Auditory Spectral Processing

EDITED BY

MANUEL S. MALMIERCA
DEXTER R. F. IRVINE

International Review of Neurobiology, Volume 70

Auditory Spectral Processing

EDITED BY

MANUEL S. MALMIERCA
Laboratory for the Neurobiology of Hearing
The Institute of Neuroscience of 'Castilla y Leon' (INCyL)
Faculty of Medicine, University of Salamanca
Campus 'Miguel de Unamuno'
37007 Salamanca, Spain

DEXTER R. F. IRVINE
Monash University
Department of Psychology
School of Psychology, Psychiatry, and Psychological Medicine
Faculty of Medicine, Nursing, and Health Sciences
Vic 3800, Australia

ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO
Academic Press is an imprint of Elsevier

AP

Elsevier Academic Press
525 B Street, Suite 1900, San Diego, California 92101-4495, USA
84 Theobald's Road, London WC1X 8RR, UK

This book is printed on acid-free paper. ∞

For all information on all Elsevier Academic Press publications visit our Web site at www.books.elsevier.com

ISBN-13: 978-0-1236-6871-4
ISBN-10: 0-12-366871-9

PRINTED IN THE UNITED STATES OF AMERICA
05 06 07 08 09 9 8 7 6 5 4 3 2 1

International REVIEW OF Neurobiology

Volume 70

International REVIEW OF *Neurobiology*

Volume 70

CONTRIBUTORS

Numbers in parentheses indicate the pages on which the authors' contributions begin.

Emily Buss (87), Department of Otolaryngology/Head & Neck Surgery, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA

Simon Carlile (399), Auditory Neuroscience Laboratory, School of Medical Sciences, University of Sydney, Sydney NSW 2006, Australia

Kevin A. Davis (169), Departments of Biomedical Engineering & Neurobiology and Anatomy & Center for Navigation and Communication Sciences, University of Rochester, Rochester, New York 14642, USA

Monty A. Escabí (207), Department of Electrical Engineering, University of Connecticut, Storrs, Connecticut 06269, USA

John H. Grose (87), Department of Otolaryngology/Head & Neck Surgery, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA

Deborah A. Hall (331), MRC Institute of Hearing Research, Nottingham NG7 2RD, United Kingdom

Joseph W. Hall III (87), Department of Otolaryngology/Head & Neck Surgery, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA

Dexter R. F. Irvine (1, 435), Department of Psychology, Faculty of Medicine, Nursing, and Health Sciences, Monash University, Victoria 3800, Australia

Enrique A. Lopez-Poveda (7), Instituto de Neurociencias de Castilla y León, Universidad de Salamanca, Salamanca 37007, Spain

Manuel S. Malmierca (1, 299), Laboratory for the Neurobiology of Hearing, The Institute of Neuroscience of 'Castilla y Leon' and Faculty of Medicine, University of Salamanca, Salamanca 37007, Spain

Russell Martin (399), Defense Science and Technology Organisation, Fisherman's Bend, Victoria 3207, Australia

Ken McAnally (399), Defense Science and Technology Organisation, Fisherman's Bend, Victoria 3207, Australia

Colette M. McKay (473), School of Life and Health Sciences, Aston University, Birmingham B4 7ET, United Kingdom

Brian C. J. Moore (49), Department of Experimental Psychology, University of Cambridge, Cambridge CB2 3EB, United Kingdom

Heather L. Read (207), Department of Psychology, The University of Connecticut, Storrs, Connecticut 06269, USA

Adrian Rees (299), School of Neurology, Neurobiology, and Psychiatry, University of Newcastle, Newcastle upon Tyne NE2 4HH, United Kingdom

Lina A. J. Reiss (135), Department of Speech Pathology and Audiology, University of Iowa, Iowa City, Iowa 52242, USA

Robert V. Shannon (121), House Ear Institute, Los Angeles, California 90057, USA

Donal G. Sinex (371), Department of Psychology, Utah State University, Logan, Utah 84322, USA

Mitchell L. Sutter (253), Center for Neuroscience and Section of Neurobiology, Physiology, and Behavior, University of California Davis, Davis, California 95616, USA

Beverly A. Wright (435), Department of Communication Sciences and Disorders, Northwestern University, Evanston, Illinois 60208, USA

Eric D. Young (135), Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland 21205, USA

Jane J. Yu (135), Department of Biomedical Engineering, Center for Hearing Sciences, Johns Hopkins University, Baltimore, Maryland 21205, USA

