The Voice Clinic Handbook
Second Edition

The Voice Clinic Handbook

Second Edition

Edited by

Tom Harris

and

David M. Howard

compton
PUBLISHING

This edition first published 2018 © 2018 by Compton Publishing Ltd.

Registered office: Compton Publishing Ltd, 30 St. Giles', Oxford, OX1 3LE, UK
Registered company number: 07831037

Editorial offices: 35 East Street, Braunton, EX33 2EA, UK
Web: www.comptonpublishing.co.uk

ISBN Hardcover: 978-1-909082-21-2 Softcover: 978-1-909082-69-4

A catalogue record for this book is available from the British Library.

Cover design: David Siddall, http://www.davidsiddall.com

Set in 11pt Adobe Caslon Pro/Myriad Pro by Stuart Brown and Kerrypress Ltd (www.kerrypress.co.uk)

1 2018

Table of Contents

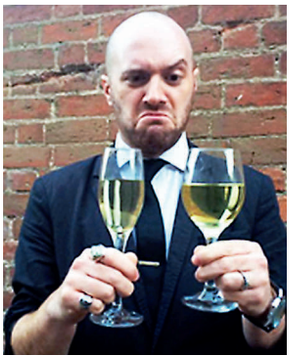
Contributors

Anthony Aymat, FRCS (Otol), FRCS (ORL-HNS)

Tony Aymat is a consultant ear, nose, and throat surgeon, working for the National Health Service at University Hospital Lewisham (UHL). He graduated from the University of Barcelona in 1990 and trained in the West Midlands, South West, and London. He specializes in voice disorders in professional and non-professional voice users and is a co-director of the UHL interdisciplinary voice clinic. In addition, he has a special interest in ear and hearing problems.

He has been involved in postgraduate medical training as well as teaching nurses and paramedical groups. He is member of the faculties in the surgical skills course at the Royal College of Surgeons, the temporal bone course in Barcelona, and the Intercollegiate Examination Course at the Royal National Nose, Throat and Ear Hospital in London.

Dane Chalfin



Dane Chalfin is a singing teacher and vocal rehabilitation coach working in both NHS and private practice in London and Manchester. His ongoing research (2006–17) with University Hospital South Manchester into the endoscopic assessment of disordered singers and the laryngopharyngeal gestures of singing and primal sounds now has over 400 subjects and may be the largest study into the singing voice to date. He also contributed to the development of an endoscopic protocol for the assessment of singers in the voice clinic (Jones, Chalfin, Garrett; 2006–17).

Dane is a past president (2015–16) and director of the British Voice Association. He is currently a vocal advisor to the Medical Committee of the British Association of Performing Arts Medicine (BAPAM) where he pioneered the creation of competencies for Vocal Rehabilitation Coaches working in multidisciplinary voice clinics in the UK. Dane has previously held the Associate Professorship of Vocal Coaching and Rehabilitation at Leeds College of Music, where he was also Principal Lecturer of Artistry and Performance (2007-17). Other academic posts include teaching fellow in Pop Voice at Liverpool Institute for Performing Arts (2003-6), Senior Lecturer in Pop Voice at the Royal Northern College of Music (2008-10), and Lecturer in Voice at University of Salford (2001-03).

Publications include TC-Helicon's *Ultimate Guide to Singing* (2016), and Bloomsbury's *The Singer-Songwriter Handbook* (2017). Clients include major recording artists, West End productions of *Wind in the Willows*, *Rock of Ages*, *Sunny Afternoon*, and *Motown*; and national tours of *Jesus Christ Superstar, Wicked, The Producers, Shout!*, and many others.

Dinah Harris, ARCM(Hons)

Dinah Harris studied at the Royal College of Music (RCM) in London with Dame Isobel Baillie and in Vienna with Rita Streich. She was a prize winner at 's-Hertogenbosch, Holland. Between the years 1971 and 1990 she was a professional opera and concert singer both in the UK and Europe.

A professor at the RCM, Dinah has worked as a singing teacher and voice coach, both privately and, from 1990 until 2010, as a member of the voice clinic team at Queen Mary's Hospital, Sidcup, Kent and Lewisham University Hospital. She has served on the faculties of the Pan Pacific Voice Conferences in San Francisco, the Vancouver Voice Conference, the Canadian Voice Care Foundation in Banff, Canada, and the Australian Voice Association. She has been an external lecturer for Goldsmiths College, University College London, Thames Valley University, Voice Care Network, the Queensland Conservatorium, the Royal Academy of Music, and RADA. Ms Harris was co-author of the first edition of *The Voice Clinic Handbook* and has served as both as a member of both the Council and Education Working Party of the BVA.

Sara Harris, FRCSLT

Sara Harris started working in voice in 1978 at the Radcliffe Infirmary, Oxford where, together with Tom Harris, she set up the Oxford Voice Clinic. In 1987, they moved to London and there they set up the Sidcup Voice Clinic where they were later joined by osteopath Jacob Lieberman and singing coach Dinah Harris. Sara was one of the co-founders of the Voice Research Society (now the British Voice Association) in 1985 and was a joint editor of *The Voice Clinic Handbook* (1998). She has contributed a number of book chapters and research articles to various text books and journals in voice.

Currently, Sara works in the Voice Disorders Unit at Lewisham Hospital and also in private practice. She was made a Fellow of the Royal College of Speech and Language Therapists in 2007 and is a past president of the BVA, where she continues to serve on the association's Education Working Party.

Tom Harris, MA, FRCS, Hon. FRCSLT

Tom Harris is a retired consultant ENT surgeon at University Hospital Lewisham, Queen Elizabeth Hospital, Woolwich, and Queen Mary's Hospital, Sidcup, and Senior Lecturer at King's College London, GKT School of Medical Education. He is a past member of the editorial board of *Journal of Voice*, *Folia Phoniatrica*, and others.

Married to Sara Harris, a voice therapist, with whom he opened one of the first multidisciplinary voice clinics in Britain (in 1982), he was the Founding Chairman of the Voice Research Society (now renamed The British Voice Association), and served as a council member and president of that association. He is the principal editor and contributor to *The Voice Clinic Handbook*, and his publications include subjects such as functional laryngeal anatomy and the importance of videostroboscopic assessment in voice disorder and phonosurgery. He wrote the official world report for the International Association of Logopedics and Phoniatrics (IALP) on *The Pharmacological Management of Voice Disorders* and is one of the team members that developed the use

of manipulative therapy in hyperfunctional voice disorders. Presently, he is actively involved in developing a new generation of superfine microlaryngeal instruments with Professor Markus Hess.

Markus M Hess, MD

Markus Hess is a professor of otolaryngology and a phoniatrician. He is the founder of the Deutsche Stimmklinik in Hamburg. His publications include more than a hundred articles in professional journals and books. Among his many international commitments he is a member of the editorial board of the *Journal of Voice*. Since 2003, he has been Secretary General and is now President of the Pan European Voice Conferences, was a founding member of the German Society for Phonosurgery and founder and board member of the European Academy of Voice. He is currently President of the International Collegium Medicorum Theatri in addition to membership of the European Academy of Voice, Union of European Phoniatrists, The Voice Foundation, and the International Association of Phonosurgery.

David M Howard, BSc (Eng), PhD (London), CEng, FIET, Senior MIEEE, FIOA

David Howard was recently appointed as the Founding Head of the Department of Electronic Engineering at Royal Holloway, University of London, where the intention is to recruit more than the national average (10–15%) of female undergraduate students. Previously, he held a Chair in Music Technology at the University of York where he was twice Head of the Department of Electronics, and before that he was a lecturer in experimental phonetics at University College London.

His main research interests are the acoustics and psychoacoustics of speech, singing, and music. David was editor-in-chief of the journal *Logopedics, Phoniatrics, Vocology* for 12 years, and he is a member of the editorial boards of the *Journal of Voice*, *Logopedics, Phoniatrics, Vocology*, *Forensic Linguistics*, and *The International Journal of Speech, Language and the Law*. He is a founding member of the International Association of Forensic Phoneticians and he was the 1997/98 president of the BVA. Outside academic research, David is an accomplished choral conductor and organist and he enjoys sailing whenever the opportunity arises.

Jacob Lieberman, MA, DO

Jacob Lieberman is an internationally-acclaimed expert on the diagnosis and treatment of muscle tension voice disorders and the management of vocal crisis.

As an osteopath and a psychodynamic therapist he developed the Lieberman protocol for the physical examination of the larynx. The application of principles taken from osteopathic medicine and psychodynamic psychotherapy led to the development of a powerful integrated treatment for a wide range of symptoms in voice and swallowing arising from tight muscles and joints. He is a member of the voice clinic team at Lewisham University Hospital, London, and the Deutsche Stimmklik in Hamburg, Germany. He has researched and co-authored papers on postural aspects in voice disorders and aspects of the mechanical properties of laryngeal function and is a contributor to leading text books in voice. His website is www.jacob-lieberman.co.uk

Julian McGlashan, MB BS, FRCS (Otol), Hon. FRCSLT

Julian McGlashan is a laryngologist and head and neck surgeon at the Queen's Medical Centre Campus at Nottingham University Hospitals. He specializes in voice and laryngeal disorders, thyroid, and head and neck surgery. He works as a part of a multiprofessional voice disorders service team. His main research interests are in imaging of the vocal folds, voice measurement, the effects of gastric reflux on the upper aerodigestive tract, and the singing voice. He lectures widely in the UK and abroad, and is actively involved in teaching on postgraduate courses. He is a council member of the British Laryngological Association and past president of the BVA.

Scott Reid Moisik, MA, PhD

Scott Moisik is currently Assistant Professor in the Division of Linguistics and Multilingual Studies at Nanyang Technological University, Singapore. Prior to this he was a postdoctoral researcher at the Max Planck Institute for Psycholinguistics in the department of Language and Genetics in Nijmegen.

Having collected some 13 major awards and distinctions for his work so far, his research interests span phonetics and phonology; anatomy, physiology, and biomechanics of speech; phonation and laryngeal articulation; voice quality in speech, personas, imitations, and characters; language and genetics. His main interest is in how these factors, along with the aerodynamic, acoustic, perceptual, and social facets of speech, conspire to shape speech sound systems.

Phiroze Neemuchwala, MA, Adv.Dip.PC, Adv.Cert.PG, BAC

Phiroze Neemuchwala is a psychotherapist and counsellor. His practice in psychosomatic psychotherapy is based in Southeast London. His training and qualifications include cognitive-behavioural, analytical psychotherapy, and Gestalt models. Since 1992 he has organized and run weekend workshops to introduce groups to the psychotherapeutic approach to dealing with symptoms. He has been associated with the Queen Mary's Hospital and subsequently the University Hospital Lewisham voice clinic since 1992.

John S. Rubin, MD, FACS, FRCS

John Rubin is a consultant ear, nose, and throat surgeon at the Royal National Throat Nose and Ear Hospital, a part of University College London Hospital (UCLH), where he is also currently chair of the consultant forum, lead clinician of the voice disorders unit and past clinical director. He is honorary senior lecturer at the Ear Institute, and honorary consultant ENT surgeon and co-chair of the voice and swallowing unit at the National Hospital for Neurology and Neurosurgery, both a part of UCLH.

His interests include voice disorders and laryngeal surgery. He has written extensively, including several books and numerous articles and chapters, and regularly lectures on voice-related topics. He has served in multiple capacities on international editorial and scientific boards and committees. Mr Rubin is past president of the *Collegium Medicorum Theatri* as well as the BVA, British Voice Association and is a founding member of the European Academy of Voice.

Acknowledgements

We would like to thank the small army of friends and helpers who, in addition to the authors, have made the publication of a second edition of *The Voice Clinic Handbook* possible. Firstly, a big thank you to our publisher, Noel McPherson, who has taken a three-year delay in receiving the finished product without visible signs of despair, anger, hostility, or similar. Instead, he has been enabling, and seemingly cheerful throughout the incubation period. Sara Harris, in addition to her own busy clinical practice, lecturing, producing material for the British Voice Association and this book, organizing the refurbishment of a derelict cottage, and coping with a somewhat derelict husband who appears to be superglued to a computer during the day and much of the night without complaint. We suckered Janet Kerr and Linda Livingstone into drawing cartoons: they seem to have a gift for making us see things in a clearer perspective and help readers forget that they are dealing with a textbook – we are very happy with their different take on clinical medicine. Jo Mills volunteered herself (again) to proof read the manuscript and correct the howlers and omissions that are wont to creep into an enterprise such as this, for which she receives our grateful thanks.

Many of the photographs in this edition are from Tony Aymat and Nick Gibbins of the UCL Voice Clinic, to whom we owe grateful thanks for their generosity.

Markus Hess and Susanne Fleischer of the Hamburg Stimmklinik have likewise been extremely helpful in providing photographs for other authors' chapters in this book.

We are deeply indebted to our friend and colleague Xing hui Hu of Laryngograph Ltd., without whose help and knowledge many thousands of laryngeal images and investigations performed over the years would have been much the poorer.

We want to thank our friend, colleague, and ally Ron Baken for all his wise and helpful advice, which is always provided – often at zero notice by telephone at ridiculous hours of the day or night – without complaint. (Even if we have woken up Jealousy, the green-eyed pit-bull terrier...)

Other people have been most generous too with their time and expertise: we must give our grateful thanks to Scott Moisik, whose contribution goes way past co-authoring two chapters; also to Neil Douglas and Kate Young for their help, and to Jacob Lieberman for allowing us to re-publish (with some modifications) his chapter from the first edition. Much of our voice education has come from discussion, arguments, agreements, and disagreements – frequently over dinner and/or beverages after a hard day's conference somewhere in the world. We know who our friends and mentors are (and, hopefully, they do too).

Finally, we must acknowledge our patients, clients, and experimental subjects, without whom there would be no Voice Clinics. They give us their trust (well, sometimes), they make us think about *what?* and *where?* . . . and then make us justify our reasons for doing what we do by finding answers to *what? why? how?* and *when?* They are our *raison d'être*.

Foreword to the First Edition

We certainly have come a long way. I first grappled with the intricacies of voice disorders back in the early 1960s - not so long ago that events are nostalgically shrouded in the softening mists of time, but far enough away that an overall perspective is possible. And my recollection is that 'voice' was clearly on the sidelines of the rehabilitation game. It wasn't that it was unimportant, exactly. There was meaningful, sometimes important work for the vocal therapist to do. Laryngectomees were plentiful in those days, and there was always the hoarse schoolteacher, clergyman, or salesman. But –at least on this side of the Atlantic—vocal rehabilitation was certainly not in the spotlight which, as I recollect, was pretty much fixed on stammering and aphasia. West, Ansberry, and Carr's *Rehabilitation of Speech*, a then-prominent American text for beginning students, devoted all of 11 of its 688 pages to 'remedial procedures for dysphonia' and another six and a half pages to rehabilitation of the laryngectomees. Therapy for aphasia got twice as much play.

There were three major groups involved in vocal dysfunction back then. The medical people dealt with, from their perspective, the "real" problems. The lesions, paralyses, infections. The psychologists laid claim to all those situations that were not the province of the medicosurgical arts, apparently believing, in those Freud-dominated years, that all –vocal nodes, aberrant vocal mutation, spasmodic dysphonia, you name it—all was psychoneurosis. And finally, there were the speech therapists who, in North America, were just starting to metamorphose into more-impressive 'speech pathologists'. Their education in voice disorders was often limited to part of a one-semester course called 'Voice and Articulation', but to them fell the role of patching up the postsurgical patient and of remediating the persistent vocal problems of the psychologized-but-still-dysphonic. Members of the three professional groups talked to each other, of course, and they sometimes referred patients back and forth, but, at least in my experience, their professional actions were perhaps better described as parallel, rather than cooperative, play. Each profession had its own set of background assumptions, its own rules of engagement, its own remedial-world view, and, of paramount import, its own turf to defend.

There were a few relatively good ways of getting an accurate assessment of vocal tract function and of obtaining better physiological insight into the patient's problem, but they were for the professional elite. The venerable Sonagraph, progenitor of almost all of today's spectral insight, was, for those times, horrifically expensive. (It could also spew clouds of noxious black grit and electrocute the fingers of the unwary.) Stroboscopy was only nascent and, in the pre-VCR era, provided no permanent record for careful analysis. Even getting an accurate reading of the fundamental frequency was difficult and perturbation measurement was out of the question. Research labs had ways of sorting out the physiological bases of a given dysphonia, but little that could be economically and conveniently applied in the average clinical setting. So, for most of us, assessment of vocal function meant not very much more than listening to vocal production.

Voice is much more 'in' today. We have, as I said, come a long way. To be sure, laryngology still doesn't have the cachet of, let us say, brain surgery, and speech pathology still has a whole lot more psycholinguistics groupies than voice specialists. But, unobtrusively, the treatment of voice disorders has moved decidedly forward on the rehabilitative stage, if not into the spotlight then at least into the brighter illumination of the

footlights. Doing something about vocal problems has become more important: witness the number of specialized voice clinics that are sprouting up. Social scientists will no doubt postulate various reasons for this: the greater importance of spoken communication in an increasingly white-collar workplace; a more affluent society indulging in more vocally-deleterious entertainments and amusements; changes in medical economics, and so on. But there are other (I like to think more important) propelling forces, and they are manifest in the construction, content, and tone of this book.

For one thing, we know a great deal more about vocal mechanisms. We've learned from models, from sophisticated probes of physiological function, in vitro and in vivo, from epidemiological and genetic analysis, from highly sophisticated acoustic analysis, from the vast array of advanced techniques made possible by explosive advances of technology. And, most important, we've been blessed with a generation of very gifted theorists.

We can also offer our patients much more. Phonosurgery comes to mind, as does an enriched palette of behavioural approaches to vocal restoration and rehabilitation. That we can do so is due to better integration of the insights of careful research, highly-structured clinical observation, and everyday exigencies of the clinical world.

And then, of course, we can understand the patient's particular problem better than we ever hoped to 35 years ago. Thanks largely to a revolution in technology interacting with advancing physiological and acoustic theory, we can obtain a much more complete picture of how and why the vocal system is malfunctioning. The clearer the picture, the more targeted the intervention can be.

Which brings me, finally, to this book, a well-endowed child of that much-more-mature science of voice. Its very tone – self-confident enough to undertake an authoritative survey of the field without being laborious or pedantic—shows how far we have indeed come. More substantially, this book is impressive in the unique way in which it weaves together insight and understandings from so many of the basic and applied sciences that inform the practice of vocal rehabilitation, and in the way in which it brings that integration to bear on such a broad spectrum of clinical problems. With clarity, conciseness, and a leavening of wit, this tome deals with whatever the question that the vocal professional might pose. (If an answer is not here the question must, I think, be very odd indeed.) Achieving this is an accomplishment of very significant magnitude. An accomplishment that merits real congratulation.

Uniquely, however, along with a thorough survey of the canonical interventions – medicosurgical, behavioural, and psychological—this text includes approaches and therapies not heretofore emphasized in the context of serious literature on vocal dysfunction. And it does so with a careful rationality and persuasiveness that have not generally been characteristic of therapeutic writing in our field. And for that, this book warrants enthusiastic applause.

But there is, at base, something very different about this text. And that is that its several authors have not simply come together (or, as is more common, remained apart!) for the purpose of writing a book. What we have here is an explanation of practice by a functioning *team*, a group of professionals who do not simply share a common understanding of their Endeavour but who do it together. There results in their book a unification of approach, a consistency of understanding, an obvious mutual respect, and a professional interdependence that are, in my experience, unparalleled in any similar publication in our profession. And for that achievement: Bravissimo!

R J Baken, PhD

Foreword to the Second Edition

And so here we are again. Not all that many professional textbooks get to have a second edition: viewpoints and (sometimes) fashions change; knowledge moves on, leaving some erstwhile experts too far from a new 'cutting edge'; authors drift to other areas of interest; sometimes authors simply get tired. Yet here we are again: the professional public, in a show of good professional judgment and stylistic discernment, has encouraged a second edition by this team of surgeons, voice rehabilitation specialists, and scientists. They're still working together, still innovating, still at the forefront of their professions, still eager and able to do such an effective job of sharing what they know, what they believe we need to know, and what information we'll require for a future that is unfolding – faster and faster—before us. Fortunately, they have not given up their basic outlook: science- and evidence-based approaches, behavioral, physical, and intensely interdisciplinary, are brought to bear on the restoration and rehabilitation of voice. Their multi-professional viewpoint is alive and well in this update and revision. We continue thereby to benefit enormously.

The enormous value of the original *Voice Clinic Handbook* lives on –and grows—in this new edition. Hats off and three cheers for that!

R J Baken, PhD
Professor Emeritus of Speech Science
Columbia University
New York, NY

Senior Lecturer
New York Medical College
Valhalla, NY

Introduction

It was never the intention that this book was to be a textbook containing everything there is to know about the management of voice problems. The intention was that we put down what voice clinic team members have found useful and what works best for them. It was not intended that any reader would read it from cover to cover; it is more for dipping into in case it can inform in your own speciality or even clarify what it is that your colleagues in other specialities have to offer.

