Dada a diversidade atual dos protocolos de RM para as variadas estruturas anatômicas e, considerando a inexistência de linhas de orientação para estudo do trato vocal durante a fala, assim como quanto às aplicações possíveis, definiram-se ainda os seguintes objetivos específicos:

- Identificar as características anatômicas e fisiológicas do trato vocal determinantes para a sua observação e análise através da técnica de RM;
- Explorar os requisitos técnicos e condicionantes da RM no estudo da articulação da fala;
- Conceber um protocolo de RM orientado e alicerçado em ensaios clínicos para diferentes resultados de análise;
- Descrever que características morfométricas do trato vocal podem ser extraídas e analisadas a partir das imagens de RM;
- Estabelecer linhas de orientação e apoio com vista à posterior utilização e implementação rotineira da RM para estudo do trato vocal na prática clínica;
- Demonstrar as várias aplicações da RM e áreas potenciais de interesse;
- Demonstrar a viabilidade da utilização da RM para caracterização qualitativa e quantitativa da articulação dos sons da fala.

Para o efeito, definiram-se duas linhas metodológicas principais, nomeadamente: 1) o reconhecimento dos passos ou etapas necessários para o desenho, implementação e otimização de um protocolo de RM dirigido ao estudo da articulação da fala, e 2) a exploração de técnicas de processamento e análise de imagem médica para caracterização objetiva e quantitativa do trato vocal durante a fala.

A metodologia seguida teve ainda por base um estudo experimental preliminar para caracterização imagiológica da RM, mediante um trabalho de campo prévio com os seguintes objetivos:

- Descrever e comparar as principais características dos equipamentos de RM de alto campo disponíveis no Hospital S. João Entidade Pública Empresarial (EPE), no Porto;
- Avaliar parâmetros geométricos, nomeadamente a relação sinal-ruído e resolução espacial, através da realização de ensaios com um fantoma criado para o efeito, de modo a reduzir a maioria das limitações ou condicionantes previamente identificadas no estudo desenvolvido no âmbito do Mestrado em Engenharia Biomédica (RUA & FREITAS, 2006a).

Do estudo realizado constatou-se ainda o valor acrescentado que advém da utilização de sistemas de RM com potência de 3.0 Tesla; nomeadamente, na superioridade destes parâmetros geométricos, em particularidade do aumento da relação sinal-ruído (ver secções 3.4.3 e 5.3).

Verificado o interesse clínico e de colaboração da instituição hospitalar acima referida, procedeu-se a uma revisão bibliográfica pormenorizada acerca do mecanismo de articulação da fala e das principais técnicas de avaliação, nomeadamente sobre as principais condicionantes existentes na aplicação da RM no estudo do trato vocal, quer do ponto de vista da técnica de Imagiologia quer do processamento das imagens bidimensionais (2D) e 3D obtidas. De seguida, procurou-se:

- 1) Instrumentar um protocolo de acordo com as necessidades da área da fala, quer do ponto de vista clínico quer da Engenharia, elaborando-se paralelamente toda a documentação de suporte e apoio às recolhas imagiológicas;
- 2) Realizar recolhas com um fantoma específico, para ajustar parâmetros de aquisição do protocolo, e/ou sempre que necessário, para solucionar problemas geométricos das imagens;
- 3) Efetuar recolhas programadas com voluntários para avaliar e otimizar o protocolo;
- 4) Utilizar técnicas estatísticas para avaliar e fundamentar o protocolo e as medidas extraídas pela análise das imagens obtidas.

No que respeita à instrumentação do protocolo de RM, inicialmente explorou-se o estudo estático e por fim, o estudo dinâmico. As recolhas imagiológicas

foram planeadas de forma sistemática e organizada, culminando na reavaliação dos resultados.

A otimização do protocolo permitiu a subsequente análise quantitativa e determinação de medidas articulatórias mediante a exploração de várias aplicações práticas.

Por fim, e no sentido de demonstrar uma aplicabilidade potencial da utilização do protocolo desenvolvido para uma situação prática concreta e ainda permitir uma análise estatística preliminar do mesmo, procedeu-se a um estudo de casos com voluntários profissionais na área da voz e sem historial de patologia oro-facial.

### **1.3 Estrutura da Dissertação**

Este documento está estruturado em oito capítulos independentes, sendo os três primeiros capítulos que se seguem centrados na introdução teórica de alguns fundamentos anatómicos e fisiológicos do trato vocal e ainda, na caracterização da técnica de ressonância magnética. Este conjunto de capítulos culmina na descrição do estado da arte relativamente aos estudos de investigação na área da produção da fala e utilização da ressonância magnética. A partir do Capítulo 5 são explorados os vários ensaios realizados e a direção metodológica seguida, assim como, demonstradas as principais aplicações desenvolvidas ao longo deste projeto.

Apresentam-se de seguida e de forma resumida o conteúdo estruturante de cada um dos restantes capítulos que constituem este documento:

- ***Capítulo 2.***

#### **Técnicas de Avaliação da Articulação da Fala**

Este capítulo está organizado em quatro secções, sendo descritos os órgãos envolvidos na produção da fala e, em particular, na articulação da fala, assim como, as suas principais técnicas de análise.

A produção da fala assenta em duas etapas fundamentais, a produção da voz e a articulação da fala, cujas especificidades anatómicas justificam diferentes técnicas de análise e abordagens de estudo descritas neste capítulo.



Por fim, são descritos os conceitos gerais de fonética, uma vez que este trabalho está centrado na caracterização da fala para a Língua Portuguesa, nomeadamente o Português Europeu que corresponde à designação dada à variedade desta Língua falada em Portugal continental, nas regiões autónomas da Madeira e dos Açores, e pelos emigrantes portugueses espalhados pelo mundo, englobando os dialectos regionais, vocabulário, gramática e ortografia. Segundo LEWIS (2009), a língua Portuguesa é a 7<sup>a</sup> língua mais falada no mundo.

• **Capítulo 3.**

**Estudo das Vias Aéreas Superiores por Ressonância Magnética**

Neste capítulo, organizado em seis secções, são apresentados os fundamentos teóricos básicos da ressonância magnética, bem como a instrumentação usada e as principais técnicas de obtenção das imagens e sincronização. Por fim, são destacadas as potenciais aplicações clínicas da ressonância magnética no estudo das vias aéreas superiores.

A produção da fala ocorre durante a expiração, seguindo-se a fonação ao nível da laringe formulando a sua fonética e por fim, dando-se a articulação destes sons pela atividade motora dos órgãos do trato vocal. Esta atividade motora é maioritariamente espontânea, mas pode ser uma imitação vocal como na repetição da fala. Assim e porque a implementação e utilização da RM no estudo das vias aéreas superiores tem vindo a crescer, atualmente já são reconhecidas várias áreas potenciais de aplicação clínica descritas neste capítulo.

• **Capítulo 4.**

**Utilização da Ressonância Magnética na Caracterização da Fala**

O desenvolvimento histórico da RM no estudo da fala, em diferentes áreas de aplicação, é explorado nas três secções que constituem este capítulo. O estado da arte relativamente aos estudos de investigação, desde a primeira utilização da RM no estudo da produção da fala por ROKKAKU e colaboradores (1986), são aqui relatados.

A utilização crescente da RM neste âmbito, descrita neste capítulo, ilustra as potencialidades técnicas e diferentes áreas de aplicação, em desenvolvimento,

assim como outros fenómenos que são alvo de estudo e aqui explanados, tais como a nasalidade, a coarticulação.

Apesar das condicionantes que ainda se verificam, a imagem por RM é, sem dúvida, uma técnica promissora no estudo da fala, inócua, não-invasiva, fornecendo informação fiável acerca da morfologia e dinâmica dos articuladores, tal como poderá ser constatado a partir deste capítulo.

- **Capítulo 5.**

- Protocolo de Estudo do Trato Vocal por Ressonância Magnética**

Este capítulo está organizado em quatro secções, onde são apresentados quais os parâmetros fundamentais necessários para o desenho de um protocolo de RM dirigido ao estudo do trato vocal. Adicionalmente, são enfatizados os aspetos técnicos da Imagiologia e benefícios da sua utilização, e descritos os vários ensaios realizados quer para caracterização estática como dinâmica dos articuladores.

Apesar das técnicas atualmente disponíveis e evolução tecnológica dos sistemas de RM, existem ainda alguns constrangimentos e desafios técnicos para avaliar a natureza funcional dos articuladores da fala. Torna-se assim imperativo o aperfeiçoamento das técnicas e instrumentação existentes de modo a aumentar a confiabilidade dos resultados.

- **Capítulo 6.**

- Processamento e análise de imagens do trato vocal**

As especificidades das imagens de RM aliadas à morfologia particular do trato vocal criam importantes condicionantes às tarefas de processamento e análise de imagem, justificando-se a exploração destas tarefas orientadas para a segmentação e caracterização do trato vocal.

Este capítulo, divide-se em quatro secções, sendo apresentadas quais as características das imagens de RM, do ponto de vista do seu processamento e análise computacional e ainda que medidas quantitativas podem ser extraídas a partir destas, usando técnicas de processamento e análise de imagem.

As várias aplicações práticas descritas destacam por um lado, a diversidade de dados morfométricos que podem ser extraídos com recurso a um protocolo

de RM adequado, e por outro, a necessidade de acompanhamento dos progressos na área do processamento e análise de imagem de modo a permitir a adequação das várias ferramentas disponíveis às particularidades deste tipo de estudo.

- **Capítulo 7.**

- **Estudo de Casos: Voz Falada e Voz Cantada**

Neste capítulo é apresentado um estudo de casos, organizado em quatro secções, exemplificativo de uma aplicação prática concreta da utilização do protocolo de RM para estudo da fala desenvolvido neste projeto. Após a descrição dos participantes e métodos usados, são destacados e discutidos os resultados obtidos.

O uso da voz falada é a função mais praticada pelo ser humano, no entanto, outras modalidades de voz profissional como a voz cantada ou a voz “teatral”, exigem domínios distintos e nem sempre são simples de executar. Neste capítulo, destaca-se assim a importância que a articulação desempenha nestas diferentes modalidades de voz e potencia um melhor entendimento e caracterização da variabilidade acústica intra e intersujeitos durante a produção da fala.

- **Capítulo 8.**

- **Conclusões finais e perspectivas futuras**

Este capítulo apresenta as conclusões finais sobre o trabalho de investigação desenvolvido, bem como as dificuldades encontradas e o posicionamento em relação aos objetivos clínico-terapêuticos previamente definidos. Destaca ainda os principais resultados experimentais obtidos e identificam-se perspectivas de desenvolvimento futuro, indicando e antecipando desafios e inovações existentes.

## **1.4 Contribuições alcançadas**

Durante todo o projeto, desenvolveram-se linhas orientadoras e metodológicas explorando-se as potencialidades da Imagiologia e da Engenharia na procura de um melhor entendimento do mecanismo da produção da fala, e em particular da articulação da fala.

Assim, podem ser destacadas as seguintes contribuições principais alcançadas neste projeto cujo enquadramento pode ser centralizado nos seguintes domínios:

*a) Estudos da fonação e do Português Europeu*

No que respeita ao mecanismo da produção da fala, foram desenvolvidas várias avaliações do trato vocal possibilitando a criação de uma base de dados morfológica para caracterização do Português Europeu, nomeadamente das vogais e consoantes fricativas (VENTURA et al., 2011a), e dinâmica, em particular no contexto de repetição de palavras e segundo uma sequência melódica, sustentada em medidas quantitativas (VENTURA et al., 2011c, d). Considerando a escassez de informação morfológica e dinâmica da produção da fala, esta Dissertação contribui assim para um melhor entendimento deste mecanismo, nomeadamente na avaliação clínica das cavidades de ressonância e dos articuladores envolvidos.

A primeira utilização da RM no estudo do Português Europeu que se realizou, para além de pioneira, teve um bom acolhimento junto da comunidade científica (publicada em VENTURA et al., 2007; 2008; 2009b; 2011b, e citada em pelo menos 5 publicações). Através das várias técnicas, quer de aquisição como de processamento e análise de imagem, foi possível a mensuração de aspetos particulares da fala, segundo diferentes abordagens.

*b) Técnica de ressonância magnética e sistematização na avaliação do trato vocal - protocolo*

Mediante o estudo exploratório das potencialidades de um sistema de RM de alto campo, utilizaram-se várias estratégias de aquisição de imagem, instrumentação e materiais de apoio, atendendo-se sempre às normas de segurança, permitindo a implementação efetiva de um protocolo de RM dirigido ao trato vocal. Espera-se que este conhecimento seja de utilidade futura, dada a atual carência de linhas de orientação para a implementação efetiva deste tipo de protocolo.

A criação do protocolo constitui *per se* um protótipo de estudo por RM, no qual figuram todas as etapas necessárias à implementação e otimização de um

método para uma finalidade de estudo, de utilização flexível, permitindo a futuros investigadores a definição de linhas metodológicas similares dirigidos a qualquer região anatômica.

*c) Processamento e análise de imagens médicas*

A diversidade de algoritmos e programas computacionais atualmente disponíveis possibilitou a definição de linhas metodológicas para a análise quantitativa do trato vocal durante a produção da fala a partir de imagens de RM, cujos dados permitiram dar sustentabilidade às opções protocolares tomadas, para além da demonstração de um conjunto de potenciais aplicações nas áreas afins da fala.

A utilização e integração dos modelos deformáveis na caracterização da dinâmica articulatória foi inovadora, pois apesar de existirem outras utilizações na área médica, no estudo da fala foram demonstradas várias potencialidades; nomeadamente, na segmentação e reconstrução automática do trato vocal para caracterização morfológica, e ainda na análise quantitativa da dinâmica articulatória (VASCONCELOS et al., 2010; 2011a, b).

*d) Aplicações práticas demonstradas*

Desenvolveu-se uma abordagem para determinar a morfometria 2D/3D do trato vocal, assim como para a caracterização da dinâmica articulatória, resultando em aplicações diversas quer para a Medicina quer para a Engenharia.

A aplicação desta metodologia, para a caracterização funcional da fala, foi na realidade um primeiro estudo descritivo e que permitirá o direcionamento do protocolo de RM de acordo com os resultados pretendidos.

A viabilidade do protocolo construído ficou demonstrada pelas variadas aplicações práticas realizadas, nomeadamente na avaliação concreta das funções da voz profissional. Esta aplicação, para além de inovadora nomeadamente no estudo das cavidades de ressonância do trato vocal, mostra a relevância que a articulação da fala tem na “facilitação” destas funções da voz no contexto profissional, intra e intersujeitos.

O estudo de casos desenvolvido, contribuiu para a definição de uma metodologia de estudo da voz profissional por RM e permitiu ainda a caracterização da articulação da fala em cantores e num ator de dobragens. Neste último caso, em referência à voz falada dos cantores e ator, o estudo foi também inovador, pois os únicos estudos descritos na literatura, realizados por ECHTERNACH e colaboradores (2010, 2011), apenas analisaram as diferenças de voz cantada na forma do trato vocal em cantores profissionais e em diferentes extensões vocálicas.

## **CAPÍTULO 2. TÉCNICAS DE AVALIAÇÃO DA ARTICULAÇÃO DA FALA**

A produção da fala é um importante mecanismo cujo interesse multidisciplinar tem motivado o desenvolvimento e aperfeiçoamento de inúmeras técnicas de análise. As especificidades anatômicas e funcionais dos órgãos envolvidos possibilitam a concretização das duas tarefas primordiais deste mecanismo: a produção da voz e a articulação da fala.

As técnicas de análise devem, por isso, atender a estas especificidades de forma a permitir um melhor entendimento, quer do ponto de vista dos estudos da fonação quer da Engenharia para a concepção de sistemas de reconhecimento e síntese da fala.

Assim, é imperativo responder às seguintes questões: Quais os órgãos envolvidos na articulação da fala? Como pode ser efetuada então a análise da dinâmica articulatória?

Neste capítulo são descritos os órgãos envolvidos na produção da fala e, em particular, na articulação da fala, assim como, as suas principais técnicas de análise.

## 2.1 Produção da voz

O aparelho fonador humano é constituído por órgãos do aparelho respiratório e do aparelho digestivo, podendo dividir-se em órgãos da fonação (responsáveis pela produção da voz) e em órgãos da articulação (responsáveis pela definição dos sons da fala), e que mutuamente se influenciam no processo da produção da fala, Figura 2.1. Os órgãos do aparelho fonador são controlados de forma independente, o que constitui uma habilidade específica do indivíduo (BENESTY et al., 2008). Os articuladores produzem ressonâncias ou modulações da fonte glotal, para a produção da voz, e também criam sons adicionais de algumas das consoantes. Os músculos constritores da faringe e da laringe também participam tanto na articulação como no controlo da qualidade da voz.

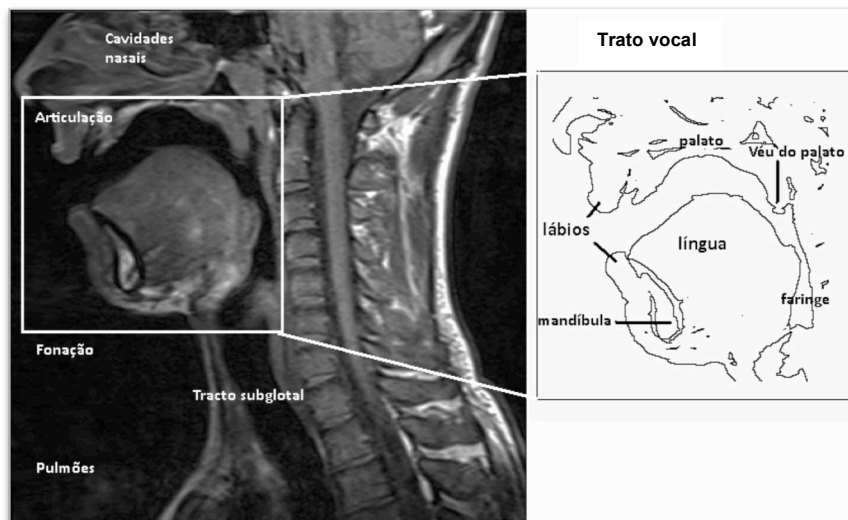


Figura 2.1 Órgãos do aparelho fonador num corte sagital de RM e respetiva segmentação e identificação do trato vocal e respetivos articuladores (VENTURA et al., 2011b).

A produção da voz requer a ação sincronizada e adequada do aparelho respiratório (pulmões e laringe) para a geração do fluxo de ar necessário, alterando-se assim o ciclo respiratório, caracterizado por uma fase inspiratória pequena em relação à fase expiratória que é maior (BENESTY et al., 2008). Para a produção de sons sonoros (vozeados), o fluxo de ar proveniente dos pulmões passa através da laringe, realizando-se a vibração das cordas vocais.

A laringe é um pequeno órgão da região cervical localizada acima da traqueia conectando-a com a faringe. Várias estruturas rígidas (cartilagens) constituem a



laringe, que mecanicamente está suspensa ao osso hióide por músculos e ligamentos. O componente harmônico mais baixo nos sons vozeados denomina-se de frequência fundamental ( $F_0$ ) e corresponde à frequência natural de vibração das cordas vocais. De um modo subjetivo, corresponde genericamente à nossa percepção da voz como aguda (“fina”) ou grave (“grossa”) e está relacionada, principalmente, com o tamanho e a espessura das cordas vocais. A frequência fundamental num homem varia entre 85 e os 400 Hz, enquanto nas mulheres é superior e varia entre 120 a 800 Hz (BENESTY et al., 2008).

O tamanho da laringe e conseqüentemente das cordas vocais, é responsável pelas diferenças de tom (do inglês *pitch*) de voz entre os indivíduos, justificando assim as diferenças de  $F_0$  de cada voz masculina, feminina e infantil (MARTINS, 1988). Este tom “natural” ou de voz falada deve-se às características intrínsecas da vibração natural das cordas vocais. A extensão vocal permite igualmente contribuir para a definição da classificação da voz cantada em baixo, barítono, tenor e contratenor para a voz masculina e contralto, *mezzo-soprano* e soprano para a voz feminina.

As técnicas de medição funcional da voz têm que permitir visualizar os órgãos internos e/ou monitorizar uma função, dado que estes órgãos não podem ser observados a olho nu. De entre as várias técnicas, descritas na Tabela 2.1, a laringoscopia é a técnica de observação direta mais usada, principalmente do ponto de vista clínico, para detetar problemas da voz, massas, estenoses, obstruções, e ainda avaliar problemas de deglutição.

**Tabela 2.1 Breve descrição das técnicas de medição funcional da produção da voz.**

Técnica	Objeto de análise	Descrição
Pletismografia	Monitorização da função respiratória	Avaliação do volume pulmonar.
Pneumotacografia		Medição do fluxo de ar através da colocação de uma máscara na boca.
Transdutor de pressão		Medição da pressão subglotal por punção através de uma agulha.
Laringoscopia	Avaliação visual da morfologia e funcional da laringe	Permite observar, através de uma imagem de vídeo, a vibração das cordas vocais durante a fala.
Glotografia	Avaliação funcional da laringe	Monitorização da vibração das cordas vocais, usando um sensor ótico (fotoglotografia) ou através de elétrodos (eletroglotografia – EGG).

A laringoscopia consiste na utilização de um instrumento de fibra ótica rígido ou flexível – o laringoscópio e permite avaliar visualmente a laringe, recorrendo a três técnicas: direta, indiretamente ou através da endoscopia (CARDING, 2000; MORRISON & RAMMAGE, 1994).

A laringoscopia indireta consiste na observação da laringe através da reflexão da luz transmitida pelo instrumento para um espelho colocado na nasofaringe. A laringoscopia direta, geralmente é utilizada quando não é possível realizar as técnicas indireta ou endoscópica, requerendo sedação e portanto impossibilidade qualquer avaliação funcional da laringe.

Na endoscopia rígida, a fonte de luz está embutida no próprio endoscópio e assim como na laringoscopia indireta, requer intubação oral não permitindo avaliar a dinâmica ou ação da laringe a não ser durante a sustentação de sons vocálicos. No entanto, quer a endoscopia rígida quer a laringoscopia indireta permitem uma melhor observação da mucosa e estrutura das cordas vocais, graças à melhor iluminação (nitidez ótica) e poder de magnificação.

A nasoendoscopia com fibra ótica flexível é, atualmente, a técnica preferencial em otorrinolaringologia, uma vez que permite uma avaliação da produção da voz, através da via nasal e sem impedimentos na dinâmica articulatória da língua. Para além disso, a nasoendoscopia flexível permite a observação das cavidades nasais e nasofaringe.

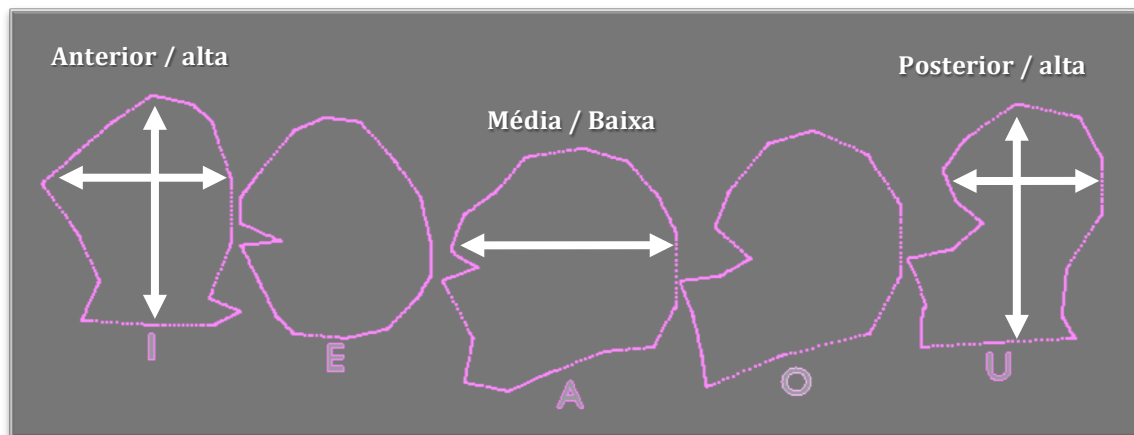
A utilização da endoscopia, rígida ou flexível, pode ser conjugada com a luz estroboscópica (pulsada e ajustada à frequência fundamental das cordas vocais), permitindo uma análise cuidada do estado da mucosa laríngea, pelo padrão de vibração durante a fonação.

## **2.2 Articulação da fala**

Funcionalmente, o mecanismo da articulação da fala representa a mais complexa atividade motora humana, utilizando apenas o movimento dos articuladores para a função de modulação da voz originando a fala (BENESTY et al., 2008). Estes órgãos, na sua maioria partes moles, como a língua, os lábios, o véu do palato e a faringe, que são extremamente ativos, mas também os órgãos

ósseos rígidos, como o palato e a mandíbula, alteram a ressonância do trato vocal de inúmeras formas e também criam fontes sonoras para algumas consoantes.

A língua é o articulador mais importante na produção da fala, pois altera grande parte da forma do trato vocal. A sua deformação determina a qualidade vocálica e permite a produção das consoantes palatais, velares e faríngeas. O movimento do ápex da língua permite a diferenciação em consoantes dentais e alveolares, e também a concretização das consoantes retroflexas, Figura 2.2.



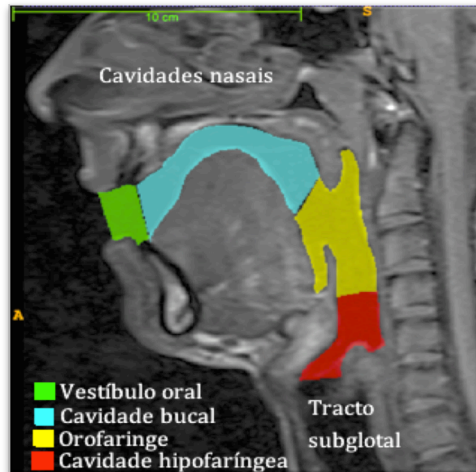
**Figura 2.2** Contornos sagitais da deformação da língua e respetiva posição na cavidade bucal extraídos a partir das imagens de RM, para a produção das vogais abertas do Português (VENTURA et al., 2008).

Os lábios deformam o término do trato vocal, contribuindo para a produção das vogais e das consoantes labiais (BENESTY et al., 2008). Quer a língua quer os lábios estão fixos à mandíbula. O véu do palato controla a abertura e fecho da passagem velofaríngea permitindo a diferenciação entre sons nasais e orais. Os músculos constritores da faringe permitem ajustar a largura lateral da cavidade faríngea, contribuindo também para a articulação das vogais e consoantes mais posteriores.

O palato e os dentes, que constituem o osso maxilar formam uma parede fixa do trato vocal e ainda que não sejam “verdadeiros” articuladores, desempenham um papel crucial pois servem de ponto de referência para muitas posições articulatórias, definindo o ponto de aproximação ou contacto com a língua.

### 2.2.1 Trato vocal e cavidade nasal

O trato vocal é um espaço acústico que serve à propagação de fontes sonoras, e inclui todas as cavidades de ar nas quais a variação de pressão acústica ocorre na produção de determinado som, Figura 2.3.



**Figura 2.3** Regiões do trato vocal propriamente dito e suas relações anatômicas numa imagem sagital de RM.

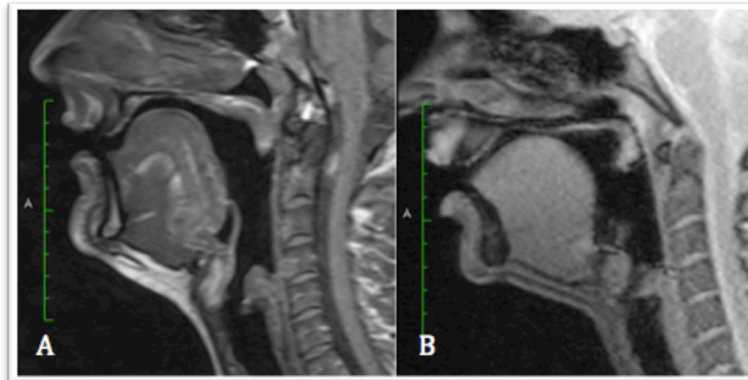
Nesse sentido, e como se pode observar pela figura acima, o trato vocal pode ser dividido em três regiões (BENESTY et al., 2008):

- Trato subglotal – que inclui todos os órgãos do aparelho respiratório localizados abaixo da glote;
- Trato vocal propriamente dito – desde a glote aos lábios cujo comprimento varia entre 15.0 cm nas mulheres e 17.5 cm nos homens. Anatomicamente, divide-se em quatro porções: na cavidade hipofaríngea (com 4 cm divididos pela cavidade laríngea supraglótica e as cavidades bilaterais, cónicas, da fossa piriforme), na orofaringe (que se estende desde as pregas ariepiglóticas ao arco anterior do palato), na cavidade bucal (do arco palatino aos dentes incisivos) e o vestíbulo oral (desde os dentes incisivos à abertura dos lábios);
- Cavidades nasais – formam um canal acessório que funciona como ressonância atribuindo características fonéticas aos sons nasais.

Tridimensionalmente, o trato vocal apresenta uma forma complexa delimitada por paredes rígidas e moles, cuja anatomia é variável, de indivíduo para indivíduo, pois depende dos seguintes fatores:

- Altura da maxila e mandíbula relativamente à altura da faringe;
- Mobilidade da mandíbula;
- Tamanho da língua.

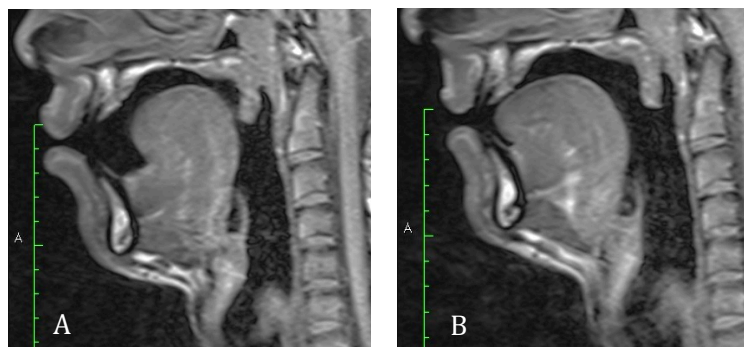
A Figura 2.4 ilustra as diferenças anatómicas entre dois sujeitos (um masculino e um feminino) e a conseqüente geometria/ configuração do trato vocal visualizadas nas imagens de RM.



**Figura 2.4** Diferenças anatómicas do trato vocal num sujeito do sexo masculino (A) e feminino (B), visualizadas numa imagem sagital de RM.

A produção das vogais resulta da ação conjunta da fonação, que gera o som, e da articulação, onde os sons são modulados atuando o trato vocal como um filtro acústico. Para os sons vocálicos, o trato vocal é um tubo fechado ao nível da glote e aberto nos lábios, Figura 2.5.

Na produção das consoantes sonoras, a vibração das cordas vocais tipicamente é contínua durante os sons, através da obstrução ou constrição do trato vocal. Esta vibração é suprimida nas consoantes surdas, devido à rápida diminuição da diferença da pressão transglotal e abdução das cordas vocais.



**Figura 2.5** Configuração do trato vocal durante a produção de uma vogal (A) e de uma consoante (B) com base em imagens sagitais de RM. Em (A) o trato vocal não apresenta qualquer obstrução/ constrição ao contrário da imagem (B).

Algumas características fonéticas dos sons da fala são o resultado da mobilidade dos órgãos articulatórios, cujo movimento é variável de órgão para órgão (BENESTY, 2008). A velocidade dos movimentos articulatórios é essencial numa sequência de sons com características distintivas. As diferenças observadas nestes órgãos com diferentes características motoras podem ser explicadas pelos seguintes fatores:

- Origem filogenética: os músculos da língua têm origem conjunta com outros sistemas motores rápidos, como o globo ocular e dedos; ao contrário dos músculos dos lábios e véu do palato que têm origem a partir de sistemas motores lentos, como a musculatura do trato digestivo.
- Inervação de cada músculo: os órgãos com movimentos rápidos são inervados por feixes nervosos finos e vice-versa, em resultado da adaptação do sistema biológico à função do órgão.
- Tipo de fibras musculares: os músculos da laringe têm uma alta concentração de fibras ultrarrápidas, ao contrário dos músculos do véu do palato. O ápex da língua executa movimentos mais rápidos que os lábios e por sua vez, do que o dorso da língua.

### 2.2.2 Técnicas de análise da dinâmica articulatória

A análise da dinâmica articulatória pode ser efetuada com recurso a várias técnicas, diretas ou indiretas, mas são sem dúvida as técnicas imagiológicas que têm tido maior utilização e desenvolvimento nas últimas décadas, face aos crescentes avanços tecnológicos, Tabela 2.2.

**Tabela 2.2 Descrição sumária das principais técnicas de análise da articulação da fala.**

Técnica	Descrição
Eletropalatografia (EPG)	Técnica que utiliza um dispositivo compacto para registar as alterações espaciais e temporais (em tempo real) dos contactos da língua com o palato, através da colocação de um palato artificial recoberto com vários eléctrodos de superfície, distribuídos por linhas e colunas.
Sistemas de seguimento com marcas ( <i>Marker tracking System</i> )	<p>Incluem vários sistemas (como o <i>X-ray Microbeam</i> e sistemas de sensorização magnética) que permitem o registo dos movimentos dos órgãos articulatórios através da colocação e seguimento de marcas em pontos específicos.</p> <p>No sistema <i>X-ray Microbeam</i>, os movimentos articulatórios são observados pela distribuição espacial de pequenas esferas de metal sobre os articuladores, através da emissão de um feixe de radiação X.</p> <p>Os sistemas de sensorização magnética (por exemplo, magnetómetro, articulografia eletromagnética – EMA) desempenham a mesma função mas sem utilizar radiação X.</p>
Técnicas imagiológicas	<p>A vídeo-fluoroscopia é uma técnica de imagem bidimensional que através da utilização de radiação X permite observar, em tempo real, o movimento dos articuladores. Os contornos das estruturas ósseas rígidas são claramente demonstrados; no entanto, no que respeita às partes moles a sua visualização é de qualidade inferior. O contorno da língua pode ser realçado através da aplicação de um produto de contraste na sua superfície. Assim como, podem também ser usadas pequenas marcas metálicas para detetar outras partes moles. Esta técnica tem vindo a entrar em desuso pela utilização de radiação X.</p> <p>A ultrassonografia (ou ecografia) é amplamente usada em investigação, principalmente no estudo da língua. Esta técnica permite obter imagens 2D em tempo real e também 3D dos órgãos de partes moles.</p> <p>A ressonância magnética (descrita no capítulo 3) é hoje uma das técnicas de imagem mais usadas no estudo da articulação.</p>

### 2.3 Estimação dos formantes e classificação fonética

Os formantes correspondem às zonas de ressonância, ou seja, às zonas amplificadas do espectro acústico da fala e são definidos por uma frequência central, uma largura de banda e uma energia. Normalmente, são indicados por F1, F2, ..., começando pela frequência mais baixa.

Estes parâmetros da fala são úteis na codificação, reconhecimento e síntese da fala, pois descrevem de forma eficiente e simples os aspetos essenciais da fala. O trato vocal pode ser modelado como um tubo acústico cujas ressonâncias básicas podem ser obtidas através do cálculo das funções de área (descritas adiante na secção 6.2.2).

A teoria fonte-filtro descrita por FANT explica este efeito acústico através de um modelo matemático, onde o trato vocal dá forma ao espectro de som gerado na laringe (FANT, 1960).

Os segmentos fonéticos são diferenciados, tanto pela presença ou ausência de vozeamento como por classes, incluindo vogais, glides e consoantes, de acordo com o seu modo e ponto de articulação. As principais características distintivas destas classes fonéticas estão descritas na Tabela 2.3.

**Tabela 2.3 Características distintivas das classes fonéticas do Português Europeu (adaptado de RIBEIRO, 2003).**

Classes	Descrição
Vogais	Existem 9 vogais orais e 5 vogais nasais, com duração maior que as glides e as consoantes. Caracterizam-se pela presença de vozeamento e forma fixa do trato vocal.
Glides ou semivogais	Ocorrem simultaneamente com uma vogal que lhe precede ou sucede, formando ditongos. Podem ser vistas como vogais mas com maior constrição e menor duração que as vogais respetivas.
Consoantes oclusivas	Constrição total do trato vocal seguida da libertação de pressão acumulada. Distinguem-se pelo ponto de oclusão e pela presença ou ausência de vozeamento.
Consoantes fricativas	Constrição parcial do trato vocal seguida de turbulência. Grande energia contida nas altas-frequências. Distinguem-se pelo ponto de constrição e pela presença ou ausência de vozeamento.
Consoantes nasais	Vozeamento e com o trato vocal totalmente fechado num determinado ponto ao longo da cavidade bucal (sendo por isso também conhecidas por oclusivas nasais). Abaixamento do véu do palato. A cavidade bucal embora fechada mantém-se acoplada à faringe e cavidade nasal, resultando numa antirressonância, ou seja, uma frequência inversamente proporcional à dimensão da constrição da cavidade bucal.
Consoantes líquidas	Estrutura marcada de formantes embora menos energéticas que as vogais. Dividem-se em laterais e vibrantes. Têm uma grande variabilidade, podendo ser ou não vozeadas e fricativando nalguns casos.



O *International Phonetic Alphabet* (IPA) é o mais conhecido alfabeto fonético para representar cada um dos segmentos fonéticos. No entanto, porque os caracteres que são utilizados neste alfabeto não são imprimíveis, o *Speech Assessment Methods Phonetic Alphabet* (SAMPA) é frequentemente utilizado para transcrever a versão para o Português Europeu (PE). Para representar o Português Europeu são necessários os subconjuntos dos alfabetos IPA e SAMPA, como indicados na Tabela 2.4.

**Tabela 2.4** Descrição das classes fonéticas do Português Europeu de acordo com os alfabetos IPA e SAMPA (adaptado de RIBEIRO, 2003).

Classe	Símbolo IPA	Símbolo SAMPA	Altura da língua	Posição da língua	Palavra
Vogais	i	i	Alta	Anterior	f <u>i</u> ta
	e	e	Média	Anterior	pe <u>r</u> a
	ɛ	E	Baixa	Anterior	se <u>t</u> a
	a	a	Baixa	Média	ca <u>r</u> a
	ɐ	6	Média	Média	ca <u>m</u> a
	ɨ	@	Alta	Média	qu <u>e</u> , se
	ɔ	O	Baixa	Posterior	co <u>r</u> da
	o	o	Média	Posterior	mo <u>f</u> o
u	u	Alta	Posterior	mu <u>d</u> o	
Vogais nasais	ĩ	i~	Alta	Anterior	ci <u>n</u> to
	ẽ	e~	Média	Anterior	se <u>n</u> to
	ã	6~	Média	Média	sa <u>n</u> to
	õ	o~	Média	Posterior	so <u>n</u> do
	ũ	u~	Alta	Posterior	mu <u>n</u> do
Glides	j	j	Alta	Anterior	pa <u>i</u>
	w	w	Alta	Posterior	pa <u>u</u>
Classe	Símbolo IPA	Símbolo SAMPA	Presença de vozeamento	Ponto de articulação	Palavra
Consoantes oclusivas	p	p	Não	Bilabial	pa <u>p</u> o
	b	b	Sim	Bilabial	ba <u>r</u> ba
	t	t	Não	Apicodental	ti <u>a</u>
	d	d	Sim	Apicodental	do <u>c</u> e
	k	k	Não	Velar	ca <u>s</u> a
	g	g	Sim	Velar	ga <u>t</u> o
Consoantes fricativas	f	f	Não	Labiodental	ba <u>f</u> o
	v	v	Sim	Labiodental	ca <u>v</u> a
	s	s	Não	Apicodental	pa <u>s</u> a
	z	z	Sim	Apicodental	ca <u>s</u> a
	ʃ	S	Não	Palatal	cha <u>p</u> éu
	ʒ	Z	Sim	Palatal	jo <u>i</u> a
Consoantes nasais	m	m	Sim	Bilabial	ma <u>r</u>
	n	n	Sim	Apicodental	na <u>d</u> a
	ɲ	J	Sim	Palatal	vi <u>n</u> ho

**Tabela 2.4 Descrição das classes fonéticas do Português Europeu de acordo com os alfabetos IPA e SAMPA (adaptado de RIBEIRO, 2003) – continuação.**

Classe	Símbolo IPA	Símbolo SAMPA	Altura da língua	Posição da língua	Palavra
Consoantes líquidas	l	l	Sim	Apicodental	sa <u>l</u> a
	ʎ	L	Sim	Palatal	fo <u>l</u> ha
	r	r	---	Apicodental	ca <u>r</u> o
	R	R	---	Velar	ca <u>r</u> ro

## 2.4 Conclusão

A respiração, a fonação e a articulação constituem os três principais mecanismos responsáveis pela produção da fala. Grande parte dos sons são produzidos durante a expiração do ar, que depois é transformado em som na laringe e depende da vibração das cordas vocais. A laringe, constitui o principal órgão responsável pela produção da voz e que devido à sua localização no pescoço “obriga” ao recurso a técnicas invasivas para a sua observação direta, através da laringoscopia.

O mecanismo da articulação da fala, isto é, responsável pela definição (modulação) dos sons da fala, resulta da atividade organizada de um conjunto de órgãos, distribuídos entre a cabeça e o pescoço, denominados de articuladores. A anatomia destes órgãos é complexa quer pelas suas características morfológicas quer pelo papel que desempenham na produção dos diferentes sons.

A língua é, sem dúvida, o órgão mais volumoso e dotado de uma enorme flexibilidade, desempenhando movimentos variados e permitindo a alteração de grande parte do trato vocal. Comparativamente, o véu do palato de pequenas dimensões e reduzida mobilidade, possibilita a diferenciação entre os sons orais e nasais, ou seja, contribui para uma das mais importantes características distintivas dos sons, principalmente no caso do Português Europeu.

O Português Europeu designa a variedade da Língua Portuguesa falada em Portugal, nas regiões autónomas e pelos emigrantes portugueses, englobando os seus dialetos regionais, vocabulário, gramática e ortografia. De acordo com o IPA, o Português Europeu é caracterizado por 35 sons e tem uma das fonologias mais

ricas das línguas românicas, incluindo vogais orais e nasais, ditongos nasais e dois ditongos nasais duplos.

Os desafios exigidos às técnicas de observação e análise da dinâmica articulatória constituem um aspeto de estudo em múltiplas áreas do conhecimento, nomeadamente para a Imagiologia, que tem hoje do ponto de vista clínico, um papel preponderante no auxílio ao diagnóstico e seguimento de inúmeras patologias. As inúmeras potencialidades da ressonância magnética têm permitido a sua utilização com sucesso na caracterização da fala humana, justificando assim a sua aplicação e maior divulgação.



### **CAPÍTULO 3. ESTUDO DAS VIAS AÉREAS SUPERIORES POR RESSONÂNCIA MAGNÉTICA**

A formação de imagens por RM difere de outras modalidades de imagem, tomográficas, como a tomografia computadorizada e consiste na localização precisa do sinal de RM de forma não invasiva e nas diversas localizações do corpo humano. A instrumentação da RM é complexa mas as potenciais aplicações clínicas têm sido crescentes, sendo hoje uma modalidade imagiológica com reconhecida superioridade em inúmeras condições patológicas, principalmente na avaliação de tecidos moles.

Para além disso, o aperfeiçoamento das técnicas de processamento de imagens aliado aos desenvolvimentos tecnológicos dos sistemas de RM têm proporcionado inúmeros avanços no conhecimento dinâmico de vários órgãos (por exemplo, o coração) ou funções (RM funcional do cérebro).

Assim, é imperativo responder às seguintes questões: Como funciona e qual a instrumentação usada em ressonância magnética? Quais as potenciais aplicações da RM no estudo das vias aéreas superiores?

Neste capítulo descrevem-se os princípios físicos básicos da ressonância magnética, bem como a instrumentação usada e principais técnicas de obtenção das imagens e sincronização. Por fim, são destacadas as potenciais aplicações clínicas da ressonância magnética neste âmbito.

### 3.1 Princípios físicos da ressonância magnética

Do ponto de vista histórico, em 1946 Block e Purcell desenvolveram os princípios da ressonância magnética, pelos quais, em 1952, receberam um Prémio Nobel da Física (PISCO & SOUSA, 1999). No período compreendido entre 1950 e 1970, a RM foi desenvolvida e utilizada para análises moleculares físicas e químicas. Só mais tarde, em 1973, surgem as primeiras publicações de Lauterbur relacionadas com a localização espacial e a técnica de reconstrução projetiva por RM (denominada então de zeumatografia).

Raymond Damadian, em 1971, demonstrou a existência de diferenças no tempo de relaxamento de diferentes tecidos e tumores, motivando assim a comunidade científica a considerar a RM como importante método de deteção de doenças.

O primeiro protótipo de um magnete com fins comerciais foi desenvolvido em 1975. No entanto, só a partir de 1980 é que surgem na literatura os primeiros estudos clínicos.

#### 3.1.1 O núcleo de Hidrogénio e o fenómeno de ressonância

Atualmente, a RM é uma modalidade imagiológica poderosíssima no auxílio ao diagnóstico médico, cujo princípio físico tem por base a interação de campos magnéticos, pulsos de radiofrequência (RF) e núcleos atômicos. Mas só são átomos “ativos” em RM os que possuem núcleos magnéticos, ou seja, que podem ser considerados como pequenos íman, sem orientação específica na ausência de um campo magnético (GIL & GERALDES, 1987).

Núcleos como o  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{19}\text{F}$ ,  $^{31}\text{P}$  são exemplos de núcleos “ativos” em RM, mas de todos, é sem dúvida o de hidrogénio ( $^1\text{H}$ ) o que possui um momento magnético mais importante e que está presente com enorme abundância no corpo humano, quer sob a forma de água, quer na forma de gordura. Para além destas características, o sinal do  $^1\text{H}$  é cerca de mil vezes superior ao de qualquer outro núcleo com momentos magnéticos semelhantes (PISCO & SOUSA, 1999).

Um paciente quando é colocado sob um potente campo magnético ( $B_0$ ) homogéneo permite que o sinal de RM dos núcleos da região em estudo seja

indistinto do sinal de qualquer outra parte do corpo. Assim, se o campo magnético estático variar ao longo da amostra, a frequência de Larmor (Equação 1) dos núcleos varia de igual forma (GIL & GERALDES, 1987).

$$\omega_0 = \gamma \times B_0 \quad (1)$$

$\gamma$  - Razão giromagnética (constante para determinados núcleos, por exemplo,  $\gamma \text{ } ^1\text{H} = 42.58 \text{ MHz}$ );

$\omega_0$  - Frequência de precessão ou de Larmor;

$B_0$  - Campo magnético principal;

Assim, cada instrumento de RM tem uma frequência de Larmor característica com base na potência do seu campo magnético (LUFKIN, 1999). Na presença de  $B_0$ , os núcleos de hidrogênio descrevem um movimento precessional e tendem para se orientar paralelamente ao campo (estado de equilíbrio) ou numa direção antiparalela ao campo (estado de excitação).

Por questões de equilíbrio, a orientação paralela será preferencialmente superior, dando-se a formação de um vetor que corresponde à soma de todas as orientações magnéticas paralelas ao campo magnético, denominado de vetor de magnetização efetivo (VME).

### 3.1.2 O sinal de ressonância magnética

Posteriormente, é aplicado um pulso de RF de  $90^\circ$ , que introduz energia ao sistema, deslocando o VME para o plano transversal – fenômeno de excitação. Para que isto aconteça é necessário que este segundo campo esteja sincronizado com a frequência de precessão dos núcleos. A amplitude e duração do pulso de RF podem ser controladas para produzir diversos graus de angulações (*flip angle* – FA) do VME, do plano longitudinal ao transversal.

Para além de introduzirem energia ao sistema, os pulsos de RF ( $90^\circ$  ou  $180^\circ$ ), fazem com que os prótons estejam em fase (coerência de fase), permitindo assim, a detecção do sinal através de antenas (o VME induz uma corrente na antena de acordo com a Lei de Faraday).

Após cessar o pulso de RF, a amplitude do sinal ao ser detectado não permanece constante declinando rapidamente até zero. Esta oscilação

amortecida é conhecida por *Free Induction Decay* (FID) e deve-se ao fenómeno de relaxação dos núcleos. A relaxação é um processo físico, dinâmico, no qual os núcleos retornam ao seu estado de equilíbrio.

Se considerarmos, por exemplo, uma amostra homogénea (água pura), sob um campo magnético uniforme, o espectro dos núcleos de  $^1\text{H}$  é um só pico (GIL & GERALDES, 1987). Aplicando-se um gradiente linear de campo magnético ao longo do eixo dos  $\text{XX}$ ,  $G_x$ , sobreposto ao  $B_0$ , as frequências de ressonância dos núcleos de  $^1\text{H}$  da água tornam-se uma função linear, ou seja, uma medida direta da coordenada  $X$  da amostra. Numa amostra contínua, o sinal de RM deixa de ser um espectro após transformada de Fourier, traduzindo-se numa projeção unidimensional da densidade dos núcleos de  $^1\text{H}$  localizados num plano perpendicular ao gradiente em função da posição  $X$  desse plano.