CONTENTS

Across-Channel Spectral Processing

JOHN H. GROSE, JOSEPH W. HALL III, AND EMILY BUSS

Speech and Music Have Different Requirements for Spectral Resolution

ROBERT V. SHANNON

Non-Linearities and the Representation of Auditory Spectra

ERIC D. YOUNG, JANE J. YU, AND LINA A. J. REISS

Spectral Processing in the Inferior Colliculus

KEVIN A. DAVIS

Neural Mechanisms for Spectral Analysis in the Auditory Midbrain, Thalamus, and Cortex

MONTY A. ESCABÍ AND HEATHER L. READ

Spectral Processing in the Auditory Cortex

MITCHELL L. SUTTER

Processing of Dynamic Spectral Properties of Sounds

ADRIAN REES AND MANUEL S. MALMIERCA

Representations of Spectral Coding in the Human Brain

DEBORAH A. HALL

Spectral Processing and Sound Source Determination

DONAL G. SINEX

Spectral Information in Sound Localization

SIMON CARLILE, RUSSELL MARTIN, AND KEN MCANALLY

Plasticity of Spectral Processing

DEXTER R. F. IRVINE AND BEVERLY A. WRIGHT

Spectral Processing in Cochlear Implants

COLETTE M. MCKAY

AUDITORY SPECTRAL PROCESSING: AN OVERVIEW

Dexter R. F. Irvine* and Manuel S. Malmierca†

*Department of Psychology, Faculty of Medicine, Nursing, and Health Sciences
Monash University, Victoria 3800, Australia
†Laboratory for the Neurobiology of Hearing, The Institute of Neuroscience of 'Castilla y Leon'
and Faculty of Medicine, University of Salamanca,
Salamanca 37007, Spain

Almost all natural sounds and sounds produced by man-made devices (such as musical instruments) are composed of multiple frequencies and their frequency composition varies over short time periods, i.e., they have complex and time-varying frequency spectra. Although a great deal of information can be derived from auditory signals in which the spectral information is considerably degraded (such as those transmitted by most telephones), it is nevertheless the case that a full perceptual appreciation of most complex sounds depends on the availability of a non-degraded representation of the spectral detail. Examples of such perceptual experiences are the appreciation of music and the extraction of so-called indexical information from speech, i.e., information about the speaker's age, gender, mood, and so on (e.g., Nygaard and Pisoni, 1995). Detailed spectral information has also been shown to be critically important in the processes of identifying the discrete auditory objects comprising complex auditory scenes, in which the sound waves generated by multiple sources sum together to form a single complex pressure wave at each eardrum, and of sound localization. The importance of detailed spectral information in these important aspects of human and animal hearing indicates that this information must be encoded in the peripheral auditory system and then made available to and processed in the central auditory brain pathways and associated structures that ultimately give rise to our auditory perceptual experience. It is the aim of this volume to provide a comprehensive overview of the neural mechanisms of auditory spectral processing and of the ways in which spectral information is used by human listeners to generate their auditory perception of the world around them.

Each point on the basilar membrane (BM) in the cochlea has a characteristic frequency (CF) at which it responds maximally, and the BM carries out a real-time spectral analysis of the pressure wave delivered to the inner ear by the outer and middle ears. The filtering properties of the BM depend partly on its passive mechanical properties, first described in the pioneering studies of von Békésy (1960), and partly on an active process of amplification (produced by forces generated by the outer hair cells [OHCs]) which feeds energy back into the cochlea at sound pressure levels (SPLs) close to threshold (Robles and Ruggero,

INTERNATIONAL REVIEW OF
NEUROBIOLOGY, VOL. 70
DOI: 10.1016/S0074-7742(05)70000-3

0074-7742/05 $35.00

2001). As a consequence of these processes, each inner hair cell (IHC) is sharply tuned to the CF of the point at the BM at which it is located, and each of the auditory nerve (AN) fibers that innervate that hair cell has the same CF and sharp frequency tuning as the hair cell from which its input is derived (Patuzzi and Robertson, 1988). One way in which spectral information is conveyed to the central nervous system (CNS), therefore, is in the distribution of activity across the AN fiber array. At frequencies below 3 to 4 kHz, the discharge of AN fibers is phase-locked to the stimulus waveform, with the consequence that information about spectral components in this frequency range is also encoded in the temporal pattern of the discharge of AN fibers. These peripheral mechanisms, and models of the processes involved, are discussed in detail by Lopez-Poveda in Chapter 1.