Like the first edition, it remains divided into three arbitrary parts: Part I provides an outline of the structure and function of the vocal tract, Part II addresses common treatment modalities, and Part III (the scientific bit) outlines the equipment for measuring voice: uses and limitations.

There has been no attempt to produce a 'house style'. The style may be a bit informal for some tastes, but we hope that, for a while at least, it is up to date and mostly evidence based. We even dare to hope that you enjoy what we have offered.

The Voice Clinic in 2018

What/who constitutes a voice clinic?

A voice clinic can be defined as a place to which a person complaining of a voice disorder can go to receive expert help in the diagnosis of their problem, and treatment appropriate to their particular needs.

This definition should be taken to imply that the clinic will offer the skills of more than one practitioner, and that the complainant will be seen for their initial appointment by a team. It should also be understood that, at the very minimum, there will always be a competent, medically qualified person present who, using appropriate techniques, can either exclude a pathological process or arrange any necessary medical treatment. There will also be a member of staff present who is able to offer at least symptomatic treatment for those dysphonias that are not due to medical disorder.

The staff

The rather stolid definition of a voice clinic offered above seeks to include all the help that is available. Of necessity, we all use what (or who) we've got. Voice clinic personnel vary internationally depending on the country's traditional management of voice disorder. In most of Europe, there is a strong tradition of medical specialization in disorders of communication, which is referred to as phoniatrics or phoniatry. A phoniatrician is a doctor who has been trained in ENT surgery, audiological medicine, and all aspects of speech and language therapy (logopedics). Although this speciality is considered essential in most of Europe and Scandinavia, for the most part the English speaking nations do not acknowledge the need for this speciality and divide the responsibility for management of voice problems between ENT surgeons and speech therapists, with the predictable wide variation in expertise available to the patient.

This situation allows ignorance to persist about the role and contribution of each team member and, not infrequently, there is confusion as to what is the most appropriate management of communication disorders and, most especially, phonation problems. There are many speech and language therapists in Britain for whom a normal letter of referral gives only the information 'I can see nothing wrong with this patient's larynx, please see and treat.' (In general, ENT surgeons are trained to detect the structural changes associated with disease, but have only the most rudimentary training as to the origins of laryngeal dysfunction.)

The situation might be acceptable if the therapist to whom the patient is referred has spent considerable time as an obligatory part of his or her formal training, learning about diagnosis and treatment. The reality, however, is that many highly competent voice specialized therapists in the UK and elsewhere have, like the ENT surgeon, learned their phoniatric skills much later in a hit-and-miss manner while actually at work. The working of most voice clinics in Britain when they first open for business could reasonably be described as 'the blind leading the blind'. This was certainly the case where we were concerned, and what is more, it is liable to remain the case in the near future unless postgraduate training courses which are currently in development are accorded proper professional recognition.

If the clinic is staffed by a phoniatrician or ENT surgeon and a speech therapist, both of whom are competent in what they do, then the system of referral will inevitably bring them patients who are harder to treat effectively and efficiently. There will for instance be the professional voice users with unusual and very demanding voicing requirements, not to mention all the patients in whom there is a significant psychological or social component in the production and maintenance of their symptoms. Although the clinic is set up to deal with acute voicing problems, one of its aims is bound to be the correction of dysphonia in such a way that it does not recur. To this end, clinics all develop their own 'short list' of voice professionals, e.g. singing teachers, voice coaches, psychotherapists et al., whom they know to be expert in their own fields, and to whom they are happy to refer the patient for further management.

The next logical step is to ask these willing professionals to join the clinic team on a regular basis, so that they may learn more about the medical/scientific background to their clients' disorders and so that the other members of the team may learn something of their skills.

There are of course certain obvious problems with this stage: What skills are available locally? How do you interest these other professionals enough in this particular enterprise to give up time on a regular basis? How do you persuade hospital management that these people have skills that are marketable in a medical context, profitable to the hospital and who therefore should be added to the payroll? The important thing to grasp is that although it is not easy to jump these hurdles, it is possible. We found that a good way to demonstrate the usefulness of these new colleagues was to 'slide them in sideways' as it were, by finding research money from unrelated sources to get things started, and by producing research results that proved that these 'additional' personnel had an important role to play in improving the efficiency of treatment. If you can persuade a director of finance that a potential member of staff's activities can be self-financing, then you are halfway there.

Our case-load and available expertise has led us to a position where we now have a 'core team' of ENT surgeon, speech therapist, voice coach/singing teacher and specialist osteopath. Other voice clinics will of course meet their patients' requirements rather differently. Configurations vary; what is important is that members of staff are prepared to see patients in joint consultation; it is not intended that patients are seen in isolation by a series of specialists who refer patients on to each other for further management. This was the system that prevailed in the past. It is slow and inefficient, it does not put patients through the system and discharge them fully recovered nearly fast enough, and it does not allow staff to learn from one another.

In today's atmosphere of contractual commitments, it is no longer possible simply to announce to a stunned world that you are going to open a new clinical service. Help has to be organized and paid for. Do not forget that as well as the additional case-load for the Appointments Desk and the Medical Records Department, and the general unhappiness about

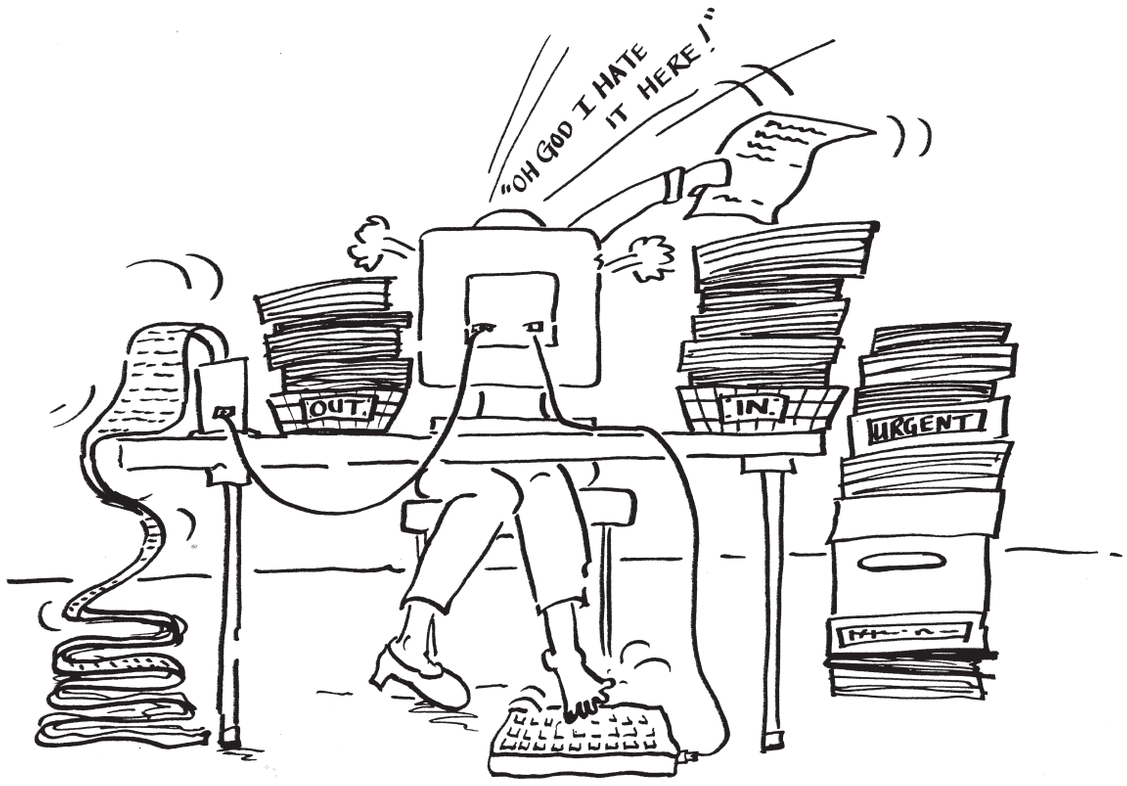
taking on an 'extra' clinic, there will of course be a requirement for additional support staff.

Patients, when they arrive, need a warm and reassuring welcome that they are in the right place and 'in good hands'. A relaxed and efficient atmosphere goes a long way towards helping to reduce anxiety and fear, especially for new patients. There must be a clinic coordinator/administrator to ensure the smooth running of the clinic for patients and staff alike – someone to make sure everything and everybody is in the right place at the right time. Important little things like tea and biscuits to soothe the nerves and reduce the tension are a valuable adjunct to any voice clinic.

Never underestimate the 'hidden' tasks required to run the clinic administratively. There has to be someone who can liaise with the patients, the voice clinic team and the other departments, as well as get together the case notes, questionnaires, letters and the forms – please don't forget the forms! It cannot rest with a professional member of the team, for instance the speech therapist or voice coach, to fulfil this role; it is too time consuming, and of course the administration goes on long after the clinic has finished for the day.

Best use of personnel

If all the professionals involved in the clinic see the patient at his or her initial interview, then, very rapidly, the team will develop its own consensus view of exactly what the patient is complaining about, what his or her needs are and how they might be most efficiently met. Clearly, after the initial interview, each involved professional may wish to assess the patient in more detail. We would not wish to pretend that objective measures relating to voice quality – tape recordings, electroglottography, spectrograms, phonetograms, airflow measurements or whatever – can all be gathered at a diagnostic and planning interview. Experience tells that examination using either rigid or flexible



chip-tip high definition endoscopes for videostrobolaryngoscopy is all there is time for on this first occasion. Quite separate time has to be found for further analysis on another occasion. This time constraint is no different from any other form of medical consultation; the patient is usually seen without a blanket battery of tests, and further management is decided upon on the basis of the history and examination.

In other traditions, voice clinics are much more hierarchical; the diagnosis and formulation of a treatment plan are decided by one person, usually a phoniatrician, and the game plan is rigidly adhered to until the patient has been seen for further review by the same individual. We feel that this takes no account of the special skills brought to the clinic by disciplines other than ENT and speech therapy. In our view, all the personnel involved in the clinic should be actively involved in formulating a treatment plan. To say that the ultimate responsibility for the patient's care rests with the doctor and therefore he or she should dictate what is to happen simply negates the possibility of any further interdisciplinary learning.

Survival as a team

> We must indeed all hang together, or, most assuredly, we shall all hang separately.
>
> Benjamin Franklin

If and when the voice clinic has become a successful fact of life, and the weeks have started becoming years, then inevitably there will be episodes of friction between staff members. For the clinic to work, its multidisciplinary nature requires that it is a clinic of equals, and the opinions of all must be heard and included in any 'game plan'. This internal rule applies regardless of any traditional hierarchical rules of healthcare provision as to 'Who is the boss'.

Problems can, and do, arise when there are sincere differences of opinion about the best course of action concerning a patient's treatment plan. No one person ever has all the right answers.

In the clinic

It is therefore imperative that a consensus is reached in each and every case. Every clinic develops its own modus vivendi to deal with this. It is obviously not acceptable to argue at length over an already confused patient. This demands a certain degree of self-discipline. Everyone should be allowed to make their own comments about the findings from the history and examination and to explain it aloud to the patient. This being done, it is almost inevitable that the team members will have a clear idea as to whether any part of the treatment is within their personal remit or not. (working together regularly develops the ability to second-guess the other members of the team quite accurately.)

If a team member thinks they have noticed a factor that may have escaped the others, he/she must of course voice the observation. If the other members agree, then it will automatically be taken into account by all, and the patient is usually much reassured by the general agreement. If, on the other hand, there is one of the rare occasions when the other team members do not agree, then the dissenting member must (a) refrain from further argument until the patient has left, and (b) equally importantly, discuss it with all the other members afterwards until all the reasons for and against have been given by all concerned and a genuine consensus has been reached.

If members take gripes and disagreements home unresolved, then resentment builds rapidly and the person feeling wronged will start looking for opportunities elsewhere and the team will die as a truly multidisciplinary entity.

Once the primary course of action is agreed, the appropriate spokesperson explains the management plan to the patient. This will normally be the therapist or surgeon who is going to organize or undertake the treatment; it is not automatically the titular 'head' of the clinic.

Beyond the clinic

If this discipline holds good within the clinic, the patients are normally impressed. They spread the word, and the clinic members will, over time, achieve a certain notoriety for their skills. All team members, of course, have their own individual careers and lives to get on with, and there will be competing commitments elsewhere, which may strain even the most loyal members of the clinic. We would strongly suggest that members air their future plans and ambitions together frequently and openly. This is not simply a matter of a chat over coffee should the opportunity arise; it is vitally important if the clinic is to continue to work as a unit. If there is no convenient time for this, make it. We would further suggest that if members are asked to talk about individual aspects of voice treatment or research that are not pertinent to their own particular domain within the clinic, they should suggest that, whenever possible, the appropriate member or members should do it instead. A truly multidisciplinary clinic cannot survive in the long term if members of the staff feel that they are being 'used' by any of the others.

The patients

Who comes to a voice clinic? Patterns of referral: service-related considerations in the UK

Once the decision has been taken to organize a specialist voice clinic, the question 'exactly who shall we see in the clinic?' arises. After the Oxford Voice Clinic was established in 1982, we found that while approximately 8% of patients passing through a general ENT outpatient department were presenting with dysphonia that required further assessment, only the more severely dysphonic patients without any apparent neoplastic pathology were being referred internally by members of the department for further management. It was not until we had gained the confidence of our colleagues that we felt able to comb out all the dysphonias referred by GPs to the department (Harris et al., 1987). When a voice clinic has been established for some time, however, it inevitably becomes well known and begins to establish its own clientele. Even in these days of commissioned healthcare, patients come via tertiary referrals from colleagues outside the area as well as from distant GPs whose patients may have bullied them into a referral for a 'second opinion' or for specialist management. The clinic runs the risk of being swamped with work, unless:

1. all the members of the clinic are able to come together for more than a single clinic per week and are also able to find additional time to deal with the treatment of the extra patients that the increased clinic capacity will engender; or
2. the clinic accepts that there will have to be some primary screening by colleagues in the general ENT service.

For all these reasons the eventual pattern of referrals within the UK seems to suggest that clinic patients will be drawn from a mix of the following.

- Internal referrals from ENT colleagues within the local department.
- Internal referrals from speech therapists who have problem patients that have been seen in an ENT department already. The patient may have self-referred for therapy, in which case the therapist needs to have the patient checked for organic pathology. (Be courteous, be careful to check that the patient's own doctor is amenable first).
- External referrals from GPs within the clinic's catchment area, who know of the voice clinic's existence and who have made a direct referral.
- External referrals from distant GPs who have been 'persuaded' to make a direct referral. In general, they are not enthusiastic about this because services are now commissioned i.e. bought and paid for by area commissioning groups of GPs whose job it is to pay for all hospital services. Hospital management is not keen on the idea of 'loss leaders' either.
- External referrals from ENT surgeons outside the catchment area for which there is a contract

using the 'special case' dispensation. There are of course other categories that do not readily fit into the above description, for instance the medico-legal referrals sent for independent assessment and others. These latter will not, however, make up the bulk of the case load.

Finance

The hospital may have a significant say as to who is on the staff of a voice clinic, but none of the above counts for anything unless you can persuade the budget holders, (i.e. the group of GP Commissioners, who hold the purse strings for the slowly withering budget that is provided by government), that what the area really wants and needs is an Integrated voice clinic as part of their healthcare provision plan. It will be noted when reading the list above, that sadly, it is no longer adequate metaphorically to hang out a sign saying 'we treat dysphonia'. In the current climate dominated by financial considerations the clinic must be seen to be financially advantageous to the unit in which it is housed.

As discussed earlier, business plans indicating the type of patients to be seen and the origin of their referral will have to be submitted to management, otherwise there will be no financial backing for the acquisition of apparatus or for hiring staff not normally found in departments of ENT surgery, judicious suggestion that the expected case-mix of patients who will be attracted to the clinic will prove 'a nice little earner' of extra revenue for the hospital does help a proposal for such a clinic to be regarded in a more favourable light. If it can be demonstrated that the throughput of such patients is more efficient than it would be if patients continue to be shuffled between a general ENT clinic and a speech therapy department, then the commissioners buying healthcare for the local population can also be satisfied at the same time (for example, the video-laryngostroboscopy routinely performed in clinic is regarded as an integral minor outpatient procedure by healthcare commissioners and private insurance companies alike). This also offers some small incentive to those who hold the financial reins to equip the proposed clinic with appropriate apparatus.

Appointments

How much time does an interview take? (How long is a piece of string?)

Do not forget that a voice clinic is one in which a minor surgical procedure – endoscopic assessment – is undertaken on every patient. Instrumental assessment always takes longer than the team would like. The patient needs to be prepared and the procedure explained. In addition to performing the endoscopy, sterilization of instruments is required afterwards and the findings must be explained to the patient, all of which takes time. Do not imagine that the clinic can perform to the best of its collective ability if much less than half an hour is allocated for every patient, new and old alike.

Running late

Eventually, even in a 'closed' clinic there will be extra patients: the emergencies; the force-booked, post-operative follow-ups; the 'walk-ins'; and the truly desperate professionals with genuinely career-threatening work commitments that cannot be ducked. The clinic will tend to become overloaded even if the original number of patient slots is adhered to as strictly as possible. We feel that as a tertiary referral centre we should endeavour to give each patient as much time as it takes to understand fully the nature of the problem and arrange a suitable treatment plan. This means that it is not always possible to stick rigidly to appointment times. This matters less than it might do as long as new patients are forewarned that the clinic may run a little late, that they should certainly bring something to read and that they should definitely not arrange other appointments timed for less than an hour after they were due to be seen in clinic. If patients are thus forewarned they do not get restless if an overloaded clinic is running late.

Tools for finding out about the patient

History-taking: freestyle question and answer versus the structured questionnaire

History-taking is a much more detailed business in a voice clinic than it would be in a more general outpatient setting. It is certainly one of the factors that requires a longer appointment time than is allocated for a general ENT clinic. In addition to the referral letter and any other related correspondence, there are areas of history-taking that are absolutely germane to the management of the patient and that must be covered in every case. According to the style of the clinic, it may be felt that a questionnaire can be given to the patient for filling out prior to the consultation, and the answers to the questions can then be 'gone over' in the clinic in less time than free-form questioning requires.

Some of the advantages of a proforma based interview are that history-taking is quicker, the form guarantees that factors are not overlooked and that negative and 'normal' findings are always written into the notes (a very important consideration in these days when 'evidence-based medicine' is the politically correct term for good clinical practice). If questions of scale are sought, then the patient is obliged to quantify his or her perceptions at a given moment in time, and this is of great value in the record as patients' memories of symptoms are notoriously unreliable. It is also true that the completed forms make any subsequent analysis for audit and research purposes very much easier.

The obvious disadvantage with the proforma is that the questions, by their very nature, always tend to limit the answer to an area within a preconceived range. The questions tend to beget yes/no type answers and it is extremely difficult to produce questions that open up a new line of thinking or that beget further searching questions. This can be more of a hindrance than is at first apparent, especially when there are specialists from several disciplines in the clinic who will pick up on even small nuances of language in the patient's answers and who may wish to probe a little further.

The Lewisham Voice Clinic is only one of many clinics that routinely use proformas. It is essential, however, always to bear in mind the convergent thinking that accompanies such questionnaires and hence in the consultation to try to expand on any areas that may be concealed by the form of the answers required by the document. Having 'ears to hear' what the patient is trying to express is exceedingly important, and for this reason there will be several further sections relating to the history in the chapters on therapy.

The purpose of paper

The purpose of a preliminary history questionnaire is merely to establish and document a number of agreed facts. This is also the aim of assessment protocols for recording the facts relating to an examination (see for example the speech therapy questionnaires and the Lieberman protocol for examination of posture and laryngopharyngeal set appended to the chapter on manual therapy). The purpose of this type of document is therefore not the same as that of specialist protocols designed to elicit information about and perhaps quantify such factors as stress and anger, of which the patient may be entirely unaware. The difference in approach for the psychological inventory is that the precise wording of each question has been demonstrated to elicit a reliable and appropriate response, and is therefore absolutely not to be changed if it is to be used for scoring purposes. We are particularly impressed by The Nottingham University Department of Psychology questionnaires for assessing levels of stress at work.

It is not necessarily a good idea to become wedded to the use of proformas, no matter how beguiling the thought that the patient does more of the work. From a clinical point of view, the data will always be recorded and accessible, and from a research point of view protocols allow the data to be shuffled into (maybe)

statistically significant bins. However, this must be offset against the extra time that this added activity will require. The next inevitable step is to consider what to do with all this painstakingly acquired data.

The trouble with data

Data is a Latin word meaning 'Things given'. Of course data are not really given, they are sort of 'extracted' with varying degrees of difficulty and tedium both for the subject and enquirer alike. When establishing a voice clinic it is a very good idea to have some long-term plans about both clinical practice and future research. If the questions 'What data do we really want or need? For now? and (perhaps) for the future?' are posed, then it will be possible to identify the path you wish the clinic to take without wasting enormous amounts of time in the acquisition of useless data (Clarke, 2004). There has been considerable debate in recent years about exactly what data constitutes a reasonable record of a patient's diagnosis and progress in the voice clinic, and we feel that the guidelines proposed by the Phoniatrics Committee of the European Laryngological Society (ELS) should be considered (Dejonckere et al., 2001).

