Psychology *of* Music

From Sound to Significance

SECOND EDITION

Siu-Lan Tan, Peter Pfordresher and Rom Harré

Psychology of Music

In *Psychology of Music: From Sound to Significance*, Second Edition, the authors consider music on a broad scale, from its beginning as an acoustical signal to its different manifestations across cultures. In their second edition, the authors apply the same richness of depth and scope that was a hallmark of the first edition of this text. In addition, having laid out the topography of the field in the original book, the second edition puts greater emphasis on linking academic learning to real-world contexts, and on including compelling topics that appeal to students' natural curiosity. Chapters have been updated with approximately 500 new citations to reflect advances in the field.

The organization of the book remains the same as the first edition, while chapters have been updated and often expanded with new topics. 'Part I: Foundations' explores the acoustics of sound, the auditory system, and responses to music in the brain. 'Part II: The Perception and Cognition of Music' focuses on how we process pitch, melody, meter, rhythm, and musical structure. 'Part III: Development, Learning, and Performance' describes how musical capacities and skills unfold, beginning before birth and extending to the advanced and expert musician. And finally, 'Part IV: The Meaning and Significance of Music' explores social, emotional, philosophical, and cultural dimensions of music and meaning.

This book will be invaluable to undergraduate and postgraduate students in psychology and music, and will appeal to anyone who is interested in the vital and expanding field of psychology of music.

Siu-Lan Tan is Professor of Psychology at Kalamazoo College in Michigan, USA. She completed degrees in Music at Pacific Union College, graduate studies at Oxford University, and a PhD in Psychology at Georgetown University. Her research focuses on musical form, music notation, and film music, and she plays piano. She is primary editor of *The Psychology of Music in Multimedia*, and appears in *Score: A Film Music Documentary*.

Peter Pfordresher is Professor of Psychology at SUNY Buffalo in New York State, USA. He completed his PhD in Psychology at the Ohio State University. His research on the relationship between perception and action in music has been published in psychology, music cognition, and music education journals, and has received support from the National Science Foundation. He sings and plays piano, guitar, and the trumpet.

Rom Harré is Emeritus Fellow of Linacre College, Oxford University, UK. He was Distinguished Research Professor of Psychology at Georgetown University, and Director of the Centre for the Philosophy of the Natural and Social Sciences at the London School of Economics.

Praise for the first edition

‘As it stands, Tan, Pfordresher, and Harré’s volume is an engaging exposition of the current state of our knowledge of the psychology of music …. To paraphrase Nietzsche, experiencing music without knowledge may not entirely be a mistake, but experiencing it with the kind of up-to-date knowledge that may be gleaned from *Psychology of Music: From Sound to Significance* is even more marvelous.’

– Aaron Kozbelt, *PsycCRITIQUES*

‘Tan et al.’s volume is an impressive achievement and merits serious consideration by anyone teaching a survey course in music cognition or seeking to recommend to a friend or colleague …. Its 300-odd pages are packed with the most detailed overview of our field we are likely to see in any text in the near term and it covers the major aspects of the field quite comprehensively.’

– Richard Ashley, *Music Perception*

‘[A]ll topics are introduced with sophistication and an interesting balance is provided between references to classical work and more recent work. Similarly, quantitative and qualitative investigations are both included …. The broadness of the book is splendid and allows for a complete introduction to the field. It will indeed be invaluable to undergraduate and postgraduate studies in the fields of psychology of music ….’

– Renee Timmers, *British Journal of Music Education*

‘The field of Music Psychology has received formative influences from many domains, thus it is no mean feat to create a representative survey of the literature. The authors have met the challenge, achieving a detailed and useful introduction to the field with this text.’

– Jessica Grahn, *Empirical Musicology Review*

‘We expect that this book will play an influential role in establishing the canonical organization for music psychology textbooks and hope to see it go through many editions in the years to come …. Meanwhile, we heartily welcome this ambitious book as a valuable new resource for teaching the psychology of music.’

– Roger Chaffin and Alexander Demos, *Psychomusicology*

Psychology of Music

From Sound to Significance

Second Edition

Siu-Lan Tan, Peter Pfordresher and Rom Harré

ROUTLEDGE Routledge
Taylor & Francis Group
LONDON AND NEW YORK

Second edition published 2018
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

and by Routledge
711 Third Avenue, New York, NY 10017

Routledge is an imprint of the Taylor & Francis Group, an informa business

First edition published by Psychology Press 2010

Cover design by Andrew Ward

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
A catalog record for this book has been requested

ISBN: 978-1-138-12466-0
ISBN: 978-1-138-12468-4
ISBN: 978-1-315-64802-6

Typeset in Times New Roman
by Keystroke, Neville Lodge, Tettenhall, Wolverhampton

Contents

About the authors

Siu-Lan Tan is Professor of Psychology at Kalamazoo College in Michigan. She completed degrees in Music at Pacific Union College, graduate studies at Oxford University, and an MA and PhD in Psychology at Georgetown University, Washington, DC. She plays the piano and taught private lessons and music classes for many years in Hong Kong and California before shifting to Psychology. Her research focuses on listeners' perceptions of musical unity, graphic representations of music, and the role of music in film and other multimedia. Her work has been published in *Music Perception*, *Psychology of Music*, *Psychomusicology: Music, Mind, and Brain*, *Empirical Musicology Review*, and *Journal of Applied Developmental Psychology*, among other journals. She often speaks on the topic of the psychology of film music, and was the Keynote Speaker for the 10th *Music and Moving Image* conference at NYU. She is also primary editor of the book *The Psychology of Music in Multimedia*, and appears in *Score: A Film Music Documentary*, which was selected as a *New York Times* Critics' Pick. Currently, she serves on the Board of Directors for the Society of Music Perception and Cognition.

Peter Pfordresher's primary training is in experimental psychology, though he is also a classically trained pianist and vocalist, and was for years a semi-professional rock musician. He holds degrees in Psychology from Georgetown University (BA), University College London (MSc), and the Ohio State University (PhD). He is currently a Professor at the University at SUNY Buffalo, New York State. His research (supported by the National Science Foundation and the National Association for Music Education) focuses on the interplay between perception and action in music, individual differences in singing accuracy, retrieval errors in performance, and how the brain processes melodic and rhythmic information from melodies. This work has appeared in prominent journals such as *Journal of Experimental Psychology: Human Perception and Performance* and *Cognition*, music cognition outlets like *Music Perception* and *Psychomusicology: Music, Mind & Brain*, and music education outlets like *Journal of Research in Music Education*. He is an associate editor for the journals *Music Perception* and *Psychological Research*, and has served on the executive board of the Society for Music Perception and Cognition, as Secretary.

Rom Harré first graduated in mathematics and held a lectureship in the University of the Punjab at Lahore in Pakistan. After graduate work in philosophy at Oxford University, he turned to philosophy of science as University Lecturer at Oxford. This work led to studies in the philosophy of psychology and ultimately to theoretical and empirical work in the field. Harré is author or editor of over 65 scholarly books in numerous fields, and has held appointments as professor in many institutions around the world. He was Distinguished Research Professor of Psychology at Georgetown University in Washington DC, and Director of the Centre for the Philosophy of the Natural and Social Sciences at the London School of Economics. Rom Harré's books include *Pavlov's Dogs and Schrodinger's Cat* (Oxford, 2009), *Key Thinkers in 20th Century Psychology* (Sage, 2006), *Wittgenstein and Psychology* (Ashgate, 2005, with M. Tissaw), *Cognitive Science: A Philosophical Introduction* (Sage, 2002), and *One Thousand Years of Philosophy* (Blackwell, 2000).

Preface to the second edition

When instructors first prepare for a new course, they often focus on rolling out as much information as possible to provide the full topography of a field of study. After a few cycles of teaching the course, they begin to sense where explanations need more clarity, and can anticipate frequently asked questions, and provide more intriguing extensions to topics to which students are naturally drawn.

And so it is with our first and second editions of *Psychology of Music: From Sound to Significance*. The first edition was inspired by the need the authors saw for a text for a course which two of us had been teaching since the late 1990s, and one had begun to teach a decade later. It took a team effort of over four years to produce the first edition. Updates and expansions to the second edition have been largely shaped through dialogues with users of the first book, both instructors and students alike (including some of our former students who are now instructors for this course), along with our own experience teaching with the text in our courses. In addition, we actively sought feedback from our classes and student focus groups involved during the preparation of this edition (please see Acknowledgments).

While continuing to aim for the breadth of scope and level of detail of the first book, our second edition updates the content with developments in the field reflecting the seven years that have passed, and also warming the content and tone to connect the subject matter to our readers' natural interests and points of curiosity. As we continue to tackle questions that lie at the heart of our field, we also consider questions we have heard our students and other readers ask:

Why do I like listening to sad music? (ch. 14)
What does my friend mean when he says he's 'tone deaf'? (ch. 11)
Can you also be 'deaf' to the beat of music? (ch. 6)
How do you get rid of a tune that's stuck in your head? (ch. 5)
Can you train yourself to have 'perfect pitch'? (ch. 5)
Are some people just born to be more musical than others? (ch. 10)
What do babies understand about music? (chs. 8 and 9)
How do I deal with my anxiety about playing in front of people? (ch. 12)
How does my neighbor, who wears a cochlear implant, experience music? (ch. 3)

We also address many other meaningful questions.

Organization of the book

The organization of the book remains the same as the first edition, while chapters have been updated and often expanded with new topics. 'Part I: Foundations' explores the acoustics of sound, the auditory system, and responses to music in the brain. 'Part II: The Perception and Cognition of Music' focuses on how we process pitch, melody, meter, rhythm, and musical structure. 'Part III: Development, Learning, and Performance' describes how musical capacities and skills unfold, beginning before birth and extending to the advanced and expert musician. And finally, 'Part IV: The Meaning and Significance of Music' explores social, emotional, philosophical and cultural dimensions of music and meaning.

In keeping with the spirit of the first edition, we intend our book to engage a wide range of readers, and have kept technical terminology for music and psychology to a minimum to ensure the readability of the text for a broad audience. Only a general familiarity with music and psychology is assumed, and key words are defined and often exemplified. (The only exception is chapter 7, which discusses prominent music theories and inevitably requires more technical knowledge than any of the other chapters.) Our illustrative figures also do not rely heavily on musical notation, and when employed, notation is usually limited to a single-line melody. Keeping to the practice of the first edition, neuroscientific studies are mainly contained within our chapter on neuroscience and music, and in separate sections toward the end of our chapters on melody, rhythm, practice, performance, and emotion in music. Depending on the emphasis of the course or particular interests of the reader, this allows the option of reading chapters while omitting these sections, without disturbing the continuity of the narrative. Where neuroscientific findings are interspersed elsewhere in the text, they are described at a more general level suited to a broad readership.

What's new about the second edition?

Our second edition updates the content with developments in the seven years that have passed, integrating over 500 new citations and references to reflect advancements in the field. Many compelling new topics have been introduced as expansions to chapters. A summary of the significant changes to chapters is provided in Table 0.1 in the 'Notes to Instructors' following this preface.

While continuing to present material with the depth and detail that we hope will be a hallmark of this book in all its eventual editions, the technical content has been clarified with more vivid and meaningful descriptions, and more illustrative examples, especially in chapters 2, 3, 4, 5, and 6. Greater emphasis has been placed on exploring topics that apply the concepts of the book to real-world contexts, and to personally meaningful applications that will engage readers. For instance, the technical content on the topic of sound and hearing (chapter 3) is balanced with new sections focusing on the practical issue of music-induced hearing loss in musicians and in listeners using personal listening devices, and recent research on music perception among individuals who use cochlear implants. As noted earlier, many of these new sections and extensions were inspired by intriguing questions from our readers.

Further, the second edition incorporates additional resources for learners. First, in response to the suggestions of students and instructors, we have employed ***bold italics*** to draw attention to key terms, accompanied by clear definitions. In general, we reserve *regular italics* for conveying emphasis and for non-English terminology (e.g., *accelerando*). In addition, in chapters containing a lot of technical terminology (such as the sections on the ear and brain), more fine-grained terms appear in *regular italics*. Second, in addition to our 'Notes to Instructors,' we have included 19 exercises and practical application projects designed to promote active learning, and to cement the knowledge gleaned from each chapter. These are included in a new appendix, 'The Chapters in Action.' Some are individual or group assignments, and others are conceptual replications of key studies that can be carried out with minimal resources.

Finally, we have also devoted more attention to social facets of music, interspersing more links to the social realm in the two developmental chapters, the chapters on practice and performance, and a significantly expanded chapter on the social psychology of music. Although the book continues to focus on classical music as a model, reflecting the focus in the existing body of research in the field, we have included more jazz and pop music and other genres in studies and examples.

Personal note

The publication of this second edition will mark the 90th birthday of our co-author, Rom Harré. Thus he has passed the torch to the two of us. Revisions for this second edition were therefore carried out by Siu-Lan Tan and Peter Pfordresher. However, it is important to note that Rom's expansive knowledge as an eminent scholar in many fields, providing greater breadth of topics than the book would have otherwise encompassed, continues to be felt in the current volume and we therefore continue to include his name as an author.

We have expressed our gratitude to our supportive colleagues, students, and publisher in our Acknowledgments section. And finally – but of *foremost significance* in our lives – we thank those who were our brightest lights throughout the journey of both editions of this book: our families. Siu-Lan gives her heartfelt thanks to Danny Kim, Khoen, Nanny, and May-Lan Tan, and Brenda Kim. Peter extends special thanks to Lyn, Emma, and Paul Pfordresher for all of our musical times together.

Siu-Lan Tan, Michigan
Peter Pfordresher, New York
Rom Harré, Oxford

Notes to instructors

The second edition of this book reflects the growing number of courses being taught in psychology of music around the world, and the increasing interest in this subject reaching a wide readership. The authors are delighted to learn that this book is currently being used in a wide range of programs, including music cognition courses in psychology and music departments, required courses in music therapy programs, and special topic courses offered as electives for a variety of majors. The book also appears on required reading lists as a qualification for entry into some music psychology graduate programs.

While the growth and vibrancy of this field is certainly exciting, there are also challenges involved in producing a book intended to give a 'snapshot' of a vital and ever-changing field of research and to serve a diverse audience. Thus, we have done our best to design a book to accommodate courses with different emphases, and readers with varied backgrounds in psychology and music. Early chapters provide general explanations of basic musical terms and concepts such as 'interval' and 'key' (chapter 5) and 'beat' and 'meter' (chapter 6), and we have expanded this basic coverage in the second edition. Throughout, key terms relating to psychology or the specific topics at hand are defined or explained, and visually accented by ***bold italics***. As mentioned in the preface, the structure of the book offers the option of reading chapters without the neuroscientific sections toward the end of some of the chapters. Readers will still gain some familiarity with findings of studies focusing on the brain, as they are also woven into other parts of the text at a level that is suitable for a broad readership.

