Tone of Voice and Mind

Norman D. Cook



PROPERTY OF
MEMORIAL LIBRARY

Tone of Voice and Mind

Advances in Consciousness Research

Advances in Consciousness Research provides a forum for scholars from different scientific disciplines and fields of knowledge who study consciousness in its multifaceted aspects. Thus the Series will include (but not be limited to) the various areas of cognitive science, including cognitive psychology, linguistics, brain science and philosophy. The orientation of the Series is toward developing new interdisciplinary and integrative approaches for the investigation, description and theory of consciousness, as well as the practical consequences of this research for the individual and society.

Series B: Research in progress. Experimental, descriptive and clinical research in consciousness.

Editor

Maxim I. Stamenov
Bulgarian Academy of Sciences

Editorial Board

David Chalmers, *University of Arizona*
Gordon G. Globus, *University of California at Irvine*
Ray Jackendoff, *Brandeis University*
Christof Koch, *California Institute of Technology*
Stephen Kosslyn, *Harvard University*
Earl Mac Cormac, *Duke University*
George Mandler, *University of California at San Diego*
John R. Searle, *University of California at Berkeley*
Petra Stoerig, *Universität Düsseldorf*
† Francisco Varela, *C.R.E.A., Ecole Polytechnique, Paris*

Volume 47

Tone of Voice and Mind: The connections between intonation, emotion, cognition and consciousness
by Norman D. Cook

Tone of Voice and Mind

The connections between intonation, emotion, cognition and consciousness

Norman D. Cook

Kansai University

John Benjamins Publishing Company

Amsterdam / Philadelphia

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Norman D. Cook

Tone of Voice and Mind : The connections between intonation, emotion, cognition and consciousness / Norman D. Cook.

p. cm. (Advances in Consciousness Research, ISSN 1381–589X ; v. 47)

Includes bibliographical references and indexes.

1. Neuropsychology. 2. Conciousness. 3. Cognition. 4. Emotions. 5. Musical perception. I. Title. II. Series

QP360 .C667 2002

153-dc21 2002074689

ISBN 90 272 5173 8 (Eur.) / 1 58811 275 6 (US) (Hb; alk. paper)

ISBN 90 272 5174 6 (Eur.) / 1 58811 276 4 (US) (Pb; alk. paper)

© 2002 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands

John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

Table of contents

Preface vii

Part I. Neuropsychology 1

Chapter 1
Cerebral specialization 5
A. Characteristically-human behaviors 5
B. A brief history of laterality 8
C. Handedness, visuospatial processing and music perception 13

Chapter 2
The central dogma of human neuropsychology 25
A. Hemispheric cooperation during language processing 27
B. The effects of callosal damage 34
C. Hemispheric encoding/retrieval asymmetry (HERA) 37
D. The central dogma 40

Chapter 3
Musical interlude 47
A. Tones, intervals and chords 48
B. Interval dissonance and consonance 57
C. Harmonic tension and resolution 60
D. Major and minor modes 83
E. Conclusions 88

Chapter 4
The coding of human emotions 93
A. Intonation and music theory 100
B. The pitch range of normal speech 107
C. Continuous or discontinuous pitches in speech? 109
D. Methodological questions 118

Chapter 5
The brain code 123
A. Cortical maps 127
B. Homotopic callosal connectivity 131
C. Intonatopic maps 138

Part II. Consciousness and cognition 151

Chapter 6
Synapses and action potentials 155
A. What needs to be explained? 158
B. Cell physiology 167
C. Implications 176

Chapter 7
Synchronization 179
A. Scandal 180
B. The neuron's two modes of contact with the external world 184
C. The temporal coordination of neuronal firing 187
D. Cellular-level and organism-level definitions 195
E. Three explanatory gaps! 199
F. Conclusions 202

Chapter 8
A bilateral neural network simulation 205
A. Bilateral cognition 206
B. The simulation 211
C. Conclusions 235

Chapter 9
Conclusions 241

Appendix 1. Musical emotions 245
Appendix 2. Calculating harmoniousness 265

References 271

Index 287

Preface

In 1986, I wrote a book with the audacious title of "The Brain Code," which summarized the findings from neuropsychology indicating a particular view of hemispheric specialization. The basic argument was that the functions of the human cerebral hemispheres are not simply different from one another, but complementary – principally due to mutually inhibitory effects acting through the corpus callosum. The topic of hemispheric specialization is just one part of human psychology, but left and right hemisphere relations simply must be addressed in any serious attempt to understand characteristically human behaviors such as language and tool-usage.

I chose the phrase "brain code" because I was convinced that the phenomena of the human mind needed to be described at the appropriate *brain* level – not at the level of the gene or the neurotransmitter, and not in terms of social psychology or politics. The lower-level "neuron code" and the higher-level "social code" are both important topics, but in my view psychology should be primarily about the functions of whole brains and individual behavior. Of course, neuron physiology underlies all psychology, and beneath that level there are relevant genetic and biochemical phenomena. Equally obvious is the fact that any true insight into human psychology will have sociological and political implications, and that sociopolitical phenomena provide much of the content of our daily mental lives. Still, I maintain that the aim of psychology proper should be the clarification of the mechanisms of thought at the level of the whole-organism in a terminology that relates, first and foremost, to the behavior of adult individuals – and in a terminology that has clear connections to "common sense" psychology and the "folk wisdom" that we are all familiar with and that still rumble through our minds and daily conversations.

My impression then and still today is that many discussions of ostensibly psychological matters end up discussions about other issues – higher or lower level topics that are perhaps of interest to anyone studying psychology, but issues that should not and cannot replace our discussion of the brain and its control over behavior. The declaration of a "brain code" was an attempt to focus on the questions of brain mechanisms – principally, how thoughts and

feelings are stored and transmitted in the brain. For many psychologists, that is the level of psychology that initially motivates the academic study of the human mind, and that is the level at which I maintain a general theory is required in order that human psychology can attain a fundamental and universal scientific foundation comparable to the genetic code of molecular biology or to quantum mechanics in atomic theory.

In fact, although I aimed for the brain code in 1986, my main message was concerned with an issue that necessarily precedes discussion of the mechanisms of the brain code – that is, a general framework for human neuropsychology. Before considering the "code," it was essential to establish the brain-level context within which such mechanisms work. In analogy with developments in molecular biology, the "framework" for brain theory is appropriately labeled "central dogma" – i.e., a first-order sketch of the channels of information flow, within which the details of the code must eventually fit. In other words, before we discuss physiological mechanisms, we must establish what kinds of psychological information are of central concern and what the macroscopic (brain-level, as distinct from neuron-level) pathways of information flow are.

In cell biology, a similar distinction is made between dogma and code. The genetic code is the very heart of the mechanisms of information transfer at the molecular level, and its understanding has led inexorably to the many applications of biotechnology. But deciphering the details of the genetic code in the 1960s was possible only after the promulgation of the central dogma by Crick in 1957. Even before molecular details of the relationship between the nucleic acids and proteins were known, the central dogma of molecular biology was a core concept about the main channels of information flow in the cell. To paraphrase that dogma, genetic information is stored in the cellular nucleus in the form of DNA, but utilized in the cytoplasm by transforming the genetic information into functional proteins using RNA messengers. It is impossible to say whether the unraveling of the genetic code would have happened as it did without this dogma, but what is absolutely clear, in retrospect, is that biological science now has two fundamental theoretical constructs. The first is the central dogma, which indicates the channels along which genetic information normally flows (or does not flow) within and between cells (DNA $\leftarrow\rightarrow$ RNA $\rightarrow$ protein). The second is the genetic code itself, which specifies the nature of that information and the molecular mechanisms of its transmission (the biochemistry of nucleotide-nucleotide binding and the relationship between nucleotide and amino acid sequences).

What was apparent some 16 years ago (and arguably many years earlier) was that human neuropsychology also had reason to claim a central dogma.

That is, before addressing the more difficult problems of the neurophysiology of the brain code, what had already been established over the preceding 150 years of clinical neurology was a general schema indicating the major pathways of information flow in the human brain. Generally speaking, most neuropsychologists today remain well aware of the importance of the brain-level pathways, but, in the expanding field known as neuroscience, those core neuropsychological insights are often overlooked amidst the many interesting neurophysiological details.

Presenting the case for a central dogma for human neuropsychology was possible already in 1986, but the discussion of the brain code itself did not, at that time, go beyond enumerating in the abstract what were the likely possibilities for callosal information-transfer between the cerebral hemispheres, given what was known about topographical cortical mapping and excitatory and inhibitory neuronal effects. As a consequence, the brain code argument lacked specificity, but that deficiency can now be partially rectified and some of the core mechanisms stated with greater certainty. Describing those mechanisms is, in any case, the main burden of the present volume. The argument for a specific central dogma for psychology remains, quite simply, central to any discussion of the brain code, but here the focus is on the brain code.

Given the basic facts about a central dogma for human psychology and even the beginnings of a brain code, we are nonetheless left with questions about how they might relate to the issues of human consciousness. Let us assume, for the sake of argument, that some firm conclusions about the channels of information flow and even the mechanisms of neuronal information storage and transfer can already be drawn. Do such conclusions lead to answers concerning the "big" problems addressed in consciousness studies?

The simple answer is "no". Many have argued that modern knowledge about brain mechanisms, in and of itself, does not shed light on the question why there is a "subjective perspective" in the human mind. Yes, we are undoubtedly biologically-evolved information-processing organic "machines" and the mechanisms are undoubtedly neuronal and knowable through scientific techniques, but there is another dimension that we must address in discussing "what it feels like" to be a conscious human being. For, unlike mechanical and electronic toys, human beings experience a subjectivity that somehow accompanies the neuronal information-processing. Why is it that human beings, at any rate – and quite likely many other animal species – enjoy, suffer and "feel" their subjectivity? No discussion of human psychology would be complete without addressing that question.

In so far as subjectivity is a topic that simply cannot be avoided when discussing the human mind, consideration of brain-level mechanisms is only part of the story. As important as the establishment of a central dogma would be for focusing on the basic issues of human cognition, and as important as the discovery of the brain code would be for achieving scientific clarity, the additional dimension of subjective feeling must be explored before we can assert that we "understand" the human mind. That dimension is generally the set of topics addressed in consciousness studies – some of which are old philosophical conundrums, but others of which are problems that have been generated by the remarkable advances in brain science and artificial intelligence over the last few decades. Questions such as: "What is the relationship between our material bodies and our immaterial spirits?" may be much the same as those asked by philosophers centuries or even millennia ago, but modern versions of these questions have, at the very least, a precision regarding the material brain that was simply unknown just a few decades ago.

So, even if we are not truly more clever than the philosophers of previous eras, when we ask questions about the mind-body problem, we do have a much clearer idea about the body-side of the equation. We can now place our perplexity concerning the existence of "mind" within an explicit, scientifically-sound "body" context. Indeed, the common phrase is no longer "mind-body problem," but rather "mind-brain problem" – in recognition of the fact that minds are products of biologically-functional neuronal networks – i.e., brains, and not merely warm bodies.

Because the central issues of human psychology necessarily involve topics concerning both (i) the brain mechanisms underlying information-processing and (ii) why we are subjectively aware of that information-processing, the present volume is divided into two main sections that address those issues separately. In the last chapter, an attempt is made to draw some connections between these two realms.

Finally, I would like to acknowledge the following funding bodies that have made this research possible: the Swiss National Fund (1989–1994), Kansai University (1994–2002), and the Research for the Future Program (Project No. JSPS-RFTF99P01401) administered by the Japan Society for the Promotion of Science (1999–2003). My personal thanks go to family and friends, colleagues and collaborators, and to about one in six academic referees.

PART I

Neuropsychology

The first two chapters summarize the evidence indicating the importance of hemispheric specialization for understanding the human mind. There are of course many other topics in the field of psychology – from the biochemistry of the neurotransmitters to the abnormalities of social interactions in psychosis, and there are stark differences of opinion on what issues should be considered top-priority. For those of an extreme reductionist bent, elucidation of cause-and-effect at the genetic or biochemical level is most important. In such a view, once it is known what a gene does or what a particular molecule does to a gene, then a necessary and sufficient understanding has been achieved of the material causality that leads, eventually and deterministically, to brain functions (including hemispheric specialization) and ultimately to social and political phenomena. The opposite extreme view is social determinism, in which higher-level social dynamics are seen as the causes that result in lower-level brain effects via countless alternative pathways. In such a view, enumerating the lower-level mechanisms is perhaps worthwhile science (occasionally leading to discoveries that can be medically exploited), but the implications for human psychology of the reductionist, physiological approach are virtually nil.

Both extreme positions are of some academic interest, but what needs to be defended in this climate of partisan determinism is the idea that there are brain-level phenomena that are not merely relevant, but are the essence of any coherent discussion of the human mind. Advocates of other views will be tempted to "explain away" all such brain-level phenomena as simply the inevitable consequences of more important "underlying" or "over-riding" phenomena, but the view defended here is that psychological phenomena must be studied at the brain level with the focus on the individual organism. Few would deny that genes (or, for that matter, chemical poisons) can have drastic effects on normal brain functions and few would deny that social interactions (or, for that matter, wars and famines) can alter one's normal psyche. But the terminology of everyday psychology is concerned with thoughts and feelings, wishes and desires, regrets and memories that, under normal circumstances, have a reality distinct from and largely independent of the influences of biochemistry and politics. We are, for the most part, individual minds interacting one-by-one with other individual minds – and it is the mechanisms of interaction and cause-and-effect at that level that should be the central focus of human psychology.

In response to this common-sense assertion, the mad-scientists among us will insist that "dopamine levels determine all emotions" or that "the collapse of the wave function in quantum mechanics is why we feel" or that "toilet-training predetermines what professions we choose" or that "the white-male military-industrial-complex conditions our every thought". As a matter of curiosity, we follow those lines of argument to see where they lead, but when we consider what it is that we actually think or feel right now, those "other-level" phenomena fall away and strike us as quite unreal. They continue to reverberate in the background and may be of academic interest, but "my interactions with you today" are visceral, tangible experiences with antecedents and consequences that can be spelled out concretely at a psychological level without jumping to the level of genes or geopolitics.

To deny the reality of "whole-organism" phenomena is to deny the very subject matter of the human condition – and a disheartening number of manifestly clever academics do indeed deny the relevance of whole-organism psychology. Our characteristically-human thoughts and feelings are, to the "other-level" theorists, merely epiphenomena – ephemeral whims and passing fancies that do nothing but obscure the individual's understanding of the "real" mechanisms of cause-and-effect lying at the level of the "hard science" of biochemistry or at the level of the "deep insight" of power politics. While there is no reason to maintain that higher- or lower-level phenomena are irrelevant or to deny that, at the extremes, both biochemistry and politics can play causal roles in human psychology, those who argue for the dominance of gene-level or politics-level causality in mental phenomena do violence to our sense of what it means to be a human being. While the views of the hard-core determinists seem to be dominant in the mass media, it is also apparent that a substantial majority of less-outspoken scholars and every new generation of students have a sense that "real psychology" should be concerned with what people think and feel for themselves as individual players in a world of people. It is for this reason that introductory psychology remains a popular course at the university level – popular, that is, until the student realizes that discussion of human psychology has been largely replaced by the irrelevancies of rodents in mazes and the diatribes of gender politics. The trenchant views of both the reductionists and the political activists have their proper places in a college education, but there is also a level of human psychology that is neither biology nor politics that every educated adult should understand. Modern neuroscience has already attained many important insights concerning how brains work, and the present era of non-invasive brain-imaging and personal computer-based behavioral experimentation promises to unlock many more secrets about mechanisms of the

mind. Today, amidst a torrent of fascinating bits and pieces, it cannot be said that the puzzle has been assembled into a coherent picture, but psychology is clearly well on its way to becoming a science. The chapters of this first section are therefore intended to bring to the fore the most important issues of a scientifically-grounded human psychology – and to fit those pieces together into a coherent view of how the cerebral hemispheres work together to produce characteristically human behavior.

Chapter 1

Cerebral specialization

Synopsis

Three behaviors most clearly distinguish human beings from other animal species – language, tool-use and music – and all three involve functional asymmetry of the cerebral hemispheres. It is argued that a proper understanding of the human psyche requires, above all else, clarification of the complementary specializations of the left and right cerebral cortices when people are engaged in these activities.

A. Characteristically-human behaviors

Ask an anthropologist what is unusual about Homo sapiens, and you will inevitably be told something about language and tools. Other topics will arise and the researcher's current obsession and grant applications will creep into the discussion, but central to any consideration of what it means to be a human being will be those two issues: the communication between people using language and the manipulation of the environment using tools. Neither language nor tool-making (and tool-usage) is totally unprecedented in the animal kingdom, but the human capacity for inventing and using words and tools is truly extraordinary. Hints and traces can be found in other species, but the complexity of language and tool phenomena in human hands is not a percentage increase over chimpanzee or dolphin capabilities, but a quantum leap. Quantify it as you will, we human beings spend a phenomenal amount of time and energy with our fancy words and fancy machines and, as a consequence, have developed over the course of thousands of years language systems and material artifacts of astounding complexity. While we share a large number of biological traits, behaviors and instincts with other species near and far, the similarities between our uses of language and tools and those of other species are weak.

They are so weak, in fact, that, as recently as 150 years ago, many educated and truly intelligent people (many of whom handled higher-level mathematics with greater ease, knew more foreign languages, had more developed musical talents, wider philosophical knowledge, and firmer moral and ethical fiber than the typical University Don today!) vociferously denied any evolutionary link with monkeys, much less mice or cockroaches. They were not stupid peo-

ple, but saw the behavioral *differences* between man and monkey as indicating a huge gap, regardless of whatever structural similarities the biologists had discovered. Already in Darwin's era and even more so today, it is fully apparent that there are links, similarities and common threads between the most sophisticated of human linguistic/manual talents and the seemingly quite different grunts and squeals of dolphins, the singing of songbirds and the twig and rock handling of apes. Yes, we are all inhabitants of one biological world – cousins in an undeniable, scientifically-sound, evolutionary sense, but it is not false pride that makes us think the differences between a gorilla and a chimpanzee are modest and those between them and us are tremendous. And the sources of those differences are, above all else, language and tools.

Now ask a neuroscientist what the brain is doing during language- or tool-usage. Again, there are many things to consider, fashionable and unfashionable topics, new and old methodologies, and personal research interests of various kinds, but let us start at the top: What is the single, most-outstanding fact about brain function during speech or goal-directed manual activity? The answer is unambiguous and no longer controversial: the left hemisphere is dominant. It is dominant for most people most of the time for the most basic language functions, and it controls the hand favored for undertaking precise manual manipulations of the external world. Issues concerning left-handers, special bimanual skills and the language-related functions of the right hemisphere must be addressed in any complete account of human brain functions, but the first and most firmly established fact is that the left hemisphere is the dominant work-horse for motor control during these crucial, quintessentially human behaviors. To ignore that one fact at the outset of an examination of the human brain is to throw the entire discussion out of proper perspective.

A good example of the kind of hemispheric asymmetry found in brain-imaging studies during language functions is shown in Figure 1-1. The reality of *bilateral activation* during word repetition and passive listening to words is perhaps the single clearest result (Figure 1-1A and B), but in all three conditions the activation of the left hemisphere is significantly stronger than that of the right hemisphere (the light areas within the darkened regions being the most highly activated).

The cerebral asymmetry underlying language functions is of such fundamental importance that the following chapter will be devoted entirely to that one topic, but it is relevant to note here that the general notion of differential functions of the cerebral hemispheres in man is well-established, principally from decades of observation and research in the field of neurology. For various reasons, the "left-brain, right-brain" story has been exaggerated in certain

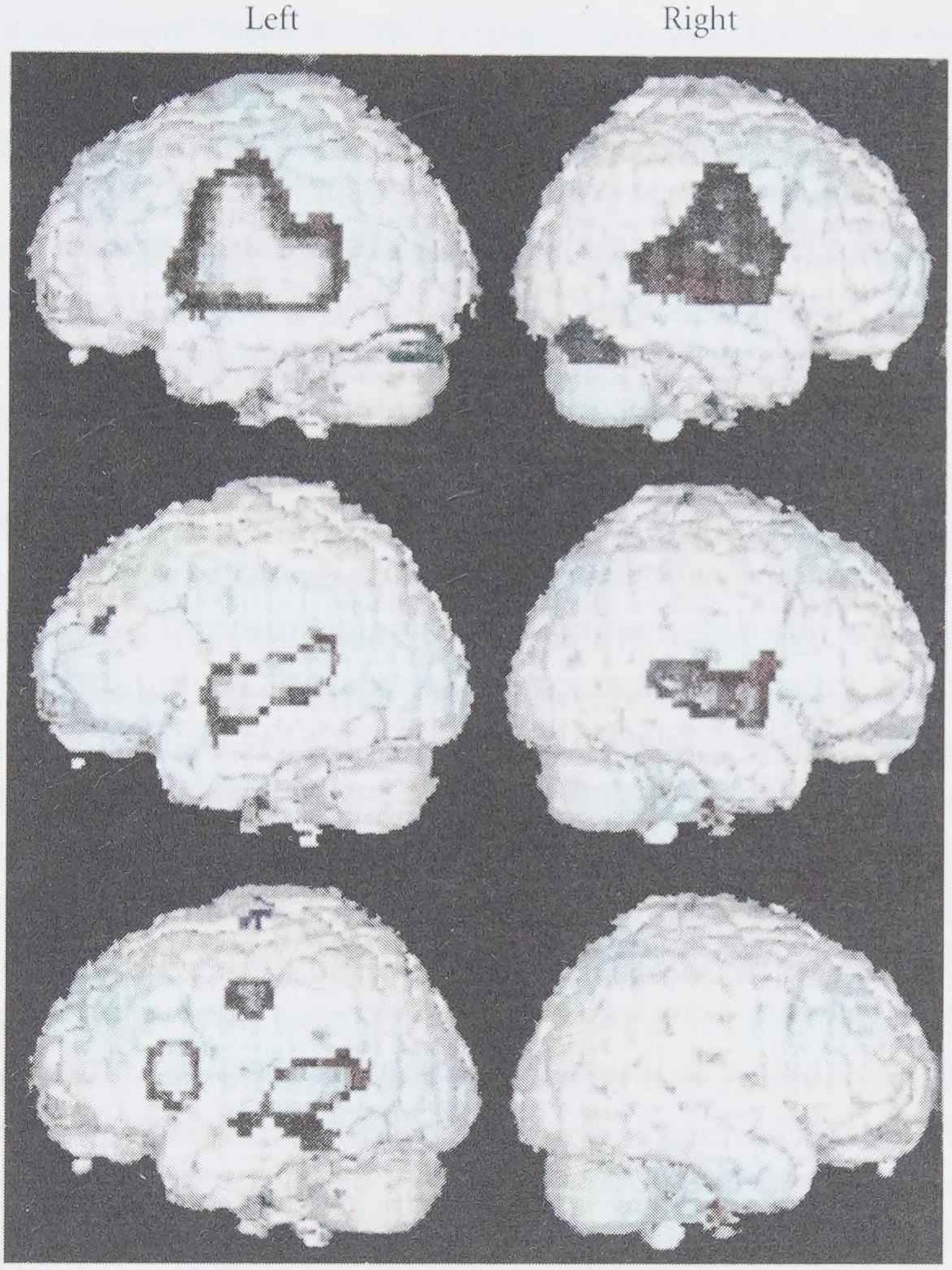


Figure 1-1. The cortical regions involved in spoken-word processing as revealed in a functional MRI experiment. The top row (A) shows the areas for repeating words relative to silent rest. The middle row (B) shows the areas for hearing words relative to silent rest. The bottom row (C) shows the areas for repeating words relative to hearing "reversed" words (Price et al. 1996). Row C is interpreted as revealing the core language areas of the brain; activation is strongest at Broca's and Wernicke's areas in the left hemisphere only.

contexts and some of the so-called scientific literature should be charitably forgotten, but there remains a core reality that no one interested in human psychology can ignore. More than anything else that has been learned from the neurology clinic, we know that brain abnormalities (strokes, tumors and traumas) involving primarily the left hemisphere have devastating effects on speech production and understanding, and on the control of the dominant hand for tool-usage and writing. That is not to say that other cognitive functions are

not lateralized or that the right hemisphere is superfluous or uninvolved in characteristically high-level human cognition, but, particularly for achieving precise behavioral output in the form of speech or tool-manipulation, the left hemisphere is the dominant control center. Delineation of the precise nature of the division of labor, left and right, will require further discussion, but there should be no uncertainty about the fact that a proper understanding of the human brain will require clarification of precisely that point.