This is a basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques.

If data are required, where will they be stored? For clinical purposes we keep the history proforma together with the downloaded images from the videostrobolaryngoscopy in the patient's notes; all other documents are kept on file and the hospital database. Voice and video recordings remain in the memory of the clinic's stand-alone computer. Experience has shown that material necessary for research should never be entrusted solely to the hospital notes which may go astray. In the Lewisham voice clinics all data relating to the patients, the voice pathology treatments and outcomes are now kept on a spreadsheet in the hospital computing system as well as being backed up in the department. Even if research is more an aspiration rather than a reality in any voice clinic, continued funding is highly dependent on its members being able to quote accurate figures for accuracy of diagnosis, time/number of appointments from start to finish, quality of outcome which are all essential data for submission to those that hold the purse-strings. Of course the same applies to clinical research. When publicly trying to persuade colleagues that there might be a better way to do something, a comprehensive, accurate database running into thousands of patients has to be persuasive.

What about research?

> We have empirical grounds for great modesty in the field of voice research.
>
> Donald C Campbell

A good definition of scientific research is: performing a methodical study in order to prove a hypothesis or answer a specific question. Why research? Finding a definitive answer is the central goal of any experimental process. It is interesting to add to the sum of human knowledge in your own area of expertise, especially when it helps the 'experts' avoid making the same mistakes that they may have made in the past or when your results confirm a clinical supposition. Because medicine is moving away from experiential expertise to 'evidence based' practice, funding of your unit will depend significantly on the amount of research that it has generated and published (vide infra).

Research must be systematic and follow a series of steps and a rigid standard protocol, but will have little value if the initial question or statement to be validated has no endpoint which increases knowledge. (See the spoof 'Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials' purportedly by GCS Smith and JP Pell (2003).)

The sort of data that one may obtain in a voice clinic may not fit comfortably into a research paradigm. Before blundering on blindly or simply giving up on the data that you have, consider the observations of (Clarke, 2004) when discussing the use of data: '...the kind of information that is available in practice (often incomplete, idiosyncratic, fragmented, partly qualitative, and rapidly changing) and works out increasingly effective ways of dealing with it on its own terms. The methods may sometimes have to work even with single-cases and in real time. Better tools are essential to extract the relevant patterns from data, to extrapolate them into the future, and to steer them away from bad outcomes. This is the purpose of the 'SERIAL' (SEquential Real-time In-depth AnaLysis) approach to research. Science is not just 'the art of the possible', but the 'art of making things possible'.'

The moral of this? Get to know a really good statistician even before you start.

Tailoring a treatment plan to individual patient requirements

> To every complex problem there is a simple solution . . . and it's wrong.
>
> Anon

Dysphonia is a symptom, not a disease entity. Hence, different patients with a similar voice disorder will not all be managed with equal efficiency by a standard treatment plan if the several factors that have played a part in producing the problem are of different orders of importance (see Morrison, 1994).

The obese patient with severe gastro-oesophageal reflux and periodic nocturnal laryngospasm may well present with the same pattern of hyperfunctional voicing as an angry super-fit athlete. The treatment plan is clearly not going to be identical for both, because the priorities are different. The same is true for different voice requirements: while an enlarging epidermoid cyst in a vocal fold is a catastrophe for an opera singer, The resultant dysphonia may be an absolute boon for a character actress of a certain age. Just because a lesion has been identified does not necessarily mean it must be fixed.

Ensuring your voice clinic survives

(Our gratitude goes to *Kate Young* FRCSLT for her very constructive thoughts about this. Eds.)

This section is primarily addressed to those working within the NHS in England or Wales, although we feel sure that those working elsewhere will find some helpful thoughts that might still be of use even in different situations.

Time is not a great healer. Over time there will inevitably be a turnover of staff, and it is inevitable that there will be erosion of quality within the clinic if this possibility is not anticipated and addressed.

Kate's Action Plan goes as follows:

If an appropriate successor (for any team member) is to be selected and the remaining team is to have any say in the matter, what are the next moves? The factors you must be aware of (and deal with as far as possible) are:

- The stakeholders and influences.
- Start communicating at the earliest opportunity. Get into "communicating and sharing the ownership". Keeping a low profile is no longer an option in today's hospital services.

Meetings

Eventually the person who is the titular head of the voice clinic service will have to give a formal presentation to whoever is the Director of Planned Care

What is 'actually' happening when a voice clinic staff member leaves and you want to reappoint a successor? People's behaviour begins to change. Suddenly, there are lots of closed door conversations which may frequently exclude the key Voice Clinic clinicians. At this juncture, it may become apparent that certain key stakeholders are taking a lead to influence their chosen outcome. Certainly the 'quietly, quietly'

approach prevails. Generally, there is a complete lack of communication and inclusion of any therapists. The clinic staff will be waved away with a 'don't worry, they'll come to you'. No 'they' wont. And even if 'they' do, it will be too late.

Who are the stakeholders and decision makers?

- Identify who this process will impact on.
- Identify who would like to influence the process.
- Who manages the clinical area and areas the Voice Clinic impacts on?
- What is the most logical and financially sensible decision for the department and trust?
- Engage with them and start talking.
- Consistency.
- Verbally agree meetings, book with secretary.
- Email the agenda ahead (not too formally).
- Take accurate notes.
- Do the minutes after meetings 'we discussed . . . '
- Don't miss anyone, keep everyone informed and equal in the process
- Include: all consultants, nurse manager, lead clinician for head & neck, etc.

Good PR helps too . . . think about possible media opportunities. For instance:

1. Internal hospital magazine.
2. Local newspaper.
3. Radio.
4. Discuss all opportunities and availability with the Media Office.
5. Social media.

Formal meetings should be agreed and involve both the senior consultants and the Joint Voice Clinic (JVC) team.

At these meetings it is strongly suggested that you make comprehensive notes with key points. Agree to discuss in departmental meeting in order to ensure dissemination and discussion among all consultants and the surgery manager, take minutes and assign actions to the person/people named in the meeting.

The bid document

- Already in process.
- Formal structure – take advice about the correct format.
- Description and rationale.
- JVC achievements and personal career success.
- Options appraisals.
- References.
- Supporting letters from consultants! (& long nights).

What should be in it?

- Executive summary.
- Contents (linked).
- Introduction.
- Local context.
  - Background.
  - Why do we need a Voice Clinic?
  - The Voice Clinic Service.
- The Voice Clinic Team.
- Referrals and case load.
- Involvement in medical training.
- Succession planning.
- Objectives of this business case.
  - Clinical drivers to maintaining a Voice Clinic Service*
  - Direct financial benefits.
  - Expertise to lead a Voice Clinic.
  - Voice Clinic challenges.
  - Options appraisal.
  - Do Nothing.
  - Service to be lead by what grade of doctor?
  - Is it possible to convert position to a substantive Consultant position?

* N.B. The service bid document should address: 'What are the clinical drivers required to maintain an Integrated Voice Clinic Service?' as fully as possible in easy to read subsections. Viz.:

The clinical drivers that support the need for a robust and effective voice clinic at [hospital group name] are as follows:

- Maintain provision of a high quality and efficient voice clinic service to the population of [area served] and beyond, nationally.
- Provide support to ENT services e.g., accepting complex voice problems, laryngeal reflux patients and other complex benign laryngeal difficulties.
- Support head and neck services e.g., voice injury due to planned surgery and predicted radiotherapy changes in the larynx.
- Support for a speech & language therapy led voice clinic, FEES (flexible endoscopic evaluation of swallowing) clinic.
- Support to laser laryngeal follow up clinic.
- Support to aging population and increase in age related voice problems.
- Support to trust wide services e.g., management of unavoidable voice problems.
- sustained during planned treatment such as thyroid surgery or radiotherapy to the neck and chest.
- Voice clinics may attract a multidisciplinary tariff and they remain highly cost effective for the Primary Care Trust.
- Economic benefits of reducing the number of work days lost due to dysphonia.
- Cost savings through avoidance of unnecessary visits to GPs and ENT departments by avoiding incorrect diagnosis, management and unnecessary follow up appointments.
- Reduction of need for (and costs and risks of) diagnostic procedures under general anaesthetic.
- Potential savings on expensive medications as correct diagnosis is made earlier.
- Voice clinics rate highly in patient satisfaction audits.

Remember to share the bid document

- Presentation(s) to the department personally and in meeting.
- Email the document and book a meeting with Director.
- Presentation to Director of Planned Care with Deputy Manager.
- Outcome: there will be 'something'.

Interviews

- Advertised job may have subspecialty interests other than laryngology/voice weighting in the description.
- You may find potential candidates names who have a limited interest in specialist voice disorders being discussed casually in conversation.
- Do any of the candidates have the expertise required to lead a voice clinic? The skills required to run a successful voice clinic service takes many years to develop and become proficient at interpreting stroboscopy and the other tests of laryngeal function in order to achieve an accurate diagnosis and to plan appropriate and cost effective management programmes.
- Prepare for the questions you may be asked before the interview.

Staffing levels having been achieved, if you are starting up an integrated voice service, you are not through yet. The money to run it comes from elsewhere outside the hospital.

'He who pays the piper calls the tune'

(For this outline we are greatly indebted to *Dr Neil Douglas*, Clinical Lead Child Health Tower Hamlets Clinical Commissioning Group, HSJ CCG of the Year 2014. Eds.)

The UK Parliament allocates the cash to be spent on the National Health Service (NHS) (£116.4 bn in 2015). Of this, NHS England and 13 area teams split £69.2 bn amongst 209 Clinical Commissioning Groups (CCG's) for them to commission and pay for all hospital, mental health and community services. Their own funding (£12.8 bn) for primary care comes directly from NHS England and the area teams. The future and fate of all NHS voice services

lies within the purview of your local family doctor or Clinical Commissioning Group (CCG).

It is therefore imperative that you are well prepared with the right material – and the right buzzwords within – when submitting a proposal for voice clinic services within the ENT department.

They will know little or nothing about your proposed specialized service, so you need to be armed with information such as:

- the number of individuals who you anticipate will require the service;
- the cost of providing the service;
- the people who are able to provide the service; and
- the financial implications for the CCG if required to try and arrange provision of service or facility themselves.

They will want to know that the proposed service is truly patient centred, and that the service design includes full documentation of outcomes that matter to the patients; i.e.

- Tier 1. Health status achieved or retained.
- Tier 2. Process of recovery time and disutility.
- Tier 3. Sustainability of health – recurrence and long term consequences of therapy.

Be aware of new buzzwords in management speak. Don't even think about calling your service a 'multidisciplinary anything'. For the foreseeable future, the buzzword is now 'the Integrated Voice Clinic/Service'. Apparently, in manager-speak 'multidisciplinary' smacks of squandering staff resources, while 'integrated' suggests fewer patient visits. Well, the latter may even be correct.

Tell them that it is coordinated across the pathway system, that it is a personalized, multidisciplinary approach, and show them that the data that you are collecting is now 'measureable patient outcome' rather than the basic catch-all 'performance' based figures. Primary Care Commissions (PCCs) like applications to be as evidence based as possible.

But, be a little careful about the latter, there are 86 indicators recognized within the NHS Outcomes framework and not one mentions voice!

Finally, don't forget the different layers of feedback that you welcome from patients, students and colleagues and system level analysis. Do remember to tell the commissioners how the voice service takes its responsibilities to educate and train colleagues and juniors very seriously, especially as the specialist training in phoniatry that is widely available in Europe and North America is not yet provided by the relevant Royal Colleges in the UK .

TMH. London, 2016

References

Clarke, DD. 2004. Structured judgement methods: the best of both worlds? In: Todd, KZ, Nerlich, B, McKeown, S, Clarke DD (Eds.) *Mixing Methods in Psychology: The Integration of Qualitative and Quantitative Methods in Theory and Practice*. London: Routledge.

Dejonckere, PH, Bradley, P, Clemente, P, et al. 2001. A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. Guideline elaborated by the Committee on Phoniatrics of the European Laryngological Society (ELS). *European Archives of Oto-Rhino-Laryngology* **258**(2): 77–82.

Harris, T, Collins, S, Clarke, D. 1987. The effectiveness of the voice clinic. In: *Transcripts of Symposium: Recent Advances in Voice Conservation*. Nærum, Denmark: Bruel & Kjær, pp. 40–51.

Morrison. M, Rammage, L. 1994. *The Management of Voice Disorders*. San Diego, CA: Singular Publishing Group, pp. 1–2.

Smith, GCS, Pell, JP. 2003 Hazardous journeys: parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. *British Medical Journal* **327**:1459.

PART I: OUTLINE OF THE STRUCTURE AND FUNCTION OF THE VOCAL TRACT

1

The structural anatomy of the larynx and supraglottic vocal tract: a review

John S Rubin

Those who have dissected or inspected many [bodies] have at least learnt to doubt; while others who are ignorant of anatomy and do not take the trouble to attend it are in no doubt at all.

Giovanni Battista Morgagni, anatomist, 1682–1771

This chapter is primarily intended as an 'aide-mémoire' for doctors and manipulative therapists; the descriptive language has therefore not been modified for more general readership.

The larynx

The larynx is an intrusion into the pharynx, rather like the rearwards-pointing ventilation funnel of an old-fashioned ship pushed into an inverted cone. It subsumes several functions.

Airway protection: this is the larynx's most important function. The larynx is a complex sphincter that acts to protect the upper airway from saliva and food particles. The supraglottic structures act both passively and actively to direct the food bolus in streams away from the glottic chink, towards the pyriform fossae and, thence, through a relaxing cricopharyngeus muscle, into the oesophagus. Many of the intrinsic laryngeal muscles act predominantly in this capacity, thereby depressing and tilting the epiglottis posteriorly. This mechanism, together with adduction and medial compression of the true vocal folds, cannot be overly emphasised as it is the critical phylogenetic role of the larynx and thus, essential to survival. Maladaptation of the involved structures is common and represents one of the more frequent causes of visits to voice clinics.

Pressure-valving: this function is also of particular importance from a survival standpoint. Here, the larynx functions in the role of closing off the airway to prevent ingress or egress of air. This allows for sudden increases in intrathoracic and intra-abdominal pressures and thus permits such activities as childbirth, defaecation, weight-lifting, forceful micturition and vomiting. It is also integral in the protective act of coughing, whereby forced laryngeal closure permits build-up of subglottic pressure followed by sudden laryngeal release with forceful expulsion of air. It has been observed that

pressure-valving involves not only closure at the level of the true vocal fold, but also at the level of the false vocal fold (Dickson and Maue-Dickson, 1982). Again, hyperfunction of this mechanism represents a common cause of patient visits to the voice clinic.

Phonation: although phonation is also important for survival (screaming to frighten off an attacking animal or to attract attention), in animals other than man this function has far less intrinsic importance than airway protection and pressure-valving, and is phylogenetically a more recent activity. The capacity to produce very complex phonatory activity is limited to human beings and has been related by Laitman et al. (1995) to the lowering of the laryngeal complex from the basicranium. This function is primarily subsumed by intrinsic laryngeal muscles related to the arytenoid cartilages by the medial portion of the thyroarytenoid muscle (the vocalis) and by one 'extrinsic' muscle – the cricothyroid muscle. More will be written of these muscles and the biomechanical activities necessary for phonation later in this chapter and in subsequent chapters.

In general, it is useful to think of the larynx as a structure consisting of cartilaginous, bony and membranous components into which a series of muscles connect, and which is draped by mucous membrane; then to visualise this structure suspended from the skull base and mandible above by muscles, membranes and ligaments, and from the trachea directly and from the clavicles and sternum by muscles below.

In infancy, the larynx is located at C2–3. It proceeds thereafter to descend throughout life to C5–6 in adulthood and perhaps as low as C8 in senescence.

We prefer to structurally review the larynx from below up.

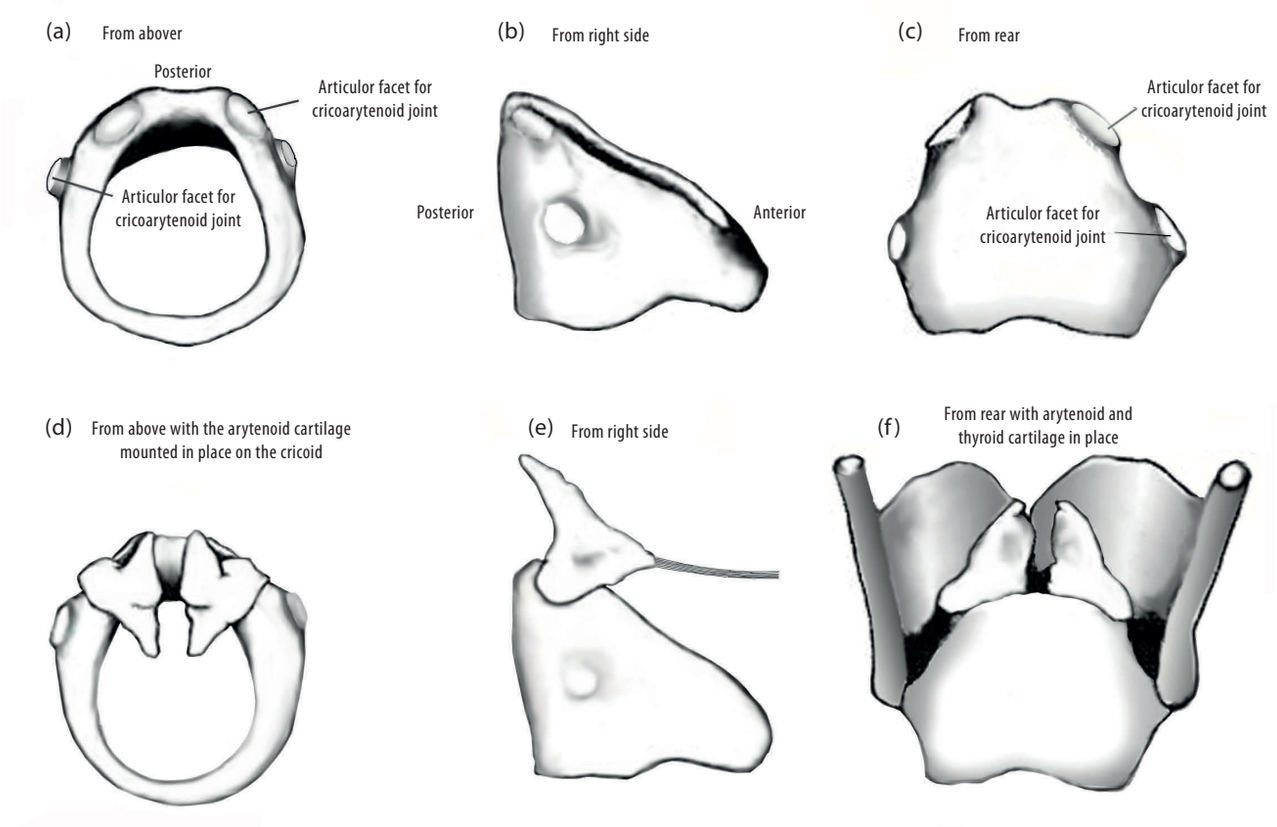


Figure 1.1 The cricoid cartilage. (a) From above. (b) From right side. (c) From rear. (d) From above, with the arytenoid cartilages mounted in place on the cricoid. (e) From right side. (f) From rear, with arytenoid and thyroid cartilage in place.

Cartilages and bones

Cricoid cartilage

The cricoid is the only completely circumferential structure in the larynx and thus, is essential for the provision of adequate rigidity of the laryngeal framework. It can be thought of as the foundation upon which the larynx is built.

The cricoid is named from the Greek description implying 'signet ring-shaped'. It consists of hyaline cartilage and is narrow anteriorly and very broad posteriorly. The narrower anterior portion is known as the arch. The posterior body when seen from the rear is like a vertical trapezoid with the lower margins of the sides chamfered off. In the posterior midline, a vertical ridge separates the lamina into two halves; inferiorly, this ridge broadens. According to Maue (1970) and Maue and Dickson (1971), the average posterior height is 25 mm in the male and 19 mm in the female.

The inferior edge of the cricoid cartilage is firmly attached to the first tracheal ring. Anteriorly, the superior aspect of the cricoid arch is attached to the thyroid cartilage via a thin, avascular membrane, and its anterior condensation, the anterior (median) cricothyroid ligament. Laterally, at or near the junction between the arch and the body of the cricoid, are the cricothyroid articular facets. These articulate with corresponding facets on the inferior horns of the thyroid cartilage to form the cricothyroid joints. The joints are plane synovial in type; the facets on each side of the larynx are frequently grossly asymmetric, one to the other. They face dorsolaterally and slightly superiorly (Dickson and Maue Dickson, 1982: p. 156). In addition to the joint capsule, there are two ligaments stabilising the joint: the posterior cricothyroid and lateral cricothyroid ligaments. The former ligament prevents spreading of the inferior thyroid horns and the latter limits, but does not abolish, posterior displacement of the thyroid over the cricoid. The functional potential of this joint is governed by the obliquity of the joint facets that somewhat limit the posterosuperior to anteroinferior motion. Rotation in a vertical plane is possible, with opening or closure of the cricoid arch and lower margin of the thyroid angle being limited by the anterior cricothyroid ligament or by contact between the same structures. Through an investigation on fresh human cadaver larynges, Dickson has identified correlative vocal fold changes in length through vertical movement of this joint of approximately 25% (Dickson and Dickson 1971; Maue Dickson 1982).