Also new to this edition is our appendix, entitled 'The Chapters in Action.' In response to requests we have received from fellow instructors since the first edition, we offer 19 brief exercises and more extended application assignments or projects that are presented in a flexible 'accordion-style' manner – adaptable to different class sizes and academic levels, individual or group projects, and courses with different emphases.

No single psychology of music book can cover all territories any longer, thanks to the vibrant growth of this field of research since the days when the focus was almost solely on perception and cognition of music. Today, the psychological study of music has expanded in developmental, social, neuroscientific, and applied directions, and numerous subspecialties. Our book is designed to give readers a firm grounding in some of the main domains of the rich and interdisciplinary field of psychology of music, to be supplemented by primary readings selected by instructors to extend to their own research topics and areas of interest, and to motivate readers to continue exploring topics that pique their curiosity.

Changes since the first edition

We recommend that instructors see the Preface for a broad overview of the book and some new features of the second edition. The organization of the book and internal structure of most chapters remain the same, although with many new topics and expansions. Chapters 5 and 14 have undergone the greatest changes due to the growth in these areas, but the chapters maintain the same general structure. The main changes and additions since the first edition are summarized in Table 0.1.

Table 0.1 Overview of changes since the first edition.

1 Introduction	Updated with new examples. Otherwise, this brief opening chapter remains much the same.
2 Acoustics	Improved clarity of discussion on acoustical properties of sound through examples relating these properties to music. Expanded discussion of Fourier analyses to emphasize relevance to music. Updated studies on acoustics of concert halls, opera houses, and other musical venues.
3 Hearing	Improved clarity of technical content on perceptual qualities of tone and physiology of the ear. New section added on music-induced hearing loss in both musicians and listeners, including risks involved in the use of personal listening devices. New section added on music perception in individuals with cochlear implants.
4 Neuroscience	Reorganized to integrate discussion of research on music and language in the brain, including the question of whether music and language constitute separate neural 'modules.' New discussions of diffusion tensor imaging and neural plasticity as related to the effects of musical training.
5 Melody	This chapter has been overhauled. Expanded discussion of basic musical concepts with application to musical examples, including the relationship between pitch height and chroma, and the distinction between absolute and relative pitch. Three new figures help illustrate these concepts. Expanded coverage of absolute pitch perception. New coverage of 'earworms.'
6 Rhythm	Revised to describe basic concepts more clearly. Updated survey of research includes beat deafness and the perception of 'groove.'
7 Structure	Clearer description of how music may be said to have a 'grammar.' Otherwise, the chapter remains much the same.
8 Development	Updated and expanded sections, especially on infant perception of pitch, meter, and memory. Inclusion of infant music cognition studies with social implications. Expanded section on infant-directed singing coordinating musical and social synchrony between caregivers and infants, including brief examples of therapeutic use of maternal voice and music with infants.
9 Learning	Expanded sections on development of singing and movement to music. Musical play topic extended to older children, including musical games in playgrounds. New section on children's composition, both solo and collaborative. More attention to the role of technology, in musical play and computer-assisted composition.
10 Practice	Expanded discussion of the role of 'talent,' genetics, and deliberate practice. Expanded section on informal practice. New section added on injuries and other conditions prevalent in musicians, particularly playing-related musculoskeletal disorders, and passage on focal dystonia added to section on music practice and the brain. Updated studies on practice strategies.
11 Performance	Increased use of examples to promote reader engagement. Discussion of new studies concerning the role of memory in improvisation, motion capture analyses of performance, and poor-pitch singing. Includes discussion of emotional communication in performance, previously in chapter 14.
12 Social psychology	New section on music performance anxiety and complementary topic of performance boost. Expanded section on conductors, including effects on performers and audience. Expanded section on gender and music, including gender stereotyping of musical instruments and genres. Greater breadth on music and consumer behavior, including discussion of possible mechanisms.

13 Meaning	Added new citations of research relevant to the conveyance of meaning in music. Otherwise, the chapter remains much the same.
14 Emotion	This chapter has been overhauled. Expanded discussion of research on both perceived and felt emotions, and broadened scope of emotions being investigated. New section added on the use of music for emotion regulation, including mood maintenance, enhancement, and rumination. Brief section added on listener engagement with music evoking sadness. Updated section on film music. (The passage on how performers express emotion was moved to Chapter 11. The passage on emotion and culture was moved to Chapter 15, where it has been expanded).
15 Culture	Updated discussion of musical abilities in nonhuman animals and emotional communication in music. New sections on how non-Westerners perceive Western music, and memory for music of other cultures.

Exercises and application assignments: 'The Chapters in Action'

For exercises and application assignments designed to accompany the chapters in this book, please see our new appendix, 'The Chapters in Action,' which can be found toward the back of this book. It includes 19 brief exercises and application assignments designed to involve students in active learning, in tandem with chapter reading. We hope that these ideas will serve as a starting base that instructors can tailor to fit their classes, and will inspire further creative ideas to make the chapters come to life for students.

Our goal was to offer a book of broad scope, to provide instructors with a range of topics to choose from, to tailor to the level and composition of the class, and to meet the goals of different courses. It is our hope that this book will continue to serve a wide array of courses and draw many more inquisitive, bright minds to this fascinating field of study, as there is still much to explore and discover.

Acknowledgments

It takes a village to write a book, and no less of a community to create the second edition. We are grateful to many colleagues and students who contributed so generously of their time and talent to this book.

The authors would like to thank Kate Gee, Timothy Justus, James Mantell, and Douglas Vipond for offering particularly instructive suggestions for our second edition. We also drew inspiration from insightful published commentaries by Richard (Ric) Ashley, Roger Chaffin, Alexander Demos, Aaron Kozbelt, Jessica Grahn, and Renee Timmers. The authors are also grateful to our colleagues Arianne Abela, Gillian Anderson, Jonathan Ashmore, Leo Beranek, Meagan Curtis, Autumn Hostetter, Timothy Justus, Andrew King, Josh McDermott, Israel Nelken, Caroline Palmer, Jessica Phillips-Silver, Christopher Plack, Jan Schnupp, Michael Schutz, Mark and Cheri Shevy, and Aaron Williamon for their generous contributions to the second edition. We also thank Kathryn Lightcap, Graphic and Multimedia Designer, for expert assistance with figures.

As with the first edition, valuable student input shaped this book from its earliest drafts to the final manuscript. We express special appreciation to Elizabeth (Elie) Penix, Mackenzie Norman, Patrick McGuire, and Karen Chow for particularly detailed suggestions and meticulous editing of selected chapters. We thank our many dedicated student assistants for valuable feedback and assistance: Christina Dandar, Ashley Schmidt, Camila Trefftz, Emma Greenspon, Jacylyn Stoebe, Tim Pruitt, Sarah Teh, Amanda Knose, Jessica Troung, Nicholas Horwood, Haley Haner, Paul Kovacs, Anthony Khoury, Ryan Rondinaro, Anthony Nagib, Tiffany Thavisin, Thomas Gadelrab, and students in our psychology of music classes.

We are grateful to the wonderful staff at Routledge, Eleanor Reedy, Michael Strang, Elizabeth Rankin, Alex Howard, and Libby Volke, for all their guidance and support and for being genuinely caring advisors to us. We would also like to thank our dedicated copy-editor Huw Jones, Kelly Winter and the Keystroke staff, and indexer Michael Hamilton, for all the care and attention they put into the preparation of this book. And finally, our special gratitude goes to publisher Lucy Kennedy at Routledge, who first came across our original manuscript and advocated for the publication of our book.

From the first edition, we continue to extend appreciation to our colleagues Leo Beranek, Steven Brown, Ginevra Castellano, Roger Chaffin, Jennifer Cox, Simone Dalla Bella, David Hargreaves, Paul Jeffries, Patrik Juslin, Timothy Justus, Rohan Krishnamurthy, Scott Lipscomb, Martin Lotze, Josh McDermott, Rod Nave, Jessica Phillips-Silver, Christopher Plack, Dirk-Jan Povel, Mari Riess Jones, Michael Schutz, Keith Swanwick, David Temperley, Michael Tenzer, Laurel Trainor, and Robert Zivadinov; and to our former students (many of whom are now professors and researchers, and some who have taught courses with this book), Matthew Bezdek, Ryan Coppolo, Emily Dayton, Timothy Griffiths, Jackie Howard, John Kulpa, James Mantell, Amanda McFarland, Lauren Mitchell, Christy Peaslee, Christina Violante, Shanti Virupannavar, Elizabeth Wakefield, Sally Warner Read, and to all our psychology of music classes; and to the dedicated staff at Taylor & Francis who guided the first edition: Lucy Kennedy, Tara Stebnicky, Sharla Plant, Dawn Harris, Nicola Ravenscroft, and our copy-editor Paula Clarke. Our gratitude to all.

1 The scope of psychology of music

People everywhere and at all times for which we have records have picked out certain patterns of sound for particular attention. Some of these patterns are the stuff of what we call 'music.' What are the characteristics of the sound patterns we recognize as music? Why is it that these sound patterns have had a special significance for human beings?

All perceptible sounds begin with a propagation of energy into the environment. It may be a light breeze setting into motion a thousand fluttering leaves, the plucking of the strings of a harp, or the striking of a bass drum. What makes the particular dance of air molecules 'musical' in some instances, while other disturbances of air molecules seem to give rise to mere sounds? Or noise?

It is not always possible or desirable to give a formal definition of the topic of a program of study. We will not try to answer the question 'What is music?' in a neat, short formula. There are paradigm cases of music that most people in a particular culture can recognize. We can start with some exemplary cases from our own culture, such as a symphonic performance, a rock concert, or an advertising jingle. Many hear church bells, ringing out a simple melody, as music. However, not all sound is music. Could we draw up a similar catalogue of sounds that everyone would agree are not music? Perhaps the sounds of busy city traffic or the whine of a vacuum cleaner might strike us as obvious cases of nonmusic. But what about the sound of waves breaking on the beach? The howl of a wolf? Or the song of a bird?

While the extremes seem clearly distinguishable, there is no sharp line to be drawn between music and sounds that are not music. Whether sounds are taken as music or not depends in part on the context in which they occur. Debussy imitated the sounds of breaking waves in *La Mer*. Respighi's orchestral score for *The Pines of Rome* included a recording of a real nightingale's song. The blasts of real car horns punctuate Gershwin's energetic orchestration in *An American in Paris*, while Edgard Varèse – who questioned the distinction between 'music' and 'noise' – included two hand-cranked fire engine sirens in his influential percussion piece, *Ionisation*. And Malcolm Arnold's *Grand, Grand Overture* included three vacuum cleaners among the orchestral instruments!

In Hollywood films, the score often doubles or produces some of the sound effects accompanying a scene. In line with the idea that in space 'there is nothing to carry sound,' Steven Price's Oscar-winning score for the 2013 film *Gravity* straddled both music and sound design, providing a soundscape that conveyed the expansiveness of space, the tension and emotion of each scene – and afforded something akin to sound effects, enabling the audience to feel the impact of the shuttle breaking apart and the massive explosions that sent debris flying through space. There is also John Cage's *4′ 33″*, a composition in which the performer does nothing for 4 minutes and 33 seconds to allow ambient and incidental sounds to define the composition. Here, the boundaries between sound and music, and the very definition of music in the absence of sound controlled by the composer or musician, are brought into question.

All-encompassing definitions of music may be elusive. Nevertheless, despite the fuzzy boundaries of the domain of music, it seems that there are auditory phenomena which we can generally agree are music, and on which there has been agreement in many cultures and historical epochs. Music is produced and perceived by human beings. Performers must learn the necessary skills to create ordered sound in meaningful patterns; through exposure or training, listeners must learn to perceive those features of ordered sound patterns as music. There is clearly room for a systematic study of all these skills, as diverse as they may be. The merging of psychology and music leads the way to such examination, and opens avenues to numerous and diverse topics for study.

The scope of the field

The psychology of music is motivated by a great many questions. Among other things, the field of ***psychology of music*** is concerned with *the processes by which people perceive, respond to, and create music, and how they integrate it into their lives*. These topics range from the way in which the ear extracts the pitch of a tone, to the way in which music is used to express certain emotions or transform moods. Though this field of study makes important use of *cognitive psychology*, it also draws on many other branches of psychology such as *sensation and perception*, *cognitive neuroscience*, *developmental psychology*, *social psychology*, and applied fields such as *educational psychology*.

The perspectives of each of these domains within psychology and more have shaped this book to some degree, as evident in the overall plan: Our exploration begins with a consideration of the physical properties of a sound wave, the transmission of sounds to the ear, and the neural bases of the perception and cognition of music. We then examine more closely the perception and cognition of melody, rhythm, and musical structure. Next, we trace the emergence and development of auditory capacities and musical abilities, and the acquisition of musical expertise, culminating in musical performance. Finally, we consider the question of meaning in music, and the social, emotional, and universal significance of music.

Interdisciplinary connections

The psychology of music attracts not only psychologists and musicians, but scholars and researchers from a wide range of other disciplines. The present volume is also informed by perspectives from fields such as *acoustics*, *neuroscience*, *musicology*, *education*, *philosophy*, and *ethnomusicology*, among other disciplines. We will provide a brief overview of these connections here.

More than 2000 years ago it was realized that the musical possibilities of sound as heard were shaped and constrained by the physical properties of sound waves as they interacted with the amazing powers of the ear. Pythagoras linked the weight of a vibrating object to pitch, while his followers extended his intuitions to include the vibrations of strings, linking pitch to the length of the string, and harmonics to the simple numerical ratios of those lengths. *Acoustics* is the science of the production, propagation, and reception of those vibrations in the air that are relevant to hearing in general and music in particular. But it is also more – as discussed in chapter 2, the way that musical instruments and the human voice shape the physical processes that reach the listener, as well as the properties of the performance venues where music is produced and enjoyed, are also parts of the science of acoustics.

With respect to *neuroscience*, recently there has been a real surge of interest in the neural underpinnings of human musicality, and the current volume considers the way that neural activity may constrain or enhance our experience of music and music-making. In order for us to experience music, the brain must pick up and import physical patterns from the auditory signal. Our studies must include an introduction to auditory neuroscience, beginning with the anatomy and physiology of the ear as presented in chapter 3, and moving on to a comparison between the neural bases of music and language processing in chapter 4. Recent discoveries from the field of neuroscience of music are also presented in sections within subsequent chapters on perception of melody and rhythm, music practice and performance, and emotion in music in chapters 5, 6, 10, 11, and 14 respectively, and sprinkled elsewhere in the book.

Insights from the field of *musicology* (the study of the structure and history of music) also continue to be essential to psychology of music, as evident in chapters 5, 6, and 7 on perception of melody, rhythm, and musical structure, and in chapter 14 on the emotional power of music. For example, in studying the power of music to 'express' emotions and its capacity to 'induce' emotional states in the listener, musicologists have addressed intriguing questions such as: How do musical structures give rise to emotions in listeners? Further, how does music come to have 'meaning' for the listener? The latter question is explored in chapter 13. A study of the basic theory of a practice involves bringing at least some of the underlying presuppositions to light as explicit principles, and subjecting them to critical examination. This is one of the ways that *philosophical* analysis is also an indispensable part of the psychology of music, bringing out certain presuppositions in the practices by which psychologists try to reach an understanding of music as a human phenomenon.