A intensidade do sinal para cada valor de  $X$  indica qual a medida da concentração de núcleos de  $^1\text{H}$  e a sua frequência é a medida da posição do plano projetado. Se o gradiente for linear, o sinal recebido com uma dada frequência provém dos centros situados num dado plano perpendicular à direção do gradiente.

Se forem excitados sequencialmente:

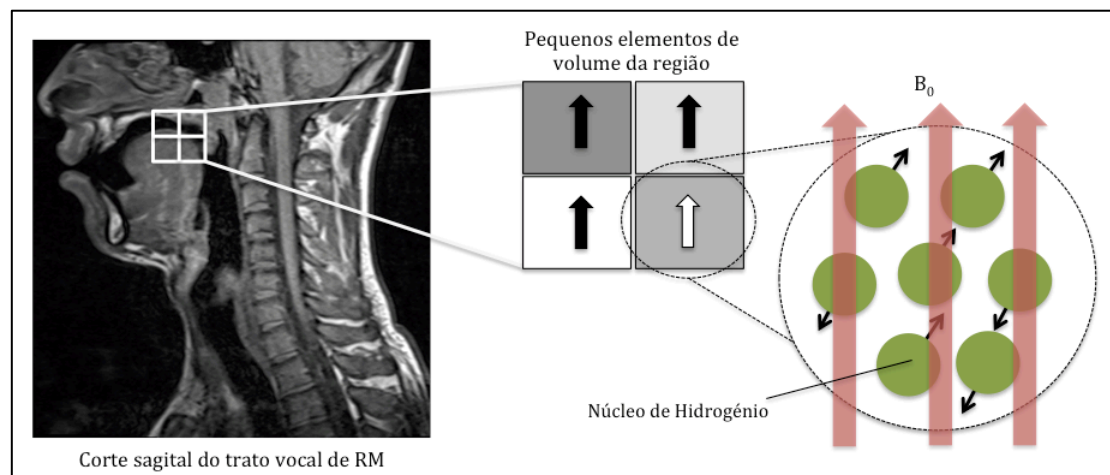
- Dois planos em intercepção por aplicação simultânea dos gradientes perpendiculares – cada frequência define uma linha; ou então,
- Três planos em intercepção – a frequência define um ponto.

Variando o campo magnético a diferentes níveis, utilizando gradientes variados é possível construir uma imagem tomográfica. Assim, a intensidade do sinal de RM na imagem é proporcional à densidade local de núcleos de  $^1\text{H}$  fortemente modulada por fatores físicos locais, tais como movimentos microscópicos e macroscópicos (por exemplo, fluxos). O facto da imagem por RM depender de um vasto leque de propriedades tecidulares, tornam esta técnica extremamente flexível, mas por outro lado, o domínio do seu conhecimento é mais difícil (BUXTON, 2009).

Na prática, as velocidades de magnetização (tempo de relaxação  $T_1$ ) e desmagnetização (tempo de relaxação  $T_2$ ) dos tecidos são representadas como diferentes escalas de cinzento nas imagens por RM (PISCO & SOUSA, 1999), onde



o branco representa uma intensidade de sinal elevada e o preto o inverso, Figura 3.1.



**Figura 3.1** Representação das diferentes velocidades de magnetização e desmagnetização de uma pequena área do trato vocal.

O processo de formação de imagem por RM engloba assim um conjunto de etapas, sequenciais, que se encontram resumidas na tabela seguinte.

**Tabela 3.1** Resumo do processo de formação de imagem por RM.

Etapas	Resultados
1) Aplicação de $B_0$	Os núcleos alinham-se e entram em precessão.
2) Seleção do corte pelo gradiente	Os núcleos realizam movimento de precessão em frequência específica.
3) Aplicação de pulsos de RF	Núcleos na área de corte entram em precessão e em fase num ângulo de inclinação maior.
4) Recepção do sinal de RF pela antena	Sinal elétrico recebido dos núcleos e enviado ao sistema.
5) Conversão do sinal em imagem	Reconstrução através da transformada de Fourier e exibição da imagem.

### 3.2 Instrumentação e segurança em ressonância magnética

Excluindo-se o computador para programação e armazenamento de dados e o monitor para visualização das imagens que são componentes comuns a qualquer equipamento de imagem médica mais avançada, fazem parte da instrumentação em RM os seguintes subsistemas (HORNAK, 1996; LUFKIN, 1999):

- Campo magnético principal;

- Bobines de gradientes;
- Bobines transmissoras e recetoras de RF.

O equipamento de RM é constituído por um magnete e uma mesa de exames. A maioria dos magnetes são do tipo supercondutor, ou seja, são electromagnetes compostos por fio condutor. A sua principal função é produzir um campo magnético uniforme, para evitar distorções geométricas na imagem e perda de resolução espacial e um campo homogéneo, cujo estado é garantido por criogénios (hélio líquido) que permitem o arrefecimento das bobines, praticamente ao zero absoluto, reduzindo assim a resistência.

Os magnetes supercondutores apresentam um elevado custo e uma tecnologia complexa, mas no entanto, permitem a obtenção de imagens com elevada intensidade, ou seja, potências de campo maiores.

Para fins clínicos, a potência de campo magnético dos equipamentos de RM disponíveis varia entre 0.2 a 3.0 Tesla (T); já para fins de investigação são utilizadas potências de campo magnético acima de 7.0 T.

Os gradientes têm como função produzir variações lineares no campo magnético principal, para localização espacial do sinal de RM, definindo assim os limites de desempenho do equipamento na obtenção de imagens. Apresentam como características principais:

- Amplitude (mT/m) – determina a resolução espacial máxima (espessura de corte e *Field of View* – FOV);
- *Slew rate* ou índice de subida do gradiente (T/m/s) – determina o tempo de eco (TE) e o tempo de repetição (TR) mínimos que podem ser obtidos;
- Linearidade – importante para o mapeamento preciso do sinal, pois a não linearidade dos gradientes origina distorções geométricas na imagem e perda de resolução espacial.

A rápida comutação (“*switch*”) dos gradientes induz correntes (conhecidas por “*Eddy currents*”) nos materiais condutores adjacentes, opondo-se aos campos de gradiente e causando a diminuição do seu desempenho e artefactos nas imagens. Esta corrente induzida pode ser minimizada através da blindagem das bobines de gradiente.

Esta comutação dos gradientes provoca também forças de Lorentz causando vibrações nos mesmos e nos componentes adjacentes, constituindo a principal fonte de ruído característico da RM.

No que diz respeito ao subsistema de RF, tem como funções a produção de pulsos de RF (excitação dos núcleos) e a detecção do sinal de RM. O magnete possui bobines transmissoras e recetoras de RF que são normalmente utilizadas em estudos de corpo (áreas de anatomia extensas). Dado que a bobine recetora tem um efeito marcante na relação sinal-ruído (do inglês, *signal-to-noise ratio* – SNR) e que o paciente é uma das principais fonte de ruído nos instrumentos de RM, é necessário a otimização da interface entre o equipamento e o paciente.

Assim, para determinadas regiões anatómicas existem antenas ou bobines dedicadas (de vários tipos) e que permitem obter uma melhor SNR, podendo ser transmissoras e/ou recetoras de RF. Estas antenas podem ser divididas em três tipos mediante o seu FOV e relação sinal-ruído:

- Antenas de superfície (têm um FOV reduzido, mas possuem elevada relação sinal-ruído);
- Antenas de volume (têm um FOV grande e uma menor relação sinal-ruído);
- Antenas *Phased Array* (série de antenas montadas num conjunto para a geração de imagens em volume).

As antenas *Phased Array* associam o conceito de antenas de superfície (elevada relação sinal-ruído) e elevado FOV das antenas de volume. Como se encontram montadas por um conjunto de sub-antenas ligadas independentemente ao sistema de aquisição, podem ser ativadas por segmentos.

De forma a garantir a proteção de eventuais interferências por RF exteriores, a sala de RM encontra-se isolada por uma gaiola de Faraday.

Segundo LUFKIN (1999) vários estudos têm sido desenvolvidos na tentativa de identificar os aspetos biológicos da RM com efeitos adversos, no entanto, nenhum determinou riscos significativos ou inesperados. Os efeitos biológicos estão relacionados com a exposição aos campos magnéticos estáticos, aos gradientes de campo e às ondas de RF, o que obriga ao cumprimento de algumas normas de segurança (LUFKIN, 1999).

Quanto aos efeitos decorrentes da exposição aos campos magnéticos estáticos, não existem estudos relacionados com a utilização de campos magnéticos de alta intensidade (acima de 1.0 T). Na maioria dos estudos realizada em animais os resultados não são unânimes, havendo alguns relatos de efeitos de temperatura, de indução elétrica e efeitos cardíacos e neurológicos (LUFKIN, 1999).

A exposição a variações rápidas dos campos magnéticos podem induzir campos elétricos e correntes elétricas em meios condutores, de acordo com a Lei de Faraday. No entanto, estes efeitos térmicos devido à ativação dos gradientes não são considerados significativos clinicamente.

No que diz respeito à exposição às ondas de RF, tem como principal efeito o depósito de energia térmica nos tecidos, sendo que o aumento efetivo da temperatura tecidual depende do sistema termorregulador do indivíduo; por exemplo, fluxo sanguíneo cutâneo, frequência de sudorese, entre outras. A taxa de absorção de RF é determinada pelo rácio em que a energia de RF se deposita nos tecidos, denominando-se de *Specific Absorption Ratio* (SAR) e é medida em watts por quilograma (W/Kg).

Para além disso, constitui um risco potencial para os pacientes, o elevado ruído acústico produzido durante a RM. O ruído está associado com a ativação e desativação dos gradientes (como abordado anteriormente), e tende a aumentar com a seleção de cortes mais finos, de um FOV menor e com a redução do TR e do TE como acontece nas sequências ultrarrápidas. Atualmente, face às preocupações crescentes da comunidade científica quanto aos níveis de ruído produzidos e pelo desconforto causado nos pacientes, os equipamentos de RM vêm dotados com protetores auditivos (auscultadores ou tampões auditivos).

### 3.2.1 Linhas de orientação da *Food and Drug Administration*

Em 2003, a *Food and Drug Administration* (FDA) dos Estados Unidos da América (EUA), após revisão das recomendações de segurança para a RM anteriormente definidas em 1982, estabeleceu os limites máximos indicados na Tabela 3.2 (HORNAK, 1996).

**Tabela 3.2 Limites de segurança em ressonância magnética (Food and Drug Administration, 2003).**

<b>Linhas de orientação da FDA (2003)</b>	<b>Limite máximo</b>
<b>Exposição a campos magnéticos:</b>	
Adultos, jovens e crianças (>1 mês)	8 T
Recém-nascidos e crianças (<1 mês)	4 T
<b>Variação dos gradientes:</b>	Sem desconforto, dor, ou estimulação nervosa
<b>Depósito de energia térmica nos tecidos:</b>	
Corpo inteiro, em média, tempo $\geq 15$ min.	4 W/Kg
Cabeça, em média, tempo $\geq 10$ min.	3 W/Kg
Cabeça ou tórax, por grama de tecido, tempo $\geq 5$ min.	8 W/Kg
Extremidades, por grama de tecido, tempo $\geq 5$ min.	12 W/Kg
<b>Nível de ruído acústico:</b>	140 dB

Dada a utilização de um campo magnético potente, antes da realização de qualquer exame, os pacientes devem ser submetidos a um rigoroso questionário de forma a averiguar eventuais contraindicações, devido à possibilidade de interferência elétrica com determinados dispositivos (por exemplo, implantes e desfibriladores cardíacos, implantes cocleares, neuroestimuladores), risco potencial de projéteis (atração de objetos ferromagnéticos) e torção de objetos.

O risco de interferência elétrica inclui o movimento, fecho ou dano, alterações da programação, interferências eletromagnéticas e produção de correntes induzidas.

Os aparelhos e materiais dentários, clips aneurismáticos, *stents*, filtros e molas intravasculares, válvulas cardíacas, próteses e material ortopédico, projéteis, balas, corpos estranhos, são exemplos de objetos ferromagnéticos usualmente utilizados ou encontrados na prática clínica. No entanto, atualmente os materiais ferromagnéticos têm vindo a ser substituídos por outros metais, como o alumínio ou a platina. Para além destas medidas de segurança, uma unidade de RM deve estar devidamente sinalizada (KANAL et al., 2007).

### 3.2.2 Controlo de qualidade em ressonância magnética

Os programas de controlo de qualidade são uma prática rotineira em imagem médica e têm como finalidade avaliar o desempenho de determinado sistema e testar a sua utilidade clínica. A RM, pelos desenvolvimentos crescentes e maior uso clínico, não é exceção, tendo sido propostos protocolos de controlo de qualidade, por várias entidades, com orientações e recomendações para os testes a realizar antes e após a instalação de um sistema de RM (NALLO et al., 2006). Estes testes devem ser realizados periodicamente pelo fabricante do sistema, e permitem a identificação da degradação da qualidade de imagem (prevenindo erros no diagnóstico) e determinar a fonte de mau funcionamento do equipamento (prevenindo manutenções).

Os fantasmas são objetos antropomórficos<sup>1</sup> usados para testar o desempenho de qualquer sistema e também em RM, para medição do limite de absorção de RF dos tecidos (SAR). No entanto, por questões de segurança e dado os princípios físicos da RM, os materiais que compõem estes objetos (fantomas) devem ser:

- Compatíveis com a utilização de  $B_0$ ;
- Dar sinal em RM, ou seja, conter hidrogénio.

A maioria dos fantasmas disponíveis são por isso, constituídos por soluções aquosas ou gelatinosas, silicone ou outras soluções paramagnéticas. A água é, sem dúvida, a substância mais frequentemente usada nestes fantasmas.

Os fantasmas podem ser divididos em dois tipos: os fantasmas de resolução e os fantasmas de homogeneidade da RF, de acordo com a sua finalidade. Os fantasmas de resolução são usados para testar as propriedades espaciais de uma imagem (por exemplo, parâmetros geométricos como a resolução espacial, a espessura de corte, a SNR e o contraste nas diferentes ponderações de imagem) e consistem num invólucro de plástico com pequenos espaços preenchidos com água. Os fantasmas de homogeneidade da RF são usados para testar a uniformidade espacial do subsistema de transmissão e receção de RF.

---

<sup>1</sup> Ou seja, com características similares aos tecidos do corpo humano.

### 3.3 Técnicas de obtenção de imagens

Para obtenção das imagens de RM foram desenvolvidas várias técnicas, prevalecendo apenas algumas para uso clínico, com base em diferentes estratégias, Tabela 3.3.

**Tabela 3.3 Análise comparativa das diferentes técnicas de obtenção de imagens por RM e principais características (adaptado de LUFKIN, 1999).**

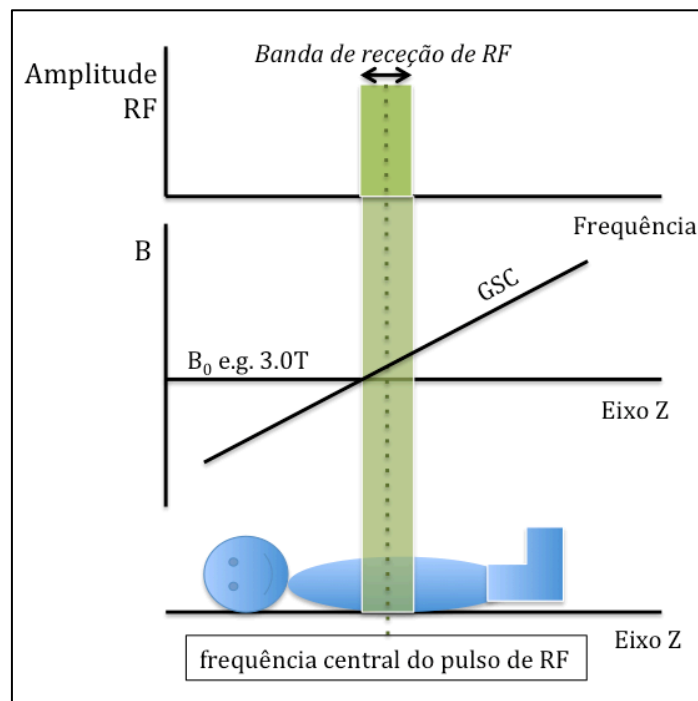
Características		Vantagens e desvantagens
<b>Pontos</b>	Excitação de um ponto através da aplicação de três gradientes alternados sobrepostos a $B_0$ .	<ul style="list-style-type: none"> <li>↑ Evita cálculos de reconstrução de imagem.</li> <li>↓ Reduzida velocidade de varrimento. Em desuso atualmente.</li> </ul>
<b>Linhas</b>	Irradiação seletiva ao longo de linhas através da aplicação de dois gradientes diferentes. O FID registado corresponde a todos os núcleos em ressonância situados numa linha de elementos definida ao longo do eixo e a sua transformada dá a densidade dos núcleos de $^1\text{H}$ segundo essa linha.	<ul style="list-style-type: none"> <li>↑ Rapidez dado que evita cálculos de reconstrução de imagem.</li> <li>↓ Problemas de baixa sensibilidade e eficiência. Em desuso atualmente.</li> </ul>
<b>Planos</b>	A imagem bidimensional da secção transversal obtém-se a partir de múltiplas projeções da secção em estudo, obtidas no mesmo plano. Neste método de projeção-reconstrução, a direção do gradiente aplicado após o pulso de $90^\circ$ é rodada no plano XY de incrementos sucessivos. O corte ou imagem planar é adquirido durante a excitação seletiva nesse local.	<ul style="list-style-type: none"> <li>↑ Técnica mais usada atualmente. Utiliza tempos de aquisição entre 2 a 10 minutos.</li> <li>↓ Necessidade de convolução ou filtragem antes da retroprojeção.</li> <li>↓ A relação sinal-ruído determinada pelos <i>voxel</i> de cada corte bidimensional não contribui para a relação sinal-ruído do corte seguinte.</li> </ul>
<b>Volume total</b>	O sinal é obtido a partir da ativação de todo o volume ou de uma "fatia" mais espessa, através da eliminação do gradiente de seleção do corte (descrito abaixo).	<ul style="list-style-type: none"> <li>↑ Obtenção de uma imagem tridimensional, podendo-se obter a partir desta qualquer secção bidimensional em qualquer plano.</li> <li>↑ Elevada relação sinal-ruído pois o sinal de cada <i>voxel</i> contribui para cada corte.</li> <li>↓ Exige uma elevada capacidade computacional e acarreta um longo tempo de aquisição dos dados.</li> </ul>

Os gradientes de campo magnético são recursos-chave que tornam possível a obtenção de imagens, permitindo a localização e codificação espacial do sinal de RM. Existem três gradientes físicos nos instrumentos de RM:  $G_x$ ,  $G_y$  e  $G_z$  que se

encontram orientados ortogonalmente no espaço. O isocentro do magnete corresponde ao ponto central do eixo dos três gradientes e do corpo do magnete, permanecendo a sua potência inalterada nesse ponto, mesmo após ativação dos gradientes (WESTBROOK & KAUT, 2000).

Um gradiente é uma variação linear do  $B_0$  ao longo de uma dimensão do paciente. O grau de variação é efetivamente muito pequeno, tipicamente cerca de duas a três ordens de grandeza inferior ao campo magnético.

Como abordado anteriormente, aquando da sobreposição de um gradiente de campo magnético ao  $B_0$ , a frequência dos núcleos varia com a localização espacial ao longo do gradiente aplicado. Este gradiente é denominado de gradiente de seleção do corte (GSC), e é ativado durante a aplicação do pulso de  $90^\circ$  para se estimular um plano de espessura finita. A localização do corte ao longo do gradiente é determinada pela frequência central do pulso de RF, Figura 3.2.



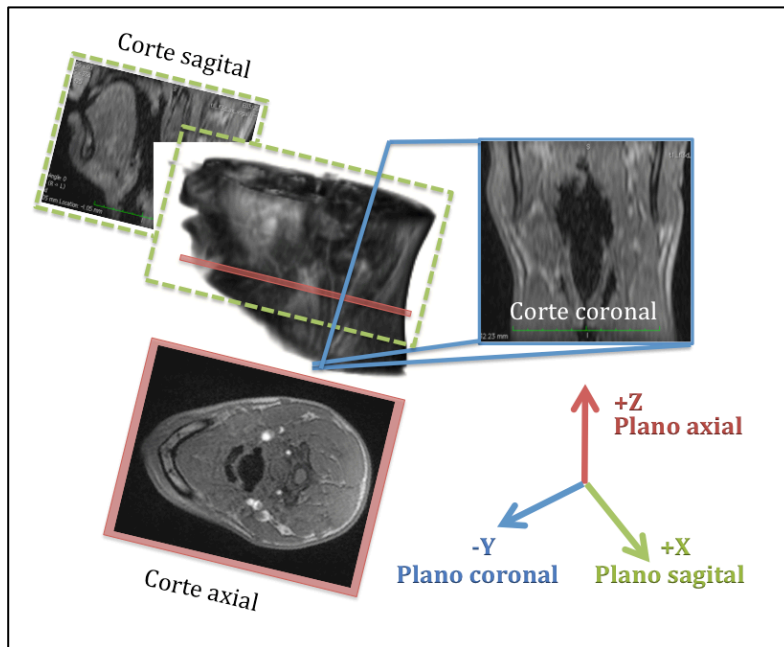
**Figura 3.2** Localização do corte ao longo do gradiente e da frequência central do pulso de RF.

A espessura de corte pode ser determinada (a) pela gama de frequências (largura da faixa ou banda de recepção) usada no pulso de  $90^\circ$  ou, (b) pela variação de amplitude do GSC. Assim, um corte fino pode ser obtido pela utilização de uma faixa estreita de RF ou pela ativação de gradientes mais



inclinados. A utilização de cortes finos acarreta, no entanto, uma menor relação sinal-ruído dado que apresentam menos núcleos de hidrogénio na amostra (LUFKIN, 1999; WESTBROOK & KAUT, 2000).

O gradiente de seleção do corte permite também determinar a orientação do plano de exame, Figura 3.3. Assim, para obter cortes sagitais, o gradiente de seleção de corte corresponderá à ativação do gradiente Gx, sendo os gradientes Gy e Gz selecionados para a codificação de espacial do sinal de RM (num passo posterior).



**Figura 3.3** Localização espacial do sinal de RM num estudo do trato vocal e sua correlação com os planos anatômicos.

Espacialmente, o sinal de RM é codificado no eixo da fase que corresponde ao menor eixo da anatomia em estudo e posteriormente no eixo da frequência que corresponde ao maior eixo da anatomia, Tabela 3.4.

**Tabela 3.4** Eixos dos gradientes para cada plano de corte.

Plano	GSC	CFA	CFR
Sagital	Gx	Gy	Gz
Coronal	Gy	Gx	Gz
Axial (cabeça)	Gz	Gx	Gy

GSC – Gradiente de Seleção de Corte; CFA – Codificação de Fase; CFR – Codificação de frequência.

A ativação do gradiente de codificação de fase permite, para além da alteração da frequência de precessão dos núcleos ao longo do mesmo, a alteração

da velocidade de precessão passando a haver uma diferença de fase ou desvio entre esses núcleos. O gradiente de codificação de fase é ativado imediatamente antes do pulso de  $180^\circ$  nas sequências spin-eco, ou entre a excitação e a leitura do sinal nas sequências gradiente-eco. Por sua vez, o gradiente de codificação de frequência produz uma diferença de frequência ou desvio de sinal e é ligado durante a leitura do sinal, sendo denominado de gradiente de leitura. O seu grau de inclinação determina a extensão da anatomia coberta ao longo do eixo de codificação de frequência, ou seja, o campo de visão. A duração do gradiente de codificação de frequência durante a leitura denomina-se de tempo de amostragem, tempo este no qual o sistema recolhe amostras de até 1024 frequências diferentes (na tecnologia atual). O tempo, a frequência de amostragem e a banda de receção regem-se pelo Teorema de Nyquist (descrito mais adiante na secção 3.4.3).

Os núcleos de  $^1\text{H}$  em cada *pixel* têm uma frequência e uma fase diferente, única, codificando espacialmente as coordenadas X e Y de cada *pixel*, respetivamente.

### 3.3.1 Formação de imagens e *espaço K*

Apesar da complexidade das ações dos gradientes necessárias à codificação do sinal que é detetado, não é suficiente para formar uma imagem. Este processo tem que ser repetido muitas vezes. Cada repetição é denominada de etapa de codificação de fase. Isto é possível mantendo constantes no tempo, os gradientes de seleção de cortes e de leitura, alterando-se a cada etapa apenas o gradiente de codificação de fase (LUFKIN, 1999).

A matriz de aquisição de dados bidimensionais – *espaço K*, tem forma rectangular com dois eixos perpendiculares entre si: o eixo de fase é horizontal e o eixo de frequência vertical, permitindo o armazenamento dos dados que posteriormente vão formar uma imagem. A inclinação do gradiente de codificação de fase vai determinar que linha do *espaço K* é preenchida.

O centro do *espaço K* contém dados com maior amplitude de sinal tanto ao longo do eixo da fase como da frequência; as linhas centrais apresentam menor resolução espacial pois são preenchidas por pequenas inclinações de gradiente.

Após preenchimento do todo o *espaço K*, os dados são submetidos à transformada de Fourier, sendo as informações de frequência e de fase mapeadas na imagem pela sua localização e as informações sobre a amplitude do sinal para cada *pixel* apresentadas como brilho. Dado que existe uma assimetria nos dois eixos da imagem, quer de codificação quer pela resolução espacial, as imagens de RM têm rotineiramente uma resolução espacial maior no eixo de codificação de frequências do que no eixo codificador de fase (por exemplo, matrizes retangulares 256x128 *pixels*).

### 3.3.2 Parâmetros de aquisição e qualidade da imagem

Os parâmetros de aquisição disponíveis para a escolha de uma sequência de pulsos são variados e definidos pelo operador (GIL & GERALDES, 1987; LUFKIN, 1999; WESTBROOK & KAUT, 2000). A combinação destes parâmetros determina o tipo de ponderação e a qualidade de imagem e estão descritos na Tabela 3.5.

A qualidade de imagem é um aspeto muito importante para o diagnóstico médico, sendo determinada pelos seguintes fatores: o contraste de imagem, a SNR, a resolução espacial e o tempo de aquisição.

**Tabela 3.5 Parâmetros de aquisição em RM determinantes do tipo de ponderação e da qualidade de imagem.**

Parâmetro de aquisição	Descrição
<b>Tempo de repetição (TR)</b>	Intervalo de tempo (medido em milissegundos) entre pulsos de excitação consecutivos.
<b>Tempo de eco (TE)</b>	Intervalo de tempo (medido em milissegundos) que decorre entre o pulso de 90° e a leitura do eco
<b>Flip Angle (FA)</b>	Ângulo de inclinação do VME após aplicação de um pulso de excitação.
<b>Tempo de inversão (TI)</b>	Intervalo de tempo (medido em milissegundos) entre a aplicação de um pulso de inversão (180°) e o pulso de 90°.
<b>Banda de receção (BW)</b>	Faixa de frequências que podem ser transmitidas.
<b>Espessura de corte (Tk)</b>	Profundidade de um volume em milímetros.
<b>Espaçamento entre cortes</b>	Do termo inglês - GAP, determina o espaço entre dois cortes consecutivos.

**Tabela 3.5 Parâmetros de aquisição em RM determinantes do tipo de ponderação e da qualidade de imagem – continuação.**

Parâmetro de aquisição	Descrição
<b>Echo train length (ETL)</b>	Conjunto de pulsos de refasamento (180°) aplicados durante uma sequência de pulsos, com o objetivo de aumentar a rapidez do exame.
<b>Field of View (FOV)</b>	Campo de visão da região anatômica e é definido pela matriz e pelo tamanho do <i>pixel</i> .
<b>Matriz</b>	Conjunto de <i>pixels</i> distribuídos em colunas e linhas; as suas dimensões são usadas para descrever a imagem. Determina a resolução espacial.
<b>Número de excitações (NEX)</b>	Número de vezes em que os dados são utilizados para a reconstrução da imagem, ou seja, a média de sinais que podem ser repetidamente observados e somados.

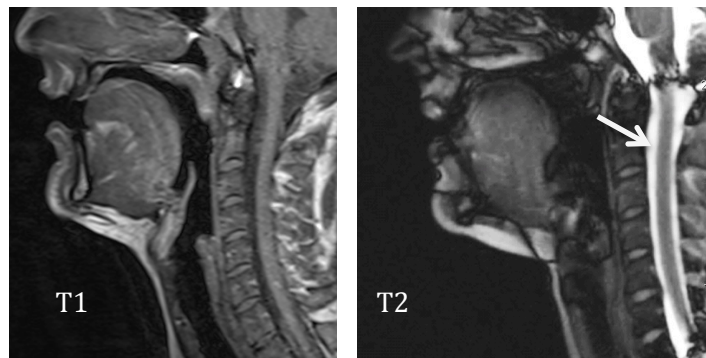
O contraste de imagem traduz a diferença de intensidade de sinal entre duas áreas adjacentes e é observado na forma de diferentes escalas de cinzentos proporcional ao sinal de RM e portanto, à amplitude de magnetização transversal. Clinicamente, é um dos fatores mais importantes de diagnóstico. O contraste de imagem depende dos parâmetros de aquisição temporais como o TR, TE, TI, do FA, do fluxo e do ETL.

O sinal FID é dado pelas diferenças de relaxação dos tecidos (que ocorre de forma natural), podendo ser de dois tipos:

- Relaxação longitudinal ou recuperação T1 – resulta da libertação de energia dos prótons excitados para o meio circundante e depende da estrutura molecular, do estado da matéria e da potência de campo magnético. Moléculas com recuperação rápida têm um T1 curto (por exemplo, gordura) e moléculas com recuperação lenta têm um T1 longo (por exemplo, líquido céfalo-raquidiano). O TR é responsável pela quantidade de sinal em T1.
- Relaxação transversal ou recuperação T2 – resulta da troca de energia entre os núcleos, que devido aos seus pequenos campos magnéticos individuais, perdem a sua coerência de fase, irreversivelmente. Moléculas que têm um decaimento rápido, têm um T2 curto (por exemplo, gordura),

e moléculas com decaimento lento têm um T2 longo (por exemplo, líquido céfalo-raquidiano). O TE controla a quantidade de decaimento T2.

Assim, valores de TR e TE curtos produzem imagens que enfatizam as características T1 dos tecidos e denominam-se de imagens “ponderadas em T1”, enquanto que TR e TE longos originam imagens “ponderadas em T2”, Figura 3.4. De realçar que, todas as imagens tanto contêm informações T1 como T2, mas dado que estes valores dos tecidos são conhecidos é possível definir-se a ponderação ideal.



**Figura 3.4** Imagens sagitais do trato vocal ponderadas em T1 e em T2. A gordura apresenta maior intensidade de sinal na imagem ponderada em T1. Na imagem ponderada em T2, observa-se maior intensidade de sinal no canal medular pela presença do líquido céfalo-raquidiano (indicado pela seta).

A relação sinal-ruído descreve a razão entre a amplitude do sinal recebido (voltagem induzida na bobine recetora) e a amplitude média do ruído (produzido pelo paciente dentro do magnete e pelo ruído elétrico de fundo do sistema). Este fator é um dos melhores critérios de qualidade de uma imagem. O ruído sobreposto ao sinal provoca uma oscilação na intensidade dos *pixels* à volta de um valor médio, diminuindo os contornos das separações entre os tecidos; logo, quanto maior for a SNR (mais sinal em detrimento do ruído) melhor será a qualidade de imagem.

A SNR é influenciada pela densidade dos prótons (na amostra), pelo volume do *voxel*, pelo NEX, pela BW e pelo tipo de antena utilizada. As soluções para aumentar a SNR não são isentas de penalizações e têm influência quer na resolução espacial como no tempo de aquisição, Tabela 3.6.

**Tabela 3.6 Fatores que aumentam a SNR e respetivas consequências.**

A SNR aumenta com:	Penalização
Aumento do volume do <i>voxel</i> (mais prótons de hidrogénio).	Diminuição da resolução espacial (isto é, a capacidade de se distinguir dois pontos como separados e distintos). A resolução espacial está relacionada com o tamanho do <i>voxel</i> .
A utilização de antenas de superfície (adaptam-se à região em estudo captando melhor o sinal).	FOV limitado.
Aumento do NEX.	Aumento do tempo de aquisição.
Bandas de receção mais estreitas.	Aumento dos artefactos.

O tempo de aquisição corresponde ao tempo necessário para se completar a aquisição de dados, sendo que quanto mais prolongado maior é a probabilidade do paciente se mover. Este fator é determinado pelo TR, pelo NEX e pelo número de codificações de fase.

### 3.3.3 Artefactos de imagem

Os artefactos correspondem a efeitos visuais presentes numa imagem e que não correspondem ao objeto real em estudo, podendo surgir sob a forma de linhas, pontos, padrões repetidos, imagens em espelho ou distorções. Em RM, assim como noutras técnicas de imagem, os artefactos são frequentes e variados de acordo com a sua origem. No caso do estudo do trato vocal, podem surgir mais frequentemente: (a) o artefacto de *aliasing* ou *wrap around*; (b) o artefacto de movimento do paciente; e (c) o artefacto de susceptibilidade magnética no caso de existirem aparelhos ou materiais ortodônticos de natureza ferromagnética ou metálicos na cavidade oral.

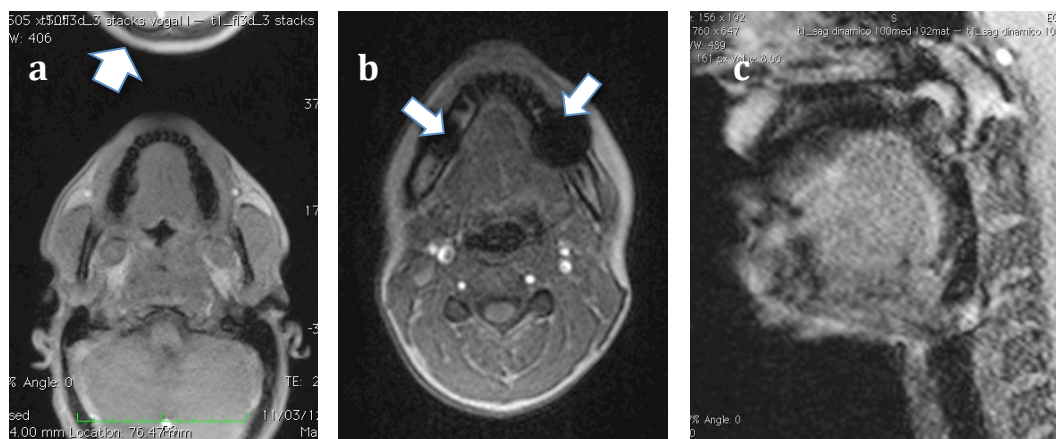
O artefacto de *aliasing* ou *wrap around* ocorre quando a imagem da anatomia fora do FOV selecionado interfere com a área em estudo, projetando-se tanto no eixo da fase como no eixo da frequência. Isto porque o gradiente é ativado em todo o seu comprimento para além do FOV selecionado. Este artefacto pode ser solucionado aumentando o FOV, usando bandas de pré-saturação (excitação dos tecidos na região adjacente ao FOV) ou não fazendo a leitura das áreas

adjacentes à região em estudo (“*oversampling*”) no eixo da fase (com penalização da SNR e do tempo de aquisição) ou da frequência.

O artefacto de movimento do paciente, pode ser voluntário ou involuntário, ou ainda por efeito do fluxo dentro de um vaso sanguíneo, traduzindo-se através de uma imagem distorcida da estrutura em movimento. Ocorre apenas no eixo da fase e pode ser solucionado com uma correta imobilização, instrução ou sedação do paciente.

Por fim, o artefacto “metálico” surge como uma perda de sinal no tecido decorrente da presença de objetos ferromagnéticos ou metálicos na região em estudo ou na sua proximidade. Ocorre em ambos os eixos da codificação e pode ser solucionada usando sequências spin-eco ou removendo os metais dessa região.

A título ilustrativo, a Figura 3.5, demonstra três imagens do trato vocal com os artefactos anteriormente descritos, obtidas no decurso das várias recolhas imagiológicas efetuadas aquando da instrumentação do protocolo de RM.



**Figura 3.5 Artefactos encontrados nos estudos do trato vocal por RM. A imagem (a) ilustra o artefacto de *aliasing* encontrando-se a parte posterior da cabeça (seta) à frente da parte anterior. Em (b) as setas ilustram dois artefactos metálicos causados por material chumbíneo existente nos dentes. E, a imagem (c) obtida na sequência do estudo dinâmico exemplifica o artefacto de movimento pela falta de definição (distorção) dos contornos de todo o trato vocal.**

### 3.4 Sequências de pulsos e a sua aplicação no estudo das vias aéreas superiores

As sequências de pulsos definem uma combinação de pulsos de RF e gradientes, que por variação dos mesmos permitem a obtenção de diferentes imagens mediante as estruturas que se pretendem estudar (HAAKE et al., 1999; HASHEMI & BRADLEY, 1997). A sequência de eventos que ocorre durante um determinado tempo de aquisição é representada por um diagrama temporal caracterizando assim uma sequência de pulsos:

- I. Aplicação de pulsos de  $90^\circ$  e  $180^\circ$  separados por um  $TE/2$ ; cada pulso de RF dura cerca de 2 a 10 ms;
- II. Após um dado TE e após cada pulso de  $90^\circ$  é originado um eco;
- III. O gradiente de seleção de corte liga durante duas emissões de RF;
- IV. Imediatamente antes da receção do eco é ativado o gradiente de codificação de fase que tem uma duração entre 1 a 5 ms;
- V. O gradiente de codificação de frequência é ativado durante o período de leitura do eco e dura cerca de 10 ms;
- VI. Toda esta sequência de eventos é repetida até o TR definido.

Os componentes essenciais, para a definição de qualquer sequência de pulsos são:

- O pulso de RF de excitação necessário para ocorrer o fenómeno de ressonância;
- Os gradientes para codificação espacial e cujo arranjo determina a forma como o *espaço K* é preenchido;
- A leitura do sinal através da combinação de um ou mais tipos de ecos determinando o tipo de contraste da imagem.

Para além destes, poderão existir componentes opcionais como outros pulsos de RF, gradientes ou outros métodos de reconstrução de imagem, com a finalidade de modificar o contraste, tornar a sequência mais rápida ou ainda reduzir artefactos.

A escolha e conjugação adequada dos parâmetros de aquisição (abordados anteriormente) permite assim, encontrar o melhor compromisso entre o



contraste, a resolução espacial e rapidez do exame face à finalidade de cada estudo.

Dada a ampla possibilidade de combinações possíveis, são várias as sequências de pulsos que podem ser utilizadas, tendo cada uma destas, uma finalidade específica. Para além disso, cada fabricante dispõe de nomenclaturas distintas (acrónimos) para as mesmas sequências.

No entanto, quanto ao seu princípio básico as sequências de pulso são divididas em dois grandes grupos: as sequências spin-eco, as sequências gradiente-eco, com numerosas variações possíveis, Tabela 3.7.

As sequências spin-eco e gradiente-eco diferem entre si no modo como é feito o refasamento dos núcleos após o pulso de excitação e conseqüentemente na produção do eco. Na sequência spin-eco, o refasamento é obtido pela aplicação de um pulso de 180°, enquanto na sequência gradiente-eco ocorre por ação dos gradientes (tornando-as mais sensíveis a artefactos).

**Tabela 3.7 Descrição sumária das sequências de pulsos básicas usadas em ressonância magnética (adaptado de LUFKIN, 1999).**

Sequência	Descrição sumária
Spin-eco (SE)	<p>É a sequência de pulsos mais comumente usada na prática clínica, pela excelente SNR das imagens que produz; no entanto, estas sequências apresentam tempos de aquisição mais longos do que as sequências gradiente-eco. Os parâmetros de aquisição de maior interesse para a sua programação são o TR e o TE.</p> <p>Estas sequências são obtidas com um pulso de excitação inicial de 90°, seguido de um ou mais pulso de 180° para refasamento dos prótons. Cada pulso de refasamento origina a produção de um eco (sinal), sendo que a produção de multiecos destina-se apenas a avaliar a evolução do sinal das estruturas em diferentes TE (ponderações).</p> <p>A sequência <i>turbo-SE</i> ou <i>fast-SE</i> (FSE) tem o mesmo princípio que a SE convencional, diferenciando-se desta última, pela aplicação de um conjunto de pulsos de refasamento de 180°, de forma a produzir o correspondente número de ecos. A este conjunto de pulsos dá-se o nome de ETL.</p> <p>Habitualmente são produzidos quatro a dezasseis ecos, o que permite uma redução do tempo de sequência. Estas sequências têm vindo a substituir as sequências spin-eco convencionais, devido aos tempos de aquisição serem substancialmente reduzidos. Têm principal aplicação nos estudos cerebrais e da coluna vertebral, pela sua rapidez e elevada ponderação T2.</p>

**Tabela 3.7 Descrição sumária das sequências de pulsos básicas usadas em ressonância magnética (adaptado de LUFKIN, 1999) – continuação.**

Sequência	Descrição sumária
Gradiente-eco (GE)	<p>O TR, o TE e o FA são os parâmetros de aquisição de maior interesse para a sua programação.</p> <p>Estas sequências utilizam pulsos de excitação com FA inferiores a 90°, o que condiciona a qualidade de imagem (menor amplitude de sinal). A restituição de fase (ou refasamento) é efetuada pela aplicação de um gradiente de campo magnético. A aplicação de gradientes não compensa os distúrbios na homogeneidade do campo magnético, podendo desta forma surgir artefactos de susceptibilidade magnética (perda de sinal nos tecidos); assim, são menos utilizadas comparativamente às sequências SE.</p> <p>A sequência <i>turboflash</i> em gradiente-eco baseia-se na redução do TE/TR dos gradientes, permitindo tempos de aquisição de imagens na ordem de um segundo. Como desta redução resulta uma perda considerável do contraste da imagem, torna-se necessário a aplicação de pré-pulsos de RF antes da aquisição, preparando assim os tecidos. Dependendo dos fabricantes pode assumir diferentes terminologias GRASS, FISP ou FLASH.</p> <p>Devido ao seu carácter ultrarrápido, tem como principais aplicações clínicas, o estudo dinâmico das estruturas (como o coração e o trato vocal durante a fala) com elevada resolução temporal, estudos em apneia de forma a superar os artefactos respiratórios e a obtenção de imagens de estruturas muito pequenas.</p>

O desenvolvimento das sequências ultrarrápidas ou ecoplanares possibilitou a obtenção de imagens em menos de 100 ms, permitindo como aplicações clínicas, o estudo da difusão e perfusão cerebral, o estudo funcional cerebral, estudos abdominais, a avaliação intrauterina do feto e ainda o estudo simultâneo da morfologia e função cardíaca. No entanto, exigem um grande esforço dos gradientes, provocando um intenso ruído acústico, por comparação às outras sequências (LUFKIN, 1999).

#### 3.4.1 Técnicas de sincronização

Os sistemas de deteção e seguimento do movimento humano em tempo-real, também conhecidos por *tracking systems*, têm despertado o interesse dos investigadores, pelo vasto leque de aplicações em diversas áreas desde a Engenharia e Computação até à Medicina. Em Imagem médica, os sistemas de

deteção e seguimento do movimento permitem, por um lado, dar feedback ao operador quanto ao movimento do sujeito durante a RM, minimizando a probabilidade de degradação da imagem devido à ocorrência de artefactos, ou por outro, detetar e usar os eventos dinâmicos de determinado órgão para sincronizar as sequências de pulso.

Para a deteção de movimento, o sistema deve procurar numa sequência de imagens, sinais que confirmem a existência de movimento no ambiente monitorizado (deteção), podendo-se assim realizar o seguimento (*tracking*) do órgão ou região em estudo (MULDER, 1994).

Os sistemas de deteção do movimento humano podem ser:

- Internos: usam sensores e fontes, ambos no interior do corpo (*“inside-in”*);
- Externos: usam sensores no corpo que detetam uma fonte artificial externa (*“inside-out”*);
- Totalmente externos: usam um sensor externo que deteta uma fonte artificial ou marcas no corpo, ou seja, o movimento é detetado externamente ao corpo (*“outside-in”*).

Das tecnologias disponíveis, os sistemas *outside-in* são menos invasivos e permitem detetar movimentos em maiores áreas do corpo. A eletromiografia, a eletro-oculografia humana e a eletro-neuromiografia são alguns exemplos de tecnologias disponíveis dos sistemas *inside-in*, cuja deteção está limitada a áreas mais pequenas. Os sistemas *inside-out* permitem detetar movimentos de grandes áreas do corpo e poderá ser conseguida, por exemplo, através da utilização de uma câmara de vídeo para captura de luz de díodos foto emissores (LEDs) reflectida. As marcas refletoras (colocadas, por exemplo, na pele), ou padrões (a cores ou preto e branco) são exemplos de sistemas *outside-in*.

Assim, para estudos da região da cabeça e pescoço por RM, os sistemas de deteção de movimento devem ter as seguintes características (LAM & GEIGER, 2002):

- Ser confortável;
- Compatível com o campo magnético e *hardware*;
- Não-invasivo;
- Seguro;

- Não interferir com a antena ou mesa de exames;
- Ter alta sensibilidade (1 mm nos três eixos).

As soluções possíveis encontradas na literatura e face às características exigidas incluem: os sistemas de detecção ótica, os sensores pneumáticos, os sensores laser e as placas de pressão, com vantagens e limitações próprias.

A detecção ótica consiste num sensor metal-óxido similar ao sistema de um rato ótico para computador pessoal. Os sensores pneumáticos funcionam como uma almofada que consiste num suporte com compartimentos de ar e sensores de pressão. Os sensores laser medem o número de movimentos a partir de pontos de referência na cabeça. As placas de pressão efetuam a sensorização da deslocação do peso da cabeça do sujeito.

Para a detecção dos movimentos dos articuladores durante a fala, as soluções mais adequadas incluem a detecção ótica ou os sensores laser. A detecção dos movimentos permitiria localizar espacialmente por exemplo, os dentes e outros articuladores complementando assim a informação morfológica obtida em RM.

O único estudo encontrado na literatura relativo à aplicação de um *tracking system* para estudo da produção da fala foi proposto por DROMEY e colaboradores (2006); estes autores desenvolveram um dispositivo de detecção e seguimento dos movimentos da mandíbula, com o objetivo de medir os movimentos da língua durante a fala, de uma forma simples e mais económica, comparativamente à articulografia eletromagnética, que é usada apenas em investigação devido ao seu elevado custo.

Para o efeito, os autores conceberam um instrumento leve, de fixação na cabeça, constituído por dois sensores magnéticos colocados a cada lado da face do sujeito. Estes sensores permitiram medir movimentos antero-posteriores, verticais e laterais, através da detecção e seguimento de marcas colocadas, um nos incisivos inferiores e outro na superfície da língua. Apesar das limitações encontradas, um dos potenciais desta aplicação, é segundo os autores, a possibilidade de ser utilizado como dispositivos de bio feedback durante o tratamento de *deficits* articulatórios.

### 3.4.2 Monitorização de parâmetros fisiológicos

Em Imagem médica, qualquer movimento do paciente é considerado um problema, encontrando-se disponíveis atualmente, várias técnicas para reduzir ou compensar os seus efeitos. Para além disso, em pacientes sedados, anestesiados, inconscientes ou incapazes de comunicar com o técnico (por exemplo, uma criança), a monitorização fisiológica é necessária para garantir a segurança destes aquando da obtenção das imagens (LUFKIN, 1999).

Atualmente, vários monitores de parâmetros fisiológicos são compatíveis com o campo magnético, contudo um sistema de RM possui dispositivos de monitorização dos seguintes dos parâmetros fisiológicos:

- Pressão arterial: medida através de um transdutor de pressão acoplado à braçadeira insuflável por um tubo que enche de forma pneumática;
- Frequência respiratória, oxigenação e trocas gasosas: medições efetuadas através de um oxímetro de pulso (para registo da saturação de oxigénio e da frequência cardíaca), um capnógrafo (para monitorizar o dióxido de carbono e gases anestésicos) e uma cinta respiratória (para detetar os movimentos respiratórios);
- Frequência cardíaca: medida através da eletrocardiografia.

Estes dispositivos de sincronismo (*gating*) permitem não só obter medidas quantitativas fisiológicas (cardíaco e respiratório) assim como compensar os artefactos de movimento, através da utilização de algoritmos especiais, *fiducial markers*<sup>2</sup>, aquisições rápidas, técnicas de registo espacial ou outras estratégias alternativas de aquisição das imagens.

No *gating* cardíaco, são recolhidas múltiplas imagens dentro do intervalo de tempo do ciclo cardíaco, correspondendo a diferentes fases da função cardíaca, através da monitorização por eletrocardiografia. Por cada codificação de fase obtida é preenchida uma linha do *espaço K*.

