Human and Machine Hearing
Extracting Meaning from Sound

*Human and Machine Hearing* describes how human hearing works and how to build machines to analyze sounds the same way people do. The details of this approach are taught at a college engineering level, in a way designed to bring a diverse range of readers to a common technical understanding. The description of hearing as signal-processing algorithms is supported by corresponding open-source code, for which the book serves as motivating documentation. Lyon shows how to understand human hearing in terms of engineering concepts, and to make those concepts into machines that can analyze sounds the way humans do, for a wide range of modern applications. With more than 35 years invested in this approach, Lyon explains how simple concepts, such as that the ear is a Fourier analyzer, have been put behind us, so that we now build machines that approach human abilities in speech, music, and other sound-understanding domains.

**Richard F. Lyon** is an engineer and scientist known for his work on cochlear models and auditory correlograms for the analysis and visualization of sound, and for analog and digital VLSI implementations of these models, starting at Xerox Palo Alto Research Center, Schlumberger Palo Alto Research, and Apple Advanced Technology Group. After a decade off to develop digital cameras and image sensors at Foveon, he moved back into hearing research, and now leads Google’s research and applications development in machine hearing. At Google, he concurrently led the team that developed camera systems for the Street View project. Lyon received a BS in engineering and applied science from Caltech and an MS in electrical engineering from Stanford University. He is a Fellow of the IEEE and a Fellow of the ACM, and is among the world’s top 500 editors of Wikipedia. He has published widely in engineering journals of the IEEE, in the *Journal of the Acoustical Society of America*, and in book chapters in diverse fields, including hearing, VLSI design, signal processing, speech recognition, computer architecture, photographic technology, handwriting recognition, computer graphics, and slide rules. He holds 58 issued United States patents for his inventions, including the optical mouse. Although he does not have a doctorate degree, he has co-advised doctoral students and served on doctorate committees at six top universities (including Caltech, Stanford, and UC Berkeley) on three continents.
Human and Machine Hearing

Extracting Meaning from Sound

RICHARD F. LYON

Google, Inc.
Un beau visage est le plus beau de tous les spectacles ; & l’harmonie la plus douce est le son de voix de celle que l’on aime.

A fine Face is the finest of all Sights, and the sweetest Musick, the Sound of her Voice whom we love.

—Jean La Bruyère (1713) from 1691 French original.

This book is dedicated to my family: my beautiful, smart, cheerful, successful, inspiring, and sweet-voiced wife Peggy Asprey, and our awesome children Susan and Erik—they are the loves of my life, and my fortune. Though this book has sometimes absorbed too much of my attention, they have all supported me in writing it, in so many ways. They are my finest of all sights, and sweetest music; they sustain me.
Contents

Foreword xv
Preface xix

Part I Sound Analysis and Representation Overview 1

1 Introduction 5
1.1 On Vision and Hearing à la David Marr 8
1.2 Top-Down versus Bottom-Up Analysis 11
1.3 The Neuromimetic Approach 13
1.4 Auditory Images 14
1.5 The Ear as a Frequency Analyzer? 16
1.6 The Third Sound 18
1.7 Sound Understanding and Extraction of Meaning 18
1.8 Leveraging Techniques from Machine Vision and Machine Learning 19
1.9 Machine Hearing Systems “by the Book” 20

2 Theories of Hearing 23
2.1 A “New” Theory of Hearing 23
2.2 Newer Theories of Hearing 26
2.3 Active and Nonlinear Theories of Hearing 27
2.4 Three Auditory Theories 28
2.5 The Auditory Image Theory of Hearing 29

3 On Logarithmic and Power-Law Hearing 33
3.1 Logarithms and Power Laws 33
3.2 Log Frequency 35
3.3 Log Power 37
3.4 Bode Plots 38
3.5 Perceptual Mappings 41
3.6 Constant-Q Analysis 44
3.7 Use Logarithms with Caution 44
Contents

4 Human Hearing Overview
4.1 Human versus Machine
4.2 Auditory Physiology
4.3 Key Problems in Hearing
4.4 Loudness
4.5 Critical Bands, Masking, and Suppression
4.6 Pitch Perception
4.7 Timbre
4.8 Consonance and Dissonance
4.9 Speech Perception
4.10 Binaural Hearing
4.11 Auditory Streaming
4.12 Nonlinearity
4.13 A Way Forward