Figure 1.1 Infant music cognition is just one of many growing areas of research that reflects the expansive scope of study in this field. Investigations in this area touch on various domains of study, such as sensation and perception, cognition, neuroscience, developmental psychology, social psychology, and applied areas such as music education.

Source: Photograph used with permission from the parent.

Within psychology of music, developmental psychologists are concerned with the emergence and maturation of musical behaviors, some of which are described in chapter 8 on music perception and cognition in infancy. How these emerging abilities may best be supported and refined is a question for *music education*, the focus of chapter 9. Musical performance requires the development of a set of highly elaborated skills, and a growing base of knowledge that allows for the sensitive interpretation of music. Innovative music education methods assume that all children are musical, and intensively immerse young children in creative and sensitive engagement with music with the aim of laying the foundations for lifelong musicality.

Finally, music is differentiated into a wide variety of musical cultures across the world, and as such, an anthropological perspective highlighting the study of distinct human cultures and their 'musics' is also relevant to an understanding of the psychology of music. These studies help us to differentiate between cultural sources of the features of a particular musical repertoire and those that may derive from the biological bases of musical perception. The application of anthropology to music is *ethnomusicology*, a major component of chapter 15. In this final chapter, we consider music as it is represented in a variety of societies around the world. But we shall not pretend to have answered entirely satisfactorily the enduring questions of what music *is*, how exactly it is created and received, and how music brings about the powerful effects that it does.

Range of research methods

Although the present discussion by no means exhausts all the questions and topics subsumed by the study of psychology of music, the richness and expansiveness of this field should already be apparent. Psychology of music is also distinguished by the scientific research methods commonly employed to study musical phenomena of interest, which leads one to a certain kind of explanation for the phenomena under investigation. There are many different kinds of explanations developed in the natural sciences, accompanied by various methods, and psychologists of music make use of many of them.

Throughout this book we will encounter a great variety of empirical studies, as various methods are needed to explore the whole gamut of musical experience. For some purposes, experiments manipulating specific variables in carefully controlled conditions are useful. A researcher may manipulate the tempo (or pace) of a song to determine if it alters listeners' interpretations of the emotion that is being expressed. For other purposes, the recording and analyzing of real-time phenomena of musical production and experience are appropriate, such as when identifying the steps a concert pianist takes to learn and memorize a complex piece of music. In some instances, physiological methods or brain-imaging techniques such as functional magnetic resonance imaging (described in chapter 4) may be employed, for example to examine changes in the body or brain while listening to or performing music. Some investigators have even monitored the heart rate, breathing rate, and brain activity of musicians and audiences during live or virtual performances. Methods such as motion-capture may be used to track the fine motor movements of a musicians' actions in real time, or to see how a musical quartet may coordinate their collaborative performance with communicative gestures. Electromyography may be used to measure electrical activity of muscles, in order to monitor subtle facial muscle movements as a singer is watching another vocalist perform. These are all examples of research topics discussed in the present book.

Qualitative studies and naturalistic observation also play an important part in studying the psychology of music. The nature of some aspects of musicality is best captured in rich observational studies of spontaneous activities in natural settings, such as observations of crowd behavior at live concerts, or of children playing musical games in playgrounds, or of parents singing to infants in their homes. These topics of study are also described in this book. Sometimes we can usefully sum up the musical experiences of a great many people in a sweeping generalization. In other cases, we need to pay close attention to music as it is perceived, appreciated, performed, or composed by an individual, if we are to reach a sufficiently detailed understanding of what is happening.

Since the days of the pioneering volumes in this field, including Carl Seashore's *The Psychology of Musical Talent* (1919) and *The Psychology of Music* (1938), Paul Farnsworth's *The Social Psychology of Music* (1958), Geza Révész's *Introduction to the Psychology of Music* (1954), and Robert Francès' *The Perception of Music* (1958/1988), a wide variety of methods have been applied to the study of psychology of music. It is a great mistake in building up a science to insist on the primacy of any one of the many methods of research that can be found across the universe of the sciences. Indeed, the psychology of music is fascinating in part because it draws on so many diverse bodies of knowledge and multiple modes of investigation! Only by adopting such breadth can we hope to understand how this remarkable human practice is possible.

Coda

Our introductory chapter has mapped the rich interdisciplinary terrain of the present volume. This book was undertaken by authors who first met some 25 years ago, when their paths first crossed. Brought together many years later by a shared passion for this field, we collaborated on this volume in response to our many enthusiastic students who have influenced us during almost two decades of teaching psychology of music classes. Clearly, the scope of inquiry of this field is vast, and no book can cover all that psychology of music encompasses. Our aim is simply to introduce readers to a broad range of classic and current research, and to ignite curiosity in the many intriguing questions and topics in this exciting and expanding field.

Considering what music meant to him after 50 years of studying the science of music, Carl Seashore – one of the important pioneers in this field – wrote in the preface to his 1938 volume of *Psychology of Music*:

> Then I was a stargazer; now I am an astronomer. Then the youth felt the power of music and gave expression to this feeling in the way he loved and wondered at the stars before he had studied astronomy. Now the old man feels the same 'power of music,' but thinks of it in the manner that the astronomer thinks of the starry heavens.
>
> (p. xi)

Like Seashore, we hope this volume will serve to bridge the gap between the 'mere love and practice of music to an intelligent conception of it' (p. xi). At the same time, while learning the various principles and practices of astronomy, we hope our readers continue to be stargazers – moved by music and constantly fascinated by its many enduring mysteries.

Recommended resources: Psychology of music

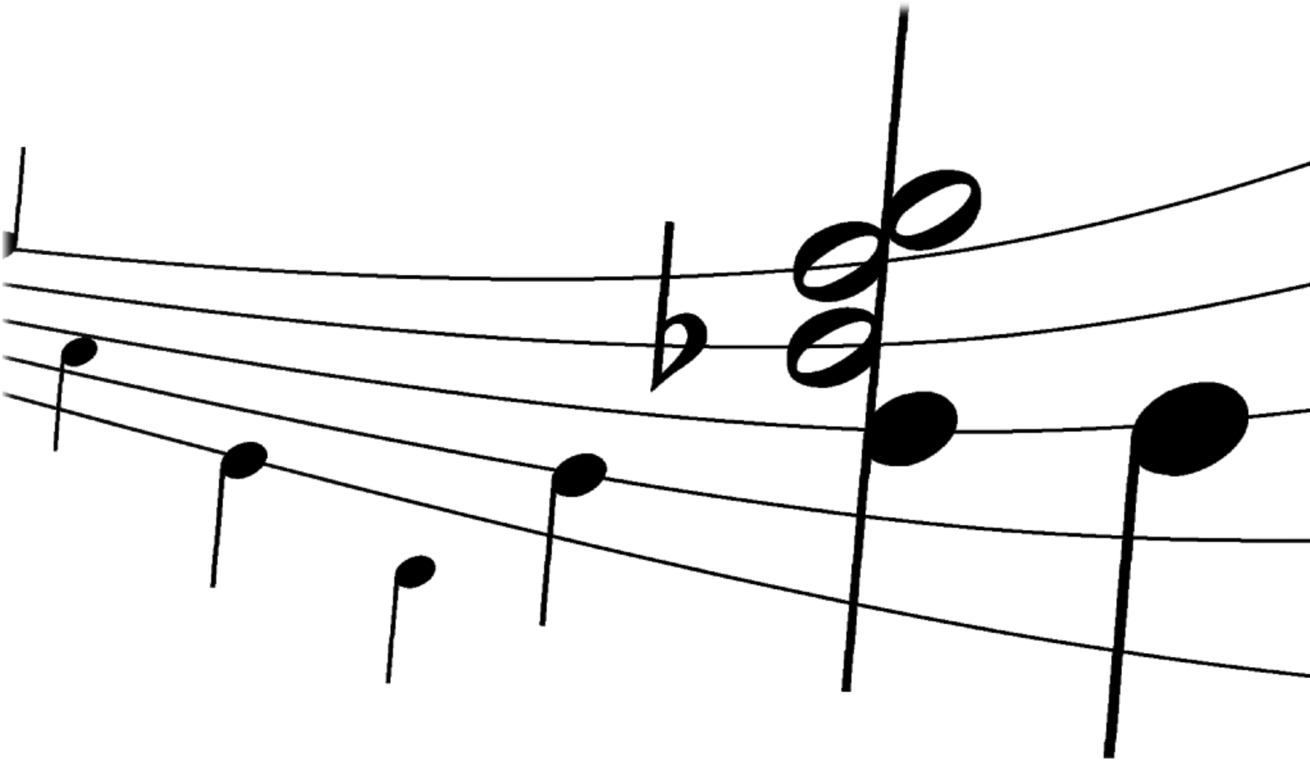
Ashley, R., & Timmers, R. (Eds.). (2017). *The Routledge companion to music cognition.* New York: Routledge.

Thompson, W. F. (2015). *Music, thought, and feeling: Understanding the psychology of music* (2nd ed.). New York: Oxford University Press.

Patel, A. D. (2015). *Music and the brain*. Chantilly, VA: The Great Courses [eighteen 30-minute lectures on topics complementing subjects in this book, available in video or audio format].

Part I

Foundations



2 The acoustics of music

Brittle twigs and dry leaves crunching underfoot, rain pellets hammering down on a tin roof, the roar of a crowd at a football game. Each of these sounds has its characteristic acoustic features and patterns that make it distinct from others. A distant flute on a hillside, the shimmering tones of a *gamelan* ensemble in a garden courtyard, the thunderous *fortissimo* chord of a full symphonic orchestra in a grand concert hall. Each of these musical sounds, too, has a distinct acoustic signature. All these auditory events – musical and nonmusical – are dispersed into various sound environments: open air, outdoor arenas, and crowded concert halls. The sound waves often travel a great distance to our ears, and the physical patterns must be transformed into neural signals in the ear and brain. It sounds like a complex process. Yet most of us quite effortlessly recognize the sounds and their sources, defining some as ‘noise’ and some as ‘music,’ and describing them as ‘sweet’ or ‘rough,’ ‘mellow’ or ‘shrill.’

Most of this book focuses on the subjective experience of music as a result of neural, cognitive, developmental, and social variables. However, the raw material that makes up music is the physical stimulus we call sound. Note that this stimulus is something distinct from either the object making the sound (e.g., a guitar), or our ear’s response to the sound. ***Acoustics*** is the science of sound, referring specifically to its production, transmission, and reception (*American Heritage Dictionary of the English Language*, 2016). The particular domain of *musical acoustics* focuses on the ‘mechanisms of sound production by musical instruments, effects of reproduction processes or room design on musical sounds, [and] human perception of sound as music’ (Hall, 2001, p. 2).

The present discussion begins with a brief description of the *physical* characteristics of sound waves (*frequency*, *amplitude*, *power spectrum*). We then illustrate a few principles of acoustics as they apply to a variety of musical instruments. Finally, we explore the acoustics of some grand concert halls and opera houses around the world. The perception of sound as music and the *perceptual* dimensions of sound (*pitch*, *loudness*, *timbre*) will be further discussed in chapter 3, accompanying a description of the auditory system.

Sound as physical stimulus

When a guitarist plucks a string, the resulting vibrations of the string lead to a corresponding disturbance in the air surrounding it. Air is important here; it functions as a medium through which sound propagates. As we will discuss later, the hollow body of an acoustic guitar is also important in amplifying the initial vibrations in the air that surround the guitar string. These physical disturbances propagate through the air and thus form a ***pressure wave***. Without some density in the air molecules, sound cannot travel and the environment is silent.

For instance, although sounds associated with the engines of giant spaceships as they streak through deep space makes for great dramatic effect in movies like *Star Wars*, in reality these battles in space would be entirely silent!

Figure 2.1 illustrates sound propagation. It shows a tuning fork, which creates sound similarly to plucking a guitar string. This figure allows us to explore the structure of a pressure wave in more detail. When the tuning fork is struck, its tines are set in motion. As a tine moves in an outward direction, the air molecules directly adjacent to it cluster together, leading to ***compression***. Then as the tine moves back in past its midpoint, the molecules spread apart, leading to ***expansion***. Of course, the tuning fork's tines, like a guitar's string, move back and forth (oscillate) rapidly over a period of time. As a comparison, think of what happens when you move your arm through the water in a swimming pool. As you move you arm to the right, your movement creates a bulge in the water that you push against. This is compression. Behind your arm, there is a kind of valley in the water – expansion.

This oscillating process leads to alternations of compression and expansion in the pressure wave surrounding the tine, and these disturbances ***propagate*** (i.e., spread out) across the surrounding air molecules. Thus, over time the pattern of compression and expansion leads to a ***sound wave*** (an acoustic pressure wave) that travels from the source of the disturbance (the tine or the guitar string) all the way to our ears, which respond to the sound wave in a manner described in chapter 3. In Figure 2.1, regions of compression are shown as dark bands to represent the clustering together of air molecules, and the spreading of molecules for regions of expansion is shown as light bands.

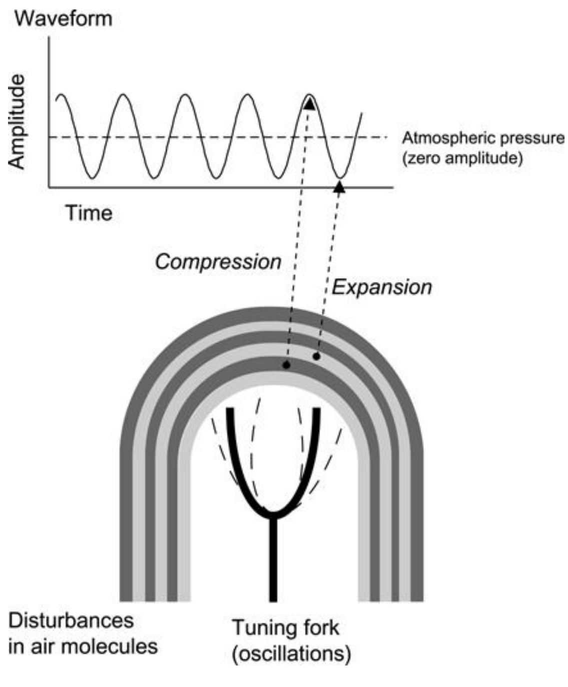


Figure 2.1 Propagation and representation of sound. Movements of the tines of a tuning fork (bottom) lead to alternating bands of compression and expansion among air molecules. For a pure tone (as in a tuning fork), the pattern of disturbances across time (for a given point in space) can be represented as a sine wave (shown at the top).