Assim, para a combinação das imagens de RM com outras técnicas de avaliação articulatória, como a EPG, ou por exemplo a EGG, é necessário:

- Assegurar a viabilidade da instrumentação destas técnicas em ambiente de RM;

---

<sup>2</sup> Termo de origem inglesa usado para definir uma marca, ponto ou objeto que é identificado no FOV da imagem para servir de referência ou medir eventos dinâmicos.

- Ajustar ou definir novos protocolos de RM adequados ao evento fisiológico a monitorizar;
- Registrar acusticamente a fala durante os estudos de RM para posterior análise do sinal.

A utilização de uma antena miniaturizada (micro-antena de RF) embutida num endoscópio de fibra ótica foi descrita por ICHIKAWA e colaboradores (2001), permitindo a combinação da RM e da endoscopia no diagnóstico de lesões gastrointestinais. Esta junção de técnicas permite por um lado superar algumas das limitações da endoscopia, como a perda de orientação, a impossibilidade de ultrapassar obstruções e por outro, a impossibilidade da RM de visualizar o aspeto da mucosa intestinal e lúmen das vísceras.

#### 3.4.3 Potenciais aplicações clínicas no estudo das vias aéreas superiores

A implementação e utilização da RM no estudo das vias aéreas superiores tem vindo a crescer, sendo atualmente já reconhecidas várias áreas potenciais de aplicação clínica, nomeadamente:

- Na avaliação de pacientes com Síndrome da Apneia Obstrutiva do Sono (SAOS), (OKUBO et al., 2006; TOGEIRO et al., 2010; ZHA et al., 2008);
- Na avaliação de pacientes traqueostomizados (FRICKE et al., 2007);
- Na avaliação pré e pós-cirúrgica do trato vocal (MÁDY et al., 2001, 2002);
- No estudo da voz cantada (ECHTERNACH et al., 2010, 2011);
- Na avaliação da etiologia de roncopatias (SCHWAB et al., 2003);
- Na avaliação de pacientes com problemas de deglutição;
- Na avaliação de perturbações articulatórias.

Por fim, o sistema de arquivo, visualização e comunicação de imagens (denominado por PACS), usado para partilha de dados não está só limitado à Radiologia, estendendo-se também a outras modalidades de imagem como a endoscopia, a patologia, a dermatologia, a dentária e a oftalmologia. Assim, a integração de informação multidisciplinar, neste âmbito, permitirá aumentar a eficácia do diagnóstico de patologias articulatórias, do tratamento e *follow-up* de lesões.

*a) Morfologia do trato vocal:*

Morfologicamente, o trato vocal apresenta uma forma tubular em “L” invertido delimitado maioritariamente por órgãos de tecidos moles (essencialmente musculares), ainda por estruturas ósseas. Este tubo é fechado ao nível da glote e termina ao nível dos lábios.

Os tecidos moles, em anatomia, designam todas as estruturas de interligação, proteção e apoio de outras estruturas e órgãos, nomeadamente músculos, tendões, gordura ou tecido adiposo, vasos sanguíneos e nervos. No caso do trato vocal, são órgãos essencialmente musculares e tecido adiposo subcutâneo, cujos tempos de relaxação são muito curtos. Num estudo realizado por GOLD e colaboradores (2004) os tempos de relaxação T1 dos tecidos do sistema músculo-esquelético, medidos *in vivo*, apresentam um acréscimo na ordem dos 20 a 22% num sistema de 3.0 T comparativamente a um de 1.5 T, sendo no músculo cerca de 1420 ms e, para a gordura subcutânea de 371 ms num sistema de 3.0 T.

As estruturas rígidas ósseas do trato vocal, a mandíbula e o palato duro, apresentam uma intensidade de sinal intermédia em ambas as ponderações de imagens.

Os dentes, a porção mais externa e rígida dos ossos (osso compacto) e o ar não são identificáveis nas imagens de RM, pela baixa ou inexistente quantidade de hidrogénio.

Assim, no estudo do trato vocal a melhor ponderação de imagem em RM, é a ponderação T1, pois proporciona um melhor contraste e diferenciação anatómica entre os tecidos moles (melhor SNR), para além de exigir tempos de eco e de repetição curtos e consequentemente um menor tempo de aquisição. A redução do tempo de aquisição é um importante fator no estudo da fala por RM, possibilitando não só a caracterização funcional como a redução do tempo de sustentação dos sons.

Uma vez que o trato vocal se estende anatomicamente por duas regiões muito distintas, a cabeça e o pescoço, justifica ainda a utilização/ combinação de duas antenas de radiofrequência, de modo a uniformizar a relação sinal-ruído.

A considerável extensão anatómica do trato vocal aliada à sua forma irregular condiciona o desenho do protocolo, pelo que a espessura e orientação dos cortes

deve ser resultado do compromisso entre o número de cortes pretendido, relacionado com a região em estudo e a relação sinal-ruído pretendida. Ou seja, cortes finos acarretam um maior número de cortes, no entanto, a relação sinal-ruído diminui pois menos prótons de hidrogénio estão presentes na amostra. Quanto à orientação dos cortes, o plano sagital permite uma visão global de todo o trato vocal e simultaneamente da sua forma e extensão. A orientação coronal permite observar a cavidade oral, enquanto a orofaringe é melhor observada no plano axial.

Segundo GOLD (2011), num sistema de RM de 3.0 T, o TR e TE devem ser programados de modo a evidenciar melhor os tecidos em estudo e a obter o contraste desejado. Na maioria das situações e no sistema músculo-esquelético, os tempos de relaxamento T1 dos tecidos são superiores comparativamente a um sistema de RM de 1.5 T, assim o TR selecionado deve ser mais longo para se obter o mesmo tipo de contraste. Similarmente, o TE deve ser ligeiramente mais curto pois os tempos de relaxação T2 são consideravelmente menores num sistema de alto campo (3.0 T).

*b) Atividade funcional do trato vocal:*

Funcionalmente, o movimento rápido dos articuladores condiciona a passagem de ar pelas diferentes cavidades de ressonância, exigindo por isso o recurso a sequências muito rápidas para a sua análise. Tratam-se, na sua maioria, de órgãos musculares estriados ou esqueléticos, responsáveis por movimentos voluntários. O articulador mais volumoso e ativo é a língua, cujos movimentos são muito diversificados e complexos. O seguimento e análise da dinâmica dos articuladores constitui assim um desafio do ponto de vista da Imagiologia.

Do ponto de vista fisiológico, os eventos dinâmicos da articulação da fala apresentam uma frequência fundamental na ordem dos 80-300 Hz, no entanto, durante uma conversa normal a velocidade da fala ou ritmo articulatorio evolui mais lentamente. Por exemplo, o valor obtido experimentalmente ouvindo e anotando o texto falado foneticamente seguido da contagem e cálculo da médias, para uma voz masculina profissional durante uma leitura, foi de 12.6 fonemas/s. Considerando ainda, o número de articuladores envolvidos e as diferentes



velocidades que apresentam (pois a sua inervação é distinta), o seu correto estudo e análise por RM exige, idealmente, o recurso:

- A um *trigger* específico para sincronismo da aquisição com os eventos articulatorios, permitindo melhorar a resolução temporal dos sistemas de RM (com as características descritas na secção 3.4.2);
- A gravação áudio simultânea durante a aquisição das imagens, permitindo correlacionar a análise acústica do som e o conjunto de imagens obtidas.

Estas soluções ainda não estão disponíveis na prática clínica. No entanto, através da utilização de sequências ultrarrápidas ou de técnicas avançadas de RM o seguimento e análise da dinâmica articulatória já é possível ainda que o rigor das medições careça de maiores aperfeiçoamentos.

A resolução temporal é um aspeto importante em múltiplas aplicações clínicas em RM; a velocidade com que os dados podem ser recolhidos está fundamentalmente limitada pelo desempenho dos gradientes (como a amplitude de gradiente e *slew rate*) e por restrições físicas e fisiológicas intrínsecas dos pacientes (por exemplo, estimulação nervosa). A utilização de estratégias rápidas de preenchimento do *espaço K* origina uma subamostragem violando o Teorema de Nyquist<sup>3</sup>. De acordo com este teorema, a reconstrução com o mínimo de perda de informação para qualquer sinal analógico, requer teoricamente pelo menos uma frequência de amostragem do sinal (isto é, uma quantidade de amostras por unidade de tempo geralmente medida em Hertz) igual ou duas vezes superior à maior frequência do espectro desse sinal. O limite da frequência do sinal imposto pelo Teorema de Nyquist deve ser, no máximo, igual a metade da frequência de amostragem. Por exemplo, num sinal com frequência de 3 Hz, a frequência de amostragem deve ser de, pelo menos, 6 Hz. Quando a frequência do sinal é superior a este limite surge o fenómeno de *aliasing* devido à subamostragem, sendo modificada a forma do sinal e perdendo-se assim a fidelidade da amostragem.

Nesse sentido e porque o tempo necessário para a aquisição rápida de imagens ainda pode ser demasiadamente longo, principalmente quando a velocidade dos eventos fisiológicos é elevada ou quando a SNR é baixa, foi

---

<sup>3</sup> Também conhecido por Teorema da Amostragem de Nyquist-Shannon, define a quantidade de amostras por unidade de tempo de um sinal, fornecendo condições suficientes (e não reais) para a reconstrução de um sinal analógico.

desenvolvida uma técnica denominada de *compressed sensing* ou *sparse sampling* (LUSTIG et al., 2007). Esta técnica, cada vez mais atrativa do ponto de vista clínico, permite uma redução substancial do tempo de aquisição num sistema de RM, reconstruindo imagens a partir de um número significativamente menor de medições, que seriam tradicionalmente consideradas necessárias. Em processamento de sinais, consiste na aquisição e reconstrução de um sinal considerado escasso ou compressível, aproveitando a redundância de muitos dos sinais. Isto acontece porque, alguns sinais contêm muitos coeficientes próximos ou igual a zero, quando representados em algum domínio (espacial ou temporal). Em RM, esta técnica tem vindo a ser utilizada principalmente no estudo cardíaco.

No que respeita aos protocolos de estudo do sistema músculo-esquelético, num sistema de RM de 3.0 T, estes podem ser concebidos, por um lado, de forma a obter uma resolução espacial e qualidade de imagem elevadas ou por outro, para privilegiar a resolução temporal e eficiência (GOLD, 2011). Assim, o compromisso necessário entre a SNR e resolução da imagem pretendida tem vindo a condicionar ainda de certa forma a capacidade de estudo *in vivo*, por RM, com elevada resolução e de um modo eficiente (com menores tempos de aquisição).

### 3.5 Conclusão

Os sistemas de RM, embora possuam especificidades técnicas próprias quer de *hardware* quer de *software*, são caracterizados essencialmente pela sua potência de campo magnético. Para fins clínicos, a potência de campo magnético dos equipamentos de RM disponíveis varia entre 0.2 a 3.0 T; já para fins de investigação, são utilizadas potências de campo magnético acima de 7.0 T. Em Portugal tem-se assistido à implementação gradual e amplificação dos sistemas de RM; no entanto, o elevado custo de instalação e manutenção dos equipamentos constitui ainda um importante condicionalismo à sua utilização. Este custo aumenta de acordo com a potência de campo, sendo os equipamentos de 3.0 T disponíveis ainda em número reduzido, diminuindo assim a sua acessibilidade.

Os princípios físicos da RM são complexos exigindo um conhecimento rigoroso do seu princípio de funcionamento, das suas técnicas de aquisição e sequências de pulso. A qualidade da imagem assume particular relevância, sendo primordial a procura de situações de compromisso entre a relação sinal-ruído e o tempo de aquisição. A adoção de protocolos de estudo dirigidos à especificidades anatómicas e funcionais dos órgãos é um fator decisivo na garantia da qualidade de imagem.

No caso particular do trato vocal, ainda que do ponto de vista morfológico apresente similaridades com os órgãos do sistema músculo-esquelético, do ponto de vista funcional, tem propriedades análogas ao sistema cardiovascular. Estas características justificam assim a implementação cuidadosa e otimização de um protocolo de RM, com vista à sua utilização quer do ponto de vista clínico quer da Engenharia.

Assim, os principais requisitos e potencialidades da RM para estudo da fala são os seguintes:

- Visualização de todos os articuladores, em qualquer plano ou orientação, em situações estáticas e dinâmicas;
- Estudo bidimensional e tridimensional do trato vocal;
- Caracterização morfológica e funcional dos articuladores, nomeadamente, em pacientes com perturbações articatórias;
- Extração de medidas quantitativas sobre a posição dos articuladores e as áreas seccionais para modelação do trato vocal e posterior síntese da fala.



## **CAPÍTULO 4. UTILIZAÇÃO DA RESSONÂNCIA MAGNÉTICA NA CARACTERIZAÇÃO DA FALA**

A comunicação através da fala é uma das atividades humanas mais utilizadas, dotando o Homem de uma forma única, inigualável e distintiva, permitindo a interação e convivência em sociedade. Aliás, é graças a esta característica que podemos exprimir muito mais do que simples sons, como os pensamentos, sentimentos e emoções.

Dentro do mecanismo da produção da fala, diferencia-se a articulação, pois graças ao movimento (dinâmica) e alteração da forma (morfologia) dos órgãos do trato vocal, são produzidos diferentes sons, com propriedades acústicas distintivas.

A junção do conhecimento morfológico à informação dinâmica do trato vocal é, por isso, muito ambicionada quer pela Fonética que se dedica ao estudo da produção dos sons, quer pela Engenharia na conceção e desenvolvimento de sistemas para síntese e reconhecimento da fala.

As técnicas de avaliação da articulação da fala são diversas, assumindo as técnicas imagiológicas particular relevância, nomeadamente a imagem por ressonância magnética, pela capacidade de aliar a imagem morfológica ao estudo dinâmico.

Assim, é imperativo responder às seguintes questões: Quais as potencialidades e condicionalismos da RM nos estudos da fala? Que técnicas têm vindo a ser desenvolvidas e que resultados foram obtidos até à data?

Este capítulo revê o estado da arte relativamente à utilização e aplicação da RM na caracterização da fala, com destaque para a descrição das potencialidades e das principais condicionantes desta técnica de imagem.

## **4.1 Perspetiva histórica: desenvolvimento da técnica e principais resultados obtidos**

Os estudos conduzidos por ROKKAKU e colaboradores (1986), e um ano mais tarde, por BAER e colaboradores (1987), impulsionaram uma das mais prementes áreas de investigação: a produção da fala, através da utilização da imagem por RM. Os rápidos avanços tecnológicos da RM deram origem ao desenvolvimento de técnicas de imagem variadas, com potencialidades em várias áreas, como a Engenharia Biomédica, a Medicina, a Terapia da Fala, a Fonética.

O reconhecimento da utilidade da RM é patenteado por vários investigadores, podendo-se observar a expansão, divulgação e as inúmeras potencialidades desta modalidade de imagem médica para diferentes aplicações (MASAKI et al., 2008).

Na década de 80 do século passado, as primeiras técnicas de imagem de RM desenvolvidas para o estudo da fala, estavam confinadas à obtenção de um número muito restrito de cortes, maioritariamente no plano sagital, em resultado dos longos tempos de aquisição. Esta condicionante obrigava, por isso, à sustentação muito prolongada, na ordem dos 40 segundos, de alguns sons, principalmente as vogais e algumas consoantes.

Esta era de estudos, denominados de estáticos por observarem posturas articatórias do trato vocal mantidas prolongadamente durante a aquisição, revelou-se insuficiente na caracterização da produção da fala, pois claramente várias condicionantes técnicas predominavam, nomeadamente:

- A insuficiente resolução temporal da RM, tanto para estudo da dinâmica articatória como para observar movimentos mais finos de alguns órgãos, como a laringe;
- A não identificação dos dentes devido ao baixo sinal que apresentam nas imagens;
- A postura imposta aos sujeitos de decúbito dorsal, traduzindo-se numa posição incomum e pouco natural desta atividade motora humana.

Para além disso, na qualidade das imagens refletia-se ainda a presença de um elevado número de artefactos de movimento e de ruído, apresentando assim uma qualidade inferior para caracterizar de forma suficiente os eventos articulatórios durante a produção da fala.

Na década de 90, surge então a primeira aplicação de uma técnica de sincronização cardíaca, desenvolvida por FOLDOVIK e colaboradores (1990), que permitia observar o movimento dos articuladores. A era dos estudos dinâmicos que se seguiu, desencadeou novas potencialidades da RM e um maior conhecimento funcional do trato vocal. Este avanço tornou possível o estudo de mais sons, em diferentes contextos articulatórios, associando-se uma melhor informação morfológica graças ao desenvolvimento tecnológico dos sistemas de RM.

O refinamento dos protocolos de aquisição e das técnicas de visualização, tem sido atualmente o principal objetivo da investigação na tentativa de estabelecer aquisições tridimensionais do trato vocal e sem a necessidade de repetição dos sons.

Para além dos resultados qualitativos morfológicos e dinâmicos obtidos pela RM, nomeadamente pela possibilidade de avaliação funcional dos articuladores e de forma não-invasiva do ponto de vista clínico, podem ser extraídos dados quantitativos como funções de área, distâncias, e volumes, importantes na reavaliação da teoria acústica e nos estudos da fonação.

#### 4.1.1 Estudos estáticos

Por volta dos anos noventa, os recentes estudos da fala através da imagem por RM abriram novos horizontes na observação dos movimentos articulatórios essencialmente com base em medições a partir de imagens médio sagitais do trato vocal. No entanto, em virtude da baixa performance dos sistemas de RM até então disponíveis, impunha-se a necessidade de incrementar a sua resolução temporal e a qualidade das imagens produzidas com elevado ruído e/ou artefactos. Os primeiros estudos estáticos resumiam-se ao estudo das vogais por serem sons fáceis de sustentar e assim, proporcionarem uma maior estabilidade

dos articuladores do trato vocal. Só posteriormente se verificou a extensão para o estudo das consoantes.

DEMOLIN e colaboradores (1996) propõem uma nova técnica de aquisição multiplanar, tendo por base a programação e aquisição de um conjunto de cortes, denominados de “*stacks*” em três orientações distintas (5 cortes coronais, 3 coronais oblíquos e 6 cortes transversais) a partir de uma imagem médio sagital de referência. Até então, na maioria dos estudos realizados, os modelos 3D do trato vocal eram obtidos através da inferência das funções de área a partir de uma imagem médio sagital (BAER et al., 1987; ROKKAKU et al., 1986).

A necessidade de melhorar o rigor destas medidas rapidamente se impôs, justificando assim um incremento no número de imagens recolhidas e em diferentes orientações considerando a forma anatômica irregular do trato vocal. No estudo descrito em DEMOLIN e colaboradores (1996) foram recolhidos catorze cortes mediante a sustentação de dez vogais orais do Francês durante cerca de 14 s, permitindo a posterior extração de áreas seccionais em diferentes localizações do trato vocal. No entanto, estes autores reportaram algumas dificuldades relacionadas com esta técnica de aquisição, nomeadamente: a) a degradação e perda de sinal nas zonas de intercepção dos diferentes cortes; b) a identificação da fronteira ar-tecidos moles e dos dentes; e c) dificuldade na segmentação do trato vocal pela similaridade de sinal que algumas estruturas adjacentes apresentam.

DEMOLIN e colaboradores (1998) aperfeiçoaram esta técnica conseguindo um aumento significativo do número de cortes: 55 cortes em três orientações distintas de forma a cobrir toda a extensão do trato vocal e respeitar a perpendicularidade à sua linha média. Contudo, tal conduziu a um maior tempo de aquisição (43 s), acarretando uma sustentação muito prolongada do som. Esta abordagem foi posteriormente melhorada por KROGER e colaboradores (2000), obtendo os mesmos resultados em cerca de metade do tempo de aquisição (21 s). A análise quantitativa das imagens por RM consistia basicamente na extração de áreas, de forma predominantemente manual.

A partir desta data, a extração manual foi substituída por análises estatísticas e mais automatizadas como pelo método de análise das componentes



principais<sup>4</sup> (PCA), como demonstraram também BEHRENS & WISMULLER (2001).

O estudo preliminar realizado por DEMOLIN e colaboradores (1998) demonstrou a fiabilidade da utilização de um modelo estatístico 3D do trato vocal, mas principalmente evidenciou o trabalho que era necessário efetuar no sentido de melhorar o viés das medições em detrimento da ausência de pontos fiáveis de referência ósseos e da dificuldade de aplicar um método de segmentação em determinadas regiões como o ápex da língua, a epiglote, entre outras.

A variabilidade intersujeitos do sinal acústico da fala é demonstrada por APOSTOL e colaboradores (1999), através da correlação entre os formantes de diferentes sujeitos e as funções de área obtidas a partir da reconstrução 3D do volume do trato vocal, obtida por interpolação dos diferentes contornos. Persistiam, no entanto, os tempos de aquisição das imagens por RM muito prolongados (43 s), obrigando a um grande esforço, a uma sustentação artificial e a paragens para respiração por parte dos sujeitos.

Ainda que estes autores tenham observado as limitações deste método, relacionadas com os padrões de correspondência da correlação e com a natureza de recursos (dados imagiológicos *versus* acústicos) de sujeito para sujeito, a segmentação do trato vocal continuava claramente com os problemas clássicos por resolver. Com o objetivo de descrever uma solução computacional que permitisse a construção de um modelo do trato vocal, de forma relativamente rápida e tendo em conta que existem diferenças intersujeitos do sinal acústico, a partir das imagens por RM e da informação acústica, NIETO-CASTANON & GUENTHER (1999) realizaram a caracterização do espaço articulatorio com recurso à PCA. Segundo estes autores, o modelo resultante pode ser utilizado na diferenciação entre os efeitos morfológicos específicos e as estratégias motoras independentes de cada sujeito.

A primeira modelação a 3D do trato vocal foi desenvolvida por ENGWALL (1999), através da criação de uma solução computacional para síntese da fala

---

<sup>4</sup> Do inglês *Principal Component Analysis* (PCA) é um dos métodos estatísticos mais usados quando a natureza dos dados é muito diversa, permitindo transformar um conjunto de variáveis originais Inter correlacionadas, num novo conjunto de variáveis não correlacionadas – as componentes principais. No estudo da produção da fala, este método foi primeiramente introduzido por MAEDA (1982) na análise de radiografias de perfil obtidas por raios-X.

pela análise visual. Segundo este investigador, é evidente a importância do rigor da extração da caracterização morfológica e dinâmica (articulatória) para o sucesso deste modelo; ou seja, através da incorporação de informação proveniente tanto de estudos estáticos como dinâmicos do trato vocal humano. Acrescentou que, a investigação deve focar-se não só em novas técnicas de medição da articulação da fala, como o raios-X, a ressonância magnética, a ultrassonografia, como também na extração de medidas quantitativas rigorosas e através da aplicação de PCA nas imagens.

A grande maioria dos autores desta área do conhecimento demonstrou a importância da utilização desta informação numa das suas potenciais áreas de aplicação: a Engenharia. Alguns dos estudos associados estão indicados a seguir.

ENGWALL & BADIN (2000) demonstraram a influência da coarticulação através do estudo de vários sons em contexto articulatório vogal-consoante-vogal, pois até então os sons eram estudados isoladamente, ou seja, fonemas.

A modelação 3D da língua usando a PCA com base nas imagens de RM foi realizada por ENGWALL (2000a) e ainda por BADIN e colaboradores (2000) aliando as imagens de vídeo dos lábios e face.

ENGWALL (2000b) com o objetivo de avaliar se as medições obtidas nos estudos estáticos por RM são representativas da dinâmica da fala, procedeu a um estudo comparativo com técnicas de avaliação da função articulatória: a EPG e a EMA. Apesar de existirem diferenças, ENGWALL concluiu que ainda assim os estudos estáticos são representativos da dinâmica articulatória, devendo ter-se em conta nos sistemas de síntese da fala que se trata de um caso de hiperarticulação.

A síntese articulatória com recurso a diversas modalidades, como EPG e EMA, conjugados com a extração das funções de área a partir das imagens por RM foi realizada posteriormente por ENGWALL (2001a, b).

Também GICK e colaboradores (2002) reiteraram a importância e as vantagens da utilização da RM na caracterização de posições articulatórias estáticas para a construção de modelos articulatórios, dada a natureza única, individual e complexa da fala humana.

Por ainda persistirem algumas limitações relacionadas com a técnica de RM, ENGWALL (2003), analisou qual o efeito da sustentação prolongada dos sons

(superior a 30 segundos) durante a RM e da posição corporal (decúbito dorsal) imposta aos sujeitos. A sustentação prolongada dos sons exige um grande esforço dos sujeitos, justificando-se por isso, a utilização de tempos de aquisição muito mais curtos e, se possível, em contexto real. Devido à força da gravidade, a posição de decúbito dorsal conduz à modificação da posição e forma da língua, obrigando a uma análise cuidadosa. Assim, a conjugação destes dois efeitos implica a hiperarticulação da fala.

Estas limitações posturais são também observadas por KITAMURA e colaboradores (2005), através da comparação de imagens médio sagitais obtidas durante a articulação de vogais em duas posições distintas: sentado e deitado. Corroborando as conclusões de ENGWALL, estes investigadores comprovaram que existe influência nos tecidos moles e ósseos em detrimento da força de gravidade e da postura corporal e que estas modificações são mais notórias nos sons mais anteriores.

Um ano depois, BADIN & SERRURIER (2006) apresentam uma extensa base de dados com modelos 3D da língua, da mandíbula, do véu do palato e cavidade nasofaríngea com base na integração da informação proveniente de duas modalidades de imagem, a RM e a tomografia computadorizada. Estes autores são responsáveis pela criação de descrições geométricas 3D para um sujeito durante a produção de 46 sons do Francês.

STORY (2008) apresentou um estudo comparativo realizado para o mesmo sujeito em dois anos distintos, com o objetivo de caracterizar a variabilidade intrasujeitos. Com base nas funções de área extraídas a partir das imagens de RM e a respetiva correlação acústica, este investigador comprovou que as diferenças morfológicas observadas em diferentes tempos podem fornecer informação adicional quanto à amostragem das funções de área e assim, permitir um melhor entendimento desta variabilidade e, no futuro, o aperfeiçoamento dos sistemas de síntese da fala.

Mais recentemente, PROCTOR e colaboradores (2010) demonstraram a construção de modelos 3D da cavidade faríngea durante a produção de consoantes fricativas do Inglês, a partir da aquisição de uma *stack* axial, e a relevância da extração de medidas quantitativas para melhor caracterização dos eventos articulatorios envolvidos na produção destes sons. Através da medição

dos volumes, estes autores constataram que em três dos quatro sujeitos, as consoantes fricativas vozeadas geram um alargamento maior da cavidade faríngea por comparação aos seus pares opostos (não vozeados). Estas diferenças observadas só foram possíveis, segundo os autores, através da análise da geometria da faringe tridimensionalmente, constituindo assim uma mais-valia na caracterização morfológica do trato vocal. Resultados similares tinham também sido anteriormente obtidos por NARAYANAM e colaboradores (1995).

Atualmente, os estudos dinâmicos são mais predominantes que os estáticos, verificando-se nestes últimos um maior investimento no refinamento da técnica de RM, de forma a, por exemplo, reduzir o tempo de aquisição sem restrições no número total de cortes. Nessa perspectiva e tendo por base uma técnica mais rápida de reconstrução das imagens de RM (*compressed sensing*), KIM e colaboradores (2009) revelam a sua primeira aplicação no estudo de alta-resolução 3D das vias aéreas superiores. A técnica de aquisição volumétrica, possibilita a aquisição de uma fatia (“*slab*”) no plano sagital com 8 cm de espessura e a sua posterior reconstrução nos planos axial e coronal com 2 mm de espessura. De acordo com o fator de aceleração utilizado, os tempos de aquisição variam entre 12, 9 e 7 segundos. Estes autores constataram que com a utilização de um fator de aceleração de cinco vezes é possível, a segmentação do trato vocal sem perdas significativas e com um tempo de sustentação razoável.

*a) Estudos da nasalidade:*

A nasalidade é uma característica complexa de alguns sons e que está relacionada com o abaixamento do véu do palato. Esta mudança de configuração introduzida pelos movimentos deste órgão, traduz-se em mudanças significativas no sinal acústico da fala. O interesse do seu estudo é salientado na literatura tanto de Engenharia como de Medicina.

O primeiro progresso na modelação do trato vocal durante a produção das vogais foi conduzido por DANG & HONDA (1996) através da incorporação dos efeitos acústicos da fossa piriforme e do acoplamento da passagem velofaríngea. Estes dois “canais”, quando incluídos nas medições das imagens por RM do trato vocal, contribuem para a criação de modelos físicos mais realistas; nomeadamente, na caracterização das duas principais formantes das vogais.

Contudo, é o investigador DEMOLIN e colaboradores (1998) o precursor do estudo dos sons nasais, realizando a mensuração das diferenças na abertura velofaríngea por comparação com as suas vogais orais opostas. Esta simples abertura introduz fortes e complexas mudanças acústicas no trato vocal.

KANE e colaboradores (2002) demonstraram um novo método para avaliar a função velofaríngea num adulto, com base na aquisição volumétrica em diferentes planos de uma sequência de imagens cinemáticas durante a repetição sucessiva de uma sílaba. Estes investigadores atestaram a razoável resolução temporal desta técnica de imagem e também a capacidade da RM em demonstrar diferenças anatómicas específicas em pacientes com fenda palatina e labial. Também ETEMA e colaboradores (2002), apresentaram medidas quantitativas da confirmação do músculo levantador do véu do palato em duas situações biomecânicas distintas: ativo e inativo, de modo a observar não só as suas características anatómicas como também funcionais. Este estudo demonstra a importância da construção de um modelo biomecânico da passagem velofaríngea na percepção e entendimento deste mecanismo e também, para o desenvolvimento do tratamento cirúrgico.

Mais tarde, FUJITA (2005) demonstrou a relevância das cavidades hipofaríngeas na construção de modelos sólidos do trato vocal. Também os investigadores SERRURIER & BADIN (2005) desenvolvem o primeiro modelo 3D do véu do palato com base nas imagens de RM combinadas com imagens de tomografia computadorizada para localização de estruturas ósseas.

#### 4.1.2 Estudos dinâmicos

Comparativamente aos estudos estáticos, o estudo dinâmico apresenta, do ponto de vista da técnica de RM, maiores desafios, pelo que se constata que são em menor número os estudos existentes.

Na vertente clínica, os estudos encontrados são menos frequentes, o que denota a sua inferior implementação; no entanto, privilegiam principalmente a caracterização dinâmica ou funcional da produção da fala. O primeiro estudo foi apresentado por CRARY e colaboradores (1996), seguindo-se MEDURI e colaboradores (1999) e posteriormente mais dois estudos realizados por MÁDY

e colaboradores (2001, 2002). Por oposição à implementação destes estudos na área da Engenharia, o número de sujeitos é consideravelmente superior, em detrimento de um *corpus* da fala menor. Como potencialidades da utilização da RM demonstradas por estes investigadores incluem-se a utilidade na avaliação de *deficits* articulatorios, no estadiamento de doenças oncológicas da orofaringe e no *follow-up* de cirurgia reconstrutiva da língua.

MÁDY & BEER (2007) demonstram a relevância clínica da utilização da RM em tempo-real durante a produção da fala, num estudo dos parâmetros articulatorios envolvidos na produção de quatro consoantes do Alemão, em oito pacientes com carcinoma da cavidade oral. Os resultados obtidos pelos autores apontam, por um lado, que existe uma ligação na avaliação pré e pós-operatória para as consoantes estudadas e, por outro lado, entre o grau de *deficit* articulatorio.

Assim, no conjunto de estudos revistos, a análise da dinâmica articulatória pode ser efetuada, basicamente, de três modos distintos: (1) usando técnicas de aquisição ultrarrápidas combinadas com estratégias rápidas de preenchimento do *espaço K*, (2) através da técnica *tagged-RM*<sup>5</sup>, ou ainda (3) utilizando métodos de sincronização das imagens de RM com o sinal acústico da fala. A primeira técnica é mais amplamente utilizada do ponto de vista clínico, ao passo que as duas restantes têm maior aplicabilidade do ponto de vista da Engenharia. Independentemente da sua aplicabilidade, estes estudos obrigam à repetição sucessiva de sons ou sequências articulatórias por parte dos sujeitos. A língua pela grande mobilidade e flexibilidade de movimentos e dada a sua importância na articulação da fala, é o órgão mais analisado nos estudos dinâmicos.

Um dos primeiros estudos dinâmico por RM foi conduzido por DI GIROLAMO e colaboradores (1996), e consistiu na determinação dos movimentos da laringe e orofaringe durante duas fases respiratórias: inspiração máxima e expiração, e na produção de vogais em voz normal e aguda. Usando um número muito restrito de imagens, estes investigadores concluíram que os cortes médio sagitais usando técnicas de aquisição rápidas por RM constituem a melhor

---

<sup>5</sup> Técnica não invasiva que consiste na aplicação de pulsos de radiofrequência de pré-saturação dos tecidos (sobreposição de linhas negras conhecidas por *tags*) e que é maioritariamente utilizada no estudo da função cardíaca, permitindo a obtenção de medidas quantitativas da deformação e movimento deste órgão.

modalidade de estudo das diferentes configurações do trato vocal podendo substituir inclusivamente as técnicas radiográficas.

Seguidamente, o estudo do fenómeno de coarticulação e compensação articulatória é efetuado por DEMOLIN e colaboradores (1997), através da recolha de imagens mediante a pronúncia lenta de uma sequência de palavras sem significado. Apesar da baixa resolução temporal dos sistemas de RM, os autores obtiveram 4 imagens por segundo, apontando como limitação ao estudo a impossibilidade de registo acústico durante a aquisição.

Nesse sentido, SHADLE e colaboradores (1999) utilizaram um método de pós-sincronização das imagens sagitais do trato vocal por RM recolhidas durante a fala. O registo acústico após a RM combinado com a sequência de imagens recolhidas permitiram observar as variações pseudotemporais da dinâmica dos articuladores.

A técnica de RM em tempo real é apresentada por DEMOLIN e colaboradores (2000, 2002) com o objetivo de estudar a coordenação dos movimentos articulatórios durante a produção de uma sequência de vogais do Francês. Apesar do elevado ruído acústico produzido pelo sistema de RM, estes autores efetuaram o registo acústico da fala, através do intercomunicador, ainda que a qualidade seja inferior, para localizar os sons na sequência de imagens obtidas. A frequência de amostragem ainda que baixa (4 a 5 *frames* por segundo) permitiu, contudo, a medição de distâncias médio sagitais e observar a posição dos articuladores envolvidos na produção da fala.

Entretanto, um estudo realizado por STONE e colaboradores (2001) referenciou a utilização da técnica *tagged-RM* para inferir a atividade muscular da língua a partir da deformação dos tecidos. Num sujeito repetindo cerca de 96 vezes uma sílaba, foi possível a utilização desta técnica e a determinação das tensões principais de deformação da língua.

Mais tarde, a medição das variações temporais do trato vocal durante várias repetições de uma sequência de vogais do Japonês, foi conduzida por

TAKEMOTO & HONDA (2003) através de um método de amostragem sincronizado<sup>6</sup> por RM (sincronismo acústico externo).

Novos incrementos na resolução temporal dos sistemas de RM, permitiram a NARAYANAN e colaboradores (2004), o estudo da dinâmica articulatória por RM com uma frequência de amostragem de 20 a 24 *frames* por segundo. A partir desta sequência de imagens, os investigadores efetuaram a análise quantitativa de distâncias entre vários pontos articulatorios ao longo de todo o trato vocal. No entanto, como limitações são ainda apontadas o elevado ruído acústico produzido durante as aquisições e a baixa qualidade das imagens. A este propósito, ENGWALL (2004) destaca que deve existir um compromisso entre a resolução temporal e espacial na aplicação dos estudos dinâmicos por RM. Recorrendo a sequências ultrarrápidas, este investigador analisou três conjuntos de palavras em contexto consoante-vogal-consoante com uma frequência de amostragem de 9 *frames* por segundo, realizando com razoável exatidão a modelação da língua.

AVILA-GARCÍA e colaboradores (2004) propuseram um método de extração automática da forma e dimensões do trato vocal durante a fala, através da aplicação do método de PCA nas imagens de RM em tempo-real. Mediante a combinação manual das imagens dinâmicas com o sinal da fala, sendo a gravação acústica previamente à aquisição por RM, os autores justificam a implementação deste método pelo facto de, as imagens de RM apresentarem muito ruído e porque a língua é um órgão altamente deformável podendo não ser facilmente representado através de um modelo paramétrico.

No seguimento do estudo anteriormente realizado, com o objetivo de estudar a compensação efetuada pela língua e laringe para neutralizar os movimentos da mandíbula, DEMOLIN e colaboradores (2006), conseguiram uma frequência de amostragem de 10 *frames* por segundo através da utilização de uma sequência rápida combinada com a transformada parcial de Fourier.

PARTHASARATHY e colaboradores (2007) utilizaram a técnica *tagged-RM* para medir os movimentos da língua e a sua deformação interna durante a fala,

---

<sup>6</sup> Do inglês *synchronized sampling method*, desenvolvido para o estudo de movimentos voluntários. Técnica que combina a sincronização da aquisição por RM através da monitorização por electrocardiograma com um *trigger* externo.



recorrendo ao método de *harmonic phase analysis*<sup>7</sup>. Este método de processamento consiste na detecção e seguimento automático das linhas negras (*tags*) a partir das imagens 2D permitindo estimar o movimento e deformação de qualquer *pixel* da imagem.

A maior frequência de amostragem observada na revisão da literatura foi alcançada por MOKHTARI e colaboradores (2007) a partir da técnica 3D de cine-RM. Estes investigadores testaram a aplicabilidade do método de PCA no estudo dos sons do Japonês, mediante a sua repetição pelo sujeito sincronizado com o ruído acústico do sistema de RM. Ainda que a resolução temporal desta técnica permitisse observar o “movimento” dos articuladores numa frequência de 35 *frames* por segundo, o sujeito teve que repetir sucessivamente cerca de 640 vezes os sons, sendo realizadas paragens para descanso. O que, do nosso ponto de vista, para além do grande esforço que exige aos sujeitos, aumenta a variabilidade acústica intrasujeitos. Acresce ainda que a gravação acústica da fala foi efetuada separadamente.

BRESCH & NARAYANAN (2009) descreveram um método para seguimento dos movimentos articulatorios em sequências de imagens de RM, tendo por base a utilização de um algoritmo automático baseado na utilização de um modelo deformável. Esta metodologia de análise permitiu a extração dos contornos do trato vocal com razoável robustez, demonstrando a “deformação” temporal destes contornos (forma) e posição dos articuladores.

Outra aplicabilidade mais recente, realizada por ECHTERNACK e colaboradores (2010), consistiu no estudo dinâmico por RM das configurações do trato vocal, em quatro jovens sopranos profissionais, para avaliação das transições em dois registos acústicos distintos: alto e baixo. O estudo foi realizado num sistema de 3.0 T, recorrendo a um microfone ótico para registo acústico da voz, obtendo um *slab* com 11 cm de espessura no plano sagital, com uma frequência de amostragem de 8 *frames* por segundo. Este conjunto de imagens permitiu realizar várias medições (distâncias e ângulos), como a abertura dos lábios, posição da mandíbula, a elevação do dorso da língua, a altura da faringe e a posição do véu do palato.

---

<sup>7</sup> É um método rápido e interativo para o processamento de imagens *tagged*-RM permitindo obter várias medidas funcionais, como linhas de fluxo, as tensões circunferenciais e radiais, mapas de velocidade, e as tensões principais.

### 4.1.3 Estudos do Português Europeu

Apesar dos vários estudos existentes sobre a produção da fala, a informação morfológica e dinâmica é ainda exígua, nomeadamente no que diz respeito aos sons do Português Europeu. O primeiro estudo do Português Europeu, com base na análise acústica e na EMA foi realizado por TEIXEIRA & VAZ (2001) culminando na construção de uma base de dados para caracterização dos sons nasais. Seguiram-se posteriormente outros estudos (TEIXEIRA et al., 2002, 2003) com a mesma linha de investigação.

A primeira utilização da RM, foi conduzida por RUA & FREITAS (2006b), resultando na caracterização morfológica de 25 sons do PE e na recolha de cine-imagens do trato vocal durante a fala em contexto articulatorio consoante-vogal (CV), para as sílabas /pa/, /ma/ e /tu/. O trabalho apresentado, para além de pioneiro, teve também grande receptividade da comunidade científica (VENTURA & FREITAS, 2007; VENTURA et al., 2008, 2009b).

MARTINS e colaboradores (2008) apresentaram uma aquisição ultrarrápida de RM, conseguindo dados volumétricos de um sujeito, para estudo das vogais orais e nasais, e das consoantes fricativas, laterais e oclusivas. As imagens bidimensionais obtidas pelos autores, permitiram o estudo comparativo entre as vogais orais e nasais, mediante a extração de contornos e, a posterior correlação com a base de dados de EMA desenvolvida anteriormente por TEIXEIRA & VAZ (2001). O estudo tridimensional realizado permitiu a determinação das funções de área de parte deste *corpus*, tendo-se mostrado particularmente útil na caracterização das consoantes palatais.

A aplicação da RM e potencialidades no estudo do PE foram expressadas por VENTURA e colaboradores (2009a), recorrendo à aquisição de imagens em diferentes orientações. Este estudo culminou na apresentação de uma base de dados de natureza morfológica do trato vocal em várias posições articulatorias e para um número considerável de sons do PE, constituindo os primeiros modelos 3D do trato vocal reconstruídos a partir de imagens de RM recolhidos em dois sujeitos jovens, voluntários e saudáveis. Este estudo demonstra o valor da técnica de RM não só em Medicina, contribuindo para o estudo de disfunções articulatorias e no planeamento terapêutico da fala, como em Engenharia Biomédica para o reconhecimento e síntese da fala.

O estudo das consoantes laterais, em diferentes contextos silábicos, apresentado por MARTINS e colaboradores (2010), descreve as características articulatórias mais comumente encontradas, apontando como principal limitação do estudo a dificuldade de registo espacial e segmentação dos modelos de gesso dos dentes.

VENTURA e colaboradores (2011c) apresentam a descrição de um estudo dinâmico por RM, realizado em dois sujeitos (do sexo masculino e feminino) usando um *pseudo-trigger* o registo eletrocardiográfico mediante sincronização de imagens cinemáticas de RM. Neste estudo, foram calculados quatro parâmetros articulatórios (usados nos sistemas de síntese da fala), demonstrando-se diferenças importantes na forma do trato vocal, comparativamente entre os dois sujeitos durante a produção dos mesmos sons.

Em 2011, VASCONCELOS e colaboradores mediante a utilização de modelos estatísticos deformáveis (modelos de distribuição pontual e modelos ativos) nas imagens de RM possibilitou a extração automática da forma do trato vocal, para diferentes sons do Português Europeu. Este estudo permitiu concluir que estes modelos permitem a extração de forma eficaz e automática das principais características do trato vocal durante os movimentos articulatórios (VASCONCELOS et al., 2011b).

## **4.2 Condicionantes da ressonância magnética na avaliação da articulação da fala**

Para além das contraindicações relacionadas com a instrumentação utilizada nestes sistemas, em particular o campo magnético e gradientes, existem algumas condicionantes que assumem grande relevância no estudo da produção da fala, nomeadamente:

- A posição de decúbito dorsal imposta aos sujeitos durante um exame de RM: resultando na alteração postural de vários órgãos (ENGWALL, 2006; KITAMURA et al., 2005);
- A colaboração e algum esforço dos pacientes exigidos: como se trata de uma aquisição durante uma atividade motora – a fala, de forma a

minimizar o cansaço dos mesmos, o número de sons a estudar, o designado *corpus* da fala, deve ser escolhido criteriosamente face à perturbação articulatória em causa;

- A não identificação dos dentes nas imagens adquiridas: pela pouca quantidade de hidrogénio que apresentam, os dentes surgem com similar sinal que as cavidades de ressonância do trato vocal, condicionando o rigor das medidas articulatórias;
- O elevado ruído acústico produzido pelos gradientes: prejudica o registo acústico durante o exame, inviabilizando a análise do sinal da fala e sua correlação com as sequências de imagens obtidas;
- Inexistência de dispositivos de sincronização ou *trigger* específicos para a fala.

Vários estudos têm sido dedicados ao registo acústico do sinal da fala durante a RM, de forma a permitir a correlação entre a forma do trato vocal e o sinal acústico produzido, nomeadamente através das técnicas de:

- Pós-sincronização, ou seja, fazendo a gravação da fala depois da aquisição de RM (AVILA-GARCÍA et al., 2004). Esta técnica é menos desejável pois a repetição de um som de fala nunca é idêntica às anteriores devido à variabilidade intrasujeitos;
- Através do intercomunicador apenas para confirmação do som produzido pois a qualidade do som produzida por este dispositivo é muito baixa (BEHRENDIS & WISMULLER, 2001; DEMOLIN et al., 2000; SHADLE et al., 1999);
- Simultaneamente à aquisição de RM: preferencialmente através de um microfone de fibra ótica (BRESCH et al., 2006; ECHTERNACK et al., 2010; NESSAIVER et al., 2006; PROCTOR et al., 2008) com recurso a técnicas de cancelamento do ruído (AALTO et al., 2011); a utilização de um microfone de fibra ótica tem a vantagem acrescida de poder paralelamente servir de *trigger* para o início das sequências dinâmicas. Alternativamente, esta sincronização pode ser efetuada com recurso a outros *pseudo-triggers* menos rigorosos, como através dos batimentos cardíacos (FOLDOVIK et al., 1990; MASAKI et al., 1999) ou do ruído acústico produzido pelo sistema de RM durante as aquisições (MATHIAK et al., 2000).

Relativamente à identificação dos dentes, alguns autores (a seguir destacados) têm proposto algumas soluções (NG et al., 2011; OLT & JAKOB, 2004; SERRURIER & BADIN, 2005; TUTTON & GODDARD, 2002); nomeadamente, através da extração indireta ou direta dos contornos dos dentes. A primeira técnica utilizada consistiu na sobreposição de imagens de tomografia computadorizada obtidas através de um molde ou marcas dentárias dos dentes dos sujeitos. De forma indireta, o espaço dentário podia ser extraído e, posteriormente adicionado nas imagens de RM. O uso de moldes ou placas dentárias causam a modificação da forma dos dentes, não conferindo assim uma mensuração rigorosa e exata do trato vocal.

A extração direta dos contornos dos dentes, a partir das imagens de RM, pode ser efetuada com recurso a líquidos para preencher a cavidade oral ou a outras técnicas de processamento de imagem (técnicas de subtração ou adição), usando estruturas ósseas para referência e registo espacial da imagem dos dentes.

TAKEMOTO e colaboradores (2004) propuseram um novo método para sobreposição das imagens 3D dos dentes diretamente nas imagens do trato vocal por RM, de forma a permitir uma correta determinação do seu volume. A extração do volume dos dentes, da maxila e da mandíbula, foi obtida através de uma aquisição inicial de RM, com o paciente em decúbito ventral e com a cavidade oral preenchida com sumo de mirtilo, que é uma substância considerada como meio de contraste oral. Após reconstruído o volume dentário, as imagens volumétricas eram realinhadas e incluídas no modelo 3D do trato vocal. Contudo, neste método, para além do desconforto imposto ao paciente pela mudança de posição e sustentação do líquido na boca, podem ocorrer artefactos de movimento da cabeça que dificultam o realinhamento dos dados obtidos. Acresce ainda, que este realinhamento requer perícia e conhecimento por parte do utilizador.

Também OLT & JAKOB (2004) seguiram a mesma linha de investigação, mas recorrendo à utilização de água e outro meio de contraste oral, com o objetivo de avaliar a viabilidade desta técnica para visualização dos dentes e mandíbula em RM. Segundo estes investigadores, de uma forma indireta e não invasiva, o preenchimento da cavidade oral com substâncias biocompatíveis permite a observação destas estruturas nas imagens de RM, para além de minimizar os

efeitos de susceptibilidade magnética das imagens na fronteira de ar-tecido mineralizado (dentes). Acresce que, a resolução espacial da técnica de RM é comparável à obtida pela técnica de tomografia computadorizada de *Dental Scan*, podendo ter um valor acrescentado em Ortodontia e Odontologia dado o seu carácter não ionizante.

Mais recentemente, NG e colaboradores (2011) propuseram a utilização de um molde gelatinoso no sentido de identificar com precisão apenas os dentes incisivos numa imagem de RM. Para garantir a correta colocação da imagem do molde dos dentes nas imagens alvo de RM do trato vocal, os autores colocaram quatro tubos de plástico preenchidos com citrato de ferro e amónio ao longo da linha média da face do sujeito em estruturas ósseas estáveis. O estudo realizado com e sem sobreposição do molde dos dentes durante a articulação de duas consoantes dentolinguais demonstra que o contorno dos dentes pode ser facilmente extraído. Segundo estes investigadores, a utilização desta técnica em trabalhos futuros poderá permitir avaliar a relação temporal e espacial entre os articuladores em pacientes com mal-oclusão dentária.