The mechanical and neural filtering that occurs in the peripheral auditory system underlies the frequency resolution ability exhibited by human and animal listeners. Psychophysical evidence on the frequency resolving ability of human listeners is reviewed by Moore in Chapter 2. As previously noted, however, most natural sounds have complex frequency spectra, and our perception of those sounds involves the integration of information in different frequency bands. This integration begins in the cochlear nucleus, where all of the AN fibers carrying information to the brain terminate, and is continued in the ascending auditory pathways that extend from the cochlear nucleus to the auditory cortex. In Chapter 3, Grose and colleagues discuss the major auditory phenomena that reflect across-channel frequency processing and consider the ways in which this processing can be understood. A number of the aspects of across-channel processing discussed in this chapter relate to the analysis of complex auditory scenes and the segregation of the separate auditory "streams" produced by different sound sources. The role of spectral information in this process is examined explicitly in Chapter 11 by Sinex, who also presents physiological evidence from the auditory midbrain nucleus, the inferior colliculus (IC) relating to some of the neural mechanisms involved. This chapter depends in part on some of the material presented in the chapters on the processing of spectral information at different levels of the auditory neuraxis and, therefore, it follows those chapters. The use of spectral information for sound localization similarly involves integration across frequency bands, and psychophysical and neurophysiological evidence on these processes is reviewed in Chapter 12 by Carlile and his colleagues.

The capacity of human listeners to perform complex perceptual tasks despite degraded spectral information was alluded to in the previous text, and is examined explicitly by Shannon in Chapter 4, in which he contrasts the limited amount of spectral information needed for the perception of speech with the detailed information required for the recognition and appreciation of music. The robustness of speech perception in the face of degraded spectral information reflects the fact that our auditory perceptual experience depends not only on the processing of spectral and other information contained in the acoustic signal but also the interpretation of this information in terms of stored information based on our auditory experience of the world (i.e., on what have been termed "bottom-up" and "top-down" processes). Although the emphasis of this book is on the former, all of the chapters dealing with the more complex aspects of auditory spectral processing point to the overwhelming importance of top-down processing.

Bottom-up processing is the focus of Chapters 5 to 9, in which the central processing of spectral information at successively "higher" levels of the ascending auditory pathway is examined: in the cochlear nucleus by Young and his colleagues (Chapter 5); in the midbrain by Davis (Chapter 6), in the thalamocortical system by Escabí and Read (Chapter 7), and in auditory cortical fields by Sutter (Chapter 8). The psychophysical evidence for integration of information across frequency bands reviewed by Grose *et al.* in Chapter 3 is complemented by neurophysiological evidence for such integration at all levels of the central auditory pathway. Much of this integration is highly non-linear, and in Chapter 5, Young and his colleagues examine in detail the non-linear integration in the circuitry in the dorsal cochlear nucleus, and the difficulties raised by this non-linearity for the task of measuring spectral sensitivity and predicting neural responses to complex stimuli. These issues are also examined in the context of the thalamocortical levels of the system in Chapters 7 and 8.

The spectral selectivity of individual auditory CNS neurons in animal preparations has been measured in a number of ways over the more than 60-year history of auditory single-unit neurophysiology. In what appears to have been the first measurement of neuronal frequency selectivity, Galambos and Davis (1943) defined what they called the "response areas" of AN fibers by determining the range of frequencies by which the fiber was excited at each of a range of SPLs. Similar procedures remain in wide use today: response amplitude (and other response characteristics) are measured over a wide range of frequency-SPL combinations, usually presented in pseudo-random order under computer control, to generate what are commonly called frequency response areas (FRAs). An alternative procedure is simply to measure threshold as a function of frequency (sometimes using an automated threshold-tracking algorithm) to generate a frequency tuning curve (FTC), which corresponds to the boundaries of the FRA. In the absence of spontaneous activity, neither of these methods is able to reveal inhibitory components of neuronal FRAs, a problem solved by the use of more complex two-tone stimuli. More recently, a range of methods using spectrally complex stimuli and reverse-correlation analytic techniques have been employed to generate spectro-temporal receptive fields (STRFs), which reveal the manner in which excitatory and inhibitory components of the response area vary over time. These methods and their applications, strengths, and weaknesses are discussed by Young and his colleagues in Chapter 5, by Escabí and Read in Chapter 7, and by Sutter in Chapter 8.

All of these methods are based on measurement of the action potentials generated by the neurons, and thus do not provide information on excitatory synaptic inputs that are below the threshold for action potential generation. Intracellular recordings from auditory cortical neurons have shown that, like neurons in other sensory cortices, they receive such subthreshold inputs in response to a range of stimuli (in this case, frequency-SPL combinations) outside the "classical" response areas defined in terms of the neurons' spike activity (e.g., Kaur *et al.*, 2004).