5 Acoustic Approaches and Auditory Influence
5.1 Sound, Speech, and Music Modeling
5.2 Short-Time Spectral Analysis
5.3 Smoothing and Transformation of Spectra
5.4 The Source–Filter Model and Homomorphic Signal Processing
5.5 Backing Away from Logarithms
5.6 Auditory Frequency Scales
5.7 Mel-Frequency Cepstrum
5.8 Linear Predictive Coding
5.9 PLP and RASTA
5.10 Auditory Techniques in Automatic Speech Recognition
5.11 Improvements Needed

Part II Systems Theory for Hearing

6 Introduction to Linear Systems
6.1 Smoothing: A Good Place to Start
6.2 Linear Time-Invariant Systems
6.3 Filters and Frequencies
6.4 Differential Equations and Homogeneous Solutions
6.5 Impulse Responses
6.6 Causality and Stability
6.7 Convolution
6.8 Eigenfunctions and Transfer Functions
6.9 Frequency Response
6.10 Transforms and Operational Methods
6.11 Rational Functions, and Their Poles and Zeros
## Contents

6.12 Graphical Computation of Transfer Function Gain and Phase 119  
6.13 Convolution Theorem 120  
6.14 Interconnection of Filters in Cascade, Parallel, and Feedback 121  
6.15 Summary and Next Steps 125  

7  Discrete-Time and Digital Systems 126  
7.1 Simulating Systems in Computers 126  
7.2 Discrete-Time Linear Shift-Invariant Systems 126  
7.3 Impulse Response and Convolution 127  
7.4 Frequency in Discrete-Time Systems 127  
7.5 Z Transform and Its Inverse 128  
7.6 Unit Advance and Unit Delay Operators 129  
7.7 Filters and Transfer Functions 131  
7.8 Sampling and Aliasing 134  
7.9 Mappings from Continuous-Time Systems 136  
7.10 Filter Design 138  
7.11 Digital Filters 138  
7.12 Multiple Inputs and Outputs 141  
7.13 Fourier Analysis and Spectrograms 142  
7.14 Perspective and Further Reading 144  

8  Resonators 145  
8.1 Bandpass Filters 145  
8.2 Four Resonant Systems 149  
8.3 Resonator Frequency Responses 152  
8.4 Resonator Impulse Responses 154  
8.5 The Complex Resonator and the Universal Resonance Curve 157  
8.6 Complex Zeros from a Parallel System 159  
8.7 Keeping It Real 163  
8.8 Digital Resonators 165  

9  Gammatone and Related Filters 169  
9.1 Compound Resonators as Auditory Models 169  
9.2 Multiple Poles 170  
9.3 The Complex Gammatone Filter 172  
9.4 The Real Gammatone Filter 175  
9.5 All-Pole Gammatone Filters 178  
9.6 Gammachirp Filters 181  
9.7 Variable Pole $Q$ 184  
9.8 Noncoincident Poles 184  
9.9 Digital Implementations 185
## Contents

### 10 Nonlinear Systems

10.1 Volterra Series and Other Descriptions ............................................. 189
10.2 Essential Nonlinearity ................................................................. 191
10.3 Hopf Bifurcation ......................................................................... 192
10.4 Distributed Bandpass Nonlinearity ................................................. 194
10.5 Response Curves of Nonlinear Systems ........................................... 195
10.6 Two-Tone Responses ................................................................. 198
10.7 Nonlinearity and Aliasing .............................................................. 199
10.8 Cautions ..................................................................................... 201

### 11 Automatic Gain Control

11.1 Input–Output Level Compression .................................................... 202
11.2 Nonlinear Feedback Control ......................................................... 204
11.3 AGC Compression at Equilibrium ................................................ 205
11.4 Multiple Cascaded Variable-Gain Stages ....................................... 207
11.5 Gain Control via Damping Control in Cascaded Resonators .......... 209
11.6 AGC Dynamics ......................................................................... 210
11.7 AGC Loop Stability ................................................................... 215
11.8 Multiple-Loop AGC .................................................................... 218