### 4.3 Conclusão

O interesse pelo conhecimento da produção da fala surge paralelamente com o desenvolvimento de novas modalidades imagiológicas, como a ressonância magnética, em finais do século XX. A observação estática dos articuladores durante a sustentação muito prolongada dos sons rapidamente foi aperfeiçoada, permitindo não só uma maior extensão de sons estudados, mas também o alargamento a várias línguas desde o Francês, Inglês, Alemão, Japonês, Sueco, entre outras.

Impôs-se também a necessidade de incrementar o rigor desta avaliação a partir de imagens bidimensionais e mais tarde, tridimensionais, quer para caracterização morfológica como dinâmica da produção da fala.

Várias técnicas de análise quantitativas foram propostas para determinação das funções de área, como para seguimento dos movimentos articulatorios ou ainda, para a construção de modelos tridimensionais.

Por oposição aos estudos da área da Engenharia com amostras (número de sujeitos) muito reduzidas em detrimento de um *corpus* da fala mais vasto, os estudos na vertente clínica são menos frequentes e assentam quase exclusivamente na caracterização funcional da fala.

É notório pela revisão da literatura, o interesse que os estudos da produção da fala despertam em várias áreas multidisciplinares, bem como pela explicação de outros fenómenos como a nasalidade, a coarticulação. De realçar também, a evolução e aperfeiçoamento da técnica de RM e o desenvolvimento de diferentes abordagens na procura de um melhor entendimento da fala.

A análise e processamento das imagens carecem de *softwares* específicos que permitam responder às necessidades dos estudos da fala; daí que a maioria dos autores tenha optado pelo método manual para as tarefas de segmentação (extração de contornos do trato vocal). A técnica de segmentação depende da qualidade das imagens de RM, do *corpus* e finalidade do estudo. A utilização de técnicas mais automatizadas tem vindo a crescer e com resultados promissores, nomeadamente nos estudos em tempo-real.

Os estudos acerca da captação e registo acústico da fala em ambiente de RM têm vindo a aumentar, ainda que continuem a subsistir algumas condicionantes para a sua implementação rotineira, nomeadamente os custos relacionados com os recursos materiais para a conceção e montagem dos sistemas (por exemplo, microfones de fibra ótica) e os custos operacionais.

O desenvolvimento de técnicas para visualização dos dentes nas imagens de RM é também uma das linhas de investigação cada vez mais visível, no sentido de aumentar o rigor das medidas volumétricas do trato vocal extraídas, para além de abrir novos horizontes de estudo para outras áreas como a Ortodontia e a Odontologia.

Em suma, apesar das condicionantes que ainda se verificam, a imagem por RM é, sem dúvida, uma técnica promissora no estudo da fala, inócua, não-invasiva, fornecendo informação fiável e com utilidade não só em Engenharia e em Medicina, como em muitas outras áreas.





## **CAPÍTULO 5. PROTOCOLO DE ESTUDO DO TRATO VOCAL POR RESSONÂNCIA MAGNÉTICA**

Em ciência da computação, um protocolo significa convenção ou padrão que controla e possibilita uma conexão, comunicação ou transferência. Também é um termo que designa uma linguagem universal passível de poder ser utilizada ou aplicada independentemente do objeto em estudo. A utilização de *hardware* e *software* em Imagiologia médica, pelas suas especificidades, requer a criação de linhas gerais de orientação que permitam garantir e reproduzir por um lado, a aquisição de imagens com qualidade diagnóstica e, por outro, a prestação de cuidados de saúde adequados à situação clínica e idade dos pacientes. Os protocolos de estudo em Imagiologia incluem linhas de orientação e referência na avaliação de várias patologias e visam, para além da melhoria da capacidade diagnóstica, um maior benefício das vantagens da modalidade de imagem em causa, como a redução do tempo de exame, uma melhor relação sinal-ruído, entre outras.

Assim, para a instrumentação, usualmente designada por arte ou modo de combinar as partes, de um protocolo de Ressonância Magnética para estudo do trato vocal, é imperativo entender: Que parâmetros fundamentais ou partes são necessários para o desenho deste protocolo face à finalidade do estudo? Que passos são necessários para a sua implementação e otimização?

Este capítulo apresenta o conjunto de etapas para a elaboração de um protocolo de imagem por RM para a avaliação da articulação da fala, enfatizando os aspetos técnicos da Imagiologia e benefícios da sua utilização.

## 5.1 Desenho do protocolo

Para a definição do protocolo de imagem por RM para estudo do trato vocal, atendeu-se, por um lado, à informação recolhida através da aplicação de questionários e, por outro, aos estudos existentes sobre este tópico que foram revistos no capítulo anterior. Esta etapa do trabalho de investigação, referente ao desenho do protocolo, foi realizada previamente ao trabalho de campo, no período compreendido entre Outubro de 2009 a Abril de 2010.

Os questionários realizados visaram a recolha de informação dos profissionais-chave para futura utilização das imagens por RM no estudo do trato vocal. A estrutura destes foi organizada de forma a identificar qual a informação que os clínicos procuram e privilegiam, assim como qual a aplicabilidade *versus* condicionantes do ponto de vista da Engenharia (Anexo 1). Deste modo, teve-se em consideração na elaboração de um protocolo de imagem por RM para estudo do trato vocal a caracterização dos seguintes parâmetros fundamentais:

- Indicações clínicas;
- Preparação do paciente/ instruções especiais;
- Antena e posicionamento do paciente;
- Planos de imagem e sequências;
- Parâmetros de aquisição;
- Resultados esperados.

Segundo o Manual de Boas Práticas da Especialidade de Radiologia aprovado pelo Despacho n.º 258/2003 da II série, no que respeita às normas de utilização da Ressonância Magnética, “não é missão destas especificações cobrir exaustivamente todo o leque de potenciais aplicações da técnica ou obrigar à realização de protocolos rígidos de exploração imagiológica os quais, pelo contrário, deverão ser variáveis e orientados segundo a indicação clínica e as capacidades técnicas incorporadas em cada aparelho”.

A determinação de um protocolo de RM para avaliação articulatória do trato vocal justifica-se pois:

- a. Até à data, não existem linhas gerais previamente definidas ou validadas;

- b. Existe uma carência de informação morfológico-dinâmica do trato vocal durante a produção da fala;
- c. As limitações e efeitos biológicos da técnica de cineradiografia não permitem uma caracterização morfológica suficiente dos articuladores e eventos articulatórios.

Portanto, do ponto de vista clínico, é difícil fazer uma avaliação do mecanismo de articulação da fala em pessoas com perturbações articulatórias. Como potenciais indicações clínicas para a prescrição deste exame imagiológico, podem ser destacadas:

- Avaliação pré e pós cirúrgica de pacientes oncológicos, como, por exemplo, tumores da cavidade oral e da faringe;
- Avaliação de pacientes laringectomizados;
- Estudo de pacientes com alterações estruturais como fendas labiopalatinas, hipertrofia das amígdalas, incompetência velofaríngea ou alterações craniofaciais;
- Estudo da cavidade faríngea no diagnóstico de SAOS;
- Estudo de pacientes com perturbações funcionais não associadas a alteração estrutural, como perturbações auditivas;
- Estudos da fonação, para obter a relação entre a geometria do trato vocal e a fala para posterior cálculo das funções de área, para estudos de acústica, entre outros.

Quer as crianças, nomeadamente, no caso de perturbações articulatórias por alteração estrutural, como os adultos podem ser grupos de interesse para recurso a este protocolo.

A preparação prévia do paciente para a realização de qualquer exame imagiológico visa não só obter a colaboração necessária à execução do mesmo, como também instruir e informar acerca dos procedimentos e dinâmica do exame. Para o efeito, elaborou-se um guia básico (Anexo 2) para os pacientes ou acompanhantes, em particular de crianças, que reúne o conjunto de instruções e informações pretendidas para a realização deste estudo. Este guia foi utilizado durante o trabalho de campo deste projeto e descreve numa linguagem simples e clara quais os procedimentos exigidos antes e durante o estudo por RM, assim

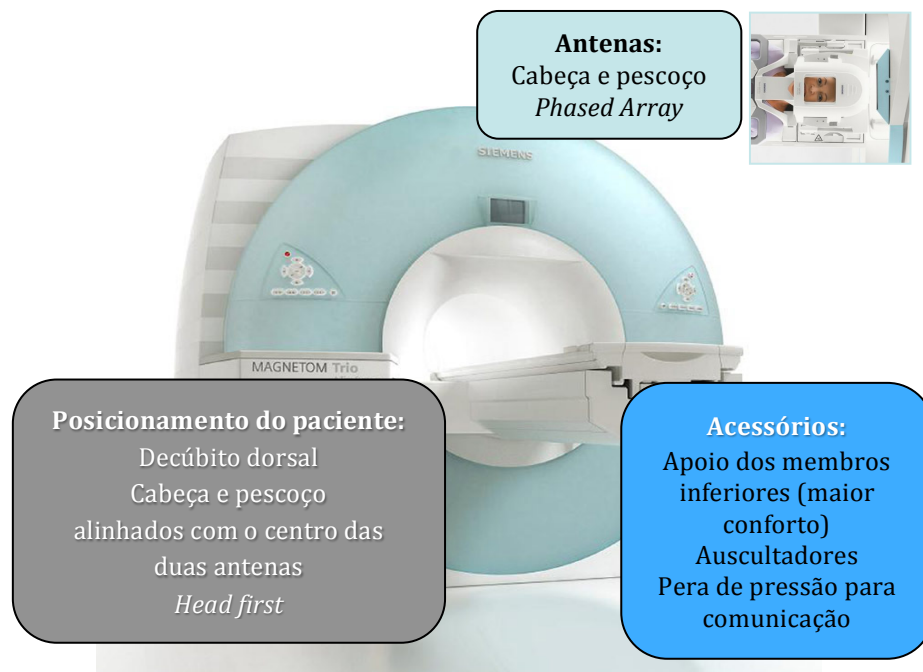
como qual a colaboração pretendida. A preparação prévia inclui essencialmente instruções quanto à preparação geral do indivíduo indicando quais os objetos a remover, por exemplo todos os objetos e acessórios metálicos, garantindo desta forma a total privacidade, o máximo conforto e a total segurança do sujeito.

Durante o estudo, para além da necessidade de imobilização são também fornecidas informações quanto às particularidades de um estudo por RM, em particular, quanto ao elevado ruído produzido e quanto à possibilidade de comunicação por um intercomunicador e a interrupção do estudo a pedido do paciente. Quanto à colaboração pretendida, é solicitado aos sujeitos, via intercomunicador, que pronunciem e sustentem determinado som, usando uma palavra como referência, durante toda a aquisição.

Num estudo por RM, acresce ainda a necessidade de excluir qualquer contraindicação à exposição a campos magnéticos, mediante a realização de um questionário nomeadamente, acautelar a presença de dispositivos ferromagnéticos e/ou elétricos, como *pacemakers*, excluir uma eventual gravidez ou historial de outras doenças e ainda a claustrofobia. Assim, antes de qualquer ensaio aplicou-se o questionário de segurança em RM existente na instituição hospitalar, permitindo averiguar eventuais contraindicações e descartar eventuais doenças ou situações clínicas que pudessem condicionar e/ou impedir a realização do estudo de RM. Adicionalmente, implementou-se um questionário de segurança (Anexo 3) com base nas linhas de orientação para práticas de segurança em RM definidas pelo *American College of Radiology*, para servir de documento de suporte ao trabalho de campo e ainda para excluir a presença de materiais metálicos e/ou ferromagnéticos na região do trato vocal ou próximo, que pudessem interferir no estudo, tais como *piercings* e aparelhos dentários.

Acresce ainda, a instrução para que os pacientes evitem qualquer movimento do corpo (principalmente a cabeça) e que mantenham uma respiração calma, durante o exame imagiológico.

De uma forma geral, independentemente das especificidades técnicas de um sistema de RM, o posicionamento e os dispositivos básicos e acessórios utilizados para estudo do trato vocal, são idênticos aos utilizados em qualquer estudo por RM da cabeça e pescoço, Figura 5.1.



**Figura 5.1** Posicionamento e dispositivos básicos para estudos de RM da cabeça e pescoço. Imagem de fundo: sistema de RM Siemens Magnetom® Trio 3.0 T.

Os parâmetros de aquisição disponíveis para a escolha de uma sequência de pulsos são variados e definidos pelo operador. A combinação destes parâmetros determina o tipo de ponderação, em particular o contraste e a qualidade de imagem devendo respeitar a indicação clínica e condição física do paciente.

No caso particular do trato vocal, a programação do estudo deve ainda atender aos seguintes fatores:

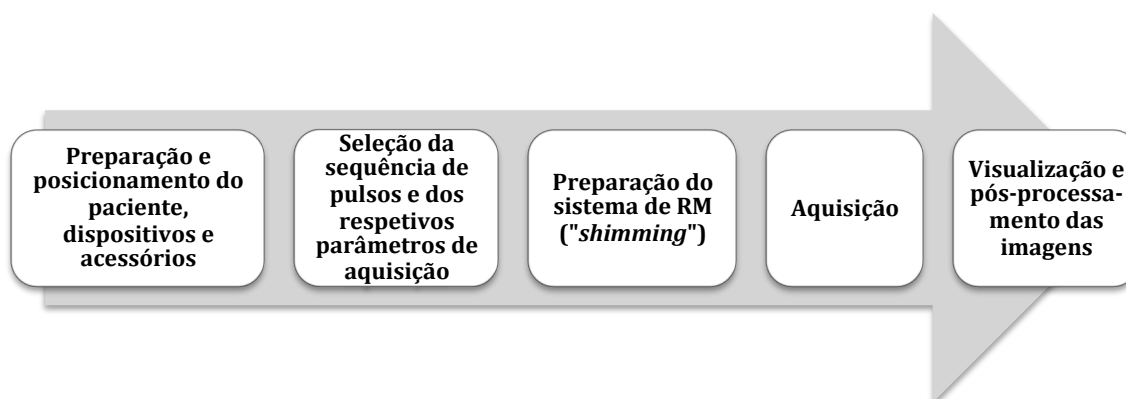
- Características anatómicas e funcionais do trato vocal: descritas anteriormente na secção 3.4.3;
- Tipo de estudo e finalidade:  
Independentemente do fim a que se destina, a caracterização morfológico-dinâmica do trato vocal pode ser obtida através de três formas: 1) de um estudo estático, mediante a sustentação prolongada de um som; 2) de um estudo dinâmico, mediante sincronização e repetição sucessiva de um som; e 3) de um estudo em tempo-real, através da captação direta de imagens em movimento. O primeiro permite obter essencialmente uma caracterização morfológica, ou seja

das posturas articulatórias do trato vocal, com interesse para posterior reconstrução 3D. A morfologia do trato vocal é caracterizada pela forma e posição dos articuladores, pelas cavidades de ressonância e pela altura da laringe. Os estudos dinâmicos e em tempo-real demonstram essencialmente os eventos dinâmicos articulatórios, permitindo também avaliar a estabilidade do fenómeno em contexto-real. A implementação de um protocolo de RM para avaliação da articulação da fala pode destinar-se a fins clínicos, como para avaliação funcional de pacientes com perturbações articulatórias ou, por outro lado, a fins acústicos, por exemplo, para modelação do trato vocal e posterior síntese da fala.

- Condicionantes relacionadas com o sistema de RM:

Deve ter-se em conta que a posição de decúbito dorsal imposta aos pacientes durante um exame de RM, tem influência no mecanismo da produção da fala resultando na alteração postural de vários órgãos (ENGWALL, 2003; KITAMURA et al., 2005). Para além disso, o estudo por RM durante a produção da fala exige a colaboração e algum esforço dos pacientes, pelo que de forma a minimizar o cansaço dos mesmos, os sons a estudar, que constituem o *corpus* da fala, devem ser escolhidos criteriosamente face à perturbação articulatória em causa.

O fluxo de trabalho adotado para a instrumentação de um protocolo de imagem por RM está representado na Figura 5.2.



**Figura 5.2** Conjunto de etapas básicas para a instrumentação de um protocolo de imagem por RM.

Quanto ao tipo de estudo, as sequências 2D de corte único ou múltiplos cortes são amplamente usadas nos estudos de produção da fala, pois permitem observar e medir a posição dos diferentes articuladores. Para além disso, servem de referência pois o tempo de aquisição é menor do que para estudos 3D. No entanto, do ponto de vista da utilização das técnicas de processamento e análise, as imagens 2D são menos atrativas pois a morfometria extraída representa uma estimativa da anatomia real do trato vocal.

Em contrapartida, o estudo 3D do ponto de vista clínico é de menor relevância devido à dificuldade de extração de medidas quantitativas e de referenciação espacial, ainda que permita um estudo volumétrico mais preciso do trato vocal, através da determinação do espaço articulatorio e também a análise das cavidades de ressonância. A construção de modelos 3D do trato vocal é, por isso, mais realista, permitindo a obtenção de medidas quantitativas mais rigorosas, como cálculo das funções de área, ainda que os tempos de aquisição sejam mais longos.

## **5.2 Implementação e otimização do protocolo**

O trabalho de campo com vista à implementação do protocolo foi realizado na Unidade de Ressonância Magnética do Hospital São João EPE, no Porto, durante os meses de Maio a Dezembro de 2010 numa primeira fase, conforme planeamento prévio e respetiva autorização, mediante a disponibilidade do serviço. Seguiu-se a fase de otimização do protocolo de RM, no período de Março a Outubro de 2011.

As recolhas imagiológicas foram organizadas de forma sistemática de acordo com a seguinte metodologia:

- 1º Implementação: Programação das sequências com o fantoma existente: fantoma esférico preenchido com óleo e com 16.7 cm de diâmetro;
- 2º Otimização: Realização de testes e ensaios com um voluntário para ajuste do protocolo;
- 3º Reavaliação dos resultados.

### 5.2.1 Estudo estático 2D e 3D

As Tabelas 5.1, 5.2 e 5.4 resumem as sequências de pulso, os planos de imagem e respectivos parâmetros de aquisição programadas para os estudos estáticos 2D e 3D.

Quanto ao estudo estático 2D, os vários ensaios realizados foram no sentido de minimizar ao máximo o tempo de aquisição e conseqüentemente de sustentação do som. De acordo com ENGWALL (2000b), num estudo comparativo entre o estudo estático por RM e as técnicas de articulografia eletromagnética e eletropalatografia, ambas as técnicas permitem o estudo da coarticulação e são representativas do mecanismo dinâmico da fala. No entanto, a informação proveniente da RM, durante a sustentação prolongada de um som, representa um fenômeno de hiperarticulação e que é naturalmente produzido pelo indivíduo com o objetivo de reproduzir claramente e tanto quanto possível cada articulação.

Na Tabela 5.1 estão descritos os principais parâmetros de RM ensaiados para o estudo 2D estático do trato vocal.

**Tabela 5.1 Comparação dos diferentes ensaios e tipos de estudo 2D realizados e respectivo efeito no tempo de aquisição versus resolução das imagens.**

Ensaio/ data	1/ 29.05.2010	2/ 26.06.2010 5/ 12.03.2011	3/ 18.09.2010 7/27.07.2011
<b>Aquisição</b>	2D		
<b>Plano de corte</b>	Sagital		
<b>Sequências de pulso</b>	<i>Turbo spin eco</i>		
<b>Tempo de repetição (TR)</b>	400	400	250
<b>Tempo de eco (TE)</b>	10	10	9.6
<b><i>Flip Angle (FA)</i></b>	160°		
<b>Número total de cortes</b>	1		
<b>Espessura de corte (mm)</b>	4	3	3
<b>Espaçamento (GAP)</b>	---		
<b>Fator turbo (ETL)</b>	5		
<b>FOV (mm)</b>	240*240		
<b>Matriz</b>	512*512	512*512	256*256
<b><i>Percent sampling (p)</i> (percentagem de amostras usadas na codificação de fase)</b>	70% (358p*512)	65% (332p*512)	50% (128p*256)



**Tabela 5.1 Comparação dos diferentes ensaios e tipos de estudo 2D realizados e respetivo efeito no tempo de aquisição versus resolução das imagens – continuação.**

Ensaio/ data	1/ 29.05.2010	2/ 26.06.2010 5/ 12.03.2011	3/ 18.09.2010 7/27.07.2011
<b>Aquisição paralela (Syngo GRAPPA)</b>			
<b>Fator de aceleração</b>	2	2	4
<b>Tempo de aquisição</b>	8.07 s	8.07 s	7.04 s
<b>Resolução (px/mm)</b>	2.133	2.133	1.067
<b>Pixel size (mm)</b>	0.469 x 0.469	0.469 x 0.469	0.937 x 0.937
<b>Display Range</b>	0 - 1393	0 - 1726	0 - 1527

Assim, e dado que este tipo de estudo serve apenas de referência para o estudo 3D, apenas um corte médio sagital é suficiente para obter uma informação mais completa acerca da morfologia do trato vocal para cada som. Pela análise da Tabela 5.1, nomeadamente nos ensaios 2 e 3 observa-se que: a espessura de corte de 3 mm aliada a uma matriz de imagem grande proporciona resultados muito superiores quanto à resolução e tamanho do *pixel*, num tempo de aquisição aceitável e passível de permitir a sustentação de um som com mínimo esforço.

Considerando os estudos revistos na literatura (ver Capítulo 4), torna-se difícil uma comparação efetiva pois alguns parâmetros de aquisição como a matriz, o FOV e a espessura de corte, que influenciam drasticamente os tempos de aquisição, são distintos. No entanto, nos estudos 2D estáticos revistos verificaram-se tempos de aquisição mais longos, como 21 segundos (KRÖGER et al., 2000), 35 segundos (BADIN & SERRURIER, 2006), 36 segundos (PROCTOR et al., 2008) e 45 segundos (SERRURIER & BADIN, 2005). Apenas o estudo realizado por MARTINS e colaboradores (2008), permite uma comparação mais aproximada, pois os autores obtiveram um tempo de aquisição de 5.6 segundos; no entanto, com parâmetros diferentes de espessura de corte (5 mm) e um tamanho do *pixel* de 0.78 x 0.78 mm.

Quanto ao estudo estático 3D, os vários ensaios realizados (Tabela 5.2) assentaram essencialmente em duas técnicas de obtenção de imagens distintas, em volume total (KIM et al., 2009; MARTINS et al., 2008) e em planos (BADIN et al., 1998; BADIN et al., 2002; BADIN & SERRURIER, 2006; DEMOLIN et al., 1996; ENGWALL, 2003; SERRURIER & BADIN, 2005).

**Tabela 5.2. Comparação dos diferentes ensaios e tipos de estudo 3D realizados e respetivo efeito no tempo de aquisição versus resolução das imagens.**

Ensaio/ data	2/ 26.06.2010	2/ 26.06.2010	3/ 18.09.2010
<b>Aquisição</b>	3D		
<b>Plano de corte</b>	Sagital	Coronal + coronal-oblívquo 45° + Axial	Sagital
<b>Sequências de pulso</b>	<i>Flash</i> gradiente-eco		
<b>Tempo de repetição (TR)</b>	10	12	6.15
<b>Tempo de eco (TE)</b>	2.17	2.17	2.45
<b>Flip Angle (FA)</b>	5°		
<b>Técnica de obtenção das imagens</b>	Volume total <i>Slab</i> com 7.8 cm	Planos <b>3 Stacks</b>	Volume total <i>Slab</i> com 7.8 cm
<b>Número total de cortes</b>	<b>26</b>	<b>54 (18+18+18)</b>	<b>26</b>
<b>Espessura de corte (mm)</b>	3	4	3
<b>Espaçamento (GAP)</b>	---		
<b>Fator turbo (ETL)</b>	1		
<b>FOV (mm)</b>	240*240	240*240	240*240
<b>Matriz</b>	256*256	256*256	512*512
<b>Percent sampling (p)</b>	70% (179p*256)	60% (154p*256)	70% (358p*512)
<b>Fator de aceleração</b>	2	4	4
<b>Tempo de aquisição</b>	13.15 s	15.82 s	13.23 s
<b>Resolução (px/mm)</b>	1.067	1.067	2.133
<b>Pixel size (mm)</b>	0.937 x 0.937	0.937 x 0.937	0.468 x 0.468
<b>Display Range</b>	0 - 406	0 - 406	0 - 1000

A obtenção de imagens pela técnica de planos permite obter uma imagem bidimensional a partir de múltiplas projeções da secção em estudo, no mesmo plano. Esta é a técnica, do ponto de vista clínico, mais usada atualmente e utiliza tempos de aquisição entre 2 a 10 minutos. Múltiplos cortes podem ser adquiridos num só plano com a mesma orientação formando uma *stack*, isto é, um conjunto de cortes 2D que “empilhados” representam um volume 3D.

A aquisição *multi-stacks* consiste na programação de duas ou mais *stacks* numa ou em várias orientações de acordo com a morfologia da região em estudo. Do ponto de vista da prática clínica, esta técnica de aquisição é comumente

usada, por exemplo, na avaliação de patologia discal da coluna vertebral, dada a orientação anatômica dos discos intervertebrais.

Na técnica de obtenção de imagens em volume total, o sinal é obtido a partir da ativação de todo o volume, usualmente denominado de *slab*, podendo-se reconstruir, posteriormente, a partir deste qualquer corte bidimensional, em qualquer plano anatômico através da divisão do volume em partições sem *GAP*. No entanto, exige uma elevada capacidade computacional e acarreta um maior tempo de aquisição dos dados. Dependendo do plano anatômico escolhido, define-se a espessura do *slab* de forma a abranger toda a região anatômica e o número de cortes *versus* espessura de cada corte efetiva a reconstruir posteriormente à aquisição. Esta técnica de aquisição, do ponto de vista clínico, é amplamente usada no estudo de estruturas vasculares dada a sua extensão e morfologia.

A Figura 5.3 ilustra uma representação esquemática da programação destas duas técnicas de aquisição de imagens por RM para estudo do trato vocal.

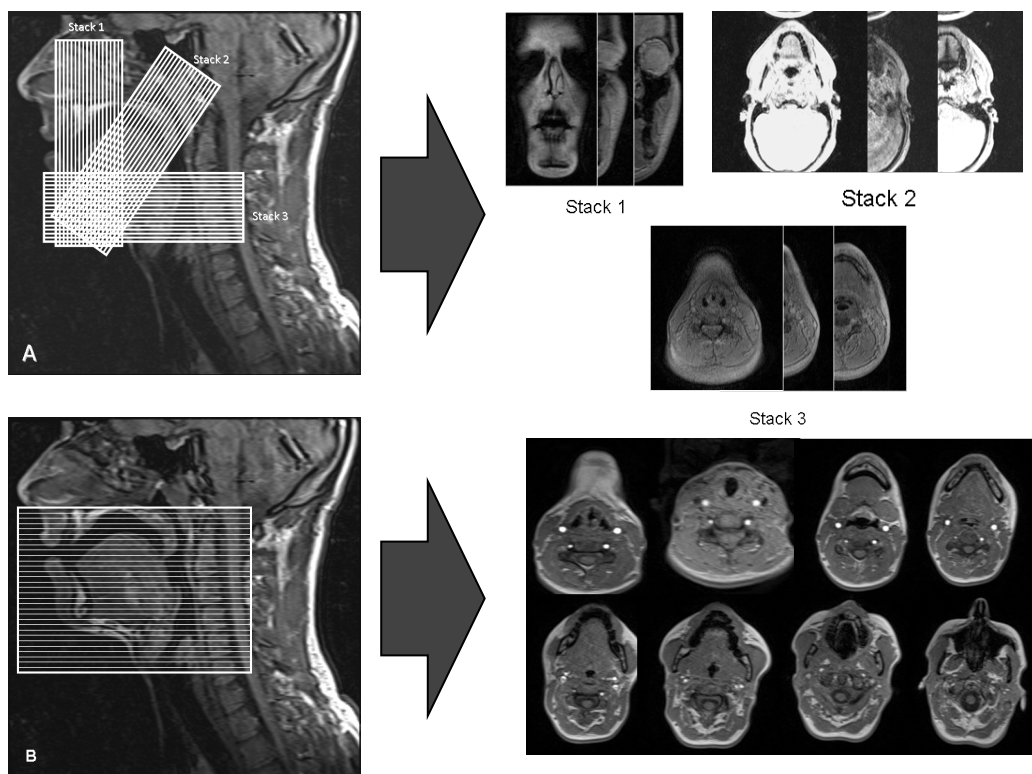


Figura 5.3 Representação esquemática para programação das técnicas de aquisição de imagens usando *multi-stacks* (A) e um *slab* (B).

Numa análise comparativa entre estas duas técnicas, a técnica de volume total apresenta globalmente melhores vantagens do que a técnica de *multi-stacks*, como se pode verificar pela Tabela 5.3, o que contribui *per se* para uma melhor modelação do trato vocal (ver secção 6.2). No entanto, a técnica de *multi-stacks* permite garantir a aquisição de cortes perpendiculares ao longo de todo o trato vocal sendo adequada para a sua geometria e conseqüentemente, um cálculo mais confiável das funções de área (APOSTOL et al., 1999; DEMOLIN et al., 1996, 1998).

**Tabela 5.3 Vantagens e limitações das várias técnicas de obtenção de imagens por RM: 3D *multi-stacks* e 3D *Slab*.**

Tipo de aquisição	Vantagens	Limitações
3D <i>multi-stacks</i>	<p>Possibilidade de orientar os cortes perpendicularmente a todo o trato vocal;</p> <p>Uma <i>stack</i> permite obter tempos de aquisição mais curtos comparativamente a um <i>slab</i>.</p>	<p>Cortes oblíquos difíceis de interpretar e altamente susceptíveis a artefactos (<i>aliasing</i> e <i>crosstalk</i><sup>8</sup>);</p> <p>Problemas de alinhamento e registo espacial entre <i>stacks</i>;</p> <p>Espessura de corte condicionada pela amplitude dos gradientes do sistema de RM; cortes finos exigem amplitudes muito grandes.</p>
3D <i>Slab</i>	<p>Possibilidade de divisão do volume sem <i>GAP</i> e com espessuras de corte mais finas;</p> <p>Obtém-se cortes contíguos sem artefactos;</p> <p>Todos os cortes estão espacialmente referenciados segundo o mesmo eixo;</p> <p>Melhor relação sinal-ruído e resolução espacial;</p> <p><i>Voxel</i> isotrópico possibilitando a reformatação em planos oblíquos sem perda significativa de resolução no pós-processamento;</p> <p>Pico de sinal (<i>display range</i>) maior.</p>	<p>Tempos de aquisição maiores (1 <i>stack</i> versus 1 <i>slab</i>);</p> <p>Maior susceptibilidade a artefactos de movimento;</p> <p>Todas as partições do <i>slab</i> estão orientadas segundo o mesmo plano.</p>

<sup>8</sup> Causado pela sobreposição de ondas de *RF* aquando da seleção do corte. Podendo ser solucionado através da seleção intercalada de cortes, pelo uso de espaçamento entre cortes, ou usando uma forma dos pulsos de *RF* quadrada.

Dadas as potencialidades da técnica de aquisição de volume total, o trabalho de campo incidiu principalmente na otimização dos parâmetros de modo a obter um compromisso ideal entre o tempo de aquisição, a relação sinal-ruído e resolução espacial, Tabela 5.4.

**Tabela 5.4 Comparação dos diferentes ensaios e tipos de estudo 3D realizados e respetivo efeito no tempo de aquisição *versus* resolução das imagens.**

<b>Ensaio/ data</b>	<b>4/ 13.11.2010</b>	<b>5/ 12.03.2011 e 7/27.07.2011</b>	
<b>Aquisição</b>	3D	3D	
<b>Plano de corte</b>	Axial	Coronal + coronal-oblívico a 30° + Axial	Axial
<b>Sequências de pulso</b>	<i>Flash</i> gradiente-eco		
<b>Tempo de repetição (TR)</b>	5.11	5.8	
<b>Tempo de <i>eco</i> (TE)</b>	2.45	2.17	
<b><i>Flip Angle</i> (FA)</b>	10°	10°	
<b>Técnica de obtenção das imagens</b>	Volume total <i>Slab</i> com 11.2 cm	Planos <i>3 Stacks</i>	Volume total <i>Slab</i> com 11.8 cm
<b>Número total de cortes</b>	56	66 (22+22+22)	60
<b>Espessura de corte (mm)</b>	2	4	2
<b>Espaçamento (GAP)</b>	---		
<b><i>Fator turbo</i> (ETL)</b>	1		
<b><i>FOV</i> (mm)</b>	270*270		
<b>Matriz</b>	280*512	256*256	256*256
<b><i>Percent sampling</i> (p)</b>	84% (227p*270)	96% (246p*256)	96% (246p*256)
<b>Fator de aceleração</b>	4		
<b>Tempo de aquisição</b>	15.69 s	5.78 s * 3	16.12 s
<b>Resolução (px/mm)</b>	1.896	0.948	
<b><i>Pixel size</i> (mm)</b>	0.527 x 0.527	1.055 x 1.055	
<b><i>Display Range</i></b>	0-1500	0-367 SNR = 1.39	0-367 SNR = 1.85

Através dos ensaios 4, 5 e 7, foi possível constatar que atualmente a técnica volumétrica 3D permite resultados comparáveis à técnica *multi-stacks* no que respeita ao tempo de aquisição *versus* número total de cortes. No entanto, o incremento da SNR e pelo facto de todos os cortes se encontrarem

“naturalmente” registados, esta técnica pode ser implementada rotineiramente no estudo do trato vocal, tanto para fins clínicos como da Engenharia.

#### *a) Alinhamento e registo espacial entre stacks*

A integração de imagens em múltiplas orientações é geralmente um processo cognitivo subjetivo, de análise diagnóstica realizado na prática clínica, dado o grande volume de imagens geradas para observação da morfologia e dinâmica dos órgãos. Ou seja, a interpretação e avaliação das imagens médicas, nomeadamente em RM, requer a análise completa de toda a informação 2D possibilitando ao clínico depois a visualização mental 3D do órgão/ patologia.

Assim, e considerando a geometria do trato vocal, poderá justificar a utilização da técnica de *multi-stacks* e conseqüentemente, a resolução dos problemas de alinhamento e registo espacial entre *stacks* com diferentes orientações.

O único método encontrado na literatura foi proposto por PROCTOR e colaboradores (2008) e consiste em importar no Matlab® os contornos resultantes da segmentação do trato vocal de cada um dos *stacks*, como um conjunto de coordenadas (x, y, z) e, posteriormente reconstruir como nuvens de pontos que representam as superfícies do tecido nas três orientações distintas. O realinhamento envolve a rotação e translação automática por coincidência dos centros de gravidade, sendo posteriormente ajustados manualmente. Esta metodologia foi especificamente desenvolvida para possibilitar a incorporação dos planos coronais-oblíquos na modelação trato vocal.

#### 5.2.2 Estudo dinâmico 2D

A Tabela 5.5 resume o conjunto de ensaios realizados para a implementação e otimização do estudo dinâmico (ou funcional) da fala, com recurso numa primeira tentativa à técnica *Grid Tagging* e posteriormente, a técnicas de aquisição ultrarrápidas. A primeira tentativa ou ensaio, em aplicar a técnica avançada *Grid Tagging* no estudo dos movimentos articulatorios não foi muito profícua, até porque na prática clínica é usada apenas em estudos cardíacos. A complexidade que apresenta constitui a principal limitação na sua utilização

clínica, mesmo no estudo do coração onde é mais usada, de forma a obter uma avaliação qualitativa da deformação do miocárdio nas diferentes fases do ciclo cardíaco. A análise quantitativa é possível, mas muito morosa e complexa, inviabilizando a sua utilidade clínica.

O recurso a sequências ultrarrápidas e estratégias de redução do tempo de aquisição em RM é, para além de mais rotineiro, mais fácil de implementar e utilizar.

**Tabela 5.5 Comparação dos diferentes ensaios e tipos de estudo realizados para a implementação e otimização do estudo dinâmico durante a produção da fala.**

Ensaio/ data	4/ 13.11.2010	5/ 12.03.2011	6/ 28.05.2011
<b>Aquisição</b>	2D dinâmico		
<b>Plano de corte</b>	Sagital		
<b>Sequências de pulso</b>	<i>Flash</i> gradiente-eco		
<b>Tempo de repetição (TR)</b>	1399.44	5.9	6.4
<b>Tempo de eco (TE)</b>	4.35	2.26	2.44
<b>Flip Angle (FA)</b>	10°	10°	10°
<b>Técnica de obtenção das imagens</b>	Planos		
<b>Número total de cortes</b>	7 / s	50	100
<b>Espessura de corte (mm)</b>	7	5	6
<b>Espaçamento (GAP)</b>	---	---	---
<b>Fator turbo (ETL)</b>	1	1	1
<b>FOV (mm)</b>	240*240	250*250	178*220
<b>Matriz</b>	564*512	256*256	156*192
<b>Percent sampling (p)</b>	100%	96% (246p*256)	90% (140p*192)
<b>Fator de aceleração</b>	4	4	4
<b>Tempo de aquisição</b>	2.80 s	1 min 29 s (0.74s + 1s entre cortes)	48s
<b>Resolução (px/mm)</b>	2.133	1.024	0.876
<b>Pixel size (mm)</b>	0.469 x 0.469	0.976 x 0.976	1.145 x 1.145
<b>Display Range</b>	0-300	0-597	0-340
<b>Aplicação avançada</b>	<i>Grid Tagging</i> <i>Trigger time</i> 1399.44 ms	37 repetições do som em média Pausa de 1s entre cada corte	30-40 repetições em média Pausa de 0.5s entre cada corte

O estudo dinâmico assentou essencialmente na captação em tempo-real dos movimentos dos articuladores, alcançando-se a recolha de 100 imagens para cada sequência de sons, no plano médio sagital, durante apenas 48 segundos.

Neste ensaio, obteve-se uma resolução superior à do estudo estático 3D, tendo-se escolhido duas sequências de sons em dois contextos articulatorios distintos:

- Sequência de sons vocálicos [a ε i ɔ u] (transcrição fonética do IPA);
- Sequência consoante-vogal da palavra [patu] (transcrição fonética do IPA).

Considerando a familiaridade dos sons e facilidade de articulação para qualquer sujeito, observou-se que a frequência de ocorrência de cada um dos sons foi de, pelo menos, 13 a 37 vezes por sequência de sons, dependendo da velocidade articulatória e da extensão do conjunto de sons. A vogal [a] ocorreu cerca de 27 e 37 vezes por sequência, o que pensamos estar relacionado com a maior duração acústica desse som, surgindo num menor número de imagens as vogais [ɔ] e [u].

Considerando a palavra [patu], cada som ocorreu, pelo menos, 13-42 vezes, dos quais com maior destaque, a vogal [a] cerca de 35 a 42 vezes, seguida da consoante plosiva [t] cerca de 13 a 16 vezes.

### **5.3 Utilização da ressonância magnética de alto campo: reflexão**

O estudo preliminar realizado anteriormente, no âmbito do Mestrado em Engenharia Biomédica, consistiu na caracterização morfológica e dinâmica da fala humana, mediante a utilização de um sistema de RM de 1.5 T no Serviço de Radiologia do Hospital S. João EPE, no Porto, culminando na construção de uma base de dados imagiológica 2D e modelos 3D do trato vocal, em situações estáticas, para 25 sons do Português Europeu (VENTURA et al., 2009a). Realizou-se também uma primeira abordagem imagiológica dinâmica mediante a combinação da sincronização cardíaca por eletrocardiografia e a repetição



sucessiva de um conjunto de três sílabas distintas, resultando numa sequência cinematográfica da fala (VENTURA et al., 2011c).

As principais dificuldades encontradas estavam relacionadas por um lado, com a inexistência de protocolos específicos para estudo do trato vocal por imagem por ressonância magnética e por outro, com as limitações do próprio equipamento de ressonância magnética, nomeadamente a baixa resolução temporal e de contraste. Estas dificuldades e, em particular, as limitações na caracterização dinâmica, vieram fundamentar a opção pela implementação deste estudo num sistema de 3.0 T existente na mesma instituição hospitalar, considerando também os objetivos agora propostos.

As melhorias alcançadas relativamente ao sistema de RM com potência de campo de 1.5 T devem-se essencialmente a dois fatores: 1) ao aumento da potência de campo; e 2) às aplicações avançadas disponíveis no *hardware* dos sistemas de 3.0 T. Estes dois fatores permitem assim um incremento da relação sinal-ruído, melhorando consideravelmente a qualidade das imagens para as tarefas de segmentação e reconstrução 3D, assim como a resolução temporal, possibilitando observar com maior acuidade os movimentos dos articuladores.

Quanto às aplicações avançadas do equipamento de 3.0 T utilizadas destacam-se:

- *Tecnologia Tim® (Total imaging matrix)* – possibilidade de integração de várias antenas e de um elevado número de canais de radiofrequência garantindo uma maior flexibilidade nos estudos, exatidão e velocidade; assim, para estudo do trato vocal é possível o uso combinado das antenas de cabeça e pescoço com melhorias significativas na relação sinal-ruído;
- Antena de cabeça com mais canais (32), antena de pescoço com 4 canais e ambas com diâmetros internos menores;
- Aquisição paralela – algoritmos que permitem reconstruir as imagens produzidas por cada canal de cada antena a partir do domínio da imagem (*Syngo SENSE*) ou do domínio da frequência do sinal transmitido pelas antenas (*Syngo GRAPPA*), acarretando uma melhor relação sinal-ruído e redução do tempo de aquisição;
- *Grid tagging* – aplicação cardíaca avançada que permite avaliar o movimento do coração, em particular a deformação mecânica do músculo

cardíaco. Tem sido também utilizada, em vários trabalhos (KANE et al., 2002; PARTHASARATHY et al., 2007; STONE et al., 2001; UNAY & OZTURK, 2000) relacionados com o estudo dinâmico durante a produção da fala.

Não obstante destas melhorias, os estudos estáticos 2D/3D podem ser efetuados com sucesso num sistema de 1.5 T. Quanto aos estudos dinâmicos, devido à natureza fisiológica extremamente rápida da produção da fala, exige uma frequência mínima de amostragem, segundo NARAYANAN e colaboradores (2004), na ordem dos 20 Hz. Assim, recomenda-se a sua implementação num sistema de 3.0 T de modo a melhorar a SNR em detrimento da redução do tempo de aquisição.

## 5.4 Conclusão

A utilização da ressonância magnética no estudo do trato vocal durante a produção da fala requer uma anamnese clínica rigorosa no sentido de justificar o potencial interesse da sua aplicação, assim como, permite orientar o protocolo de estudo e a seleção do *corpus* da fala em função da indicação clínica.

Assim, tanto do ponto de vista da sua utilidade clínica como da Engenharia, o tipo de estudo *versus* finalidade pretendida devem ser criteriosamente escolhidos, assim como deve ser garantida a preparação e instrução prévia dos pacientes de modo a garantir a qualidade diagnóstica das imagens e o seu processamento e análise *a posteriori*.

Estruturalmente, o protocolo de RM para caracterização funcional da fala é constituído pelo estudo estático 2D de cada som em estudo seguido de um estudo estático 3D e/ou dinâmico, de acordo com a finalidade pretendida.

Os estudos estáticos 2D são indispensáveis pois permitem a observação qualitativa detalhada da morfologia de todo o trato vocal com elevada resolução e num tempo de aquisição muito curto (requerendo um tempo mínimo de sustentação do som), servindo de imagem padrão para o respetivo som em estudo, quer para posterior comparação com as imagens sequenciais 2D do estudo dinâmico, quer para o estudo 3D.

Os estudos estáticos 3D permitem a obtenção de um considerável número de cortes seccionais em qualquer orientação, com uma resolução e SNR suficientes para posterior extração de características dimensionais, como áreas transversais ou volumes. Estes têm particularmente interesse quer para avaliação das cavidades de ressonância quer para a posterior modelação acústico-articulatória da fala. Para além disso, permitem uma caracterização morfológica mais completa das cavidades de ressonância.

O estudo dinâmico 2D permite a captação em tempo-real dos movimentos dos articuladores, no plano médio sagital, possibilitando o estudo de uma sequência de sons, sílabas ou palavras. A resolução temporal das imagens depende da velocidade articulatória e da extensão do conjunto de sons utilizada. A utilização de técnicas de aquisição ultrarrápidas combinadas com estratégias rápidas de preenchimento do *espaço K* é preferível, pois para além de disponíveis em todos os equipamentos de RM estas técnicas permitem a recolha de um número considerável de imagens.

Dada a natureza funcional dos articuladores da fala, a captação direta de imagens 2D sequenciais apresenta ainda algumas condicionantes técnicas de instrumentação num sistema de RM. Nesse sentido, impõem-se à técnica de RM conjuntamente com a Engenharia, importantes desafios nomeadamente, na procura e desenvolvimento de um *trigger* específico da fala e na gravação acústica da fala simultaneamente com as aquisições das imagens, de modo a permitir não só um maior rigor na análise das imagens através da correlação dos eventos articulatórios com o sinal da fala, como aumentar a confiabilidade dos resultados.



## **CAPÍTULO 6. PROCESSAMENTO E ANÁLISE DE IMAGENS DO TRATO VOCAL**

A área do processamento e análise de imagem médica é atualmente uma das mais promissoras e em maior desenvolvimento. A técnica de segmentação de imagem é uma tarefa muito utilizada, permitindo a individualização ou separação de um órgão, estrutura ou região em estudo da restante anatomia presente numa dada imagem. É por isso, determinante para o sucesso das tarefas seguintes, exigindo por parte do operador conhecimentos acerca das características das imagens, da anatomia em estudo e ainda, do tipo de métodos de processamento existentes.

As especificidades das imagens de RM aliadas à morfologia particular do trato vocal criam importantes condicionantes às tarefas de processamento e análise de imagem. No entanto, vários algoritmos computacionais estão atualmente disponíveis, integrados em ferramentas interativas e acessíveis a qualquer utilizador.

Assim, justifica-se a exploração das várias tarefas de processamento e análise de imagem médica, particularmente a sua aplicação na segmentação e caracterização do trato vocal. Importa assim indagar sobre: Que passos são necessários para a concretização das tarefas de processamento e análise de imagens de RM? Que aplicações práticas surgem em detrimento da escolha do protocolo e finalidade do estudo?

Neste capítulo são inicialmente descritas as principais características, do ponto de vista computacional, das imagens de RM. Seguidamente, apresentam-se as medidas quantitativas que podem ser extraídas na caracterização morfológica e funcional da fala usando técnicas de processamento e análise de imagem.

## 6.1 Processamento e análise de imagens de ressonância magnética

O sinal de RM é codificado espacialmente graças à variação linear dos gradientes de campo magnético, ao longo do paciente e em diferentes tempos (YOO, 2004). O sinal que é produzido pela aplicação de uma série de pulsos de RF é transformado do domínio da frequência para o domínio da imagem por um processo matemático denominado de transformada rápida de Fourier.

A imagem seccional reconstruída pode ser orientada em qualquer plano e depende das características tecidulares em estudo, ou seja, da correlação entre a energia absorvida (pulsos de RF) *versus* a frequência de precessão dos núcleos de hidrogénio na amostra. Uma imagem de RM consiste numa matriz numérica na qual os dados da amostra se encontram posicionados (localização no espaço) e discretizados (quantização da densidade do *pixel*). No entanto, porque cada elemento da imagem, isto é, cada *pixel* não está calibrado a uma determinada escala, estes vão depender das características da aquisição, do paciente e do equipamento. Acresce ainda que durante a aquisição, estes valores não são constantes devido às inhomogeneidades de campo magnético, constituindo assim um dos fatores mais preocupantes quando se pretende efetuar operações de segmentação nestas imagens.

A resolução das imagens de RM é definida pela sua matriz e usualmente pode abranger 512 x 512 (*m* linhas por *n* colunas) ou 256 x 256 *pixels*, dependendo da frequência de amostragem ou do número de codificações de fase. A resolução espacial depende da potência do campo magnético e da performance dos gradientes.

Tipicamente as imagens de RM apresentam elevado contraste (12 *bits*, ou seja, 4096 níveis de cinza por *pixel*), muito ruído e são altamente suscetíveis à ocorrência de artefactos, dificultando as tarefas de processamento e análise. Dado que um protocolo por RM, inclui a utilização de várias sequências de pulso e ponderações, em vários planos, originam um grande volume de dados (cerca de 50 a 60 imagens por sequência de pulsos num tempo de aquisição variável entre 5 a 10 minutos) e um tempo total de exame variável entre 50 a 60 minutos. O

peso computacional é considerável dependendo da conjugação de todos os parâmetros técnicos de aquisição.

Assim, de uma forma geral, os estudos por RM do trato vocal envolvem as seguintes etapas (DESERNO, 2011) para a sua concretização:

- I. Formação das imagens: obtenção e reconstrução das imagens de RM com o mínimo de artefactos e ruído;
- II. Visualização das imagens e, se necessário, a utilização de filtros de processamento de imagem para remoção do ruído;
- III. Análise das imagens: segmentação das imagens para extração de características, como contornos, orlas, linhas e pontos; e posterior extração de parâmetros ou características mensuráveis do trato vocal;
- IV. Organização e manipulação das imagens: conjunto de técnicas para arquivo eficiente, comunicação, transmissão e compressão dos dados.