In Chapter 9, Rees and Malmierca review the ways in which neurons at different levels of the auditory pathway respond to frequency modulations. Dynamic changes in frequency, whether as modulations of single tones or clusters of harmonics, are important information bearing elements of natural sounds, and acoustically distinguishing attributes in the vocal repertoire of many species. Specificity for parameters like the direction of frequency change is apparent at all levels above the auditory nerve. At more central locations, there is evidence for increased specificity, topographic mapping of frequency modulation parameters, and differences in response properties between cortical subdivisions.

The single unit recording techniques which have provided such rich information about spectral processing mechanisms in animal models are not applicable to the human auditory CNS, except under exceptional conditions, such as when such recordings are made in association with surgery and electrical recording to identify epileptogenic foci (e.g., Brugge *et al.*, 2005). However, the study of human auditory cortical activity, and specifically of spectral processing in human auditory cortex, has been revolutionized in recent years by the development of a range of functional imaging techniques, culminating in functional magnetic resonance imaging (fMRI). These methods, and the data on spectral processing in human auditory cortex that they have yielded, are examined by Hall in Chapter 10. In contrast to the dependence of most animal electrophysiological recording techniques on neural spiking activity, the signals generated by imaging techniques such as fMRI reflect predominantly population synaptic activity.

The processing of spectral information occurs in auditory circuitry that until relatively recently was thought to be modifiable by experience early in development, but stable in adults. It has now been realized that sensory cortical and subcortical structures retain substantial capacity for plasticity in adult animals as a result of altered input consequent on either injury (e.g., Kaas and Florence, 2001) or various types of learning (e.g., Weinberger, 2004). Much of the evidence for such plasticity in the auditory system has been derived from psychophysical and electrophysiological studies of the processing of spectral information, and this evidence is reviewed by Irvine and Wright in Chapter 13.

Almost certainly the most interesting application of our knowledge of spectral processing is provided by the cochlear implant, which has now been used to restore functional hearing to tens of thousands of profoundly deaf people. The design of cochlear implants and the processing strategies utilized in them have been based on our knowledge of peripheral spectral processing, i.e., understanding of cochlear place and temporal coding mechanisms was applied to the task of electrically stimulating surviving auditory nerve fibers in the manner that was most effective in evoking normal auditory percepts. But the flow of information has been bi-directional: the perceptual experiences of implant users have also provided valuable information about basic mechanisms of spectral processing. Furthermore, the improvement in performance of implant users over the post-implantation period stands as perhaps the most dramatic practical example of plasticity in auditory processing. These issues are reviewed in detail by McKay in Chapter 14, and are also examined by Shannon in Chapter 4.

A final comment is required concerning the relationship between the chapters comprising this volume. It is our hope that the chapters together make up an integrated whole which gives a comprehensive overview of auditory spectral processing, although as previously noted more attention is given to bottom-up than to top-down processes. It is realized, however, that the interest of many readers will be focused on particular issues in the field. Therefore, each chapter has been designed to stand alone, even though this might involve some overlap in the content of various chapters, or even disagreement between authors as to the interpretation of particular effects or the value of certain techniques. The overlap is an unavoidable consequence of making the chapters stand-alone; the disagreement is a characteristic of any dynamic research area.

References

Brugge, J. F., Volkov, I. O., Reale, R. A., Garell, P. C., Kawasaki, H., Oya, H., Li, Q., and Howard, M. A. (2005). The posterolateral superior temporal auditory field in humans: Functional organization and connectivity. *In* "Auditory Cortex—A Synthesis of Human and Animal Research" (R. König, P. Heil, E. Budinger, and H. Scheich, Eds.), pp. 145–162. Erlbaum, Mahwah, N.J.

Galambos, R., and Davis, H. (1943). The response of single auditory-nerve fibers to acoustic stimulation. *J Neurophysiol* **6,** 39–58.

Kaas, J. H., and Florence, S. L. (2001). Reorganization of sensory and motor systems in adult mammals after injury. *In* "The Mutable Brain" (J. H. Kaas, Ed.), pp. 165–242. Harwood Academic Publishers, Amsterdam.

Kaur, S., Lazar, R., and Metherate, R. (2004). Intracortical pathways determine breadth of subthreshold frequency receptive fields in primary auditory cortex. *J Neurophysiol* **91,** 2551–2567.