To state this issue in the context of evolution, the unusual behavior of human beings cannot be explained simply in terms of some small (10–20%) increase in numbers of nerve cells compared to other primates – and certainly cannot be explained by the 10% decrease in neurons relative to the dolphin. We inhabit a far more interesting psychological universe and enjoy a far wider range of behaviors than other species for reasons having little to do with the numbers of neurons, and everything to do with the functional organization of those neurons. It is known that the left hemisphere is involved in the skilled "manipulation" of (verbal and manual) behavior, and the right hemisphere is involved in "other" activities (to be discussed below). Some hints of left-right specialization in other species are known, but are far less significant for animal behavior than are hemispheric differences for human behavior. The reality of cerebral dominance is thus the first outstanding fact underlying a brain-level framework for human neuropsychology. To understand the uniqueness of Homo sapiens, we must examine language and tool-usage. And to understand what is happening in the brain to allow such unusual behavior, we must examine left and right cerebral activity. Lateral specialization is not the whole story, but it is arguably *the central theme* underlying all aspects of specifically human neuropsychology.

B. A brief history of laterality

Cerebral "dominance" was a major theme of research in human psychology in the 1960s and 1970s. For many, the central question was essentially: Which hemisphere is better (faster or more accurate) at processing various types of sensory information? In retrospect, it is clear that much of the early laterality research was less profound than some of the newspaper headlines suggested, but such research did demonstrate the reality of left-right asymmetries in a vast number of psychological tasks. Not only was the dominance of the left hemisphere confirmed and reconfirmed in many linguistic tasks using a variety of technologies and methodologies, but the right hemisphere's own "dominance"

in other realms was also proven. The neglected, "non-dominant" right hemisphere was found to be superior to the left in a range of cognitive, auditory, visual and somatosensory tasks, while displaying consistent inferiority in most motor tasks.

Many discussions in the 1970s went well beyond the facts – as hemisphere differences were invoked to explain, in one fell swoop, all of the puzzles of human psychology, including the subconscious mind, creativity, and parapsychological phenomena – but the inevitable backlash was also exaggerated. Within academic psychology, Efron's *The Decline and Fall of Hemispheric Specialization* (1990) was an attempt to turn a critical eye on the methodologies of laterality research and to reject those claims for left-right differences that could be as easily explained by response biases as by differences in cognition of the cerebral hemispheres. Unfortunately, the title of Efron's book was taken too literally as a rejection of not only some cases of poor methodology, but also of several core ideas in neuropsychology. Today, a surprising number of psychologists – from Department Chairmen to best-selling authors – consider the entire issue of "laterality" to be a meaningless Age of Aquarius dead-end, but there is perhaps no clearer indication of misplaced priorities in the study of the human brain than the dismissal of the functional specialization of the cerebral hemispheres. Most of the hemispheric differences in language processing known from clinical neurology, and confirmed in countless ways in the psychological laboratory, are demonstrably real and statistically reliable; they remain the "bedrock of modern neuropsychology" (Weiskrantz 1980) – no matter how recklessly visual split-field or dichotic listening techniques may have been occasionally applied.

Historically, the issue of hemispheric specialization and its implications for human psychology have repeatedly fallen into and out of fashion over the course of more than 150 years (Harrington 1987). On balance, one can only note that some care is needed in separating the important findings that contribute to our understanding of the human psyche from the peripheral issues that boil down to rather uninteresting debates about artifacts, wishful thinking, and details of research methodology. For researchers, even these peripheral issues are of some relevance, but the cooperation of the cerebral hemispheres in cognition is far too important a topic for an understanding of human psychology to be left to academic fashions.

In recent years, the lateralized functions of the cerebral hemispheres in the normal brain have again been the focus of much serious research (e.g., Hellige 1993) that has led gradually to a more comprehensive understanding of hemispheric differences, if not yet a full understanding of hemispheric interactions.

As discussed below, it is now known that (i) both hemispheres are actively involved in high-level, characteristically-human cognition, and that (ii) the content of the cognition differs between the cerebral hemispheres. The changing emphasis on various laterality themes in the recent past can be summarized as in Figure 1-2.

The late 20th Century focus on cerebral laterality was due largely to the so-called split-brain patients. Academic studies followed by sensational stories in the popular press noted that very different functions of the two hemispheres can sometimes be measured in patients who had undergone severance of the corpus callosum (Sperry 1966, 1968; Sperry et al. 1969). The fact that all people contain two potentially independent "brains" in one skull became widely known, and attention was focused on precisely how those "two brains" – and possibly those "two personalities" – differ. While some of the speculation prompted by the split-brain research was excessive, the basic findings on the split-brain patients have stood the test of time: such patients do show signs of internal contradictions as a direct consequence of the loss of cortico-cortical connections between the cerebral hemispheres (Gazzaniga 1995). It can therefore be inferred that the corpus callosum in the intact human brain acts to resolve contradictions between the hemispheres and integrate the cognition of the left and right to produce more-or-less unified behavior.

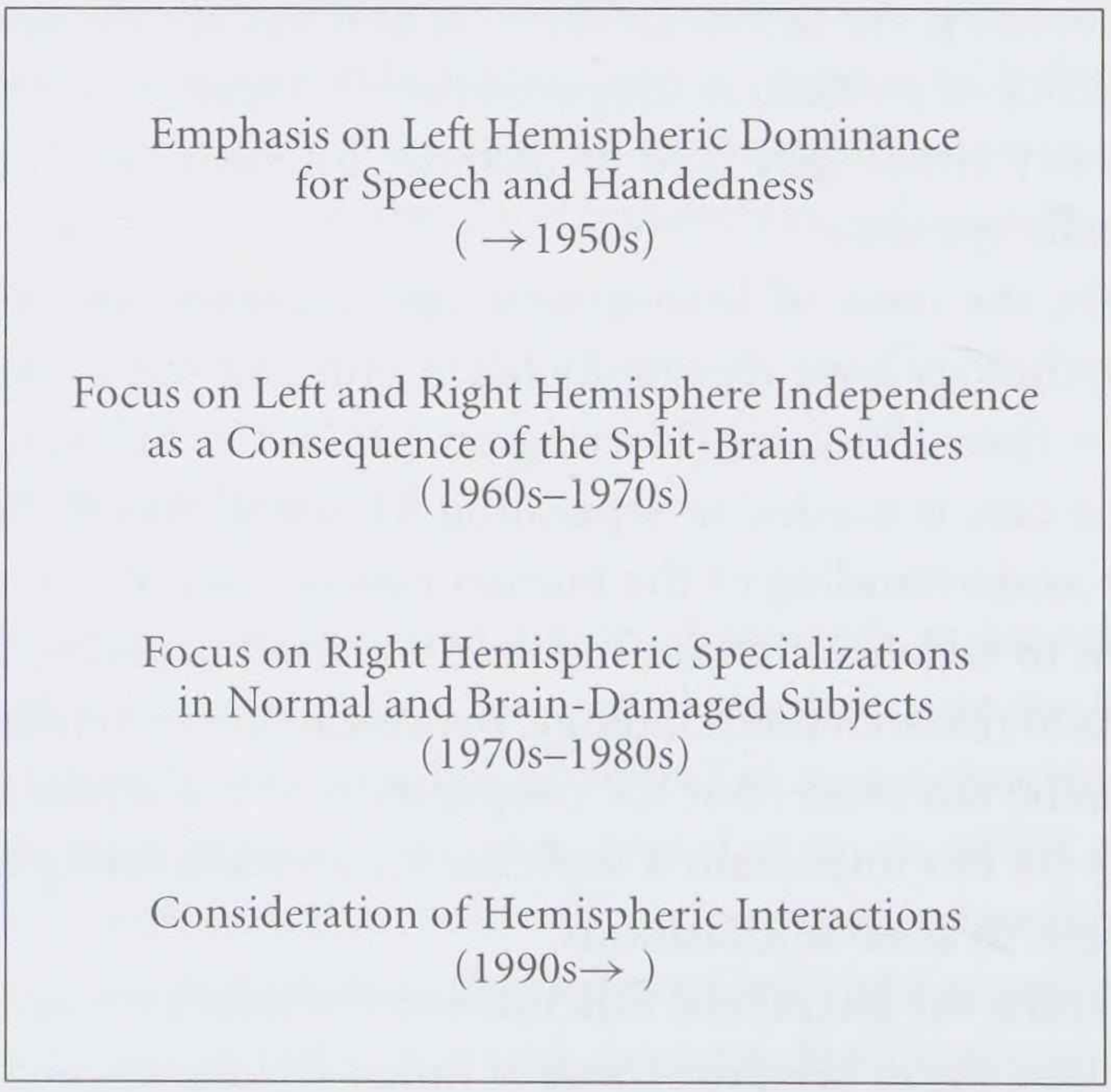


Figure 1-2. The recent evolution of the main themes concerning human laterality.

Subsequent to the initial split-brain work, neuropsychological studies on brain-damaged patients provided numerous examples of hemispheric specialization. Related, if generally much weaker, results in normal subjects using tachistoscopic and dichotic techniques were also reported, and gradually many methodological issues have been addressed – and some resolved. There has even been a resurgence in interest in the controversial topic of the "two personalities" of the cerebral hemispheres and its implications for psychotherapy (Schiffer 1998). Most recently, brain-imaging techniques have made it possible to measure directly the cortical activity in normal subjects and these new methodologies have again invigorated laterality research.

Progress has been real, but, despite the fact that many bold dichotomies of hemisphere function have found their way into the textbooks, virtually none of the dualities has survived the harsh glare of empirical research. "Verbal" and "visuospatial" remain the most popular terms to describe the specializations of the left and right hemispheres, but studies of unilateral brain-damage, continued investigation of the split-brain patients, behavioral studies of normal subjects, and recent brain-imaging work are unanimous in showing that *both* hemispheres are involved in their own ways in *both* verbal and visuospatial information-processing.

Most importantly, ever since the earliest EEG studies, a consistent finding in all types of brain-imaging studies has been the approximately bilateral activation of the cerebral hemispheres during most cognitive processes (e.g., Figure 1-1). Unfortunately, a cursory look at published research findings will often *not* reveal that fact, because research papers emphasize the statistically significant differences between left and right activation. Such differences will in fact be a major concern in the following chapters, but the underlying similarity of overall hemispheric activation in most experimental situations bears some emphasis: *Bilateral brain activation in response to sensory stimulation is the normal state*. When visual stimuli are given, the occipital cortex is activated bilaterally, and when auditory stimuli are provided, the bilateral temporal cortices respond. Both the mechanisms of and the evolutionary reasons for bilateral activation are understood. Although sensory stimulation may be asymmetrical to begin with, the ascending reticular activating system of the brainstem will be activated – in addition to sending the information along the thalamocortical tract. Because of dense cross-connections in the brainstem itself, the brainstem in turn produces "non-specific" cortical activation, so that unilateral or bilateral sensory input has bilateral effects on the cortex (Jones 1985; Steriade et al. 1990). The evolutionary sense of this mechanism is simply that if the animal is experiencing salient visual or auditory information anywhere

in the sensory field, activation of the entire visual or auditory system may aid survival. So, while small asymmetries of cortical activation are often the focus of research interest, unilateral "left brain" or "right brain" activation is virtually never the case.

The mystery of the cerebral hemispheres that was brought into focus as a consequence of the split-brain research in the 1960s is the fact that every normal individual has two more-or-less complete "brains" within one skull. Each "brain" has a full set of sensory, cognitive and motor capabilities – and yet, despite this neurological duality, we behave in a unified way – typically pursuing one goal at a time. Above all else, the split-brain patients revealed that the two hemispheres can perceive, think and behave in ways that are at least somewhat independent and at least sometimes contradictory. The paradox of the "dual brain" is that, in the callosally-connected normal state, we are not cognizant of this duality and, quite frankly, we do not know if our internal conflicts, dilemmas and the weighing of choices are the result of hemispheric differences, or merely the result of shifting priorities and consideration of different logical possibilities (with no relevant hemispheric localization). It may be the case that the mental duality that is empirically known in the split-brain patients has a correlate in the normal mind, but this cannot be known through introspection.

What is known for a certainty is that callosal connections in the intact brain allow communication between the two cerebral cortices, such that the stark contradictions occasionally demonstrated by the split-brain patients (e.g., dressing with one hand and undressing with the other) do *not* occur. In some way, hemispheric cooperation is achieved in virtually all situations in the normal brain. As a consequence, one important question for human psychology concerns how this is achieved – how the hemispheres communicate to produce the coherent, unified behavior of the normal individual – and indeed how a lack of coordination may lead to characteristic psychopathology[1] (Crow 1997, 1998). On the one hand, "neuronal coordination" is a tremendously difficult question that demands resolution of many basic issues, such as how the firing of many neurons can be integrated for goal-directed behavior and how the sense of a unitary consciousness is obtained. On the other hand, with or without answers to those difficult questions, the anatomical simplicity of the bilateral symmetry of the brain and its callosal connections suggest that there may be a small number of ways in which the hemispheres can interact.

In principle, any perceptual or motor task might be employed to answer questions about hemispheric coordination, but the most relevant types of behavior are those for which functional asymmetry of the cerebral hemispheres

has been firmly established in the neurological clinic. For this reason, study of sensory thresholds and motor reaction times is of only limited relevance to problems concerning characteristically-human psychology. The main candidates for experimental study are therefore high-level cognitive functions involving language – for which classical neuropsychology has provided abundant indication of hemispheric functional asymmetry. Evidence indicating that both hemispheres are actively engaged in language processing is the topic of Chapter 2, but before addressing the language issue, three other topics often covered in the laterality literature should be briefly mentioned.

C. Handedness, visuospatial processing and music perception

The asymmetrical bias that most people have for using their hands in a skilled fashion is of course familiar. Since motor nerve tracts from the brain cross the midline to innervate the musculature of the contralateral arm and hand, it has been known for many decades that the left hemisphere is "dominant" when the favored right hand is used. It is, however, far from certain that simply switching a tool to the left hand means that the right hemisphere immediately becomes the executor of motor control. To examine that issue, Kim and colleagues (1993) used magnetoencephalography to measure the amount of activation at motor cortex during a "skilled" manual task. The task was simply the repeated, unilateral touching of the thumb to each of the other four fingers of the hand in succession, while brain activity was recorded. A summary of the results is shown in Figure 1-3.

Of interest is the fact that the left hemisphere was active during the skilled movement of *either* hand. It was of course expected that the left hemisphere would be dominant during movement of the fingers of the right hand, while the right motor cortex was relatively silent, but, even during the left-handed task in the same (right-handed) subjects, the left hemisphere was active. As seen in the figure, the right hemisphere was activated when the left hand was involved (in comparison with the levels of right hemisphere activity during ipsilateral hand movement), but the activation of the left hemisphere was not significantly different during either left or right hand tasks.

Brain-imaging does not normally provide information on the neuronal pathways in use, but it does provide crucial information concerning which brain regions are relatively active or inactive. The left hemisphere was found to be activated during the execution of both left-handed and right-handed skilled motor tasks. The implication of this very simple experiment is that the normal

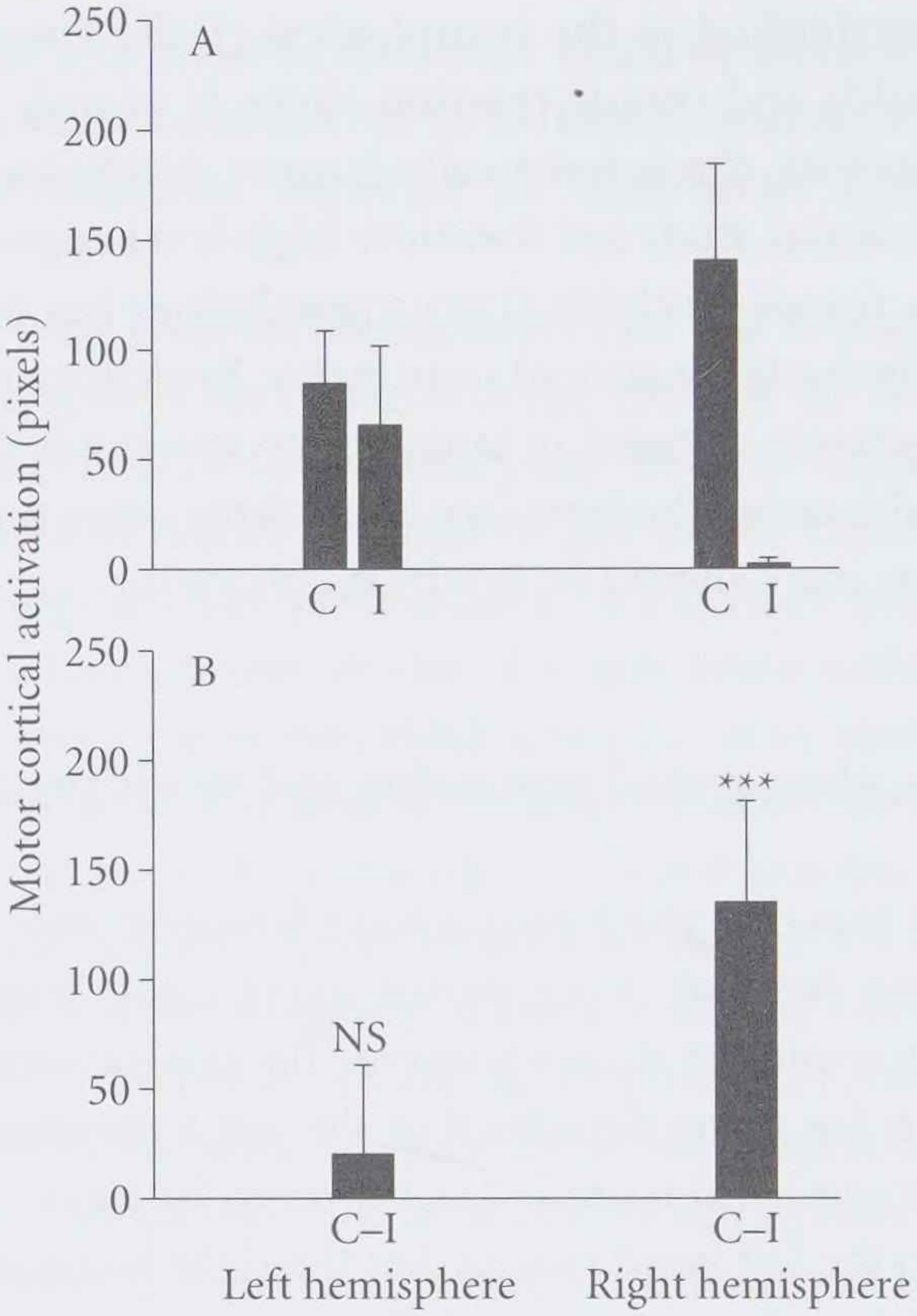


Figure 1-3. Relative activation of the left and right motor cortex during a motor task of the left or right hand. (A) shows the hemispheric activation during ipsilateral, I, and contralateral, C, hand movements. (B) shows that, although the right hemisphere is significantly activated during the contralateral left-hand task, the left hemisphere is *equally* activated whichever hand is used (after Kim et al. 1993).

pattern of executive motor control exercised by the left hemisphere is not easily shifted over to the right hemisphere – even when the non-dominant hand is the actual somatic structure involved. It would appear that, instead of relying on the unpracticed motor control skills of the right hemisphere, the brain retains its normal mode of dominance by having the left hemisphere control the right hemisphere – presumably using fibers that cross in the corpus callosum.

Such results are consistent with other findings in neuropsychology suggestive of executive control and "response selection" mechanisms being located in the left hemisphere (Rushworth et al. 1998). Together, they are indication that the ease with which we may be able to switch hands in a motor task is deceptive with regard to the actual control mechanism. If it is an even mildly skilled motor task, the brain is probably smart enough *not* to hand control over to the naïve, untrained right hemisphere, but rather maintains left hemisphere exec-

utive control, while using the right hemisphere as a "slave" processor under the direct control of the left hemisphere. Related questions are often raised with regard to the control mechanisms in bimanual motor skills: When both hands are simultaneously involved in a complex motor function, are the left and right motor cortices working independently? If the above experimental results can be generalized, then it is likely that the dominant left hemisphere is the one mastermind controlling the motor functions of both hands.

Experimental results such as these are clear indication of true unilateral "dominance" with regard to the motor control of manual behavior, but the relationship between the hemispheres during perceptual processes are better described as complementary than "dominant" or "subdominant." The complementarity for language skills will be discussed in the next chapter, but it is worth mentioning here a related complementarity with regard to how the cerebral hemispheres process visuospatial and musical information.

In the early days of laterality research, many clinical studies indicated that the right hemisphere was important for the processing of both geometrical-visuospatial and musical stimuli. Patients with right-sided brain-damage showed deficits in what would otherwise be rather simple drawing tasks and melody recognition tasks. Such results suggested a role of the right hemisphere in art, creativity and holistic information-processing, but subsequent attempts to delineate more precisely the capabilities of the right hemisphere indicated a more complex picture.

In the visual modality, the right hemisphere showed superior performance in a variety of perceptual tasks – particularly the completion or identification of geometrical shapes and part-whole relations. What came as a surprise was the finding that certain visuospatial tasks were better performed by the left hemisphere. Among these was mental rotation – often held up as a classic "right hemisphere" function! Recent brain-imaging studies have found clear indication that when the visuospatial task is an active one, involving the generation of visual images (Kosslyn 1994; D'Esposito et al. 1997) or the active, but purely "mental" manipulation or rotation of images (Alivisatos & Petrides 1997; Gill et al. 1998), the left cerebral cortex is more strongly activated than the right (see Figure 1-4).

In picture completion tasks, on the other hand, where active manipulation of the visual information is *not* required, patients with left hemisphere damage perform better than those with right hemisphere damage (Mehta et al. 1987). In fact, this simple left/right story is probably incomplete in so far as any task of this kind consists of several subtasks. In mental rotation, there are subtasks such as image generation, rotation and comparison, each of which

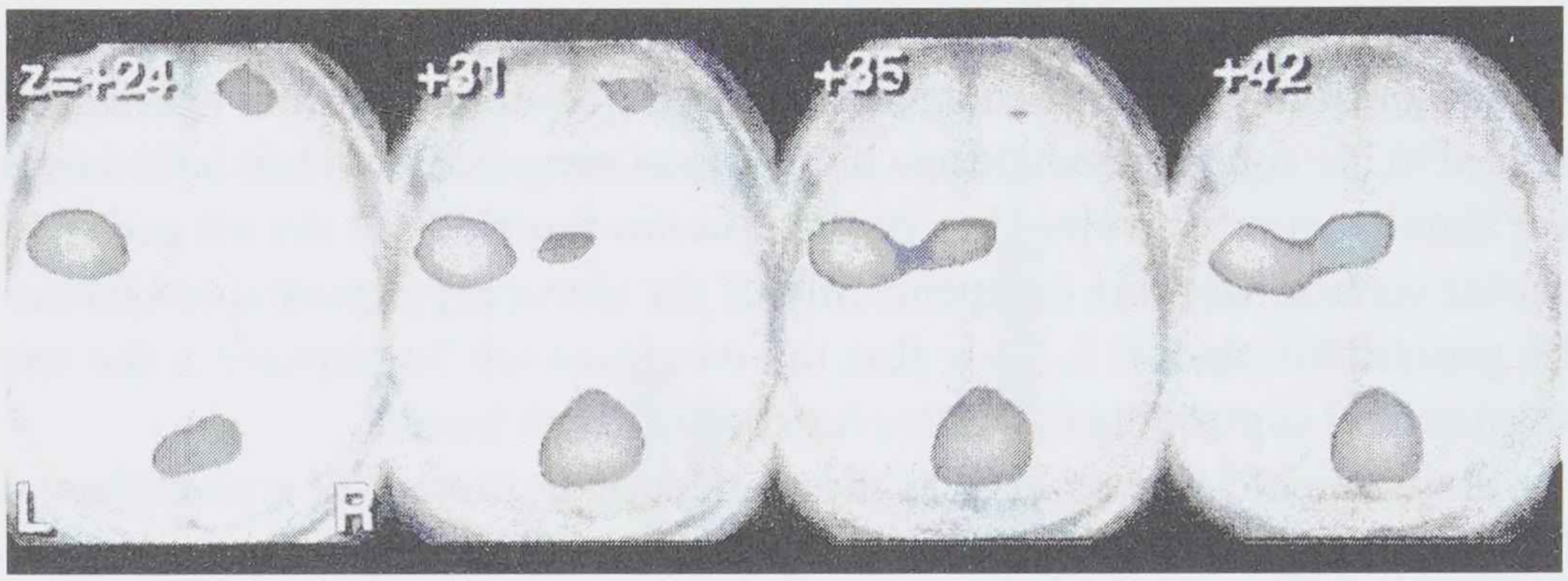


Figure 1-4. Brain activation in a PET study of image generation (from D'Esposito et al. 1997). In addition to bilateral visual cortex (bottom), there is strong activation in the left inferior temporal lobe (and a small focus of right frontal cortex).

may be lateralized to the left or right (Gill et al. 1998). In retrospect, the idea that the hemispheres have complementary talents is hardly surprising, but the fact that both hemispheres contribute to such tasks clearly indicates that gross hemispheric dichotomies focusing on separate sensory modalities for separate hemispheres (e.g., auditory versus visual processing) are almost certainly mistaken.