Thyroid cartilage

The thyroid cartilage is shield-shaped (thus its name in Greek). It consists of hyaline cartilage, which begins to ossify in the third decade of life. It is constituted by two pentagonal laminae that meet in the midline, the angle at the join being more acute (90°) in the male than the female (120°) and accounting for the 'Adam's apple'. Superior and inferior horns arise from the posterior edges of the laminae. The inferior horns have already been noted to articulate with the cricoid cartilage. The superior horns, as well as the entire superior edge of the cricoid cartilage, are connected by the thyrohyoid membrane to the hyoid bone. Laterally, this membrane condenses to form the paired lateral thyrohyoid ligaments. Medially, it gives rise to the medial thyrohyoid ligament. Laterally the thyrohyoid membrane is pierced by the superior laryngeal nerve (internal branch) and vessels. On the lateral aspects of the two pentagonal laminae of the thyroid are the two oblique lines. These serve as the origin of the deep layers of the strap muscles (the thyrohyoid and sternothyroid muscles).

The thyroid cartilage is covered externally by heavy perichondrium and internally by thinner perichondrium that is dehiscent over a small prominence where the anterior commissure of the true vocal folds attaches.

Hyoid bone

The thyrohyoid membrane connects the thyroid cartilage to the hyoid bone. Although not truly a laryngeal structure, and originating from a different embryologic

anlage, the hyoid bone will be mentioned here because it is important to laryngeal fixation and, thus, function.

The hyoid bone is a more or less horseshoe-shaped bone consisting of a central body anteriorly with two lateral projections, the greater and lesser cornua, the greater being the arm of the horseshoe and the lesser being a small posterosuperior projection from the superior surface of the junction between body and greater cornu. The hyoid bone attaches the larynx to the tongue musculature, the mandible and to the skull base via a series of muscular attachments. These muscles are important in laryngeal elevation given their capacity to pull the tongue body posteriorly and larynx either anteriorly or posteriorly. They include such muscles as the myoglossus, hyoglossus and geniohyoid, not to mention the mylohyoid, digastric muscles, stylohyoid, etc. The hyoid bone is also attached to most of the strap muscles (sternohyoid, omohyoid, thyrohyoid), which, upon contraction, act as laryngeal depressors (sternohyoid, omohyoid) or elevators (thyrohyoid). Although these muscles have a significant but minor role in the microcontrol of laryngeal pitch (Vilkman et al., 1996), by controlling the absolute length and configuration of the supraglottic vocal tract, they are more important in the production of the formants (enhanced resonances) of the supraglottic vocal tract.

These muscles are frequently misused in singing, and persistent hyperfunction may lead to a pitch-locked, fatigued voice with loss of appropriate vocal timbre. (See Chapters 5, 6, 7 and 8 for detail.)

Epiglottis

The epiglottis is a large, leaf-shaped structure made up of fibroelastic cartilage. The vocal folds insert – as previously mentioned – onto the inner surface of the thyroid cartilage. Just above this insertion is the petiole or attachment of the epiglottis via the thyroepiglottic ligament. The epiglottis extends superoposteriorly to the level of the base of the tongue where it overhangs the larynx. Its perichondrium is more tightly bound on its laryngeal than its lingual surface. Mucous glands are present on both surfaces although more are located on the laryngeal surface.

Superiorly, above the level of the hyoid bone, the epiglottis forms the vallecula with the base of the tongue via a series of mucosal condensations. Inferiorly, it defines the laryngeal inlet together with the arytenoid cartilages, via a series of mucosal folds (the aryepiglottic folds). As will be noted later in this chapter, a series of intrinsic laryngeal muscles insert into the epiglottis and are concerned mainly with laryngeal protection during swallowing. These muscles frequently contribute to the persistent hyperfunctional state.

Arytenoid cartilages

The paired arytenoid cartilages perch atop the posterior signet ring portion of the cricoid cartilage. Each cartilage is a dynamic pyramidal-shaped structure, apex up, which is integral to vocal fold motion. According to Maue (1970), the average height in adult males is 18 mm and in adult females 13 mm. Antero-posterior dimensions are, respectively, 14 mm and 10 mm.

The arytenoid cartilage is somewhat conical in shape, although its medial surface is flat, its posterior and inferior surfaces are concave and its antero-lateral surface is irregularly rounded. It has three pronounced angles: a sharp anterior angle, the vocal process (which makes up 40% of vocal fold length in the adult); a blunt postero-lateral angle, the muscular process; and a recurved superior angle – the apex (Boileau Grant, 1972).

The cricoarytenoid joint is a synovial load-bearing joint. An elliptical joint facet, 6 mm in length, sits on the posterior aspect of the rounded upper border of the cricoid. It slopes laterally, downwards and forwards. The deeply grooved base of the arytenoid cartilage has a corresponding facet that articulates with it, with the long diameter of the cartilage being set at right angles to that of the cricoid (Boileau Grant, 1972). This facet is on the inferior surface of the muscular process. This allows the arytenoid cartilage to function much like a sesamoid, turning the direction of pull of the muscles around the long axis of the joint (viz. the patella).

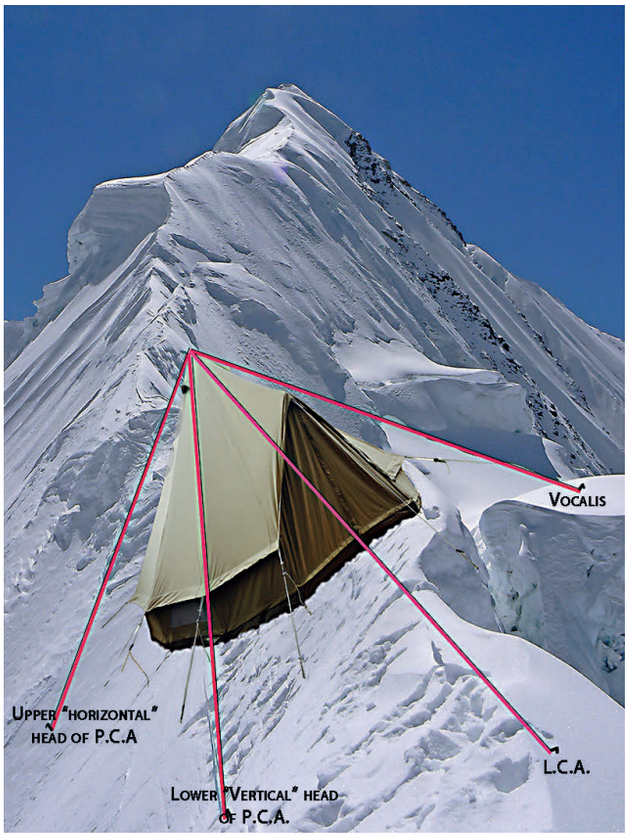


Figure 1.2 The arytenoid cartilage perches on the cricoid cartilage 'like a tent on the ridge of a mountainside'. The red guy-ropes indicate the approximate vectors of pull of several arytenoid muscles. The floor of the tent represents an approximation of the cricoarytenoid joint.

Two ligaments affect motion of the cricoarytenoid joint, the posterior cricoarytenoid ligament and the vocal ligament. The former is contiguous with the joint, attaches to the superior rim of the cricoid lamina between the two cricoarytenoid facets and extends anteriorly to the medial surface of the arytenoid cartilage. The latter attaches to the vocal process and extends anteriorly to insert onto the thyroid cartilage. There is also a tight fibrous articular capsule enclosing the cricoarytenoid joint (Dickson and Maue Dickson, 1982). We think that the primary function of the posterior cricoarytenoid ligament is prevention of lateral dislocation of the arytenoid on forced abduction of the vocal folds.

Motion permitted in the cricoarytenoid joint includes sliding along the long flat axis of the cricoid facet and rocking about the short convex axis of the cricoid facet. In fresh cadaver specimens, Dickson and Maue Dickson (1982) have demonstrated that only a few millimetres of sliding is possible due to the limiting joint capsule. Rocking (rotation), however, is quite free with the ligaments acting as guide wires.

Corniculate/cuneiform cartilages

The corniculate cartilages are very small pieces of elastic cartilage, roughly conical in shape, lying on, and frequently fused with, the apex of each arytenoid cartilage. The cuneiform cartilages (cartilages of Wrisberg) are very small bits of cartilage lying within the aryepiglottic fold, extending from the arytenoid to the epiglottis.

Membranes and spaces

Thyrohyoid membrane

The thyrohyoid membrane has already been described in its relation to the thyroid cartilage (see page 5).

Cricothyroid membrane

The cricothyroid membrane is a bilateral triangular-shaped structure; thus it has been referred to as the triangular membrane. To confuse the issue further, it has also been called the cricovocal ligament and the cricothyroid ligament. For the purposes of this discussion, the cricothyroid ligament will be considered to be the anterior (median) condensation of the cricothyroid membrane that extends between the cricoid arch and the deep inferior border of the thyroid cartilage back to the level of the cricoarytenoid facet. The borders of the cricothyroid membrane can be defined as follows: superior border, vocal ligament, anterior border, anterior cricothyroid ligament; inferior border, superior rim of the cricoid arch; posterior border, posterior end of the vocal ligament, attaching to the vocal process and inferior fossa of the arytenoid cartilage (Dickson and Maue Dickson, 1982). The body of this membrane is

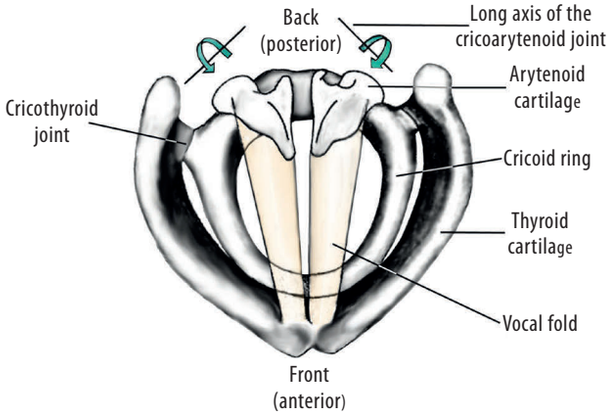


Figure 1.3 Principal cartilages of the larynx seen from above.

known as the 'conus elasticus' (Reidenbach, 1996) and is the supporting structure for the deeper surfaces of the vocal folds below the vocal ligaments.

Vocal ligaments

The vocal ligament is the condensation of the superior edge of the cricothyroid membrane. It extends from the end of the vocal process of the arytenoid and passes horizontally forward. Its anterior attachment to the thyroid cartilage is known as Broyle's ligament. It maintains the positional integrity of the true vocal fold.

Quadrangular membrane

The quadrangular membrane is a fibroelastic sheath. It is more delicate than the cricothyroid membrane. It extends bilaterally from either side of the epiglottis and curves backwards to the lateral border of the arytenoid cartilage. Its free upper edge is slightly thickened to form the aryepiglottic ligament. Its free lower border condenses to form the vestibular ligament, the basis for the false vocal (vestibular) fold. Its anterior vertical height is greater than its posterior vertical height.

Mucous membranes

Folds of mucous membrane clothe the cartilaginous and membranous structures of the larynx. Reviewing this on a coronal basis, mucosa lines the thyrohyoid membrane and a portion of the thyroid cartilage: folding upward, it ascends over the lateral border of the quadrangular membrane and then descends, draping over the medial aspect of the quadrangular membrane, thus creating the aryepiglottic fold superiorly and the vestibular fold inferiorly. Under the vestibular fold it pouches laterally to form the laryngeal ventricle. It then turns around medially along the floor of the ventricle to cover the conus elasticus and the vocal ligament and thence inferiorly to create the true vocal fold. It then continues inferiorly to become tracheal mucosa. Superiorly, it also drapes over the lingual surface of the epiglottis, creating the valleculae.

The epithelial lining of the larynx consists of respiratory epithelium (pseudo stratified, ciliated, columnar epithelium) and stratified squamous epithelium. Areas having direct contact with saliva and food particles, and those directly involved with phonation, are lined with the latter epithelium. These include the lingual surface of the epiglottis, pyriform sinuses, superior half of the laryngeal surface of the epiglottis, superior surface of the false vocal folds, the true vocal folds and subglottic aspects of the true vocal folds.

In the clinical/surgical literature, all the structures of the larynx above the vocal folds are generally collectively referred to as the supraglottis. This title is not very helpful as it is imprecise and refers to all the structures above the vocal folds and below the aryepiglottic folds/epiglottis as well as the airspace that it contains. For this reason we prefer to discuss the structures under the title of *epilarynx*, although we acknowledge that it is often referred to as the *laryngeal vestibule* (Fink, 1975: p. 36; Painter, 1986, 1991; Zemlin, 1998: p. 117; Titze, 2008: p. 2734). It will be further discussed in Chapter 2.

The muscles

The intrinsic muscles of the larynx

All the intrinsic muscles of the larynx – with the exception of the vocal fold tensing muscles, the cricothyroid muscles – have a common function of origin: they are

all originally parts of sphincteric mechanisms for sealing the larynx and lower respiratory tract off from the digestive tract. With the exception of the cricothyroid muscles and the most lateral (and vertical) fibres of the thyroarytenoid – the thyroepiglottic muscle, which are inserted instead into the epiglottis – the rest all have attachment to the arytenoid cartilages. Fortunately, the intrinsic laryngeal muscles have descriptive names outlining their attachments:

1. posterior cricoarytenoid
2. lateral cricoarytenoid
3. oblique interarytenoid
4. aryepiglotticus
5. transverse interarytenoid
6. thyroarytenoid, functionally subdivided into two parts: medially, vocalis muscle fibres and laterally, lateral thyroarytenoid or muscularis fibres.

The thyroarytenoid muscles also send fibres almost vertically to attach not to the arytenoid but to the epiglottis. These are the thyroepiglottic fibres of the thyroarytenoid – now generally considered to be a muscle in its own right (the 'thyroepiglotticus' muscle) – and the cricothyroid.

Note: the cricothyroid muscles are situated outside the laryngeal cartilages and have a different innervation to that of all the other muscles, the external laryngeal nerve. All the other intrinsic muscles are innervated by the recurrent laryngeal nerve. It therefore seems probable that, unlike all the other muscles, the cricothyroid muscle was never an evolutionary requirement for sealing the larynx. (Suzuki, M., et al. 1970)

All of the intrinsic laryngeal muscles are paired in a broadly mirror-image arrangement to right and left of the midline sagittal plane. The only exceptions are the transverse fibres of the interarytenoid, which are attached to identical sites on both arytenoids at each end of the muscle. Even after a unilateral recurrent laryngeal nerve palsy, these fibres will continue to cause some abduction of the vocal fold.

The posterior cricoarytenoid (PCA) muscles

These muscles are shaped like the fanning out of two scallop shells. At their broad ends they are inserted

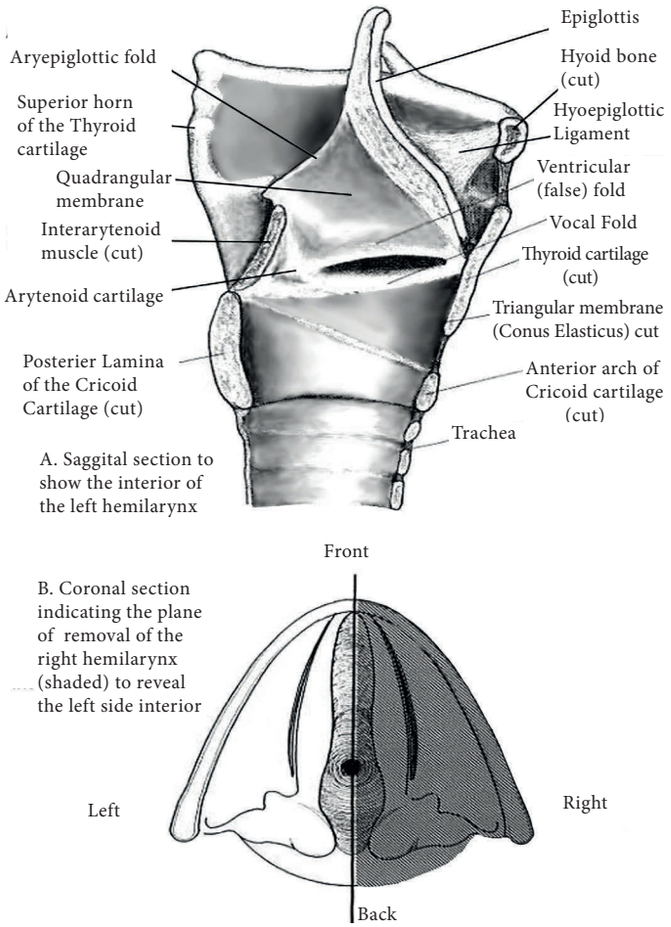


Figure 1.4 Sagittal section of the larynx showing the interior of the left hemilarynx. (a) Sagittal section to show the interior of the left hemilarynx. (b) Coronal section indicating the plane of removal of the right hemilarynx (shaded) to reveal the left side interior.

into the back of the body of the cricoid cartilage, to the left and right of the midline. The muscle fibres travel outwards and upwards converging like the ribs on the shell to form the strong musculotendinous attachment covering the whole superior surface of the muscular process of the arytenoid cartilage.

The action of the upper or 'horizontal' bundle of fibres of the posterior cricoarytenoid is to pull the arytenoid cartilage backwards, lengthening the vocal ligament and vocalis muscle. The direction of pull around

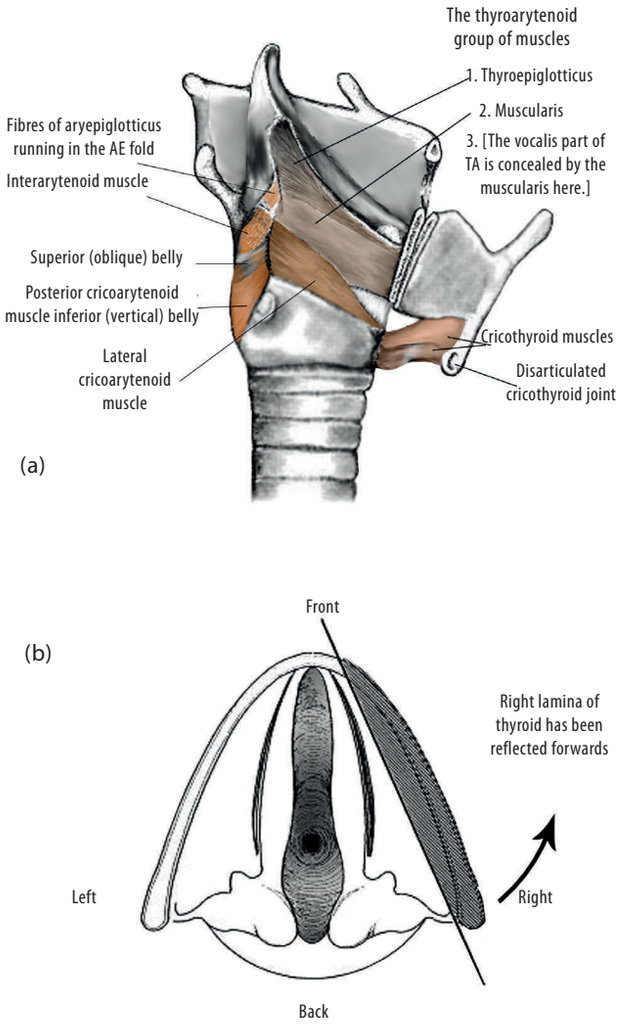


Figure 1.5 Right lateral view of laryngeal intrinsic muscles with the right thyroid lamina reflected forward. (a) Right lateral view of laryngeal intrinsic muscles with the right thyroid lamina reflected forwards. (b) Coronal view to show the plane of the right thyroid lamina which has been rotated forward.

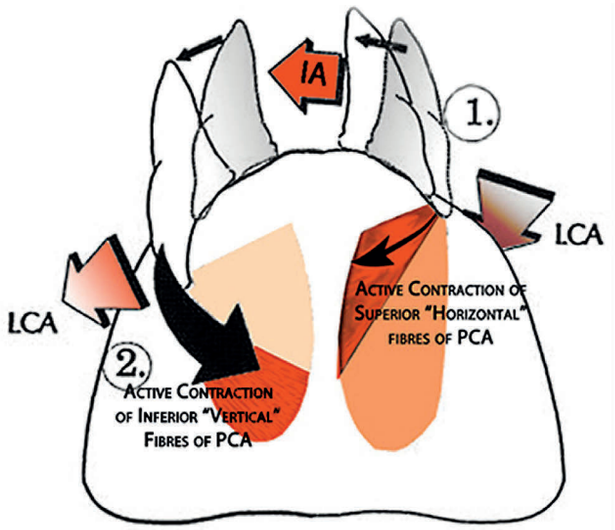


Figure 1.6 Diagram of the posterior aspect of the cricoid body and arytenoid cartilages to show the action of the posterior cricoarytenoid muscles. Left side: muscles involved in medial compression (IA and LCA). Right side: muscles involved in abduction (PCA and LCA).

the long axis of the cricoarytenoid joint stretches the vocal fold backwards in a neutral 'cadaveric' position with the free borders of the vocal folds lying slightly separated. It is the only group of muscle fibres in the larynx capable of tightening the vocal folds, all the others to a greater or lesser degree pull the arytenoid cartilage forward into the larynx, shortening the vocal ligament and slackening the vocalis muscle. The inferior or 'vertical' bundle of muscle fibres rotate the cartilage upwards and outwards away from the glottic lumen. This action is classically described as abduction (drawing apart) of the vocal folds. Recent observation suggests that PCA's parting of the folds may be assisted by another muscle, the lateral cricoarytenoid, which probably acts as a stay, thereby altering the vector of pull of the PCA into a downward and outward direction.