### 12 Waves in Distributed Systems

12.1 Waves in Uniform Linear Media .................................................... 219
12.2 Transfer Functions from Wavenumbers ........................................ 226
12.3 Nonuniform Media ..................................................................... 230
12.4 Nonuniform Media as Filter Cascades .......................................... 234
12.5 Impulse Responses ..................................................................... 235
12.6 Group Velocity and Group Delay .................................................. 235

### Part III The Auditory Periphery

13 Auditory Filter Models ..................................................................... 237

13.1 What Is an Auditory Filter? ............................................................. 241
13.2 From Resonance to Gaussian Filters .............................................. 243
13.3 Ten Good Properties for Auditory Filter Models ......................... 244
13.4 Representative Auditory Filter Models ......................................... 246
13.5 Complications: Time-Varying and Nonlinear Auditory Filters .... 252
13.6 Fitting Parameters of Filter Models ................................................. 255
13.7 Suppression .................................................................................. 257
13.8 Impulse Responses from Physiological Data ................................. 260
13.9 Summary and Application to Cochlear Models ............................. 264
## Contents

14 **Modeling the Cochlea**  
14.1 On the Structure of the Cochlea  
14.2 The Traveling Wave  
14.3 1D, 2D, and 3D Hydrodynamics  
14.4 Long Waves, Short Waves, and 2D Models  
14.5 Active Micromechanics  
14.6 Scaling Symmetry and the Cochlear Map  
14.7 Filter-Cascade Cochlear Models  
14.8 Outer Hair Cells as Active Gain Elements  
14.9 Dispersion Relations from Mechanical Models and Experiments  
14.10 Inner Hair Cells as Detectors  
14.11 Adaptation to Sound via Efferent Control  
14.12 Summary and Further Reading  

15 **The CARFAC Digital Cochlear Model**  
15.1 Putting the Pieces Together  
15.2 The CARFAC Framework  
15.3 Physiological Elements  
15.4 Analog and Bidirectional Models  
15.5 Open-Source Software  
15.6 Detailing the CARFAC  

16 **The Cascade of Asymmetric Resonators**  
16.1 The Linear Cochlear Model  
16.2 Coupled-Form Filter Realization  

17 **The Outer Hair Cell**  
17.1 Multiple Effects in One Mechanism  
17.2 The Nonlinear Function  
17.3 AGC Effect of DOHC  
17.4 Typical Distortion Response Patterns  
17.5 Completing the Loop  

18 **The Inner Hair Cell**  
18.1 Rectification with a Sigmoid  
18.2 Adaptive Hair-Cell Models  
18.3 A Digital IHC Model  

19 **The AGC Loop Filter**  
19.1 The CARFAC’s AGC Loop  
19.2 AGC Filter Structure
## Contents

19.3 Smoothing Filter Pole–Zero Analysis 332  
19.4 AGC Filter Temporal Response 335  
19.5 AGC Filter Spatial Response 337  
19.6 Time–Space Smoothing with Decimation 338  
19.7 Adapted Behavior 341  
19.8 Binaural or Multi-Ear Operation 341  
19.9 Coupled and Multistage AGC in CARFAC and Other Systems 342

### Part IV The Auditory Nervous System 345

20 Auditory Nerve and Cochlear Nucleus 347  
20.1 From Hair Cells to Nerve Firings 347  
20.2 Tonotopic Organization 350  
20.3 Fine Time Structure in Cochleagrams 351  
20.4 Cell Types in the Cochlear Nucleus 352  
20.5 Inhibition and Other Computation 353  
20.6 Spike Timing Codes 354

21 The Auditory Image 355  
21.1 Movies of Sound 355  
21.2 History 356  
21.3 Stabilizing the Image 357  
21.4 Triggered Temporal Integration 360  
21.5 Conventional Short-Time Autocorrelation 365  
21.6 Asymmetry 367  
21.7 Computing the SAI 367  
21.8 Pitch and Spectrum 369  
21.9 Auditory Images of Music 369  
21.10 Auditory Images of Speech 369  
21.11 Summary SAI Tracks: Pitchograms 371  
21.12 Cochleagram from SAI 373  
21.13 The Log-Lag SAI 376