A simple manifestation of the uneven hemispheric involvement in mental rotation can be easily experienced. In Figure 1-5 are shown two geometrical objects. Neither is in a "gravitationally stable" position, i.e., sitting squarely on a flat surface, but rather both are precariously askew. To become a subject in this experiment, place the central fixation cross directly in front of your nose at a distance of about 40 centimeters, and try to determine if the two objects are geometrically the same or different.

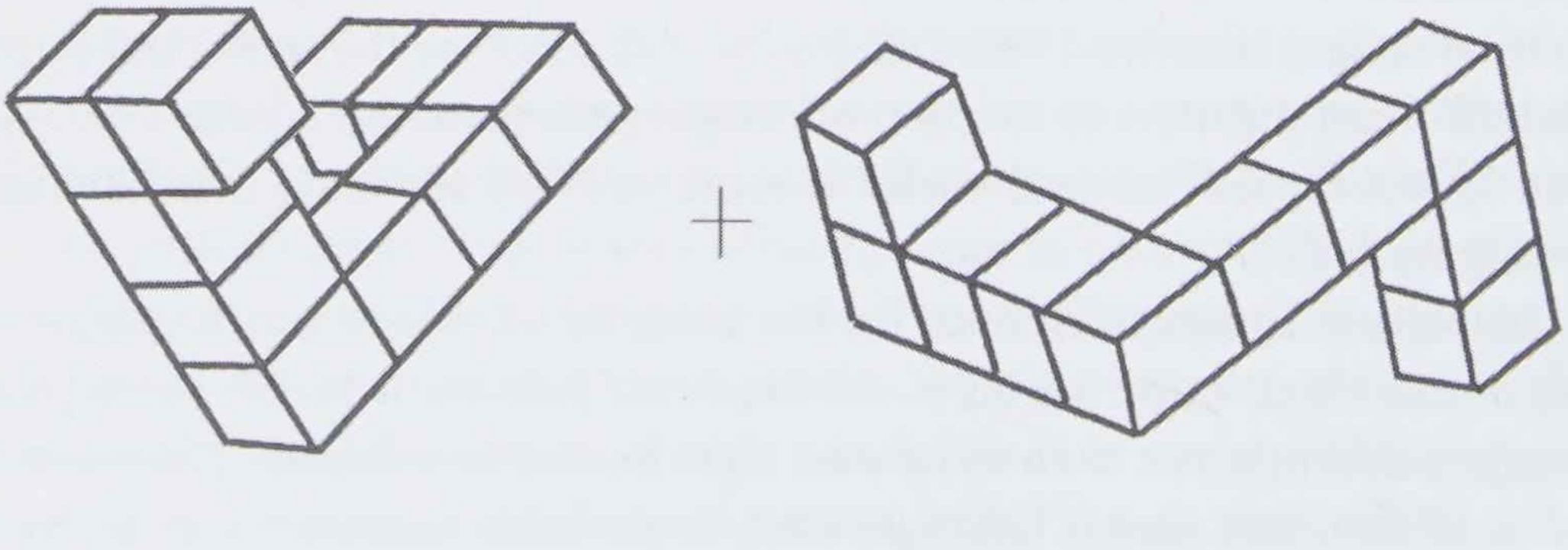


Figure 1-5. Do these two objects have the same or different geometrical configurations?

Once you are confident of the correct answer, try the task again – this time asking yourself how you make the judgment concerning the similarity of the objects. With a little introspection, most people can state how they arrived at an answer. In a more controlled experiment using a series of stimuli similar to those in Figure 1-5, we have found that most subjects will adopt a strategy of mentally rotating one or the other object to the position of the unrotated object, and then making the comparison. Of course, most people can, if asked, rotate either object and various other strategies are possible. Some people will make a judgment by first rotating one object and then rotating the other object; some claim to rotate both simultaneously; and some people use other strategies – such as imagining their arms and legs twisted into similar configurations. So, it cannot be said that there is one and only one correct strategy, but, in a large population of subjects, we found that right-handers have an overwhelming tendency to rotate the object on the right, and left-handers to rotate the object on the left (Cook et al. 1994) (Figure 1-6).

A behavioral experiment of that kind does not in fact demonstrate anything about hemisphere involvement, but there are indirect implications.

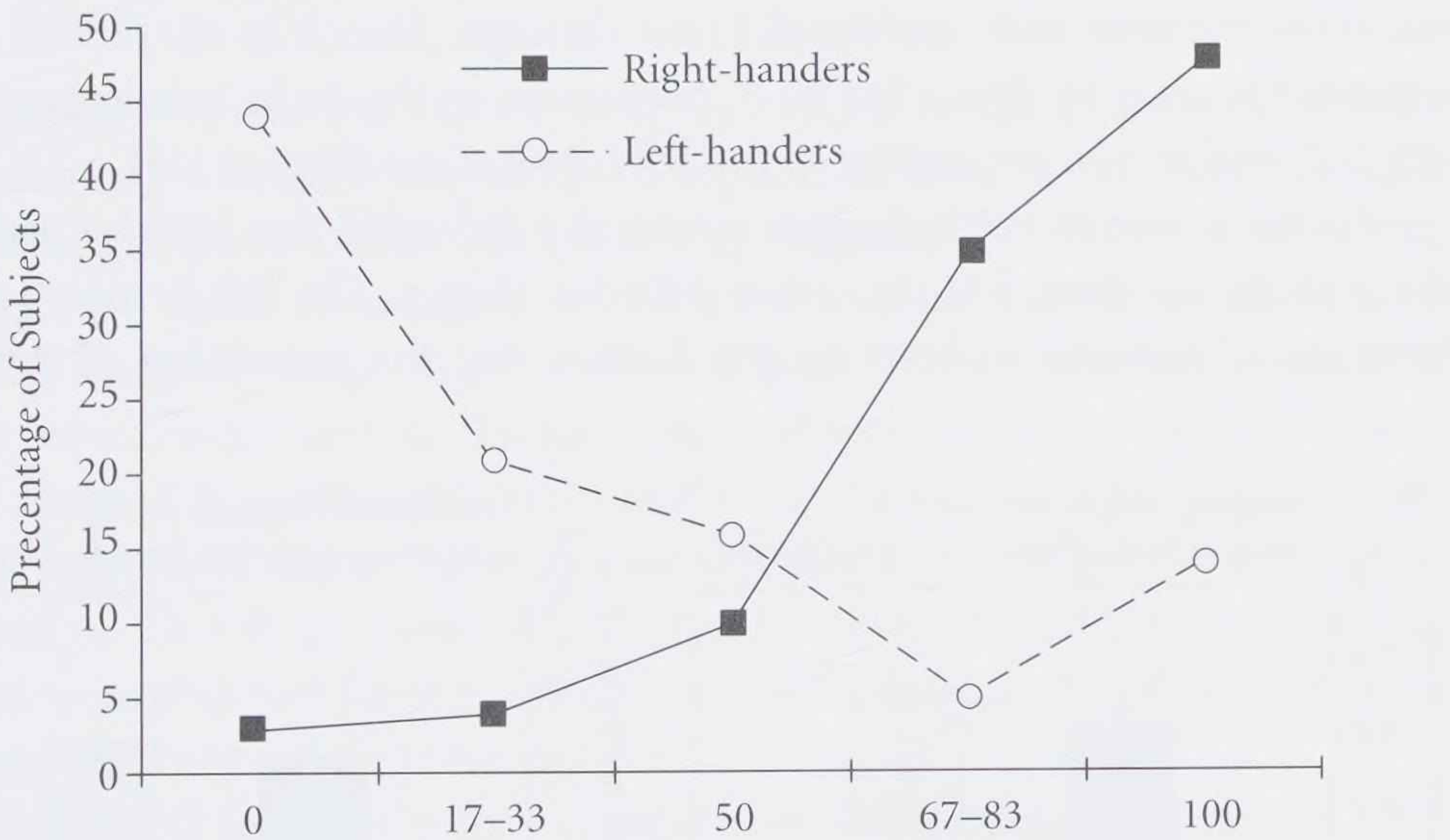


Figure 1-6. The direction of rotation among right- and left-handers. There is a clear tendency for right-handers to rotate the object on the right and vice versa for left-handers (Cook et al. 1994). Mental rotation is particularly interesting as a handedness test because it is a skill that most people have mastered, but which virtually no one has ever been trained in. Unlike writing and throwing and other skills with strong lateral biases, probably no one has ever been told how to mentally rotate.

Specifically, the image on the right of the fixation cross will be projected more strongly to the left visual cortex and that on the left to the right visual cortex (uniquely to the left and right hemispheres if stimulus presentation is rapid and eye fixation remains on the cross, but in any case more strongly to one side or the other with central presentation). Right-handers who are more practiced in manipulating objects with the right hand will generally find it easier and "more natural" to mentally rotate the right-sided object and to leave the left-sided object alone as the reference to compare against, and vice versa for left-handers. This is arguably an ingrained habit – where the dominant hemisphere functions as the active manipulator and sequentializer of information, and the non-dominant hemisphere acts as the reference – maintaining an image in mind for comparison. The important point is that for hemispheric specialization to be useful, clearly *both* hemispheres must be simultaneously engaged, because information concerning the configuration of *both* objects is essential for making the same-different judgment.

With regard to the processing of musical stimuli, case reports on brain-damaged patients have clearly indicated for many decades that the right hemisphere plays an important role. In a recent study on a large number of non-musician patients with unilateral brain-damage, Alcock et al. (2000) found remarkably clear evidence for pitch processing in the right hemisphere (Figure 1-7) and tempo processing in the left hemisphere (Figure 1-8).

Dichotic tests in normal subjects have also indicated that both hemispheres are actively involved, but do rather different things. The left hemisphere has been found superior in both the production and the perception of rhythmic

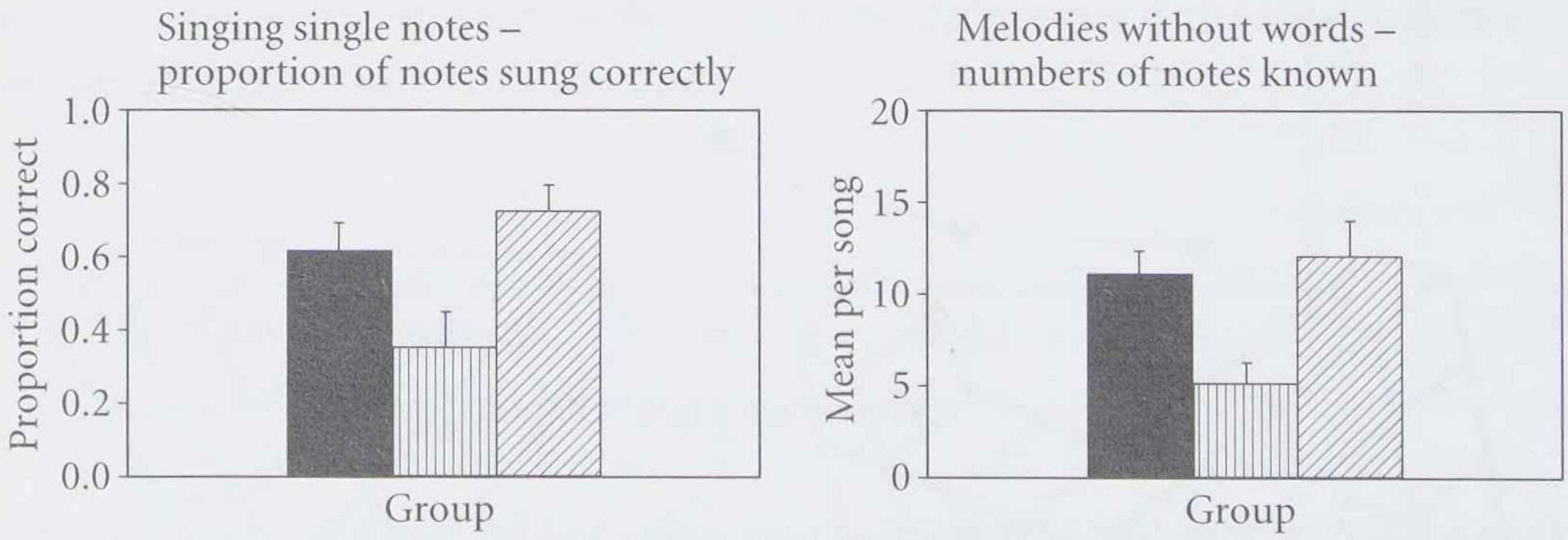


Figure 1-7. A comparison of the pitch (left) and melody (right) capabilities of left hemisphere-damaged patients (solid bars), right hemisphere-damaged patients (vertical stripe bars) and controls (diagonal stripe bars). Left hemisphere damage had only slight effects on pitch and melody (Alcock et al. 2000).

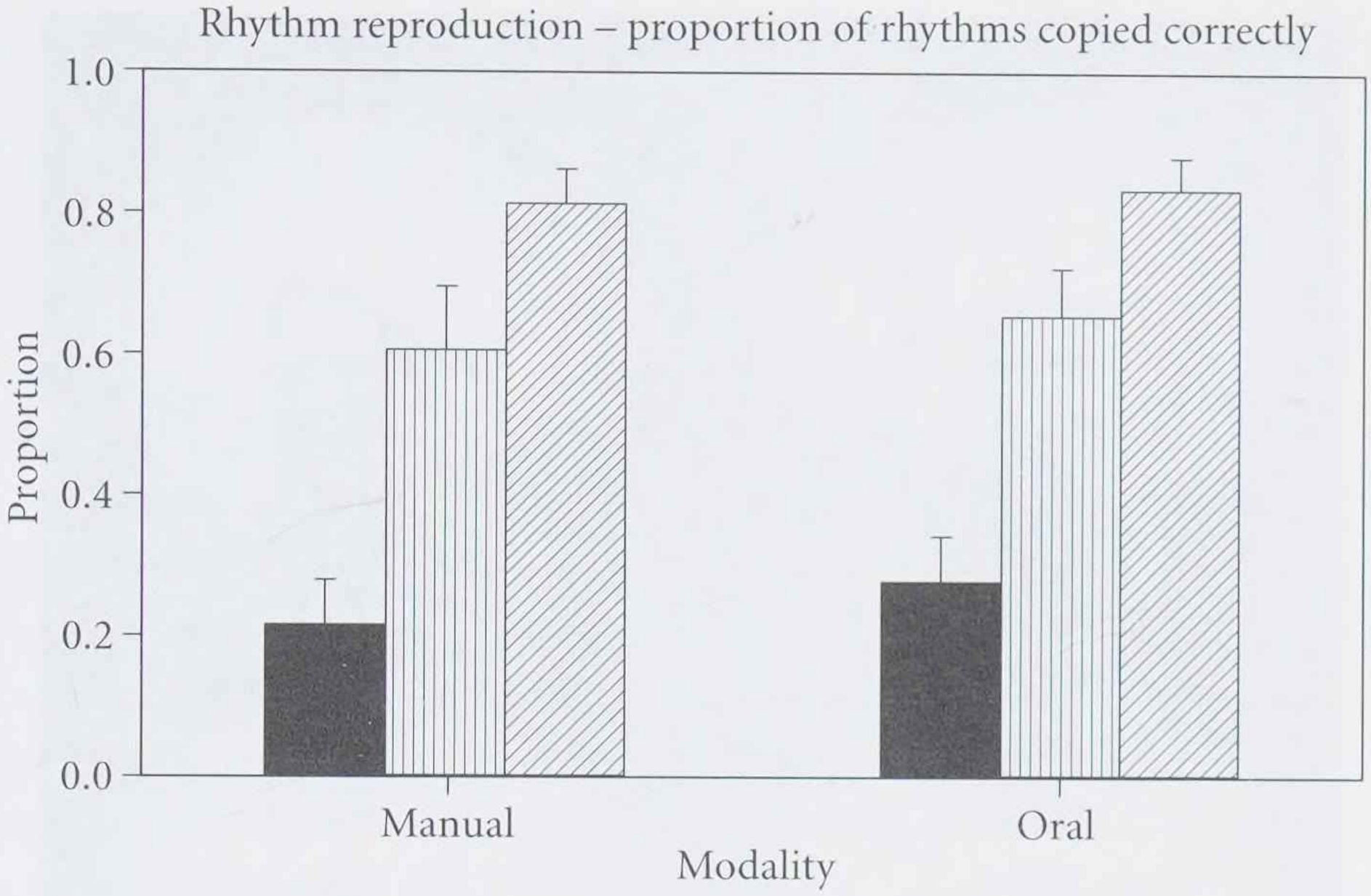


Figure 1-8. In the reproduction of rhythms, either manually or orally, left hemisphere damage (solid bars) had more severe effects than right hemisphere damage (vertical stripe bars) (from Alcock et al. 2000).

phenomena, whereas the right hemisphere excels in the processing of pitch information, particularly harmony. Perception of consonant-vowel combinations versus musical sounds was studied by Hugdahl et al. (1999) using a PET technique in which blood flow in the brain and behavioral measures were obtained simultaneously. They concluded that, "the CV-syllables resulted in greater neural activation in the left temporal lobe while the musical instruments resulted in greater activation in the right temporal lobe" (Figure 1-9). Behaviorally, more correct answers were obtained for the musical stimuli when delivered to the left ear (primarily the right hemisphere), while more correct answers were obtained for the consonant-vowel syllables when delivered to the right ear (left hemisphere) (Figure 1-10).

Using MEG techniques in a comparison of phonemes and chords, Tervaniemi et al. (1999) have also found clear indication of hemispheric specialization (Figure 1-11). The idea behind their study was to match the linguistic and musical stimuli as closely as possible in terms of pitch components. Therefore, for the phonemes, they used "e" and "o" vowels that differed only at the second formant, whereas for the chords, they used four tones that differed only at the second tone that produced either a major or minor third. Despite the strong similarity of both types of stimuli, they found clear lateralization effects.

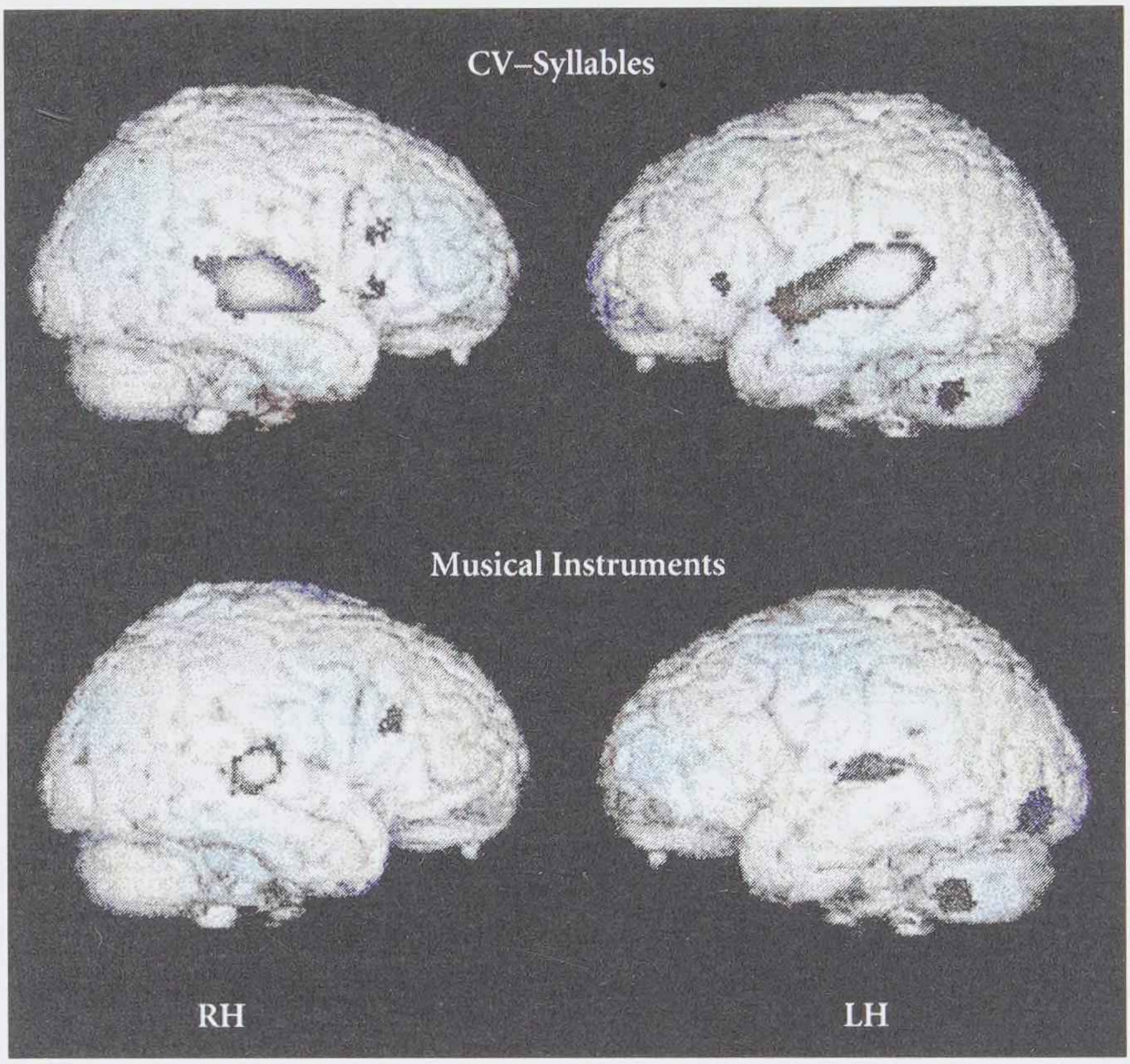


Figure 1-9. Although there was noteworthy activation of the auditory cortex bilaterally in both the linguistic and musical conditions, significant differences in blood flow levels were found (LH > RH for identification of consonant-vowel syllables; RH > LH for identification of musical instruments) (Hugdahl et al. 1999).

Experimental evidence of hemispheric specialization during music processing similar to that presented here continues to accumulate and fully supports the conclusions drawn from brain-damaged patients. The right hemisphere is particularly strong at the perception of pitch, pitch combinations (harmony) and pitch sequences (melody) (Zatorre 1985, 1997, 2001; Zatorre et al. 1994, 2002), while the left hemisphere shows strengths in processing the rhythm or tempo of musical tones, and of course in the processing of phonetic sequences.