Despite all the published literature in many totally respectable sources suggesting that this muscle alone is responsible for vocal fold abduction, even simple inspection of the structural conditions operating in the larynx shows that this simplistic statement cannot be wholly correct (see 'Muscle fibre types' at the end of this chapter).

The lateral cricoarytenoid (LCA) muscles

These muscles are small and originate on the superior rim of the cricoid cartilage. The muscle fibres lie in a posterosuperior and medial direction, crossing the top of the cricoid to attach to the lateral half of the muscular process of the cricoid cartilage. The classical description of this muscle's action is that of 'medial compression' or approximation of the tips of the arytenoid vocal processes. A more logical description is given below. It is pointless to describe the activity of this muscle in isolation, that is, a muscle that slides

the arytenoid a little bit laterally along the long axis of the cricoarytenoid joint. Physiologically its activity is always in tandem with other muscles. With simultaneous activity of the oblique interarytenoid fibres – and maybe even the contralateral aryepiglottic muscle – then the action becomes sphincteric, causing the arytenoid to rotate anteromedially about the long axis of the cricoarytenoid joint, thus bringing the tips of the vocal processes together ('medial compression'). With simultaneous activity in the PCA, the result will be to negate the adductive capacity of the most superior and medial fibres of the PCA and to alter the pull of the most lateral and oblique fibres into a more strongly abductive direction. This combined activity appears to produce a complex combination of posterior rotation about the long axis of the cricoarytenoid (although much less than by the unrestricted activity of PCA alone), plus some lateral slide along the same axis together with lateral tilting. The overall effect when seen from above is abduction of the vocal fold with limited elongation of the folds (Figure 1.1).

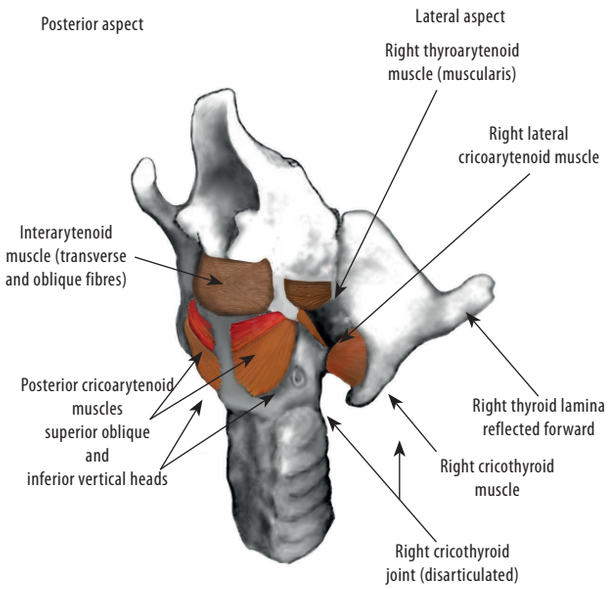


Figure 1.7 Intrinsic muscles of the larynx: right posterolateral view with the right thyroid lamina swung forward and outward to reveal the muscles.

(For a more detailed explanation, read the section on 'pulleys' in chapter 5, pp. 61–2.)

The aryepiglotticus muscle

This is really the prolongation of some of the fibres of the oblique arytenoid muscle running anteriorly in the aryepiglottic fold from the arytenoid apex to attach to the epiglottis.

The transverse arytenoid or interarytenoid muscle

This muscle is attached at both ends to each arytenoid cartilage. The attachments extend over the entire length of the dorsolateral ridge and dorsomedial concave surface of the arytenoids. Despite being an unpaired muscle, there is a double innervation from both recurrent laryngeal nerves indicating a double embryological origin of the muscle. It would appear that it opposes the natural tendency for the arytenoids to separate by sliding downwards and out when the vocal folds have been adducted and tension is being applied to the vocal fold/PCA by cricothyroid activity.

The interarytenoid muscle oblique fibres

The oblique arytenoid muscles originate on the medial half of the muscular process of the arytenoid. The muscle fibres lie in a similar plane to those of the LCA; i.e. they run upwards in the direction broadly following the circumference of the cricoid, crossing the midline to attach to the apex of the contralateral arytenoid.

The thyroarytenoid muscle

The thyroarytenoid muscle is a complex structure with two distinct portions of differing function. It arises from the deep surface (back) of the thyroid lamina lateral to the insertion of the vocal ligament. Almost L-shaped in cross-section, the upper, lateral portion is a flat sheet lying deep to the quadrangular membrane within the aryepiglottic fold. There is a considerable increase in muscle bulk inferiorly where its medial

(vocalis) fibres travel posteriorly with the vocal ligament forming the shelf-like body of the vocal fold. The posterior ends of both parts of the muscle are widely attached to the arytenoid cartilage; the fibres of vocalis (the lower, medial portion) attach to the inferior fossa and vocal process of the arytenoid cartilage, and the upper lateral thyroarytenoid fibres attach to the whole of the dorsolateral ridge of the cartilage.

That there is an evolving differentiation of function of the two parts of this muscle is borne out by recent comparative studies of the muscle fibres in human larynges and those from other species. In canine larynges, both parts of the muscle appear to be innervated by a mixture of 'fast-twitch' and 'slow-twitch' fibres, such as would be suitable for a mixed function of rapid sealing of the larynx and primitive voicing. In man, the lateral thyroarytenoid retains the twitch capacity in the lateral fibres but the superior vocalis portion has become expanded and is entirely populated by 'slow-twitch' fibres in very small motor pools suitable for the sustained and very accurate muscle contractions needed for voicing (Sanders, 1995).

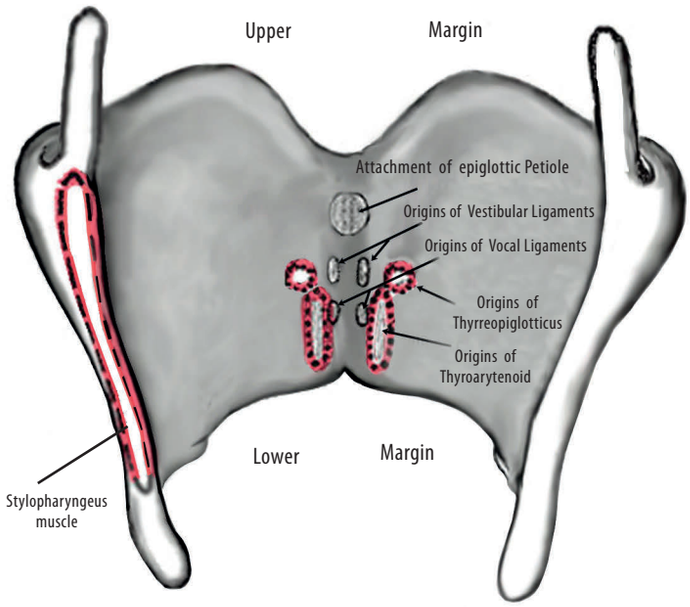


Figure 1.8 Posterior view of thyroid cartilage showing anterior muscular and ligamentous attachments. (After Gray).

The thyroepiglotticus muscle

The thyroepiglotticus muscle, although it is still sometimes considered simply as a separate head of the thyroarytenoid muscle, has quite a distinct function in sealing the larynx by drawing down the epiglottis.

Gray describes a separate origin lateral to the origin of the ventricular ligament and superior to the origin of the thyroarytenoid muscle on the interior of the thyroid cartilage. Some fibres appear to originate together with the superior fibres of the thyroarytenoid. The thyroepiglotticus ascends posterosuperiorly within the aryepiglottic fold. Together with the muscularis belly of the thyroarytenoid, it helps control the configuration and degree of constriction of the laryngeal inlet (aryepiglottic folds) and ventricular bands (false vocal folds). Sometimes these are collectively referred to as the muscles of the quadrangular membrane (see Figures 1.4 and 1.5a and b earlier in this chapter, and the the section on the epilarynx in Chapter 2 for additional comments). The thyroepiglotticus muscle, although still sometimes considered simply as a separate head of the thyroarytenoid muscle, has quite a distinct function in sealing the larynx by drawing down the epiglottis. Gray describes a separate origin lateral to the origin of the ventricular ligament and superior to the origin of the thyroarytenoid muscle on the interior of the thyroid cartilage. Some fibres appear to originate together with the superior fibres of the thyroarytenoid. It ascends posterosuperiorly within the aryepiglottic fold. Together with the muscularis belly of the thyroarytenoid it is thought to help control the configuration and degree of constriction of the laryngeal inlet (aryepiglottic folds) and ventricular bands (false vocal folds).

The extrinsic muscles of the larynx

These will be outlined in functional groups.

Anterior strap muscles

The larynx is unique in the body because it is a semi-rigid articulated structure, and yet it lacks the passive support of a joint or joints that might connect it somewhere to the rest of the skeleton. As every voice user is aware, it is highly (and controllably) mobile in the neck in a broadly vertical plane, and to a lesser degree in the anteroposterior plane. The control of laryngeal position is not achieved by the tugging of the pharyngeal musculature alone, but by the coordinated activity of all the extrinsic strap muscles as well. The muscles can be broadly divided into functional groups.

Muscles that lower laryngeal position

Sternothyroid: This has its origin on the deep surface of the manubrium sterni and ascends to insert into the oblique line of the thyroid cartilage. It lowers the larynx.

Sternohyoid: The sternohyoid overlies the sternothyroid. This flat muscle also takes origin on the posterolateral surface of the manubrium sterni, and passes up over the thyroid cartilage to attach to the lower border of the body of the hyoid bone. It lowers the hyoid bone and, consequently, the larynx. Both the sternohyoid and the sternothyroid pull in a slightly anterior direction.

Omohyoid: This muscle originates on the superior border of the scapula. It ascends anteriorly to a position superior to the medial end of the clavicle, where the cervical fascia binds it in position. From there it turns superiorly to insert into the inferior border of the body of the hyoid, lateral to the insertion of the sternohyoid. Why the belly of this muscle is tethered in cervical fascia is not clear. The effect would seem to be to render a more anterior direction to its laryngeal depressor action.

Muscles that raise laryngeal position

Thyrohyoid: The thyrohyoid raises the thyroid cartilage with respect to the hyoid bone. This flat muscle originates at the oblique line on the thyroid lamina. It runs upwards deep to the other strap muscles to insert into the lateral part of the inferior border of the body and greater cornu of the hyoid. It maintains the stability of the thyrohyoid membranes and is the sole muscular suspension of the larynx under the hyoid bone. It can rotate the thyroid posteriorly at the cricothyroid joints if the cricoid has been prevented from passively following due to other antagonising forces. The posterior part of the lower attachment, however, the portion arising at the oblique line, would seem to have a vector of pull running through the cricothyroid joint; in other words the posterior part of the muscle has the capacity to draw the thyroid cartilage towards the hyoid with very little rotational force at the cricothyroid (CT) joints.

Geniohyoid: The geniohyoid muscle raises the hyoid bone with respect to the mandible (jaw) with additional anterior pull on the hyoid bone. It has attachments to the front of the body of the hyoid and the lower genial tubercle at the midline inside the lower margin of the mandible. It also pulls the hyoid anteriorly, assisting in the creation of space in the hypopharynx.

Mylohyoid: The mylohyoid muscle raises the hyoid bone with respect to the mandible (jaw) and elevates the body of the hyoid. It is a flat sheet of muscle arising from the entire mylohyoid line on the inner surface of the mandible. It is attached to its opposite number in a fibrous raphe stretching between the body of the hyoid and the symphysis (midline) of the mandible. It forms the floor of the mouth. Contracting the muscle not only raises the anterior body of the tongue, it also elevates the body of the hyoid.

Digastric: The digastric muscles raise the hyoid bone with respect to the mandible (jaw), with neutral elevation regardless of anteroposterior laryngeal position. There are two distinct muscles connected by a common tendon. The anterior belly arises from the lower border of mandible near the midline; the posterior belly arises on the skull base medial to the mastoid process. Their common tendon is bound to the body of the hyoid by a fibrous tunnel at the junction between greater horn and body. These muscles are important for laryngeal raising during swallowing.

For further details of the role of these muscles in the positioning of the larynx, and its effect on fundamental frequency, see Chapter 6.

Muscles of the supraglottic vocal tract

Pharyngeal constrictors

These will be reviewed from the bottom upwards.

The inferior constrictor

This muscle is actually in two parts.

Cricopharyngeus: This arises (a) between the origin of the cricothyroid muscle and the cricothyroid joint, and (b) from the tendinous band between the inferior thyroid tubercle and the cricoid. Fibres are more or less contiguous with the muscle from the opposite side; i.e., the median raphe is minimal and they are continuous and blend with fibres of the upper oesophagus. The resting tone of this muscle is high in order to prevent the intrathoracic oesophagus inflating like the lungs with inspired air. The muscle relaxes with the approach of a swallowed bolus of food. The muscle anchors the cricoid cartilage posteriorly, and therefore assists in the closure of the cricothyroid visor if the thyroid cartilage is tilted forward above it. The upper fibres may anchor the cricoid against the tug of the trachea and oesophagus.

Thyropharyngeus: This muscle arises from the oblique line and inferiorly from the lateral border on the thyroid lamina. The fibres insert posteriorly into the median raphe. The muscle contracts immediately after the passage of a swallowed bolus of food, effectively 'stripping' the bolus onward towards the stomach. The lower fibres are almost horizontal while the upper fibres are set obliquely, ascending posteriorly to their insertion into the raphe behind the pharynx. Because of the oblique setting of the fibres, contraction on swallowing or voicing produces much less tendency to rotate the thyroid cartilage backwards than might be expected; indeed, if the hyoid is being held forward, then the thyropharyngeus may actually tend to rotate the thyroid cartilage forwards on the cricoid.

Note that there is a weakness in the muscular coat of the pharynx between the cricopharyngeus and thyropharyngeus muscles known as Killian's dehiscence. Normally, the two parts of the muscle function independently but in a coordinated manner. Should the swallow become dyscoordinate for any reason, then any resultant waves of high pressure may produce a pulsion diverticulum or pharyngeal pouch through the weak point at Killian's dehiscence.

The middle constrictor

This is also a fan-shaped muscle arising from the lesser cornu of the hyoid and lower part of the styloid ligament together with the whole of the upper border of the greater cornu of the hyoid. The lower fibres, as before, insert into the median raphe to form a cone within the inferior constrictor. The upper fibres similarly overlap the superior constrictor so that the arrangement is broadly one of concentric funnels inserting into lower ones. This arrangement not only assists the smooth progress of a wave of contraction to progress down the pharynx on swallowing, but when contracted independently of the inferior constrictor also permits an inverted U-shaped constriction in the pharynx, which is essential for the generation of a particular resonance characteristic, 'the singer's formant'. (Selective production of constrictions at different positions in the supraglottic vocal tract dictates the frequencies at which the formants of a sung or spoken vowel sound will occur. For a description of the resonance characteristics of the vocal tract, see Chapter 13.)

The superior constrictor

This is a complex, more or less quadrilateral sheet arising anterior to the pterygoid hamulus and on occasion to the adjoining posterior margin of the medial pterygoid plate, to the pterygomandibular raphe, to the posterior end of the mylohyoid line of the mandible

and to the side of the tongue. The fibres curve back to insert into the median raphe up as far as the pharyngeal tubercle of the occipital bone.

Unsurprisingly, the four parts are described according to their origin:

- pterygopharyngeal
- buccopharyngeal
- mylopharyngeal
- glossopharyngeal.

The constrictors are all supplied by the pharyngeal plexus; the motor and sensory fibres coming from trigeminal, glossopharyngeal and vagus (motor fibres of accessory origin). Pharyngeal rami come from the superior (jugular) ganglion. The autonomic sympathetic fibres are from the ganglion, and the parasympathetic fibres are postganglionic, mostly via the glossopharyngeal nerve.

Note: all of the above muscles do exactly what their names suggest: they constrict the pharynx, or in the context of this book, the major part of the supraglottic vocal tract. None of them is capable of actively dilating (increasing the volume of) the supraglottic vocal tract. Only the extrinsic strap muscles of the larynx – which, together with the suprahyoid muscles, are responsible for raising the larynx and drawing it forward (in the phase of swallowing immediately prior to a food bolus being squeezed from the back of the tongue into the pharynx) – are capable of active pharyngeal dilatation. This is of great importance when considering the shaping of the supraglottic vocal tract to make a tube that resonates correctly for production of vowel sounds or higher resonances ('formants') that give brilliance and carrying power to the singing voice.

The stylopharyngeus

This muscle originates on the medial side of the styloid process. It is a long thin muscle that initially travels down the outside of the pharynx, then passes inside between the superior and middle constrictors to spread out beneath the mucous membrane, some of the fibres blending into the constrictors, lateral glossoepiglottic fold and, with the palatopharyngeus, into the posterior border of the thyroid cartilage.

The stylopharyngeus is supplied by the glossopharyngeal nerve, it is a pharyngeal elevator and widener.

You cannot talk about swallowing or the modification of the supraglottic vocal tract for particular resonating characteristics without including the palatal muscles.

The palatal muscles

Levator veli palatini

This muscle rises from:

- the base of skull (petrous temporal bone) just anterior to the opening of the carotid canal;
- the fascia hanging from the tympanic temporal bone that forms the carotid sheath;
- the cartilaginous part of the eustachian tube.

It therefore lies inferior to the eustachian tube. It passes though the upper margin of the superior constrictor in front of salpingopharyngeus and spreads out in the soft palate between two layers of palatopharyngeus in the palatine aponeurosis. It elevates the soft palate.

Tensor veli palatini

This arises from the scaphoid fossa of the pterygoid process, the lateral lamina of the eustachian tube and the medial side of the spine of the sphenoid. The muscle descends, becoming tendinous to turn round the pterygoid hamulus, passes through the origin of buccinator, and inserts into the anterior palate in the palatine aponeurosis and palatine crest of the horizontal part of the palatine bone. It tenses the anterior, relatively unmuscular, part of the palate.

This muscle alone is innervated by the mandibular nerve; all the other palatal muscles are innervated by the pharyngeal plexus.

Palatoglossus

The palatoglossus arises from the palatine aponeurosis to the side of the tongue. It forms the anterior arch of the fauces.

Palatopharyngeus

This muscle arises in the palate from two fasciculi separated by the levator veli palatini. The posterior fasciculus arises jointly with its opposite number in the median plane on the pharyngeal side of the palate. The anterior bundle arises at the posterior margin of the hard palate from the palatine aponeurosis. It passes between the levator and tensor veli palatini and unites with the posterior bundle and also the salpingopharyngeus. This conjoint flat muscle passes laterally and downwards and is inserted into the posterior border of the thyroid cartilage together with the stylopharyngeus. It forms an internal longitudinal coat for the pharynx, and it also forms the posterior arch of the fauces.

The palatopharyngeus pulls the free border of the palate down. This may be the reason why increased nasality is often noted with a persistently raised, backed laryngeal position.

Salpingopharyngeus

The salpingopharyngeus originates on the inferior part of the eustachian tube cartilage. It travels downwards to blend with the palatopharyngeus. It may simply represent a slip of palatopharyngeus.

References

Boileau Grant JC. 1972. *Grant's Atlas of Anatomy* 6ed. Baltimore, MD: Williams and Wilkins.

Dickson D, Dickson W. 1971. Functional anatomy of the human larynx. *Transactions of the Pennsylvania Academy of Ophthalmology and Otolaryngology* **25**(1): 29–38.

Dickson DR, Maue-Dickson W. 1982. *Anatomical and Physiological Bases of Speech*. Boston, MA: Little, Brown.

Fink BR. 1975. *The Human Larynx: A Functional Study*. New York, NY: Raven Press.

Gray H. 1918. The cartilages of the larynx. Fig. 950 In: *Anatomy of the Human Body*. 1918 et seq. XI: Splanchnology. Section 1a: The larynx.

Laitman JT, Noden DM, Van De Water TR. 1995. Formation of the larynx: from homeobox genes to critical periods. In: Rubin, JS, Sataloff, RT, Korovin, GS, Gould WJ (eds) *Diagnosis and Treatment of Voice Disorders*. New York, NY: Igaku-Shoin.

Maue WM. 1970. *Cartilages, ligaments and articulations of the adult human larynx*. University of Pittsburgh, PhD dissertation.

