Acoustics of Speech Production

R. S. McGowan

CReSS
Books

Published by CReSS Books 2018.

ISBN: 978-0-9997574-0-6

Cover silhouette: Paul Dirac and Richard Feynman in discussion.

Cover design by Rebecca McGowan and R. S. McGowan. We gratefully acknowledge the R. P. Feynman Estate for permission to use this image.
Printed by Harvard Book Store.

for Winifred and Rebecca

Table of Contents

Chapter 1: Air: The Acoustic Material in the Vocal Tract, and Some Physics

Chapter 2: Acoustic Air Motion in a Straight Tube

Chapter 3: Acoustic Air Motion in a Tube of Finite Length

Chapter 4: Mass-Spring Systems

Chapter 5: Standing Waves and Normal Modes

Chapter 6: Applications of Normal Modes: Acoustic Perturbation Theory and Green's Functions

Chapter 10: Multiple Sub-Tubes

Chapter 11: Damping and Losses

Chapter 12: Helmholtz Resonators and Side Branches

Chapter 13: Fluid Mechanics and Aeroacoustic Sources

Chapter 14: Scaling, Curvature, and Speech Development

Chapter 15: Layered Structure Model for Vocal Fold Vibration

Preface

Acoustics is the study of sound. Acoustics, as a part of mathematical physics, is a theory of small disturbances in an otherwise quiescent medium, like air. This means that we take the conservation equations describing air motion, and linearize them. There are other assumptions made in the *acoustic approximation*, but thinking of acoustics in air as linearized fluid mechanics is not far off the mark. While general mechanics of air in the human respiratory system is complex, and often non-acoustic, we can usually make the acoustic approximation for pertinent air motion in most of the supraglottal vocal tract during speech production. When we cannot, the approximation usually breaks down in localized regions, such as at the glottis itself, or in other highly constricted regions. These regions often act as sources of sound for the regions of the vocal tract where pertinent air movement is well described by the acoustic approximation.

We make an important distinction immediately regarding measurable quantities and the various approximations that can be used to describe air motion. This distinction has become clouded when air motion described in the acoustic approximation has been related to the motion described by alternative approximations to the conservation equations, such as the Bernoulli equation. These relations have often been made in an ad hoc manner that seem to make it necessary refer to such things as “acoustic pressure” or “Bernoulli pressure”. When considering the physics of air, we understand air has certain physical, measurable properties, such as density, pressure, particle velocity, volume velocity, and so on. These quantities are defined in such a way that they do not depend on the particular theoretical approximation used to describe their changes in space and time. For instance, there is no such thing as “acoustic pressure” or “Bernoulli pressure”, but there are instances and regions where the behavior of pressure is best described using the acoustic approximation, and others when its behavior is best described by Bernoulli’s equation. The motion of air does not accommodate itself according to the theoretical or mathematical approximations that are intended to describe the air’s motion.

There were several motivations that I had to write this book. I believe, that to make progress in understanding sounds that are propagated in the atmosphere as a result of speech, the physics of acoustics must be understood by, at least, some researchers. Further, previous books on speech acoustics, such as those of Fant (1960), Flanagan (1965), and Stevens (1998) were written using electrical analogues. Electrical analogues are fine for many calculations, but they assume that the linearized acoustic approximation is valid. It is important to examine acoustics in the broader context of fluid mechanics to understand the use of acoustics for calculation and its connection to fluid motions in the vocal tract that do not conform to the acoustic approximation.

Other motivation came from questions that I have received from linguists, psychologists, and speech clinicians regarding acoustics. One of these questions regarded filter banks used to synthesize speech. If I remember correctly, the question was how is it that frequencies of a source, $\mathcal{F}$, such as the harmonics of the voice source, drive the filters in a filter bank, but do not follow the relation $\mathcal{F} = c/\lambda$, where c is the speed of sound and λ is the wavelength of the resonance represented by a particular filter in the filter bank? Questions like this made it clear that the way acoustics is traditionally presented to speech researchers is as a computational tool. The physics has often been lost in our understanding when computing outputs from inputs with transfer functions and filter banks. I hold to the idea that "The primary purpose of theory is understanding, not calculation."