22 Binaural Spatial Hearing 379  
22.1 Rayleigh’s Duplex Theory: Interaural Level and Phase 379  
22.2 Interaural Time and Level Differences 385  
22.3 The Head-Related Transfer Function 386  
22.4 Neural Extraction of Interaural Differences 389  
22.5 The Role of the Cochlear Nucleus and the Trapezoid Body 392  
22.6 Binaural Acoustic Reflex and Gain Control 394  
22.7 The Precedence Effect 395
22.8 Completing the Model 397
22.9 Interaural Coherence 397
22.10 Binaural Applications 398

23 The Auditory Brain 400
23.1 Scene Analysis: ASA and CASA 400
23.2 Attention and Stream Segregation 402
23.3 Stages in the Brain 407
23.4 Higher Auditory Pathways 410
23.5 Prospects 415

Part V Learning and Applications 417

24 Neural Networks for Machine Learning 419
24.1 Learning from Data 419
24.2 The Perceptron 420
24.3 The Training Phase 421
24.4 Nonlinearities at the Output 423
24.5 Nonlinearities at the Input 426
24.6 Multiple Layers 428
24.7 Neural Units and Neural Networks 428
24.8 Training by Error Back-Propagation 429
24.9 Cost Functions and Regularization 432
24.10 Multiclass Classifiers 434
24.11 Neural Network Successes and Failures 436
24.12 Statistical Learning Theory 437
24.13 Summary and Perspective 439

25 Feature Spaces 441
25.1 Feature Engineering 442
25.2 Automatic Feature Optimization by Deep Networks 443
25.3 Bandpass Power and Quadratic Features 444
25.4 Quadratic Features of Cochlear Filterbank Outputs 445
25.5 Nonlinearities and Gain Control in Feature Extraction 446
25.6 Neurally Inspired Feature Extraction 448
25.7 Sparsification and Winner-Take-All Features 448
25.8 Which Approach Will Win? 449

26 Sound Search 450
26.1 Modeling Sounds 451
26.2 Ranking Sounds Given Text Queries 457
26.3 Experiments 461
# Contents

26.4 Results 463  
26.5 Conclusions and Followup 465  

**27 Musical Melody Matching** 467  
27.1 Algorithm 469  
27.2 Experiments 475  
27.3 Discussion 478  
27.4 Summary and Conclusions 480  

**28 Other Applications** 481  
28.1 Auditory Physiology and Psychoacoustics 481  
28.2 Audio Coding and Compression 482  
28.3 Hearing Aids and Cochlear Implants 483  
28.4 Visible Sound 489  
28.5 Diagnosis 491  
28.6 Speech and Speaker Recognition 493  
28.7 Music Information Retrieval 493  
28.8 Security, Surveillance, and Alarms 494  
28.9 Diarization, Summarization, and Indexing 495  
28.10 Have Fun 495  

*Bibliography* 497  
*Author Index* 545  
*Subject Index* 557
Foreword

*Human and Machine Hearing* is a book for people who want to understand how the auditory system and the brain process sound, how to encapsulate aspects of our hearing knowledge in computer algorithms, and how to combine the algorithms into a machine that simulates the role of hearing in some aspect of everyday life—such as listening to the melody of a song or talking to a friend in a noisy restaurant. This is what Dick Lyon means by “Machine Hearing.” The applications typically involve the segregation and identification of sound sources in everyday environments where there are competing sources and background noises—applications where there is reason to believe that the auditory form of sound analysis and feature extraction will be more effective and more robust than that provided by the traditional combination of the Fourier magnitude spectrum and MFCCs (mel-frequency cepstral coefficients). To construct a hearing machine and apply it to a real-world problem is an enormous undertaking; the latter half of the book documents the construction of a sophisticated auditory model and how it was integrated with machine learning algorithms to produce two hearing machines—an auditory search engine and an auditory melody matcher. The first half of the book describes the basic science that underpins machine hearing; it sets out the problems of constructing a stable, computationally efficient system, and it explains how to deal with each problem in turn. So the book is a comprehensive reference work for machine hearing with an ordered set of worked problems that culminate in two impressive demonstrations of machine hearing and its potential. This combination makes the book ideal both as a reference manual for experts working in the field of machine hearing and for graduate-level courses on machine hearing.