Maue WM, Dickson DR. 1971. Cartilages and ligaments of the adult human larynx. *Archives of Otolaryngology* **94**: 432.

Painter C. 1986. The laryngeal vestibule and voice quality. *Archives of Oto-Rhino-Laryngology* **243**: 329–37.

Reidenbach MM. 1996. The attachments of the conus elasticus to the laryngeal skeleton: physiologic and clinical implications. *Clinical Anatomy* **9**(6): 363–70.

Sundberg J. 1974. Articulatory interpretation of the "singing formant". *Journal of the Acoustical Society of America* **55**(4): 838–44.

Suzuki M, Kirschner JA, Murakami Y. 1970. The cricothyroid as a respiratory muscle: its characteristics in bilateral recurrent laryngeal nerve paralysis. *Annals of Otology Rhinology and Laryngology* **79**(5): 976–83.

Titze IR. 2008. Nonlinear source–filter coupling in phonation: theory. *Journal of the Acoustical Society of America* **123**(5): 2733.

Vilkman E, Sonninen A, Hurme P, Körkkö P. 1996. External laryngeal frame function in voice production revisited: a review. *Journal of Voice* **10**: 78–92.

Zemlin WR. 1998. *Speech and Hearing Science: Anatomy and Physiology* 4ed. Boston, MA: Allyn and Bacon, p.117.

2

The functional anatomy of the vocal tract: some additional considerations

Tom Harris and Scott Moisik

Since the first edition of this book the authors, as part of the general re-think, have decided to re-work some of the more important functional aspects of the anatomy of the vocal apparatus.

One of the problems encountered when trying to comprehend the functioning of the laryngeal and pharyngeal musculature is that the images that can be acquired in a voice clinic in the clinical situation, using endoscopy with videostroboscopy or high speed video capture, etc., are generally views of the larynx as seen from above. These only produce two-dimensional visual information about three-dimensional events.

Everyday problems when viewing laryngeal function in the voice clinic

Without accurate knowledge of the nature of the cricoarytenoid joints, and the three-dimensional spatial position of the long axes of the joints, it is simply not possible to describe or interpret the actions of any of the muscles attached to the arytenoid cartilages accurately. This information may be summed up in the following diagrams. Disregarding these important details has allowed significant errors to persist in the textbook literature. For example: 'The posterior cricoarytenoid muscle is the only abductor of the vocal fold' – no mention that it is also the only muscle that can tension the vocal fold by opposing all the other musculature that pulls the arytenoid forwards.

If the declivity of the cricoarytenoid joint (CAJ) and the angle of its deviation from the sagittal midline are taken into account, then it becomes clear that the lower, vertical fibres of PCA will indeed *ab*duct the arytenoid, while the upper horizontal fibres *tension* the vocal folds and prevent their anterior collapse. (Zemlin et al., 1984).

Aggregation of the morphology of multiple larynges has given us accurate models of cartilages, joints and muscles, and these give us a precise formulation as to how different gestures may be obtained with vectors of pull by different muscle combinations. (Moisik, 2006, 2013; Moisik and Esling, 2014). The muscles themselves have become highly adapted both histochemically and morphologically to meet their specific functional requirements.

Muscle fibre types histochemically

Skeletal muscle is often divided biochemically into three broad types of fibre.

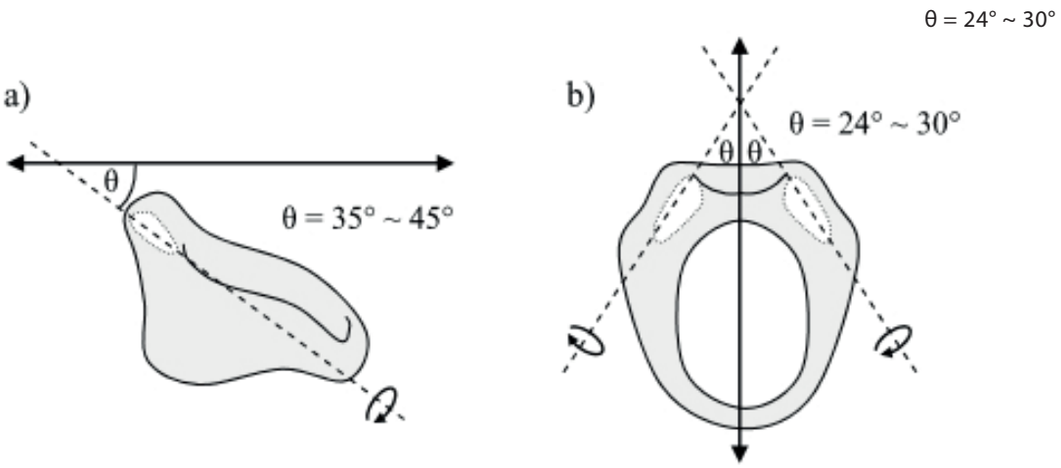


Figure 2.1 Angles designating the rotational axis of the arytenoid cartilage based upon the longitudinal axis of the articulatory facet for the cricoarytenoid joint on the cricoid cartilage. Image (a) illustrates the declivity of the axis, corresponding to about 35° ~ 45°. Image (b) shows the deviation from the sagittal midline of the cricoid cartilage, occurring at 24° ~ 30° (both sides are shown). Image reproduced with permission from Moisik 2008, MA Thesis, p. 59, fig. 3.6.

Type I fibres

Type I fibres are also known as slow-twitch fibres. They are red in colour due to the presence of large quantities of intracellular myoglobin together with high numbers of mitochondria. Due to this fact they are very resistant to fatigue and are capable of producing repeated low-level contractions by producing large amounts of ATP through an aerobic metabolic cycle.

For this reason, the muscles containing mainly type I fibres are often postural muscles – such as those in the neck and spine – due to their endurance capabilities. Also, athletes such as marathon runners have a high number of this type of fibre, partly through genetics, partly through training.

Type IIa fibres

Type IIa fibres are also sometimes known as fast oxidative fibres and are a hybrid of type I and II fibres. These fibres also contain a large number of mitochondria and large quantities of myoglobin, hence their red colour. They manufacture and split ATP at a fast rate by utilising both aerobic and anaerobic metabolism and so produce fast, strong muscle contractions, although they are more prone to fatigue than type I fibres. Resistance training can turn type IIb fibres into type IIa due to an increase in the ability to utilise the oxidative cycle.

Type IIb fibres

Often known as fast glycolytic fibres, they are paler in colour due to lower levels of myoglobin, and also contain fewer mitochondria. They produce ATP at a slow rate by anaerobic metabolism and break it down very quickly. This results in short, fast bursts of power and rapid fatigue. As mentioned above, this type of fibre can be adapted into type IIa fibre with training. This adaptation produces the superior resistance to fatigue shown by type IIa fibres.

In the larynx, the posterior cricoarytenoid muscles have two functionally distinct bellies. The horizontal belly has 80% slow (type I) fibres and 20% fast (type II) fibres. The vertical belly contains equal amounts of slow and fast fibres (~55:45%) and contains relatively equal amounts of fast and slow myosin heavy chain (MyHC) protein, clearly distinguishing

between the two compartments. Detailed information regarding the vertical and horizontal bellies of the posterior cricoarytenoid (PCA) muscle is essential information for understanding both normal voice production and a variety of vocal pathologies, such as vocal fold immobilisation and spasmodic dysphonia (Sciote, 2003).

A teleological, biomechanical analysis of the different bellies of the PCA allows inference of completely different functions for each of these muscle bellies (Brandon et al., 2003: Asanau et al., 2011).

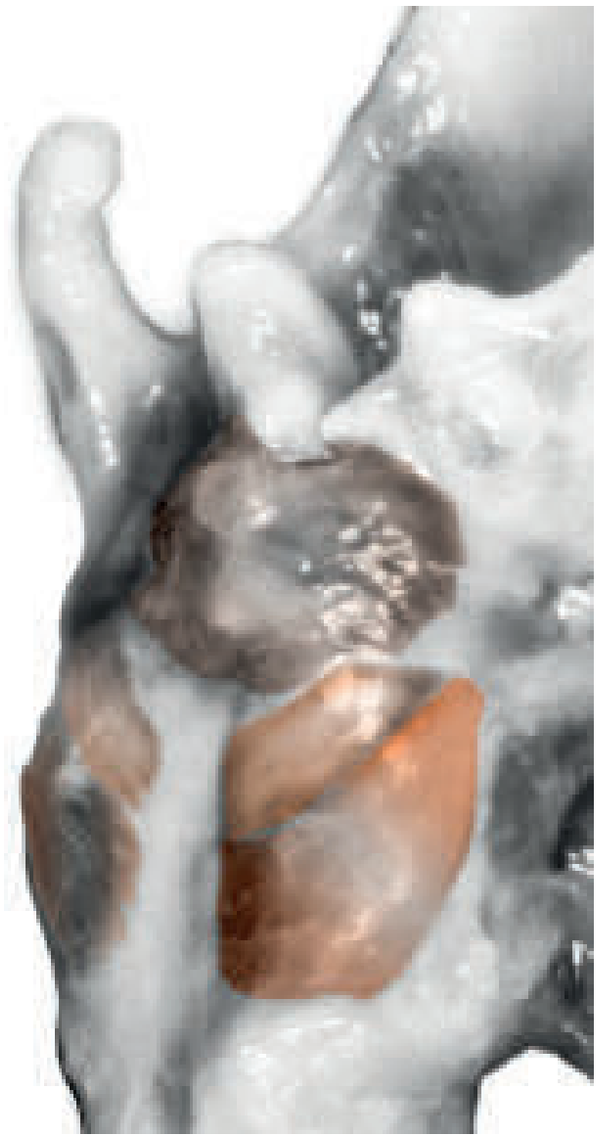


Figure 2.2 The contrasting appearance of upper horizontal and lower vertical bellies of PCA.

The difference in muscle fibre type supports the observations concerning the different vectors of pull of the two bellies, which suggest that the horizontal belly has adapted for long-term tensioning of the vocal ligament while the vertical belly would seem to be perfectly adapted for fast 'twitch' abductions of the vocal fold.

The morphology of individual laryngeal muscles also shows adaptation to different functional requirements. Zemlin et al. (1984) noted that the PCA is fan-shaped with two functional subunits, which significantly differ in their muscle fibre-type composition. It might also suggest that when a neuromuscular re-innervation technique is proposed to deal with a rough breathy voice following unilateral recurrent laryngeal nerve palsy, the superior belly of the PCA might provide the optimum site for implantation if the arytenoid on the affected side shows much anterior collapse.

The richness of the innervation of the motor units of muscle fibres is a good indicator of the type of activity required in a specific muscle. The thyroarytenoid and cricothyroid muscles present the smallest (9.8 +/- 0.2) and largest (20.5 +/- 0.9) motor unit size, respectively, suggesting that thyroarytenoid muscle has a greater capacity to fine tune its total force compared with the other intrinsic laryngeal muscles (Santo Neto and Marques, 2008).

For a much more detailed account of the microanatomy of the vocal fold musculature and functional subspecialization, the reader is referred to Sanders, 2014.

The glottis/vocal folds

The vocal folds are two composite folds comprising layers of muscle, connective tissue, and epithelium. All the layers are distinct and modified to suit the evolutionary functions of the larynx. They are the principal, but by no means the only, components of the laryngeal sound source:

Table 2.1: Different descriptions of the characteristics of vocal fold layers (after Hirano, 1981).

<table>
<tr><th>Three layer description
of physical characteristics</th><th>Five layer description
of histological characteristics</th><th>Two layer 'cover–body'
of functional characteristics</th></tr>
<tr><td>Mucosa</td><td>Epithelium</td><td rowspan="3">Cover</td></tr>
<tr><td rowspan="2">Ligament</td><td rowspan="2">Superficial layer
lamina propria
Intermediate layer
lamina propria
Deep layer
lamina propria</td></tr>
<tr><td rowspan="2">Body</td></tr>
<tr><td>Muscle</td><td>Muscle</td></tr>
</table>

The covering epithelium

From the surface inwards, the covering epithelium consists of a durable but flexible layer of non-keratinising stratified squamous epithelium between 0.05 and 0.1 mm thick, similar to that of the skin covering the body. It will permit a degree of stretching sideways, but is relatively incompressible. (In physical jargon it is said to be anisotropic.) This relatively waterproof layer covers and encapsulates a complex layer that appears to be largely made up of semifluid gel-like substance, the lamina propria.

The lamina propria

This second layer is actually much more sophisticated than initial impressions suggest: there are at least three major divisions of the lamina propria, a catch-all description for all the non-muscular material between the epithelium and the fibres of the vocalis muscle.

The superficial layer

This consists of a mixture of elastin fibres with a few dense birefringent fibres of type I collagen embedded as a deformable network within an interstitial fluid matrix composed of mucoproteins and mucopolysaccharides (Hirano et al., 1981). It is the fibroblasts of the superficial layer of the lamina propria (SLLP) that are responsible for the production and maintenance of the ground substance. This fluid-retaining material is amorphous but retains its structure as it is interspersed

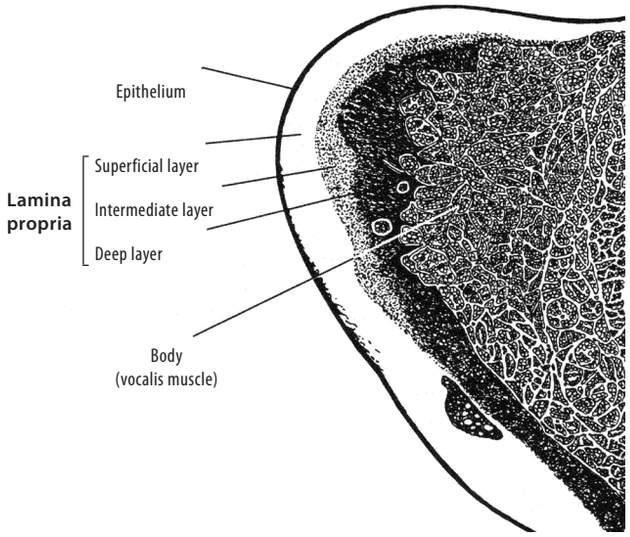


Figure 2.3 The vocal fold in cross-section: coronal view. After M Hirano.

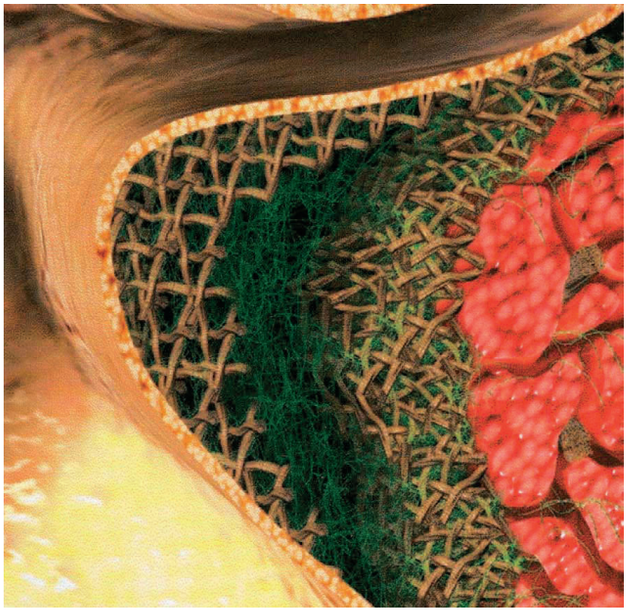


Figure 2.4 Types of collagen within the vocal fold (Madruga de Melo et al., 2003). Brown fibres represent type I collagen and green fibres type III. Reproduced by permission of John Wiley and Sons.

in the elastin/collagen framework of Reinke's space. The elastin fibres can stretch longitudinally to almost double their resting length and because they are not longitudinally aligned, the superficial layer is readily deformable in all directions (isotropic). The SLLP in health is a resilient, thin gel, which can be very considerably deformed and yet retain its structural integrity under conditions of great physical stress, and it is essential for the generation of mucosal waves on phonation. A possible reason for the traditional lack of surgical regard for this structure is that in health, it measures only approximately 0.5 mm in thickness in the mid-section of the fold.

The intermediate layer

Like the superficial layer, this is a blend of elastin and collagen fibres. The collagen is structured and arranged differently, however. The collagen fibres are finer (type III), densely packed and, although they anastomose and branch freely in a sort of wicker basket arrangement, the majority are aligned longitudinally between the anterior and posterior attachments of the vocal fold. There are aggregations of elastin fibres at the anterior and posterior ends of the vocal ligament known as the anterior and posterior maculae flava. They appear to function as shock absorbers at points of maximum impact on phonation. Collagen fibres are also present as a minor component in this layer.

The deep layer

The deep layer immediately underneath the intermediate layer is made up principally of the dense birefringent, relatively inextensible type I collagen fibres that are also aligned longitudinally in a coil-like fashion. This tendinous structure is intimately connected to the vocalis muscle with collagen fibres which penetrate its surface. The alignment alters inferiorly where the layer becomes contiguous with the fibres of the triangular membrane of the conus elasticus. Together, the intermediate and deep layers of the lamina propria are between 1 and 2 mm thick and make up the vocal ligament. The main stress applied to the ligament is one of extension, and the deformation under stress is described as being anisotropic.

The body of the vocal fold

The vocalis muscle

The vocalis muscle, the medial portion of the thyroarytenoid muscle, is 7–8 mm thick and lies deep and lateral to the lamina propria in a manner akin to muscle underlying a layer of fascia. This represents the major part of the muscle and is histochemically different from the muscularis portion. In the human larynx, the innervation arrangements are significantly different from other species: there is a gradient from the most medial fibres nearest the ligament (the superior vocalis) where the densely-innervated fibres are almost exclusively of the slow-twitch variety, through to the muscularis portion where the fibres are entirely of the fast-twitch variety (Sanders, 1995). This is reflected by the human capacity to produce a very sustained phonatory tone. The fibres are gathered in small bundles, or fascicules, and the bundles in the superior vocalis

appear, from observation of phonation, to be capable of differential activity for changing the shape of the vibrating free border to produce the variable thickness required for different registers and pitches. The inferior part of the vocalis is hypothesised to be the functional equivalent of the lower mass in the two-mass model of vocal fold activity.

Although the classical description of the vocal folds' laminated epithelial, lamina propria and muscular layers satisfies the morphological requirements, it does not lend much understanding to the mechanical ability of the vocal folds' ability to vibrate at widely varying pitches and intensities, without damage, under normal conditions. It is the interplay between all the layers, the isotropic layers being firmly sandwiched between the anisotropic, which allows for the generation of complex mucosal waves. It is the ability to apply tension longitudinally along the intermediate and deep layers of the lamina propria that permits thinning of the fold at its free border while at the same time producing an increase in the firmness of the fold. It is also the ability of the vocalis muscle to selectively stiffen in different configurations that permits a wide range of vocal timbres.

All the descriptions in Table 2.1 may therefore be regarded as valid; which one is used depends chiefly on whether the user is speaking or writing from a physical, morphological/histochemical or functional point of view.

From the standpoint of the surgeon operating on the vocal folds, the system falls nicely into the first, physical, category; the tissue planes naturally divide between the mucosa and the ligament. Unfortunately, tissue damage is no respecter of surgical planes, and the results of damage are reflected in the disruption of the normal arrangements of any of the five principle, histologically-discrete layers. Hirano's description of vocal fold vibration in a two-part system, the 'cover–body' theory, makes it quite obvious that all microsurgery must be designed to minimise any disruption of the vocal fold cover, while the aim of any 'framework' surgery has to be to produce as little damage as possible to the function of the vocal fold 'body'. This is why Isshiki type I thyroplasty is to be preferred, from a functional point of view, to injection of Teflon® paste or similar preparations. It also demonstrates clearly why any Teflon or collagen paste left between the intermediate and deep layers of the lamina propria can produce such unfortunate results when the patient attempts to voice post-operatively.

With all the stresses produced within the layers of the vocal fold cover during phonation, how is it that the epithelium does not simply strip off the superficial layer of the lamina propria? The answer appears to lie in the ultrastructure of the basement membrane zone, or the transitional area between the epithelium and superficial layer of the lamina propria. Underneath the basal cells of the epithelium, the electron microscope reveals that their plasma membranes contain structures known as attachment plaques. The basal cells overlie a layer known as the lamina lucida, which in turn contains aggregations corresponding to the attachment plaques known as sub-basal dense plates. These dense plates are anchored by microfibrils to a layer of type IV collagen/fibronectin, the lamina densa. It is this lamina densa of the basement membrane that maintains the structural integrity of the epithelium. Beneath it, things are even more complex. There is no evidence of collagen from the lamina propria being simply embedded into the lamina densa to secure adhesion; instead, there is a very subtle and flexible arrangement of type VII collagen fibres that helps to maintain the attachment of the basal cells/basement membrane to the SLLP. Both ends of these fibres arise in the lamina densa, and the midportions form loops that project into the superficial layer of the lamina propria. The thicker type III collagen fibres of that layer pass through the loops, as do other more deeply set loop fibres (Gray et al., 1994). Their appearance and position almost suggest a sort of ultramicroscopic layer of inverted staples attaching a web of fibres in the SLLP to the basal layer of the epithelium

Table 2.2: Interstitial component proteins of the extracellular matrix in the lamina propria (after Colton et al., 2011).