Cada uma destas etapas é desenvolvida e otimizada separadamente, o que torna assim, estes estudos muito exigentes do ponto de vista operacional. Os avanços nas tarefas de obtenção e reconstrução das imagens e da sua segmentação poderão conseqüentemente contribuir para um maior sucesso destes estudos.

No que diz respeito às tarefas de segmentação de imagem, nos vários estudos encontrados, vários métodos foram usados, cada qual com vantagens e limitações, nomeadamente:

- Método manual: não requer a utilização de nenhum *software* específico para a sua utilização e por isso, é fácil de implementar. No entanto, por ser dependente do operador e ser uma tarefa muito árdua e morosa, produz alguma incerteza nos resultados (BADIN et al., 2000; BIRKHOLZ & KRÖGER, 2006; DEMOLIN et al., 1996; ENGWALL, 2006; SERRURIER & BADIN, 2005; SOQUET et al., 2002; TAMEEM & MEHTA, 2004);
- Método semiautomático: forma interativa e de fácil utilização, menos moroso comparativamente ao método manual e usualmente de fácil implementação, nomeadamente na diferenciação anatômica do trato vocal, como fronteiras ar/tecidos moles. Exige uma menor dependência do operador, é um método ainda relativamente simples, no entanto,

necessita de ajustes manuais. Surgem ainda dificuldades, como fuga de contornos pela proximidade de estruturas com similaridade de intensidade de sinal e os resultados dependem da homogeneidade da imagem que está dificultada pelo ruído e artefactos das imagens de RM (BAER et al., 1991; BEHRENDTS & WISMULLER, 2001; ENGWALL, 2000b; ENGWALL, 2004; KENNETH et al., 2001; MOORE, 1992; STORY et al., 1996; TAKEMOTO & HONDA, 2003);

- Método automático: método dinâmico, menos moroso que os anteriores e que permite a adaptação automatizada a todo o volume. Requer alguma análise cuidadosa dos resultados e usualmente correções manuais. Apresenta dificuldades em delimitar áreas com oclusões, como acontece nalguns sons em que o trato vocal apresenta interrupções (BRESCH et al., 2006; NARAYANAN et al., 2004).

Dependendo da finalidade pretendida, nomeadamente o tipo de análise e dimensão do *corpus* da fala e quantidade de imagens obtidas, a escolha do método de segmentação varia, apesar de que a automatização é um recurso crucial face às características quer das imagens de RM quer pelas especificidades anatómicas e dinâmicas do trato vocal.

## 6.2 Morfometria 2D/3D do trato vocal

A análise da forma ou morfometria do trato vocal pode ser efetuada a partir de imagens bidimensionais como também a partir de aquisições volumétricas tridimensionais, tanto de forma qualitativa mais subjetiva como quantitativa. A análise qualitativa em Imagiologia médica é a mais rotineiramente empregue na prática clínica e tem por base a visualização das diferenças de intensidade de sinal dos tecidos, isto é, o contraste nas diferentes ponderações e sequências de pulso e a apreciação geral da morfologia dos órgãos. Dessa primeira análise e observando-se a alteração dos padrões anatómicos normais da estrutura ou órgão em estudo, do ponto de vista clínico justifica-se uma caracterização quantitativa rigorosa. A análise qualitativa está sujeita a variabilidade inter e



intra-observadores justificando-se por isso, o desenvolvimento de sistemas computacionais auxiliares de análise das imagens adquiridas.

Quanto mais dados forem apreciados e recolhidos da anatomia em estudo maior será a precisão e exatidão das medidas quantitativas extraídas e conseqüentemente do diagnóstico clínico, ou no caso da Engenharia, da performance, por exemplo, dos sistemas de síntese articulatória criados a partir de tais dados.

As medidas quantitativas para caracterização da forma do trato vocal incluem a determinação de áreas seccionais e volumes do ponto de vista clínico, e o cálculo das funções de área para modelação do trato vocal e posterior simulação acústica.

As tarefas de processamento e análise de imagem atualmente disponíveis através de vários programas comerciais de domínio público, dispõem de um conjunto de algoritmos e ferramentas automatizadas que permitem não só a redução do tempo despendido como ainda a sua fácil utilização e implementação. No entanto, na escolha do programa a usar deve ter-se em conta determinadas características do mesmo, de forma a ajustar-se às necessidades do estudo, nomeadamente:

- Plataforma – compatibilidade com o sistema computacional a usar, em particular o sistema operativo do mesmo;
- Ficheiros de *input* – formatos de imagem que consegue abrir, especialmente o formato *Digital Imaging and Communications in Medicine* (DICOM);
- Ficheiros de *output* – formatos de imagem que produz;
- Aplicações – em que áreas médicas podem ser úteis.

Nas secções que se seguem descrevem-se algumas das aplicações práticas realizadas no decurso deste projeto, que exemplificam por um lado, as várias técnicas de processamento e análise de imagem e as medidas quantitativas que podem ser extraídas do trato vocal e que, por outro lado, suportaram o processo de decisão na construção e otimização do protocolo de RM.

### 6.2.1 Interesse Clínico

Do ponto de vista morfológico, a morfometria do trato vocal inclui a determinação de áreas seccionais e a modelação 3D do trato vocal com base nas sequências de pulso 2D e 3D.

Como referido anteriormente no Capítulo 4, a maioria dos estudos revistos na literatura recorreu à técnica *multi-stacks* para aquisição das imagens de RM, verificando-se apenas mais recentemente a utilização da técnica de aquisição volumétrica durante a produção da fala (AALTO et al., 2011; KIM et al., 2009; MARTINS et al., 2008). No sentido de avaliar a adequabilidade atual de cada uma destas técnicas de aquisição e pela inexistência de estudos comparativos na literatura, o desenvolvimento e otimização dos estudos estáticos 3D teve por base estas duas metodologias de aquisição de imagens e, na posterior, análise comparativa e confiabilidade dos resultados (VENTURA et al., 2011a). No estudo participaram três jovens voluntários, sem história de patologia oro-facial, com idades de  $22.3 \pm 1.5$  anos, treinados previamente para garantir a produção correta dos onze sons do Português Europeu.

Procedeu-se então à definição dos parâmetros de aquisição e estruturação de cada uma das técnicas de aquisição, a partir de uma sequência de pulsos rápida *flash* gradiente eco, de forma a:

- Obter um número comparável de imagens, para abranger todo o trato vocal e com razoável SNR e resolução espacial de modo a facilitar a morfometria do trato vocal e respetiva modelação 3D;
- Reduzir o tempo de aquisição de modo a reduzir o efeito de hiperarticulação dos sons, evitar artefactos de movimento e facilitar a sustentação dos sons.

A análise comparativa das imagens obtidas pelas duas técnicas de aquisição e intersujeitos, assentou em duas fases distintas, considerando o elevado número de imagens e dimensão do *corpus* da fala:

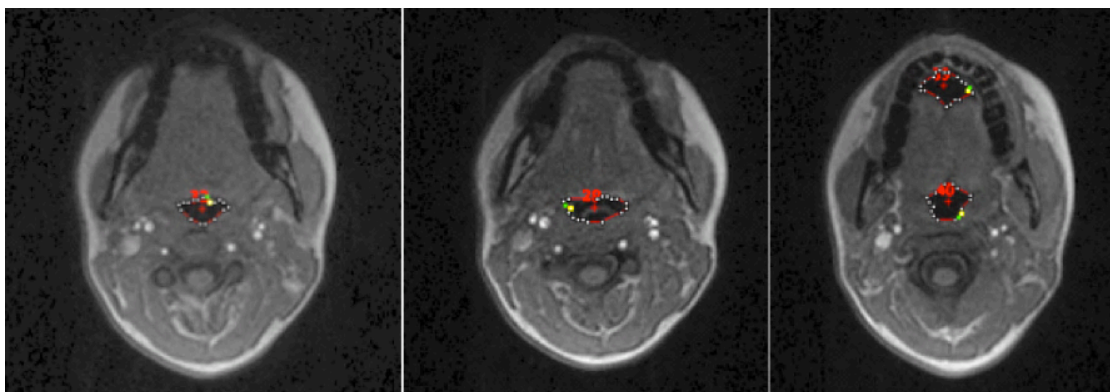
- I) Para a caracterização dos cinco sons vocálicos, procedeu-se à análise cuidadosa das imagens adquiridas para extração dos contornos e áreas seccionais do trato vocal e medição do volume total;
- II) Para a caracterização das consoantes, uma vez que a segmentação do trato vocal é mais difícil de realizar dado que existe uma obstrução ao longo do

trato vocal, apenas se procedeu à reconstrução tridimensional do trato vocal e respetiva estimativa dos volumes.

*a) Caracterização das vogais*

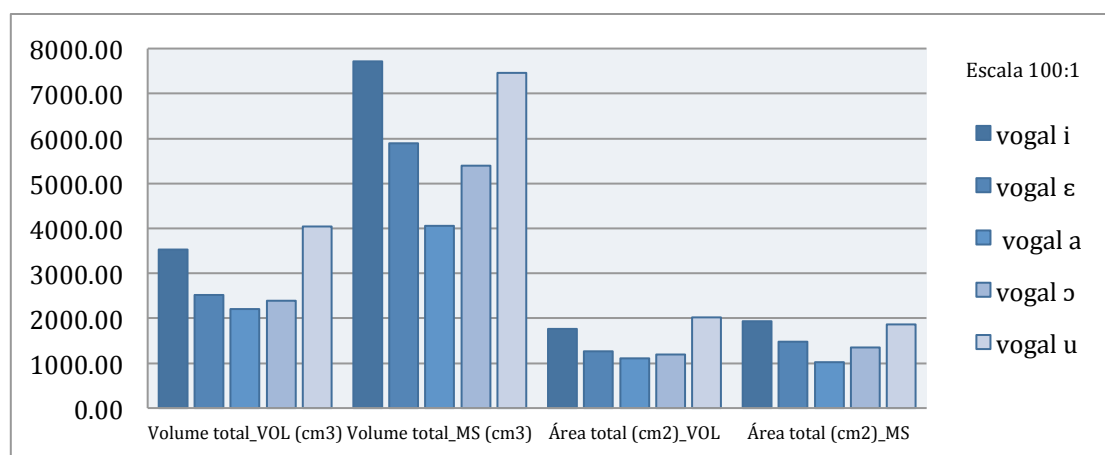
Na primeira tarefa, recorreu-se ao programa de processamento de imagens médicas, análise e visualização - MIPAV (MCAULIFFE et al., 2001), desenvolvido pelo Centro de Tecnologia da Informação e pelo *National Institutes of Health* (Bethesda, Maryland, EUA). Para a segmentação do trato vocal utilizou-se o método de *Level-set* originalmente desenvolvido por OSHER & SETHIAN (1988) e, posteriormente usado na segmentação de imagens por MALLADI e colaboradores (1995). Os modelos de *Level-set* são modelos deformáveis geométricos permitem a segmentação com base na deformação de superfícies, sendo a curva inicial o *Level-set* de nível zero, evoluindo de forma a convergir para a forma de interesse a segmentar. Estes modelos têm tido uma utilização crescente em Imagem médica e visão computacional pela sua eficiência (ROUSSON & PARAGIOS, 2008). Estas curvas podem ser facilmente aplicadas a formas complexas e diferentes topologias, assegurando o sucesso da segmentação e exigindo baixa intervenção do utilizador.

Os contornos segmentados do trato vocal em cada imagem foram posteriormente agrupados, como se pode observar na Figura 6.1, permitindo o cálculo das áreas transversais e do volume total que resultou da soma de todas as áreas transversais e multiplicação pela espessura de corte.



**Figura 6.1** Segmentação dos contornos do trato vocal ao longo de várias imagens de RM, pelo método de *Level-set* e respetivas regiões de interesse obtidas.

As áreas seccionais calculadas usando técnica de *multi-stacks* apresentam uma distribuição bastante heterogénea, não-linear, devido às três orientações de corte e são globalmente maiores para as vogais em estudo. Os contornos do trato vocal foram segmentados onde se encontrou um contorno fechado em cada um dos cortes, desde a glote até aos lábios. A título ilustrativo, a Figura 6.2 apresenta os volumes e área totais obtidas para cada vogal para um dos sujeitos do sexo feminino, a partir das imagens adquiridas com base nas técnicas de aquisição volumétrica e *multi-stacks*.



**Figura 6.2** Análise comparativa da morfometria do trato vocal durante a produção de 5 sons vocálicos extraída a partir da técnica de aquisição volumétrica (VOL) e *multi-stacks* (MS).

### b) Caracterização das consoantes

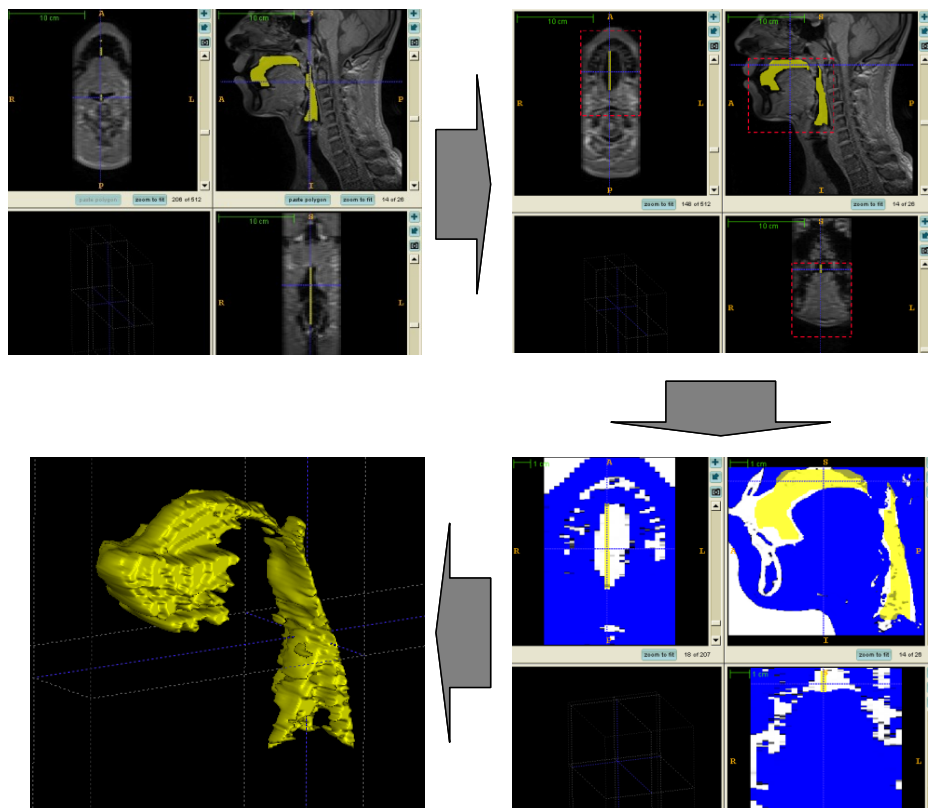
A reconstrução 3D do tacto vocal a partir das imagens de RM foi realizada com recurso ao programa *ITK-SNAP*® (versão 2.1.4-rc1) desenvolvido no *Penn Image Computing and Science Laboratory*, de domínio público para segmentação de estruturas 3D em imagens médicas, utilizado frequentemente em imagens de tomografia computadorizada e ressonância magnética (YUSHKEVICH et al., 2006). A utilização desta ferramenta no estudo do trato vocal foi anteriormente referida também por outros autores (FUCHS et al., 2008; MARTINS et al., 2008).

Este programa possibilita, de forma bastante interativa, a segmentação manual e automática através de métodos de contornos ativos, usualmente

conhecidos por *snakes*<sup>9</sup>, e janelas de visualização para navegação 3D. Para além disso, permite o recurso a várias aplicações, nomeadamente:

- Cursor de ligação para navegação em 3D permitindo a segmentação em simultâneo nos três planos ortogonais;
- Compatível com diferentes formatos de imagem, como DICOM, *NiftI*, e *Mayo Analyze*.
- Segmentação manual em simultâneo nos três planos ortogonais;
- Interface do utilizador simplificada para a segmentação semiautomática;
- Ferramentas de pós-processamento para visualização dos resultados;
- Amplo apoio tutorial.

A Figura 6.3 ilustra a sequência de tarefas de processamento de imagens para segmentação e reconstrução 3D do trato vocal.



**Figura 6.3** Sequência de tarefas de processamento de imagem realizadas para modelação 3D do trato vocal. As quatro janelas de visualização permitem executar a segmentação simultânea nos três planos ortogonais e visualizar o respetivo modelo 3D reconstruído.

<sup>9</sup> O modelo original de *snakes*, desenvolvido por KASS et al. (1987) é, basicamente, uma curva contínua parametrizada, no domínio da imagem, que se pode deslocar sob a influência de forças internas, definidas em função da própria curva e forças externas, definidas em função da imagem. As *snakes* são modelos de contornos deformáveis que permitem a sua utilização de forma interativa por minimização de energia, sendo também usadas no seguimento de objetos deformáveis em sequências de imagens.

Dado o considerável número de imagens que se obtiveram na aquisição volumétrica de RM do trato vocal e de forma a, minimizar o tempo despendido e sistematizar as tarefas de segmentação realizadas para todos os conjuntos de imagens, procedeu-se da seguinte forma:

- i) Abertura do ficheiro DICOM, através da operação *loading greyscale image*;
- ii) Pré-segmentação manual do trato vocal de forma a minimizar o número de iterações;
- iii) Definição de uma caixa delimitativa da região de interesse de forma evitar a fuga de contornos para estruturas adjacentes;
- iv) Segmentação 3D em simultâneo nos três planos ortogonais usando o algoritmo de *snakes*;
- v) *Volume rendering* (renderização de volume) – para visualizar os modelos 3D criados a partir da segmentação das imagens 2D.

Os modelos 3D do trato vocal obtidos para cada par de consoantes fricativas do Português Europeu, vozeados e não vozeados, encontram-se ilustrados na Figura 6.4.

Comparando os volumes medidos dos modelos tridimensionais do trato vocal gerados a partir de cada técnica de aquisição de imagens, constatou-se que os valores relativos às consoantes não vozeadas são menores por comparação com a sua consoante homóloga vozeada. Apenas no sujeito masculino se observou o oposto, durante a produção dos dois pares de consoantes fricativas, especialmente nos pontos de articulação alveolar e pós-alveolar. Resultados semelhantes também foram observados num estudo realizado por PROCTOR e colaboradores (2010), mas apenas considerando o volume da faringe durante a produção das consoantes fricativas do Inglês Americano: deve-se notar que as diferenças entre o Inglês Americano e o Inglês Britânico são relativamente semelhantes às diferenças entre o Português Brasileiro e o Português Europeu.

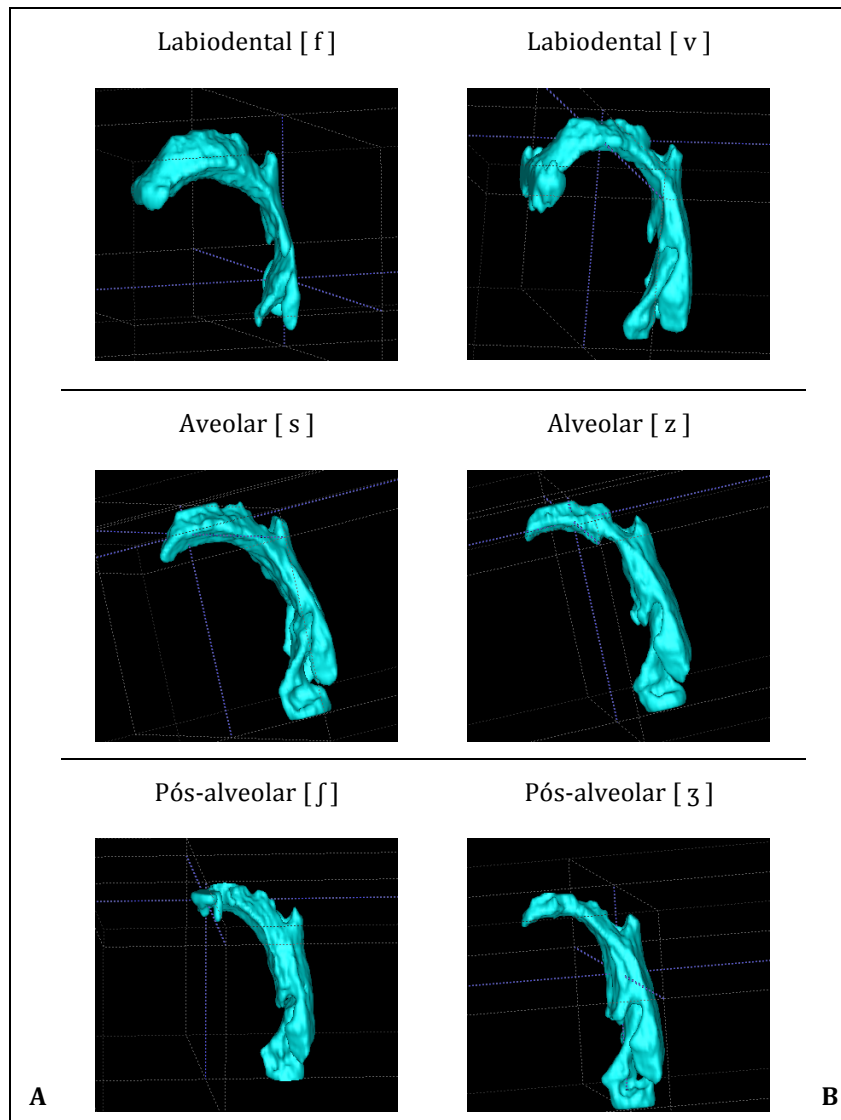


Figura 6.4 Modelos 3D do trato vocal reconstruídos a partir da aquisição volumétrica durante a produção de consoantes fricativas não vozeadas (A) e vozeadas (B).

Constatamos ainda que em relação a cada ponto de articulação, a forma global 3D do trato vocal foi semelhante entre cada grupo de consoantes e para cada sujeito.

Deste estudo, foi possível observar que, utilizando a mesma metodologia de análise para cada técnica de aquisição, os resultados obtidos são distintos na avaliação quantitativa do trato vocal, pelo que a técnica de aquisição *multi-stacks* gera uma sobrestimação dos resultados nas diferentes medidas realizadas. A diferença é particularmente notória no cálculo dos volumes e intersujeitos. No entanto, estes resultados devem ser analisados com cuidado pois a sua

interpretação não pode ser feita literalmente dadas as diferenças inerentes à multidimensionalidade de cada técnica de aquisição.

Assim, constatou-se que:

- A utilização de conjuntos de cortes (*stacks*) orientados de acordo com a forma do trato vocal permite obter informações úteis acerca das dimensões laterais do trato vocal, em particular, na região velofaríngea; no entanto, os cortes coronais e oblíquos são muito difíceis de segmentar e interpretar;
- Ainda que sejam comparadas somente as medidas quantitativas extraídas a partir dos cortes transversais de cada técnica de aquisição e durante a produção da mesma vogal, os resultados são consideravelmente muito distintos. Os valores mensurados das áreas seccionais do trato vocal são mais elevados na técnica de aquisição *multi-stacks*; além disso, não foram analisadas um número significativo de imagens por não se identificar um contorno fechado, particularmente na região dos lábios;
- Os problemas de alinhamento e registo entre os diferentes *stacks* causam inúmeras dificuldades durante a modelação 3D, obrigando à análise individual de cada conjunto de cortes, condicionando assim a precisão dos volumes calculados; acresce ainda que a reconstrução dos modelos 3D do trato vocal impõe um ajuste manual;
- A não-identificação dos dentes nas imagens obtidas por ressonância magnética dificultam o processo de segmentação e, conseqüentemente, interferem na precisão dos resultados. Esta situação é particularmente notória, nos cortes coronais devido à fuga dos contornos, resultando em modelos 3D do trato vocal ligeiramente sobrestimados para ambas as técnicas de aquisição;
- A sobreposição de cortes nas diferentes *stacks* leva ao aumento das medidas quantitativas extraídas;
- A epiglote não foi tida em consideração na análise dos cortes transversais independentemente da técnica de aquisição usada.

Em suma, consideramos que tendo em vista a morfometria do trato vocal, a aquisição volumétrica é preferível pois as medidas quantitativas obtidas, numa comparação intersujeitos, são mais consistentes, fáceis de extrair e analisar e



menos dependente dos conhecimentos do operador, em particular do ponto de vista do reconhecimento anatômico dos contornos do trato vocal. No entanto, na caracterização morfométrica de uma cavidade de ressonância em particular, a aquisição de *stacks* perpendicularmente orientados, permitem uma análise dimensional acrescida e mais rigorosa.

Deste modo, a técnica de aquisição de ressonância magnética deve ser selecionada de acordo com o propósito do estudo, objetivos clínicos e precisão dos resultados pretendidos.

### 6.2.2 Interesse em Engenharia

Ainda que atualmente nenhum sistema de síntese articulatória seja capaz de reproduzir com elevada qualidade o sinal acústico da voz falada ou cantada, devido à complexidade “natural” deste mecanismo de produção e escassez de dados morfológico-dinâmicos, estes sistemas computacionais cada vez mais se aproximam da realidade uma vez que se baseiam nos modelos do trato vocal humano e no processo de articulação (KROGER & BRIKHOLZ, 2009).

Estes sistemas consistem em três módulos destacando-se a função primordial do modelo do trato vocal, que consiste em gerar informação geométrica do trato vocal relativamente à forma e posição dos articuladores e funcional acerca da dinâmica articulatória, Figura 6.5. Estes parâmetros mensuráveis de forma, posição e movimento são gerados em função do tempo para todos os parâmetros definidos pelo modelo.

A relação entre as características geométricas e acústicas do trato vocal não é simples, mas no que respeita às vogais, as formantes podem ser determinadas com base nas funções de área, ou seja, nas áreas seccionais ao longo de todo o trato vocal. A análise destas áreas exige medições rigorosas, podendo ser extraídas diretamente a partir de imagens volumétricas obrigando a uma postura articulatória estática, ou indiretamente, a partir de imagens sequenciais mas que se limitam a um único corte sagital.

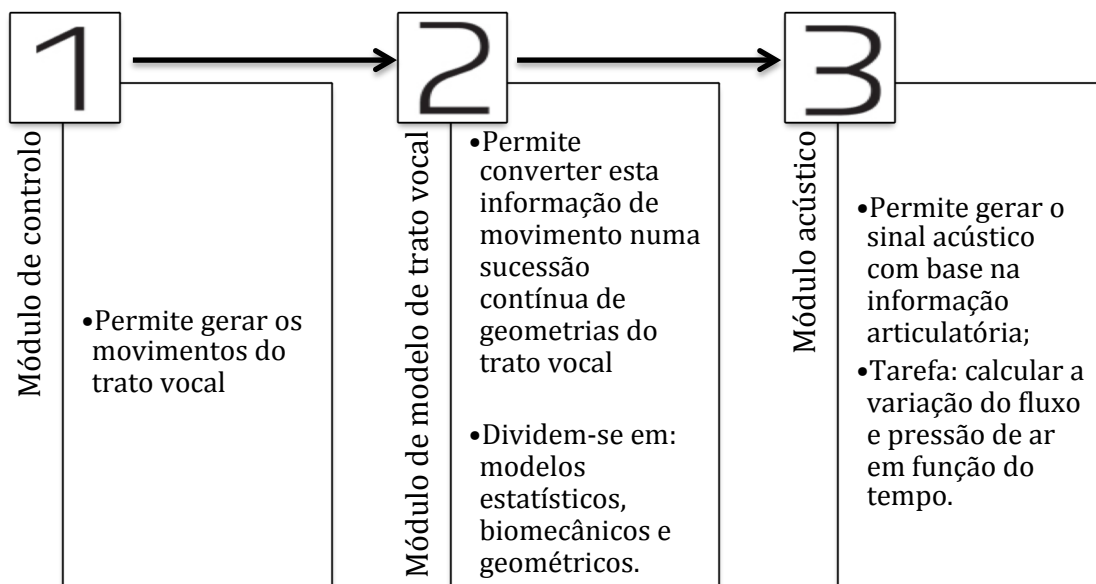


Figura 6.5 Módulos constituintes de um sistema de síntese articulatória.

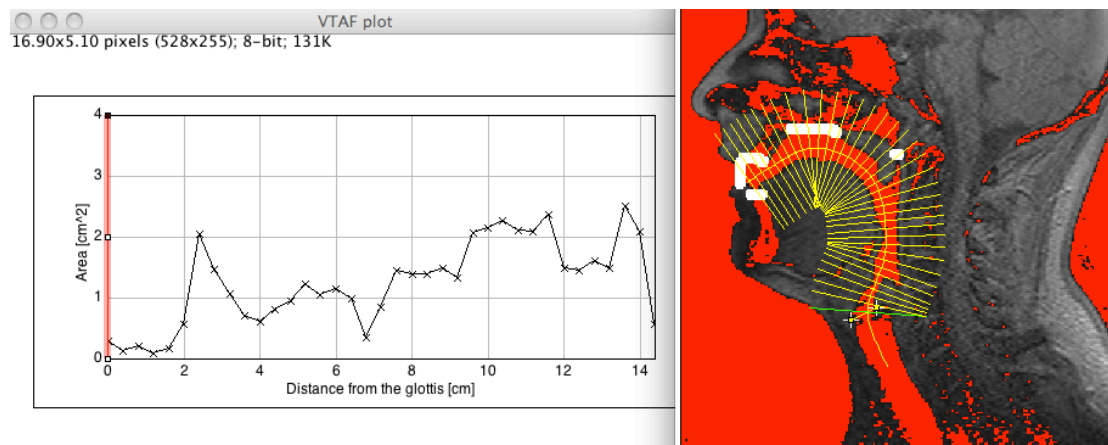
Estas áreas seccionais são localizadas e obtidas nos planos perpendiculares ao longo da linha central do trato vocal que se estende desde a glote até aos lábios (KROGER et al., 2000; STORY et al., 1996).

Considerando a importância que a RM apresenta na caracterização morfológica dos diferentes órgãos, nomeadamente de tecidos moles e a utilização crescente que se tem vindo a verificar nos estudos da produção da fala, TAKEMOTO e colaboradores (2006) desenvolveram o algoritmo “VTAF\_extractor” para extrair as funções de área diretamente a partir de uma aquisição volumétrica de RM. Este algoritmo encontra-se disponível numa funcionalidade (*plugin*) do programa de domínio público, ImageJ (versão 1.44o) em linguagem Java, desenvolvido por Wayne Rasband do *National Institutes of Health* (Bethesda, Maryland, EUA).

O algoritmo é de fácil utilização, obrigando no entanto, à aquisição volumétrica de imagens apenas no plano sagital; este possibilita a determinação da linha central do trato vocal, a posterior reconstrução dos cortes seccionais perpendicularmente a esta e a medição das áreas. Assim, em apenas 9 passos é possível a extração das funções de área, exigindo apenas do utilizador conhecimentos da anatomia em estudo para i) a marcação da linha correspondente à cordas vocais no corte mais central; e ii) limitar a região de

interesse ao nível dos lábios e outros pontos, dos dentes, véu do palato em que possa ocorrer “fuga” de segmentação, em todas as imagens sagitais.

A partir dos ensaios anteriormente descritos no capítulo anterior, foi possível a experimentação deste “VTAF\_extractor”, num sujeito do sexo masculino e para a vogal [a] (transcrição fonética do IPA) do PE. As funções de área extraídas estão representadas na Figura 6.6.



**Figura 6.6** Gráfico com a distribuição das áreas seccionais extraídas desde a glote até aos lábios e respetiva localização dos cortes perpendiculares ao trato vocal na imagem médio sagital de RM.

### 6.3 Medidas articulatórias a partir de imagens sequenciais

A caracterização funcional da fala a partir das imagens de RM, impõe um conjunto de desafios, quer do ponto de vista das técnicas de aquisição e performance de um sistema de RM, quer também do ponto de vista da utilização de métodos de análise computacional.

A rápida evolução tecnológica e divulgação da RM, nomeadamente dos sistemas de alto campo, veio abrir novos horizontes na avaliação de inúmeros órgãos e sistemas, mas originou principalmente, o desenvolvimento e aperfeiçoamento de técnicas avançadas, em órgãos para os quais, além da caracterização morfológica era necessário o estudo dinâmico ou funcional.

Os eventos dinâmicos da produção da fala requerem uma frequência de amostragem mínima, exigindo do ponto de vista da técnica de RM, o

desenvolvimento de estratégias de preenchimento do *espaço K* ou *trigger* mais eficientes.

Do ponto de vista computacional e dado que a quantidade de imagens obtidas ao longo do tempo é considerável, o recurso a técnicas automatizadas é crucial para a extração eficiente de medidas quantitativas dos eventos articulatórios observados.

A importância da aplicação dos modelos estatísticos nos estudos da produção da fala foi demonstrada por vários investigadores, nomeadamente por HARSHMAN e colaboradores (1977) que utilizaram o método de PCA para seguimento e análise articulatória da língua. Também MAEDA (1988), com recurso à análise fatorial, descreveu a forma lateral do trato vocal, seguindo-se YEHIA & TIEDE (1997) que demonstraram a utilização dos modelos PCA na observação da forma do trato vocal a partir de imagens de RM.

Estes modelos utilizam pontos-chave da região em estudo para, a partir destes localizar ou segmentar a forma do trato vocal em imagens posteriores. Este processo de etiquetagem obriga ao conhecimento da anatomia em análise para a marcação dos pontos articulatórios pretendidos. As imagens de referência geradas neste processo, vão assim servir de imagens de treino para o processo de modelação estatística *a posteriori*.

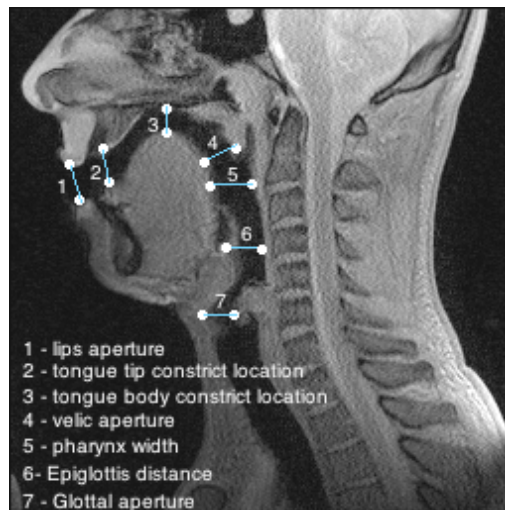
Os modelos de distribuição pontuais (COOTES et al., 1992; VASCONCELOS et al., 2010) são assim construídos a partir de um conjunto de imagens de treino, onde a forma geométrica do trato vocal é representada por um conjunto de pontos-chave, isto é, pontos característicos do objeto em questão. Os conjuntos de pontos-chave de cada imagem de treino são etiquetados considerando a anatomia dos pontos articulatórios pretendidos e posteriormente alinhados. Finalmente, recorre-se ao método de análise das componentes principais com o objetivo de seguir a variabilidade da forma geométrica do trato vocal.

Para além da forma anatómica, na construção destes modelos pode-se utilizar os níveis de cinzento que rodeiam cada ponto-chave. Deste modo, através da combinação destes dois modelos – forma geométrica e dos níveis de cinzento, é possível obter modelos de forma ativos e modelos de aparência ativa, que conduzem a modelações mais realísticas e robustas, para caracterizar e simular o

trato vocal, bem como para posterior segmentação e reconhecimento automático em novas imagens.

A utilização destes modelos nas imagens sequenciais de RM permitem a caracterização dos eventos articulatórios, quer através da extração da forma do trato vocal e posterior reconstrução 3D, quer pela extração de medidas quantitativas como, a distância euclidiana entre os principais pontos de articulação (VASCONCELOS et al., 2011a, b).

Do desenvolvimento e otimização da técnica de RM para estudo dinâmico dos articuladores durante a fala, descrito no capítulo anterior, foram recolhidas 200 imagens para cada sujeito. Para a caracterização funcional da fala, procedeu-se à medição de sete distâncias, Figura 6.7, cujos pontos de articulação escolhidos incluem todos os articuladores e permitem a classificação dos sons (VENTURA et al., 2011d).



**Figura 6.7** Medidas articulatórias realizadas imagens sequenciais de RM durante a produção da fala (VENTURA et al., 2011d).

A extração automática dos pontos através dos modelos de aparência ativa foi realizada separadamente para cada sequência de sons em estudo e para cada sujeito, utilizando-se somente as primeiras 20 imagens sequenciais obtidas para a etiquetagem das restantes.

A identificação do som produzido foi realizada mediante a análise visual da postura articulatória, tendo-se excluído as posições de repouso ou imagens com muitos artefactos. A análise quantitativa consistiu na determinação da média e

desvio-padrão das sete distâncias extraídas para cada sujeito e, para cada grupo de imagens que caracterizam cada um dos sons.

Como principais resultados, e considerando as duas sequências de sons, observou-se uma distribuição dos valores medidos com menores diferenças para a vogal [a]. Numa análise comparativa entre sujeitos para esta vogal, e considerando os dois contextos articulatórios, verificaram-se valores médios inferiores das distâncias quando a vogal foi produzida no contexto articulatório consoante-vogal, Figura 6.8.

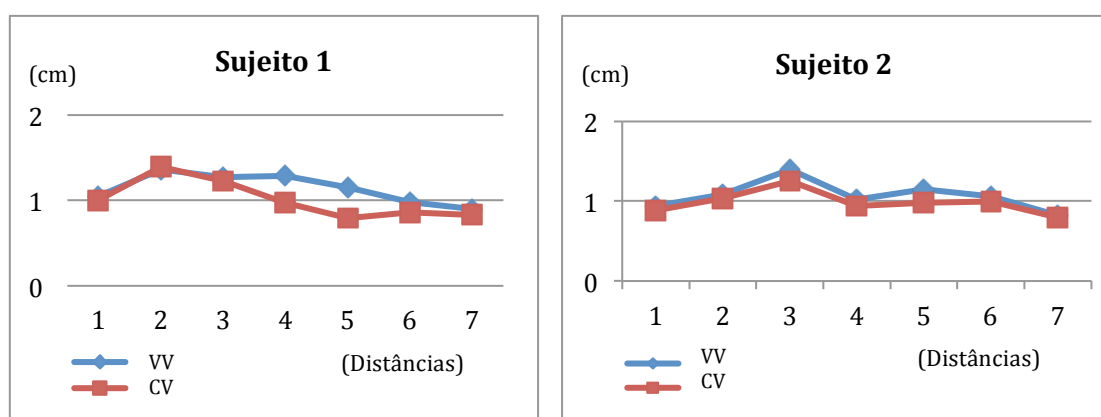


Figura 6.8 Análise comparativa das distâncias medidas para a vogal [a] quando produzida em contexto articulatório vogal-vogal (VV) ou consoante-vogal (CV), para cada sujeito.

## 6.4 Identificação dos dentes em imagens de RM

A identificação dos dentes em imagens de RM constitui um importante desafio nos estudos da fala, quer pelas condicionantes impostas pela instrumentação dos sistemas de RM, quer mesmo por questões de biocompatibilidade, o que tem motivado estudos de investigação por parte da comunidade científica (ver secção 4.2). Acresce que, o estudo dos dentes por RM, considerando que as técnicas imagiológicas atualmente disponíveis utilizam radiação X, abre novas perspetivas diagnósticas e terapêuticas, nomeadamente de visualização 3D avançada, com potencial nas áreas de Odontologia e Ortodontia (OLT & JAKOB, 2004).

Neste projeto esta problemática não foi desprezada, tendo sido realizados vários ensaios experimentais e contactos, quer com técnicos de farmácia como com médicos dentistas.

Para além de terem sido exploradas algumas das técnicas já descritas na literatura, para melhorar a informação obtida nas imagens de RM, como o preenchimento da cavidade bucal com água (OLT & JAKOB, 2004), procurou-se também a utilização de materiais com os seguintes requisitos:

- Biocompatibilidade na cavidade bucal: substâncias inócuas para os dentes, sem riscos no caso de deglutição e de fácil utilização e manuseamento;
- Compatibilidade com o ambiente usual dos sistemas de RM: respeitando as normas de segurança e suas contraindicações;
- Facilmente identificável nas imagens de RM: substância com elevado sinal e sem causar artefactos nas imagens para posterior extração da localização e forma dos dentes;
- Possibilidade de serem utilizadas no protocolo de RM dirigido ao estudo do trato vocal desenvolvido, isto é, durante a fonação;
- Disponibilidade e acessibilidade do material.

Com base nestes pressupostos, realizaram-se ensaios experimentais com substâncias para revestimento dos dentes, como verniz de flúor, um adesivo ortodôntico, uma substância de branqueamento dentário com moldeira e uma moldeira em silicone efetuada por uma médica dentista.

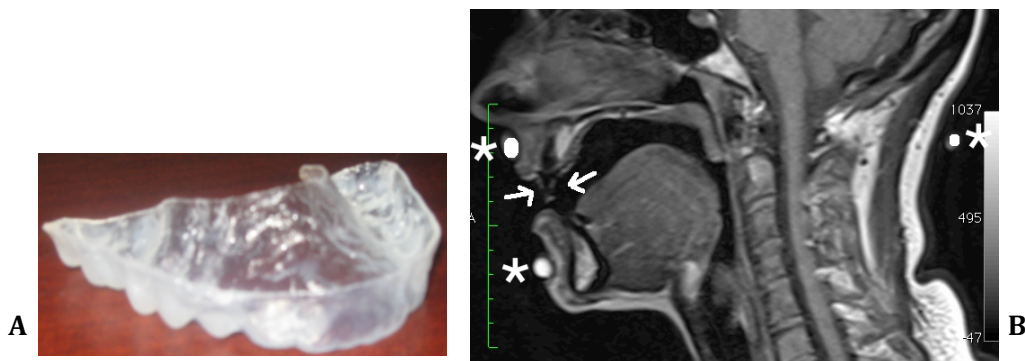
Apesar das várias soluções tentadas, nenhum destes materiais foi identificável nas imagens de RM. Depois de vários testes e contactos, definiram-se duas novas opções:

- 1) Utilizar água para preenchimento da cavidade bucal quer na posição corporal de decúbito ventral quer dorsal;
- 2) Utilizar a moldeira de silicone, na arcada dentária superior, impregnada com vaselina sólida.

Em ambas as opções, procedeu-se à colocação de cinco marcas de óleo de fígado de bacalhau, localizadas em pontos anatómicos de superfície: 1) protuberância occipital externa, 2-3) articulações temporo-mandibulares, 4) limite superior da sínfise mentoniana, 5) espinha nasal anterior de modo a permitir posteriormente a referência e alinhamento espacial das imagens dos dentes ou dos modelos 3D dos dentes, nas imagens de RM obtidas durante a fala.

Quanto ao preenchimento da cavidade bucal com água, para além do desconforto que causa, independentemente da posição corporal, a extração dos contornos dos dentes, tem que ser feita manualmente, resultando numa geometria grosseira para a anatomia em causa. Acresce ainda a impossibilidade de realização desta tarefa em fonação, tratando-se assim de uma solução unicamente interessante para extração direta da morfologia dentária numa situação estática.

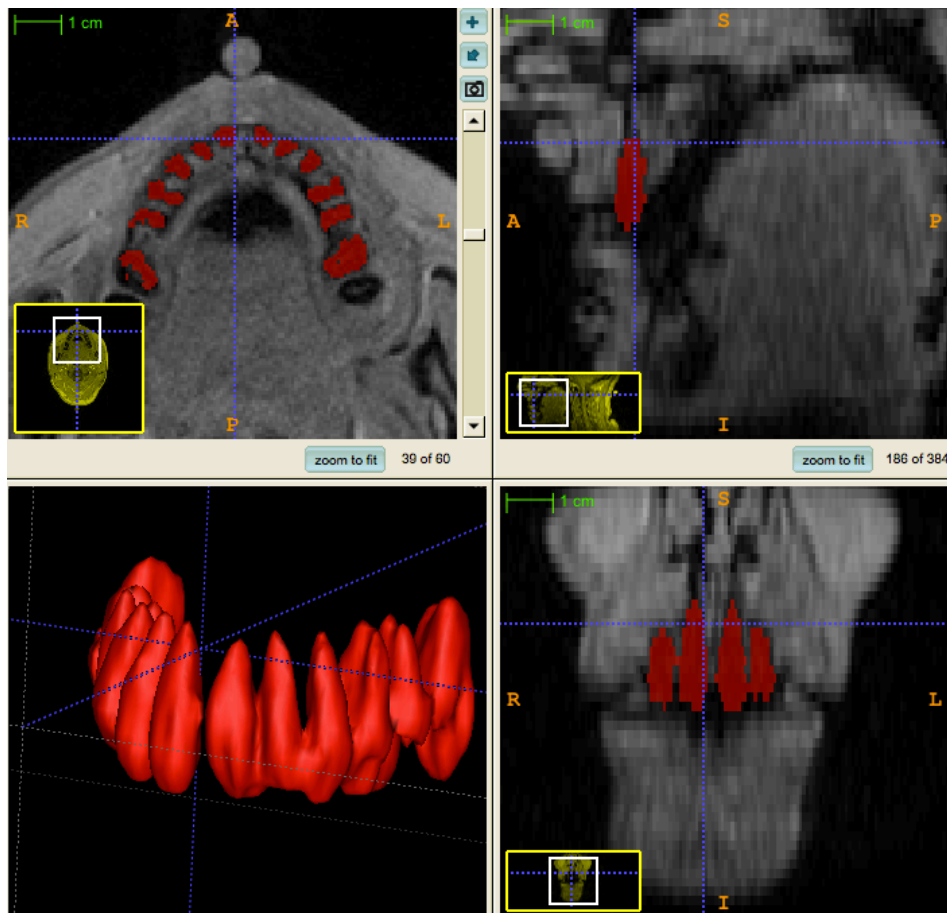
A utilização de vaselina impregnada na moldeira é viável, menos desconfortável que a opção anterior, possibilitando em contrapartida a sua utilização durante a fonação, podendo ser assim relevante quer na caracterização morfológica como dinâmica da produção da fala, Figura 6.9. O elevado sinal produzido pela vaselina, permite a separação eficaz dos dentes da cavidade bucal, assim como a utilização de algoritmos de processamento e análise de imagem automáticos.



**Figura 6.9 (A) Fotografia da moldeira. (B) Imagem sagital 2D de RM obtida durante a produção da vogal /a/, com utilização da moldeira impregnada com vaselina (indicada pelas setas) e marcas de óleo de fígado de bacalhau (sinalizadas por asteriscos).**

Na Figura 6.10 apresentam-se as tarefas de processamento e análise realizadas nas imagens de RM obtidas a partir do estudo 3D usando a moldeira impregnada de vaselina e o respetivo modelo 3D reconstruído dos dentes da arcada dentária superior. Através do modelo 3D dos dentes é possível visualizar com rigor, a morfologia da raiz dentária e a posição de cada um dos dentes.





**Figura 6.10** Tarefas de processamento e análise de imagem realizadas para a extração da morfologia 3D dos dentes da arcada superior a partir de imagens de RM, usando uma moldeira impregnada com vaselina.

Em suma, é possível considerar que se trata de uma solução viável, acessível, e que pode ser utilizada quer em estudos estáticos 3D como dinâmicos, desde que seja assegurada a colaboração de especialistas em Ortodontia ou Odontologia para a correta concepção da moldeira em silicone para o indivíduo em estudo.

## 6.5 Conclusão

A exploração das várias tarefas de processamento e análise de imagem constitui uma fase primordial para a caracterização morfológica e dinâmica do

trato vocal e conseqüentemente, para um conhecimento mais aprofundado da natureza funcional da fala humana a partir das imagens de RM.

A multidiversidade de programas computacionais e das suas funcionalidades obriga necessariamente à escolha criteriosa dos métodos de análise face aos resultados pretendidos e ainda, à definição de linhas de orientação práticas para as diferentes áreas multidisciplinares da fala.

No entanto, destacaram-se algumas funcionalidades na caracterização da fala humana para fins em Medicina e em Engenharia e que expressam as suas potencialidades comprovada pela crescente divulgação, relevância e reconhecimento da comunidade científica que as mesmas têm tido.

A automatização da técnica de segmentação é um requisito fundamental para a posterior análise das imagens de RM, dada a diversidade de medidas quantitativas obtidas e que podem ser usadas na caracterização do trato vocal.

Neste âmbito porém, destaca-se como principal conclusão de que não existe uma técnica de processamento “ideal” ou preferencial, exigindo por isso o acompanhamento dos progressos nesta área e a adequação das várias ferramentas disponíveis face às particularidades do estudo.

No que respeita à identificação dos dentes nas imagens de RM, a opção pela vaselina impregnada na moldeira é viável, pouco desconfortável, podendo ser utilizada durante a fonação, nomeadamente para avaliar os sons produzidos neste ponto de articulação. Acresce ainda a possibilidade de recurso a técnicas automáticas de processamento e análise, para a reconstrução e visualização da morfologia 3D dos dentes, a partir das imagens de RM obtidas.