Details concerning functional specialization for language remain to be discussed, but it can be said that the era of all-or-none functional dichotomies has come to a crashing end with the advent of brain-imaging techniques. Multiple brain regions are activated – sometimes in parallel, sometimes in rapid sequence – suggesting the importance of ideas of complementary hemispheric talents. In part, the complementarity of the hemispheres is simply a reflection

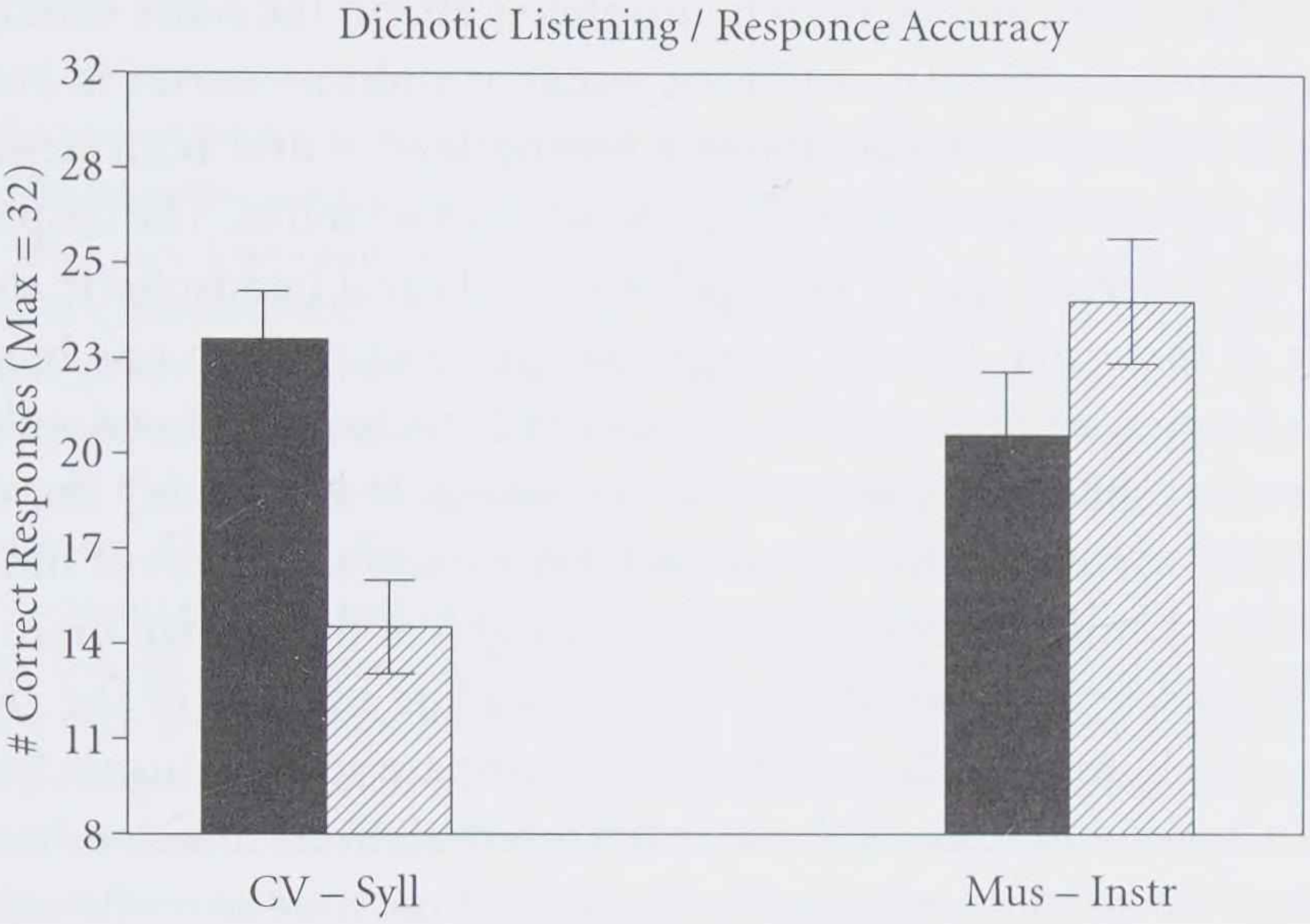


Figure 1-10. Behavioral measures (response accuracy) are consistent with the blood flow data indicating left hemisphere (solid bars) specialization for language and right hemisphere (striped bars) specialization for music (Hugdahl et al. 1999).

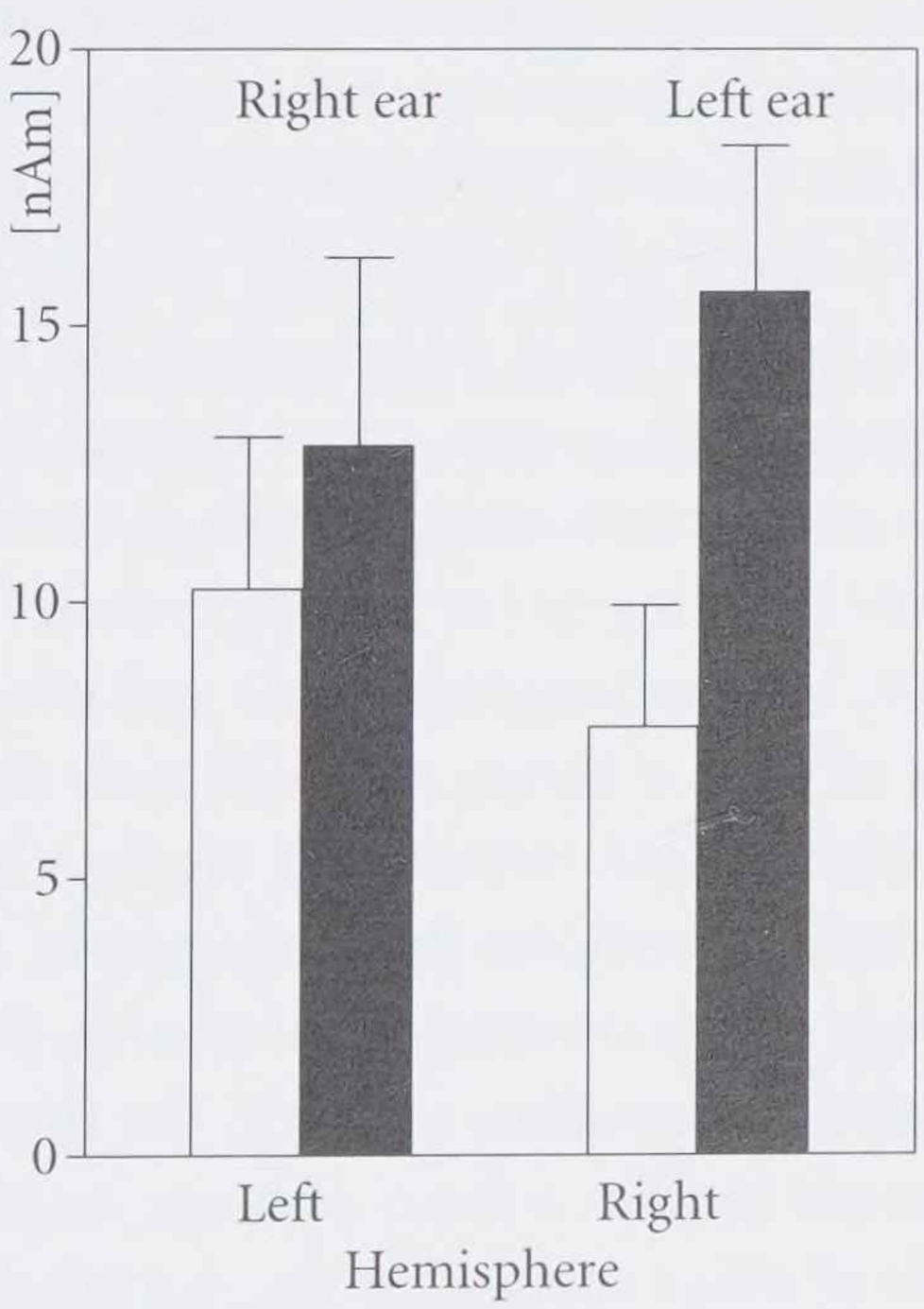


Figure 1-11. The mean strength of equivalent current dipoles indicates relative left hemisphere activation when normal subjects listen to CVC phonemes (white bars) and relative right hemisphere activation when listening to chords (black bars) (from Tervaniemi et al. 1999).

of the fact that, when awake, both hemispheres attend the same external stimuli. If the stimulus is visual or auditory, visual or auditory cortex in *both* hemispheres will become activated above a resting level – and both hemispheres will extract relevant information from the sensory stimulus. The dolphin is rumored to be different, and to be capable of unilateral hemispheric sleep, but, living out of the water, human beings have no need to keep one hemisphere awake at all times, and have apparently evolved mechanisms for simultaneous, complementary, bilateral information-processing. It is arguably the case that this capability for simultaneous, parallel, yet somewhat different functioning of the cerebral hemispheres is the unique strength of the human brain.

The neuroscientist specializing on the caudate nucleus of the cat or the monkey cerebellum will object: Of course, there are some unusual features of the human brain, and some of those indicate left-right asymmetry, but the human brain is essentially a mammalian brain and most neurons function in us just the way they function in any other animal! To which, the reply is psychological: Although we are totally incapable of introspecting on the where and how of neuronal phenomena inside our heads, objective scientific techniques have provided clear evidence concerning the reality of cerebral specialization during language, tool-usage and certain other kinds of musical and visuospatial behavior in human beings, but in few other species. How overwhelmingly important such functional cerebral asymmetry is for human psychology (and, conversely, how unimportant it is for most other species) can be understood by answering one simple question: What percentage of the waking, conscious day do we spend using tools (pencils, knives, screw-drivers, sewing-needles, tennis racquets, paintbrushes and so on), using language (reading, writing, talking and listening) or actively listening to (or playing) music? For most people, the answer is close to 100%. We use language, tools and music all day, everyday – often in combinations of two or three, and with only occasional breaks! The bulk of traditional neurology and virtually all modern brain-imaging studies on higher cognitive functions indicate that engaging in any of these activities implies differential usage of the cerebral hemispheres. So, while our neurons are physiologically typical mammalian neurons, the functional asymmetry of the human cerebral cortex implies a form of brain activation probably never experienced by animals of other species. At the neuronal level, we are just another mammal, but at the brain level, we are in another universe – the nature of which we must understand if we are to understand what it means to be a human being.

Notes

1. The pathology of hemispheric asymmetry is too complex an issue to pursue here, but it is worth noting that the normal pattern of functional asymmetry appears to be disturbed in the major psychoses. Disturbances, particularly hyperactivity, of the right hemisphere have been found in depression and abnormalities of the left hemisphere in schizophrenia. It remains the case today that, unlike various minor psychiatric disorders, there is no satisfactory animal model of the human psychoses. This is to be expected if, neurologically, the psychoses are disorders of the uniquely-human balance of information-processing between the cerebral hemispheres. As Crow (1997) has emphasized, schizophrenia may be the price Homo sapiens as a species pays for utilizing an unusual – and apparently delicate – asymmetry of hemispheric control for language functions. The complete unraveling of this complex knot would, more than anything else, establish human psychology as a science and would imply some measure of genuine therapy for psychotic patients who today can only be administered palliatives without curative effects. Crow (1997, 1998, 2002) has been the major force in explicating the evolutionary origins of this view by pursuing the genetics of schizophrenia. He and colleagues have identified a gene transposition, present on the human X and Y chromosomes, that (i) occurred after the separation of the chimpanzee and human lineages, (ii) shows abnormalities in schizophrenics, and (iii) codes for a polypeptide, cadherin, which is expressed solely in the human neocortex. The exact function of cadherin remains to be established, but the "brain code" hypothesis outlined here implies that it is involved in the functional asymmetry of the human brain (most likely, transforming the excitatory callosal impulses at certain regions of association cortex into inhibitory effects – thus promoting and/or maintaining hemispheric functional differences). The psychiatric implication of this view of hominid evolution is that disturbance of the healthy *functional asymmetry* in the normal brain is caused by a loss of the normal inhibition left-to-right and/or right-to-left. The latter (leading to LH abnormalities) has been implicated as the mechanism underlying auditory hallucinations (Jaynes 1976) and the former (leading to RH abnormalities) is arguably the mechanism through which the contextual and affective disturbances occur in schizophrenia.

Chapter 2

The central dogma of human neuropsychology

Synopsis
The hallmark of humanity is the ability to communicate through language. Although the dominance of the left hemisphere has been known for more than a century, neuropsychological studies over the past few decades have shown important language-related functions of the right hemisphere as well. The non-dominant hemisphere's contributions to language are reviewed here, and their implications for a central dogma of psychology are discussed.

Since most mammalian species show few, weak or simply no indications of lateral specialization, why would the human brain show such unmistakable and pervasive functional asymmetry? Why are we so different? Speculations about mechanisms abound, but most suggestions about the origins of laterality suffer from over-prediction. If human left-right asymmetry has its origins in the parity of fundamental particles, the directionality of DNA twisting, the unequal importance of molecular stereoisomers, the off-center position of the heart, or something asymmetrical about fetal life, then the functional asymmetry of the human brain should certainly have analogs in the brains of a great many animal species, primitive and advanced. That is simply not the case.

If, however, we remain at the neurological level and ask what is special about the neural control of speech, then one obvious candidate mechanism presents itself. What is most unusual about speech is that it is produced by motor apparatus located along the midline of the body. It is specifically this midline location that is a potential source of neurological difficulties, because motor neurons from *both* hemispheres reach the midline organs (Passingham 1981). As a consequence, the bilateral innervation creates the potential for hemispheric conflict in the control of speech that is much greater than in the (almost purely) unilateral control of either hand or foot. In other words, functional "dominance" for speech (by either the left or right hemisphere) makes sense simply to facilitate the integrity and coordination of speech-related motor behavior. Although the independence of the motor control of the left and right hands also implies the possibility for a certain kind of incoordination,

the lack of unity of hand movements is due to the independence of two separate effector mechanisms, whereas the potential problem in controlling speech involves sending contradictory commands to one and the same motor apparatus. True conflict is a possibility on the midline that cannot occur at the lateral periphery.

Any animal species that has developed mechanisms of communication using the midline speech organs should therefore show asymmetries of motor control to the extent that complex communication is possible. The barking of dogs is unlikely to be asymmetrical in that the phonetic output is not highly differentiated. The song of some songbirds, in contrast, is known to be a complex communicative system and is therefore more likely to demand integrated unilateral motor control (Nottebohm 1994).

In this view, the asymmetric control of the organs of speech is more important than the uneven hemispheric control of the hands, i.e., handedness – and indeed a much greater proportion of people shows unilateral left hemispheric control for speech (99% of right-handers and 70% of left-handers) than for handedness (90% of the general population) (Annett 1985). Moreover, brain-imaging studies have shown abnormal *bilateral* activation of the motor cortex during stuttering in chronic stutterers, and disappearance of the bilateral hemispheric activation with trained remission of the stuttering. In a PET study by Fox et al. (1996), control subjects showed the normal pattern of left hemispheric activation in both auditory and motor cortex during speech, whereas stutterers showed abnormal right hemisphere activation (Figure 2-1). They concluded that "stuttering induced widespread overactivations of the motor system in both cerebrum and cerebellum, with right cerebral dominance."

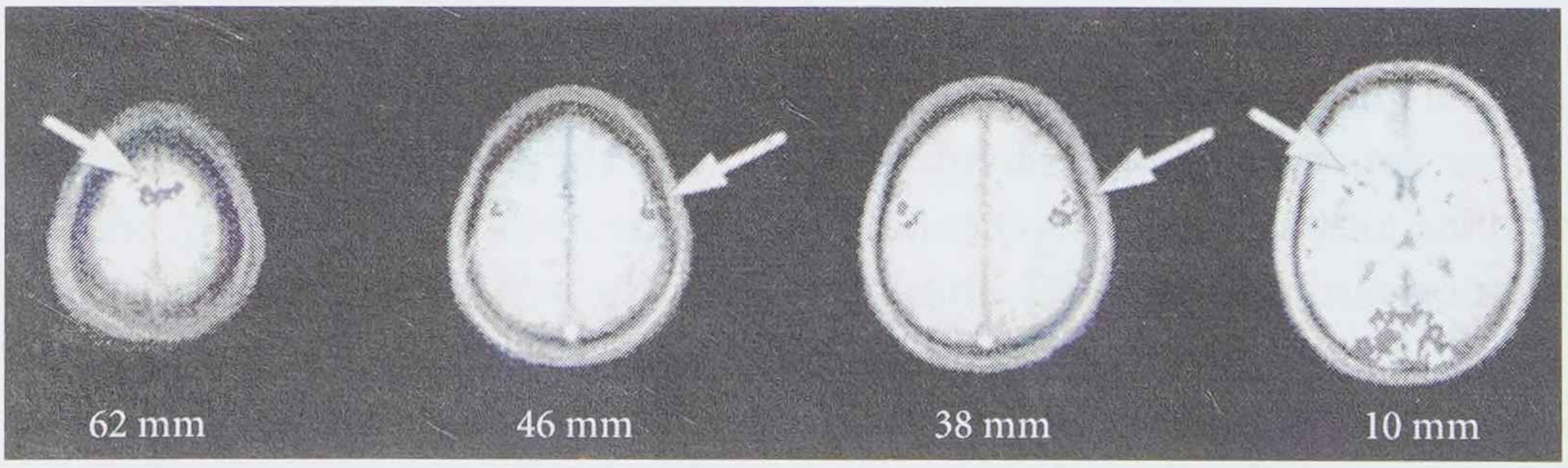


Figure 2-1. In a comparison with normal subjects, chronic stutterers characteristically show bilateral hemispheric activation during speech. Here, the arrows indicate regions of abnormally high blood flow in the supplementary motor area (bilaterally) and right motor and auditory cortex, as well as the left insula (from Fox et al. 1996).

Such evidence cannot be considered definitive proof with regard to the evolutionary origins of cerebral asymmetry, but it provides important insight into the neurology of stuttering and clearly illustrates the advantage of having functional asymmetries in the control of midline motor organs. For fluent, coherent, coordinated control of the complex motor activity required for speech, it is best to have one "dominant" executive to set the priorities and determine the sequence of motor events. Whether or not there are equally strong grounds for having one hand clearly dominant for skilled manipulation is somewhat less clear. Although most people do indeed have marked functional asymmetries with regard to fine motor control of the hands, that fact can be explained as a consequence of a small innate, genetic asymmetry followed by decades of unimanual practice, rather than being a necessity for coherent control. However the handedness issue may be resolved, one important implication of (left) unilateral control of the speech organs is that the corresponding motor and premotor cortical regions of the right hemisphere are thereby freed to become involved in "other functions." The initial, perhaps small, functional asymmetry needed for unambiguous control of the speech apparatus may have then subsequently "snowballed" into many other lateralized functions (Kosslyn 1994; Zatorre et al. 2002).

A. Hemispheric cooperation during language processing

The need for unilateral motor control of speech is a rather straight-forward issue concerning speech fluency, but the bulk of the literature in neuropsychology is concerned with syntactic and semantic issues: During language understanding or generation, what kinds of information-processing take place preferentially in the left or right hemisphere? Most questions concerning neuronal mechanisms remain essentially unanswered, but a great deal is already known both from clinical studies and experiments with normal subjects concerning the relative capabilities of the hemispheres. Some of the most interesting research in this realm are reviewed below (see also Beeman & Chiarello 1998).

The analysis of language processes in normal and brain-damaged subjects has been undertaken from the level of the smallest segments (morphemes, graphemes, phonemes) through small units (words, phrases) to complete utterances and coherent messages (sentences, jokes, short-stories, etc.). At each level, a fairly consistent pattern of hemispheric functional asymmetry has been found and indicates the active involvement of both hemispheres at multiple levels of language processing[1] (Figure 2-2).

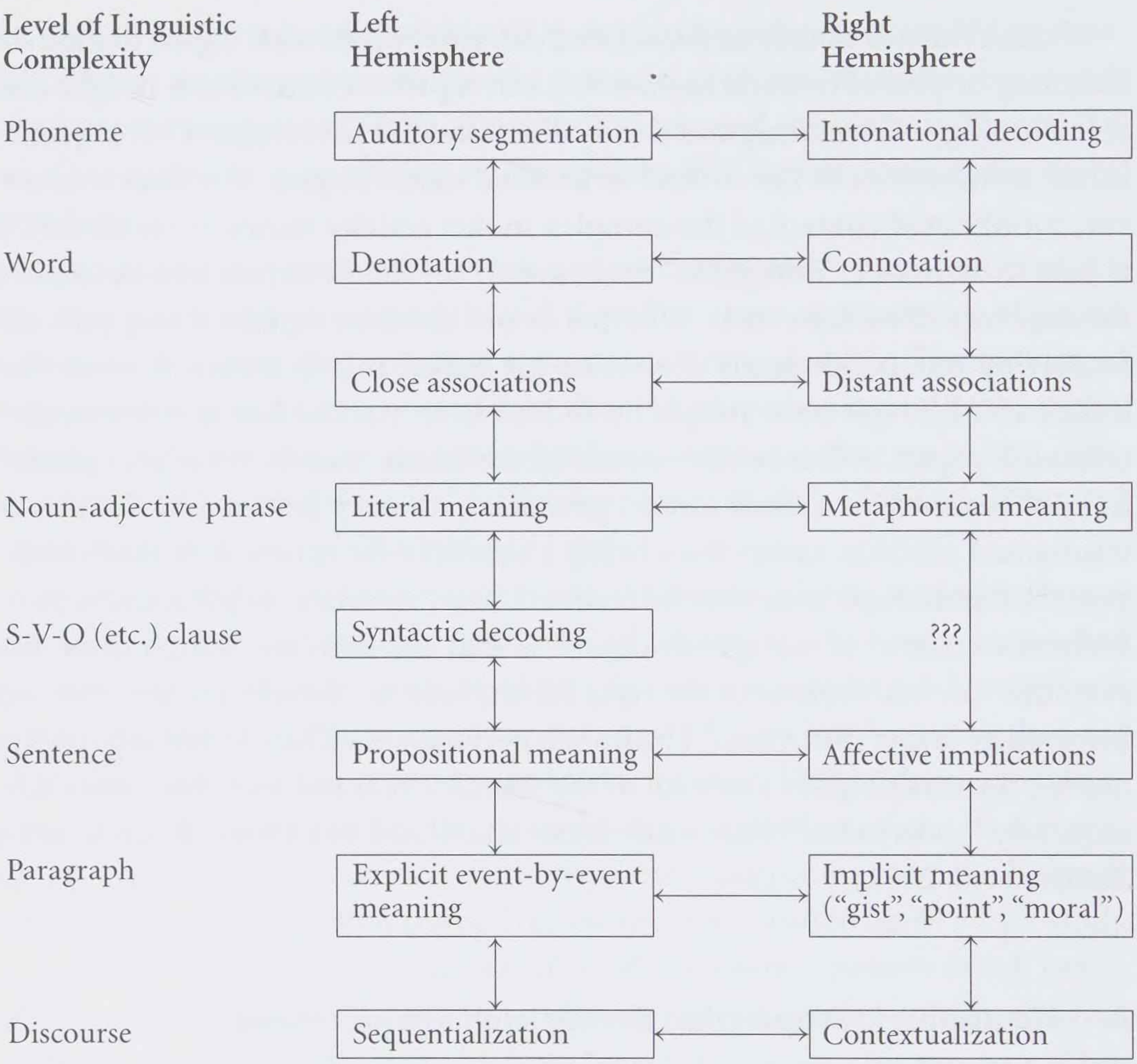


Figure 2-2. A schema of the multiple levels of bilateral language processing. The horizontal arrows indicate levels at which the cerebral hemispheres may interact via the corpus callosum. The vertical arrows indicate the bottom-up sequential processing from smaller to larger linguistic units, as well as the top-down effects from larger to smaller units. It is uncertain whether there is any type of right hemispheric processing comparable to the syntactic decoding of the left hemisphere[2] (Cook 2002a, b).