How does a sound wave propagate in order to reach our ears? A useful (though imperfect) analogy is to think again of water. When you throw a stone in a pond, the stone's impact creates ripples (compression and expansion) that radiate out from the source of its impact, though the extent of these waves diminishes with increasing distance from the source. Like a sound wave, the movement of these oscillations is longitudinal. In any ***longitudinal wave***, the movement of the oscillation is parallel to the direction of movement overall (as opposed to a 'transverse' wave, characteristic of light, in which the oscillation is perpendicular to the overall motion).

The movement of longitudinal waves can be demonstrated with a spring or Slinky toy. When a momentary force is applied to the first coil, it nudges the coil adjacent to it, which in turn nudges the next coil, and so on. The disturbance travels through to the other end like a wave, but the coils themselves do not travel from end to end. Each coil of a spring makes an oscillating motion that contributes to the wave and seems to pass it along, and then returns back to its place. In a sound wave, molecular movement affects adjacent molecules so that a periodic pressure wave is propagated through the medium (i.e., air) – but the constituents of the medium (i.e., the molecules) move only locally.

Although the essence of sound is a disturbance, a single discrete acoustic disturbance – such as a book falling on the floor – will not lead to the experience of music. Rather, the disturbances that lead to music follow regular patterns, resulting from vibrations in some sound-producing source. This is true of the plucked guitar string. Although the physical gesture is discrete (a single plucking motion), the string's vibration will continue for some time until the vibrations either diminish or the sound is stopped by another physical motion (e.g., damping the string with your hand). When the physical vibration is sustained, the surrounding air molecules begin to vibrate and the disturbance travels away from the source of sound, in an ever-widening wave of compressions and expansions, as shown in Figure 2.1. As it turns out, patterns of vibrations are integral to the link between sound and music. In the section that follows, we discuss different properties of these vibrations, and the sound waves (ultimately music) that result from them.

The properties of sine waves

One reason we used a tuning fork in Figure 2.1 rather than a guitar is that the tuning fork's vibration pattern is of the simplest sort: a sine wave (also called a 'pure tone'). A ***sine wave*** has a simple defining feature: Its pattern can be described fully using one frequency of vibration, which comes from the timing of alternation between compression and expansion phases. This vibration takes the smooth curved shape of a sinusoid because this is the natural pattern that a mass-spring system follows, and sound events fit this general model. Aside from the tuning fork, hardly any other sounds we hear in the real world can be fully described as a pure tone. Nevertheless, a sine wave is a clear way of representing important characteristics of sounds. Two parameters (i.e., measurable features) of sine waves that are critical for our discussion of musical sound are illustrated in Figure 2.2.

One parameter has to do with the rate at which a sine wave alternates between compression and expansion: ***frequency (f)***, the rate at which the crests or troughs of the wave pass a point in a given measure of time. Every time the sine wave completes a full cycle from compression to expansion and back again is considered a single period or 'cycle.' The number of cycles per second is referred to as ***hertz (Hz)***. Thus a sound source that is vibrating at 50 cycles per second is said to have a frequency of 50 Hz. In Figure 2.2, sine waves plotted in the top row have a lower frequency (1 cycle per second) than those in the bottom row (2 cycles per second; note that both signals would be below the frequency range of human hearing: see chapter 3). Many

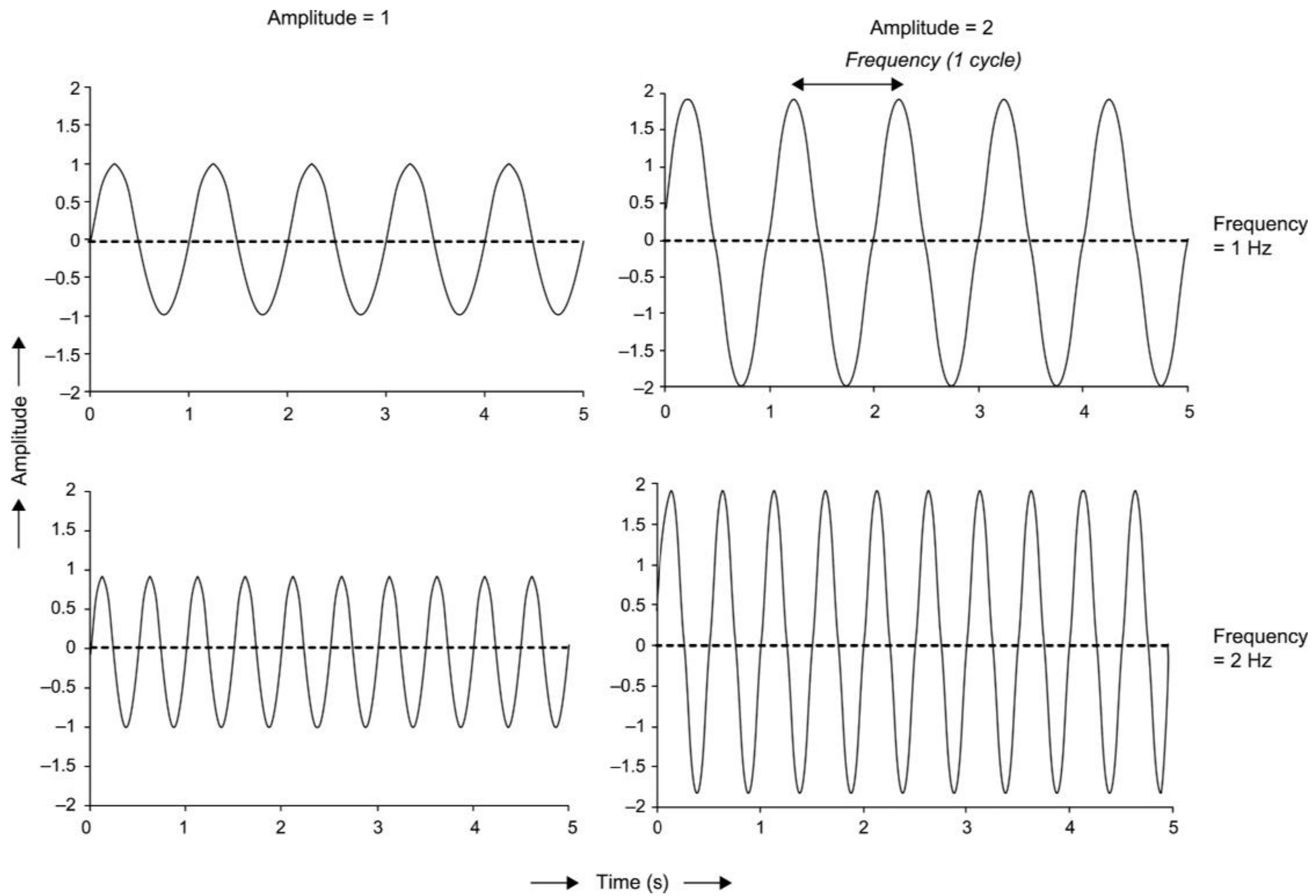


Figure 2.2 Four sine waves, varying in frequency and amplitude.

Source: Copyright © Peter Pfordresher.

physical variables can control frequency; a particularly important one is tension. When you pluck an open string on the guitar, then turn the tuning knob to tighten the string, and then play the same string again, you will find the pitch on the second pluck is higher than the first because the tuning adjustment has increased the tension in the string. A similar phenomenon happens in brass instruments when one tightens the lips (embouchure) while making the lips buzz, or when singers increase pitch by tightening the vocal folds (see chapter 11). Increasing tension causes sound frequency to increase. This in turn influences a perceptual variable: pitch.

In 1563 the mathematician-physicist Giovanni Battista Benedetti discovered that the frequency of vibration of the source of a sound gives rise to the aural sensation we call ***pitch***. We will discuss pitch at length in chapters 3 and 5; for now it is enough to define pitch as related to the experience of sounds as varying from high (a soprano singer) to low (a bass). When the frequency is high, there are many cycles in a given span of time, and the sound is perceived as high-pitched. A tone with a low frequency has few cycles in a given span of time, and the sound is perceived as low-pitched. Extending the physical to the musical, consider the fact that changes in this simple parameter, frequency, can enable us to experience the opening phrases of 'Mary Had a Little Lamb' or 'Twinkle, Twinkle, Little Star' as distinct melodies. It is important to note that pitch, though related to the physical frequencies of sounds, refers to subjective experience, and not to physics. Thus, whereas frequency refers to the physical signal, pitch relates to our experience of frequency information. We elaborate more on this important distinction in chapter 3.

The second parameter of a sine wave that is relevant to a discussion of music is ***amplitude***. The amplitude of a wave is the maximum displacement compared to the resting state, and corresponds to the perceptual dimension of sound that we experience as ***loudness***. Generally, the greater the amplitude of a wave, the more energy it transmits. Sound waves with more energy are generally perceived to be louder than sound waves with less energy, though the subjective effect depends to some extent on the frequency of the sound. In Figure 2.2, sine waves on the left side have lower amplitudes than those on the right side. In terms of physical gestures that lead to sound, the amount of force involved in a gesture roughly leads to changes in amplitude of the resulting vibration, which in turn affects the subjective experience of loudness. On the guitar string, a more forceful plucking action will lead to a greater initial displacement of the string, and thus a larger magnitude of alternating compression and expansion in the surrounding air.

Complex sound waves

Sine waves, as shown in Figures 2.1 and 2.2, have only a single frequency of vibration, which is why they are often called 'pure' tones (just as light with only one wavelength is called a pure color). But, as we will see, most sounds are best understood as a combination of many frequencies that occur simultaneously. The pure tone produced by tuning forks, beeps produced by automated teller machines, and old computers are well described by sine waves, but a friend's voice, the sultry saxophone, and other sounds of nature require a more complex description. Even a single plucked guitar string does not produce a sound wave that looks like a sinusoidal alternation between compression and expansion. Although the resulting sound wave may have a recognizable recurring pattern, it will not be sinusoidal and not be reducible to a single frequency of vibration. Not surprisingly, such sounds are referred to as ***complex tones***. An example is shown at the top of Figure 2.3.

If most natural sounds are complex, why discuss sine waves at all? The sine wave formula ends up being critical in order to make sense of the physical structure in a complex tone.

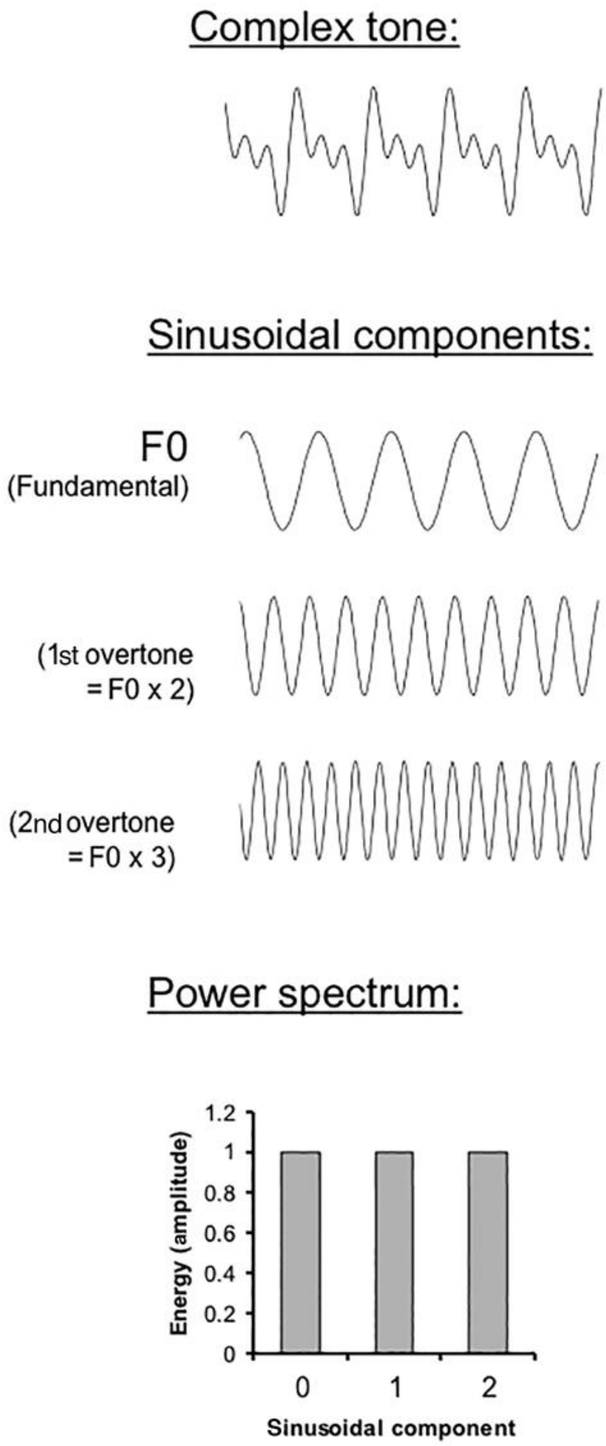


Figure 2.3 A complex wave, and below it a Fourier analysis.

Source: Copyright © Peter Pfordresher.

Specifically, any complex tone can be understood as a combination of sine waves. The process by which a complex wave is decomposed (broken up) into a set of component sinusoids is referred to as ***Fourier analysis***, named after French mathematician and physicist Jean-Baptiste-Joseph Fourier (1768–1830), who developed this technique.

Figure 2.3 illustrates how a complex tone can be decomposed into a series of sine waves. In this case, we need three sine waves (or 'components'), which are shown below the complex tone. The sum of these components yields the complex tone. Although the Fourier analysis was originally developed with engineering applications in mind, its application to sound does yield some important psychological implications because different components of a complex tone play different roles in perception. The component with the lowest frequency (longest wavelength) is commonly referred to as the ***fundamental frequency***. This frequency is often (though not always, as described in chapter 3) associated with the pitch of a complex tone. Higher-frequency components are often referred to as ***overtones***. The pattern formed by the overtones, called the overtone series, leads to a quality of sound referred to as ***timbre*** (conceptualized as sound quality or 'color,' also described further in chapter 3). Timbre helps us differentiate between the sounds of a flute and a clarinet, or to identify a voice on the telephone. A critical concept for understanding timbre is the 'overtone series' of a complex tone, which we turn to next.

Frequency relationships in the overtone series

The ***overtone series*** refers to the specific set of frequencies that are higher than the fundamental in a complex tone. The overtone series influences timbre in part based on the mathematical relationship among the frequencies in the overtone series. In many cases, the frequency of each overtone is an integer multiple of the fundamental frequency. For instance, in Figure 2.3 the first overtone is double the frequency of the fundamental and the second overtone is triple the frequency of the fundamental. This frequency pattern is similar to a harp string. When it is plucked, it does not just vibrate as a whole, but also in halves, thirds, quarters, fifths, and so on – each creating successively higher frequencies based on these integer ratios (2:1, 3:1, etc.). An overtone series like the one shown here defines a certain kind of complex tone, referred to as a ***harmonic complex tone***, and for these kinds of tones the overtones are often simply called ***harmonics***, with the term implying a certain mathematical relationship among the overtones. Harmonic complex tones are critically important for music because they convey a clear pitch, as we will discuss in chapter 3.