This is a book about settled science, as one of my colleagues has said. Almost all the the material in this book comes from nineteenth and early twentieth century physics. The newest topic that is in the book, other than the two research topics in Chapters 14 and 15, is in Chapter 13, where we discuss how air motion that is not described by the acoustic approximation can serve as a source of acoustic energy for acoustic wave motion. This is the study of aeroacoustics, and its modern beginnings came just after the first half of the twentieth century with the publication of Lighthill's paper on jet noise in 1952.

The book is divided into four parts. Chapters 1 through 7 provide a thorough explanation of acoustic motion in a tube of finite length and constant cross-sectional area with a moving piston at one end. Some discussion of another type of source is also included. In the first part of the book, almost all of the results are presented in what we consider to be the time domain. This part lays out the fundamental mathematical physics of the situation, so that further developments in variable-area tubes in Chapters 8 through 12 become more a matter of computation. However, we never switch completely to just computation of physical quantities, but continue to provide physical understanding with the derivation of equations. Much of the second part of the book involves frequency domain considerations. The third part of the book, which is in Chapter 13, is an introduction to fluid mechanics that does not follow the acoustic approximation. This leads to a short study of aeoroacoustics and the way that the fluid motions that do not satisfy the acoustic approximation can provide energy for acoustic wave motion in the vocal tract. Finally, Chapters 14 and 15 outline two research projects that have not hitherto been published. The first involves a hypothesized relation between tongue surface curvature constraints applied to young children and their resulting acoustic output during speech. The second work is a proposed model for vocal fold vibration that uses some of the ideas introduced in Chapter 13. It is intended as a replacement for lumped element models, such as the two-mass model. Biomechanical parameters can be more easily related to the proposed model than to lumped element models. This final topic uses advanced mathematical tools.

This book is targeted at people interested in the physical aspects of acoustics that arise during the act of speaking. This includes people involved directly with the science of speech production, but without advanced mathematical knowledge, and those in the physical sciences with some advanced mathematical knowledge. [Here, we consider anything more than a smattering of calculus to be advanced knowledge of mathematics.] It is written to be accessible by those who are not trained in mathematics, while not holding back on presenting important physical ideas that are expressed mathematically. There are many simulations presented in plots and with numerical values to aid understanding. The book is something of a narrative about the physical acoustics encountered during speech production. It is intended to be read in sequence. This is not a text book, or even a reference book, in the traditional sense. For one thing, there are no problems for the student. However, if the reader wishes a more thorough understanding, he or she could derive some of the mathematical expressions presented in the book.

The reader should just forge ahead if he or she finds that some mathematical expressions are too difficult to understand. There is always a discussion or simulation that should still be illuminating. We would also recommend that readers use software to simulate some of the results, and turn to references for more in-depth discussion of topics of interest.

Equations are numbered separately in each chapter. Further, Chapters 1, 2, and 8 have short mathematical appendices with equations numbered as, for example (A1.2), where A1 refers to the Appendix to Chapter 1, and 2 refers to the equation number in that appendix. In some derivations, equation numbers are written in square brackets next to the steps of the derivation, so the reader can justify the step. We do this only for the first few uses of equations in the appendices.

There are numerous people to thank. There are friends, those who I have worked with in speech production research, and those across the United States who have either hosted me or encouraged me on my cross-country treks to teach in the American southwest. The number of people is too great to mention them by name. I want to thank my colleagues at Imperial Valley College, and my friends on both sides of the American-Mexican border in the college's region. All of these people helped me realize that a pedagogical vocation may not be out of reach for me.

Finally, I thank the people who read earlier versions of this book: Michael Howe, Khalil Iskarous, Lynn McGowan, Philip Rubin, and Reiner Wilhelms-Tricarico. They helped to improve this book immensely and are encouraging friends as well.

R. S. McGowan
Lexington, Massachusetts

References

Fant, G. (1960). *Acoustic Theory of Speech Production.* Mouton, The Hague.

Flanagan, J.L. (1965). *Speech Analysis, Synthesis, and Perception.* Springer-Verlag, New York.

Lighthill, M.J. (1952). On sound generated aerodynamically, Part 1: General Theory. *Proc. Roy. Soc. A*. **211**, 564-587.

Stevens, K.N. (1998). *Acoustic Phonetics.* MIT Press, Cambridge, MA.