Intonation

The loss of the affective prosody of speech following RH damage is well-known clinically (Weniger 1984; Behrens 1989). As in music perception/production, the two principal dimensions of the prosody of speech concern the temporal dimension – i.e., the rhythm and timing of speech output, particularly consonant stops – and the pitch dimension – i.e., the fluctuation in auditory frequency of particularly vowel sounds. Abnormalities of timing and fine temporal discrimination are found following LH damage, and may be a key factor

responsible for dyslexia (Tallal et al. 1996). Following RH damage, the pitch dimension appears to be most disturbed (Behrens 1989; Ross et al. 1997); callosal damage has similar effects (Klouda et al. 1988). The range of pitch fluctuations is reduced, the frequency of changes in the direction of pitch intervals (melodiousness) decreases (Schirmer et al. 2001), and what prosody there is often seems inappropriate to the linguistic content. Although sometimes dismissed as "para" linguistic, the production and understanding of prosody is clearly important for normal verbal communication and is a function for which the RH is dominant.

In dichotic listening experiments with normal subjects, Bulman-Fleming and Bryden (1994) and Grimshaw (1998) have studied intonation using a design that allows the measurement of affective and linguistic understanding simultaneously. They demonstrated the superiority of the RH in detecting emotional prosody (happy, sad, etc.) and the LH in detecting linguistic meaning. In studies on patients with unilateral brain lesions, Van Lancker and Sidtis (1992) showed a double dissociation between pitch and rhythm perception, and Alcock et al. (2000) reported similar effects (unrelated to speech) that suggest LH specialization for rhythm and RH specialization for pitch. Ross et al. (1981, 1997) have defended the idea that there are types of aprosodia due to focal lesions of the RH – aprosodias that are as specific as the varieties of aphasia that occur with focal lesions of the LH. Most importantly, they present clinical evidence for a distinction between two forms of aprosodia related to the understanding and to the production of prosody – a distinction that mirrors sensory and motor aphasia. Zatorre et al. (1994, 2002) have repeatedly found evidence of RH involvement in pitch perception and have emphasized the role of the RH in both speech prosody and music. McNeely and Parlow (2001) found complementary linguistic and prosodic processing in normal subjects, and Pell (1999) found that right hemisphere-damaged patients produced significantly less F0 variation in comparison with normals.

Study of the effect that tone of voice can have on lexical decisions has demonstrated a clear connection between the two. Wurm et al. (2001) have shown that the "emotional context" provided by the tone of voice enhances word perception. This effect is thought to be related to the influence that affective expectations can have on word meanings (Kitayama 1991). Similarly, Nygaard and Queen (1999) found faster lexical decisions when the voice intonation matched the semantic content of individual words. Like Wurm et al. (2001), they argued that the emotional tone of voice is a form of contextual information that is processed in parallel with the semantic content of lexical units.

Words

In line with the idea that the LH and RH functions can be summarized as verbal and visuospatial, respectively, the possibility that concrete, easily-visualized words might be more competently processed in the RH, and abstract, less-easily-visualized words by the LH has often been studied. Results have been mixed, probably reflecting differences in stimulus materials. In a recent fMRI study designed to examine specifically the abstract/concrete aspects of hemispheric processing, Kiehl et al. (1999) found (a) bilateral activation in temporal, parietal and frontal regions during all verbal processing, with (b) more activation in right temporal cortex for abstract words, and (c) more activation of left temporal cortex for concrete words. This pattern is the exact opposite of what many psychologists would have predicted, and is indication that the concrete/abstract dimension is problematical for distinguishing between LH and RH processing.

With the notable exception of the so-called function words, most nouns, adjectives and verbs have connotative meanings, in addition to their dictionary denotations. The connotation is related to the affective state and the larger cognitive context within which the words are typically used and normally has implications beyond the literal meaning of the phrase or sentence. Since both cerebral hemispheres in most individuals will be exposed to the same words in the same contexts for an entire lifetime, it is of extreme interest that unilateral brain damage can lead to deficits of linguistic processing at either the denotative or the connotative level. In a classic study by Brownell and colleagues (1984) the understanding of words by the LH and RH was compared by having patients with unilateral brain-damage group words according to their similarity. The dimension along which "similarity" was to be determined was not specified, but the test design forced a choice between denotative and connotative grouping. Patients with an intact RH, but a damaged LH preferred metaphoric or connotative pairings, whereas patients with an intact LH, but a damaged RH preferred antonymic or denotative pairings. Gainotti and colleagues (1983) also found specific lexical-semantic deficits in patients with RH damage suggestive of semantic processing by the RH that differs from that of the LH. More recently, Taylor et al. (1999) showed qualitative hemispheric differences in semantic category matching in normal subjects. The dimensions of RH semantic processing are complex and may reflect individual differences, rather than neuropsychological universals, but it cannot be said that semantics is exclusively a LH function.

Word associations

The idea of a semantic network within which the words known to an individual are organized along various semantic dimensions has a long history in both psychology and artificial intelligence, and the possibility that the two hemispheres contain similar lexicons, but are organized differently has been frequently studied. Particularly in light of the demonstration of the independence of the hemispheres in simple word recognition tasks in both normal subjects (Iacoboni & Zaidel 1996) and split-brain patients (Zaidel 1985), it is of interest to know if the semantic organization of the LH and RH differ. In a split visual field study, Rodel et al. (1992) found the LH to favor close associations, and the RH to favor distant associations. Using an ERP technique, Kiefer et al. (1998) examined the hemispheric response to closely related words, distantly related words and unrelated words. Both hemispheres responded to closely related words, but only the RH responded to distantly related words. Such findings are viewed as support of the idea that the RH maintains more associations than the LH, with the latter focusing on one of several possible trains of thought. Using a semantic priming technique, Beeman et al. (1994) have provided evidence suggestive of a relatively "coarse" semantic organization of the RH.

Phrases

Judgments of metaphoric meaning show a similar laterality, with the RH preferring the metaphoric to the literal and the LH preferring the reverse. In a study by Winner and Gardner (1977), patients were asked to select one of four pictures that depicts the meaning of a phrase, such as "lending a hand". With one picture depicting a literal loaning of a disembodied hand and another picture depicting an individual giving help to another, the LH-damaged patients more often chose the metaphoric meaning, and the RH-damaged patients more often the literal meaning. Similar results were obtained by Van Lancker and Kempler (1987) and Anaki et al. (1998). Bottini et al. (1994) compared literal and metaphoric sentences in a PET study using normal subjects. The relevant comparison revealed right-sided frontal, temporal and parietal activations during the metaphoric sentences and left-sided activations during the literal sentences. Finally, Burgess and Chiarello (1996) have shown that an intact RH is essential for metaphor comprehension.

Paragraphs and stories

The construction of a coherent paragraph consisting of several, individually coherent (grammatically and semantically) sentences necessitates the sequencing of the sentences in the semantically correct order. In a comparison of the abilities of brain-damaged patients to do so, Gardner et al. (1983) found the RH-damaged patients to perform more poorly than LH-damaged patients. Schneiderman et al. (1992) found that RH damage significantly disrupts the patient's ability to arrange sentences into coherent paragraphs. The understanding of short stories requires one to grasp not only the individual actions of story participants, but also the consistency, overall coherence and sequence of events. Wapner et al. (1984) presented short stories to groups of brain-damaged subjects and had them retell the stories immediately upon completion of presentation. The stories were constructed with various logical anomalies, including temporally- or causally-anomalous events, and counter-intuitive actions. Regardless of the type of anomaly, RH-damaged patients were generally capable of retelling the story including its main elements, but failed to detect the nature of the anomaly. In contrast, the LH-damaged patients with intact RH, despite more diverse language problems, were capable of detecting the unusualness of the stories. These and related results have led Gardner and colleagues to consider the RH as an "anomaly detector" (1983). In related work on the understanding of jokes, Winner et al. (1998) have reported deficits in understanding following RH damage, and have argued that it represents a loss of second-order mental states. Nichelli and colleagues (1995) undertook a PET study to determine where in the brain the "moral" of a story is perceived. They found that right temporal and right prefrontal regions were activated most strongly.

In one of Geschwind's (1982) most neglected papers, he noted that the single most common consequence of diffuse RH-damage, as seen in the neurological clinic, is the "confusional state." He defined this condition as one in which the ability for speech production is normal, but the coherence of verbal output is degraded, leading to unwitting humor, paramnesias and an inability to carry a train of thought to its logical conclusion. Being based on clinical observations, Geschwind's argument that the RH normally prevents "confusion" must be considered anecdotal, but raises the interesting question of what state we are in when we are "not confused". Whether in conversation or in a monologue of speaking or writing, when ideas fall into place and lead to coherent conclusions, it might be said that each word, thought or statement is "in context" – and that cognition as a whole is contextually-grounded. If such coherency is a function of the RH, it may be that the highest level contribu-

tion of the RH to language functions is the construction or maintenance of cognitive contexts. [Note, however, that Leonard and colleagues (1997) have failed to demonstrate a contextual role of the RH in brain-damaged patients. This might be attributable to the syntactic nature of the task (the resolution of ambiguous pronouns), but, in any case, highlights the need for a more precise definition of "context".]

What is significant about the above findings is that they indicate that, within the linguistic realm, both cerebral hemispheres are engaged in information processing at approximately the same level of complexity, but with apparently different strategies. Unlike hemispheric dichotomies in which very unlike processes are contrasted (verbal/visuospatial, etc.), the multiple levels of bilateral language-processing summarized in Figure 2-2 suggest a complementary competence of the "two brains". Both hemispheres process linguistic information, but manage *not* to duplicate their processing, despite the fact that their life-long experience of all language input is identical. The complementarity, as distinct from dissimilarity, of the two modes of cognition has been a recurring theme in the laterality literature (e.g., Landis et al. 1984; Kinsbourne 1982; Regard et al. 1994). Bogen (1997), in particular, has been a persistent defender of the idea that the RH is capable of high-level cognition, and has shown convincingly that both hemispheres in most of the split-brain patients have linguistic competence – provided only that one does not insist on a definition of "linguistic competence" that is solely syntax-based.

While the ability to respond literally to simple questions and to produce syntactically-coherent propositional statements is a prerequisite to more complex language usage, verbal exchanges among normal people rarely remain at the literal level. If you don't laugh at my jokes, don't respond appropriately to my metaphors, don't grasp the "gist" of my argument or if you giggle in response to my unhappy news, we do not understand one another in the sense that we normally use the word "understand". It may be the case that literal language use and non-metaphoric information exchanges constitute the foundation on which metaphoric language is built, but the syntactic and literal semantic issues that have been the primary topic of traditional linguistics – and are the language strengths of the LH – are closer to the starting point than the completion of an understanding of characteristically human communications. In an extensive review of the cognitive psychology of non-literal language use, Gibbs (1994) has argued that:

> Metaphor, metonymy [part-whole metaphors], irony and other tropes [figures of speech] are not linguistic distortions of literal mental thought but constitute basic schemes by which people conceptualize (p. 1).
>
> Metaphor is a fundamental mental capacity by which people understand themselves and the world through the conceptual mapping of knowledge from one domain onto another (p. 207).

Clearly, in so far as we are engaged in verbal communication more complex than asking directions to the nearest bus stop, the understanding of language requires the contributions of both literal understanding, on the one hand, and metaphoric/connotative/affective processes, on the other. Since linguistic and paralinguistic information must be brought together to obtain the benefits of both literal and non-literal modes, the question of "integration" is an important issue still facing cognitive psychology. This general problem has been understood for many years, and felt acutely by researchers in artificial intelligence who have been able to implement a variety of literal language-understanding processes and supra-human logical inference mechanisms, but have utterly failed to build intelligent machines. Given the nature of psychological research and the underlying assumptions of a scientific methodology, it is inevitable that definition of the identifiable components of cognition should precede discussion of the integration of those components, but the gap between robotic language-processing and the level of normal human metaphoric language use is as great as ever. From a neuropsychological perspective, the bridging of the gap between the realm of literal language and that of non-literal language means addressing questions of the relationship between the language functions of the LH and RH.

B. The effects of callosal damage

The results concerning the language specializations listed in Figure 2-2 have come predominantly from patients with unilateral damage to the cerebral cortex, but a remarkable fact is that callosal damage alone can produce effects similar to those following RH damage (e.g., the loss of affective intonation [Klouda et al. 1988; Ross et al. 1997] and infrequent use of affect-related words following callosal section [TenHouten et al. 1985]). In general, the language abnormalities of the split-brain patients are mild when tested in a non-lateralized fashion, but already in the earliest discussions of these patients, Sperry (1968; Sperry et al. 1969) noted that their spontaneous speech was affectively flat or inappropriate, and unusually concrete with a tendency toward literalism.

Those comments are particularly noteworthy as they were made before most of the neuropsychological studies on the affective, contextual and higher-order contributions of the RH to language understanding.

Another remarkable acute effect of callosal section is mutism. Cutting the corpus callosum results in the complete loss of speech for days, weeks or months in most callosotomy patients (Ross et al. 1984). The effect is not permanent, but remains unexplained. Why would the speech-competent LH require input from the RH to initiate speech? RH damage itself does not normally produce mutism – indicating that, when the integrity of the RH itself is compromised, the LH is *not* prevented from acting on its own. Paradoxically, following severance of the corpus callosum when the RH is intact and presumably capable of normal information processing, the presence of two functioning cerebral hemispheres that have been suddenly disconnected means that the LH can no longer undertake its most usual and perhaps least effortful behavior, speech. The implication is that the LH, prior to callosotomy, normally *awaits* cognitive input from the RH before initiating verbal behavior.

The phenomenon of mutism is perhaps not surprising in light of the effects summarized in Figure 2-2. That is, if the multilevel hemispheric division of labor shown in the figure is typical of the normal brain, then mutism might be understood as a consequence of the loss of the cognitive (connotative, metaphorical, contextual) input that motivates most *normal* speech behavior. It is relevant to note that, in response to perceived speech, even when the literal meaning is entirely clear, normal people do not necessarily respond if the "point" of the speech, the "moral" of the story, or the emotional impact of the words are not perceived. When affective information, implications and underlying significance are missing, many normal people are *not* talkative and are reticent about engaging in "meaningless" verbal discourse except for reasons of social conventions and politeness. Outside of the classroom, many people normally do not talk unless motivated by the concerns listed above under right hemisphere functions! It would therefore not be surprising if mutism followed the disconnection of the RH from the LH motor programs capable of driving the speech apparatus simply because the semantic motivation normally provided by the RH is missing. Such an interpretation of post-callosotomy mutism is entirely post hoc, but clearly a loss of spontaneous speech would not be paradoxical if the various "motivations" that normally trigger verbal behavior reside in the RH and are disconnected from their normal outlet through the LH.

In so far as callosal damage produces linguistic deficits similar to those of RH damage, the obvious inference is that, even with both hemispheres fully functional, RH cognition does not influence behavior if information is not sent

across the corpus callosum for use by the executive, "dominant" LH. In other words, for the purposes of motor utilization of the "para-linguistic" information of the RH, callosal connections are essential (or, if not essential, at least the most efficient route over which information can flow between the left and right cerebral cortices). Less clear is the influence that an intact corpus callosum has on RH language understanding. Either the RH has inherent mechanisms that differ from those in the LH and that allow the RH to process linguistic information in its own unique (affective, connotative) way or, alternatively, the linguistic processing of the RH is influenced transcallosally by LH mechanisms. In brief, does the RH need the syntactical decoding undertaken by the LH to correctly understand even propositional speech? Clearly, the answer is "no" with regard to word meaning alone, since the RH of the split-brain patients comprehends words and has a lexicon of a typical 8 year-old (Zaidel 1985; Iacoboni & Zaidel 1996; Bogen 1997), but the sequential order and implicit causality of events that is clarified through syntax might require LH processing.

Whatever the case may be with regard to hemispheric cooperation during language *understanding*, what is known about language *expression* is that the RH does *not* act as a language-processor capable of *independent* action. It relies on the LH for verbal expression, and when its access to the LH is prevented by callosal section, RH information is simply not expressed verbally. The affective state of the RH may be "leaked" through limbic mechanisms (blushing, giggling) or somatically through gestures or facial expressions, but the RH remains verbally silent if direct transfer to the LH is not possible. The only apparent exception to this rule is verbal expression through singing. The clearest demonstration of this effect comes from the Wada test: following left carotid artery injection of sodium amytal, the LH is temporarily incapacitated, but singing is not disrupted. Contrarily, right-sided anesthesia has little effect on propositional speech, but disrupts singing. The capability of the RH to sing provides an interesting insight into the nature of RH language capabilities. While it seems likely that the pitch contour of the song aids the RH in its verbal expression, what is most remarkable about RH singing is that, with the help of pitch cues, the RH is indeed capable of correct pronunciation, correct syntax and appropriate timing of speech output, at least for the overlearned language of song lyrics. Propositional speech may not be its strength, but the RH is *not* non-verbal!

The role of the corpus callosum remains something of an enigma, but the puzzle is clearly a direct consequence of the fact that basic questions remain unanswered concerning the semantic organization of information stored in large expanses of association cortex, where callosal fibers are most abundant.

In cortical regions where the nature of the information is known, callosal functions can be quantitatively studied, but in many frontal and associations areas, the nature of cortical processing itself is uncertain, and the added effects of callosal activity remain unknown. On the one hand, the exchange of sensory hemifield information certainly occurs via the corpus callosum, but the number of callosal fibers increases remarkably from primary to secondary to tertiary sensory cortical regions. Low-level sensory functions are thus apparently *not* the main chore of this nerve tract. The explosive evolutionary growth of the corpus callosum is strong indication that it serves useful purposes in the mammalian brain and, in the human brain, it is the largest nerve tract present. Firm conclusions concerning the "callosal code" are not yet possible, but the comments of Creutzfeldt (1995) after a lifetime of study of the cerebral cortex are worth noting:

> The functions of the corpus callosum may be more appropriately summarized as allowing cooperation between the two hemispheres in controlling behavior, rather than being responsible for the interhemispheric transfer of all available information (p. 484).

Some ideas on the neuronal mechanisms of cooperation will be discussed in Chapter 5.

C. Hemispheric encoding/retrieval asymmetry (HERA)

Starting in 1993, several laboratories reported a consistent, but unexpected prefrontal asymmetry of activation in various short-term memory tasks; Tulving et al. (1994) labeled this phenomenon the hemispheric encoding/retrieval asymmetry (HERA). The basic effect is that LH prefrontal regions are relatively active during the encoding phase of stimulus memorization, whereas RH prefrontal regions are activated in the recall or retrieval phase. This was found using various brain-imaging techniques, including EEG, ERP, PET and fMRI, and reported by diverse groups. Debate continues regarding the influence of the nature of the stimuli and whether or not recall success and/or effortfulness are important factors (e.g., Nyberg 1998), but the reality of the effect using meaningful verbal stimuli is not in doubt (Fletcher et al. 1998a, b; Heun et al. 1999). While these issues remain unsettled, the asymmetry of cortical activation during verbal information-processing in short-term memory tasks is an unambiguous indication of some form of hemispheric differentiation *plus* collaboration.

In a typical HERA experiment, word-pairs, such as category-exemplar combinations (furniture-bookcase, tool-hammer, fruit-papaya), are presented during the encoding phase and the recall of the exemplars in response to the category label is demanded in the retrieval phase. Brain activation during these two phases are remarkably different (Figures 2-3 and 2-4).

The familiarity of the words, their concreteness and ease-of-visualization are factors that might influence the strength of activation, but the most robust effects have been found in tasks requiring a verbal response to a verbal stimulus (notably, the use of semantically "empty" stimuli, i.e., pronounceable non-words, does not elicit the HERA effect [Lee et al. 2000]). Although activation of the LH during the encoding of verbal material is unremarkable and expected solely on the basis of LH dominance for language, the activation of

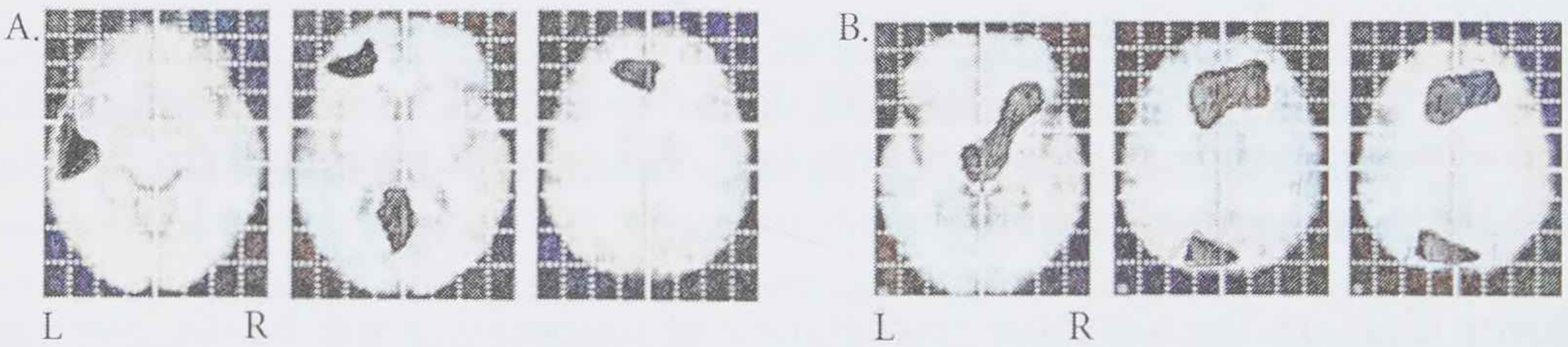


Figure 2-3. Brain activation during the two phases of a short-term memory experiment: encoding (A) and retrieval (B). The darkened regions indicate increased neuronal activation relative to a control condition (Shallice et al. 1994). During encoding left temporal cortex and left frontal cortex are highly activated. During retrieval the activation of right frontal cortex is predominant.