The numbering convention for frequency components varies depending on whether one adopts the more general distinction between fundamental and overtones (which can have any numerical relationship with each other), or that of a series of harmonics, including the fundamental (for the more restrictive class of harmonic complex tones). Figure 2.4 shows an example of the frequency components for C2 using music notation. There are 12 frequency components, starting with the fundamental frequency, which conveys the pitch at C2. By convention, the numbering of frequencies based on the fundamental and overtones starts at 0, so the fundamental is often referred to as F0. By contrast, numbering based on 'harmonics' (for a harmonic complex tone) starts at 1 (the fundamental being the first harmonic), so that the number of each component reflects its numerical relationship to the fundamental frequency (the second harmonic is double the fundamental, the next is triple, etc.).

Of course, not all complex tones are harmonic. As mentioned before, harmonic complex tones are characteristic of sounds with a pitch-like or 'melodious' timbre. On the other hand, sounds with a rough or 'noisy' timbre have overtones that are not so simply related to the

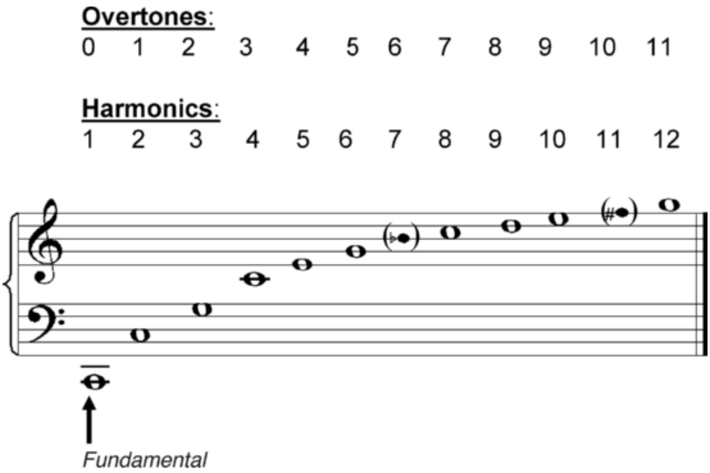


Figure 2.4 The harmonic or overtone series based on the fundamental C2, with harmonic and overtone numbers shown. Notes shown in parentheses denote approximate pitches. Numbering conventions for frequencies when considered as overtones or harmonics relative to the fundamental are both shown.

fundamental, called ***anharmonic complex tones***. Figure 2.5 shows an example of a complex tone (represented as in Figure 2.3) made up of sine waves that are not based on integer relationships. If played, the sound in Figure 2.5 would probably sound 'rough.' For instance, the static noise of a radio that is not tuned to a station comes from frequencies that are entirely unrelated to each other. Likewise, many percussion instruments (like drums and bells) include harmonics that do not conform to integer relationships.

In reality, the sounds we hear are almost never entirely harmonic, or entirely anharmonic (which would be white noise). However, this distinction provides a useful way to distinguish between different sorts of timbres. For instance, in the human voice, a smoother timbre (such as that of crooner Michael Bublé) is closer in structure to the kind of sound wave shown in Figure 2.3, whereas a rougher-sounding singing voice (perhaps Louis Armstrong or Tom Waits) would have more components that do not conform to such integer relationships.

Amplitude relationships in the overtone series

Differences among timbres are not only based on numerical relationships among sound frequencies. The second way that the overtone series determines timbre is through the amplitude associated with each frequency. Not all overtones necessarily emerge with the same intensity. If a harp string is plucked at its center, it is mainly the fundamental and the even-numbered overtones that will sound. Clarinet tones in the low register are almost completely lacking the first overtone. The sound each instrument produces is made up of a different composition of frequencies that in part give it a characteristic sound. This is another way in which Fourier analysis contributes to our understanding of sound. The bottom section of Figure 2.3 shows a representation of sound referred to as a ***power spectrum***, which plots the amount of energy (or power) associated with each frequency component in a tone (expressed as amplitude). Because energy is related to amplitude, the power spectrum can be thought of as a way of showing the amplitude of sine waves in a complex tone. The artificial

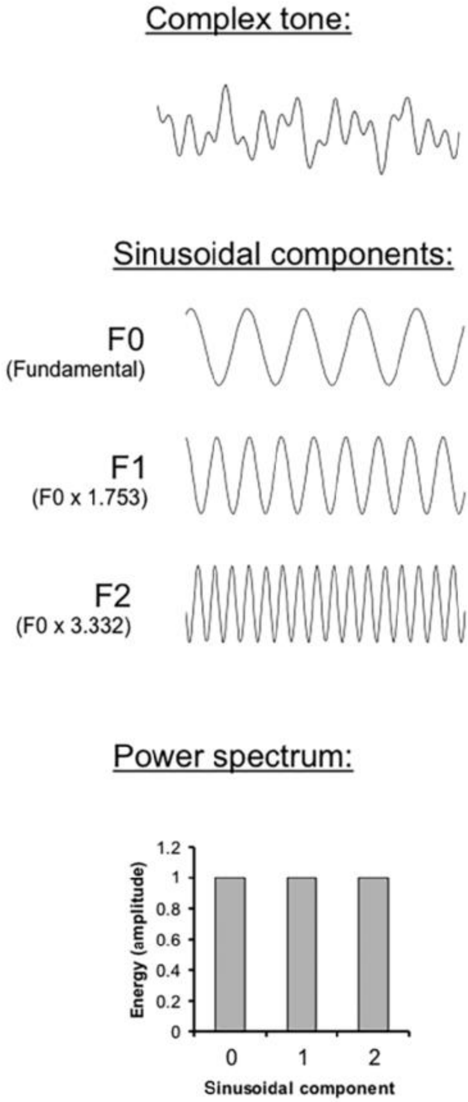


Figure 2.5 Waveform for an anharmonic complex tone, with its Fourier analysis below.

tone in Figure 2.3 consists of components that are equally 'powerful.' What happens when components are not similarly powerful?

How the amplitudes of different partials are distributed defines an important measure of timbre: a sound's ***spectral centroid*** (McAdams, 2013). The spectral centroid reflects the 'center of gravity' in the sound's spectrum. Two examples are shown in Figure 2.6. When the spectrum is dominated by low-frequency energy, as in the left-hand sound wave in Figure 2.6, the centroid is based at a frequency close to the fundamental, and the subjective quality is usually considered to be 'rich' or 'dull.' By contrast, a higher centroid, as in the right side of Figure 2.6, can lead to sounds having a 'bright' or 'nasal' quality.

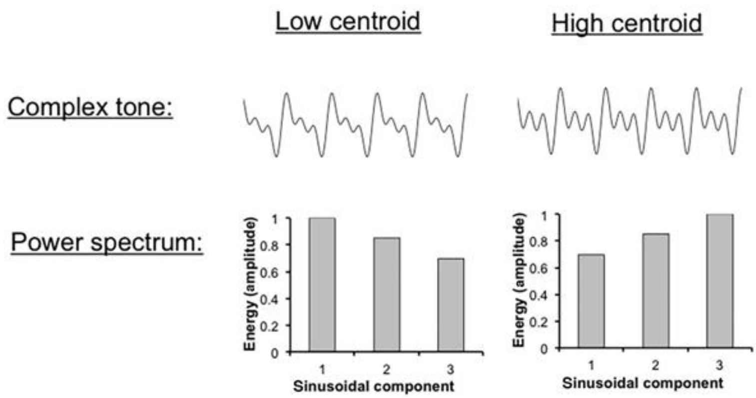


Figure 2.6 Waveforms for two complex tones differing in spectral centroid. The Fourier analysis is shown below each waveform.

Source: Copyright © Peter Pfordresher.

Thus far we have emphasized acoustical differences *across* instruments (and singers). However, it is important to note that the spectral structure of a single instrument is not constant, but can change as a function of pitch, loudness, duration, or expressive intentions.

An example of how pitch influences the power spectrum is shown in Figure 2.7. This figure was generated by recording a trombonist producing B♭ in three different octaves. The left side of the graph shows (complex) waveforms associated with each pitch, and the right side shows their power spectra (cf. Figures 2.3 and 2.5). Each B♭ is shown as F1 in the power spectra, along with its higher overtones. In principle, the power spectrum for each produced pitch could be identical because frequencies are plotted by harmonic number (1 for fundamental), so the fundamental frequency is always to the far left of the X-axis. However, this is not the case. The acoustic structure emphasizes relatively lower frequencies within the range of frequencies presented when the performer produces a higher pitch, and emphasizes relatively higher frequencies when the performer generates a lower pitch. These differences ultimately reflect the fact that a certain (intermediate) range of frequencies is likely to have a higher level of energy across many different produced pitches in the trombone. This range ends up being relatively high when one plays a low pitch and relatively low when one plays a high pitch. Such bands of high energy that are enhanced by the structure of a sound medium are known as ***resonances***, and may help the listener identify a specific instrument. We discuss the concept of resonance more in the next section.

Fourier analysis involves complex mathematics for which we have given a verbal sketch. Readers interested in this topic may wish to refer to Benson's (2006) *Music: A Mathematical Offering*, which includes a good chapter on Fourier analysis in a musical context, or Kammler's (2008) *A First Course in Fourier Analysis* (which, despite its title, is an in-depth exploration with applications to music and other areas).

The acoustics of musical instruments

Musical instruments (and we include the human voice among them) are devices for producing sound waves. Each species of instrument has its own acoustic characteristics, and the physical structures of instruments vary widely. However, almost all musical instruments depend on the

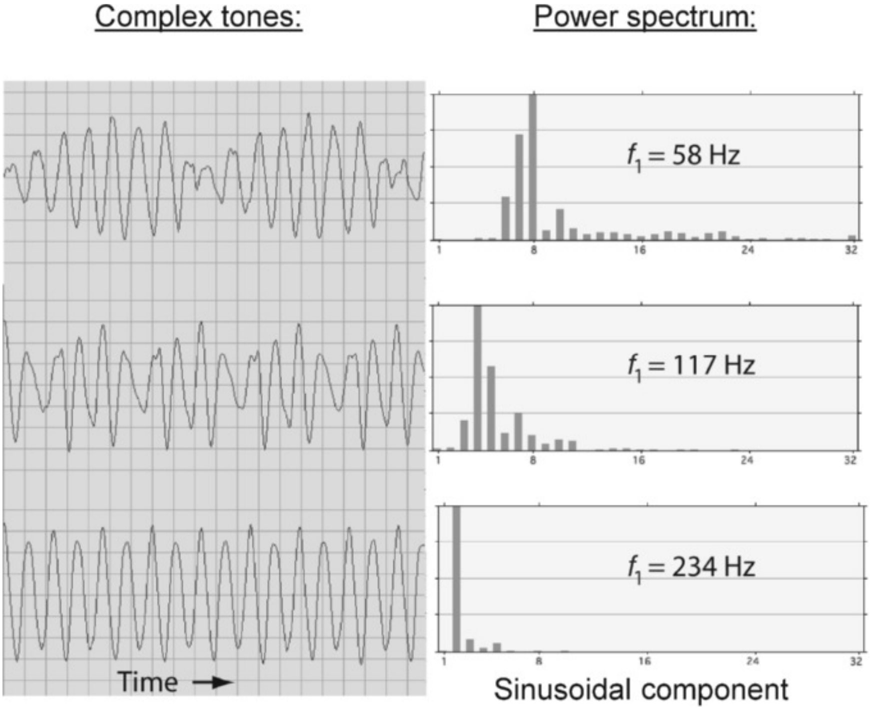


Figure 2.7 Waveforms (left) and power spectrum plots (right) for three tones produced by a trombone. (Power spectrum plots show frequency values as harmonic number rather than Hz.)

Source: Figure created by Matt Segars and composited by Rod Nave at the Department of Physics and Astronomy at Georgia State University, used with permission.

principle of ***coupled acoustics***. That is, in most musical instruments there are two vibrating devices: One *generates* the sound wave, and the other is ***coupled*** to it, and *amplifies* it.

The initial vibration that is activated by the performer – for instance, by blowing into an oboe or bowing a violin – is often of insufficient energy to be properly heard in a large space or at a distance. A ***resonator*** (such as a hollow tube, an air chamber, or a flexible soundboard) is ‘coupled’ to the basic device. The importance of coupling can easily be demonstrated by blowing the reed of a clarinet mouthpiece detached from the instrument, which will only produce a thin squeak. In the clarinet, the vibrations of the reed excite oscillations in the column of air in the body of the instrument. The air column is the resonator. In the double bass, the vibrations of the strings resonate in the hollow body of the instrument. The similarity between the terms ***resonator*** and ***resonance*** (introduced earlier) is no coincidence. Any given resonator is likely to serve as a better conduit for certain frequencies than others. These resonances, as mentioned before, help give a particular instrument or voice a characteristic timbre across many different pitches and loudness levels.

The groundwork for what we know about sound resonators and many other acoustic phenomena was laid by Hermann von Helmholtz in 1863, with his publication of *On the Sensations of Tone*. One of the leading physicists of the 19th century, Helmholtz’s interest in music led him to propose a research paradigm for bringing the physical sciences into a working relationship with the arts. Realizing that the discoveries of physical acoustics (such as the ratios of modes of vibration of a taut string) were correlated with experiences of consonance, harmony, and other phenomena that were the basis of music, he looked for a bridging science. He found this in physiological acoustics, the anatomy and physiology of the ear, and neural links to the brain.

Most pertinent to the present discussion is Helmholtz's experimentation with sound resonators consisting of a large collection of hollow spheres (of many different sizes) made of glass or metal, with two tiny openings on each end. One opening was to be placed snugly inside the ear for an airtight fit, and the other was to be pointed at the source of a sound. Helmholtz discovered that if the sound contained an overtone with the same or similar frequency as the natural resonant frequency of the hollow space inside the resonator, the overtone would be amplified and would sound inside the resonator. Just as 'Helmholtz resonators' of different sizes and materials amplify different overtones, the properties of the resonators of musical instruments in part determine which overtones are amplified or suppressed. 'Brighter' tones occur when the higher overtones are present and amplified, and a 'duller' sound is produced when upper overtones are weak or missing.

Strings

We can group instruments in which sound is produced by setting a string into motion into three main groups. In one group there are 'bowed strings.' These are the instruments such as violins and double basses in which a string is continuously supplied with energy from a moving bow, and the vibration is maintained at the same pitch to give an enduring tone. In another group such as harpsichords, guitars, and mandolins, a string is set in vibration after brief contact with a plectrum or some other device that plucks the string (in the case of Jimi Hendrix, sometimes his teeth!). Then there is the third group of instruments in which the string is struck with a hammer, as in the piano, which also belongs to the family of percussion instruments. In most of these instruments, the hammer rebounds almost immediately off the string. Whether or not the string continues to vibrate depends on whether it is damped by contact with some material such as felt.