Lyon’s idea of a machine hearing system has four “layers.” The first two simulate auditory frequency analysis in the cochlea and auditory image construction in the brain stem. Together they form an auditory model that is intended to simulate all of the mechanical and neural processing required to produce your initial auditory image of a sound, that is, the internal auditory representation of sound that is thought to provide the basis for perception, streaming, auditory scene analysis, and all subsequent processing. The third layer applies application-dependent feature extraction to the auditory image and reduces the mass of features to a sparse form for the fourth layer, which extracts meaning with machine learning techniques. Together the third and fourth layers make the auditory model into a specific form of hearing machine, designed to perform a particular listening task.
The compact, authoritative introductions to auditory physiology, auditory perception, the acoustics of sound, and the mathematics of auditory filtering and auditory signal processing include the essential facts and functions, along with brief sketches of the people and experiments associated with milestones in the history of hearing research. This part of the book is a delightfully readable reference manual for machine hearing. Lyon’s involvement with the field over the years gives the chapters real authority. The central chapters describe Lyon’s preferred auditory model, which has two distinct stages: the first simulates the operation of the cochlea; the second simulates the conversion of the cochlear output into your initial auditory image of a sound in the neural centers between the cochlea and auditory cortex. The cochlear processing section is a transmission-line filter bank that simulates basilar membrane motion with a “cascade of asymmetric resonators” (CAR). The gains of the resonators are continuously adjusted by a distributed AGC (automatic gain control) network whose action is applied separately to each CAR stage through the outer-hair-cell component of that stage. The resulting system exhibits the “fast-acting compression” (FAC) characteristic of auditory processing, as well as longer-term adaptation characteristic of mid-brain efferents. This stimulus-specific adaptation is intended to make machine hearing robust to interference in the way that human hearing is. The CARFAC model provides an accurate, stable simulation of cochlear processing across the full dynamic range of hearing—an enormous engineering achievement. These chapters are supported by some wonderful figures illustrating how the AGC network adjusts filter gain and shape across the complete set of CARFAC frequency responses as the level and content of a sound varies.

The neural processing section of the auditory model is relatively simple; it applies a form of “strobed” temporal integration (STI) separately to each channel of information flowing from the cochlear section of the model. STI automatically stabilizes sections of the neural activity that repeat, much as the trigger mechanism in an oscilloscope makes a stable picture from an ongoing time-domain waveform. The result for the complete set of cochlear channels is referred to as a stabilized auditory image (SAI)—a series of two-dimensional frames of real-valued data that form an “SAI movie” when presented in real time. Each frame is indexed by cochlear channel number on the vertical axis and “lag relative to strobe time” on the horizontal axis (see many examples in the figures in Chapter 21). The vocal sounds of animals (including speech) contain periodic segments that distinguish animate sources from environmental noises in the natural world, and the SAI presents a detailed, stable view of each repeating neural pattern for as long as it persists in the sound. In this way, STI and the SAI facilitate feature extraction and source segregation in everyday listening where the signals (speech, music, animal calls) are commonly mixed with interfering noises.

Together, the CARFAC cochlear model and the SAI encapsulate much of what we now know (and hypothesize) about auditory processing, and they provide a representation of sound that emphasizes the features and distinctions of everyday listening.

What is needed, then, is a digital version of the auditory brain that can put the auditory model to work in the service of machine hearing. This is the topic of the remaining chapters of the book. Lyon concludes that auditory scene analysis (ASA) and the algorithms used to perform computational ASA (CASA) are not, as yet, able to simulate the
auditory brain, primarily because we do not understand the cortical processing behind the auditory brain. Similarly, he concludes that the neural networks commonly used in machine learning to train a nonlinear mapping from a large set of input patterns to outputs defined by a set of training data are unlikely to provide the basis for a successful model of the auditory brain, in this case because they are unlikely to be able to take SAI frames as input patterns due to the size of the frames and the frame rate. Some form of auditory feature extraction will have to be applied to the SAI frames to concentrate the auditory information in them and reduce the magnitude of the categorization problem for the machine learning systems used to implement machine hearing tasks. Lyon also believes that fine timing information is involved in the construction of human auditory features at a fundamental level, and that hearing machines will have to include fine temporal structure in some form or other.