ECM constituent	Function	Localisation in normal lamina propria
Collagen (fibrillar protein)	Strengthens the lamina propria (Gray et al., 2000) .	Density increases across superficial layer of lamina propria to deep layer of lamina propria (Gray et al., 2000).
Elastin (fibrillar protein)	Provides the stretch and recoil of the lamina propria (Gray et al., 2000). Very hydrophobic.	Highest density in the intermediate layer of the lamina propria (Gray et al., 2000).
Hyaluronan (glycocaminoglycans)	Affects tissue viscosity, osmosis (very hydrophilic) and tissue dampening (Laurent et al., 1995).	Found throughout the vocal fold with the highest density being in the intermediate layer of the lamina propria (Butler et al., 2001).
Decorin (proteoglycan)	Inhibits collagen formation and promotes lateral association of collagen fibrils to form fibres and fibre bundles in the ECM. Binds to fibronectin (Iozo, 1997).	Found throughout the lamina propria. Highest density in the superficial layer of the the lamina propria (Gray et al., 1999).
Fibronectin (glycoprotein)	Induces cell migration and ECM synthesis. May be involved in the development of fibrosis (Ehrlich, 2000).	Found throughout the lamina propria including the BMZ (Gray et al., 1999).

in a manner that allows for gliding, bending and compression of the layers of tissue without shearing a plane between them.

It is now thought that loss or damage to these loops allows the shearing stresses produced by phonation to damage the vocal fold cover, blistering the epithelium off the SLLP and thus inducing the inflammatory repair that produces nodules.

This hypothesis is not incompatible with the standard teaching concerning the pathogenesis of polyps and polypoid degeneration (Dikkers, 1999): it may be that it is damage to this structure that permits unusual stresses to the microvasculature of the vocal fold cover, causing extravasation of blood and subsequent development of polypoid change if the process is repeated and becomes chronic.

It is the different stiffness characteristics of these layers which result in the mechanically de-coupled groupings of the layers to form:

- the mucosal cover – composed of epithelium and the superficial layer of the lamina propria;
- the vocal ligament (transition) – composed of the intermediate and deep layers of the lamina propria;
- the body of the vocal fold – the vocalis muscle.

Traditional descriptions of the vocal folds have centred around the cellular components of tissue (histology). However, in recent years advances in

technology have made the investigation of the extracellular components of tissue possible, and these extracellular components are collectively termed the Extracellular Matrix (ECM). The ECM suspends cells, chiefly fibroblasts, which are derived from mesenchymal stem cells and which retain the ability to differentiate along multiple cell lines. They are thought to be important in suppressing activation/proliferation of immune and inflammatory cells, thus allowing them to engage in the regeneration of tissue and repair of damage, with either minimal or significant scarring.

Besides the cellular component, the ECM is made up of connective tissue (mostly the structural) fibrillar proteins collagen and elastin, and interstitial (mainly adhesive) fibrillar glycoproteins, fibronectin and laminin. Critical to the physical properties of the lamina propria are large molecules of glycocaminoglycans (GAGs, unbranched chains of sugar molecules – (GAGs, e.g. hyaluronan) and proteoglycans (GAGs attached to a protein). The hyaluronan in particular has the ability to sustain great mechanical deformation without damage . (Thibeault, 2004; Hahn, 2006)

What other structures form part of the sound source?

To quote Esling et al. (2015), "the 'laryngeal articulator' – consisting of the glottal mechanism, the supraglottic epilaryngeal tube, the pharyngeal/epiglottal mechanism, and including three levels of folds: the vocal folds, the ventricular folds and the aryepiglottic folds – is responsible for the generation of multiple source vibrations and for the complex modification of the epilaryngeal and pharyngeal resonating chambers that account for a wide range of contrastive auditory qualities".

In the clinical/surgical literature, all the structures of the larynx above the vocal folds are generally collectively referred to as the supraglottis. This title is not very helpful as it is imprecise and refers to all the structures above the vocal folds and below the aryepiglottic folds/epiglottis as well as to the airspace that it contains. For this reason we prefer to discuss the structures under the title of epilarynx, although we acknowledge that it is often referred to as the laryngeal vestibule (Fink, 1975: p. 36; Painter, 1986, 1991; Zemlin, 1998: p. 117; Titze, 2008: p. 2734).

When reviewing the anatomical geometry of the epilarynx (Moisik 2013, PhD Dissertation: p. 1, fig. 1.1), we must consider two spaces: the laryngeal vestibule and the laryngeal ventricle (Fink, 1975: p.36;; Zemlin, 1998: p. 117).

The ventricle, on the other hand, is the cavity immediately above the vocal folds,bounded above by the caudal surface of the ventricular folds.

The vestibule communicates with this space superiorly, and the spaces can be (arbitrarily) separated by reference to the plane defined by the margin of the ventricular folds.

Such a bipartite division of laryngeal space is useful insofar as it correctly conveys that these two spaces become separated by the adduction of the ventricular folds. In aero-acoustic literature, however, the spaces are lumped together to comprise the epilaryngeal cavity: for example, both Sundberg (1974) and Titze (2008) regard the base of the epilarynx as being defined by the cephalad surface of the vocal folds.

The tube-shaped epilarynx is found at the bottom of the pharynx tube; together these structures define the lower vocal tract. The actions of these tubes are:

1. pharyngeal constriction, and
2. epilaryngeal constriction.

The epilarynx is important in the production of a wide range of voice qualities used in singing and speech. It helps a voice compete with an orchestra by assisting in the production of the singer's formant, and the anterior and posterior parts are both capable of narrowing/closing independently. The anterior compartment does so by squeezing the periepiglottic fat pad by the thyroepiglottic fibres of the thyroarytenoid (TA) muscle (Reidenbach, 1997). It can actively narrow to produce certain vocal styles such as 'belting' or 'twang' (Sundberg, 1974; Yanagisawa et al.,, 1989; Honda et al.,1995; Schellenberg, 1998;

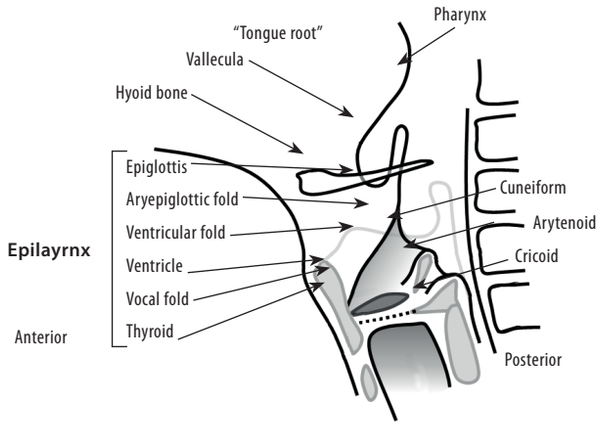


Figure 2.5 Basic epilaryngeal anatomy. (Reproduced courtesy of Scott Moisik 2013, PhD Dissertation, 1, fig. 1.1.)

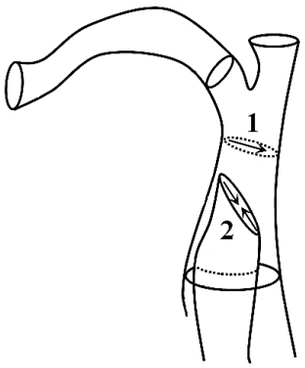


Figure 2.6 The epilarynx as a tube-in-a-tube. (Reproduced courtesy of Scott Moisik 2013, PhD Dissertation, p. 18, fig. 2.1.)

A short digression about 'twang'

'Twang' voice quality was first described by G. Oscar Russell in 1931 when he speculated (accurately) that there is physical, mechanical contact between the ventricular folds and the vocal folds. Subsequently, Jo Estill in her *Voice Craft* primer (1992) described the aryepiglottic sphincter as being 'the Twanger'. Additionally, 'twang' voice quality is produced with a high laryngeal position, and Estill felt that the aryepiglottic constriction plus high larynx position produced another resonator within the vocal tract (within the epilaryngeal space) which enhanced the 2–4 kHz band of the spectrum. Titze & Story (1997) and Titze et al. (2003) broadly concurred with this view, stating that 'the inertance of the vocal tract facilitates vocal fold vibration by lowering the oscillation threshold pressure. This has a significant impact on singing. Not only does the epilarynx tube produce the desirable singer's formant (vocal ring), but it acts like the mouthpiece of a trumpet to shape the flow and influence the mode of vibration' (1997). And: 'Overall, the perception of *twang* increased directly with pharyngeal area narrowing, vocal tract shortening, and decreased open quotient' (2003).

The authors of this chapter have been disturbed by the view that the creation of the singer's formant is all down to the constriction of the aryepiglottic folds. What we have observed consistently when performing flexible laryngoscopy in singers is not only a variable degree of laryngeal raising and narrowing of the epilarynx, but also a distinct narrowing of the pharynx at the same level as the aryepiglottic folds (see Lindqvist-Gauffin, 1972; Sundberg and Nordström, 1976; Laver, 1980)). One would expect to see this gesture for formant tuning of the singer's formant.

The muscles involved in these clinical observations would seem to be middle constrictor of the pharynx +/- palatopharyngeus, suggesting that tuning the singer's formant is oropharyngeal – just as it is for other formants. We agree entirely that 'twang' is due to epilaryngeal constriction, but would argue that in contrast to production of the singer's formant, this is an adjustment of the sound source to produce the distinct, longer, closed phase timbral quality and not merely a quirk of formant tuning of the resonators (see Moisik, 2013: pp. 194–203).

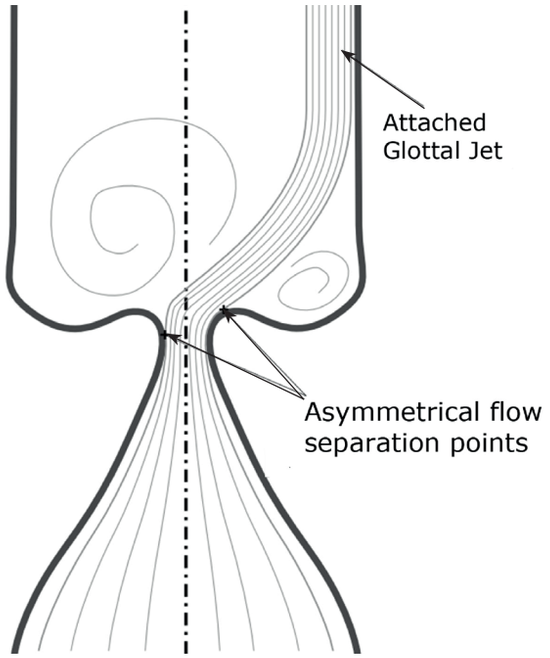


Figure 2.7 Flow separation points.
(After Šidlof et al., 2011.)

Hillel, 2001; Titze, 2008; Honda et al., 2010; Esling and Edmondson, 2011).

The interaction of the vocal folds and the boundary configuration of the epilaryngeal structures are inextricably intertwined (Titze, 1997; Esling et al., 2007) and are beyond the remit of this small chapter, but for those interested in modelling the (non-linear) aerodynamic behaviour of the airflow above the glottis and the concept of fluid–structure–acoustic interaction between the airflow, elastic vocal folds, and sub- and supraglottal acoustic spaces, we recommend authors such as Khosla et al., (2008), Šidlof et al. (2011), Moisik and Esling (2014), and Moisik and Gick (2013)

From a clinical point of view, knowing that there are flow separation points on the superior margins of the left and right vocal fold cover (where the glottal jet airstream separates, producing a laminar flow in the middle and turbulent recirculating vortices at the edges), left/right asymmetry has the potential to affect the vocal outcome following microsurgery. Not only the visco-elastic properties of the mucosa, but also the symmetry in height of the left and right vocal folds have been shown to be critically important in this respect.

Šidlof et al. (2013) have also produced a mathematical model in which the jet deflection angle, flow rate, glottal velocity and pressure, drag and lift force, and jet contours (including phase-averaged data) are compared for convergent and divergent glottic configurations.

Thus it is important to be aware of the airflow consequences in the epilarynx in the following cases.

Recurrent and superior laryngeal nerve paresis

It is widely accepted that the cricothyroid muscle is responsible for the paramedian position of the vocal fold in recurrent laryngeal nerve paralysis. However, support in the literature for this theory is not conclusive, and the cadaveric vocal fold position expected after lesions of the vagus nerve has also been reported in patients with an intact superior laryngeal nerve. The study by Woodson (1993) compared the configuration of the glottis in patients with unilateral paralysis due to known lesions of either the recurrent laryngeal or vagus nerve. Normal subjects were studied as controls. Results indicated that the alteration of glottic configuration in laryngeal paralysis cannot be adequately characterised by standard terms of vocal fold position (see also Hillel, 2001). The paralysed vocal fold is shortened, with anteromedial rotation of the arytenoid and resultant bowing and thinning of the membranous fold with some degree of weakening of the ipselateral supraglottic structures.

What research in recent decades has made abundantly clear is that the relationship between the degree of neurological damage and the appearance of the larynx is very unreliable at any stage following the onset of symptoms. If precise assessment is difficult where there is a recurrent laryngeal nerve problem, it may be even harder to diagnose, assess and prognosticate for superior laryngeal nerve problems. Appearances may vary considerably and may include apparent laryngeal rotation, and torsion within the larynx due to weakening of the cricothyroid muscle on the affected side,

thereby permitting a wider range of movement about the cricothyroid joint. All that the clinician may observe is a (false) impression that the vocal fold on the affected side seems slightly shorter than on the other. There may well be a slight disparity in the height of the superior surfaces of the vocal folds, but this may be very difficult to pick up in the clinic if high-speed video is not available. Studies such as that by Chhetri et al. (2013) have shown that vocal-fold tension asymmetry produces vibratory phase asymmetry: the side with greater tension leading in the opening phase. Clinical observations of vocal fold lag, reduced vocal range and aperiodic voice in superior laryngeal paresis and paralysis were also noted.

Yet another problem which may bother performers is the non-linear ('chaotic') glitch in voicing that may well occur under these conditions. (For those interested in treatment, see Chapter 6, Speech therapy for dysphonia.)

Vibration of the epilarynx occurs in a vast array of musical styles, from blues and jazz to hard rock and death metal (Borch et al., 2004; Eckers et al., 2009). However, despite its evident importance, a fully integrated account of its functions is still lacking.

Mongolian singing styles of the Altai Mountains (kargyraa voice) (Sakakibara et al., 2004) and Tibetan mantra chanting (Lindestad, 2001) have been examined using high-speed laryngoscopic video which clearly shows the ventricular folds engaged in oscillation that is half the frequency of vocal fold vibration. The activity of the ventricular folds produces alternating compression/rarefaction of the air within the epilaryngeal air space, which alternately reinforces/cancels the waves produced by the oscillating vocal folds. This can be demonstrated by displaying the sound signal and the electroglottography (EGG) produced by the patient simultaneously. The sound signal is exactly half the frequency of that of the EGG plot.

A number of researchers point out that laryngealisation may also be accompanied by ventricular fold adduction and dropping the vocal processes forward and downward. Effectively, this increases the functional mass of the vocal folds (e.g. Moore, 1971: p.72). Under these conditions, the vibratory unit becomes massive, although not very tense, and vocal fold vibration is damped yet responsive to low subglottal pressures (Hollien et al., 1966: p. 247). In the clinical situation, it certainly appears that this 'bass register' singing can only be achieved when the singer is using what sounds to Western ears like 'pressed' voicing.

The pre-epiglottic space

The pre-epiglottic space or, more accurately, the 'peri-epiglottic space' and adipose tissue extend around the epiglottis in a horseshoe fashion (Maguire and Dayal, 1974; Reidenbach, 1997). This space consists of a medial part anterior to the epiglottis, below the level of the hyoepiglottic ligament, and two lateral parts extending beyond the lateral margins of the epiglottis. These pre-epiglottic and paralaryngeal components are separated by collagenous fibrous septa that extend from the inner surface of the thyroid cartilage to the lateral border of the epiglottic cartilage and are continuous inferiorly with the quadrangular membrane (Sakai, et al., 1990; Reidenbach, 1997). It seems probable that the fat pad, being incompressible, plays a significant role in the 'bunching' of anterior laryngeal structures for high, loud, tenor notes, as well as being required for certain linguistic gestures and, as may be seen in voice clinics, as a rather strenuous attempt to compensate for vocal fold incompetence. It is also likely that the mechanism involves 'locking' the epiglottis in an upright position by moving the hyoid bone forward (see Chapter 1, Muscles that raise laryngeal position). There is much thryroarytenoid activity, especially in the thyroepiglottic muscle component, which pulls the epiglottis downwards (Moisik and Gick, 2013), compressing the periepiglottic fat pad inferoposteriorly causing the epiglottic petiole to conceal the anterior commissure and the anterior third of the ventricular folds to bunch medially.

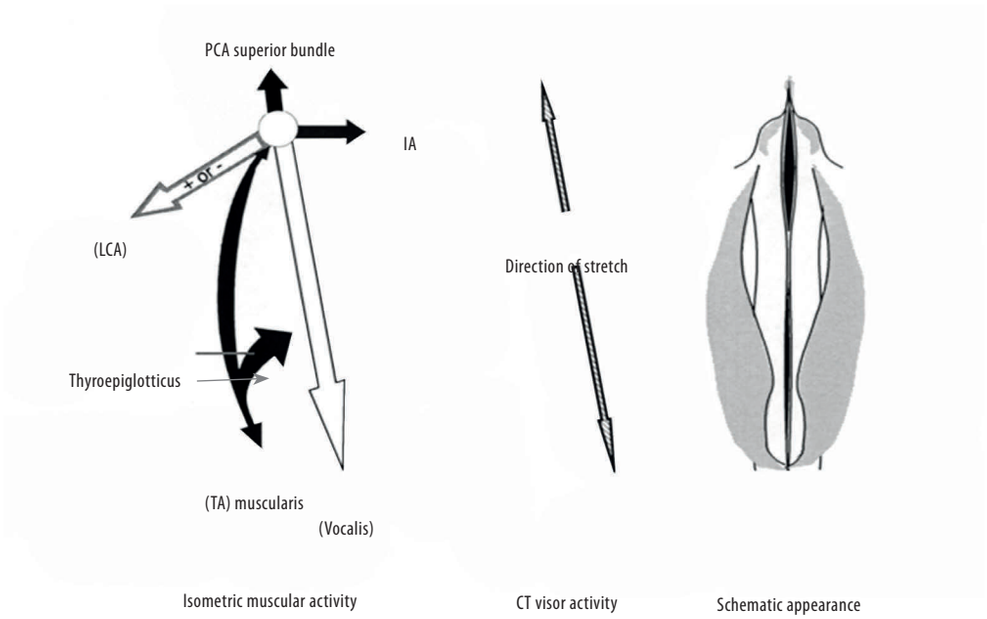


Figure 2.8 Epilaryngeal compression.

References

Asanau A, Timoshenko AP, Prades JM, et al. 2011. Posterior cricoarytenoid bellies: relationship between their function and histology. *J Voice* **25**(2): 67–73.

Borch DZ, Sundberg J, Lindestad, PÅ, et al. 2004. Vocal fold vibration and voice source aperiodicity in "dist" tones: a study of a timbral ornament in rock singing. *Logopedics Phoniatrics Vocology* **29**(4): 147–53.

Butler JE, Hammond TH, Gray SD. 2001. Gender-related differences of hyaluronic acid distribution in the human vocal fold. *Laryngoscope* **111**(5): 907–11.

Brandon C, Rosen C, Georgelis G, et al. 2003. Muscle fiber type composition and effects of vocal fold immobilization on the two compartments of the human posterior cricoarytenoid: a case study of four patients. *J Voice* **17**: 63–75.

Chhetri DK, Neubauer J, Bergeron JL, et al. 2013. Effects of asymmetric superior laryngeal nerve stimulation on glottic posture, acoustics, vibration. *Laryngoscope* **123**(12): 3110–16.

Colton RH, Casper, JK Leonard R. 2011. Morphology of vocal fold mucosa: histology to genomics. In: *Understanding Voice Problems: A Physiological Perspective for Diagnosis and Treatment*, 4th edn. Philadelphia, PA: LWW/Wolters Kluwer Health, pp. 63–75.

Dikkers, FG & Nikkels PG. 1999. Lamina propria of the mucosa of benign lesions of the vocal folds. *Laryngoscope* **109**(10): 1684–89.

Eckers C, Hütz D, Kob M, et al. 2009. Voice production in death metal singers. In: *NAG/DAGA Procedings of the International Conference on Acoustics*, 23–26 March 2009. Rotterdam: The Netherlands, pp. 1747–50.

Esling JH, Zeroual C, Crevier-Buchman L. 2007. A study of muscular synergies at the glottal, ventricular and aryepiglottic levels. In: *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken, Germany: ICPhS, pp. 585–88.

Esling JH, Benner A, Moisik, SR. 2015. Laryngeal articulatory function and speech origins. In: *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, Scotland: ICPhS.