O futuro das técnicas de processamento e análise das imagens de RM do trato vocal passará pelo alargamento dos algoritmos integrados e suas aplicações na rotina médica diária e pelo aperfeiçoamento dos protocolos de RM. Estes avanços previstos no apoio ao diagnóstico, ao planeamento terapêutico e à terapia poderão assim ser de grande utilidade, não só para os clínicos como para os terapeutas da fala.

Tornar-se-á por fim necessária a sua standardização de forma a assegurar a interoperabilidade necessária para a sua implementação.

## **CAPÍTULO 7. ESTUDO DE CASOS: VOZ FALADA E VOZ CANTADA**

A articulação da fala é, sem dúvida, uma das atividades motoras humana frequentemente mais utilizada, servindo para as demais funções da voz falada e da voz cantada. Os sons da fala humana são, na sua maioria, vozeados ao nível da laringe e modulados pelos articuladores, havendo porém um pequeno grupo de sons que são produzidos apenas graças à articulação, ou seja, sem voz.

A fala é para muitos profissionais da voz, como professores, cantores, atores e terapeutas da fala, mais do que uma ferramenta de comunicação ou interação, dependendo desta função para o desempenho da sua profissão. O uso da voz falada é a função mais praticada pelo ser humano; no entanto, outras modalidades de voz profissional como a voz cantada ou a voz “teatral”, exigem domínios distintos e nem sempre são simples de executar. A voz falada é natural e inconsciente não necessitando de treino ou adaptações específicas ao contrário da voz profissional.

Questionamo-nos então de que modo a articulação se adequa às demais modalidades da voz ou afins? Existirão diferenças articulatórias para os diferentes contextos de voz falada e cantada?

Este capítulo ilustra uma aplicação prática concreta da utilização do protocolo de RM para estudo da fala desenvolvido neste projeto, descrevendo-se os participantes e métodos usados, terminando na apresentação e discussão dos resultados obtidos.

## 7.1 Introdução

Ao longo do desenvolvimento deste projeto, realizaram-se várias aplicações práticas, que foram resumidamente descritas no capítulo anterior, orientadas para a demonstração da utilidade das técnicas de processamento e análise de imagem consideradas e ainda, para a obtenção de resultados específicos, como medidas quantitativas do trato vocal, com relevância para as demais áreas multidisciplinares, tais como a Medicina e a Engenharia, contribuindo ainda para a Fonética, a Linguística e a Imagiologia.

Dada a extensão de aplicabilidades que advém da utilização da RM na caracterização funcional da fala, justificava-se do nosso ponto de vista, um estudo de casos, orientado para uma avaliação concreta da função da fala, em diferentes modalidades da voz, no sentido de demonstrar a aplicabilidade e potencialidades do protocolo de RM desenvolvido.

Assim, surgiram motivações pessoais e profissionais para o estudo da voz em diferentes contextos ou modalidades de voz profissional: falada, cantada e teatralizada. O objetivo principal deste estudo de casos, a seguir descrito, foi avaliar qual a adaptação da articulação em cada um dos contextos da voz profissional, intrasujeitos e intersujeitos.

Definiram-se ainda como objetivos mais específicos:

- Caracterizar a morfologia e dinâmica dos articuladores em diferentes contextos de voz;
- Extrair medidas articulatórias para avaliação das cavidades de ressonância e posição dos articuladores;
- Observar qual a adaptação da articulação nos diferentes contextos de voz profissional.

Para além do interesse para o ensino, quer da Fonética quer artístico, este estudo é também relevante do ponto de vista clínico, fornecendo linhas de orientação e conhecimentos acerca da adaptação da articulação da fala nos diferentes contextos de voz profissional.

### 7.1.1 Diferenças entre os contextos de voz profissional

De acordo com BEHLAU & REHDER (1997), os parâmetros distintivos mais comuns entre os diferentes contextos de voz profissional (falada e cantada) são: a respiração, a fonação, a ressonância e projeção de voz, a qualidade vocal, o vibrato, a articulação dos sons da fala, as pausas e a postura corporal. Além disso, acresce a influência do fator emocional na fala. Num estudo recente, LEE e colaboradores (2006) demonstraram a utilização da RM no estudo do efeito das emoções na produção da fala. Estes investigadores concluíram que na generalidade a articulação é mais ativa, em particular o modo de articulação; nomeadamente, na variação da forma e extensão do trato vocal.

A voz falada utiliza geralmente uma maior percentagem de sons não vozeados, ao contrário da voz cantada em que os sons produzidos são maioritariamente vozeados.

O canto é um fenómeno sensório-motor que requer competências físicas muito próprias e equilibradas (BUNCH, 1997). Já a voz teatralizada obriga ao domínio de ambos os contextos de voz falada e cantada, e é na voz e também no gesto que o ator cria determinada personagem.

A Tabela 7.1 resume as principais diferenças encontradas entre os dois contextos de voz profissional. Considerando as potencialidades da utilização da RM e as suas condicionantes na caracterização funcional da fala, este estudo está focado em duas das características distintivas dos contextos de voz falada *versus* cantada, ou seja, a articulação e a ressonância.

**Tabela 7.1 Características distintivas dos contextos de voz falada *versus* cantada/teatralizada (adaptado de BEHLAU & REHDER, 1997).**

Voz falada	Voz cantada/ teatralizada
Não apresenta vibrato.	O vibrato é um recurso vocal estético muito valorizado em determinados estilos de canto, como o lírico.
A identidade dos sons é mantida e a articulação é bem definida.	A mensagem a ser transmitida na voz cantada está além das palavras e, portanto, pode-se privilegiar os aspetos musicais e não os verbais (subarticulação). Esta situação é mais evidente no canto lírico.

**Tabela 7.1 Características distintivas dos contextos de voz falada versus cantada/teatralizada (adaptado de BEHLAU & REHDER, 1997), – continuação.**

Voz falada	Voz cantada/teatralizada
Necessidade de articulação dos sons de forma precisa para percepção do conteúdo verbal que se quer transmitir.	Necessidade de controlo da qualidade da voz podendo ocorrer reduções articulatórias, prolongamentos e outros recursos específicos (como o vibrato e o formante do cantor <sup>10</sup> e o formante do ator <sup>11</sup> ).
A respiração é natural havendo variações do ciclo respiratório de acordo com as emoções e a extensão das frases. A inspiração é relativamente lenta e nasal nas pausas longas e rápida e bucal durante a fala.	A respiração é treinada e os ciclos respiratórios são pré-programados de acordo com as frases musicais. Existe uma maior coordenação entre a fonação e a respiração, que também é mais precisa e específica.
A ressonância é geralmente média, ocorrendo com o trato vocal em condições naturais, sem uso particular de alguma cavidade e sem necessidade de grande projeção vocal.	A ressonância é geralmente alta e concentra-se na parte superior do trato vocal. No canto, a ressonância pode ser utilizada com forma de enriquecimento estilístico. No teatro, pode ser variável de acordo com a qualidade vocal que se pretende dar à personagem.
A qualidade vocal pode ser neutra ou com pequenos desvios mas que não são considerados sinais de disfonia.	Alguns cantores e atores podem ter desvios expressivos da qualidade vocal apresentando diferentes tipos de voz, até mesmo disfónicos, e muito bem aceites pelo público. Os cantores líricos têm um modelo de perfeição vocal que rejeita qualquer desvio. O canto coral representa uma situação intermédia.
As pausas são intrínsecas do orador.	As pausas são pré-programadas e definidas por uma série de aspetos. As pausas de hesitação não são aceitáveis no canto.
A postura é variável com mudanças constantes.	A postura é menos variável, excetuando-se quando o canto é acompanhado da performance de dança.
A velocidade e ritmo da voz são pessoais e geralmente ocorrem independentemente da consciência do orador. Para além disso, variam de acordo com as diferentes realidades profissionais, por exemplo, um professor e um locutor de rádio.	A velocidade e ritmo da voz dependem do tipo de música, harmonia, melodia e andamento utilizado pelo cantor. São usualmente controladas, ensaiadas e pré-programadas.

<sup>10</sup> O vibrato consiste na modulação repetida da frequência e/ou intensidade (*pitch*) da voz. Requer treino, produzindo um efeito estético evidente e um papel primordial conferindo à voz, uma riqueza expressiva, leveza e poder emocional. O formante do cantor corresponde ao agrupamento dos formantes 3, 4 e 5, à volta de 2500 Hz, típico em cantores masculinos e femininos contralto e é desenvolvido para projeção de voz sem amplificação acima do som de uma orquestra.

<sup>11</sup> O formante do ator corresponde ao agrupamento dos formantes 4 e 5, à volta de 3500 Hz, típico em atores masculinos com técnica vocal e é um indicador de boa qualidade a voz.

## 7.2 Metodologia

Nesta secção, são descritos os participantes e os métodos utilizados, assim como os recursos materiais e as técnicas de análise. Este estudo foi realizado mediante a autorização e disponibilização de recursos humanos e materiais da Unidade de Ressonância Magnética do Serviço de Radiologia do Hospital S. João, EPE no Porto e seguindo as normas operacionais desse serviço.

### 7.2.1 Participantes e *corpus* da fala

Da amostra fizeram parte quatro profissionais da voz, saudáveis e que voluntariamente aceitaram participar, mediante consentimento informado relativamente ao estudo a realizar (Anexo 4). Por questões de segurança e de forma a despistar eventuais contraindicações à RM, realizou-se um questionário previamente à recolha imagiológica (como descrito no Capítulo 5).

Dada a especificidade da técnica de RM e do estudo em particular, procedeu-se à informação (através do guia elaborado descrito na secção 5.1), preparação e treino prévio de cada sujeito com auxílio de um terapeuta da fala.

Assim, os sujeitos apresentavam idades compreendidas entre os 30 e 47 anos (média  $35.5 \pm 8.0$  anos) e um peso médio de  $63.0 \pm 8.3$  Kg e representavam as seguintes modalidades de voz profissional:

- Uma cantora lírica (LIR);
- Uma cantora jazz (JAZ);
- Um ator de dobragens (ATD);
- Um terapeuta da fala (TF).

Nenhum dos sujeitos apresentava historial de patologia oro-facial ou perturbação articulatória, tendo essa avaliação sido realizada durante o processo de seleção por um terapeuta da fala.

Para constituição do *corpus* da fala escolheram-se as três vogais [i a u] (transcrição fonética do IPA para o PE), pela familiaridade dos sons a qualquer um dos profissionais da voz em estudo, independentemente do tipo de voz e pela facilidade com que são produzidos garantindo quer a utilização nos estudos estáticos quer nos dinâmicos.

Acresce ainda que, uma vez que estas vogais constituem as extremidades do triângulo das vogais, ou seja, ocupando as posições mais extremas da articulação de vogais, figuram como representativas da mobilidade global do sistema.

No contexto de voz falada, as vogais foram pronunciadas utilizando uma palavra de referência em contexto articulatorio CV, usando-se a consoante oclusiva [p] por não influenciar a produção dos sons vocálicos e assim, permitir uma produção mais natural do som. As palavras de referência foram as seguintes: [papa], [pipa] e [pupa]. Acresce ainda que, para além destas vogais se caracterizarem por posições articulatorias estáveis e claramente distintas na voz falada, são sons frequentemente usados no treino vocal da voz cantada, nomeadamente no aquecimento vocal.

*a) Caracterização da modalidade de voz profissional:*

No sentido de otimizar o protocolo de RM e extrair informações morfológicas e dinâmicas comparáveis entre os três participantes: LIR, JAZ e ATD, procedeu-se previamente a um conjunto de procedimentos, a seguir descritos, para melhor caracterização de cada modalidade de voz profissional.

Antes da recolha das imagens e para determinação dos tons a utilizar no estudo estático, realizou-se uma avaliação da extensão vocal de cada cantora, ou seja, da gama de tons avaliada em semitons entre o tom mais grave do registo modal e o mais agudo do registo de falsete (MENDES et al., 2003). Neste sentido, selecionaram-se três tons localizados em diferentes pontos da extensão vocal de cada cantora: (1) tom grave, localizado a 25%; (2) tom intermédio, localizado a 50%; (3) tom agudo, localizado a 75%. Atendendo que o objetivo era recolher imagens em regiões da voz que o cantor utilizasse em performance vocal, teve-se em consideração o critério “conforto vocal” de forma a complementar a seleção dos tons. Assim, sempre que se verificou que o tom pré-definido era realizado em esforço aceitou-se um ajuste do mesmo para o tom mais próximo produzido dentro da tessitura vocal do cantor, ou seja, em “conforto vocal” (BEHLAU, 2004).

Para a caracterização da performance vocal do ator realizou-se uma recolha de dados áudio, mediante a leitura de um texto usado frequentemente em protocolos de avaliação vocal, “A história do rato Artur”, na sua voz natural e



personificada, em duas personagens distintas que o ator considerasse serem extremos vocais e que lhe fossem familiares; ou seja, usados anteriormente na sua atividade profissional. Garantidas as condições técnicas adequadas à posterior análise acústica, procedeu-se à gravação da voz de cada participante, em voz natural e em performance. De modo a caracterizar a performance dos três registos das cantoras e do ator, foi realizada por um Terapeuta da Fala, a avaliação acústica e perceptiva dos participantes no estudo, Tabela 7.2.

**Tabela 7.2 Avaliação perceptiva e acústica dos três participantes para caracterização do registo de voz.**

Modalidade de voz profissional	Avaliação perceptiva e acústica
LIR	<p>Indivíduo do género feminino, com 35 anos. <i>Mezzo/Soprano</i>.</p> <p><u>Voz falada</u>: altura tonal adequada (<math>F_0</math> média = 220 Hz), grau ligeiro de soprosidade, ressonância difusa, variações tonais adequadas (<math>F_0</math> máximo = 421 Hz, <math>F_0</math> mínimo = 142 Hz, 20 semitons).</p> <p><u>Voz cantada</u>: extensão vocal de 32 semitons (146,83 Hz – 880 Hz), canto com predominância de registo fisiológico de falsete e adaptação na transição de registos, presença de adaptação dos mecanismos de articulação e ressonância com evidência na duração das vogais e respetivas formantes, presença de prolongamento dos segmentos vocálicos em final de frase melódica e de vibrato. Ausência de formante do cantor (como esperado por se tratar de uma soprano).</p>
JAZ	<p>Indivíduo do género feminino, com 30 anos. Soprano.</p> <p><u>Voz falada</u>: altura tonal adequada (247 Hz), qualidade de fonte glótica normal, ressonância difusa, variações tonais adequadas (<math>F_0</math> máximo = 440 Hz, <math>F_0</math> mínimo = 196Hz, 15 semitons).</p> <p><u>Voz cantada</u>: extensão vocal de 37 semitons (174,61 Hz – 1396,9 Hz), canto com predominância de registo fisiológico modal havendo ausência de adaptação na transição de registos, reduzida adaptação dos mecanismos de articulação e ressonância em tonalidades mais graves, presença de vibrato. Ausência de formante do cantor (como esperado considerando o estilo de canto).</p>
ATD	<p>Indivíduo do género masculino, com 47 anos.</p> <p><u>Voz natural</u>: altura tonal aguda/adequada (134 Hz), grau ligeiro de tensão (qualidade de fonte glótica), grau moderado de ressonância faríngea; débito variável (moderado a rápido), variação de altura tonal adequada (<math>F_0</math> máximo = 345 Hz, <math>F_0</math> mínimo = 92 Hz, 24 semitons) e articulação precisa (adequada).</p> <p><u>Personagem 1</u>: altura tonal aguda/exagerada (240 Hz), grau acentuado de tensão (qualidade de fonte glótica), grau acentuado de ressonância faríngea, variação de altura tonal muito acentuada (<math>F_0</math> máximo = 240 Hz, <math>F_0</math> mínimo = 97 Hz, 26 semitons), débito rápido e articulação precisa (exagerada).</p> <p><u>Personagem 2</u>: altura tonal grave/exagerada (110 Hz), grau moderado de tensão e acentuado de rouquidão (qualidade de fonte glótica) com características de diplofonia (fonação ventricular) e uso recorrente de registo basal, grau ligeiro de ressonância hiponasal, variação tonal pouco acentuada (<math>F_0</math> máximo = 275 Hz, <math>F_0</math> mínimo = 65 Hz, 26 semitons), débito lento e articulação precisa (adequada).</p>

### 7.2.2 Equipamento e protocolo de RM

Para a aquisição das imagens utilizou-se a seguinte instrumentação: um sistema de RM Magnetom Trio (Siemens AG, Germany) com uma potência de campo de 3.0 T, e duas antenas *Phased Array*, para a emissão e captação do sinal da região da cabeça e pescoço, com 32 canais e 4 canais, respetivamente. A gravação da fala não foi possível dado o elevado ruído acústico produzido pelo sistema de RM durante o processo de aquisição.

Os sujeitos foram posicionados em decúbito dorsal e com um apoio sob os membros inferiores para maior conforto.

O protocolo de RM incluiu o estudo estático e dinâmico de modo a permitir uma caracterização morfológica das cavidades de ressonância e dos articuladores e, por fim, uma caracterização da dinâmica articulatória. Todos os sujeitos realizaram um protocolo de RM para estudo da voz falada, seguindo-se de um protocolo direcionado para (1) o estudo da voz cantada e teatralizada para as cantoras e o ator respetivamente, ou (2) um protocolo de RM para avaliação de padrões articulatórios patológicos simulados somente pelo terapeuta da fala, Tabela 7.3.

**Tabela 7.3 Tipos de estudo realizados para avaliação das diferentes modalidades de voz profissional.**

Tipo de estudo	Objetivos/ Características
<b>1. Estudo da voz falada para as três vogais</b> 2D Estático 3D Estático	Servir de referencial para os sons da voz falada de cada sujeito. Aplicados aos quatro sujeitos.
<b>2a. Estudo da voz cantada</b>  3D Estático de cada vogal em três tons 2D Dinâmico vogal /i/_arpejo <sup>12</sup> oitava 2D Dinâmico de uma melodia com a sílaba /pi/	Avaliar a articulação dos sons em contexto sustentado e em diferentes tons. Variação tonal em arpejo durante a recolha de imagens sequenciais. Variação tonal numa melodia no registo vocal do cantor. Aplicado às cantoras.
<b>2b. Estudo da voz teatralizada</b> 3D Estático para cada vogal e usando duas personagens distintas 2D Dinâmico de uma melodia com a sílaba /pi/ e usando duas personagens distintas	Avaliar as diferenças articulatórias entre as duas personagens criadas pelo ator. Variação tonal numa melodia de acordo com o registo vocal de cada personagem. Aplicado ao ator de dobragens.

<sup>12</sup> Arpejo, designa a execução sucessiva de notas tocadas uma a uma num acorde musical.

**Tabela 7.3 Tipos de estudo realizados para avaliação das diferentes modalidades de voz profissional – continuação.**

Tipo de estudo	Objetivos/ Características
<b>2c. Estudo de padrões patológicos</b> 3D Estático para avaliação de sigmatismo 3D Estático para avaliação de voz tensa 2D Estático para avaliação de escape nasal	Simulação de situações patológicas. Aplicado apenas ao terapeuta da fala.

No estudo dinâmico da voz cantada, o andamento musical foi definido através de um metrônomo de pêndulo, cujo pulsos sonoros foram transmitidos através do intercomunicador unidirecional do sistema de RM antes de cada aquisição durante cerca de 30 a 40 segundos. Os tons previamente definidos na caracterização da voz de cada cantora, foram transmitidos através de um piano e usando a mesma via de comunicação.

No arpejo, o andamento musical fixado foi de 6 tempos, permitindo assim, no estudo dinâmico por RM o agrupamento das imagens em cinco acordes crescentes da vogal /i/. Durante as aquisições, verificou-se que não era possível solicitar esta tarefa às cantoras com tons pré-definidos, devido ao ruído produzido pelo equipamento; pois o arpejo exige um controlo auditivo-vocal apurado. Por este motivo, optou-se por utilizar como referência o mesmo tom do próprio equipamento de RM, ou seja um Mi bemol da 3<sup>a</sup> oitava, facilitando a referência harmónica entre os tons a produzir e a execução da tarefa propriamente dita.

A melodia escolhida foi “parabéns a você” com letra adaptada com a sílaba /pi/ com o objetivo de avaliar simultaneamente a dinâmica articulatória e a variação tonal. A sincronização temporal da melodia foi efetuada de acordo com o ritmo do ruído produzido pelo sistema de RM durante a aquisição, correspondendo a um andamento musical de 4 tempos. Os sujeitos iniciaram a melodia após uma sequência de seis tempos de espera; ou seja, a partir da sexta imagem recolhida.

Durante as recolhas utilizou-se uma folha de registo previamente elaborada para o efeito. Os parâmetros de aquisição encontram-se resumidos na Tabela 7.4.

**Tabela 7.4 Protocolo de RM para estudo da voz falada versus cantada.**

Tipo de estudo	Parâmetros
<b>2D Estático</b>	TR = 400 ms, TE = 10 ms, FA = 160°, um corte sagital com Tk = 3 mm, FOV = 240 x 240 mm, Matriz 512 x 512, ETL = 5, resolução = 2.133 <i>pixel</i> /mm, <i>pixel size</i> 0.47 x 0.47 mm, tempo de aquisição de 8.07 s.
<b>3D Estático</b>	Slab no plano sagital com 78 mm de espessura. TR = 7.6 ms, TE = 2.87 ms, FA = 5°, Tk = 3 mm, FOV = 240 x 240 mm, Matriz 256 x 256, ETL = 1, resolução = 1.067 <i>pixel</i> /mm, <i>pixel size</i> 0.94 x 0.94 mm, total de imagens = 26, tempo de aquisição de 10.06 s.
<b>2D Dinâmico</b>	TR = 6.4 ms, TE = 2.44 ms, FA = 10°, Tk = 6 mm, FOV = 195 x 240 mm, Matriz 140 x 192, resolução = 0.800 <i>pixel</i> /mm, <i>pixel size</i> 1.25 x 1.25 mm, total de imagens = 30 e 60, num tempo de aquisição de aproximadamente 48 s.

### 7.2.3 Processamento e análise das imagens

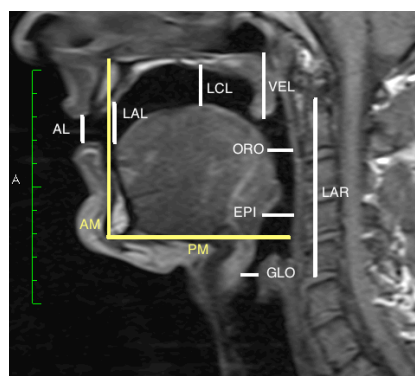
A análise das cavidades de ressonância e morfometria 3D do trato vocal foi efetuada a partir das aquisições volumétricas com recurso ao programa *ITK-SNAP*® (versão 2.1.4-rc1), cujas funcionalidades estão descritas na secção 6.2.1, realizando-se para o efeito a segmentação e reconstrução 3D do trato vocal e das duas principais cavidades de ressonância – a cavidade bucal e a faringe e, por fim, a determinação dos respetivos volumes. A dinâmica articulatória foi caracterizada a partir das imagens sequenciais do estudo dinâmico, mediante a determinação de dez medidas articulatórias sendo as primeiras cinco utilizadas nos sistemas de síntese articulatória, descritas na Tabela 7.5.

A medida EPI foi introduzida neste estudo, por ser uma estrutura “esquecida” nos estudos da produção da fala, mas cuja localização nos parece importante, principalmente pelo relacionamento com a cavidade de ressonância faríngea. Quanto às restantes medidas articulatórias ORO, LAR, AM e PM foram definidas tendo por base os estudos realizados por ECHTERNACK e colaboradores (2010, 2011) na avaliação do registo vocal em cantores a partir da RM.

**Tabela 7.5** Parâmetros articulatórios definidos para avaliação da dinâmica articulatória e respectiva medição pretendida.

Medidas articulatórias	Articuladores envolvidos e medição pretendida
<b>Abertura labial (AL)</b>	Lábios (superior e inferior). Distância vertical entre os dois lábios para avaliar a abertura labial.
<b>Localização do ápex da língua (LAL)</b>	Ápex da língua. Distância vertical do ápex da língua ao palato para avaliar a localização do ápex da língua relativamente ao palato anterior.
<b>Localização do dorso da língua (LCL)</b>	Corpo da língua. Distância vertical do dorso da língua ao palato, para avaliar a posição e altura da língua.
<b>Abertura véu do palato (VEL)</b>	Véu do palato. Distância vertical em toda a extensão do véu do palato para avaliar a sua localização e abertura.
<b>Abertura glotal (GLO)</b>	Glote. Distância horizontal entre os dois pontos de tecidos moles na glote para medir a abertura glotal.
<b>Largura da orofaringe (ORO)</b>	Orofaringe e língua. Distância horizontal entre a parede posterior da orofaringe e a língua para medir a largura desta cavidade.
<b>Localização da epiglote (EPI)</b>	Epiglote e faringe. Distância da epiglote à faringe para localização da epiglote.
<b>Altura da laringe (LAR)</b>	Glote. Distância vertical entre a linha que passa pela glote e a apófise odontoide da 2ª vértebra cervical, para medir a posição da laringe.
<b>Altura da mandíbula (AM)</b>	Sínfise mentoniana da mandíbula. Distância vertical entre a linha que passa pelo ponto mais inferior da mandíbula e o ponto mais alto do palato, para medir o deslocamento supero-inferior da mandíbula.
<b>Protrusão da mandíbula (PM)</b>	Sínfise mentoniana da mandíbula. Distância horizontal entre o ponto mais inferior da mandíbula e a parede posterior da faringe, para medir o deslocamento antero-posterior da mandíbula.

A Figura 7.1 indica as medidas articulatórias realizadas nas imagens sagitais de RM e respetivos articuladores usados na determinação das mesmas.



**Figura 7.1** Parâmetros articulatórios definidos para caracterização da dinâmica articulatória a partir das imagens sequenciais de RM.

Estas medidas articulatórias dinâmicas, foram realizadas manualmente com recurso ao programa OsiriX® (versão 3.9.4), desenvolvido por Antoine Rosset (ROSSET et al., 2004), sendo posteriormente analisadas estatisticamente através do programa SOFA® (versão 1.1.1) desenvolvido pela Paton-Simpson & Associates Ltd (PAIS), para permitir a caracterização funcional da fala em diferentes contextos de voz profissional. Para a caracterização estatística da variabilidade intersujeitos realizaram-se análises da covariância e do coeficiente de correlação de Pearson e Ró de Spearman e, por fim, Testes de Kruskal-Wallis.

#### 7.2.4 Incerteza das medições

A incerteza da medição constitui uma indicação quantitativa da qualidade dos resultados obtidos nos estudos realizados; nomeadamente, quando se pretende comparar resultados ou implementar um procedimento e, neste caso em concreto, um protocolo de RM para estudo do trato vocal.

Neste sentido, para aferir a qualidade das medições procedeu-se por um lado, à análise e determinação da incerteza da aquisição associada à qualidade das imagens adquiridas e, por outro, à determinação da confiabilidade do avaliador relativamente às medidas articulatórias realizadas na caracterização estatística dos diferentes contextos de voz profissional associada à qualidade das medições, uma vez que todas as medições foram realizadas apenas pela principal investigadora deste projeto.

Relativamente ao processo de aquisição e uma vez que se realizaram vários testes com o fantoma existente na Unidade de RM do Hospital S. João EPE, no Porto durante a fase de implementação e otimização do protocolo de RM dirigido ao trato vocal, selecionaram-se aleatoriamente quatro imagens provenientes de aquisições 2D e 3D, procedendo-se de seguida à mensuração do raio do fantoma. Considerando que o fantoma é esférico com um raio igual a 16.70 cm, a incerteza de medição obtida foi de  $16.83 \pm 0.06$  cm. A variância estimada foi de 0.003, indicando que a dispersão estatística do resultado e portanto, o afastamento da medida relativamente ao resultado esperado é , como se pode observar, muito reduzido.

No que respeita à determinação da confiabilidade do avaliador, recorreu-se ao cálculo do *Intraclass Correlation Coefficient* (ICC) tendo por base os seguintes pressupostos:

- Todas as medidas articulatórias foram efetuadas pelos mesmos observadores;
- Quatro observadores (incluindo a principal investigadora deste projeto) realizaram manualmente, com recurso ao programa OsiriX®, as dez medidas articulatórias nas primeiras dez imagens do estudo dinâmico por RM para o mesmo sujeito, correspondendo a 12.5% da amostragem.

De seguida, através do programa MedCalc® *statistical software* (versão 12.1.4) calculou-se: 1) a consistência das medições ou seja, assumindo que não é realmente necessário que dois observadores compartilhem de um significado comum quanto à escala de classificação; e 2) a concordância absoluta dos avaliadores, assumindo que as diferenças sistemáticas nas medições são relevantes; os resultados estão indicados na Tabela 7.6.

**Tabela 7.6 Determinação da confiabilidade do avaliador pelo cálculo do ICC, quanto à consistência das medições e concordância absoluta dos avaliadores.**

Número de observadores = 4		
Consistência das medições	ICC (2,4) <sup>a</sup>	Intervalo de Confiança: 95%
Medição única	0.9910	0.9780 de 0.9936
Medição média	0.9977	0.9969 de 0.9984
Concordância absoluta dos avaliadores	ICC (2,4) <sup>a</sup>	Intervalo de Confiança: 95%
Medição única	0.9905	0.9869 de 0.9933
Medição média	0.9976	0.9967 de 0.9983

Número de medidas (n) = 100

<sup>a</sup> Two-way medidas médias aleatórias (consistência/ Concordância absoluta)

Relativamente à escala de ICC, esta deve-se situar entre 0.0 e 1.0, sendo que quanto mais próximo de zero maior é a aleatoriedade das medições e o valor tenderá para a unidade quando o efeito estimado de erro é zero. Assim, observou-se que os valores de ICC obtidos em ambos os parâmetros avaliados, a consistência e concordância absoluta, situaram-se acima de 0.99, o que significa que a força da concordância ou confiabilidade do avaliador é quase perfeita (MCBRIDE, 2005).

## 7.3 Resultados e discussão

De acordo com o protocolo de RM utilizado para avaliação das diferentes modalidades de voz profissional, foram realizadas duas técnicas de aquisição distintas, primeiramente para estudo estático e, posteriormente para estudo dinâmico. Do estudo estático, resultaram três imagens 2D no plano sagital para referência dos sons vocálicos e observação da anatomia dos articuladores e 26 imagens sagitais provenientes das aquisições volumétricas 3D para cada um dos sons, ambas as aquisições (2D/3D) com a mesma espessura de corte e durante a sustentação do som produzido no contexto de voz falada. A partir deste estudo, realizou-se a caracterização morfológica das cavidades de ressonância e dos articuladores.

Para o estudo dinâmico foram recolhidas cerca de 30 e 60 imagens sequenciais mediante a captação direta de imagens 2D, com vista à caracterização da dinâmica articulatória no contexto de voz cantada, em particular para avaliar a variação tonal em arpejo e, no contexto de uma melodia em especial na sucessão coerente de sons e pausas na voz do sujeito, respetivamente.

### 7.3.1 Caracterização morfológica do trato vocal

Na Figura 7.2 apresentam-se as imagens 2D dos quatro sujeitos recolhidas no contexto de voz falada. Como se pode observar, a morfologia global do trato vocal para cada som é condicionada maioritariamente pela língua; no entanto, é possível destacar algumas características distintivas e que poderão dever-se aos domínios distintos de cada profissional da voz, nomeadamente:

- Todos os sujeitos apresentam uma epiglote saliente;
- O véu do palato é claramente mais saliente nos sujeitos JAZ, ATD e TF;
- No sujeito LIR observa-se uma cavidade bucal e faringe amplas neste plano de corte, assim como uma ligeira elevação da laringe, comparativamente aos outros sujeitos.



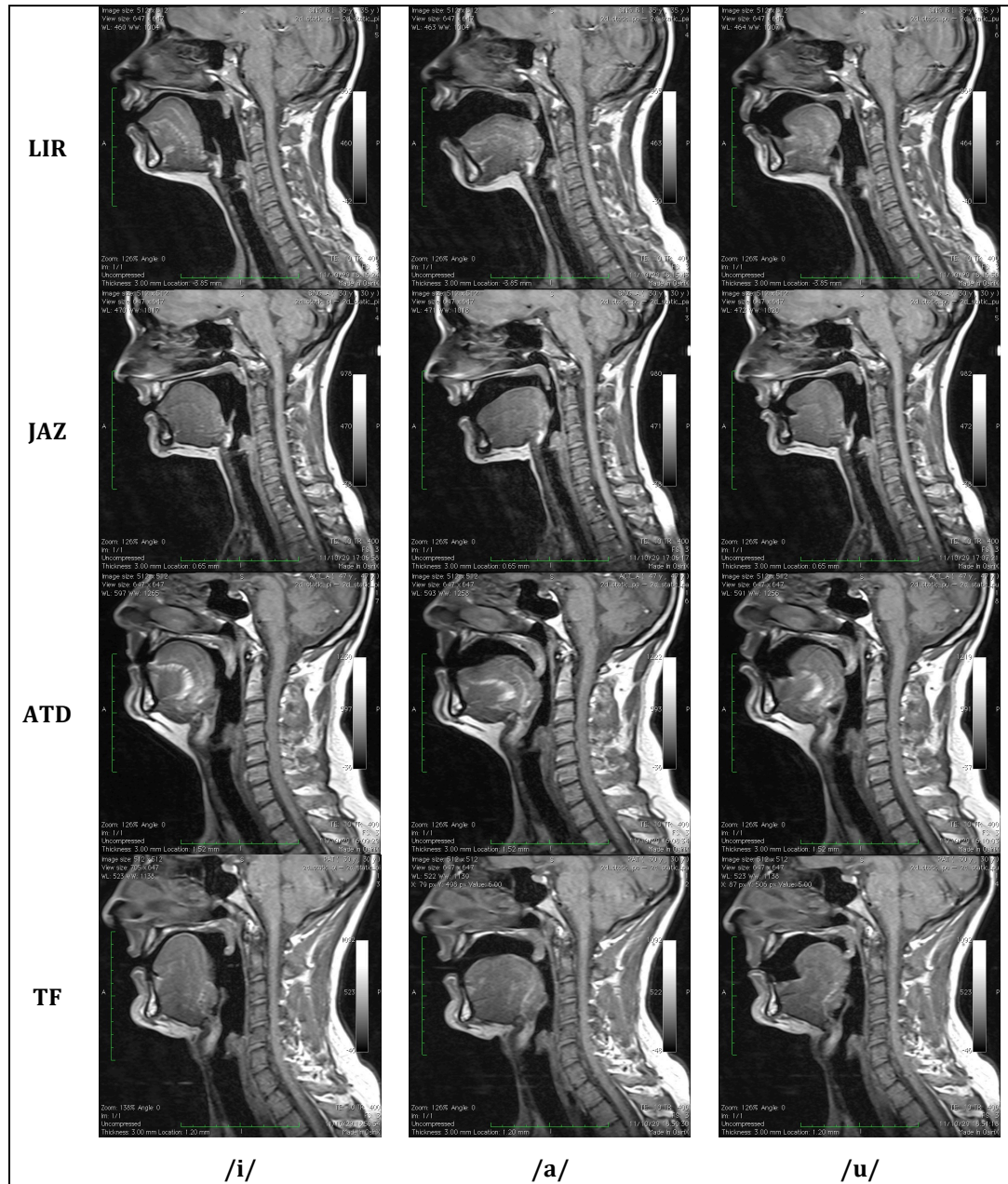
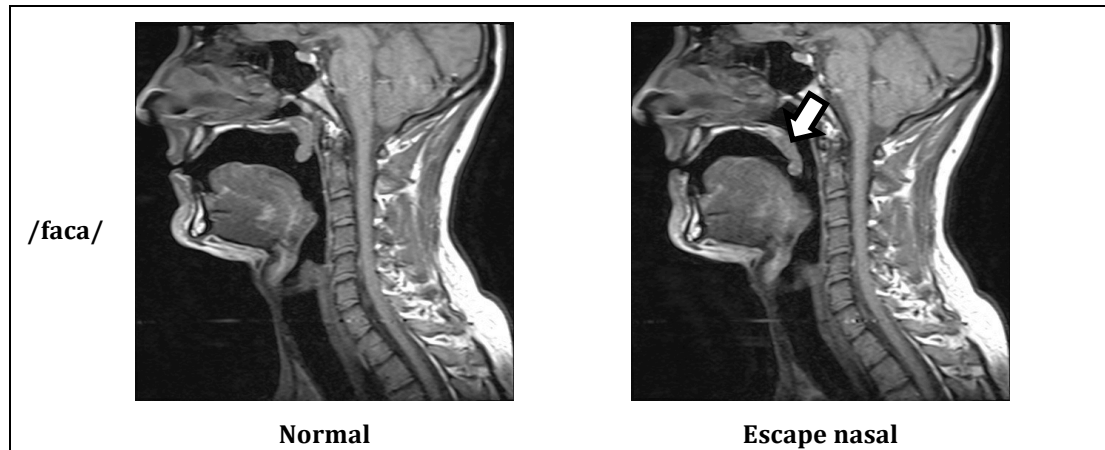


Figura 7.2 Imagens sagitais 2D de RM obtidas durante a sustentação de três vogais, no contexto de voz falada de cada um dos sujeitos.

Na situação patológica de “escape nasal”, obtiveram-se duas imagens sagitais, durante a sustentação da sílaba /fa/ da palavra /faca/, a primeira no contexto “normal” do sujeito e a segunda, mediante a simulação dessa patologia, Figura 7.3.



**Figura 7.3** Estudo estático 2D do padrão patológico “escape nasal” por RM, na palavra /faca/, observando-se a abertura do véu do palato (indicado pela seta representada) provocando a saída dos sons pela cavidade bucal e nasal, e ainda a modificação da morfologia global do trato vocal em resultado dessa alteração.

As imagens 2D obtidas permitem com elevado detalhe morfológico e num tempo de aquisição muito curto, a observação de todo o trato vocal, possibilitando assim a deteção desta situação patológica, uma vez que se tratam de sons orais e que no contexto sustentado não seria de esperar a abertura do véu do palato.

No que respeita ao estudo estático 3D, após segmentação e reconstrução das principais cavidades de ressonância e de todo o trato vocal, determinaram-se os respetivos volumes associados a cada vogal e para cada um dos sujeitos, Tabela 7.7.

Neste contexto de voz, foi notória a variabilidade nos volumes estimados do trato vocal para cada som intersujeitos, destacando-se no entanto uma tendência comum para a vogal /u/ cujos volumes medidos foram consideravelmente maiores para todos os sujeitos por comparação com as outras vogais. Seguiram-se maiores volumes do trato vocal para a vogal /i/ e, por fim, a vogal /a/, à exceção do sujeito LIR, em que se verificou o oposto.

**Tabela 7.7 Volumes 3D do trato vocal e das principais cavidades de ressonância durante a produção das três vogais no contexto da voz falada de cada um dos sujeitos.**

Voz Falada	Sujeitos			
	1	2	3	4
	LIR	JAZ	ATD	TF
<b>Vogal /a/ volumes (mm<sup>3</sup>)</b>				
<b>Trato vocal</b>	<b>48958.6</b>	<b>28779.8</b>	<b>41320.0</b>	<b>45275.1</b>
<b>Cavidade bucal</b>	20542.7	23696.2	29589.2	29077.8
<b>Faringe</b>	9858.7	5476.5	12980.6	18617.9
<b>Vogal /i/ volumes (mm<sup>3</sup>)</b>				
<b>Trato vocal</b>	<b>36864.0</b>	<b>35303.0</b>	<b>45080.0</b>	<b>49557.2</b>
<b>Cavidade bucal</b>	16447.8	5637.3	12070.9	4176.6
<b>Faringe</b>	28887.2	30216.8	55455.4	43851.3
<b>Vogal /u/ volumes (mm<sup>3</sup>)</b>				
<b>Trato vocal</b>	<b>69554.0</b>	<b>47518.9</b>	<b>71331.1</b>	<b>76188.1</b>
<b>Cavidade bucal</b>	21281.0	17328.5	33652.4	43416.2
<b>Faringe</b>	33288.6	22451.7	35553.5	40070.2

Quanto aos volumes das principais cavidades de ressonância verificou-se o seguinte:

- A vogal /a/ acarretou volumes maiores da cavidade bucal, enquanto que na vogal /i/ os volumes foram maiores na faringe, para todos os sujeitos;
- Para a vogal /u/ ainda que o volume do trato vocal tenha sido maior para qualquer um dos sujeitos, para o sujeito TF foi resultado de um maior volume na cavidade bucal, enquanto que nos restantes sujeitos, a faringe apresentou volumes mais elevados.

Considerando o contexto de voz cantada, os resultados estão indicados na Tabela 7.8. Numa análise comparativa, verificou-se que os volumes medidos, tanto do trato vocal como da cavidade bucal para ambos os sujeitos e para a vogal /a/, foram claramente superiores a 45%; ou seja, cerca do dobro do volume, em qualquer uma das variações tonais relativamente ao contexto de voz falada.

No que respeita à vogal /u/, as diferenças entre os volumes do trato vocal no contexto de voz falada *versus* cantada foram para qualquer um dos tons superiores ao volume de referência para o sujeito LIR, enquanto que para o sujeito JAZ foram inferiores ao volume de referência.

Para ambos os sujeitos, foi na produção da vogal /a/ e no tom agudo onde as diferenças entre os volumes do trato vocal foram maiores, verificando-se um

aumento do volume de aproximadamente o dobro da sua dimensão por comparação ao contexto de voz falada.

**Tabela 7.8 Comparação dos volumes 3D do trato vocal e das principais cavidades de ressonância durante a produção das três vogais nos contextos da voz falada *versus* cantada.**

		Sujeitos			
		JAZ		LIR	
		Volumes do trato vocal (mm <sup>3</sup> )			
Vogal /a/	Voz falada	28779.8		48958.6	
	Tom grave	46949.4	Dif. 63.13%	87283.3	Dif. 78.28%
	Tom intermédio	42089.9	Dif. 46.25%	85748.7	Dif. 75.15%
	Tom agudo	75051.6	Dif. 160.78%	97076.1	Dif. 98.28%
Vogal /i/	Voz falada	35303.0		36864.0	
	Tom grave	35569.3	Dif. 0.75%	47099.7	Dif. 27.77%
	Tom intermédio	36697.9	Dif. 3.95%	37528.4	Dif. 1.80%
	Tom agudo	45475.5	Dif. 28.81%	37562.7	Dif. 1.90%
Vogal /u/	Voz falada	47518.9		69554.0	
	Tom grave	37633.9	Dif. -20.80%	94186.2	Dif. 35.41%
	Tom intermédio	32262.9	Dif. -32.11%	81432.4	Dif. 17.08%
	Tom agudo	40795.3	Dif. -14.15%	70389.8	Dif. 1.20%
		Volumes da cavidade bucal (mm <sup>3</sup> )			
Vogal /a/	Voz falada	23696.2		20542.7	
	Tom grave	38833.6	Dif. 63.88%	66205.4	Dif. 222.28%
	Tom intermédio	39012.9	Dif. 64.64%	50427.2	Dif. 145.48%
	Tom agudo	50754.2	Dif. 114.19%	74471.5	Dif. 262.52%
Vogal /i/	Voz falada	5637.3		16447.8	
	Tom grave	12284.5	Dif. 117.91%	15991.7	Dif. -2.77%
	Tom intermédio	21927.0	Dif. 288.96%	17368.1	Dif. 5.60%
	Tom agudo	30731.0	Dif. 445.14%	20039.1	Dif. 21.83%
Vogal /u/	Voz falada	17328.5		21281.0	
	Tom grave	16697.3	Dif. -3.64%	46983.7	Dif. 120.78%
	Tom intermédio	24663.9	Dif. 42.33%	50920.3	Dif. 139.28%
	Tom agudo	33341.3	Dif. 92.41%	55650.6	Dif. 161.50%
		Volumes da faringe (mm <sup>3</sup> )			
Vogal /a/	Voz falada	5476.5		9858.7	
	Tom grave	8276.7	Dif. 51.13%	5500.2	Dif. -44.21%
	Tom intermédio	8142.2	Dif. 48.68%	11852.0	Dif. 20.22%
	Tom agudo	5687.4	Dif. 3.85%	16776.9	Dif. 70.17%
Vogal /i/	Voz falada	30216.8		28887.2	
	Tom grave	23965.1	Dif. -20.69%	36835.0	Dif. 27.51%
	Tom intermédio	19630.4	Dif. -35.03%	15724.8	Dif. -45.56%
	Tom agudo	10829.0	Dif. -64.16%	11253.5	Dif. -61.04%
Vogal /u/	Voz falada	22451.7		33288.6	
	Tom grave	19946.8	Dif. -11.16%	36650.4	Dif. 10.10%
	Tom intermédio	8303.0	Dif. -63.02%	28107.4	Dif. -15.56%
	Tom agudo	6111.9	Dif. -72.78%	14813.1	Dif. -55.50%

Dif. - Diferença.

No que respeita às principais cavidades de ressonância do trato vocal, verificou-se o seguinte:

- As variações nos volumes de acordo com cada variação tonal provocaram, na sua maioria, o aumento do volume de referência da cavidade bucal para cada vogal;
- Na faringe, verificou-se o oposto para as vogais /i/ e /u/ sendo que as diferenças de volume levaram à redução do volume dessa cavidade para a maioria dos tons.

No contexto de voz teatralizada, observaram-se também volumes com diferenças notáveis, principalmente no que respeita à personagem 2, em que para qualquer uma das vogais produzidas, os volumes do trato vocal foram superiores a 50% da voz falada do ator e assim como no contexto de voz falada, essa diferença foi consideravelmente maior para a vogal /a/, Tabela 7.9.

**Tabela 7.9 Comparação dos volumes 3D do trato vocal e das principais cavidades de ressonância durante a produção das três vogais no contexto de voz teatralizada.**

		ATD	
		Volumes do trato vocal (mm <sup>3</sup> )	
Vogal /a/	Voz falada	41320.0	Dif.
	Personagem 1	42656.8	3.24%
	Personagem 2	107496.0	160.15%
Vogal /i/	Voz falada	45080.0	Dif.
	Personagem 1	29990.0	-33.47%
	Personagem 2	92933.7	106.15%
Vogal /u/	Voz falada	71331.1	Dif.
	Personagem 1	37589.0	-47.30%
	Personagem 2	115380.0	61.75%
		Volumes da cavidade bucal (mm <sup>3</sup> )	
Vogal /a/	Voz falada	29589.2	Dif.
	Personagem 1	27195.1	-8.09%
	Personagem 2	88854.7	200.29%
Vogal /i/	Voz falada	12070.9	Dif.
	Personagem 1	10641.8	-11.84%
	Personagem 2	22029.8	82.50%
Vogal /u/	Voz falada	33652.4	Dif.
	Personagem 1	19846.0	-41.03%
	Personagem 2	33810.6	0.47%

Dif. – Diferença.

**Tabela 7.9 Comparação dos volumes 3D do trato vocal e das principais cavidades de ressonância durante a produção das três vogais no contexto de voz teatralizada - continuação.**

		ATD	
		Volumes da faringe (mm <sup>3</sup> )	
Vogal /a/	Voz falada	12980.6	Dif.
	Personagem 1	11738.7	-9.57%
	Personagem 2	27282.1	110.18%
Vogal /i/	Voz falada	55455.4	Dif.
	Personagem 1	17499.9	-68.44%
	Personagem 2	83272.8	50.16%
Vogal /u/	Voz falada	35553.5	Dif.
	Personagem 1	25702.7	-27.71%
	Personagem 2	76804.9	116.03%

Dif. - Diferença.

Os volumes medidos para a personagem 1 foram menores para qualquer uma das vogais por comparação à voz falada do ator; ou seja, a ressonância das cavidades diminuiu e as diferenças não foram tão expressivas. Nesta personagem, o volume da cavidade bucal diminuiu principalmente durante a produção da vogal /u/, enquanto que o volume da faringe foi consideravelmente menor durante a produção da vogal /i/.

No que respeita à vogal /a/ todos os volumes medidos tanto do trato vocal como de qualquer uma das suas cavidades de ressonância foram superiores a 110%, no contexto de voz do personagem 2, tendo sido a ressonância da cavidade bucal a mais expressiva.

A Tabela 7.10 apresenta os volumes estimados para o sujeito TF durante a simulação de dois padrões patológicos: voz tensa e sigmatismo. Quanto à voz tensa, os volumes do trato vocal e principais cavidades de ressonância foram claramente inferiores e mais expressivas para a vogal /u/; ou seja, condicionando a redução dos volumes comparativamente à voz falada do sujeito.

No padrão patológico de voz tensa, as diferenças entre os volumes medidos durante a produção da vogal /a/ foram menores; ou seja, existiu uma redução do volume de ressonância por comparação à voz falada do sujeito.

No caso do sigmatismo interdental, os volumes medidos para qualquer uma das cavidades de ressonância e para o trato vocal foram aproximadamente o dobro dos volumes obtidos na voz natural do sujeito. Já para o padrão patológico

de sigmatismo lateral, as diferenças foram pequenas, excetuando-se para a cavidade bucal, onde se verificou uma redução considerável do volume.