This thinking leads to the intriguing idea of feature engineers and machine hearing engineers—people who use auditory knowledge, on the one hand, and knowledge about machine learning, on the other hand—designing mappings that convert auditory representations of sound with high dimensionality into forms that are suited to machine learning systems. Where possible, the engineers would identify auditory features that humans use and design algorithms to extract them from streams of SAIIs. Lyon argues, however, that the development of machine hearing does not require the successful identification of the auditory features used by humans to solve listening problems. Rather, the engineer just needs to build a good interface between what we know about hearing and what we know about a machine learning system that might address the listening task. Indeed, it is argued that the mapping should not remove more information than absolutely necessary to get the machine hearing task running. The machine learning algorithms might find nonintuitive features that actually perform better than the ones designed by a feature engineer to simulate human feature extraction. In summary, Lyon concludes that we will need to be careful about the problems we take on in the near future. We do not know enough about the auditory brain to simulate it. To make machine hearing a reality, we need intelligent mapping procedures to connect the very sophisticated CARFAC–SAI model of hearing to good learning machines—procedures that may, or may not, extract features the way humans do. This discussion of the options currently available to machine hearing engineers is fascinating, and his conclusions about how to proceed are very convincing.

Lyon is a great teacher and he has a deep understanding of the science and art of machine hearing. The reader will be greatly rewarded for engaging with any and all sections of the book.

– Roy D. Patterson, 2016, Cambridge, UK
Preface

If we understood more about how humans hear, we could make machines hear better, in the sense of being able to analyze sound and extract useful and meaningful information from it. Or so I claim. I have been working for decades, but more intensely in recent years, to add some substance to this claim, and to help engineers and scientists understand how the pieces fit together, so they can help move the art forward. There is still plenty to be done, and this book is my attempt to help focus the effort in this field into productive directions; to help new practitioners see enough of the evolution of ideas that they can skip to where new developments and experiments are needed, or to techniques that can already solve their sound understanding problems.

The book-writing process has been tremendous fun, with support from family, friends, and colleagues. They do, however, have a tendency to ask two annoying questions: “Is the book done yet?” and “Who is your audience?” The first eventually answers itself, but I need to say a few words about the second. I find that interest in sound and hearing comes from people of many different disciplines, with complementary backgrounds and sometimes incompatible terminology and concepts. I want all of these people as my audience, as I want to teach a synthesis of their various viewpoints into a more comprehensive framework that includes everything needed to work on machine hearing problems. That is, electrical engineers, computer scientists, physicists, physiologists, audiologists, musicians, psychologists, and others are all part of my audience. Students, teachers, researchers, product managers, developers, and hackers are, too.

The book’s treatment of various aspects of hearing and engineering may be too deep for some, too shallow for others; many will find that something they know is missing, but hopefully all will also find useful things they didn’t know. In particular, the system theory in Part II is taught with the aim of bringing this diverse audience to a common understanding of the math, physics, engineering, and signal-processing concepts needed to design, analyze, and understand the hearing models and applications taught in the later parts. Many aspects of the later parts of the book can be appreciated without mastering the system theory of Part II, but I recommend at least reading it through to get familiar with the terminology and to know where to refer later if more depth of understanding on particular points is desired.

Hearing has perhaps the most deep and elegant combination of linear and nonlinear aspects of any biological system. Readers will learn why the concepts of linear systems are so important in hearing, and also why these concepts are not nearly enough to explain hearing. Understanding nonlinear systems is always challenging, and we address
that challenge by compartmentalizing the important nonlinearities of hearing into well-defined simple mechanisms that are individually not that hard to understand. We develop auditory models in terms of continuous-time systems, and implementations in terms of discrete-time systems with efficient implementations on computing machines; here again, having the nonlinearities compartmentalized is important.