Not all vibrations of a string are controlled by the actions of the performer. One of the three strings of the long-necked string instrument the Indian *sitar* is not plucked, but vibrates in sympathy with the vibrations of the melody and drone strings. A subtler effect of *sympathetic vibration* occurs when an object vibrates not because it is stimulated by a direct action such as plucking a string, but because vibrations in a sound wave cause the object to vibrate due to the resonant frequencies of that object. The phenomenon of sympathetic vibrations can be demonstrated on a grand piano when playing one or more keys while pressing the sustain pedal (the rightmost pedal). The pedaling action removes the dampers from the strings, freeing many strings you are not striking with the piano hammer to vibrate in sympathy with the strings corresponding to the depressed keys.

Despite the variety of ways the strings are set in motion, all of these musical sounds follow the same principle of coupled acoustics. The vibrating strings radiate a weak sound initially, but the energy is transferred to the bridge and top plate, and into the body of the instrument and the back plate. In addition to other factors such as the thickness, length, mass, and tension of the strings, the particular characteristics of the resonating body play a critical role in the resulting sound's timbre.

Woodwinds and brass

The same coupling principle that lies behind the production of musical tones by stringed instruments also applies to wind instruments. A vibration produced in the mouthpiece sets a column of air resonating. It is this vibration that is transmitted to the listener by the air. There are several ways in which musical sounds are produced in woodwind instruments, a few examples of which are discussed next.

Edge tones are created by an action such as blowing across a hole and onto a sharp edge. The tones are produced because the upper and the lower planes of an edge produce asymmetrical vortices as the air stream passes over the edge. This asymmetry induces an oscillation in pressure, which is picked up and amplified by the adjacent air column in the body of the instrument. The effect only works if the incoming air stream is narrow, as for example from the pursed lips of the flute player (essentially a hole), or through the embouchure of the clarinetist (which is essentially a slit). This is another example of coupled acoustics, because the physical process at the edge is quite different acoustically from the wave transmission in the air column. But the instrument only produces musical tones because the two processes are coupled.

Reed tones are produced by the player setting one or two reeds vibrating, which in turn produces sound waves in the adjacent air columns. This is the system to be found in single-reed instruments such as clarinets and saxophones, and double-reed instruments such as oboes and bassoons. In the Scottish bagpipe, the player does not put the reed into the mouth, but blows into a reservoir in the form of a bag; when the bag is squeezed, sound is created as air passes through the reeds of three pipes creating drones and a fourth 'chanter' pipe providing the melody. Brass instruments use 'reed tones,' but the 'reeds' are the human lips; the vibrating motion that creates the sound is independent of the instrument, being freely controlled by muscle tension. Brass instrumentalists must expend a lot of effort developing the muscles of their lips as a way of controlling pitch height and intonation (good brass instrumentalists can play music simply by 'buzzing' their lips). Finally, we should notice that the human voice is a reed tone instrument, the vibrations of the vocal folds (also known as vocal chords) resonating in the air chambers of the respiratory system. The acoustics of the vocal resonators are complex; imagine the constantly changing shape of the resonating chambers of a vocalist singing a lyric such as 'Hallelujah!'

There are interesting variations in tone quality depending on whether the body of a wind instrument is of conical bore or even bore throughout, and how it flares at the end. A flared bell radiates the high overtones, which in part gives the trumpet its 'bright' sound. The length of the resonating columns of wind instruments also varies widely; generally, longer pipes produce lower or deeper tones. The tubing of a French horn when unrolled into a straight pipe, for instance, measures about 2.7 to 3.7 meters (depending on the type of horn). This is approximately two to three times the length of the pipe of a trumpet, accounting in part for the warm, rich tones it can produce. It is the length, shape, and materials of the pipe resonator that, to a large extent, account for the distinct sounds produced by different wind instruments.

Percussion

Unlike the violin or the clarinet, in which the performer continuously supplies energy to the system, percussion instruments receive energy in short bursts. A drumstick strikes the membrane of the drum in a single stroke. This stroke sets the relevant parts of the instrument vibrating at their natural frequencies, depending in part on the size and rigidity of the materials with which it is made. As the surface of the drum is flexible, the impact also sets into motion a wave disturbance that travels to other parts of the surface that were not directly struck. Thus a single strike of a drum generates many frequencies that bear no simple relationship to each other, producing an 'unpitched' sound. The induced vibrations then die away until the performer makes another stroke.

Drums, chimes, marimbas, and many percussion instruments have resonating chambers or tubes. The coupled acoustics of the *saron*, one of the pitched percussion instruments in the Indonesian *gamelan* orchestra, can easily be deconstructed as it consists of a few loose metal keys, simply placed over a shallow wooden box, which serves as a trough resonator. Without it, the *saron* would barely be heard over the full-percussion *gamelan* that often performs in outdoor venues. The piano's main resonator is the soundboard; essentially, it repeats the vibratory motions of the piano strings and must be carefully crafted if the instrument is to produce a rich sound. However, not all percussion instruments have resonators. For instance, vibrating plates such as cymbals and gongs have no built-in 'amplifiers.'

Across the huge diversity of musical instruments which human ingenuity has created, the basic principles are the same. With very few exceptions, instruments produce pitched sound waves, from which the listener extracts the basic musical qualities of pitch, loudness, timbre, and variations in duration (rhythm). Almost all instruments depend on the principle of coupled acoustics to generate sound waves that are clearly audible at some distance from the performer, and the various resonators contribute their characteristic properties to the sound wave that ultimately emerges. However, the quality of sounds produced by the instruments and voices is also affected by the characteristics of the physical environments into which they are released. In the final section in this chapter, we consider the sound fields in which music is commonly performed and appreciated: concert halls, opera houses, and other performance venues.

The acoustics of musical venues

In 1895, Wallace Clement Sabine, professor of mathematics and philosophy at Harvard University, was summoned to address a practical problem: Listeners were having a hard time understanding what speakers were saying in a lecture hall in Harvard's Fogg Art Museum. After performing a series of tests, Sabine determined that sounds in the hall were sustained within an audible range for a long duration before decaying, making speech indistinct. Sound-absorbent materials were placed in the room, and the problem was remedied. Prior to this time, buildings were not usually constructed with acoustics in mind. The first music hall to be designed with the help of acoustic engineering was the New Boston Hall (Boston Symphony Hall) in 1900, for which Sabine applied formulas he developed for calculating ideal reverberation times. He was soon regularly consulted to assist in building design and address acoustic problems of completed structures. Sabine's work laid the foundations for ***architectural acoustics*** (a field that encompasses the analysis, design, and control of sound in a building or other structure), which was later continued in the realm of music performance venues by Beranek (2004, 2007).

Today, the study of the acoustics of a concert hall or opera house has come to encompass more than just the (rather complex) physics of sound propagation in an enclosed space. It has come to include the consideration of features that affect the perception of music, and often also the subjective evaluations of the listeners and performers. The best conditions for hearing an orator can be fairly well defined in acoustic terms, as the focus is often on the clarity and intelligibility of speech. However, a host of other factors, such as the type of music being performed and even the reputation of the music hall, are relevant to how music sounds to a particular listener or performer. Following important work by Yoichi Ando (1985) and others on the *subjectivity* of architectural acoustics, much of the research on the acoustics of music performance venues takes listeners' and performers' subjective preferences into account (e.g., see Gade, 2015). Although not the focus of this chapter, there are also many social variables (such as the degree to which patrons feel socially comfortable in the venue and familiar with the orchestra), which cannot be isolated from the perceived musical quality of the concert experience (Pitts, Dobson, Gee, & Spencer, 2013).

Studies addressing the quality of the experience of music under different conditions can be carried out in several ways. It is possible to create artificial sound fields with properties that mimic the structures of the sound fields one would find in a concert hall, church, or other enclosed space. For instance, different materials can be placed inside an ***anechoic chamber*** (an insulated room designed to completely absorb reflections of sound and thus be free of echoes) to shape the way sound fields are created, and listeners may be asked to judge the quality of the sound piped in through loudspeakers. Another method is to create ***binaural musical recordings*** (by employing a model of a human head with stereo microphones inserted in each ear, in order to record sound traveling to both ears) to quite faithfully reproduce musical performances in different halls. Listeners can then compare these recordings inside an anechoic chamber.

Increasingly, researchers are also using ***computer modeling*** (involving computer programs to simulate real-world situations and processes) to simulate how sound is propagated in a room or concert hall. For instance, the LIVE (large interactive virtual environment) lab at McMaster University uses computer modeling to create a virtual acoustic environment that can be controlled by researchers to simulate how a piece of music sounds in different performance spaces (such as a jazz club, subway station, or cathedral), while measuring heart rate, breathing rate, brain activity (by electroencephalography), and many other responses of the performers and audience.

None of these procedures, however, captures the complete experience of sitting in the performance hall during a live concert! Therefore, an alternate method that has been used is to ask listeners to give their impressions of the acoustical qualities of live performances in music performance halls of which the sound fields and acoustic parameters have been studied. Leo L. Beranek (2004) used this approach in an extensive study of 50 of the major concert halls and opera houses of the world, originally reported in 1962 and subsequently revised in an expanded volume in 2004. He set about interviewing conductors, music critics, experienced listeners, and performers about their experience of music in various venues. Then, using standard acoustical concepts and techniques, he also measured the physical attributes of the halls in order to also collect more objective data. As we shall see later in this chapter, he found some consistency in listeners' and performers' preferences in some of the world's great music halls.

Direct and reflected sound

Our discussion of the acoustics of venues for musical performance focuses on a few basic concepts as they apply to the quality of listeners' and performers' experience in music performance venues: directed and reflected sound, sound absorption, and reverberation time. Throughout, both physical (or objective) and subjective parameters will be considered.

Direct sound travels directly from the source to the listener; it contains auditory information in an uncontaminated form. The clarity of the sound of an orchestra depends heavily on direct sound. ***Reflected sound*** reaches the listener by 'bouncing off' one or more surfaces such as walls, ceilings, pillars, and sound baffles. Depending on the time interval between the arrival of direct and reflected sound at the ears of the listener, the added reflected sound may add a pleasant richness to the musical tones – or it may 'muddy' up the sound and even create distracting echoes. At high reflection levels, the delay associated with these echoes is around 50 milliseconds (50/1000ths of a second; Gade, 2007). This time delay may seem small, but for complex musical passages it can interfere considerably with the experience of music.

An important measure is the ***initial time delay gap (ITDG)***, defined as ‘the time at which the first reflection is heard after the direct sound’ (Beranek, 2007, p. 4). An ITDG of less than 25 milliseconds has been found in the best concert halls, including the Concertgebouw in Amsterdam and Boston Symphony Hall. This allows listeners to detect clear onsets of musical tones, to hear successive tones distinctly, and to localize where sounds are coming from on the stage. If the ITDG is greater than about 35 milliseconds, ‘the hall will sound like an arena, with a lack of intimacy’ (p. 4).

There is also the clarity of sound to consider, or the extent to which musical tones or a singer’s lyrics sound clear and distinct. The ***clarity index*** is a measure of the ratio of early sound energy (from ***early reflections***, arriving within about the first 80 milliseconds of the direct sound) to late sound energy (***late reflections*** typically arriving after 80 milliseconds). The preferred clarity index of a performance venue varies with the characteristics of the music being played; contrapuntal Baroque music demands higher clarity than music of the romantic period (Reichardt, Alim, & Schmidt, 1974, p. 243) so that the interlaced lines or ‘voices’ can each be heard distinctly.

Two-thirds of the world’s concert halls that were most highly rated by conductors, music critics, and experienced concert-goers in Beranek’s (2004) study are ‘shoebox’-shaped. The narrow width of the halls allows for stronger ***lateral reflections*** (from side to side) to complement the sounds coming from directly in front of the listener (from the stage). Strong lateral reflections tend to produce a sensation of being ‘bathed’ in the sound, and have been shown to enhance musical dynamics (i.e., expressive variations between soft and loud) for listeners (Ando, 1985; Pätynen, Tervo, Robinson, & Lokki, 2014). If a shoebox-shaped hall is too wide (more than 25 meters across), the sound produced will tend to be muddy. However, if the hall is too narrow (less than 15 meters wide), the direct sound and first reflected sound may merge so that the sound reflected from the closest wall could mask and diminish the vibrance of the sound coming directly from the stage (Beranek, 2015).

The ‘shoebox’ design should not imply a long narrow hall with flat walls and plain ceilings. Sound is dispersed more uniformly throughout a space when there are some irregularities in the ‘box,’ such as coffered (ornamental, recessed) ceilings, niches in the walls, balconies, statues, and textured surfaces that can diffuse or diffract sound waves. Irregularities in the lower part of shoebox-shaped halls diffuse reflected sound and create the impression of ‘warmer’ tones, and those in the upper part of halls enrich the reverberant sound (Beranek, 2015).

Fan-shaped halls are often more problematic. The fan shape leads to progressively widening walls toward the back of the hall, which directs lateral reflections away from the listeners. This may weaken the sense of ***acoustic intimacy*** (the degree to which sounds seem to be coming from nearby rather than remote surfaces), ***listener envelopment*** (the sense of being surrounded by reverberant sound coming from all directions), and ***warmth*** (provided by transmission of low frequencies such as bass tones). Further, while the first reflections arrive from the narrow parallel walls in shoebox-shaped halls, the first reflections from halls that fan out from the stage are likely to come from the more distant path of the tall ceilings – thus making for a longer reverberation time. Concave walls behind the audience also tend to send reflected waves back to a point on the stage, and can create echoes – an annoyance to many performers (especially instrumentalists).

To give an interesting case study, a notable exception is the fan-shaped *Aula Magna* at the University of Caracas in Venezuela, which receives good reviews from performers and critics. To address the potential acoustical problems stemming from the nonparallel walls and curved features of the architectural design, it was decided during construction that sound-reflecting panels often referred to as 'clouds' (covering an area equivalent to 70% of the ceiling) were to be hung below the ceiling and along the side walls. Rather than using standard rectangular panels, sculptor Alexander Calder was commissioned to create suspended panels in the dynamic abstract shapes characteristic of the geometric-shaped mobiles for which he is known. The result, shown in Figure 2.8, is both acoustically and visually pleasing!

However, pure laws of physics alone do not always predict audience preferences. Eight of the nine top-rated opera houses in Beranek's (2004) study are horseshoe-shaped. He noted that the horseshoe shape is not acoustically ideal (the curved wall behind the audience directs sound back to the center of the stage, and balconies create areas of low resonance underneath them), and yet seemed to be preferred by listeners and performers for other qualities. For example, this arrangement brings performers and audience closer together, which increases ***visual clarity*** so that facial expressions and gestures can be clearly seen, and ***acoustical clarity*** for the intelligibility of the lyrics and speech. The slight echo common in halls of this shape provides feedback to singers, and does not seem to distract listeners.