Esling JH, Edmondson, JA. 2011. Acoustical analysis of voice quality for sociophonetic purposes. In: M Di Paolo, M Yaeger-Dror (eds). *Sociophonetics: A Student's Guide*. London & New York, NY: Routledge, pp. 131–48.

Estill J. 1996. Primer of *Compulsory Figures (Level One)* Sant Rosa, CA. Estill Voice Training Systems.

Fink, BR (1975) The human larynx: a functional study. New York, NY: Raven Press.

Gick B, Moisik SR. 2015. Quantal biomechanics in an embodied phonetics. *The Journal of the Acoustical Society of America* **137**(4). doi: http://dx.doi.org/10.1121/1.4920399.

Gray SD, Pignatari SSN, Harding P. 1994. Morphologic ultrastructure of anchoring fibers in normal vocal fold basement membrane zone. *Journal of Voice* **8**: 48–52.

Gray SD, Titze IR, Alipour F, Hammond, TH (2000) Biomechanical and histologic observations of vocal fold fibrous proteins. *Annals of Otology Rhinology and Laryngology* **109**(1): 77–85.

Hahn MS, Kobler J, Zeitels S, et al. 2005. Midmembranous vocal fold lamina propria proteoglycans across selected species. *Annals of Otology Rhinology and Laryngology* **114**(6): 451–62.

Hahn MS, Kobler J, Starcher B, et al. 2006. Quantitative and comparative studies of the vocal fold extracellular matrix I: elastic fibers and hyaluronan. *Annals of Otology Rhinology and Laryngology* **115**(2): 156–64.

Hillel, AD. 2001. The study of laryngeal muscle activity in normal human subjects and in patients with laryngeal dystonia using multiple fine-wire electromyography. *The Laryngoscope* **111**(s97): 17–47.

Hirano M, Kurita S, Nakashima T. 1981. The structure of the vocal folds. In: Stevens KN, Hirano M (eds.) *Vocal Fold Physiology*. Tokyo: University of Tokyo Press.

Hirano M, Kakita Y. 1985. Cover-body theory of vocal fold vibration. In: Daniloff RG (ed.) *Speech Science*. San Diego, CA: College-Hill Press, pp. 1–46.

Hirano S, Minamiguchi S, Yamashita M, et al . 2009. Histologic characterization of human scarred vocal folds. *Journal of Voice* **23**(4): 399–407.

Honda K, Kitamura T, Takemoto H, et al. 2010. Visualization of hypopharyngeal cavities and vocal-tract acoustic modelling. *Computer Methods in Biomechanics and Biomedical Engineering* **13**(4): 443–53.

Khosla S, Murugappan S, Gutmark E. 2008. What can vortices tell us about vocal fold vibration and voice production? *Current Opinion in Otolaryngology & Head and Neck Surgery* **16**(3): 183–7.

Kotby MN, Kirchner JA, Kahane JC, et al. 1991. Histo-anatomical structure of the human laryngeal ventricle. *Acta Oto-laryngologica* **111**(2): 396–402.

Laurent TC, Laurent UB, Fraser JR. 1995. Functions of hyaluronan. *Annals of Rheumatic Diseases* **54**(5): 429–32.

Laver J. 1980. *The Phonetic Description of Voice Quality*. Cambridge: Cambridge University Press.

Lindestad P-Å, Södersten M, Merker B, Granqvist S. 2001. Voice source characteristics in Mongolian "throat singing" studied with high-speed imaging technique, acoustic spectra, and inverse filtering. *Journal of Voice* **15**: 78–85.

Lindqvist-Gauffin J. 1972. A descriptive model of laryngeal articulation in speech. *Quarterly Progress and Status Report, Speech Transmission Laboratory, Royal Institute of Technology* **13**(2–3): 1–9.

Madruga de Melo EC, Lemos M, Aragão Ximenes Filho J, et al. 2003. Distribution of collagen in the lamina propria of the human vocal fold. *The Laryngoscope* **113**(12): 2187–91

Maguire A, Dayal VS. 1974. Supraglottic anatomy: The pre- or the peri-epiglottic space? *Canadian Journal of Otolaryngology* **3:** 432–45.

Moisik, SR. 2008. *A three-dimensional model of the larynx and the laryngeal constrictor mechanism: visually synthesizing pharyngeal and epiglottal articulations observed in laryngoscopy*. Master's thesis, University of Victoria, Canada.

Moisik SR, Gick B. 2005. The quantal larynx revisited. *Acoustical Society of America Proceedings of Meetings on Acoustics*, Vol. 19, 060163 (2013).

Moisik, SR. 2013. *The epilarynx in speech*. Doctoral dissertation, University of Victoria, Canada, Ch. 4 & Ch. 5., p3.

Moisik SR, Esling, JH. 2014. Modelling the biomechanical influence of epilaryngeal stricture on the vocal folds: a low-dimensional model of vocal–ventricular fold coupling. *Journal of Speech, Language, and Hearing Research - Supplement: Select Papers from the 8th International Conference on Vocal Physiology and Biomechanics* **57**: S687–S704.

Painter C. 1986. The laryngeal vestibule and voice quality. *Archives of Oto-Rhino-Laryngology* **243**: 329–37.

Painter C. 1991. The laryngeal vestibule, voice quality and paralinguistic markers. *European Archives of Oto-Rhino-Laryngology* **248**: 452–8.

Reidenbach, MM. 1997. Anatomical considerations of closure of the laryngeal vestibule during swallowing. *European Archives of Oto-Rhino-Laryngology* **254**: 410–12.

Reidenbach, MM. 1998a. The muscular tissue of the vestibular folds of the larynx. *European Archives of Oto-Rhino-Laryngology* **255**: 365–7.

Reidenbach, MM. 1998b. Aryepiglottic fold: normal topography and clinical implications. *Clinical Anatomy* **11**: 223–35.

Sakai F, Gamsu G, Dillon WP, et al. 1990. MR imaging of the larynx at !.5 T *Journal of Computer Assisted Tomography* **14**(1): 60–71.

Sakakibara K. I, Kimura M, Imagawa H, Niimi S, & Tayama N. 2004a. Physiological study of the supraglottal structure. *International Conference on Vocal Fold Physiology and Biomechanics 2004*, Marseille.

Sakakibara K. I, Fuks L, Imagawa H, & Tayama N. 2004b. Growl voice in pop and ethnic styles. In: *Proceedings of the International Symposium on Musical Acoustics*.

Sanders I. 2014. The microanatomy of the vocal fold musculature.

In: Rubin JS, Sataloff RT, Korovin GK (eds) *Diagnosis and Treatment of Voice Disorders* 4th edn. San Diego, CA: Plural Publishing. .

Santo Neto H, Marques MJ. 2008. Estimation of the number and size of motor units in intrinsic laryngeal muscles using morphometric methods. *Clinical Anatomy 21*(4): 301–6.

Schellenberg M. 1998. *Acoustic effect of vocal warm-ups on singers' voices*. Master's thesis, University of Victoria, Canada.

Sciote JJ. 2003. Muscle fiber type composition and effects of vocal fold immobilization on the two compartments of the human posterior cricoarytenoid: a case study of four patients. *Journal of Voice* **17**(1): 63–75.

Šidlof P, Doaré O, Cadot O, Chaigne A. 2011. Measurement of flow separation in a human vocal folds model. *Experiments in Fluids* **51**(1): 123–36.

Šidlof P, Horáček J, Řidký V. 2013. Parallel CFD simulation of flow in a 3D model of vibrating human vocal folds. *Computers & Fluids* **80:** 290–300.

Sundberg J. 1974. Articulatory interpretation of the "singing formant". *Journal of the Acoustical Society of America* **55**(4): 838–44.

Sundberg J, Nordström PE. 1976. Raised and lowered larynx – the effect on vowel formant frequencies. *Quarterly Progress and Status Report. Speech Transmission Laboratory, Royal Institute of Technology* **17**, pp. 35–39.

Thibeault SL. 2004. Hyaluronan biology in vocal fold morphology and biomechanics. In: Garg HG, Hales CA (eds.) *Chemistry and Biology of Hyaluronan*, Oxford: Elsevier, ch. 16, pp. 339–48.

Titze IR, Story BH. 1997. Acoustic interactions of the voice source with the lower vocal tract. *Journal of the Acoustical Society of America* **101**(4): 2234–43.

Titze IR, Bergan CC, Hunter EJ, Story B. 2003. Source and filter adjustments affecting the perception of the vocal qualities twang and yawn. *Logopedics Phoniatrics Vocology*. **28**(4): 147–55.

Titze, IR (2008) Nonlinear source–filter coupling in phonation: theory. *Journal of the Acoustical Society of America* **123**(5): 2733.

Woodson, GE. 1993. Configuration of the glottis in laryngeal paralysis. I: clinical study. *The Laryngoscope* **103**(11): 1227–34.

Yanagisawa E, Estill J, Kmucha ST, Leder S. 1989. The contribution of aryepiglottic constriction to "ringing" voice quality – a videolaryngoscopic study with acoustic analysis. *Journal of Voice* **3**(4): 342–50.

Zemlin WR. 1998. *Speech and Hearing Science: Anatomy and Physiology* 4th edn. Boston, MA: Allyn and Bacon.

Zemlin WR, Davis P, Gaza C. 1984. Fine morphology of the posterior cricoarytenoid muscle. *Folia Phoniatria* **36**(5): 233–40.

speech that is based on human hearing. A review of the human hearing system can be found in Pickles, (1982), Moore (1982), Plack (2005) or Howard and Angus (2016).

The frequency analysis capability of each ear can be modelled as a bank of band-pass filters operating in parallel whose bandwidths vary as a function of their centre frequencies as shown in figure 13.36. The human hearing system effectively has narrow band filters at low frequencies, which become wider as frequency increases. Thus the human hearing system exhibits 'good' frequency resolution at low frequencies and 'good' time resolution at high frequencies.

Figure 13.37 shows the output in spectrographic form from the real-time hearing model (Swan et al., 1994) which consists of a bank of 64 filters whose bandwidths are determined by the curve shown in figure 13.36 for the vowel of spa spoken by a healthy adult male using (A) modal, (B) breathy, (C) falsetto, (D) creaky, (E) harsh and (F) whispered voice qualities as used in figures 13.15 and 13.32 (upper plot) and 'Speech matters to you' spoken by a healthy adult male as used in figure 13.33.

The 'good' frequency and time resolution of the GammaTone spectrogram can be seen at low and high frequencies as clear harmonics and striations respectively. This is particularly clear in the modal voice quality (A) example in the upper plot of figure 13.37 as well as throughout the spectrogram of the sentence 'Speech matters to you' in the lower plot. The frequency axis is related to the filter bandwidths (see figure 13.36), with each filter output receiving an equal distance on the Y-axis. This has the effect of compressing the high frequencies with respect to the low frequencies in a manner, which is fairly similar to a logarithmic scale. The perceptual 'weight' of the different frequency

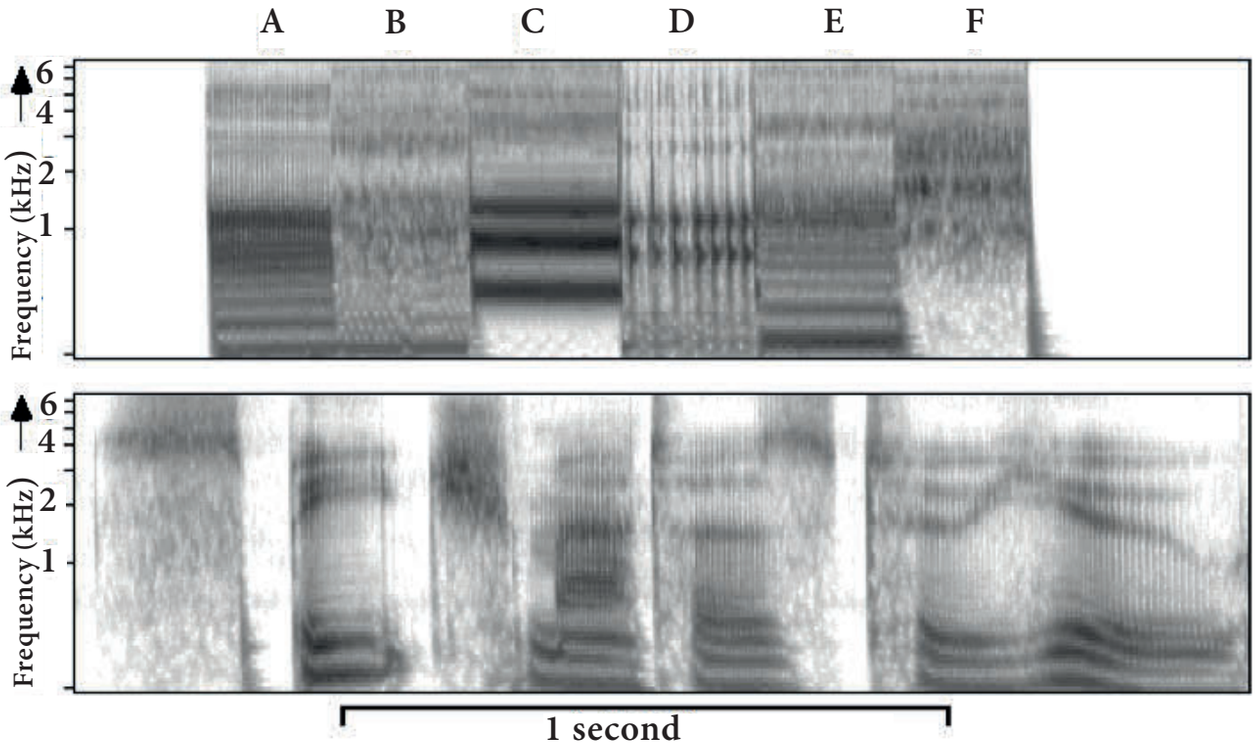


Figure 13.37 Spectrograms from the human hearing model based on ERB spaced GammaTone filters for: (UPPER) the vowel in spa spoken by a healthy adult male using (A) modal, (B) breathy, (C) falsetto, (D) creaky, (E) harsh and (F) whispered voice qualities, the same data as in figures 3.14 and 3.29, and (LOWER) 'Speech matters to you' spoken by a healthy adult male, the same data as in figure 3.30.

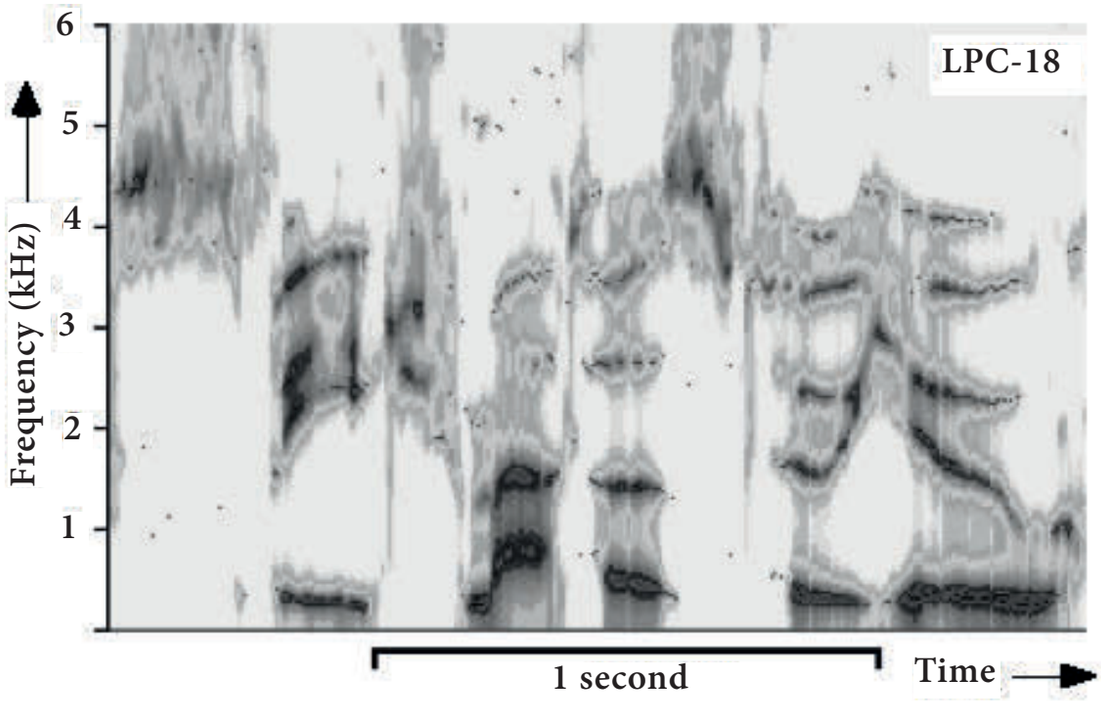


Figure 13.38 Eighteenth order LPC analysis of 'Speech matters to you' spoken by a healthy adult male. This can be compared directly with the spectrograms in figures 13.30 and 13.34.

regions is represented in these spectrograms, which has the effect of spreading the lower formants out and compressing the high frequency information. The lower three harmonics in the falsetto example are clear and the striations in the creak example become thinner at high frequencies, demonstrating the increased time resolution towards high frequencies. Another interesting example of this is shown at the very end of the spectrogram where the final 'click' of the editing process is smeared in time at the bottom but not at the top of the spectrogram. Other features can be compared with the traditional spectrograms of these voice qualities plotted in figure 13.32.

In the sentence, the unvoiced fricatives at either end of the first word 'speech' are very clear, and the different frequency balance of each can be seen. Throughout the voiced regions, the lowest three to five harmonics can be clearly seen, with striations becoming apparent above approximately 1kHz. The last fricative of 'matters' is devoiced which can be seen as the harmonics at its start gradually disappear. The formant transitions are clearly illustrated, especially in the final voiced section of the words 'to you'.

By way of comparison, figure 3.38 shows an eighteenth order LPC spectrogram of the same utterance 'speech matters to you' for which a spectrogram based on a human hearing model is plotted in figure 3.37 and traditional narrow and wide band spectrograms are plotted in figure 13.33. The frequency scales in figures 13.33 are linear. It is left to the reader to make comparisons between these various representations of the acoustic patterns of the utterance, providing an opportunity to see the potential benefit of using an analysis based on the human hearing system in terms of how it combines the features found in narrow and wide band spectrograms.

Coda

To date, the use of hearing modelling spectrograms for the analysis of speech and singing is not widespread, but it is very likely that they will become more popular as greater in-depth understanding of human

perception of speech, singing and other audio signals is desired. As more is becoming known about the human hearing system in centres beyond the peripheral hearing system in terms of neural firing patterns (Plack, 2005; Patterson et al., 1995) and the result of combining the outputs from both ears at higher centres, it is likely that representations of acoustic signals will continue to change to reflect this increased knowledge. A key aspect of human hearing is timing and how this is reflected in the various representations of the input signal that are proposed; the correlogram (Slaney, 1998) is one such representation.

The output from spectrographic analysis is a trade-off between time and frequency, hence the need for wide and narrow band analyses in its traditional form as described above. The hearing system takes account of this by having variation in its bandwidth with centre frequency. Understanding why this is the case and why temporal/frequency acuity is better at high/low frequencies is of considerable interest in this regard, and probably links to species survival. Whatever the outcome of such work, the importance of exploring these areas cannot be overemphasised.

Notes

1. An earlier system for describing the frequency components of a periodic waveform was in terms of 'overtones', where the first overtone refers to the first frequency component that is 'over' or above f0, and this is the second harmonic. The second overtone is the third harmonic, the third overtone the fourth harmonic, and so on. This can sometimes be a source of confusion and we shall use the harmonic system here. Further discussion can be found in Howard and Angus (1996).

2. An octave contains 12 semitones.

3. The ratio of two pressure values P1 and P2 is expressed in decibels (dB) as follows:

   $$dB = 20*\log_{10}\left\{\frac{P1}{P2}\right\}$$

   If P1 = 4*P2, then this is equivalent to: $20*\log_{10}\left\{\frac{P1}{P2}\right\}$
   dB = 20*(-0.6021) dB = -12.042dB

4. One way of thinking about this is that a non-repeating waveform has a period which is infinite since it consists of just one cycle when extended for ever. The Fourier theorem states that the spectrum will consist of harmonics that are spaced by the f_0, where f_0 = (1/period). In this case, f_0 = (1/infinity) = zero. Therefore, the spacing between its harmonics must be zero Hz, or they must be infinitesimally close together, which is the nature of a 'continuous' spectrum.

5. This field is often referred to in the literature as pitch extraction, but since pitch is a subjective judgement (which none of the algorithms to be discussed are making) the term 'f_0 estimation' is preferred and adopted here by the author.

6. As discussed earlier, the second harmonic has a frequency which is twice that of f_0 by definition, which is equivalent to a pitch increase of one octave.

7. This description applies equally well to the formants themselves when they are excited by the acoustic pressure pulses arising from vocal fold closures, since a single formant has same characteristics as the analysis filter of the spectrograph. It too can be described in terms of its centre frequency, or 'formant frequency', and a bandwidth, and its output resulting from each vocal fold closure is a sinewave at the formant frequency which decays in amplitude depending on the formant bandwidth, otherwise described