**Tabela 7.10 Comparação dos volumes 3D do trato vocal e das principais cavidades de ressonância durante a produção das três vogais em dois padrões patológicos.**

		TF	
		Volumes do trato vocal (mm <sup>3</sup> )	
Vogal /a/	Voz falada	45275.1	Dif.
	Voz tensa	41502.0	-8.33%
Vogal /i/	Voz falada	49557.2	Dif.
	Voz tensa	41093.3	-17.08%
Vogal /u/	Voz falada	76188.1	Dif.
	Voz tensa	48080.6	-36.89%
		Volumes da cavidade bucal (mm <sup>3</sup> )	
Vogal /a/	Voz falada	29077.8	Dif.
	Voz tensa	26193.2	-9.92%
Vogal /i/	Voz falada	4176.6	Dif.
	Voz tensa	6649.8	59.22%
Vogal /u/	Voz falada	43416.2	Dif.
	Voz tensa	22422.7	-48.35%
		Volumes da faringe (mm <sup>3</sup> )	
Vogal /a/	Voz falada	18617.9	Dif.
	Voz tensa	16163.1	-13.19%
Vogal /i/	Voz falada	43851.3	Dif.
	Voz tensa	37734.1	-13.95%
Vogal /u/	Voz falada	40070.2	Dif.
	Voz tensa	20624.4	-48.53%
		Volumes do trato vocal (mm <sup>3</sup> )	
/saco/	Voz normal	48673.9	Dif.
	Sigm_interdental	74060.2	52.16%
	Sigm_lateral	51919.7	6.67%
		Volumes da cavidade bucal (mm <sup>3</sup> )	
/saco/	Voz normal	12922.6	Dif.
	Sigm_interdental	18926.4	46.46%
	Sigm_lateral	4601.1	-64.39%
		Volumes da faringe (mm <sup>3</sup> )	
/saco/	Voz normal	41676.0	Dif.
	Sigm_interdental	68726.1	64.91%
	Sigm_lateral	52510.3	26.00%

Dif. - Diferença; Sigm - Sigmatismo.

Em suma, deste estudo estático 3D por RM e considerando quer os diferentes contextos de voz profissional quer o estudo dos padrões patológicos, numa análise intersujeitos observaram-se tendências comuns para a vogal /a/. Assim, e a título ilustrativo, a Figura 7.4 permite observar as diferenças morfológicas

encontradas; nomeadamente, o expressivo volume da cavidade bucal nos contextos de voz cantada/ teatralizada, por oposição ao maior volume faríngeo que se verifica no sigmatismo.

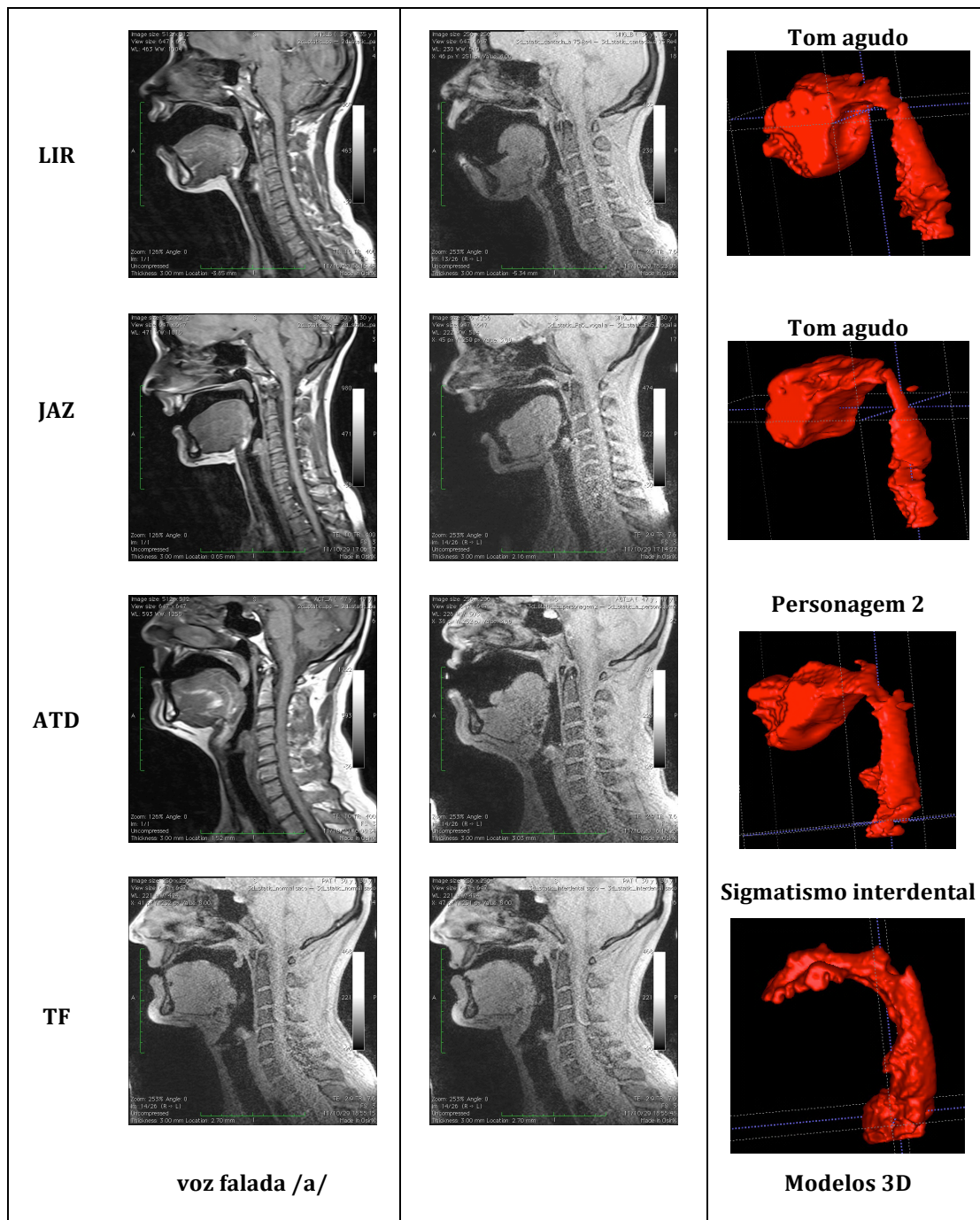


Figura 7.4 Imagens sagitais 2D de cada um dos sujeitos obtidas durante a produção da vogal /a/ nos contextos de voz falada *versus* cantada/teatralizada e no padrão patológico de sigmatismo e respetivos modelos 3D.



Não se encontrou na literatura revista nenhum estudo acerca da avaliação volumétrica das cavidades de ressonância do trato vocal a partir da RM, na caracterização de diferentes contextos de voz profissional. A extração de volumes foi somente descrita por PROCTOR e colaboradores (2010), para caracterização da cavidade faríngea durante a sustentação de consoantes fricativas. O que permite aferir da importância das observações obtidas com o estudo realizado.

### 7.3.2 Caracterização da dinâmica articulatória

O estudo da voz cantada de acordo com o registo vocal de cada profissional da voz (LIR, JAZ e ATD) consistiu na análise da variação tonal para a vogal /i/ em arpejo apenas para as cantoras e por fim, na variação dessa mesma vogal num contexto de melodia, no tom “habitual” das cantoras e para as duas personagens criadas pelo ator, a partir de imagens sequenciais de RM.

Relativamente à análise das imagens de RM em arpejo, os resultados estão apresentados nas Tabelas 7.11 e 7.12. Numa análise comparativa das medidas articulatórias obtidas independentemente da variação tonal, ou seja, considerando a média e desvio padrão das medidas extraídas das 30 imagens de RM, através da análise do coeficiente de correlação de Pearson da Tabela 7.11, foi possível concluir que:

- A largura da orofaringe é a medida que apresenta maior variabilidade intra e intersujeitos;
- Relativamente à tendência dos resultados, para ambos os sujeitos, verifica-se que é comum para a maioria das medidas, à exceção da abertura glotal e altura da laringe, cuja negatividade da covariância indica uma tendência divergente dos resultados ainda que a relação não seja forte;
- Existe uma associação linear alta e significativa entre sujeitos para a medida articulatória LAL,  $r = .82$ ,  $\rho < .001$ ; verificando-se assim que 67% da variação na medida LAL é explicada pelo número de imagens recolhidas;

- As medidas articulatórias VEL ( $r = .67$ ), ORO ( $r = .63$ ) e AM ( $r = .61$ ) estão moderadamente associadas e cujas associações são estatisticamente significativas (todas com  $\rho < .001$ ), o que significa que aumentam de acordo com o número de imagens recolhidas.

**Tabela 7.11 Análise quantitativa das medidas articulatórias (independentemente da variação tonal) em arpejo para as duas cantoras a partir das 30 imagens de RM obtidas.**

Medida	Média e desvio padrão (cm)	Valor mínimo (cm)	Valor máximo (cm)	Covariância	Correlação (r)
AL_LIR	1.06 ± 0.37	0.34	2.12	0.03	$r = .17$
AL_JAZ	1.04 ± 0.44	0.15	1.69		$\rho = .370$
LAL_LIR	0.96 ± 0.39	0.49	1.95	0.15	$r = .82$
LAL_JAZ	0.77 ± 0.49	0.19	1.85		$\rho < .001^*$
LCL_LIR	0.85 ± 0.41	0.26	1.80	0.08	$r = .57$
LCL_JAZ	0.82 ± 0.27	0.39	1.34		$\rho = .001$
VEL_LIR	1.71 ± 0.19	1.39	2.09	0.05	$r = .67$
VEL_JAZ	1.05 ± 0.41	0.49	2.15		$\rho < .001^*$
ORO_LIR	2.28 ± 0.67	1.23	3.02	0.20	$r = .63$
ORO_JAZ	2.24 ± 0.50	1.00	2.82		$\rho < .001^*$
EPI_LIR	1.73 ± 0.32	1.23	2.62	0.01	$r = .19$
EPI_JAZ	1.66 ± 0.24	1.15	1.97		$\rho = .310$
GLO_LIR	1.43 ± 0.27	0.90	2.18	-0.01	$r = -.14$
GLO_JAZ	1.10 ± 0.23	0.77	1.78		$\rho = .452$
LAR_LIR	5.72 ± 0.25	5.00	6.10	-0.01	$r = -.10$
LAR_JAZ	6.57 ± 0.38	6.05	7.43		$\rho = .588$
AM_LIR	7.20 ± 0.32	6.80	7.85	0.09	$r = .61$
AM_JAZ	6.14 ± 0.46	5.29	6.83		$\rho < .001^*$
PM_LIR	7.75 ± 0.23	7.06	8.14	0.01	$r = .24$
PM_JAZ	7.44 ± 0.20	7.02	7.83		$\rho = .198$

*r* – coeficiente de correlação de Pearson entre cada sujeito para a mesma medida articulatória ( $-1 \leq r \leq +1$ ).  
 N = 30; \* A diferença é estatisticamente significativa pois o valor-  $\rho < .001$ .

Numa análise comparativa das medidas articulatórias obtidas considerando as cinco variações tonais executadas pelas cantoras, Tabela 7.12, ou seja, considerando a média e desvio padrão das medidas extraídas de cada conjunto de 6 imagens de RM, observou-se uma maior discriminação dos resultados e consequentemente da correlação.

Num estudo intersujeitos, através da análise do coeficiente de correlação de Spearman, foi possível observar que:

- Existe uma associação linear positiva perfeita para as variáveis LCL e VEL e estatisticamente significativa ( $r_s=1.00$ ,  $p < .01$ ), o que significa que para ambos os sujeitos, estas medidas articulatórias aumentam de acordo com aumento do tom de voz;
- Para a variável LAL existe uma associação muito forte e estatisticamente significativa ( $r_s=.90$ ,  $p < .05$ ) comparativamente às demais medidas articulatórias pois o coeficiente Ró de Spearman é próximo de 1 (um), o que significa que esta medida articulatória aumenta de acordo com aumento do tom de voz; verificando-se assim que 81% da variação na medida LAL é explicada pela variação tonal;
- Para a variável ORO existe uma associação muito forte ( $r_s=.80$ ) comparativamente às demais medidas articulatórias verificando-se assim que 64% da variação na medida ORO é explicada pela variação tonal; no entanto, esta associação não é estatisticamente significativa;
- Existe uma relação positiva moderada para a variável AM ( $r_s=.50$ );
- As variáveis GLO e LAR apresentam uma correlação forte e negativa ( $r_s=-.70$  e  $r_s=-.60$ , respetivamente) indicando que estas medidas se movem em direções opostas para cada sujeito, pois são mais próximas de -1; esta correlação significa que estas medidas diminuem de acordo com aumento do tom de voz.

**Tabela 7.12 Análise quantitativa da variação tonal em arpejo para as duas cantoras a partir do estudo dinâmico de RM.**

Medida	Média e desvio padrão (cm)					Correlação
	TOM 1	TOM 2	TOM 3	TOM 4	TOM 5	
AL_LIR	0.93 ± 0.62	1.48 ± 0.20	1.00 ± 0.11	0.94 ± 0.13	0.93 ± 0.25	$r_s = -.21$ $\rho = .741$
AL_JAZ	0.37 ± 0.35	1.03 ± 0.29	1.05 ± 0.16	1.26 ± 0.22	1.48 ± 0.12	
LAL_LIR	0.72 ± 0.34	0.69 ± 0.08	0.84 ± 0.21	0.94 ± 0.15	1.58 ± 0.20	$r_s = .90$ $\rho = .037^*$
LAL_JAZ	0.31 ± 0.15	0.48 ± 0.31	0.64 ± 0.19	0.87 ± 0.23	1.57 ± 0.16	
LCL_LIR	0.59 ± 0.27	0.53 ± 0.05	0.66 ± 0.08	0.94 ± 0.15	1.54 ± 0.21	$r_s = 1.00^{**}$
LCL_JAZ	0.70 ± 0.11	0.53 ± 0.13	0.76 ± 0.26	0.93 ± 0.21	1.18 ± 0.10	
VEL_LIR	1.87 ± 0.16	1.76 ± 0.20	1.75 ± 0.17	1.60 ± 0.12	1.55 ± 0.14	$r_s = 1.00^{**}$
VEL_JAZ	1.52 ± 0.51	1.25 ± 0.24	0.95 ± 0.11	0.93 ± 0.16	0.60 ± 0.12	

**Tabela 7.12 Análise quantitativa da variação tonal em arpejo para as duas cantoras a partir do estudo dinâmico de RM – continuação.**

Medida	Média e desvio padrão (cm)					Correlação
	TOM 1	TOM 2	TOM 3	TOM 4	TOM 5	
ORO_LIR	1.84 ± 0.54	2.96 ± 0.08	2.71 ± 0.06	2.58 ± 0.06	1.30 ± 0.07	$r_s = .80$
ORO_JAZ	1.43 ± 0.44	2.55 ± 0.14	2.62 ± 0.20	2.44 ± 0.19	2.18 ± 0.18	$\rho = .104$
EPI_LIR	1.59 ± 0.48	1.95 ± 0.13	1.89 ± 0.36	1.68 ± 0.19	1.56 ± 0.13	$r_s = -.20$
EPI_JAZ	1.29 ± 0.19	1.60 ± 0.07	1.72 ± 0.08	1.90 ± 0.04	1.81 ± 0.11	$\rho = .747$
GLO_LIR	1.23 ± 0.40	1.53 ± 0.13	1.62 ± 0.29	1.31 ± 0.11	1.46 ± 0.20	$r_s = -.70$
GLO_JAZ	1.28 ± 0.13	1.08 ± 0.37	1.03 ± 0.19	1.12 ± 0.19	1.00 ± 0.13	$\rho = .188$
LAR_LIR	5.71 ± 0.42	5.83 ± 0.18	5.87 ± 0.13	5.58 ± 0.16	5.63 ± 0.17	$r_s = -.60$
LAR_JAZ	6.80 ± 0.30	6.29 ± 0.33	6.48 ± 0.48	6.72 ± 0.39	6.58 ± 0.22	$\rho = .285$
AM_LIR	7.00 ± 0.28	7.16 ± 0.30	7.05 ± 0.05	7.14 ± 0.28	7.65 ± 0.15	$r_s = .50$
AM_JAZ	5.56 ± 0.37	5.94 ± 0.23	6.08 ± 0.24	6.45 ± 0.21	6.70 ± 0.11	$\rho = .391$
PM_LIR	7.89 ± 0.14	7.68 ± 0.28	7.82 ± 0.18	7.64 ± 0.29	7.71 ± 0.21	$r_s = .30$
PM_JAZ	7.64 ± 0.17	7.53 ± 0.17	7.47 ± 0.12	7.30 ± 0.17	7.29 ± 0.13	$\rho = .624$
n	6	6	6	6	6	N = 5

$r_s$  – coeficiente de correlação entre cada sujeito para a mesma medida articulatória de acordo com a variação tonal ( $-1 \leq r_s \leq +1$ ).

\* A diferença é estatisticamente significativa pois o valor-  $\rho < .05$ .

\*\* A diferença é estatisticamente significativa pois o valor-  $\rho < .01$ .

Relativamente à caracterização da dinâmica articulatória no arpejo, para a cantora LIR, verificou-se que a medida articulatória:

- AL: aumenta apenas do 1º ao 2º tom, decrescendo gradualmente nos restantes tons;
- LAL e LCL: ambas aumentam apenas a partir da 3ª variação tonal;
- VEL: diminui gradualmente e a variabilidade das medições diminui com a variação tonal indicando uma maior elevação do véu do palato;
- ORO: aumenta inicialmente nos primeiros 2 tons, diminuindo gradualmente com a variação tonal;
- EPI e GLO: aumentam gradualmente até ao 3º tom e depois decrescem;
- LAR: a medida aumenta até ao 3ª tom, diminuindo a partir da quarta variação tonal, correspondendo assim à subida da laringe para os tons agudos;
- AM: vai aumentando gradualmente com a variação tonal, à exceção da transição do 2º para o 3º tom;

- PM: variação intertonal de aumento e diminuição da medida e que poderá dever-se a uma estratégia individual da cantora.

Para a cantora JAZ foi possível verificar que a medida articulatória:

- AL: aumenta gradualmente a medida e a variabilidade da medição reduz com a variação tonal;
- LAL: aumenta gradualmente com a variação tonal;
- LCL: diminui inicialmente nos dois primeiros tons, aumentando a partir do 3º tom;
- VEL: diminui gradualmente com a variação tonal indicando uma maior elevação do véu do palato. A variabilidade dos resultados para cada variação tonal é considerável mas vai diminuindo gradualmente;
- ORO: aumenta inicialmente nos primeiros 3 tons, diminuindo gradualmente nos restantes;
- EPI: aumenta sempre a distância da epiglote à faringe, diminuindo ligeiramente apenas no quinto tom;
- GLO: variação subtil intertonal de aumento e diminuição da medida;
- LAR: variação intertonal de aumento e diminuição da medida: descida e subida da laringe, respetivamente; o que poderá dever-se a uma estratégia individual da cantora.
- AM: aumenta gradualmente com a variação tonal;
- PM: diminui gradualmente com a variação tonal.

Em suma, constatou-se uma maior mutabilidade dos resultados num maior número de medidas articulatórias para a cantora LIR, o que poderá indicar a necessidade de uma maior adaptação da articulação face ao contexto de voz profissional e nesta tarefa em particular.

No que respeita à avaliação da variação da vogal /i/ num contexto de melodia, no tom “habitual” das cantoras e para as duas personagens criadas pelo ator, a partir das 60 imagens sequenciais de RM, os resultados foram agrupados considerando as quatro vozes em estudo e apenas para os três tons vocais presentes no conjunto dos 16 sons musicais recolhidos (dos 27 que compõem a melodia).

Deste modo, os 15 conjuntos imagens, constituídos por 4 imagens por cada tom da melodia, foram reordenadas em três conjuntos principais: grave, intermédio e agudo, Figura 7.5.

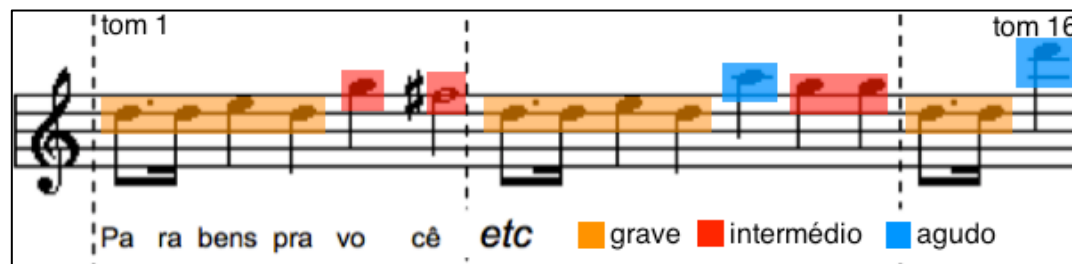


Figura 7.5 Notação musical da melodia “Parabéns a Você” e respetivo agrupamento das três variações tonais para o conjunto de sons musicais recolhidos.

Os resultados obtidos encontram-se indicados na Tabela 7.13. Numa análise intrasujeitos, pode-se constatar que:

- A maioria das medidas articulatórias apresentaram valores mais elevados nos sons em tom intermédio da melodia para a cantora LIR; excetuando-se a protrusão mandibular cujo valor medido foi maior nos sons em tom grave, e a abertura do véu do palato e altura da laringe foram superiores nos sons agudos;
- Neste contexto de melodia, a cantora JAZ apresentou na maioria das medidas articulatórias valores superiores no tom agudo; as medidas de protrusão mandibular e abertura do véu do palato, foram mais elevadas nos sons em tom grave;
- Na personagem 1 criada pelo ator de dobragens, grande parte das medidas articulatórias exibiram valores superiores na variação tonal intermédia;
- A distribuição das medidas articulatórias foi mais equilibrada para a personagem 2, dividindo-se os valores mais elevados entre as variações tonais grave e intermédia.

**Tabela 7.13** Medidas articulatórias obtidas para cada modalidade de voz dos sujeitos agrupadas pelas variações tonais: grave, intermédio e agudo, no contexto de uma melodia.

Sujeitos	Medidas	Média e desvio padrão (cm)		
		Tom Grave	Tom Intermédio	Tom Agudo
LIR	AL	0.80 ± 0.28	0.87 ± 0.34	0.62 ± 0.30
	LAL	0.55 ± 0.21	0.71 ± 0.32	0.53 ± 0.20
	LCL	0.64 ± 0.16	0.78 ± 0.20	0.77 ± 0.52
	VEL	1.64 ± 0.20	1.68 ± 0.23	1.72 ± 0.21
	ORO	2.95 ± 0.33	2.75 ± 0.75	3.00 ± 0.19
	EPI	1.83 ± 0.26	1.88 ± 0.34	1.79 ± 0.20
	GLO	1.45 ± 0.18	1.53 ± 0.19	1.50 ± 0.25
	LAR	5.60 ± 0.22	5.76 ± 0.38	5.73 ± 0.54
	AM	7.07 ± 0.19	7.18 ± 0.26	7.12 ± 0.28
JAZ	PM	7.92 ± 0.14	7.84 ± 0.17	7.88 ± 0.11
	AL	0.72 ± 0.32	0.93 ± 0.18	1.05 ± 0.26
	LAL	0.26 ± 0.15	0.40 ± 0.19	0.58 ± 0.46
	LCL	0.26 ± 0.15	0.41 ± 0.15	0.63 ± 0.49
	VEL	1.39 ± 0.19	1.23 ± 0.22	1.00 ± 0.12
	ORO	2.68 ± 0.49	2.66 ± 0.15	2.86 ± 0.28
	EPI	1.55 ± 0.11	1.60 ± 0.12	1.60 ± 0.07
	GLO	1.10 ± 0.20	1.21 ± 0.16	1.12 ± 0.25
	LAR	6.60 ± 0.23	6.63 ± 0.47	6.86 ± 0.35
ATD Personagem 1	AM	5.68 ± 0.14	5.79 ± 0.17	5.81 ± 0.15
	PM	7.53 ± 0.14	7.37 ± 0.21	7.31 ± 0.18
	AL	1.24 ± 0.52	1.42 ± 0.09	1.01 ± 0.66
	LAL	0.39 ± 0.19	0.39 ± 0.08	0.29 ± 0.21
	LCL	0.27 ± 0.15	0.29 ± 0.05	0.23 ± 0.17
	VEL	2.99 ± 0.48	3.10 ± 0.40	2.94 ± 0.34
	ORO	2.14 ± 0.29	2.24 ± 0.16	2.01 ± 0.59
	EPI	1.84 ± 0.18	1.99 ± 0.17	1.78 ± 0.37
	GLO	0.71 ± 0.48	0.56 ± 0.50	1.13 ± 0.55
ATD Personagem 2	LAR	6.26 ± 0.41	6.07 ± 0.33	6.26 ± 0.32
	AM	7.07 ± 0.37	7.26 ± 0.14	7.05 ± 0.47
	PM	8.09 ± 0.22	8.20 ± 0.19	8.27 ± 0.13
	AL	0.67 ± 0.32	0.62 ± 0.12	0.42 ± 0.40
	LAL	0.19 ± 0.20	0.12 ± 0.14	0.02 ± 0.06
	LCL	0.26 ± 0.21	0.27 ± 0.13	0.23 ± 0.22
	VEL	2.99 ± 0.17	3.05 ± 0.10	3.07 ± 0.15
	ORO	3.26 ± 0.40	3.33 ± 0.09	2.86 ± 0.79
	EPI	2.01 ± 0.15	2.05 ± 0.11	2.03 ± 0.45
GLO	0.66 ± 0.35	0.54 ± 0.29	0.85 ± 0.50	
n	LAR	7.43 ± 0.30	7.26 ± 0.24	7.12 ± 0.20
	AM	7.23 ± 0.19	7.17 ± 0.08	6.93 ± 0.40
	PM	9.19 ± 0.39	9.28 ± 0.12	8.81 ± 0.56
	n	37	16	7

De forma a avaliar a diferença entre as médias dos resultados obtidos no conjunto dos três grupos de tons produzidos pelas quatro vozes, realizaram-se Testes de Kruskal-Wallis para cada uma das medidas articulatórias, Tabela 7.14.

**Tabela 7.14 Testes de Kruskal-Wallis para avaliar a diferença entre as médias dos resultados obtidos para cada medida articulatória no conjunto dos três grupos de tons produzidos pelas modalidades de voz no contexto da melodia.**

Medida articulatória	H (3)	$\rho$
AL	8.45	.038*
LAL	8.46	.037*
LCL	7.63	.054
VEL	9.43	.024*
ORO	8.81	.032*
EPI	9.39	.025*
GLO	8.44	.038*
LAR	10.42	.015*
AM	6.26	.099
PM	10.39	.016*
N	3	

\* A diferença é estatisticamente significativa quando o valor-  $\rho < .05$ .

H(3) – Estatística do teste (número da amostra).

Após análise da Tabela 7.14 foi possível constatar que as medidas articulatórias AL ( $H(3)=8.45, p < .05$ ), LAL ( $H(3)=8.46, p < .05$ ), VEL ( $H(3)=9.43, p < .05$ ), ORO ( $H(3)=8.81, p < .05$ ), EPI ( $H(3)=9.39, p < .05$ ), GLO ( $H(3)=8.44, p < .05$ ), LAR ( $H(3)=10.42, p < .05$ ) e PM ( $H(3)=10.39, p < .05$ ) são significativamente afetadas pela variação tonal; isto é, existem diferenças nestas medidas de acordo com a variação tonal dos sons no contexto de uma melodia. As medidas articulatórias da altura da laringe (LAR) e protrusão da mandíbula (PM) apresentam maior poder estatístico de significância.

O estudo realizado por ECHTERNACH e colaboradores (2010) com o objetivo de analisar as diferenças na forma do trato vocal em sete contratenores profissionais numa tessitura correspondente à do alto, utilizado para voz masculina com extensão vocal imediatamente acima do tenor, observaram uma associação significativa de várias medidas articulatórias na transição entre o tom mais grave do registo modal e o mais agudo do registo de falsete numa escala musical ascendente e descendente. Estes autores encontraram diferenças nomeadamente na abertura labial, no estreitamento da faringe e na diminuição da altura da laringe para esta tarefa de voz cantada.

Ainda que o estudo anteriormente descrito apresente diferenças metodológicas, nomeadamente no método de medição e tipo de medidas usadas, e ainda diferenças no contexto de voz profissional e as modalidades de voz avaliadas, os resultados vão de encontro aos obtidos neste projeto para esta



tarefa de melodia, pois também se observou a alteração destas medidas articulatórias com a variação tonal.

## 7.4 Conclusão

A utilização deste protocolo de RM, na avaliação de diferentes modalidades de voz profissional, permitiu a recolha de um conjunto de dados morfológicos e dinâmicos para a caracterização de três vogais do Português Europeu. Em termos morfológicos foram reconstruídos, a partir das imagens de RM, cerca de 27 modelos 3D do trato vocal, para além dos modelos 3D das principais cavidades de ressonância. Para a caracterização da dinâmica articulatória foram realizadas, manualmente, cerca de 600 medições para a avaliação da variação tonal intra e intersujeitos no arpejo e 2400 medições para estudo no contexto de uma melodia.

As diferenças nas cavidades de ressonância observadas entre os contextos de voz falada *versus* cantada/ teatralizada foram avaliadas mediante o processamento e análise de um conjunto de imagens sagitais de RM provenientes das aquisições volumétricas 3D, culminando no cálculo dos volumes da região do trato vocal e das suas principais cavidades de ressonância.

Por comparação ao contexto de voz falada, observou-se uma tendência comum para todos os sujeitos, independentemente da modalidade de voz profissional, para um aumento considerável dos volumes estimados, nomeadamente:

- Na voz cantada: as diferenças foram particularmente maiores, em cerca de duas vezes, na produção sustentada da vogal /a/ em tom agudo para ambos os sujeitos. Na produção da vogal /u/, observaram-se estratégias articulatórias distintas, pois em qualquer tom de voz, os volumes medidos foram superiores para o sujeito LIR, enquanto que para o sujeito JAZ, foram inferiores aos volumes de referência. Na vogal /i/, os volumes medidos foram superiores no tom grave produzido pelo sujeito LIR, e no tom agudo para o sujeito JAZ.
- Na voz teatralizada: as diferenças foram particularmente maiores, ou seja, superiores a 50%, para a personagem 2 criada pelo ator. A personagem 1,

por contraste, registou valores inferiores comparados com os volumes de voz falada do ator, diminuindo assim o volume das cavidades de ressonância. De modo similar às cantoras, os volumes estimados na produção da vogal /a/ para as personagens criadas pelo ator, foram superiores, nomeadamente a personagem 2, cujo volume foi cerca de duas vezes maior que no contexto da voz falada.

O estudo estático 3D possibilitou ainda a caracterização, em contexto simulado, de dois padrões patológicos. No caso da voz tensa, observaram-se volumes claramente inferiores e mais expressivos para a vogal /u/, indicando que existe uma redução dos volumes das cavidades de ressonância e do trato vocal neste padrão patológico. Esta alteração morfológica em termos volumétricos foi igualmente observada na produção das outras vogais, ainda que não de forma tão expressiva. No caso do sigmatismo interdental, verificou-se um aumento de aproximadamente para o dobro dos volumes obtidos na voz natural do sujeito. Já no padrão de sigmatismo lateral as diferenças foram mais subtis.

O estudo dinâmico por RM da voz cantada na produção da vogal /i/, quer numa tarefa concreta somente para as cantoras, quer numa tarefa de melodia para os três sujeitos (LIR, JAZ e ATD), permitiu a análise e extração de um conjunto de medidas articulatórias, observando-se características intra e intersujeitos.

No que respeita à tarefa do arpejo, observou-se uma associação linear perfeita e estatisticamente significativa para as variáveis LCL e VEL ( $r_s = 1.00$ ,  $p < .01$ ), o que significa que em ambas as cantoras, estas medidas articulatórias aumentam com a variação tonal. Também se verificou uma associação muito forte e estatisticamente significativa para a variável LAL ( $r_s = .90$ ,  $p < .05$ ). Com tendências opostas, observou-se que as variáveis GLO e LAR apresentaram uma correlação forte e negativa ( $r_s = -.70$  e  $r_s = -.60$ ), o que indica que estas medidas diminuem com a variação tonal.

Relativamente às diferenças entre as médias dos resultados obtidos para cada medida articulatória no conjunto dos três tons produzidos, no contexto de uma melodia em quatro vozes distintas, foi possível constatar que existem oito variáveis (em dez) que são significativamente afetadas pela variação tonal.

Destas, destacaram-se a altura da laringe (LAR) e a protrusão da mandíbula (PM) que mostraram um maior poder estatístico de significância.

Concluiu-se assim que do ponto de vista morfológico e dinâmico existem diferenças de articulação da fala distintivos para cada modalidade de voz profissional, podendo assim justificar as diferenças acústico-perceptivas da voz dos sujeitos.

Adicionalmente, ficou demonstrada a adaptação da articulação nos diferentes contextos de voz profissional, sustentada em medidas morfométricas variadas, a partir do protocolo de RM desenvolvido neste projeto.



## **CAPÍTULO 8. CONCLUSÕES FINAIS E PERSPETIVAS FUTURAS**

## 8.1 Conclusões finais

Neste projeto, através do desenvolvimento de protocolos e de metodologias experimentais, foram exploradas técnicas e várias valências, quer na vertente da técnica de RM quer das técnicas computacionais de processamento e análise de imagem. A abordagem seguida impôs o contacto com várias áreas afins à da Imagiologia e necessitou da colaboração e disponibilidade de vários profissionais e investigadores de diferentes áreas complementares.

Ao longo desta publicação foram apresentados os trabalhos realizados, sustentados numa exaustiva revisão da literatura e em ensaios de campo, usando os recursos disponibilizados e atendendo sempre às considerações e objetivos iniciais estipulados para este projeto.

Do ponto de vista da estrutura, procurou-se a interligação das diferentes temáticas relacionadas com o tema central da Dissertação, sintetizando-se as principais considerações de cada capítulo e realçando-se os pontos essenciais sempre dirigidos à utilização da RM para a caracterização funcional da fala.

De seguida, apresentam-se as conclusões finais do trabalho desenvolvido e que assumem agora um maior destaque.

Ficou demonstrada no segundo capítulo, a complexidade anatómica e funcional dos órgãos envolvidos na articulação da fala, assim como o papel de cada articulador na produção dos sons.

O ênfase atribuído à descrição do Português Europeu através da caracterização fonética, assim como às várias técnicas de observação e análise da dinâmica articulatória, permitem reforçar a necessidade da implementação e otimização da RM ao estudo da fala humana, assim como a identificação dos requisitos necessários para a sua utilização de forma mais efetiva.

Como ficou claro, as várias potencialidades da RM permitem sustentar a sua utilidade nos estudos da fala, principalmente por permitir reunir informação tanto morfológica como dinâmica e de uma forma não-invasiva.

No capítulo 3 foi apresentada uma descrição concisa do princípio de funcionamento e instrumentação da RM, enfatizando-se os principais requisitos e potencialidades desta técnica no estudo da fala.

Embora a RM possua especificidades técnicas muito próprias, a sua implementação e otimização para um determinado propósito clínico, exige um conhecimento rigoroso e domínio dos seus princípios físicos, de modo a garantir a qualidade das imagens obtidas e conseqüentemente, contribuir para um diagnóstico mais eficiente e preciso.

No entanto, considerando as características anatómicas e sobretudo funcionais dos articuladores, é primordial e essencial a procura de situações de compromisso. Neste estudo em particular, trata-se do compromisso entre a relação sinal-ruído, que é fundamental para o sucesso das técnicas computacionais de processamento e análise das imagens e o tempo de aquisição, que deve ser o mais reduzido possível, quer para minimizar o tempo de sustentação dos sons no estudo estático, quer para aumentar a resolução temporal da observação da dinâmica articulatória.

Na definição de um protocolo de RM dirigido ao trato vocal acresce ainda a ambivalência morfológica *versus* dinâmica, pois se por um lado os articuladores apresentam similaridades morfológicas aos órgãos do sistema músculo-esquelético, por outro, possuem propriedades funcionais análogas ao sistema cardiovascular; ou seja, os eventos fisiológicos são extremamente rápidos, requerendo uma frequência mínima de amostragem na ordem de 20 Hz para a articulação da fala (como descrito na secção 5.3).

Do ponto de vista clínico e terapêutico, a fala quando observada numa perspectiva funcional, é considerada uma função secundária em relação às funções primordiais à vida humana, como a mastigação, a deglutição, a respiração e a sucção. Deste modo, ainda que os órgãos anatómicos envolvidos nestas funções sejam comuns, a avaliação da função da fala é usualmente abordada de forma paralela. Nesse sentido, a imagem por RM reúne muitos requisitos técnicos cuja utilidade é extensiva à avaliação das demais funções e patologias associadas à fala, como roncopatias, lesões tumorais, Síndrome da Apneia Obstrutiva do Sono, bem como a outras perturbações.

No que respeita aos estudos de investigação da fala, tem-se assistido a um desenvolvimento crescente, motivado também pelos avanços tecnológicos e informáticos da Imagiologia, especialmente ao nível da RM.

É notório o interesse pelo conhecimento da produção da fala ao nível do Português Europeu, assim como, a necessidade de incrementar o rigor da avaliação deste mecanismo a partir da RM, bem como para as demais línguas faladas.

Para além dos desenvolvimentos observados para ultrapassar algumas das condicionantes da RM, outros fenómenos têm sido estudados, tais como a nasalidade, a coarticulação e mais recentemente o canto.

Do ponto de vista clínico, os estudos são exíguos, justificando-se assim a criação de um protocolo de RM dirigido à caracterização funcional da fala, de modo a permitir a sua implementação e utilização rotineira na avaliação das demais situações patológicas.

Como ficou demonstrado no capítulo 4, subsistem ainda algumas limitações não só técnicas, como as relacionadas com o processamento e a análise das imagens, como também pela carência de soluções próprias em termos de *hardware* que permitam responder às necessidades mais específicas dos estudos da fala.

A construção de um protótipo de estudo por RM, no qual figuram todas as etapas necessárias à implementação e otimização de um protocolo para uma determinada finalidade de estudo constitui, *per se*, uma importante contribuição, permitindo a futuros investigadores a definição de linhas metodológicas similares dirigidos a outras regiões anatómicas.

No estudo do trato vocal e tendo sempre como objetivo central a caracterização funcional da fala, realizaram-se vários ensaios sustentados sempre numa análise quantitativa dos resultados obtidos, culminando na definição de um protocolo de estudo fundamentado e que permite responder às necessidades tanto clínicas como computacionais.

No capítulo 5, destacaram-se principalmente os aspetos de preparação e instrução prévia dos pacientes de modo a garantir o sucesso da utilização do protocolo desenvolvido, assim como a necessidade de uma anamnese clínica rigorosa para melhor orientação do protocolo de estudo e seleção do *corpus* da fala.



No que respeita à aquisição, processamento e análise de imagens médicas por RM, como foi descrito no capítulo 6, cada etapa é geralmente desenvolvida e otimizada separadamente, exigindo um estudo aprofundado, quer das características das imagens produzidas quer das técnicas computacionais que melhor se ajustam a tais características. Do ponto de vista operacional, é importante não só um conhecimento aprofundado da morfologia e dinâmica do trato vocal, mas também sobre as potencialidades e aplicações dos algoritmos computacionais comumente existentes.

A utilização de técnicas automáticas de processamento e análise de imagem, permite a obtenção de medidas quantitativas com maior precisão e é particularmente vantajosa, quando existe um grande volume de dados para analisar, reduzindo assim o tempo, o esforço e a subjetividade do operador.

Nos estudos da fala, o recurso a técnicas manuais ainda predomina, justificando-se assim a necessidade do desenvolvimento de algoritmos adequados face às especificidades anatómicas do trato vocal; ou seja, persistem ainda algumas dificuldades, do ponto de vista do processamento e análise de imagem, quer pelo elevado ruído típico das imagens de RM, quer pela similaridade de sinal das estruturas vizinhas ao trato vocal que dificultam em particular as tarefas de segmentação.

Espera-se que os avanços nas tarefas de aquisição e reconstrução das imagens e da sua segmentação contribuam para um maior sucesso e fiabilidade do estudo da produção da fala por RM.

No capítulo 7 foi apresentado um estudo de casos para avaliação de diferentes modalidades de voz profissional e demonstrada a utilização do protocolo de RM desenvolvido de forma dirigida ao estudo da articulação da fala. Procurou-se assim a caracterização da voz de vários profissionais, desde o canto ao teatro. Ilustraram-se ainda, em contexto simulado, alguns padrões articulatorios patológicos.

Por outro lado e, no sentido de, permitir um estudo comparativo e sustentado, o protocolo de RM foi integralmente explorado, culminando num conjunto de imagens representativas, quer da morfologia quer da dinâmica do trato vocal.

Ficou demonstrado o modo como a articulação da fala se adequa às demais modalidades de voz profissional, por comparação à voz falada de cada sujeito, nomeadamente no que respeita à dinâmica dos articuladores.

Tanto no canto como no teatro, as diferenças nas cavidades de ressonância estimadas pelos volumes extraídos do estudo estático 3D realizado, apresentam uma tendência comum. Ou seja, de aumento das dimensões do trato vocal, ainda que nas cantoras seja na sua maioria determinado pela cavidade bucal, enquanto que para o ator é também determinado pela faringe.

Apesar da amostra considerada não ter uma dimensão significativa, foi possível a observação de associações entre as várias medidas articulatórias extraídas, em dois contextos dinâmicos específicos: o arpejo aplicado às cantoras e a melodia para as cantoras e o ator de dobragens.

Durante o ensaio houve a necessidade de proceder a alguns reajustes relativamente às tarefas solicitadas aos sujeitos, principalmente pelo elevado ruído produzido pelo equipamento de RM durante as aquisições. A despeito destas dificuldades, foram recolhidas um número considerável de imagens e realizadas várias tarefas de processamento e análise de imagem.

No contexto da melodia, foi possível observar as diferenças obtidas para cada medida articulatória e para as quatro vozes em estudo, ficando estatisticamente demonstrados os parâmetros articulatórios afetados pela variação tonal.

Para a obtenção destes resultados, ao longo deste projeto, foi necessário contornar várias dificuldades nomeadamente, a inexistência de linhas de orientação para a criação de protocolos e a complexidade subjacente quer à técnica de RM quer ao mecanismo da produção da fala. Procuraram-se várias soluções para ultrapassar tais dificuldades com sucesso, em particular as várias condicionantes da RM para o estudo do trato vocal, ficando por ultrapassar a possibilidade da captação acústica da fala simultânea com a aquisição das imagens de RM. Ainda que neste projeto, tenha sido garantida a correta produção dos sons em estudo mediante o recurso a sujeitos treinados e à gravação acústica em diferido.

A captação acústica da fala simultânea com a aquisição das imagens de RM será de particular interesse nos estudos dinâmicos e para confirmação dos sons produzidos, aquando da sua implementação na prática clínica.

No que concerne às técnicas computacionais de processamento e análise de imagem, para além da exploração de programas computacionais existentes, desenvolveram-se algumas soluções específicas resultantes da colaboração com outros investigadores, nomeadamente na utilização de modelos deformáveis para segmentar as imagens de RM e caracterizar e simular os sons representados.

Neste projeto, realizaram-se ainda várias investigações e experiências para identificação dos dentes nas imagens de RM, propondo-se uma solução satisfatória quer para extração da morfologia dentária quer para localização posicional dos dentes durante a produção da fala. Esta abordagem para além de permitir a determinação mais precisa da morfometria do trato vocal, abre novas perspetivas diagnósticas nas áreas de Odontologia e Ortodontia (como descrito na secção 6.4).

As várias aplicações práticas realizadas permitem assim concluir que a RM possibilita bons resultados na caracterização funcional da fala e que quando apoiada numa metodologia e análise quantitativa consistente, demonstra uma clara utilidade e adequação.

## **8.2 Perspetivas futuras**

Ainda que as linhas de investigação desenvolvidas neste projeto apresentem aplicabilidade imediata e inclusivamente não exista, em termos clínicos, outro método não invasivo para caracterização funcional da fala, podem ser perspetivados alguns aperfeiçoamentos e desafios futuros, relacionados, por um lado, com a técnica de RM e novos algoritmos computacionais de processamento e análise de imagens e, por outro, com a implementação do protocolo desenvolvido em situações patológicas concretas.

No que respeita à imagem por RM, perspetivam-se progressos no processo de aquisição das imagens, através do desenvolvimento de sequências de pulso mais

rápidas, permitindo assim um incremento na resolução temporal do equipamento. Na produção da fala, esta evolução permitirá reforçar as mais-valias da RM na caracterização da dinâmica articulatória e aumentar a resolução das imagens, nomeadamente reduzindo os artefactos de movimento.

Esperam-se ainda inovações ao nível do desenho das antenas, quer através da sua miniaturização quer do desenho e geometria adaptada a determinadas regiões, conhecidas do ponto de vista técnico como antenas “dedicadas”, resultando na melhoria da relação sinal-ruído.

Estes avanços da técnica de RM, possibilitarão no futuro um posicionamento da RM na avaliação diagnóstica de eventos dinâmicos e/ou funcionais e ainda, a nível terapêutico, no apoio aos clínicos em melhores procedimentos de intervenção.

No que respeita ao progresso da criação artificial da fala por métodos acústico-articulatórios, este dependerá do conhecimento tridimensional e dinâmico do trato vocal para uma simulação acústica mais exata, contribuindo também para tal os aperfeiçoamentos da técnica de RM.

O alargamento dos algoritmos integrados de processamento e análise de imagens e a extensibilidade a outras aplicações médicas, vão permitir aumentar a acuidade diagnóstica dos métodos de imagem, aumentando a confiabilidade dos avaliadores e consequentemente dando melhor resposta às várias necessidades clínico-terapêuticas.

A otimização dos resultados do estudo dinâmico por RM depende do desenvolvimento de algoritmos automáticos, de modo a permitir uma maior rapidez na análise e extração de medidas quantitativas e consequentemente a sua aplicação num maior número de participantes.

Tornar-se-á ainda necessária no futuro realizar uma validação estatística com dimensão mais significativa, isto é, usando amostras de maior dimensão, de modo a comprovar a eficiência dos resultados para as diferentes áreas de aplicação descritas na secção 3.4.3.

Quanto aos desenvolvimentos dos estudos da produção da fala por RM, perspectivam-se melhorias, quer das técnicas de aquisição, nomeadamente dinâmicas, quer no que respeita às principais condicionantes relacionadas com a RM: a identificação dos dentes e a captação acústica da fala. Prevê-se assim, o

desenvolvimento de novas metodologias para extração da morfologia dentária e a utilização e maior desenvolvimento de microfones óticos e de sistemas de cancelamento de ruído, que permitam a recolha sincronizada das imagens de RM e da voz durante a produção da fala.

No domínio da Fonética, a utilização deste protocolo de RM possibilitará um melhor entendimento da forma como os sons produzidos e fornecerá documentação imagiológica para o ensino da fonética articulatória.

Finalmente, a implementação do protocolo desenvolvido durante este projeto em situações patológicas concretas vai dotar, num futuro próximo, os clínicos e terapeutas da fala, de uma ferramenta de avaliação objetiva, consistente e não-invasiva, aumentando assim a confiabilidade, o posicionamento e o sucesso da técnica de RM neste âmbito.