Sound absorption

Not all sound waves reach the ear directly (direct sound) or bounce back from dense surfaces (reflected sound) in a contained space. Some become trapped in materials such as ceilings

Figure 2.8 The interior of the Aula Magna at the University of Caracas in Venezuela, featuring 'clouds' by sculptor Alexander Calder.

Source: Published courtesy of Leo Beranek, *Concert Halls and Opera Houses* (Springer, 2004).

and walls, stage curtains, carpeting, and seat upholstery – and reflect little energy back. This sound has been *absorbed*.

Absorption materials commonly found in music performance venues can be classified into two kinds. ***Porous absorbers*** (curtains, theater seats, carpets) absorb high-frequency sounds more efficiently than bass sounds. ***Resonant absorbers*** (such as wood panels) are set into vibration by the energy released by the sound source and respond to low-frequency sounds. By conducting numerous tests of the sound absorption of different materials, Sabine determined that overall ***absorbing power***, the key to reverberation times, varies not only with the material reflecting sound, but with the frequency (pitch) of the instrument. For example, the absorbing power of a given surface for the higher register of the violin is nearly double that of the same surface for the lower register of the double bass (Sabine, 1922).

Audiences also absorb sound; imagine the heavy fabrics of formally attired concert audiences of the 1700s and 1800s! Audience absorption is a tricky variable, as one cannot always predict the number of occupants who will attend a performance, nor the exact arrangement of occupants filling the available seats. The audience area is larger if the audience is seated on a slope, and thus absorption is affected not only by the size of the crowd, but by the dimensions and plane of the seating space. The seats themselves also absorb a lot of sound, which is one reason why back rests in most concert hall seats are usually low (ending below the shoulder) and chairs are not as plush as seats in a movie theater. Chairs with lightly cushioned seats and wooden arm rests reflect more sound and absorb less high-frequency sounds, so the delicate tones of chamber music are more intense and brilliant. On the other hand, heavily upholstered chairs can keep the music from sounding too loud in small performance venues by absorbing more sound (Beranek, 2015).

Reverberation time

Sabine was a pioneer in identifying reverberation time as a critical factor in indoor acoustics. ***Sabine's formula*** (i.e., his original equation for calculating reverberation time, *RT*) is written as:

$$RT = 0.161\mathrm{V}/\mathrm{A}$$

where *V* is the volume of the room in cubic meters, and *A* is the total absorbing power in sabins. Sabine's measure of reverberation time was the time it takes a sound to decay until it is barely audible, as sensitive methods to determine sound levels had not yet been devised. He made painstaking systematic measurements of the time it took sound to decay, starting from 1,000,000 times the first audible sound level to silence (Sabine, 1922, pp. 60–68). Today, ***reverberation time (RT)*** is defined as the length of time for a sound to decay by 60 decibels. Although alternate measures such as ***early decay time*** (***EDT***, based only on the initial part of the decay) have since been devised, RT is still regarded as one of the most important basic acoustic parameters.

Performance halls range from reverberant to dry – that is, from those in which there are many complex reflections to those with absorbent walls and ceilings, which reduce the loudness and longevity of these reflections. From the perspective of performers, two important concerns are ***ease of ensemble*** (the degree to which performers can hear themselves and others playing together) and ***support*** (the degree to which the room facilitates the musicians' efforts to create tones and fill up the space). Early sound reflections in the stage area are critical for ease of ensemble, while both early and late reflections are important for good support. When performing in a 'dry' space, it is difficult for instrumentalists and vocalists to feel like they are 'filling up' the room with music, as the sound disappears quickly. They may then sing or play with more force, and the tone quality may be compromised (Gade, 2015).

Instrumental and vocal ensembles must also give more precise performances in a room with a short reverberation time, as asynchronies are more discernible. A choir singing lyrics such as 'to tell a tale of tragedy' must really stay together to sound crisp in a dry space! On the other hand, long reverberation times in large spaces may make it difficult for performers to hear themselves and each other, as ***horizontal clarity*** (the distinctiveness of tones played successively) and ***vertical clarity*** (the distinctiveness of tones played simultaneously) are diminished. Indeed, long reverberation times have been found to affect the tempo of performance for choral singers, as they tend to sing more slowly and with less precise timing in reverberant rooms, although intonation (pitch accuracy) may not be affected (Fischinger, Frieler, & Louhivuori, 2015).

The highest-rated concert halls in Beranek's (2004, 2007) studies have a reverberation time of 1.8 to 2.0 seconds, which seems ideal for orchestral music, while a shorter RT of 1.24 to 1.6 seconds seems to be ideal for the best opera houses. Venues designed for speech (such as oration, and plays) generally require the shortest RTs of 0.7 to <1.0 second, as it is critical for the audience to perceive rapidly unfolding speech sounds. In listening to music, however, the *loss* of some sounds (the rasping of the cello bow or the hiss of air expelled from a clarinet) due to longer reverberation times may actually enhance the listener's appreciation of the music. Further, some degree of blending together of the sounds of an orchestra or choir is essential to a music performance. Auditoriums with modifiable size and flexible reverberation times (e.g., with chambers that can be opened or sealed, or removable panels) are practical designs for performance venues that must accommodate speech, as well as vocal and instrumental performances.

Optimal reverberation times vary for different instruments and vocal ranges, and musical repertoire. Even within the same 'family' of instruments, ideal reverberation times differ somewhat; pianists prefer halls with shorter reverberation times, whereas pipe organists prefer more reverberant spaces (Veneklasen, 1975). La Scala in Milan has an RT of 1.2 seconds, while Wagner's Festspielhaus in Bayreuth has a much longer RT of 2.2 seconds. It is not hard to see why this difference should emerge, considering the characteristics of Italian and German opera! One might compare the lucidity of Verdi's operas and his focus on the singers and text with the lush, expansive orchestration of Wagner's operatic works.

Very large auditoriums are rarely effective for music performance. In his study of architectural structures for the performance of music from the 17th century to the present, Forsyth (1985) notes that the Royal Albert Hall in London (built in 1871) was *10 times* larger in volume than most concert halls of its time, and attributes many acoustical problems it encountered through the years to its immense proportions. In very large modern spaces such as indoor arenas which may house audiences as large as 10,000 or 15,000 or more for 'pop' and rock concerts, performers must rely on electronic amplifiers and loudspeakers to carry and disperse sound, and low-frequency reverberation is particularly difficult to control. Generally, low-frequency absorbers are not as efficient as those that absorb high frequencies, which is why audiences in such venues often find themselves in a wash of persistent low-frequency sounds.

Bidirectional influence: Music and architecture

Our discussion has focused on how considerations about music performance have shaped the way we design and construct buildings, but the effects may have been *bidirectional.*

Sabine once argued that the acoustics of a sound space are so critical to effective performances that the architectural traditions of different eras may have fundamentally shaped the development of music (cited in Forsyth, 1985). The cavernous, highly resonant stone buildings of the Romanesque period allowed vocal tones to linger, supporting the exploration of rich vocal harmonies characteristic of choral music of that time. As the classical outdoor amphitheater evolved into the roofed horseshoe-shaped concert building, the improved horizontal clarity may have lent itself to the development of ornate contrapuntal Baroque music with its complex interplay of melodies (Forsyth, 1985). An analysis of symphonic scores suggests that some classical composers took the effects of reverberation in performance venues into consideration in their composition and orchestration (Meyer, 2015). It is intriguing to consider the possible *mutual influence* of the construction of musical buildings and the construction of musical works.

Coda

At first glance, the science of acoustics does not seem very daunting. Wave trains in the air and the overtones of vibrating strings are not difficult concepts to master. Acousticians have succeeded in extracting many listener-relevant properties of such sound fields, so that we can form fairly clear ideas of how musical experience is related to the physical properties of the field. The complexity of sound fields is directly related to the structure and properties of the bounded spaces in which music is played and heard. But here, mathematical analysis and ascertainable principles begin to part company. Using the broadly defined acoustical concepts described in this chapter, it is possible to draw some correspondences between architecture and audience. However, one of the lessons to be learned from the research on the acoustics of performance venues is how difficult it has turned out to be to design a 'perfect' acoustic environment from the principles of the science of acoustics! There is much about the experience of performing and listening to music that cannot be captured by physical laws alone. It is to the psychological and more subjective qualities of musical sound, and the pathway from the source of a sound to the ear and auditory cortex, that we turn our focus in our next chapter.

3 Auditory perception and the neurophysiology of hearing

Passport photographs for many countries require both ears to be clearly visible. This is because the ears, though not quite as unique as fingerprints or snowflakes, are distinctive features of a person's appearance. If you were to gather a dozen of your friends and peer closely at their ears, you would be likely to find great variability in the shape, size, and particularly in the convolutions or patterns of ridges of their ears! In fact, outer ears are so different with respect to their shapes and irregularities that if your outer ears were switched with someone else's, you would probably have trouble accurately identifying exactly where sounds were coming from in the environment around you.

This chapter focuses on the workings of the marvelous ear, to reveal the mechanisms in the auditory system that enable people to hear the musically salient aspects of sound. First, we begin with an examination of pitch, loudness, duration, and timbre as the four dominant perceptual properties of the heard sounds from which music emerges as an auditory experience. In the second section, we describe the main structures of the ear and the pathways leading from the ear to the brain. How do the ears and the relevant parts of the nervous system extract these features of sound? At the end of the chapter, we conclude with a brief discussion on the topic of cochlear implants and implications for the perception of music.

Perceptual qualities of sound

Chapter 2 described how a sound wave is propagated, and explained that sound waves can be characterized by their ***physical*** properties: *frequency*, *amplitude*, and the *power spectrum*. As discussed briefly in that chapter, these physical properties of a sound wave correspond to the ***perceptual*** (i.e., subjective) qualities: *pitch*, *loudness*, and *timbre* respectively, as shown in Table 3.1. In addition, there is also the *duration* of the acoustic or auditory event, as sounds must unfold in time.

The distinction between physical stimuli and our perceptual experience of those stimuli is an important one in psychology. Our brains do not reconstruct all the physical properties of incoming stimuli. If they did, we would see X-rays, and hear radio waves, and the result

Table 3.1 Physical and perceptual properties of a tone.

Physical properties of sound waves	*Perceptual properties of tone*
Frequency (hertz)	Pitch
Amplitude (decibels)	Loudness
Power spectrum	Timbre

would likely be confusion. Instead, our brains have adapted to process those aspects of the physical world that benefit our survival, and they do so in a way that highlights those features that matter. In this chapter, we will discuss some basic ways that auditory perception involves a transformation of the physical sound wave described in chapter 2. The chapters that follow further explore how our brains reframe the experience of sound in a musical context.

Pitch

The perceptual experience of ***pitch*** is related to the ***frequency*** of vibrations in sounds. The ability to perceive pitch thus relies on the ear's ability to encode frequencies from physical stimuli. The pitches that most human beings can detect range from about 20 to 20,000 Hz (based on the unit ***hertz***, which refers to *the frequency of a wave expressed in cycles or oscillations per second*). This capacity varies widely from person to person and declines with age. Starting earlier, though often not really noticeable until about age 60, ***presbycusis*** (i.e., hearing loss associated with aging) results in a loss of sensitivity especially to high-frequency sounds.

While the range of intact human hearing is about 20 to 20,000 Hz, most musical pitches fall within the range of approximately *20 to 4000 Hz*. For instance, the fundamental frequencies of a grand piano with 88 keys range from 27.5 Hz to 4186 Hz. Towards the extreme ends of the spectrum of musical tones, the lowest pipe in a pipe organ is about 16 Hz (and is often felt more as a rumble or vibration than a tone), and the highest note of the piccolo may reach around 4500 Hz, producing a shrill sound that is uncomfortable to some listeners. However, the threshold of human sensitivity to pitch extends far beyond the piccolo's top register, enabling us to also appreciate some of the high overtones in its timbre. To consider the highest pitch that a college student with intact hearing can detect (20,000 Hz), imagine a grand piano extended by a little over two more octaves of keys. Sadly, such pristine hearing among young people is not common these days, a point we will return to later.

In the previous chapter, we discussed how most musical sounds come from a complex pattern of vibration that can be broken down into many component frequencies. Even a single musical tone, created by depressing one key on a piano, includes many different frequencies because of the many mathematically related ***overtones*** (or ***harmonics***) that also sound (at different relative loudness for different instruments, as described in chapter 2). Yet the perceived quality is of one single pitch. Relatedly, one of the mysteries of auditory perception, directly relevant to music, is that we typically hear one tone as having one pitch, but can still make out distinct pitches when multiple tones form a cluster, for instance when a pianist plays a chord. Much research has been devoted to understanding how the auditory system extracts a single pitch from a complex combination of frequencies – the reverse of Fourier analysis (described in chapter 2) – leading to theories of pitch processing in the auditory system. Such theories have largely been motivated by the physiology of the cochlea. We will therefore discuss them later in this chapter, accompanying a description of how the cochlea transduces the sound wave.

The simplest account of pitch processing, which works in many but not all cases, is to say that perceived pitch is linked to the ***fundamental*** (i.e., the lowest frequency in a complex tone), while the higher overtones contribute to the timbre of that tone (see chapter 2). However, this generalization does not always apply. Consider listening to someone singing a song on the telephone. The frequency range of most telephones is limited, and as such does not go low enough for many fundamental frequencies of low voices. Yet a listener will have the impression of hearing a voice at about the same pitch (though not the same timbre) as if the singer were physically present. This experience is connected with a curious effect, the phenomenon of ***residue pitch*** or ***virtual pitch***: that is, under some conditions, a missing fundamental will seem to be heard solely on the basis of hearing its overtones (Schouten, 1940; Terhardt, 1974). The same phenomenon is at play when listening to a pocket-size transistor radio. The fundamental frequencies of the lower tones are often not audible, and yet the proper pitches are usually heard as the fundamental is 'inferred' from the overtones.

Consonance and dissonance

Another important concept related to pitch perception, motivated by physiology but grounded in behavioral studies, is the ***critical bandwidth*** (or ***band***). The critical band refers to a range of frequencies that evoke a similar response in the auditory system, and bears on the degree to which our auditory system responds selectively to different frequencies. Frequency is a continuous variable, with an infinite number of possible values, and so it would be inefficient for our auditory system to respond selectively to every possible value. So instead, it seems that our auditory system responds similarly (e.g., with a similar cochlear response) to frequencies in close proximity to each other, and frequencies that evoke a similar response are said to fit within the same critical band. In some cases, this 'shortcut' of the auditory system can influence our perception of music. The most prominent example is in the case of musical dissonance.