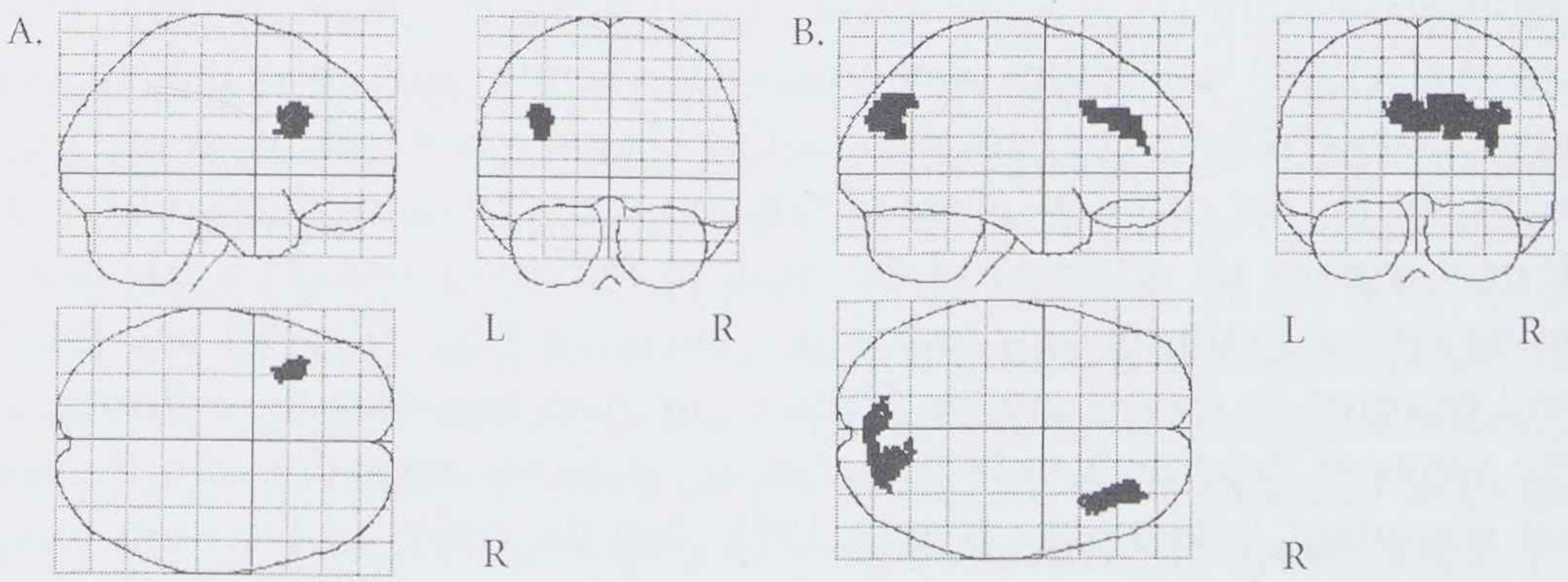


Figure 2-4. Brain activations in a short-term memory experiment by Fletcher et al. (1998a, b). A focus of left frontal activation is seen during the encoding phase (A), and foci of right frontal and bilateral occipital activation are seen during the retrieval phase (B).

the RH during recall to produce a verbal response is a nonsensical, inexplicable effect if the possibility of interhemispheric communication is not considered. That is to say, in an extreme "independent-hemispheres" model, HERA simply cannot be explained: if the information were initially encoded solely in the LH, recall should not involve RH activation at all, much less activation more robust than that of the LH regions involved in encoding. Presumably, the prefrontal RH activation during recall is indicative of retrieval of information in response to the verbal (category) stimulus, and transfer of that information to the LH to initiate the appropriate speech response. That the corpus callosum (CC) may be involved in this interhemispheric communication is the obvious first consideration, so that a working hypothesis might be summarized as:

The encoding phase:

sensory organs (→ posterior LH) →
prefrontal LH (→ CC → prefrontal RH)

followed by the retrieval phase:

sensory organs (→ posterior LH → prefrontal LH →CC) →
prefrontal RH (→ CC → prefrontal LH) → **speech organs**

The structures in bold-type are known to be involved, while the involvement of the structures in parentheses is empirically uncertain, but theoretically necessary, whenever the basic HERA effect is obtained. In other words, if retrieval occurs from structures to which encoding must necessarily have first taken place, then we must postulate a relatively "diffuse" (or delayed) involvement of the RH subsequent to LH encoding. Similarly, the "silent" involvement of the LH during retrieval must be assumed, even if not evident in brain-images – initially to register the verbal category to which an exemplar must be matched, and subsequently to provide the exemplar speech response. Using subtraction techniques in brain-imaging, the subtraction of a "control" condition which involves both language understanding and a verbal response, effectively erases the activation of the anterior and posterior language centers of the LH, while only the cognitive activity required for the category-exemplar processing is unique and remains visible. Interhemispheric communication is thus a *necessary* component of any viable explanation of the HERA phenomenon.

D. The central dogma

In the discussion of the hemispheric contributions to language processing, a remarkable similarity between the effects due to RH damage and callosal damage was noted (Sections A and B); that similarity suggests that normal RH contributions to language are realized through transmission of RH information to the LH across the corpus callosum. Moreover, the most obvious possibility for explaining HERA (Section C) was that the RH contributes to language-processing by communicating callosally with the LH, rather than using its information for direct control over the relevant somatic musculature (the organs of speech). The HERA effect is a clear illustration of hemispheric cooperation that would not be predicted by a simple "LH-dominance" model of language processing. This pattern of information processing in the human brain can be summarized in what I have referred to as the "central dogma" of human neuropsychology (Cook 1986, 1989, 2002a, b):

$$RH \longleftrightarrow LH \rightarrow \text{striate musculature}$$

It is "central" because it involves truly fundamental language and handedness functions that are characteristic of the human brain. And it is "dogma" because it is a fundamental tenet of clinical neuropsychology – fully established in that realm, but remaining scientifically "incomplete" in so far as neuron-level mechanisms remain uncertain.

The central dogma for psychology suggests that the function for which the left hemisphere has become specialized is the sequentialization (Greenfield 1991; or perhaps "response selection," Rushworth et al. 1998) of motor behavior. With practice, all forms of motor activity become at least partly automatized and the sequence of flexion and extension of individual muscles is executed with little or no conscious attention. As a consequence, it is easy to forget the motor control difficulties inherent to even "simple" manual tasks. Truly unfamiliar tasks – such as eating with chopsticks – are surprisingly difficult until sufficient practice has been had. But even eating soup with the non-dominant hand can be sloppy – not because we are unfamiliar with soupspoons, but because the rather trivial balancing act needed to keep the spoon horizontal is not something we normally do with the left hand. For anyone still unimpressed with the complexity required for "simple" motor control tasks, try enunciating a new sentence in a foreign language learned in high school. Overlearned phrases like "Comment allez-vous?" may sound almost like French, but any brand-new concatenation of consonants and vowels is likely to leave your Parisian friends dumbfounded! Both speech and hand movements require se-

quential control with millisecond accuracy, but it is particularly with regard to speech that unilateral dominance is important to achieve coherent sequentialization and to avoid hemispheric conflict. Since motor nerves extend from both hemispheres, an unhealthy competition could ensue if dominance is not established and maintained and it is for this reason that the strongest demand for unilateral control is thought to be in the control of speech.

What, then, is the motor cortex of the right hemisphere involved in? It is apparently *not* sequentialization *nor* fine-grained temporal control – at which the left hemisphere is competent. For example, right hemisphere damage or damage to the corpus callosum will leave basic rhythmic control of the right hand unaffected, but the tapping of even simple rhythms with the left hand becomes impossible (Kashiwagi et al. 1989). The right hemisphere is nonetheless involved in the maintenance of certain kinds muscular tension where sequential ordering is not an issue – in a word, "tonic" muscular control. In typical bimanual motor functions, the different roles of the hemispheres are clearly seen in the ways the hands cooperate. Indeed, truly unimanual motor activity is the unusual; normally, the right hand will perform skilled, sequential movements while the left hand maintains a stable "framework" or "context." The right hand writes while the left hand holds the paper immobile; the right hand hammers, screws or saws while the left hand holds the nail, screw or wood in place; the right hand inserts the knife or needle while the left hand holds the material in position; the right hand applies the tape or glue while the left hand maintains pressure on the pieces; the right hand picks around for a melody while the left hand maintains pressure on the strings or keys for a given chord; the right hand carves out a shape while the left hand holds the object still, and so on. Call the right hand "dominant" if you like, but undertaking any of those tasks singlehandedly with the dominant right hand alone is difficult or impossible, unless the supportive role of the left hand is replaced by mechanical means.

If both hands are in fact actively involved, why is the left hemisphere referred to as "dominant" in classical neurology? The answer lies in the nature of the discombobulation of the system when either set of functions is disturbed. With regard to handedness, the supporting role of the RH can often be replaced or eliminated altogether without totally preventing the executive LH from carrying on at some level of competence. One-handed writing is possible if the paper is immobilized; and one-handed sawing, hammering, sewing and so on can be managed, if less efficiently. In speech, if the tonic vocal cord tension that is needed to produce recognizable vowels and intonational contours is somehow imperfect, the voice will lack appropriate prosody, distinctions between vowels may be unclear, and the musicality of singing will certainly suffer. But,

in so far as the sequence of consonants is correctly produced – even with rather imperfect vowel production, speech will be comprehensible to others. Neither a monotone voice nor a foreign accent is a serious impediment to understanding! But if there are problems in sequentialization, there is no possibility for the listener to sort out the intended words from the combinatorial zoo that an improperly sequentialized set of speech sounds will imply. If you "meese-pronoonce" your vowels, your friends will tease you about your regional accent, but they will understand you well enough to keep the conversation going. But if you put your consonants in the wrong place, they won't understand you and will probably consider you a "doggamned cuffing loof"! Sequentialization is simply more important than tone (so much so that in Hebrew only the sequence of consonants is written, with the vowels omitted altogether).

Even though in adulthood most people have mastered the sequentialization of speech and are at least competent at some kinds of skilled motor control of the dominant hand in writing and tool-usage, the inherent complexity of millisecond control of striate musculature can always be experienced again and we can refamiliarize ourselves with the problems of fine motor control whenever we attempt to speak a foreign language or whenever we try to use the non-dominant hand to execute an otherwise familiar motor skill. Sequential-ordering is a task that the left hemisphere normally performs quite competently, but it is remarkable how asymmetrical the human cerebral hemispheres are in the procedure of turning ideas into actions. Although both cerebral hemispheres are heavily connected to sensory and motor organs, it is known that the right hemisphere does *not* actively control behavior. If the left hemisphere is damaged during adulthood or the right hemisphere is forced to work on its own in the split-brain condition, the right hemisphere is quite incapable of understanding even mildly complex propositional speech, despite its rich lexicon (Zaidel 1985). The rather simple sequential regularities that are inherent to any linguistic grammar are already too complex for the right hemisphere to figure out. Despite the fact that it contributes in characteristic ways to language understanding and production, it is not "independent" and is clearly not a competent executor of sequentialized motor commands in speech. Instead, a large part of the flow of linguistic information to and from the right hemisphere appears to go through the left hemisphere.

This fundamental fact about the neuropsychology of language is as frequently misunderstood as it is widely acknowledged. On the one hand, the cerebral hemispheres are each virtually "whole brains" – complete information-processing neural structures with extremely similar sensory experiences of language (and everything else) over an entire lifetime. If one hemisphere is sur-

gically removed at a young age, the remaining hemisphere alone is capable of a nearly-normal range of sensory, cognitive, affective and motor processing – including those involving language. On the other hand, in the normal intact brain, the hemispheres become specialized during the developmental process to such an extent that, subsequent to acute brain-damage in adulthood, the undamaged right hemisphere often cannot take over lost left hemisphere functions, and vice versa. It is this paradox of equipotentiality, but functional specialization (and consequent asymmetry of information flow) that lies at the heart of the "central dogma."

Analogous to the central dogma of molecular biology, this three-element dogma concerning brain functions does *not* address the important issues of mechanisms – i.e., the neurophysiology of interhemispheric information flow and the translation from cortical coding to somatic behavior. The central dogma is therefore *not* the "brain code" – a deciphered neuron-level code with a scientific clarity comparable to the genetic code, where the nucleotide-nucleotide and nucleotide-amino acid pairing provide detailed mechanisms of information flow. The psychological dogma does, however, express the first-order reality of left-right localization of language functions. For focusing on the most important aspects of language input and output, it is a useful summary of the division of labor that the human brain employs. There may be other flow-charts for non-linguistic processes or for the cognitive processes of other species, but for the all-important usage of language (and tools) the human brain has evolved a functional asymmetry that entails not only functional differences, but a specific asymmetrical pattern of information flow (Figure 2-5).

Regardless of the label given to this pattern of hemispheric interaction, its general validity with regard to linguistic processing has been partially known since the mid-1800s, when the dominance of the left hemisphere for speech became established. Precisely what the right hemisphere does during language processing is a long-standing puzzle that has been only partially solved in recent years. However that issue may eventually be worked out, it is clear that most of the confusion about hemispheric specialization has come from attempts to summarize all types of hemispheric functions with a single dichotomy of appropriate adjectives: verbal/visuospatial, digital/analog, arithmetic/geometric, or whatever. Unfortunately, even when there is strong empirical support for a given dichotomy in a given context, the dichotomy is defined by the nature of stimulus materials and does not necessarily generalize to other modalities. What is needed therefore is *not* the perfect set of philosophical or psychological adjectives, but a *neuronal mechanism* – or, more likely, a small set of neuronal

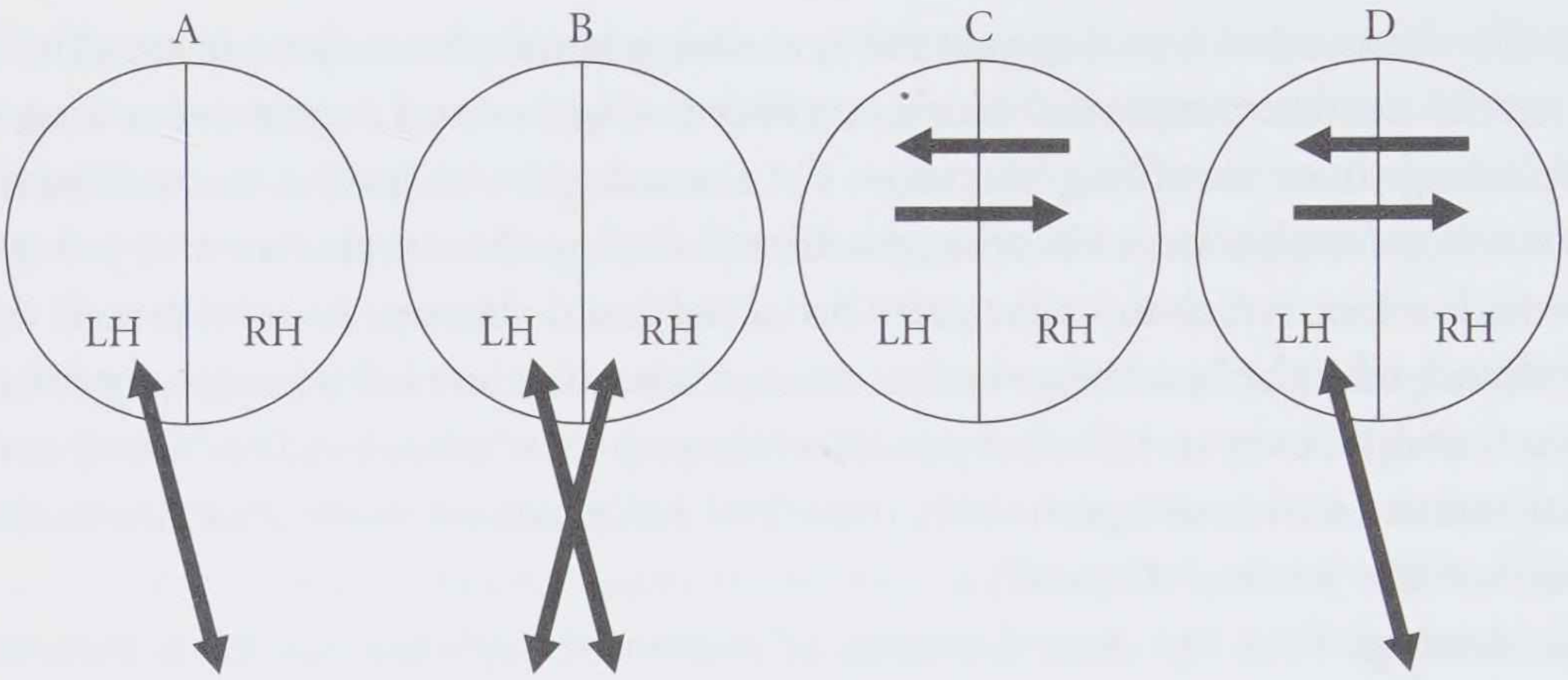


Figure 2-5. The evolution of ideas on human laterality (Cook 2001b). An emphasis on LH "dominance" for speech and handedness was the main theme of early laterality studies (A), while hemispheric independence and the distinct modes of LH and RH cognition were the focus of most split-brain and split-brain-motivated studies (B). Hemispheric interaction has more recently been researched (C), but hemispheric interaction models still need to be reconciled with basic facts indicating LH dominance, particularly with regard to language (D) (Cook 2002b).

mechanisms – that can be understood to operate on various kinds of cortical information.

The establishment of neuronal mechanisms to replace (and explain) a host of psychological dichotomies may seem unachievable at the present time, but the relative simplicity of callosal anatomy (largely homotopic connections between cortical areas that process information of the same modality or modalities) and the two-dimensional organization of the cortical surface greatly constrain the types of callosal mechanisms that must be considered. Already by the 1980s, all four of the most obvious possibilities (i.e., assuming predominantly excitatory or inhibitory synaptic effects, and predominantly diffuse or focal callosal fiber termination) had been defended in the laterality literature as "the crucial mechanism" underlying human laterality (reviewed in Cook 1986). Since then several variations on the basic excitatory and inhibitory models have been advocated (Chiarello & Maxfield 1996; Burgess & Lund 1998; Querné et al. 2000). When taken as the "one-and-only" mechanism of hemispheric interaction, no single hypothesis alone can account for the diverse nature of interhemispheric information flow, but the partial validity of each model may well be demonstrable both behaviorally and in brain-imaging studies. Final answers concerning neuronal mechanisms are not at hand, but the HERA pattern of cortical activation provides an indication that something as simple as

the "central dogma" flow-chart for interhemispheric communications may be a useful framework into which the neuronal mechanisms will eventually need to be inserted. (See Chapter 5 for further discussion.)

Notes

1. Although Figure 2-2 summarizes the known pattern of high-level hemispheric cooperation during language functions, it does not address questions about how the two hemispheres actually communicate. The mere presence of the corpus callosum, its explosive growth in primate evolution, and the high density of callosal fibers in association cortex argue strongly for some kind of direct multi-level hemispheric interaction, but the alternative view is also worth considering: the two hemispheres may veer in different cognitive directions solely as a function of the affective bias at an early stage in (auditory) processing, the consequence of which being that the RH will always be more sensitive to the emotional implications of language – whatever the linguistic complexity.

2. The second puzzle of bi-hemispheric language processing concerns whether there is an unpaired module for syntactic processing in the LH or a paired module in the RH. The intonation argument developed in later chapters suggests that the three-element processing that lies at the heart of syntax (e.g., SVO word order) might have an analog in three-pitch processing in the RH. In the latter case, the cognitive capability that is unique to human beings would not be language per se, but rather the ability to perceive three-body interactions, whether in the form of basic syntactic structures or three-tone harmonies. The implications for music perception are discussed in detail in Chapter 3, but the interesting issue of the interaction between pitch and syntax in relation to interhemispheric communications remains to be explored.

Chapter 3

Musical interlude

Synopsis

The most firmly-established paralinguistic talent of the right hemisphere is the perception and generation of voice intonation. In order to determine what kind of information is contained in the tone of voice, basic concepts in psychoacoustics are reviewed and relevant findings in music psychology are discussed. It is shown that relative "dissonance" is the salient feature of two-tone pitch combinations and relative "tension" is the salient feature of three-tone combinations. A general model of the affect of pitch is then developed.

The main argument of the first two chapters has been that the left and right cerebral hemispheres contribute different types of information to characteristically human cognition. Although both hemispheres are exposed to essentially the same world over the course of a lifetime, they become specialized for different ways of thinking. With regard to language, the two forms can be summarized in somewhat overlapping dichotomies: literal versus figurative meaning, denotation versus connotation, focal versus contextual processing, and so on. There is empirical support for such characterizations, but in fact it has been the use of these and similar dichotomies that has given much of the laterality literature a bad reputation. The secret admirer of "hard science" hidden in even the fuzziest of psychologists insists on asking what precisely is meant by "connotation" or "context"! Without unambiguous definitions such terms remain problematical, and their diverse usage by various researchers makes the problem only worse. Further clarity on the meaning of the hemispheric dichotomies will be attempted in Chapter 5, but even acknowledging some inherent ambiguity here, the experimental and clinical support for the various characterizations of left-right specializations cannot be lightly discarded without ignoring much of what has been learned in neuropsychology over the past century. Primarily from the fact that the left hemisphere is dominant in guiding most forms of skilled motor behavior, and yet the right hemisphere shows characteristic contributions to cognition and ultimately behavior, a central dogma has been proposed, as follows:

$$RH \longleftrightarrow LH \rightarrow \text{striate musculature}$$

To give some substance to the bare skeleton of this dogma, however, it is necessary to state with greater precision what the nature of right hemisphere processing is – and ultimately to specify the form in which information travels between the cortices of the left and right hemispheres. For that purpose, the most firmly-established paralinguistic functions of the RH, i.e., intonation and prosody, need to be examined in greater detail, and therefore the basic terminology of pitch phenomena must be discussed. Some of the terminology is commonly used by linguists in the description of intonation, but there is still other terminology related to music that must be introduced. This is necessary for two reasons. The first is simply that music itself is a characteristically human activity; certain kinds of hemispheric specialization related to music are already well-known and must be included in a general discussion of human cognition. The second is that the pitch phenomena during speech demand a more precise theoretical framework than is currently available in linguistics – and music theory provides such a framework. Once the basic features of sounds – in music and speech – have been described, it will then be possible to delineate the relationship between music and speech intonation, and to show where these phenomena enter into the "brain code".

A. Tones, intervals and chords

There is a known range of acoustic phenomena that the human auditory system can comfortably handle. Generally, tones with auditory frequencies above 60 cycles per second (Hertz) and below 12,000 Hertz and amplitudes above 20 decibels and below 100 decibels lie within that range. The extremes vary with age and are very different for different animal species, but generally these limits define what might be referred to as a (human) "physiological" range. Tones that are near to either extreme – low-pitched buzzing noises or high-pitched squeaks – and any sound that approaches the upper limit of auditory amplitude strike us as unpleasant or even painful (Figure 3-1).

This known range of acceptable tones is simply a matter of low-level physiology related to the types of acoustic vibrations that stimulate, but do not damage, the cochlear membrane. It is of psychological interest, however, that there are also mildly "unpleasant" sound phenomena that fall well within the physiological auditory range. Although the unpleasantness is clearly orders of magnitude weaker than that of extremely loud, low and high frequency tones, there are two-tone intervals and three-tone chords that sometimes produce an irritating "ring". In contrast with the unpleasantness of extremely loud tones,

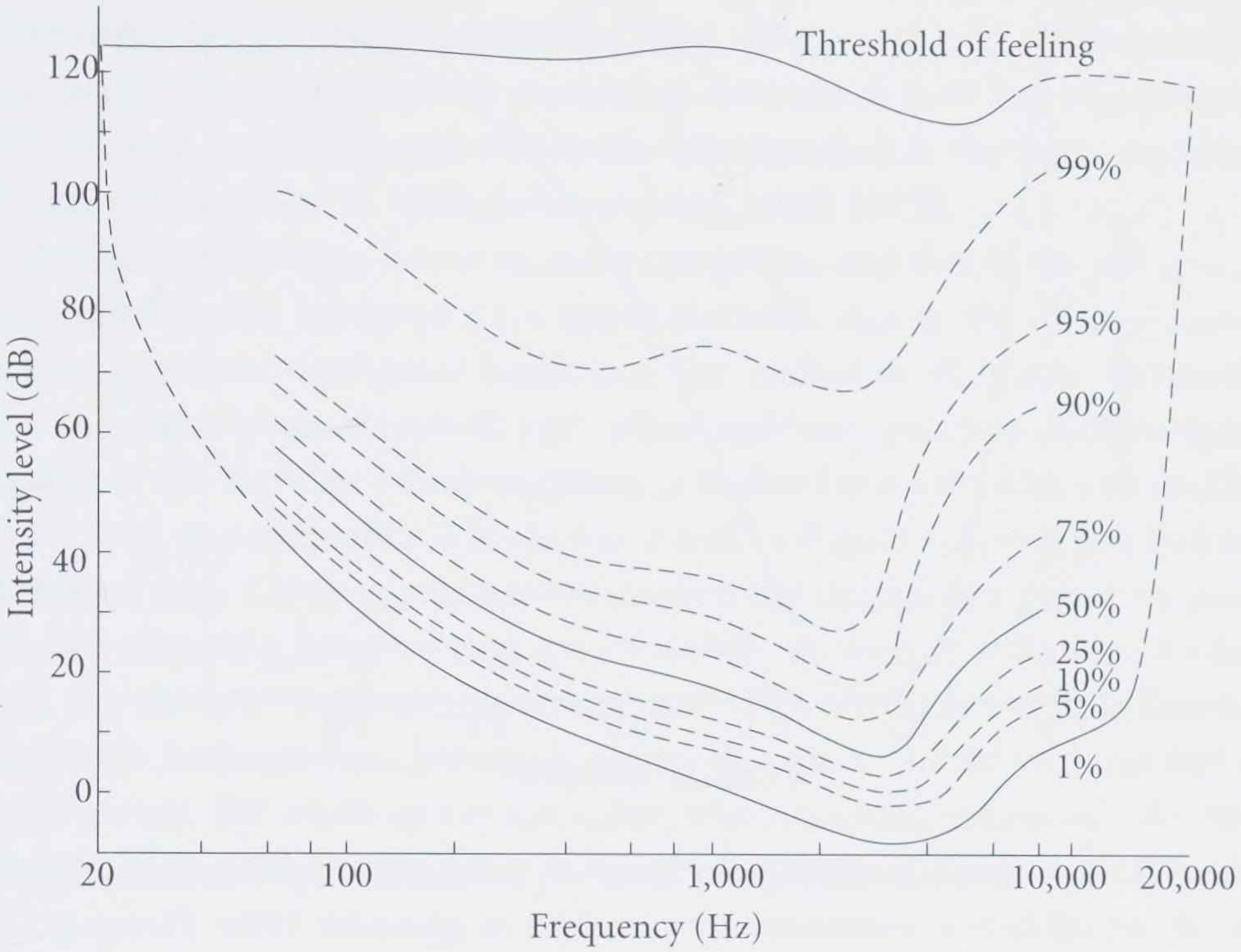


Figure 3-1. The range of pitch frequencies that the adult auditory system can handle. Below 20 Hertz or above 12,000 Hertz, sounds are inaudible to the human ear. The fundamental frequencies of most music and speech fall between 60 and 2000 Hertz. About 90% of the normal population can hear those frequencies when sounded with an intensity of more than 20 decibels (figure from Pierce 1992).

we would not normally describe unpleasant musical intervals or chords as "painful", except in a metaphorical sense. Nevertheless, what is striking is the commonality of the perception of the pleasantness or unpleasantness of certain tone combinations. Acquired musical tastes and individual likes and dislikes that are a consequence of learning and exposure to various musical traditions make the discussion somewhat complex, but there is nonetheless some agreement on the sounds that we as a species find to be particularly pleasant and unpleasant.