## **REFERÊNCIAS BIBLIOGRÁFICAS**

- AALTO, D., MALINEN, J., PALO, P., AALTONEN, O. – Recording speech sound and articulation in MRI. In *Proc. Biodevices 2011 - International Conference on Biomedical Electronics and Devices*, Rome, Italy, 2011, p. 168-173.
- APOSTOL, L., PERRIER, P., RAYBAUDI, M., SEGEBARTH, C. – 3D geometry of the vocal tract and inter-speaker variability. In *Proceedings of the 14th International Congress of Phonetic Sciences (ICPhS99)*, San Francisco, USA, 1999, p. 443-446.
- AVILA-GARCÍA, M.S., CARTER, J.N., DAMPER, R.I. – Extracting tongue shape dynamics from magnetic resonance image sequences. *Transactions on Engineering, Computing and Technology*. Vol. 2 (2004) p. 288-291.
- BADIN, P. & SERRURIER, A. – Three-dimensional modeling of speech organs: Articulatory data and models. *IEIC Technical Report (Institute of Electronics, Information and Communication Engineers)*. Vol. 106, n.º 177 (2006) p. 29-34.
- BADIN, P., BAILLY, G., REVÉRET, L., BACIU, M., SEGEBARTH, C., SAVARIAUX, C. – Three-dimensional linear articulatory modeling of tongue, lips and face, based on MRI and video images. *Journal of Phonetics*. Vol. 30 (2002) p. 533-553.
- BADIN, P., BOREL, P., BAILLY, G., REVÉRET, L., BACIU, M., SEGEBARTH, C. – Towards an audiovisual virtual talking head: 3D articulatory modeling of tongue, lips and face based on MRI and video images. In *5th Speech Production Seminar*, Munchen, Germany, 2000, p. 261-264.
- BADIN, P., POUCHOY, L., BAILLY, G., RAYBAUDI, M., SEGEBARTH, C., LEBAS, J.F., TIEDE, M.K., VATIKIOTIS-BATESON, E., TOHKURA, Y. – Un modèle articulatoire tridimensionnel du conduit vocal basé sur des données IRM. In *Actes des 22èmes Journées d'Etude sur la Parole*, Martigny, Suisse, 1998, p. 283-286.
- BAER, T., GORE, J.C., BOYCE, S., NYE, P.W. – Application of MRI to the Analysis of Speech Production. *Magnetic Resonance Imaging*. Vol. 5 (1987) p. 1-7.
- BAER, T., GORE, J.C., GRACCO, L.W., NYE, P.W. – Analysis of vocal tract shape and dimensions using Magnetic Resonance Imaging: Vowels. *Journal of the Acoustic Society of America*. Vol. 90, n.º 2 (1991) p. 799-828.
- BEHLAU, M. – **Voz - O Livro do Especialista**. Rio de Janeiro: Livraria e Editora RevinteR, Lda, 2004.
- BEHLAU, M. & REHDER, M.I. – **Higiene Vocal Para o Canto Coral**. Rio de Janeiro: Livraria e Editora RevinteR, Lda, 1997.
- BEHRENDTS, J. & WISMULLER, A. – A segmentation and analysis method for MRI data of the human vocal tract. In *Proceedings of the Symposium on Human and Machine Perception in Acoustic and Visual Communication*, Tutzing, Germany, 2001, p. 179-189.
- BENESTY, J., SONDHI, M.M., HUANG, Y. – **Springer Handbook of Speech Processing**. Heidelberg: Springer, 2008.
- BIRKHOLZ, P. & KRÖGER, B.J. – Vocal tract model adaptation using magnetic resonance imaging. In *7th International Seminar on Speech Production (ISSP'06)*, Ubatuba, Brazil, 2006, p. 493-500.



- BRESCH, E., & NARAYANAN, S. – Region segmentation in the frequency domain applied to upper airway real-time magnetic resonance images. IEEE Transactions on Medical Imaging. Vol. 28, n.º3 (2009) p. 323-338.
- BRESCH, E., NIELSEN, J., NAYAK, K., NARAYANAN, S. – Synchronized and noise-robust. Audio recordings during realtime MRI scans. Journal of the Acoustical Society of America. Vol. 120, n.º4 (2006) p. 1791-1794.
- BUNCH, M. – **Dynamics of the Singing Voice**. 4<sup>th</sup> Edition. New York: Springer-Verlag, 1997.
- BUXTON, R.B. – **Introduction to Functional Magnetic Resonance Imaging: Principles and Technique**. 2 edition. USA: Cambridge University Press, 2009.
- CARDING, P. – **Evaluating Voice Therapy: measuring the effectiveness of treatment**. London: Whurr Publishers, 2000.
- CHAN, S. – Radiology The Importance of Strategy for the Evolving Field of Radiology. Radiology. Vol. 224, n.º3 (2002), p. 639-648.
- COOTES, T.F., TAYLOR, C.J., COOPER, D.H., GRAHAM, J. – Training models of shape from sets of examples. In *Proceedings of the British Machine Vision Conference*, Leeds, UK, 1992, p. 9-18.
- CRARY, M.A., KOTZUR, I.M., GAUGER, J., GORHAM, M., BURTON, S. – Dynamic magnetic resonance imaging in the study of vocal tract configuration. Journal of Voice. Vol. 10, n.º 4 (1996) p. 378-388.
- DANG, J. & HONDA K. – An improved vocal tract model of vowel production implementing piriform resonance and transvelar nasal coupling. In *Proceeding of Fourth International Conference on Spoken Language Processing (ICSLP'96)*, Philadelphia, USA, 1996, p. 965-968
- DEMOLIN D., HASSID S., METENS T., SOQUET A. – Real-time MRI and articulatory coordination in speech. Comptes Rendus Biologies. Vol. 325, n.º4 (2002) p. 547-556.
- DEMOLIN, D., GEORGE, M., LECUIT, V. – Coarticulation and articulatory compensations studied by dynamic MRI. In *Proceedings of 5th Eurospeech 97 Conference*, Rhodos, Greece, 1997, p. 31-34.
- DEMOLIN, D., LECUIT, V., METENS, T., NAZARIAN, B., SOQUET, A. – Magnetic Resonance Measurements of the Velum Port Opening. In *Proceedings of 5th International Conference on Spoken Language Processing (ICSLP 98)*, Sydney, Australia, 1998, p. 425-429.
- DEMOLIN, D., METENS, T., SOQUET, A. – Real time MRI and articulatory coordinations in vowels", In *Proceedings 5th Speech Production Seminar*, München, Germany, 2000, p. 86-93.
- DEMOLIN, D., METENS, T., SOQUET, A. – Three-dimensional Measurement of the Vocal Tract by MRI. In *Proceedings of the 4th International Conference on Spoken Language Processing (ICSLP 96)*, Philadelphia, USA, 1996, p. 272-275.

- DEMOLIN, D., SAMPAIO, A., METENS, T. – An MRI Study of Articulatory Compensation. In *7th International Seminar on Speech Production*, Ubatuba, Brazil, 2006, p. 467.
- DESERNO, T.M. – **Biomedical Image Processing**. 1st Edition. New York: Springer-Verlag, 2011.
- DI GIROLAMO, M., CORSETTI, A., LAGHI, A., FERONE, E., IANNICELLI, E., ROSSI, M., PAVONE, P., PASSARIELLO, R. – Assessment with magnetic resonance of laryngeal and oropharyngeal movements during phonation. *La Radiologia Medica*. Vol. 92, n.º 1-2 (1996) p. 33-40.
- DROMEY, C., NISSEN, S., NOHR, P., FLETCHER, S.G. – Measuring tongue movements during speech: Adaptation of a magnetic jaw-tracking system. *Speech Communication*. Vol. 48, n.º5 (2006) p. 463-473.
- ECHTERNACH, M., SUNDBERG, J., ARNDT, S., MARKL, M., SCHUMACHER, M. – Vocal Tract in Female Registers — A Dynamic Real-Time MRI Study. *Journal of Voice*. Vol. 24, n.º 2 (2010) p. 133-139.
- ECHTERNACH, M., TRASER, L., MARKL, M., RICHTER B. Vocal tract configurations in male alto register functions. *Journal of Voice*. Vol. 25, n.º 6 (2011) p. 270-277.
- ENGWALL, O. – A 3D tongue model based on MRI data. In *Sixth International Conference on Spoken Language Processing*, Beijing, China, 2000a, p. 1-4.
- ENGWALL, O. – A revisit to the application of MRI to the analysis of speech production— testing our assumptions. In *Proceedings of the 6th International Seminar on Speech Production*, Sydney, Australia, 2003, p. 43-48.
- ENGWALL, O. – Are static MRI representative of dynamic speech? Results from a comparative study using MRI, EPG and EMA. In *Proceedings of the 6th International Conference on Spoken Language Processing (ICSLP)*, Beijing, China, 2000b, p. 17-20.
- ENGWALL, O. – Assessing MRI measurements: Effects of sustentation, gravitation and coarticulation. In Harrington, J., & Tabain, M. (Eds.), **Speech production: Models, Phonetic Processes and Techniques**. New York: Psychology Press, 2006.
- ENGWALL, O. – From real-time MRI to 3D tongue movements. In *Proc. 8th International Conference on Spoken Language Processing (ICSLP-2004)*, Jeju Island, Korea, 2004, p. 1109-1112.
- ENGWALL, O. – Making the Tongue Model Talk: Merging MRI & EMA Measurements. In *Proceedings of 7th EuroSpeech Conference*, Aalborg, Denmark, 2001a, p. 261-264.
- ENGWALL, O. – Modeling of the vocal tract in three dimensions. *Environment*. Vol. 4, (1999) p.113-116.
- ENGWALL, O. – Synthesizing static vowels and dynamic sounds using a 3D vocal tract model. *Technology*, Citeseer (2001b) p. 1-6.
- ENGWALL, O. & BADIN, P. – An MRI study of Swedish fricatives: coarticulatory effects. In *Proceedings of the 5th Seminar on Speech Production: Models and Data*, München, Germany, 2000, p. 297-300.

- ETTEMA, S.L., KUEHN, D.P., PERLMAN, A.L., ALPERIN, N. – Magnetic resonance imaging of the levator veli palatini muscle during speech. *The Cleft Palate-Craniofacial Journal*. Vol. 39, n.º 2 (2002) p. 130-144.
- FANT, G. – **The Acoustic Theory of Speech Production**. The Hague: Mouton, 1960.
- FOLDVIK, A.K., HUSBY, O., KVAERNESS, J., NORLI, I.C., RINCK, P.A. – MRI Film of Articulatory Movements. In *Proceedings of the 1st International Conference on Spoken Language Processing (ICSLP 90)*, Kobe, Japan, 1990, p. 421-422.
- FOOD AND DRUG ADMINISTRATION – **Criteria for Significant Risk Investigations of Magnetic Resonance Diagnostic Devices**. [online]. Document issued on: July 14, 2003 [Disponível em: <http://www.imrser.org/PDF/FDA.guidance.2003.pdf>]. [Acesso em: 30/08/2009]
- FRICKE, B.L., ABBOTT, M.B., DONNELLY, L.F., DARDZINSKI, B.J., POE, S.A., KALRA, M., AMIN, R.S., COTTON, R.T. – Analysis Using Cine MRI Findings in Children with Tracheostomy Tubes. *Korean Journal of Radiology*. Vol. 8, n.º 6 (2007) p. 506-511.
- FUCHS, S., WINKLER, R., PERRIER, P. – Do Speakers' Vocal Tract Geometries Shape their Articulatory Vowel Space? In *8th International Seminar on Speech Production*, Strasbourg, France, 2008, p. 333-336.
- FUJITA, H. – An experimental study of acoustic characteristics of hypopharyngeal cavities using vocal tract solid models. *Acoustical Science and Technology*. Vol. 26, n.º 4 (2005) p. 353-357.
- GICK, B., KANGB, A.M., WHALENC, D.H. – MRI evidence for commonality in the post-oral articulations of English vowels and liquids. *Journal of Phonetics*. Vol. 30, n.º 3 (2002) p. 357-371.
- GIL, V.M. & GERALDES, C.F. – **Ressonância Magnética Nuclear: Fundamentos, Métodos e Aplicações**. Lisboa: Fundação Calouste Gulbenkian, 1987.
- GOLD, G.E. – Clinical Protocol Challenges in MSK High Field (3T and 7T). In *Proc. Intl. Soc. Mag. Reson. Med. 19*, Montréal, Canadá, 2011.
- GOLD, G.E., HAN, E., STAINSBY, J. – Musculoskeletal MRI at 3.0 T: relaxation times and image contrast. *American Journal of Roentgenology*. Vol. 183, n.º 2 (2004) p. 343-351.
- HAAKE, E.M., BROWN, R.W., THOMPSON, M.R., VENKATESAN, R. – **MRI: Physical Principles and Sequence Design**. United States of America: Wiley Blackwell, 1999.
- HARSHMAN, R.A., LADEFOGED, P. & GOLSTEIN, L. – Factor analysis of tongue shapes. *Journal of the Acoustical Society of America*. Vol. 62, n.º 3 (1977) p. 693-707.
- HASHEMI, R.H. & BRADLEY, W.G. – **MRI: The Basics**. United States of America: Lippincott Williams & Wilkins, 1997.
- HORNAK, J.P. – **The Basics of MRI** [Online Book]. Chapter 9, 1996-2007. Disponível em: <http://www.cis.rit.edu/htbooks/mri/inside.htm>. [Acesso em: 03/09/2011].

- ICHIKAWA, T., KITAHARA, F., ARAKI, T., NANBU, A., HORI, M., KUMAGAI, H., AOKI, S., FUJINO, M. – MR endoscopy: Results in Applying to the Upper Gastrointestinal and the Pancreatobiliary Tracts. In *Proc. Intl. Soc. Mag. Reson. Med* 9, 2001, p. 2043.
- KANAL, E., BARKOVICH, A.J., BELL, C., BORGSTEDT, J.P., BRADLEY, W.G., FROELICH, J.W., GILK, T., GIMBEL, J.R., GOSBEE, J., KUHN-KAMINSKI, E., LESTER, J.W., NYENHUIS, J., PARAG, Y., SCHAEFER, D.J., SEBEK-SCOUIMIS, E.A., WEINREB, J., ZAREMBA, L.A., WILCOX, P., LUCEY, L., SASS, N. – ACR guidance document for safe MR practices: 2007. *American Journal Of Roentgenology*. Vol. 188, (2007) p. 1447-74.
- KANE, A.A., BUTMAN, J.A., MULLICK, R., SKOPEC, M., CHOYKE, P. – A new method for the study of velopharyngeal function using gated magnetic resonance imaging. *Plastic and Reconstructive Surgery*. Vol. 109, n.º 2 (2002) p. 472-481.
- KASS, M., WITKIN, A., TERZOPOULOS, D. – Snakes: active contour models. *International Journal of Computer Vision*. Vol. 1, n.º 4 (1987) p. 321-331.
- KENNETH, T., INGO, T., HOFFMAN, E., STORY, B. – Three-dimensional vocal tract imaging and formant structure: Varying vocal register, pitch, and loudness. *The Journal of the Acoustical Society of America*. Vol. 109, n.º 2 (2001) p. 742-747.
- KIM, Y.-C., NARAYANAN, S.S., NAYAK, K.S. – Accelerated three-dimensional upper airway MRI using compressed sensing. *Magnetic Resonance in Medicine*. Vol. 61, n.º 6 (2009) p. 1434-1440.
- KITAMURA, T., TAKEMOTO, H., HONDA, K., SHIMADA, Y., FUJIMOTO, I., SYAKUDO, Y., MASAKI, S., KURODA, K., OKU-UCHI, N., SENDA, M. – Difference in vocal tract shape between upright and supine postures: Observations by an open-type MRI scanner. *Acoustical Science and Technology*. Vol. 26, n.º 5 (2005) p. 465-468.
- KRÖGER, B. & BIRKHOLZ, P. – Articulatory synthesis of speech and singing: State of the art and suggestions for future research. In Esposito A, Hussain A, Marinaro M (eds). *Multimodal Signals: Cognitive and Algorithmic Issues*. LNAI 5398, Springer, Berlin, (2009) p. 306-319.
- KRÖGER, B.J., WINKLER, R., MOOSHAMMER, C., POMPINO-MARSCHALL, B. – Estimation of Vocal Tract Area Function from Magnetic Resonance Imaging: Preliminary Results. In *Proceedings of the 5th Seminar On Speech Production: Models And Data*, München, Germany, 2000, p. 333-336.
- LAM, Y. & GEIGER, B. – **Head Motion Detection for MRI Trainer**. Report. Biomedical Engineering, University of Wisconsin-Madison, 2002.
- LEE, S., BRESCH, E., ADAMS, J., KAZEMZADEH, A., NARAYANAN, S. – A study of emotional speech articulation using a fast magnetic resonance imaging technique. In *Ninth International Conference on Spoken Language Processing*, Pittsburgh, Pennsylvania, 2006, p. 2-5.
- LEWIS, M.P. – **Ethnologue: Languages of the World** [online]. Sixteenth edition. Dallas, Tex.: SIL International. Disponível em: <http://www.ethnologue.com/>. [Acesso em: 03/10/2011].
- LUFKIN, R.B. – **Manual de Ressonância Magnética**. 2ª Edição. Rio de Janeiro: Guanabara Koogan, 1999.

- LUSTIG, M., DONOHO, D., PAULY, J.M. – Sparse MRI: the application of compressed sensing for rapid MR imaging. Magnetic Resonance in Medicine. Vol. 58, n.º 6 (2007) p. 1182– 1195.
- MÁDY, K. & BEER, A. – Articulatory parameters in consonant production after tumour surgery: a real-time MRI investigation. Archives of Acoustics. Vol. 32, (2007) p. 1-8.
- MÁDY, K., SADER, R., ZIMMERMANN, A., HOOLE, P., BEER, A., ZEILHOFER, H., HANNIG, C. – Use of real-time MRI in assessment of consonant articulation before and after tongue surgery and tongue reconstruction. In *Proceedings of the 4th International Speech Motor Conference*, Nijmegen, The Netherlands, 2001, p. 142–145.
- MÁDY, K., SADER, R., ZIMMERMANN, A., HOOLE, P., BEER, A., ZEILHOFER, H., HANNIG, C. – Assessment of consonant articulation in glossectomee speech by dynamic MRI. In *Proceedings of 7th International Conference on Spoken Language Processing (ICSLP)*, Denver, 2002, p. 961–964.
- MAEDA, S. – A Digital Simulation Method of the Vocal-Tract System. Speech Communication. Vol. 1, n.º 3-4 (1982) p. 199-229.
- MAEDA, S. – Improved articulatory model. Journal of the Acoustical Society of America. Vol. 84, S1 (1988) p. S146.
- MALLADI, R., SETHIAN, J.A., VEMURI, B. – Shape modeling with front propagation: A level set approach. IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 17, n.º 2 (1995) p. 158-175.
- MARGULIS, A.R. & SUNSHINE, J.H. – Radiology at the turn of the millennium. Radiology. Vol. 214, n.º 1 (2000) p. 15-23.
- MARTINS, M.R. – **Ouvir falar – Introdução à Fonética do Português**. 3ª Edição. Lisboa: Editorial Caminho, 1988.
- MARTINS, P., CARBONE, I., PINTO, A., SILVA, A., TEIXEIRA, A. – European Portuguese MRI Based Speech Production Studies. Speech Communication. Vol. 50, n.º 11-12 (2008) p. 925-952.
- MARTINS, P., OLIVEIRA, C., SILVA, A., TEIXEIRA, A. Articulatory Characteristics of European Portuguese Laterals: a 2D & 3D MRI Study. In *VI Jornadas en Tecnología del Habla and II Iberian SLTech Workshop*, Vigo, Spain, 2010, p. 33-36.
- MASAKI, S., NOTA, Y., TAKANO, S., TAKEMOTO, H., KITAMURA, T., HONDA, K. – Integrated magnetic resonance imaging methods for speech science and technology. Journal of the Acoustical Society of America. Vol. 123, n.º 5 (2008) p. 3734.
- MASAKI, S., TIEDE, M. K., HONDA, K., YASUHIRO, S., FUJIMOTO, I., NAKAMURA, Y., NINOMIYA, N. MRI-based speech production study using a synchronized sampling method. Journal of Acoustical Society of Japan E. Vol. 20, n.º5 (1999) p.375-379.
- MATHIAK, K., KLOSE, U., ACKERMANN, H., HERTRICH, I., KINCSES, W.E., GRODD, W. – Stroboscopic articulography using fast magnetic resonance imaging. International journal of language & communication disorders. Vol. 35, n.º 3 (2000) p. 419-25.

- MCAULIFFE, M.J., LALONDE, F.M., MCGARRY, D., GANDLER, W., CSAKY, K., TRUS, B.L. – Medical Image Processing, Analysis & Visualization In Clinical Research. In *Proceedings of the IEEE Computer-Based Medical Systems (CBMS)*, Bethesda, MD, USA, 2001, p. 381-386.
- MCBRIDE, G.B. – **A proposal for strength-of-agreement criteria for Lin's Concordance Correlation Coefficient**. National Institute of Water and Atmospheric Research Ltd, Hamilton, New Zealand, 2005.
- MEDCALC® statistical software** [online]. Disponível em: <http://www.medcalc.org/>. [Acesso em: 30/11/2011].
- MEDURI, S., BAZZOCCHI, M., ZUIANI, CH., FALCONE, B., BERTINO, G., MARIONI, G. – Functional MR with use of FLASH sequences in the evaluation of the phono-articulatory tract. *Magnetic Resonance Materials in Physics, Biology and Medicine*. Vol. 9, n.º 1-2 (1999) p. 5-15.
- MENDES, A.P., HOWARD, ROTHMAN, B., SAPIENZA, C., BROWN, W.S. – Effects of vocal training on the acoustic parameters of the singing voice. *Journal of Voice*. Vol. 17, n.º 4 (2003) p. 529-543.
- MOKHTARI, P., KITAMURA, T., TAKEMOTO, H., HONDA, K. – Principal components of vocal tract area functions and inversion of vowels by linear regression of cepstrum coefficients. *Journal of Phonetics*. Vol. 35, n.º 1 (2007) p.20-39
- MOORE, C.A. – The correspondence of vocal tract resonance with volumes obtained from Magnetic Resonance Images. *Journal of Speech, Language, and Hearing Research*. Vol. 35, n.º 5 (1992) p. 1009-1023.
- MORRISON, M. & RAMMAGE, L. – **The Management of Voice Disorders**. San Diego: Singular Publishing Group, Inc., 1994.
- MULDER, A. – **Human Movement Tracking Technology**. Hand Centered Studies of Human Movement Project. Technical Report. Simon Fraser University, 1994.
- NALLO, A.M., ORTENZIA, O., D'ARIENZO, M.D., COMIGLIO, D., BENASSI, M. – MRI Quality Control Tools for Procedures and Analyses. *Journal of Experimental & Clinical Cancer Research*. Vol. 25, n.º 1 (2006) p. 121-127.
- NARAYANAN, S., NAYAK, K., LEE, S., SETHY, A., BYRD, D. – An approach to real-time magnetic resonance imaging for speech production. *The Journal of the Acoustical Society of America*. Vol. 115 (2004) p. 1771-1776.
- NARAYANAN, S.S., ALWAN, A.A., HAKER, K. – An articulatory study of fricative consonants using magnetic resonance imaging. *Journal of the Acoustical Society of America*. Vol. 98, n.º 3 (1995) p. 1325-1347.
- NESSAIVER, M., STONE, M., PARTHASARATHY, V., KAHANA, Y., KOTS, A., PARITSKY, A. – Recording high quality speech during tagged Cine MRI studies using a fiber optic microphone. *Journal of Magnetic Resonance Imaging*. Vol. 23, n.º 1 (2006) p. 92-97.
- NG, I.W., ONO, T., INOUE-ARAI, M.S., HONDA, E., KURABAYASHI, T., MORIYAMA, K. – Application of MRI movie for observation of articulatory movement during a

- fricative/s/and a plosive/t. The Angle Orthodontist. Vol. 81, n.º 2 (2011) p. 237-244.
- NIETO-CASTANON, A. & GUENTHER, F.H. – Constructing speaker-specific articulatory vocal tract models for testing speech motor control hypotheses. In *14th International Congress of Phonetic Sciences (ICPhS 99)*, San Francisco, USA, 1999, p. 2271-2274.
- OKUBO, M., SUZUKI, M., HORIUCHI, A., OKABE, S., IKEDA, K., HIGANO, S., MITANI, H., et al. – Morphologic analyses of mandible and upper airway soft tissue by MRI of patients with obstructive sleep apnea hypopnea syndrome. Sleep. Vol. 29, n.º 7 (2006) p. 909-915.
- OLT, S., & JAKOB, P.M. – Contrast-enhanced dental MRI for visualization of the teeth and jaw. Magnetic Resonance in Medicine. Vol. 52, n.º 1 (2004) p. 174-176.
- OSHER, S. & SETHIAN, J.A. – Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. Journal of Computational Physics. Vol. 79, n.º 1 (1988) p. 12-49.
- PARTHASARATHY, V., PRINCE, J.L., STONE, M., MURANO, E.Z., NESSAIVER, M. – Measuring tongue motion from tagged cine-MRI using harmonic phase (HARP). Journal of the Acoustic Society of America. Vol. 121, n.º1 (2007) p. 491-504.
- PISCO, J.M. & SOUSA, L. – **Noções Fundamentais de Imagiologia**. Lisboa: Lidel Edições Técnicas, 1999.
- PROCTOR, M.I., SHADLE, C., ISKAROUS, K. – A Method of Co-registering Multiple Magnetic Resonance Imaged Vocal Tract Volumes for Fricatives. Journal of the Acoustical Society of America. Vol. 123, n.º 5 (2008) p. 3735.
- PROCTOR, M.I., SHADLE, C., ISKAROUS, K. – Pharyngeal articulation in the production of voiced and voiceless fricatives. The Journal of the Acoustical Society of America. Vol. 127, n.º 3 (2010) p. 1507-1518.
- RASBAND, W. – **ImageJ** [Online]. National Institute of Mental Health, Bethesda, Maryland, USA. Disponível em: <http://rsbweb.nih.gov/ij/index.html>. [Acesso em: 03/10/2007].
- RIBEIRO, C.M. – **Processamento Digital de Fala** [online]. Sebenta de Apoio. Disponível em: <http://www.deetc.isel.ipl.pt/comunicacoesep/disciplinas/pdf/index.html>. 2003. [Acesso em: 05/04/2009].
- ROKKAKU, M., HASHIMOTO, K., IMAIZUMI, S., NIIMI, S., AND KIRITANI, S. – Measurements of the Three-dimensional Shape of the Vocal Tract Based on the Magnetic Resonance Imaging Technique. Annual Bulletin of the Research Institute of Logopedics and Phoniatics. Vol. 20 (1986) p. 47-54.
- ROSSET, A., SPADOLA, L., RATIB, O. – OsiriX: An Open-Source Software for Navigating in Multidimensional DICOM Images. Journal of Digital Imaging. Vol. 17 (2004) p. 205-216.
- ROUSSON, M. & PARAGIOS, N. – Prior knowledge, level set representations and visual grouping. International Journal of Computer Vision. Vol. 76 (2008) p. 231-243.

- RUA, S.M. & FREITAS, D.R. – **Estudo Morfológico-dinâmico do Tracto Vocal Humano** [Dissertação]. Mestrado de Engenharia Biomédica da Faculdade de Engenharia da Universidade do Porto, Setembro, 2006a.
- RUA, S.M. & FREITAS, D.R. – Morphological dynamic imaging of human vocal tract. In *Proceedings of the Computational Modelling of Objects Represented in Images. Fundamentals, Methods and Applications (CompIMAGE)*, Coimbra, Portugal, 2006b, p. 381–386.
- SCHWAB, R.J., PASIRSTEIN, M., PIERSON, R., MACKLEY, A., HACHADOORIAN, R., ARENS, R., MAISLIN, G., PACK, A.I. – Identification of upper airway anatomic risk factors for obstructive sleep apnea with volumetric magnetic resonance imaging. *American Journal of Respiratory and Critical Care Medicine*. Vol. 168, n.º 5 (2003) p. 522-530.
- SERRURIER, A. & BADIN, P. – Towards a 3D articulatory model of velum based on MRI and CT images. *Papers in Linguistics* . Vol. 40 (2005) p. 195-211.
- SHADLE, C.H., MOHAMMAD, M., CARTER, J.N., JACKSON, P.J.B. – Multiplanar dynamic magnetic resonance imaging: new tools for speech research. In *Proceedings of the 14th International Congress of Phonetic Sciences (ICPhS)*, San Francisco, USA, 1999, p. 623-626.
- SOQUET, A., LECUIT, V., METENS, T., DEMOLIN, D. – Mid-sagittal cut to area function transformations: Direct measurements of mid-sagittal distance and area with MRI. *Speech Communication*. Vol. 36 (2002) p. 169-180.
- STONE, M., DAVIS, E., DOUGLAS, A., NESSAIVER, M., GULLAPALLI, R., LEVINE, W., LUNDBERG, A. – Modeling the motion of the internal tongue from tagged cine-MRI images. *The Journal of the Acoustical Society of America*. Vol. 109, n.º 6 (2001) p. 2974-2982.
- STORY, B.H. – Comparison of Magnetic Resonance Imaging-based vocal tract area functions obtained from the same speaker in 1994 and 2002. *The Journal of the Acoustical Society of America*. Vol. 123, n.º 1 (2008) p. 327-335.
- STORY, B.H., TITZE, I.R., HOFFMAN, E.A. – Vocal tract area functions from magnetic resonance imaging. *The Journal of the Acoustical Society of America*. Vol. 100, n.º 1 (1996) p. 537-554.
- TAKEMOTO, H. et al. – **VTAF extraction algorithm** [online]. JASA, 2006. Disponível em: [http://basil.is.konan-u.ac.jp/vtafextractor/issp2011\\_poster.pdf](http://basil.is.konan-u.ac.jp/vtafextractor/issp2011_poster.pdf). [Acesso em: 05/01/2011].
- TAKEMOTO, H. & HONDA, K. – Measurement of Temporal Changes in Vocal Tract Area Function during a continuous vowel sequence using a 3D Cine-MRI Technique. In *Proceed. 6th Int. Seminar on Speech Production*, Sydney, Australia, 2003, p. 284-289.
- TAKEMOTO, H., KITAMURA, T., NISHIMOTO, H., HONDA, K. – A method of tooth superimposition on MRI data for accurate measurement of vocal tract shape and dimensions. *Acoustical Science and Technology*. Vol. 25, n.º 6 (2004) p. 468–473.
- TAMEEM, H.Z. & MEHTA, B.V. – Solid modeling of human vocal tract using magnetic resonance imaging and acoustic pharyngometer. In *Proceed. 26th Annual International Conference of the IEEE*, San Francisco, USA, 2004, p. 5115-5118.



- TEIXEIRA A. & VAZ F. – European Portuguese nasal vowels: an EMMA study. In *Proceedings of the Eurospeech 2001*, Institut EETA, Aveiro, Portugal, 2001, P.1483-1486.
- TEIXEIRA, A., MOUTINHO, L.C., COIMBRA, R.L. – Production, acoustic and perceptual studies on European Portuguese vowels height. In *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, 2003, p. 3033-3036.
- TEIXEIRA, A., SILVA, L., MARTINEZ, R., VAZ, F. – SAPWindows – towards a versatile modular articulatory synthesizer. *Proceedings of the 2002 IEEE Workshop on Speech Synthesis*, Aveiro, Portugal, 2002, p. 31–34.
- TOGEIRO, S.M., CHAVES, C.M., PALOMBINI, L., TUFIK, S., HORA, F., NERY, L.E. – Evaluation of the upper airway in obstructive sleep apnoea. The Indian Journal of Medical Research. Vol. 131, (2010) p. 230-235.
- TUTTON, L.M. & GODDARD, P.R. – Pictorial review MRI of the teeth. The British Journal of Radiology. Vol. 75 (2002) p. 552-562.
- UNAY, D. & OZTURK, C. – Real-time tagged-MRI for Tongue Motion Analysis. In *Proc. of 1st National Congress of Magnetic Resonance*, Izmir, Turkey, 2000.
- VASCONCELOS, M.J. VENTURA, S.R., FREITAS, D.R., TAVARES, J.M.R.S. – Inter-speaker speech variability assessment using statistical deformable models from 3.0 Tesla magnetic resonance images. Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine. ISSN: 0954-4119 (print) - 2041-3033 (online), Professional Engineering Publishing, (2011a), *in press*.
- VASCONCELOS, M.J., VENTURA, S.R., FREITAS, D.R., TAVARES, J.M.R.S. – Using Statistical Deformable Models to Reconstruct Vocal Tract Shape from Magnetic Resonance Images. Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine. Vol. 224, n.º 10 (2010) p. 1153-1163.
- VASCONCELOS, M.J., VENTURA, S.R., FREITAS, D.R., TAVARES, J.M.R.S. – Towards the Automatic Study of the Vocal Tract From Magnetic Resonance Images. Journal of Voice. Vol. 25, n.º 6 (2011b) p. 732-742
- VENTURA, S.R. & FREITAS, D.R. – Modelos Morfológicos Tridimensionais por IRM do Tracto Vocal para as Principais Vogais do Português Europeu. Livro de *Actas do Congresso de Métodos Numéricos e Computacionais em Engenharia e Congresso Ibero Latino-americano sobre Métodos Computacionais em Engenharia (CMNE/CILAMCE)*, Porto, Portugal, 2007, p. 527.
- VENTURA, S.R., FREITAS, D.R., RAMOS, I.M., TAVARES, J.M.R.S. – Magnetic Resonance Imaging-based Vocal Tract Assessment: a Comparative Study. Submetido ao Computer and Speech Language, ISSN: 0885-2308, (2011a), ELSEVIER.
- VENTURA, S.R., FREITAS, D.R., RAMOS, I.M., TAVARES, J.M.R.S. – Requisitos e condicionantes da imagem por ressonância magnética no estudo da fala humana. In *Congresso de Métodos Numéricos em Engenharia - CMNE 2011*, Coimbra, Portugal, 2011b, p. 1-12.
- VENTURA, S.R., FREITAS, D.R., RAMOS, I.M., TAVARES, J.M.R.S. – Three-dimensional modeling of tongue during speech using MRI data. In *Proceedings of the 8th*

*International Symposium on Computer Methods in Biomechanics and Biomedical Engineering*, Porto, Portugal, 2008.

- VENTURA, S.R., FREITAS, D.R., TAVARES, J.M.R.S. – Application of MRI and biomedical engineering in speech production study. Computer Methods in Biomechanics and Biomedical Engineering. Vol. 12, n.º 6 (2009a) p. 671-681.
- VENTURA, S.R., FREITAS, D.R., TAVARES, J.M.R.S. – Magnetic resonance imaging of the vocal tract: techniques and applications. In *Proceedings of the 1st International Conference on Imaging Theory and Applications (IMAGAPP)*, INSTICC Press, Lisboa, Portugal, 2009b, p. 105-110.
- VENTURA, S.R., FREITAS, D.R., TAVARES, J.M.R.S. – Toward Dynamic Magnetic Resonance Imaging of the Vocal Tract During Speech Production. Journal of Voice. Vol. 25, n.º 4 (2011c) p. 511-518.
- VENTURA, S.R., VASCONCELOS, M.J., FREITAS, D.R., RAMOS, I.M., TAVARES, J.M.R.S. – Speech Articulation Assessment Using Dynamic Magnetic Resonance Imaging Techniques. In *VipIMAGE 2011 - III ECCOMAS Thematic Conference on Computational Vision and Medical Image Processing*, Olhão, Algarve, Portugal 2011d, p. 225-231.
- WESTBROOK, C. & KAUT, C. – **MRI in Practice**. 2ª Edição. Rio de Janeiro: Guanabara Koogan, 2000.
- YEHIA, H. & TIEDE, M. – A parametric three-dimensional model of the vocal-tract based on MRI data. Society. Vol. 3, (1997) p. 1619-1622.
- YOO, T.S. – **Insight into Images: Principles and Practice for Segmentation, Registration and Image Analysis**. Canada: A K Peters, 2004.
- YUSHKEVICH, P.A., PIVEN, J., HAZLETT, H.C., SMITH, R.G., HO, S., GEE, J.C., GERIG, G. – User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. Neuroimage. Vol. 31, n.º 3 (2006) p. 1116-1128.
- ZHA, D., QIAO, L., GE, Y., DENG, Z., XUE, T., LIU, Y., QIU, J. – The MRI study of upper airway on obstructive sleep apnea hypopnea syndrome. Journal of Clinical Otorhinolaryngology, Head, and Neck Surgery. Vol. 22, n.º 13 (2008) p. 593-596.

## **ANEXOS**

## ANEXO 1

**QUESTIONÁRIO**

O presente questionário visa a recolha de informação útil para a instrumentação de um protocolo de imagem por ressonância magnética de estudo do tracto vocal (para avaliação da articulação da fala) para futura aplicação quer do ponto de vista clínico (com utilidade, por exemplo, para radiologistas, otorrinolaringologistas e terapeutas da fala) como da engenharia (com interesse para síntese da fala, computação gráfica, visão computacional, entre outras).

**Identificação dos destinatários:**

- Médicos (otorrinolaringologistas e radiologistas)
- Terapeutas da fala

**Indicações clínicas possíveis:**

- Pessoas com perturbações articulatórias. Quais:
- Outras:

**Tipos de estudo por RM:**

- Estático (através da sustentação de um som durante a aquisição de imagens)  
Em que situações clínicas?
- Dinâmico (através da repetição sucessiva do mesmo som durante a aquisição de imagens)  
Em que situações clínicas?
- Tempo-real (através da aquisição directa de imagens do tracto vocal em movimento)  
Em que situações clínicas?

**Informação necessária para avaliação da articulação da fala por RM:**

- Morfológica (anatômica)  
Especifique:
  - Forma dos articuladores
  - Posição dos articuladores
  - Cavidades de ressonância
  - Outras:

- Fisiológica (dinâmica da articulação da fala)
- Contexto consoante-vogal
  - Contexto consoante-vogal-consoante
  - Palavras sem significado
  - Outras:

**Tipos de imagem relevantes para avaliação da articulação da fala:**

- Cortes sagitais (2D)
- Cortes coronais (2D)
- Cortes axiais (2D)
- Cortes oblíquos (2D)
- Modelos tridimensionais do tracto vocal
- Análise cinemática (modo cine) em 2D
- Outras:

**Corpus da fala/ finalidade:**

Sons	Tipo de estudo por RM	Finalidade
	Estático	
	Dinâmico	
	Tempo-real	
	Outro:	

**Informações adicionais relevantes:**

Muito obrigada pela sua indispensável participação e colaboração.

## ANEXO 2



FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Programa Doutoral em Engenharia Biomédica (PRODEB)

Título da tese: A Utilização da Ressonância Magnética para a Caracterização Funcional da Fala

**GUIA BÁSICO – INSTRUÇÕES DE PREPARAÇÃO E COLABORAÇÃO NO ESTUDO****A. PREPARAÇÃO****A1) Antes do estudo:**

1. Remova todas as jóias (como, colares, brincos, anéis);
2. Remova ganchos de cabelo, cintos e outros acessórios de vestuário;
3. Remova dentaduras, dentes falsos, placas dentárias removíveis;
4. Retire próteses auditivas;
5. Retire os óculos ou lentes de contacto;
6. Retire o seu relógio, telemóvel, cartões bancários e todos os outros cartões com banda magnética;
7. Retire todos os objectos metálicos (piercing, etc.);
8. Para seu conforto use uma bata ou as suas roupas desde que não possuam fechos de metal ou correntes;
9. Deixe os seus objectos pessoais no vestuário;
10. Retire o calçado e coloque as protecções plásticas nos pés;

**A2) Durante o estudo:**

1. Na sala vai permanecer deitado na mesa de exames do equipamento e colocado um dispositivo (antena) na região da cabeça e pescoço;
2. Para seu conforto vão ser colocados apoios, nos membros inferiores e na cabeça, pois terá que permanecer imóvel durante o estudo, tanto quanto possível;
3. Não será administrado qualquer tipo de fármaco ou medicamento;
4. Devido aos níveis de ruído acústico produzidos pelo equipamento de RM, podem ser fornecidos auscultadores, se achar necessário;
5. Será observado pelo técnico de Radiologia durante a realização do exame, havendo sempre a possibilidade de comunicação por um intercomunicador e a interrupção do estudo a seu pedido (accionada através de uma campainha).

**B. COLABORAÇÃO PRETENDIDA:**

1. Todas as instruções vão ser fornecidas previamente pelo investigador de forma clara;
2. É crucial que permaneça calmo e atento;
3. Durante o estudo ser-lhe-á pedido que pronuncie e/ou repita determinados sons, após a seguinte instrução verbal: **“Respire fundo, e diga ‘som X’ em 1...2...3”**;
4. Existirão pausas para respirar normalmente durante o estudo, para programação do exame.

Declaro que as informações deste guia foram perceptíveis e de fácil compreensão. Li e entendi todo o conteúdo e tive a oportunidade de esclarecer quaisquer dúvidas.

---

(Assinatura do participante)

## ANEXO 3



FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Programa Doutoral em Engenharia Biomédica (PRODEB)

Título da tese: A Utilização da Ressonância Magnética para a Caracterização Funcional da Fala

**QUESTIONÁRIO DE SEGURANÇA EM RM**

O presente questionário visa a pesquisa de eventuais contra-indicações à utilização da ressonância magnética (RM) para fins de imagem médica e investigação clínica. Este foi elaborado com base nas linhas de orientação para práticas de segurança em RM, definidas pelo *American College of Radiology (ACR)*, [1]. Leia atentamente cada uma das questões abaixo antes de responder.

Sexo: M  F  Data de Nascimento: \_\_\_/\_\_\_/\_\_\_ Peso: \_\_\_ Kg

Código

Possui algum destes dispositivos dentro do corpo?	Sinalize com (x)	SIM	NÃO
Qualquer tipo de implante eléctrico, mecânico ou magnético			
Pacemaker cardíaco ou outro dispositivo cardíaco implantado			
Válvula cardíaca, implante artificial, stent, filtro ou eléctrodo interno			
Clip aneurismático			
Neuroestimulador			
Bombas internas de infusão (e.g. insulina, quimioterapia)			
Implante colinear ou ocular, ou fragmentos metálicos nos olhos			
Prótese metálica interna (e.g. anca, joelho) ou parafusos metálicos decorrentes de cirurgia ortopédica			
Próteses dentárias, implantes dentários, placas ou outro dispositivo ortodôntico			
Dispositivos metálicos como balas, fragmentos de balas ou metálicos			
Dispositivos de fixação, válvulas de pressão, shunts			
Implantes mamários ou penianos			
Dispositivo intra-uterino			
Penos para medicação tópica (e.g. estrogénio, nicotina, nitroglicerina, contraceptivo)			
Agrafos cirúrgicos ou outros dispositivos de fixação			
Piercings, tatuagens ou maquilhagem permanente			

Se respondeu afirmativamente a qualquer uma das opções acima, indique:

Tipo de material/dispositivo: \_\_\_\_\_

Localização do material/dispositivo: \_\_\_\_\_

Data da colocação: \_\_\_/\_\_\_/\_\_\_

OBS.: As próteses metálicas não são uma contra-indicação, aconselhando-se contudo a realização da RM decorridos 2 a 4 meses após a sua colocação, para estabilização da mesma.

Se é do sexo feminino:	Sinalize com (x)	SIM	NÃO
Está ou pensa está grávida?			
Está a amamentar?			

OBS.: Uma vez que o estudo dos possíveis efeitos adversos da RM no feto é ainda limitado, este exame, por precaução, deve ser evitado nos primeiros 3 meses de gravidez.



FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Programa Doutoral em Engenharia Biomédica (PRODEB)

Título da tese: A Utilização da Ressonância Magnética para a Caracterização Funcional da Fala

**QUESTIONÁRIO DE SEGURANÇA EM RM**

Possui historial de alguma destas doenças?	Sinalize com (x)	SIM	NÃO
Diabetes			
Insuficiência renal ou outra doença renal			
Doença hepática			
Doença cardíaca			
Claustrofobia			
Doença alérgica a algum medicamento ou alimento, qual? _____			

Declaro que compreendi a informação constante neste questionário, tendo-me sido dada a oportunidade de fazer as perguntas que julguei necessárias.

Porto, ..... de ..... de 20\_\_

\_\_\_\_\_ (Assinatura)

(campo reservado à investigadora responsável)

Registo do SAR: \_\_\_\_\_ W/Kg

Complicações decorrentes do estudo:

---



---



---



---

\_\_\_\_\_ (Assinatura da investigadora responsável)

[1] E. Kanal, A.J. Barkovich, C. Bell, J.P. Borgstede, W.G. Bradley, J.W. Froelich, T. Gilk, J.R. Gimbel, J. Gosbee, E. Kuhni-Kaminski, J.W. Lester, J. Nyenhuis, Y. Parag, D.J. Schaefer, E.A. Sebek-Scoumis, J. Weinreb, L.A. Zaremba, P. Wilcox, L. Lucey, and N. Sass, "ACR guidance document for safe MR practices: 2007." *American journal of roentgenology*, vol. 188, 2007, pp. 1447-74.



## ANEXO 4

**FORMULÁRIO DE CONSENTIMENTO INFORMADO**

Conforme a “Declaração de Helsíquia” da Associação Médica mundial  
(Helsíquia 1694; Tóquio 1975; Veneza 1983; Hong Kong 1989; Somerset West 1996; Edimburgo 2000)

**DESIGNAÇÃO DA TESE:** A Utilização da Ressonância Magnética para a Caracterização Funcional da Fala

**DESIGNAÇÃO DO ESTUDO:** Estudo da Voz Falada e Voz Cantada: estudo de casos

**Investigadora:** Sandra Moreira Rua Ventura

**E-mail:** [sandra.rua@eu.ipp.pt](mailto:sandra.rua@eu.ipp.pt)

**Orientador:** Prof. Dr. Diamantino Freitas (FEUP)

**Co-orientador:** Prof. Dr. João Manuel Tavares (FEUP)

**Âmbito:** Programa Doutoral em Engenharia Biomédica

**Introdução**

Foi solicitado para participar no estudo acima mencionado pois não possui historial de patologia orofacial ou deficit articulatorio, nem qualquer contra-indicação à utilização da ressonância magnética (RM) para fins de imagem médica e investigação clínica.

Este formulário de consentimento informado fornece-lhe as informações necessárias sobre este estudo e que serão discutidas consigo. Assim que tiver compreendido o estudo e que concordar em participar no mesmo, solicita-se que assine o formulário de consentimento. Ser-lhe-á fornecida uma cópia para seu uso pessoal.

Antes de ser informado a respeito do estudo, é importante que tome conhecimento dos seguintes aspectos:

- A sua **participação é inteiramente voluntária**;
- Poderá decidir não participar ou cancelar a sua participação neste estudo, em qualquer altura.

**Objetivos do estudo**

As recolhas imagiológicas por RM do trato vocal durante a fala servirão para caracterizar o mecanismo da produção da fala, do ponto de vista morfológico e dinâmico em diferentes contextos de voz falada e de voz cantada. **O objectivo principal é avaliar qual a adaptação da articulação nas diferentes funções associadas à fala ou afins.**

Deste estudo, vão participar cerca de quatro a cinco pessoas, jovens, voluntárias e com diferentes conhecimentos e experiências profissionais desde o canto, a interpretação, e ainda terapeutas da fala.

**Procedimentos**

Se resolver participar neste estudo e assinar o respectivo formulário de consentimento, ser-lhe-á também pedido que preencha um questionário de segurança em RM e fornecido guia com instruções de preparação e colaboração no estudo. Todas as recolhas serão supervisionadas pelo Médico Radiologista envolvido no estudo.

Durante as recolhas:

- Será efectuado uma preparação e treino prévio do corpus da fala em colaboração com um Terapeuta da Fala;
- Será instruído devidamente de modo a facilitar a sua colaboração;
- Ser-lhe-á solicitado que pronuncie diferentes sons, em sequências devidamente estruturadas;
- Deverá permanecer imóvel durante o estudo, tanto quanto possível;
- Não será administrado qualquer tipo de fármaco.

### Riscos e/ou desconforto

A RM é um método de imagem, que permite a obtenção de imagens multiplanares de qualquer parte do corpo humano, e que não utiliza radiação ionizante. A imagem por RM tem por base a interacção de um campo magnético e pulsos de radiofrequência com os núcleos do corpo humano.

**Não existem estudos científicos que comprovem a existência de danos biológicos resultantes da exposição a campos magnéticos como o utilizado em RM.**

Os pulsos de radiofrequência aplicados causam variações de temperatura, devido ao depósito de energia nos tecidos. Será, por isso, respeitado o limite de absorção (dose média  $\leq 3W/kg$ , para um tempo  $\geq 10min.$ ) para a região da cabeça bem como o aumento da temperatura corporal não será superior a  $1^{\circ}C$ .

Durante o exame verificam-se níveis de ruído acústico variáveis produzidos pelo equipamento de RM, sendo fornecidos auscultadores, para maior conforto e instrução durante o estudo.

Dada a utilização de um campo magnético, existe a possibilidade de interferência eléctrica, risco potencial de projecteis, torção de objectos, logo serão tomadas todas as normas de segurança necessárias.

A necessidade de permanecer imóvel durante as recolhas e o próprio posicionamento poderão causar algum desconforto.

### Benefícios previstos

Dada a escassez de informação morfológico-dinâmica da produção da fala, será possível através da sua participação:

(1) Demonstrar a aplicabilidade do protocolo de ressonância magnética do trato vocal, em situações estáticas e dinâmicas para a caracterização da fala;

(2) Extrair medidas quantitativas do trato vocal em diferentes contextos de voz falada e cantada, e analisar quais as diferenças articulatórias entre sujeitos.

Este conhecimento contribuirá para uma melhor compreensão dos mecanismos subjacentes às disfunções articulatórias da fala, assim como possibilitará o desenvolvimento de algoritmos para síntese e reconhecimento da fala.

### Custos

Este estudo não implicará qualquer custo para si.

### Sigilo

Será garantida a confidencialidade dos seus dados pessoais:

- A sua identificação será efectuada através de um código;
- Não será identificado pessoalmente em nenhuma publicação referente a este estudo.

### Declaração de Consentimento

Eu, abaixo-assinado, ..... (nome completo),

compreendi a informação que me foi fornecida acerca do estudo em que irei participar, tendo me sido dada a oportunidade de fazer as perguntas que julguei necessárias.

Por isso, consinto participar no estudo respondendo a todas as questões propostas e permitindo a execução das recolhas imagiológicas por ressonância magnética do trato vocal.

Porto, 29 de Outubro de 2011

\_\_\_\_\_  
(Assinatura)

Pela Investigadora Responsável,

\_\_\_\_\_  
Código