The two aspects that best characterize the book’s auditory models are ideas that I have pursued for many years, with many collaborators: the filter-cascade structure with embedded nonlinearities to model the cochlea; and the stabilized auditory image, or auditory correlogram, to capture and display the temporal fine structure in the signals that the cochlea sends to the brain. These two aspects are on opposite ends of the auditory nerve, and support my strategy to “respect the auditory nerve.” We know so much from auditory physiologists about the properties of sound representation on the auditory nerve, that to build models and systems that either do not produce or do not use the cochlear nerve’s rich information about sound seems indefensible. The book shows some of the ways we have used such information productively.

The auditory models of Parts III and IV of the book are supported by open-source code, which should enable readers to get a good start on building machine hearing systems. Part V of the book introduces a very open-ended future of interesting applications, and I fully expect readers will become contributors to growth in this field of applications.

On History and Connection Boxes

While there are historical comments, and comments on connections to related concepts in other fields, throughout many chapters, I have segregated some of them into boxes, both to highlight them and to keep them out of the way. In many cases, my aim is to honor the sources of the ideas we use, while trying to make the literature more accessible by saying a few words about how it connects. I trust that my mention of old technologies such as vacuum tube (valve) amplifiers and Helmholtz resonators and flame manometers will be received as intended: as clues to a very interesting heritage from generations of giants whose shoulders we stand upon, in both human and machine hearing.

My own EE training was in the era of transistors and early integrated circuits, when courses like “Circuits, Signals, and Systems” were all about analog continuous-time technology. In modern times, signals and systems are taught from the beginning with discrete-time concepts, for good reasons having to do both with pedagogy and the modern medium of implementation in digital computers. Although modern engineers may view sound naturally as the kind of discrete-time sampled data that they work with in computers, I have chosen to stick with continuous time as the primary conceptual domain in this work, since sound and the ear really exist in that domain. I hope that readers will not view the continuous-time domain as something out of history—it is the real world.
I mostly use the editorial “we” in the book, referring not only to myself as author but also to others who contribute to the ideas, including our readers. In a few places I switch to using “I,” for more personal comments.

Though I paid my friends and colleagues a dollar for each bug or suggestion that I acted on, I owe them much more than that in thanks. Through their effort, the book has been much improved. I hope others will send suggestions for improving the next edition, and will earn a few dollars, too. I’m sure we have left some more errors for them to find.

**Online Materials**

Find errata, and links to code and other resources, at machinehearing.org.

**Thanks**

There are many people who have cared enough about this work to spend time helping and encouraging me. First among them is Roy Patterson, without whose encouragement I could never have even started, and who has continued to inspire me through the slow process.

Among my readers who have given me actionable feedback, Ryan “Rif” Rifkin stands out; he found me more bugs than everyone else combined. Others who contributed, whether by carefully reading chapters or giving feedback on overall impressions, include: Jont Allen, Peggy Asprey, Fred Bertsch, Alex Brandmeyer, Peter Cariani, Wan-Teh Chang, Sourish Chaudhuri, Brian Clark, Lynn Conway, Achal Dave, Bertrand Delgutte, Dick Duda, Diek Duifhuis, Dan Ellis, Doug Eck, Dylan Freedman, Jarret Gaddy, Daniel Galvez, Dan Geisler, Pascal Getreuer, Chet Gnegy, Alex Gutkin, Yuan Hao, Thad Hughes, Aren Jansen, James Kates, Nelson Kiang, Ross Koningstein, Harry Levitt, Carver Mead, Ray Meddis, Harold Mills, Channing Moore, Stephen Neely, Eric Nichols, Fritz Obermeyer, Ratheet Pandya, Brian Patton, Justin Paul, Manoj Plakal, Jay Ponte, Rocky Rhodes, David Ross, Mario Ruggero, R. J. Ryan, Bryan Seybold, Shihab Shamma, Phaedon Sinis, Jan Skoglund, Malcolm Slaney, Daisy Stanton, Rich Stern, John L. Stewart, Ian Sturdy, Jeremy Thorpe, George Tzanetakis, Marcel van der Heijden, Tom Walters, Yuxuan Wang, W. Bruce Warr, Lloyd Watts, Ron Weiss, Kevin Wilson, Kevin Woods, Ying Xiao, Bill Yost, Tao Zhang, and probably others that I have missed. Many thanks to all!

And finally, huge thanks to Lauren Cowles, my editor at Cambridge University Press, for her years of patience in helping to make this book happen.