The perception of two-tone pitch intervals was intensively studied in the 1960s by psychophysicists. The basic findings (Plomp & Levelt 1965) were that extremely small intervals (less than about one quarter-tone) are perceived as more-or-less one tone, a unison, and are not unpleasant, whereas intervals of 1–2 semitones are perceived as being a mildly unpleasant, dissonance (Figure 3-2). Beyond an interval of 1–2 semitones, most two-tone pitch combinations are perceived as more-or-less consonant, i.e., not notably unpleasant and

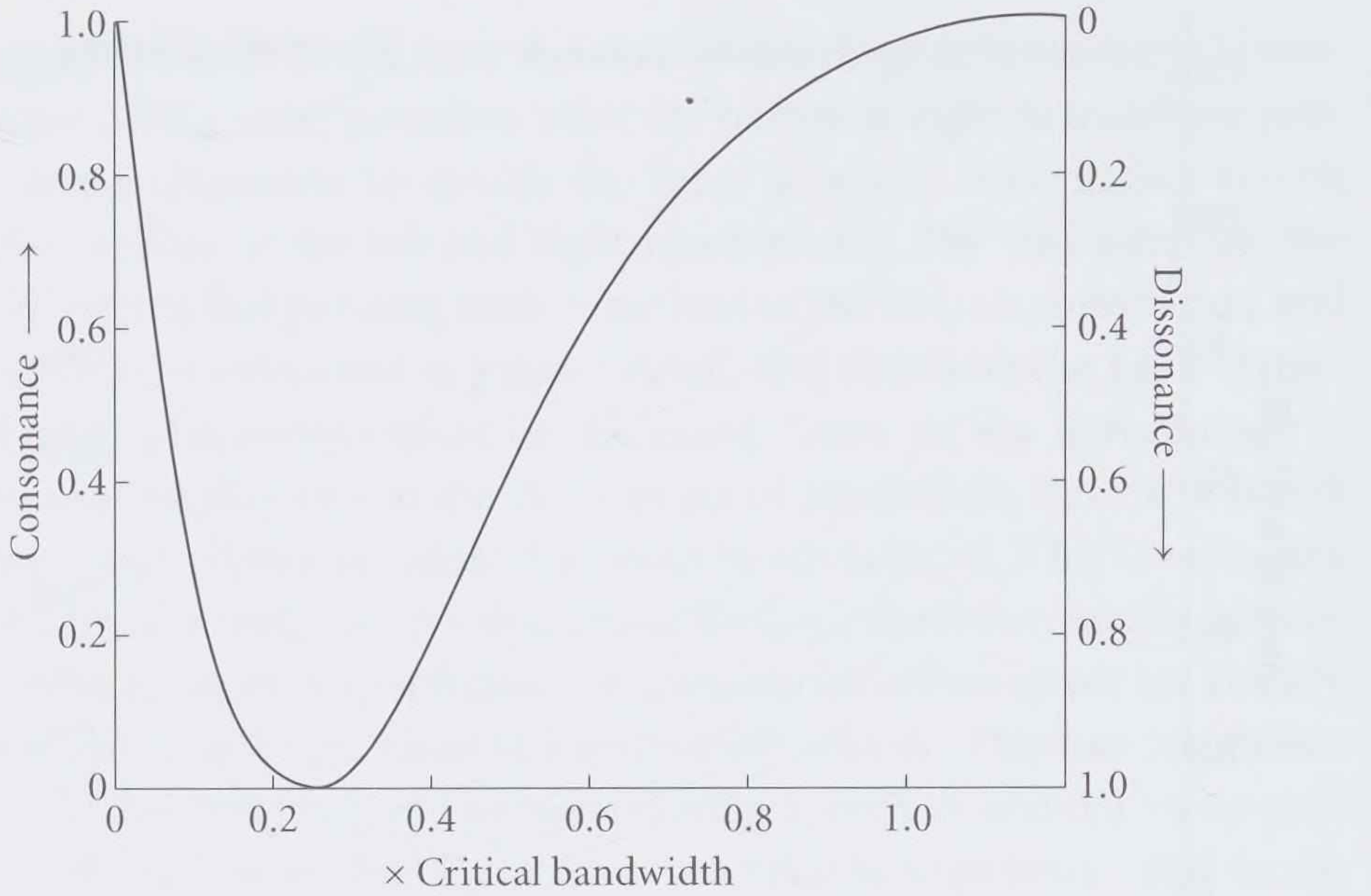


Figure 3-2. The most outstanding empirical fact about interval perception is that small intervals (roughly 1–2 semitones) are perceived as dissonant (after Plomp & Levelt 1965).

not a source of auditory irritation. In fact, the nature of interval perception is somewhat more complex than this simple picture implies, but the basic phenomenon that Plomp and Levelt reported was essentially the "common sense" idea that small (1–2 semitone) intervals are dissonant.

Although Figure 3-2 captures an important aspect of interval perception by indicating that there is a range where unmistakable dissonance is heard, the psychology of the perception of auditory intervals is not quite that simple for three main reasons. The first, also studied by Plomp and Levelt, is that the location of the trough of dissonance in Figure 3-2 will slide left or right somewhat depending on the frequency range. At lower frequencies – say, the first octave at the far left on a piano keyboard – even an interval of three semitones can sound dissonant. In contrast, at high frequencies, a whole-tone interval may strike us as somewhat less dissonant than a similar interval played near middle-C. This influence of mean frequency is known and quantifiable and, in any case, is relatively small within the pitch range of most music and speech (50–3000 Hertz).

The second factor influencing interval perception is the relative amplitude of the two tones. The curve shown in Figure 3-2 – and similar curves with peaks at slightly higher or lower frequencies due to mean frequency – is valid for two tones of equal amplitude. Play one tone or the other a little louder, and the

relative dissonance will change. Both of these effects can be taken into account, and appropriate psychophysical models of dissonance have been constructed in which both frequency and amplitude are included in the basic equations (Kameoka & Kuriyagawa 1969; Sethares 1993, 1998, 1999).

One further effect is, however, more complex – and that is the influence of the so-called higher harmonics (or upper partials). Again, the effect is known and has a firm mathematical basis, but the inclusion of higher harmonics makes the entire issue of interval perception suddenly complex. The complexity is due to the fact that vibrating physical bodies have a tendency to oscillate at more than one frequency. This is illustrated in Figure 3-3, and can be easily understood with a little experimentation with the strings of a guitar or piano. If the two ends of a taught string are immobile, striking it will cause a vibration at the string's "fundamental frequency" (F0), which is a simple function of its length and tautness. Although not easily seen or heard without technical equipment, the same string has other, simultaneous modes of vibration. Higher frequency vibrations occur at integer multiples of the fundamental fre-

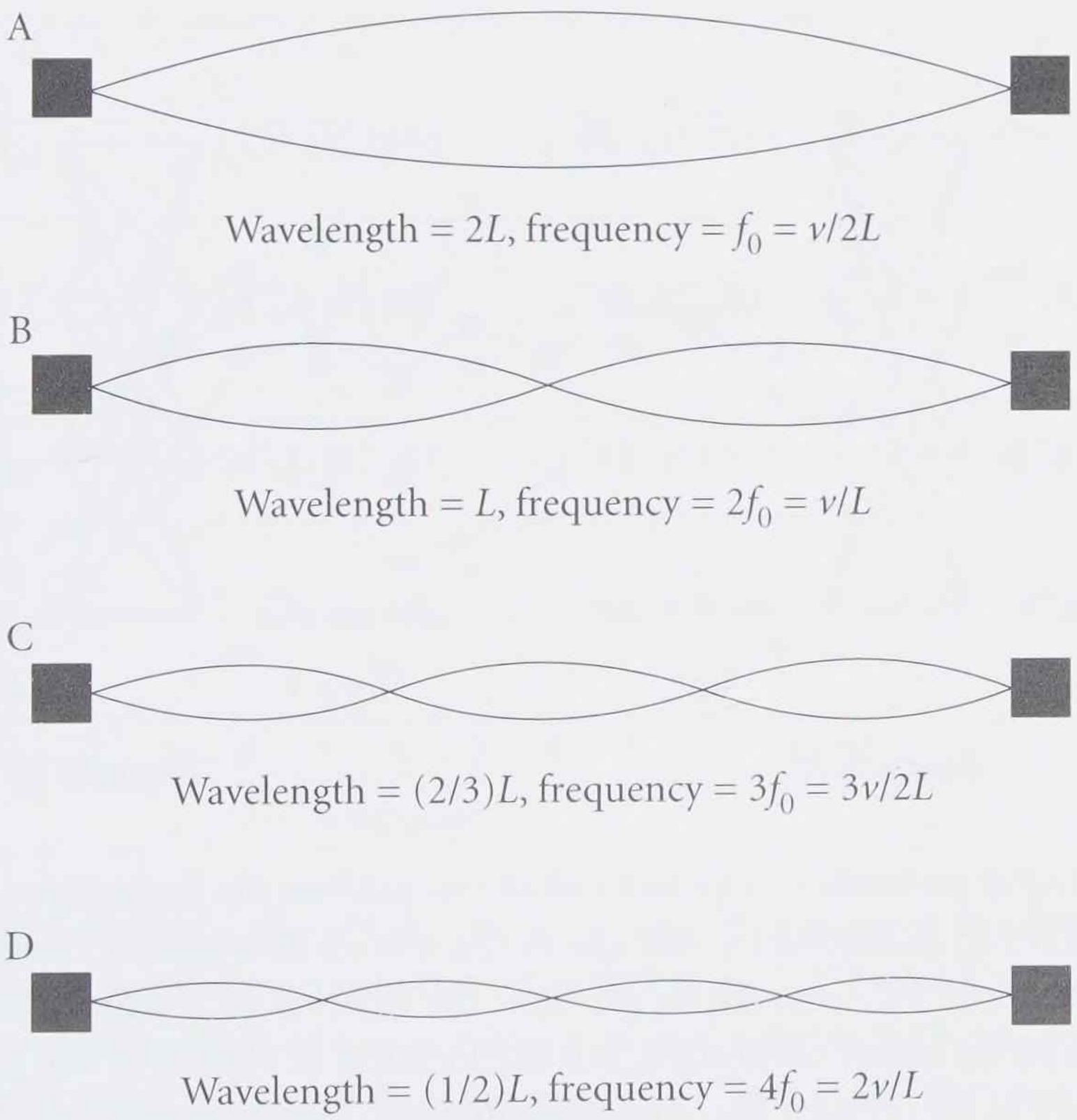


Figure 3-3. A vibrating string will cause sounds at a fundamental frequency and at multiples of that frequency. The intensities of the upper partials (B, C, D) gradually weaken (figure from Pierce 1992).

quency (Figure 3-3). If the F0 is 110 Hz, corresponding to a low A on the piano, then there will also occur vibrations at 220 Hz (F1), 330 Hz (F2), 440 Hz (F3), 550 Hz (F4), and so on. In fact, the amplitudes (intensities) of the higher harmonics drop off rapidly, so that it is difficult to hear more than one or two higher harmonics unless the musical instrument is specifically designed to amplify them, but the presence of higher harmonics is real and an inherent part of all music using instruments that have vibrating strings or tubular resonating bodies.

The upshot of the phenomenon of higher harmonics is that, in listening to music, we rarely hear pure tones consisting of only one frequency and we rarely hear pure intervals consisting of only two distinct fundamental frequencies. Even if only two notes are sounded on the musical instrument, there are at least faint echoes of the higher harmonics – and that means that our evaluation of the pleasantness/unpleasantness or consonance/dissonance of two-tone combinations is actually the evaluation of many tone combinations with a variety of frequencies and amplitudes. A simple example is illustrated in Figure 3-4. There, we see that the "total dissonance" of an interval of 6 semitones (e.g.,

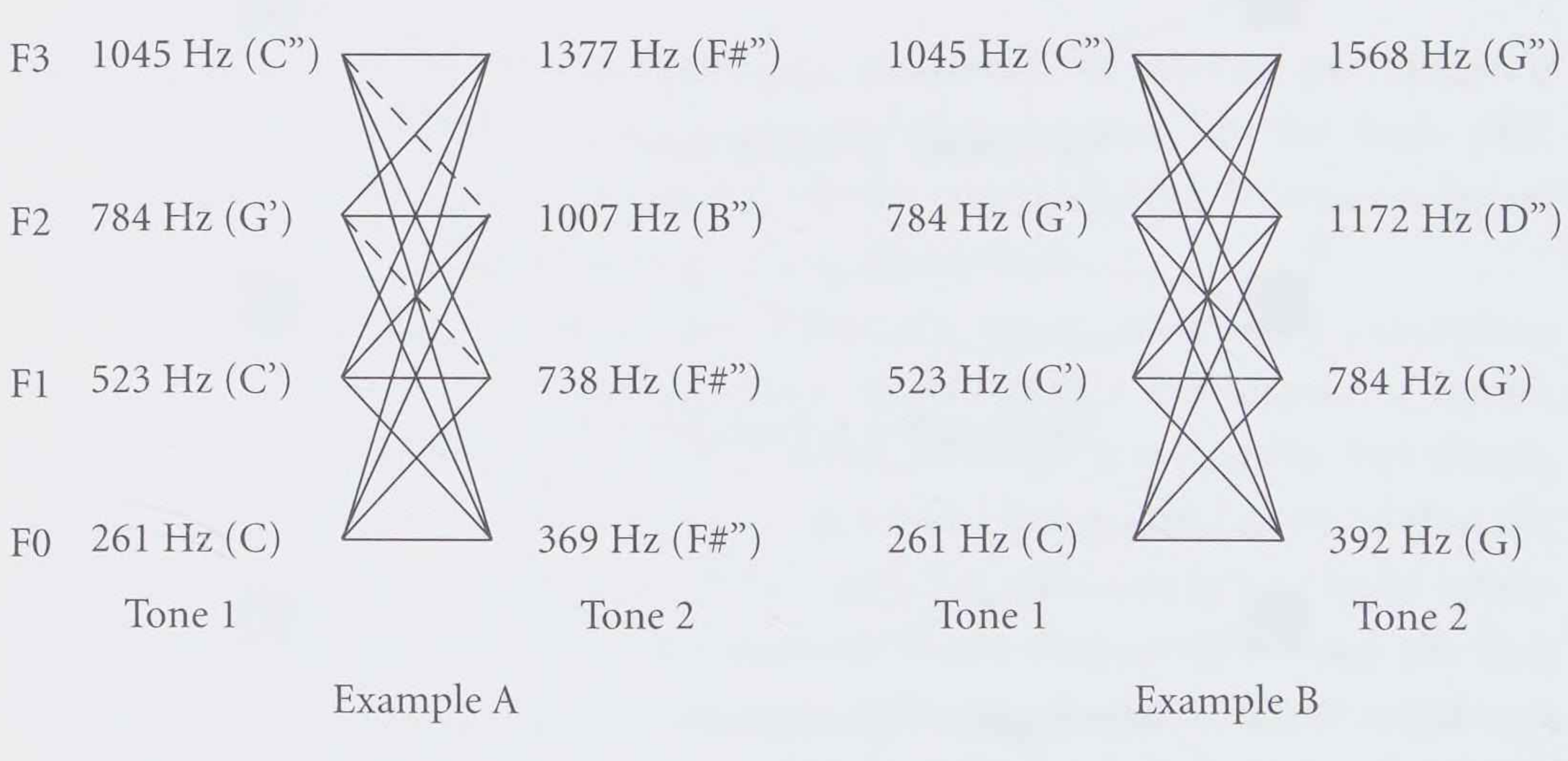


Figure 3-4. In the evaluation of an interval of two pitches, the frequencies of the upper partials can have an influence. In example A, the tritone (6 semitone) interval between the F0s (e.g., C and F#) is much larger than the interval of a minor third. On their own they would be rather consonant, but the location of certain of the upper partials (e.g., 784 and 738 Hz) is such that some dissonance is heard. In example B, the slightly larger interval of a fifth (e.g., C and G) is notably more consonant – a fact that can be explained by the absence of dissonant intervals among the upper partials. Solid lines indicate consonant intervals; broken lines indicate dissonant semitone intervals.

C and F#) is greater than we might otherwise expect, simply because higher harmonics are involved.

As a consequence of the physical reality of higher harmonics in virtually all natural sounds (i.e., not produced as sine waves in a computer), interval evaluation is a somewhat more complex story than what Figure 3-2 suggests. Indeed, experimental results on interval perception normally show a multi-peak dissonance curve, as in Figure 3-5, but even this more complex structure of the dissonance curves can be explained if the presence of higher harmonics is taken into consideration. Most importantly, the troughs at intervals of both 6 and 11 semitones can be understood as results of the dissonance between various combinations of the upper partials, despite the fact that the intervals between the fundamental frequencies are themselves consonant. The modern experts on the psychoacoustics of intervals (Terhardt 1974; Terhardt et al. 1982a, b; Parncutt 1989) make a more subtle case for the role of the upper partials in interval perception – taking into account the masking and salience of individual tones in a tone complex. As discussed below, however, the argument remains

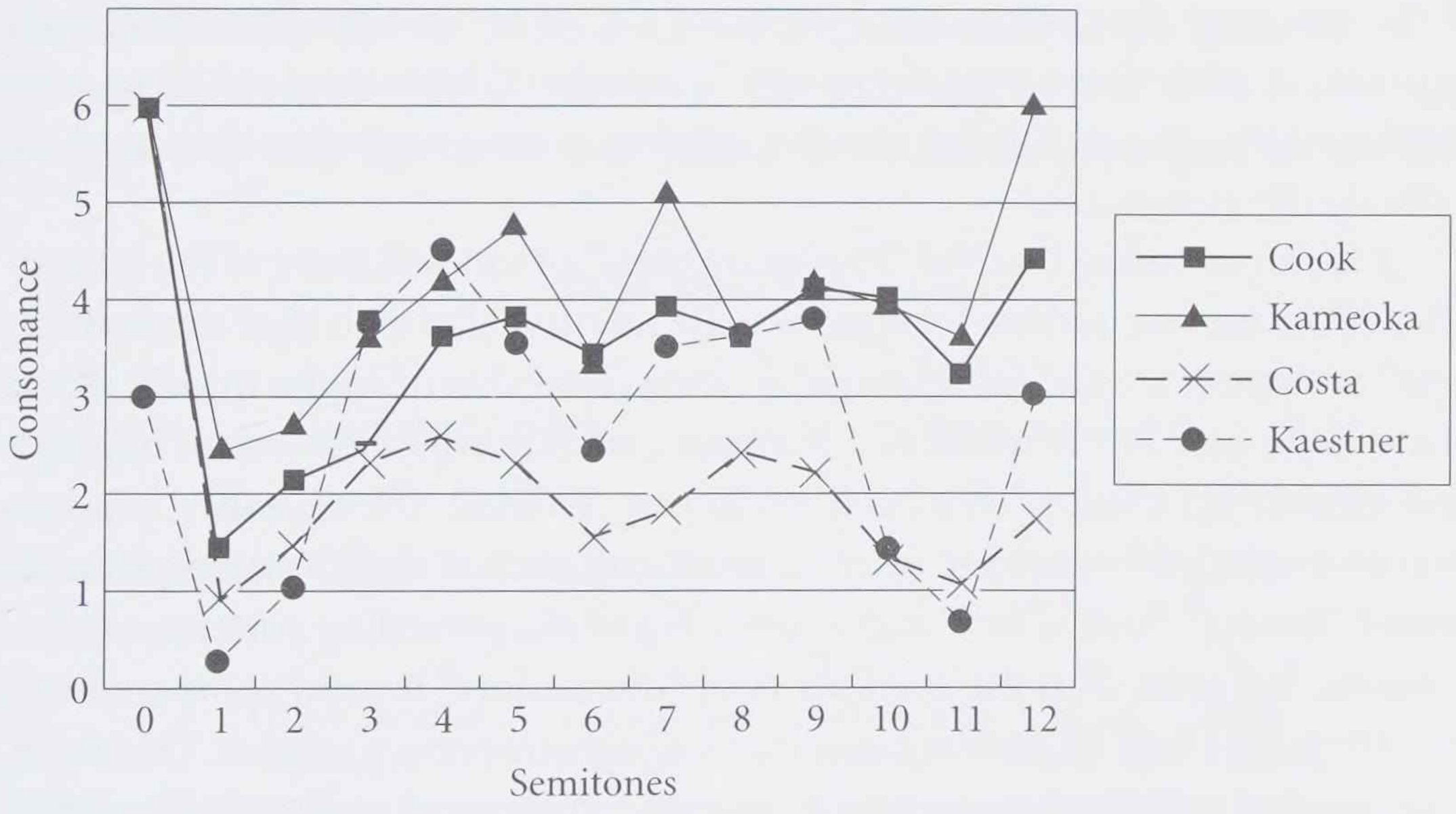


Figure 3-5. Experimental dissonance curves using either pure or complex tones typically show more complex structure than that of Figure 3-2. Notable troughs of dissonance are found at intervals of 1–2 semitones, 6 semitones and 10–11 semitones (data from Kaestner 1909; Kameoka & Kuriyagawa 1969; Costa et al. 2000 and Cook 2000b have been normalized to a common scale). Differences among the curves are presumably due to differences in tone quality (the number of upper partials) and perhaps due to differences in instructions to the subjects to evaluate the intervals in terms of their "consonance," "smoothness," "stability" or "pleasantness."

that the upper partials of pure tones have a decisive influence on the perception of intervals.

Weaker, but still clearly audible is the "unpleasantness" of certain three-tone combinations. Again, the possible influence of lower-level factors – pitch range and pitch intervals – must be considered, but if none of the three tones lies near the upper limits of the physiological range and none of the intervals are 1–2 semitones, then the three-tone combination is a chord that is potentially "harmonious." In such a case, no matter how inappropriate the chord might be in a particular musical context, it would be an exaggeration to use the term "painful" – and even "unpleasant" seems to overstate the case. In the terminology of traditional music theory, such chords are often referred to as unresolved. More precisely, among the many possible three-tone combinations, there are three general classes of chordal phenomena: some chords are inherently "dissonant" because of the presence of a dissonant *interval*, some are inherently stable, final and "resolved," and some are tense, unstable, and referred to as "unresolved." Although chords containing a dissonant interval are "unresolved" as chords, they have an "unresolved" ring primarily because of the two-tone dissonance, not because of the three-tone configuration itself. In contrast, other three-tone chords are "unresolved" because of the three-tone configuration, despite the fact that the intervals contained within the chord are individually consonant.