The terms ***consonance*** and ***dissonance*** refer to a basic and important subjective continuum in music – reflecting the degree to which tones that are played simultaneously (called a *harmonic interval*) sound pleasing and relaxed, or displeasing (Helmholtz, 1863). In a highly influential paper, Plomp and Levelt (1965) demonstrated that the degree of perceived dissonance for intervals formed by complex tones may come from interactions among the upper overtones. Specifically, they found that perceived dissonance is maximal when there are overtones that fit *within the same critical band*, while not being identical in frequency. The greatest dissonance is associated with approximately 25% of the bandwidth, which happens for highly dissonant intervals like the tritone (e.g., C and F#). In contrast, the overtones for highly consonant intervals (e.g., C and G, a perfect fifth) are either identical in frequency or do not fit within the same critical band.

Thus, one of the key insights imparted by Plomp and Levelt's foundational work is that in order to understand consonance and dissonance, we cannot merely consider the relationship between the two fundamental frequencies that make up an interval – but must look further to the relationships between their overtones. When the fundamentals of the tones, or their constituent overtones, are so close together that they fit within the same critical band, they pose a challenge for the auditory system to reconcile; this tends to produce a sense of roughness that we may perceive as 'dissonance.'

Although Plomp and Levelt's contribution offers a good starting point for the understanding of how we experience consonance and dissonance, more recent work suggests familiarity may play a role that is at least as strong – or perhaps even stronger – than the role of critical bands. Whereas the critical band is a common feature of all human auditory systems, McDermott, Lehr, and Oxenham (2010) found large individual differences in the degree to which this construct predicts the experience of consonance across individuals. Perhaps most telling was the fact that musically trained individuals perceived consonance or dissonance based on whether two pitch categories – and not necessarily their overtones – are commonly considered to be consonant in musical practice. That is, a musician tends to hear a C-major chord (consisting of C, E, and G) as 'consonant' even if the component frequencies of each note are re-organized to create roughness. By contrast, other listeners (mainly nonmusicians) were more strongly influenced by the composition of overtones, and would not always hear C major as 'consonant.'

Such findings introduce the possibility that familiarity with musical rule systems influences our perception of consonance, perhaps even more strongly than purely physical factors (McLachlan, Marco, Light, & Wilson, 2013). More exciting new work on crosscultural differences in consonance perception is discussed in chapter 15, again suggesting that our knowledge adds a lot to the basic processing discovered by Plomp and Levelt.

Loudness

It may be surprising to learn just how sensitive the auditory system is to differences in the amount of energy that is transferred by a sound wave (i.e., in the ***amplitude*** of the physical sound wave). In fact, if the ear were more sensitive than it is, we would actually perceive the collisions among air molecules! How loud a sound is perceived to be matters a good deal in music, as in other contexts in which what we hear is of significance. The more energetic the pressure wave – that is, the greater its amplitude – the louder the sound heard will be as the wave impacts on the ear.

The perceived loudness of a sound is based on its ***intensity***. The most common way to represent different levels of intensity is through the decibel scale, which is based on a formal mathematical definition. Specifically: *Intensity in decibels (dB) of a sound is defined as 10 × the logarithm of ratio of the energy of that sound to a measure of static sound pressure*. Thus, if we give sound pressure (usually estimated as having an energy of 10^{-12} watts per centimeter squared) a value of 1, the decibel measure is 10 times log 1/1, that is 10 × log 1. Log 1 is 0 dB. The faintest discernible sounds approach this measure. Suppose we double the energy of the impinging sound. The measure will now be 10 × log 2/1, that is log 2. Log 2 is 0.3, so the decibel measure is 3 dB. If we double the energy again, leading to a measure of 10 × log 4/1, we get 6 dB, and 10 × log 8/1 gives 9 dB.

Although the decibel scale better approximates perceived loudness than would the raw energy ratio (Moore, 2012), it is important to note that the decibel scale measures the *physical signal*, not a *perceptual experience*. One's experience of loudness is only roughly proportional to the energy of the impinging sound wave (e.g., perceived loudness approximately doubles for every 10 dB, but only for moderate intensities; Hellman, 1976). In some instances, the physical measure of intensity in decibels does not correlate perfectly with perceived loudness in real-world listening conditions. This is because how loud a sound is perceived to be depends not only on the amplitude of the wave, but also on its frequency. For this reason, ***A-weighted decibels (dBA)*** are typically used to express relative loudness as perceived by the human ear; these are based on weighting curves that accommodate for how the ear is less sensitive to sounds at very high and very low frequencies, for instance. (For a discussion of other loudness scales such as *phon* and *sone*, see Rasch & Plomp [1999].)

To give some reference to 'real-world' sounds in decibels (hereafter assumed to be A-weighted): A whisper at a distance of 3 feet away is about 10 to 15 dB, the ambient noise inside a restaurant is about 45 to 55 dB, conversation at a distance of 3 feet is approximately 65 dB, and a hammer hitting a steel plate at about 2 feet away is 114 dB (Pierce, 1992, pp. 119–120). The human auditory apparatus allows people to discriminate loudness differences of approximately 1 dB for frequencies between 200 and 4000 Hz (Yost, 2000). Bats, dogs, and many other animals have much more sensitive auditory systems.

Application: Music-induced hearing loss

Many countries set regulations for noise exposure in the workplace. For instance, the Occupational Safety and Health Administration (OSHA, 1983) in the US recommends a

permissible exposure limit (PEL) of no more than 90 dB over an eight-hour period for working environments, as prolonged exposure to loud noise has been shown to lead to hearing loss. Orchestral performances have been measured at sound levels of up to 112 dB (Sataloff, 1991), and vocal performances during opera can reach 105 dB for voices alone without accompanying instruments (Laitinen, Toppila, Olkinuora, & Kuisma, 2003). Even the average sound level of a single instrument or solo voice music produced by a music student in a small practice room has been recorded at 87 to 95 dB (Phillips & Mace, 2008). Thus, it is not surprising that some studies have found the incidence of hearing loss to be greater among musicians than nonmusicians. For instance, Axelsson and Lindgren (1981) found that 42% of adult orchestral musicians exhibit greater hearing loss than expected for their ages. Further, 45% of undergraduate music students have detectable hearing loss, suggesting that the damage may already be manifested during the college years (Phillips, Henrich, & Mace, 2010).

The phrase ***music-induced hearing loss (MIHL)*** refers to loss primarily caused by engagement in playing music, listening to music, or to exposure to music in the environment such as in a club. As it is impossible to separate this from exposure to other sounds in musicians' lives, it is also more broadly referred to as a subset of ***noise-induced hearing loss (NIHL)***. MIHL may be instrument-specific, as some studies have found greater hearing loss in the left ear for violinists and in the right ear for flute players (e.g., Axelsson & Lindgren, 1981; Ostri, Eller, Dahlin, & Skylv, 1989) though other studies have not shown clear patterns corresponding to instruments (e.g., Phillips et al., 2010). Those seated in front of loud sections of the orchestra (especially brass or percussion) may also be at greater risk for developing MIHL. Rock and pop musicians often suffer more severe damage in the ear closest to the source of sound, such as a loudspeaker or amplifier (Sataloff, 1991).

At high intensity, even brief exposure to very loud music can lead to permanent damage to one's hearing. For instance, Park (2003) found maximum noise levels in *karaoke* singing environments in Korea to exceed 95 to 115 dB, and found up to 8 dB of significant hearing loss (centered at around 4000 Hz) in listeners after less than two hours of exposure! Unlike hearing loss due to aging (presbycusis), which tends to affect sensitivity to higher frequencies, hearing loss due to exposure to loud music affects sensitivity to music, as it begins much closer to the region of frequencies within the musical range. Typically, music-induced hearing loss is found at around 4000 Hz or 6000 Hz (Phillips et al., 2010). As musicians pride themselves on their ability to hear tonal nuances, this presents a significant loss to both production and reception of music. For instance, even tones in mid-register may lose their fullness or richness, as the high overtones would be reduced or lost.

A study showed that early noise exposure in mice produces pathological changes in the auditory system that make the inner ear more susceptible to the effects of aging (Kujawa & Liberman, 2006). If this also holds for humans, this finding suggests that early exposure to loud noise (including music) could predispose a person for more deleterious effects of presbycusis later in life.

Personal listening devices and MIHL

A growing body of research focuses on the risks involved in the widespread use of portable listening devices (PLDs) such as MP3 players and iPods, due to their high output levels and loud listening habits reported by users. For instance, about one-third of college students reported occasionally listening to MP3 players at 100% listening level (Hoover & Krishnamurti, 2010), which has been measured at over 125 dB on some PLDs (Breinbauer et al., 2012).

PLD users are at greater risk for developing irreversible hearing loss, especially when using devices in environments in which the background noise is louder than 65 dB (Jiang, Zhao, Guderly, & Manchaiah, 2016), because listeners tend to turn the music up to compensate for the ambient noise. For instance, commuters in the New York City transit system listened to their PLDs at a maximum of 106 dB when standing on subway platforms and 112 dB while riding inside subway cars, in order to hear the music (Gershon, Neitzel, Barrera, & Akram, 2006).

Most forms of noise-induced hearing loss are permanent, due to irreparable damage to hair cells in the inner ear (as discussed later in this chapter). Indeed, studies show evidence of irreversible hearing loss manifested in young people who use PLDs. For instance, an audiometric study found diminished hearing in the 3000 to 8000 Hz range in university students who had used PLDs for 1 to 5 years, compared to peers who did not use PLDs (Peng, Tao, & Huang, 2007). Earbud users tend to experience more damaging effects than those using earphones that fit over the ear, as studies show that earbuds allow more ambient noise to leak through, so users' listening levels tend to be significantly higher (e.g., Hodgetts, Rieger, & Szarko, 2007).

Aside from limiting one's use of PLDs and monitoring loudness levels of devices, using earphones that isolate ambient noise can be helpful for volume control (Portnuff, 2016). Knox (2009) also found that adding a simple visual feedback accessory that helps listeners monitor their listening levels on PLDs was effective in reducing listening levels. For general environmental noise or when attending musical events, ear plugs are an inexpensive method for preventing MIHL. More expensive ear plugs can be custom-made for orchestra musicians, for instance to allow low-frequency sounds of a double bass to be heard while blocking higher-frequency sounds of other instruments close by. More public education on MIHL is needed via entertainment channels, health care professionals, college courses, and through corporations that design and manufacture PLDs, headphones, and earbuds.

Duration

Musical tones also differ in duration. Such differences are represented in musical notation by the symbols for whole note, half note, quarter note, and so on (semibreve, minim, crotchet, quaver, etc.), each representing a tone of one half the duration of the preceding note value. However, these are simply the subdivisions of note durations that conveniently organize tones in most Western music. In practice, musical performance involves a lot of finer-grained variations in timing, and our ability to sense these small variations in timing is important to our ability to perceive expressivity in music. (Expressive timing is discussed in greater detail in chapter 11 on performance.)

The duration of the acoustic signal may not solely determine the perceived length of a tone. Schutz and Lipscomb (2007) asked internationally acclaimed percussionist Michael Burritt to play a series of tones on a marimba – using either long graceful gestures or short clipped gestures to strike the keys. Their acoustic analysis revealed that the resulting tones were acoustically indistinguishable in duration. The long- and short-gesture tones were also indistinguishable to listeners hearing an audio recording of those tones. However, participants tended to perceive the sound of the tones to be significantly longer when accompanied by a video showing Burritt playing with long and graceful gestures, and shorter in duration when accompanied by a video showing short, choppy gestures.

Figure 3.1 shows time-lapse images of the videos used in this study, and the videos themselves may be viewed at www.maplelab.net/illusion.[1] Notice that the 'long' stroke traces

Figure 3.1 Time-lapse images of the videos used in Schutz and Lipscomb's (2007) study, which may be viewed at www.maplelab.net/illusion. Captured 200 milliseconds apart, they show that acclaimed percussionist Michael Burritt's striking mallet (held in his right hand) continues moving for a longer time after impact for the 'long' stroke (top row) versus the 'short' stroke (bottom row). Although the difference in post-impact motion does not affect the tone's acoustic duration, it does affect our *perception* of the duration of the tone.

Source: Copyright © Michael Schutz and Scott Lipscomb, printed with permission of Michael Burritt.

a smooth arc after impact, whereas the 'short' choppy stroke stops abruptly. Schutz and Lipscomb concluded that the 'difference' participants 'heard' between tones created by long-slow and short-swift strokes was perceptual rather than physical, caused by 'visual artifacts of the performer's acoustically inconsequential gesture' (2007, p. 896). There is some similarity to the 'McGurk effect,' which can be viewed at https://vimeo.com/200587571,[2] in which watching the lip movements of a speaker alters the listener's perception of the syllable being spoken (McGurk & MacDonald, 1976). In a subsequent study, Schutz and Kubovy (2009) showed that this illusion persisted even when the video of Burritt was reduced to a point-light display consisting of only a single white dot tracking the moving mallet head against a black background, as can be seen at www.maplelab.net/virtualmarimbist. This is just one illustration of how the pure physical signal and the perception of a sound may differ. Other examples, pertaining to pitch perception, will be discussed in chapter 5.

Our perception of tone durations is also circumscribed by the physical limitations of the auditory system. For example, clicks are perceived as if occurring simultaneously if separated by less than 2 milliseconds, as shown in pioneering research by Exner (1875). Similarly, there is a threshold of order: Sound events are heard as separate, but their temporal order is not discernible for separations less than 15 to 20 ms (Hirsch, 1959). Sounds of very short duration may not even be perceived as a tone. For instance, a very short tone burst (2 to 4 cycles in length) is usually heard as a click without a discernible pitch (Pierce, 1992). The higher in pitch the sound is, the longer the tone burst needs to be if there is to be no click. As we shall see in the following section, there is also a temporal (i.e., time-based) dimension to the perception of timbre.

Timbre

Timbre or 'tone color' is the intrinsic and distinctive quality of the sound produced by a musical source such as a certain instrument or a particular human voice. The brilliant sound of the trumpet contrasts in timbre with the mellow, reedy sound of the clarinet. Timbre perception is a very complex topic that poses many methodological challenges to researchers (see Hajda, Kendall, Carterette, & Harshberger, 1997). In fact, the definition of timbre given by the American National Standards Institute is based on timbre being *any* quality of a tone other than its pitch, loudness and duration! For the purposes of this brief discussion, we focus on two aspects: *frequency relationships* (also discussed in chapter 2) and *changes across time*.

Frequency relationships

The quality we refer to as 'timbre' corresponds with the ***Fourier analysis*** of the waveform, which reflects the various frequencies it contains. A pure tone, such as the ringing of a tuning fork, is produced by a sound wave of a single frequency of vibration. However, as mentioned earlier, such a smooth wave would be more characteristic of the sound made by a well-crafted tuning fork, as opposed to the more complex sound of an acoustic instrument. Each tone produced by most musical instruments comprises not only the fundamental frequency, but also its overtone series (as shown in Figure 2.4 in the previous chapter). The particular combination of overtones sounded by an instrument is responsible, in part, for its distinctive tone quality or 'timbre,' and provides important cues for the accurate identification of the sound of musical instruments by both humans and machines (Brown, Houix, & McAdams, 2001; Martin & Kim, 1998).