It is worth noting that the "unpleasantness" of tones that are at the boundaries of the human auditory system is quite intense relative to the "unpleasantness" of dissonant musical intervals or unresolved chords in the middle of the piano keyboard. In the midst of a Baroque concerto where dissonant intervals and unresolved chords have been studiously avoided, the mistaken production of a sustained semitone interval or an augmented chord would probably sound "wrong," "out of key" and destructive of the preceding patterns of consonance, but even then the metaphorical "discomfort" is mild compared to a low-frequency buzz of similar loudness or a high-frequency screech. Of course, if an interval includes tones that are near the extremes of auditory perception, the interval will strike us as extremely unpleasant, but the unpleasantness is not primarily because of the proximity of the tones to one another, but because of their proximity to the limits of the "physiological" range. In such a case, the perception of the interval itself will be overwhelmed by the unpleasantness of one or both of the high- or low-frequency tones. Similarly the stability of a chord will be difficult to evaluate as a chord, if it contains extremely high or low pitches or markedly dissonant intervals.

The classification of all possible combinations of three pitches is illustrated in Figure 3-6.

Among the three-tone chords commonly used in traditional Western classical or popular music, there are four types that are referred to as triads, and have generally well-understood musical effects. These are the resolved (major and minor) triads and the unresolved (augmented and diminished) triads. Because of the effects of chordal inversions – i.e., the same combinations of pitches with different pitches as the lowest tone – the classification becomes more complex, but the underlying structures of the chords are similar. The structure of the four kinds of triads in their three possible configurations is shown in Figure 3-7.

For the musician, the above discussion of chords is already an insult to his intelligence and for the non-musician it may already be testing his patience for dealing with irrelevant details! In fact, however, it is precisely this level of harmonic complexity (three-tone combinations) that must be understood to achieve even a rudimentary understanding of both musical and intonational phenomena. Musically, the effects of four-tone combinations (and still more complex chords) can be (at least partially) understood as the summation of two- and three-tone effects. Moreover, the harmonic structure of music that develops over time can be (at least partially) understood as sequences of rela-

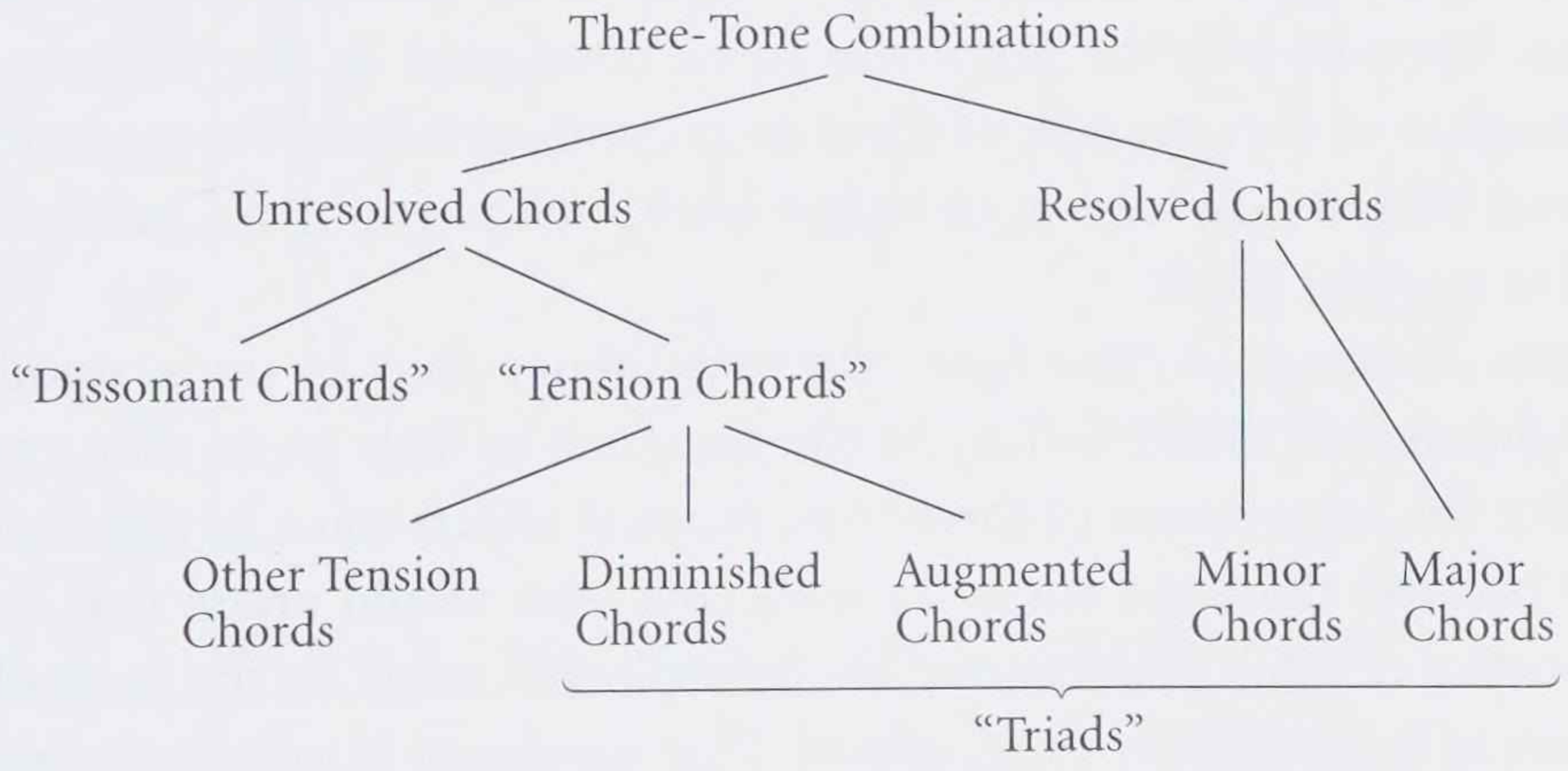


Figure 3-6. The classification of three-tone combinations. The unresolved chords that do not contain dissonant intervals are referred to here as "tension chords." In traditional music theory, the distinction between the (fixed, stable, sonorous and harmonious) "resolved" triads and the (unsettled, unstable, rough and inharmonious) "unresolved" triads is frequently noted, but the "tension" label is not often used with regard to chordal structure. The phenomenon of harmonic tension is, however, important in the discussion that follows.

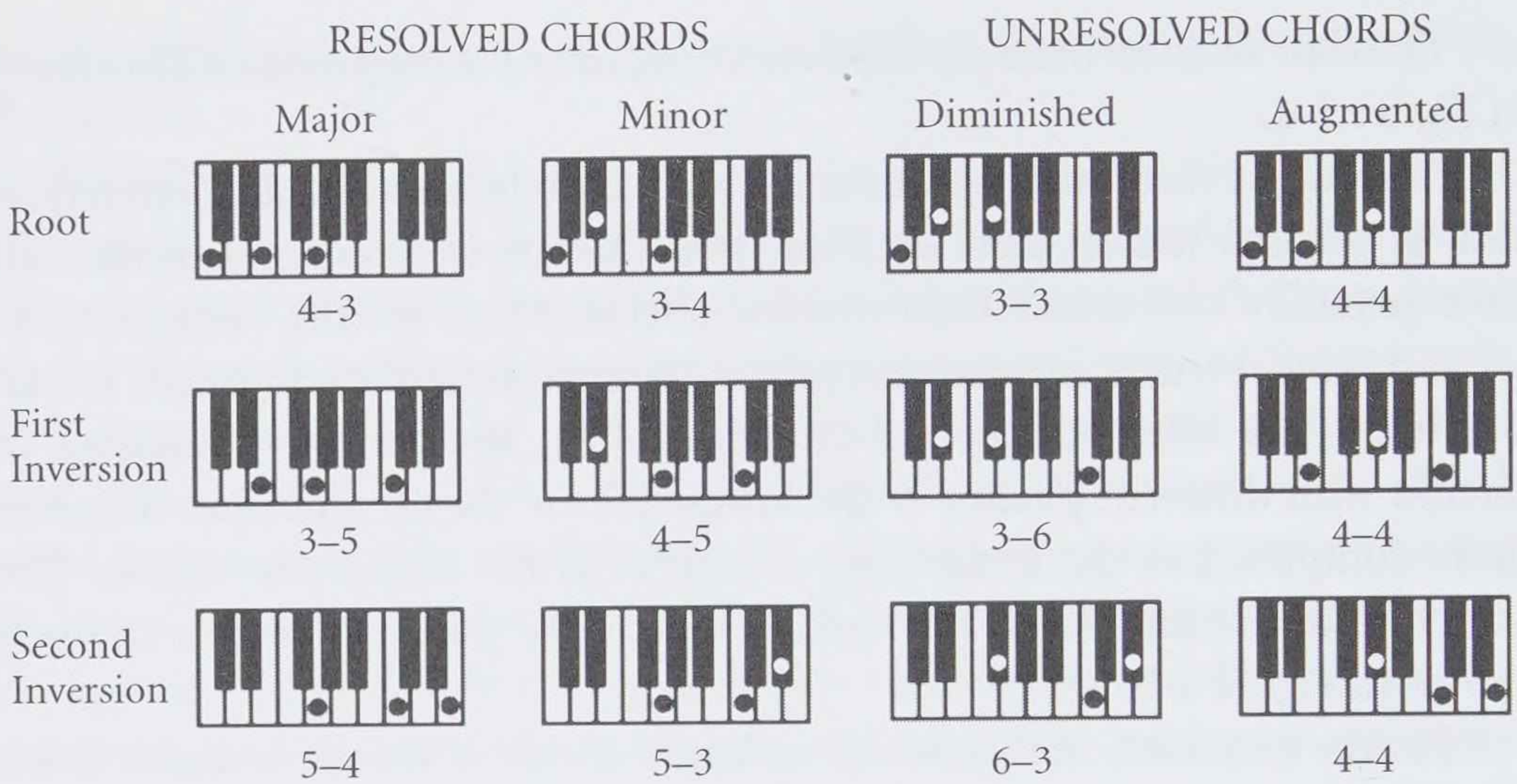


Figure 3-7. The three inversions of the four types of "triad." The musical effects of the inversions are relatively subtle, but the distinction between the resolved major and minor chords versus the unresolved augmented and diminished chords is unambiguous, and perceptible to the ear of the non-musician. Numbers denote the size of the intervals (in semitones) in each chord.

tively simple (two- and three-tone) pitch phenomena. Of course, as additional tones and larger time segments are brought into consideration, treating music as "simply" the summation of these small units becomes more and more dubious. Nevertheless, the argument to be developed in this chapter is that consideration of the structure of three-tone combinations is a necessary (if not sufficient) basis for discussing all higher-level pitch phenomena, particularly as related to musical mode.

Again, the assertion that basic harmony theory must be understood in order to understand music will strike the musician as little more than common sense, but the importance of three-tone musical effects must be stated and defended because there are some theoreticians who would argue that the psychoacoustics of pitch phenomena can be explained *solely* on the basis of combinations of interval (two-tone) effects. That argument is manifestly false with regard to music (as discussed below), and is also untrue with regard to the pitch phenomena of speech, so that anyone with an interest in speech intonation and prosody simply must come to grips with the basic phenomena of harmony theory (three-tone effects).

B. Interval dissonance and consonance

Let us return briefly to the issue of musical intervals and the effects of upper partials. In the late 1960s, Kameoka and Kuriyagawa (1969) undertook an ambitious experimental and theoretical study of dissonance perception. Using a model of interval perception similar to the curve shown in Figure 3-2, they showed how the addition of more and more upper partials changes the theoretical dissonance curve for interval perception (Figure 3-8). What is remarkable about those curves is that, once six upper partials have been included, theoret-

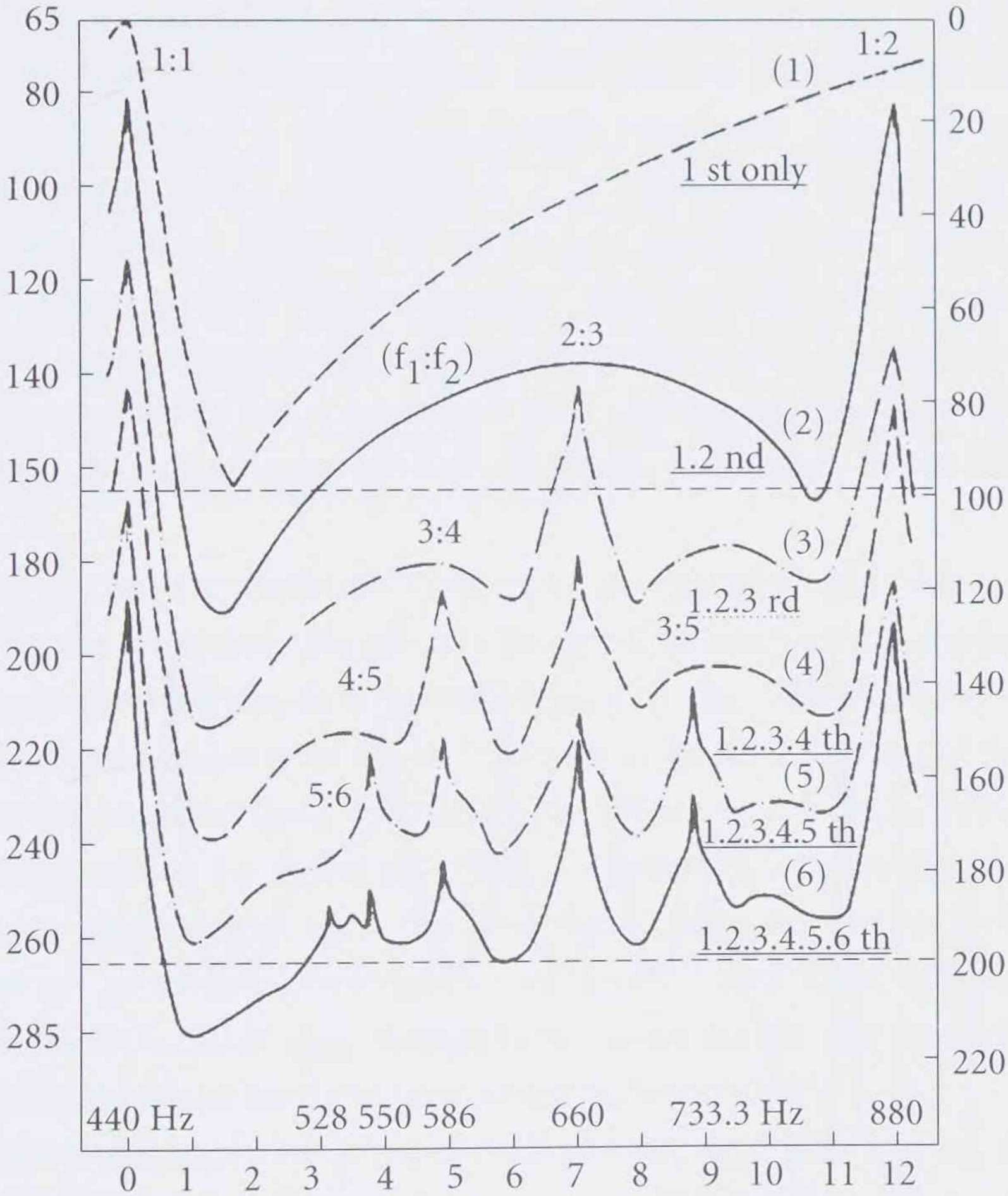


Figure 3-8. The effects of upper partials on the theoretical dissonance curve (from Kameoka & Kuriyagawa 1969). The upper dashed curve is the "model" of interval dissonance. By calculating the effects of that model among all pairs of upper partials, gradually more and more complex dissonance curves are obtained. The bottom curve that includes 6 upper partials shows three large peaks of consonance at unison, at an interval of a fifth and at an interval of one octave. Smaller peaks are also seen at an interval of a fourth, a major sixth and at both the major and minor thirds.

ical peaks of consonance arise at the musically most important intervals of the major and minor scales. Notable also are the two dips at a tritone (six semitones) and eleven semitones – both of which are known dissonances – even though they lie far from the trough of dissonance at 1–2 semitones, as shown in Figure 3-2. Upper partials thus appear to explain why diatonic scales have emerged in various musical cultures and why certain tone combinations are particularly pleasing (Parncutt 1989; Sethares 1999). As shown in Figures 3-8 and 3-9, the "total dissonance" (or alternatively "total consonance") curve becomes gradually more complex as the effects of more and more upper partials are included in the calculation. Small differences in the dissonance curves are produced by the different tuning systems – just, Pythagorean, and equitempered – and by different assumptions concerning the strength of the upper partials, but the basic structures shown in Figures 3-8 and 3-9 are a consequence of the strong dissonance at 1–2 semitones. Other scalar divisions of the octave are of course known and no music theorist would declare that the diatonic scales are the one and only "natural" way of making music, but historically those scales and folk traditions using 5 of the diatonic tones have become dominant throughout the world. That dominance has at least a partial explanation in the set of consonant intervals that the summation of upper partials implies.

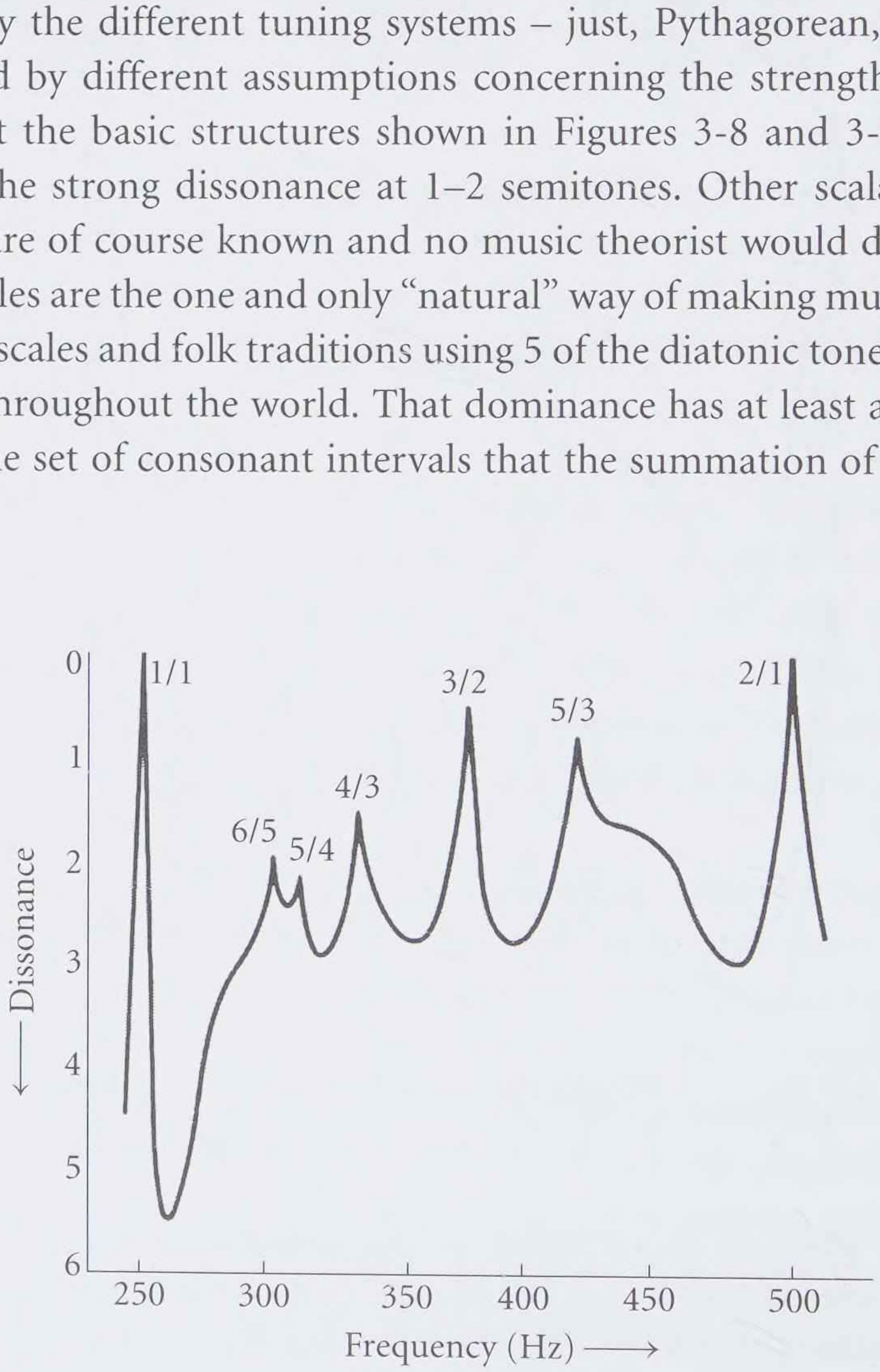


Figure 3-9. Peaks of consonance are seen at intervals corresponding to the division of a vibrating string into integral sections (figure from Pierce 1992).

What is of interest from the point of view of music history, music theory and ethnomusicology is the location of the peaks in the dissonance curve. Specifically:

1. The maximal peaks of consonance correspond to unison and the octave. Singing in unison or at an interval of an octave is the easiest, most common form of spontaneous singing for two voices.
2. The next highest consonance peak corresponds to an interval of a fifth, which is the next easiest interval to sing and, historically, was the next two-part interval for chanting in early church music (ca. 1200)
3. The next highest peaks correspond to intervals of a fourth and a sixth. These were also recognized early on as "musical" and therefore suitable for use in church hymns (ca. 1300).
4. The next two peaks of consonance correspond to intervals of a major third and a minor third. Although different tuning systems place them at slightly different pitch locations, these intervals play important roles in defining mode. Eventually (ca. 1400), both intervals were acknowledged as musical, the major third prior to the minor third. Together with the idea of building intervals on top of intervals, all of the twelve tones of the so-called chromatic scales can then be produced.

It is thus clear that the theoretical dissonance curves might play a role in explaining why the spectrum of auditory frequencies has historically been divided into 12 approximately equal intervals, and why the vast majority of musical cultures have used subsets of these 12 tones (usually employing 5- or 7-tone scales). The explanation lies primarily in the fact that, if human beings find 1–2 semitone intervals to be dissonant (for, presumably, some reason due to human auditory physiology), then the complex dissonance curves in Figures 3-8 and 3-9 are an inevitable consequence of the fact that most vibrating objects (and particularly tubular and stringed instruments) vibrate at various higher frequencies, in addition to their fundamental frequencies.

The above summary of interval consonance and dissonance, and its implications with regard to musical scales, merely scratches the surface of the psychoacoustics of intervals, but there are further phenomena that need to be addressed – specifically, three-tone harmony (see Grout 1973; Parncutt 1989; Pierce 1992; Jourdain 1997; Roederer 1995; Scruton 1997; and Huron 1991, 2001, for further discussion).

C. Harmonic tension and resolution

The major and minor chords are known to have characteristic, recognizable moods, associated with primary emotions (e.g., Hevner 1936; Wedin 1972), even in subjects as young as three years old (Trainor & Trehub 1992). Most musicians will insist that the "ring" of three-tone combinations is more than the sum of the "ring" of the intervals contained therein. The musician's intuition here is not irrelevant, but we must consider the possibility that the feeling inherent to simple three-tone harmonies might be simply a consequence of intervals and their upper partials. In other words, is it true that the perception of three-tone harmony is something more than and distinct from the perception of two-tone interval effects? Particularly since it is apparent that (barely audible) upper partials influence the perception of simple intervals, it is conceivable that the most parsimonious explanation for the resolvedness or unresolvedness of some chords, and the bright or dark mood of other chords might be that three-tone combinations produce effects not evident in the underlying intervals due again to the contributions of (barely audible) upper partials. If that were the case, then the difference between resolved and unresolved chords and between major and minor chords should also lie simply in differences in the patterns of dissonant upper partials.

It must be stated that such a view has gained some popularity in the field of music psychology, but parsimony runs into at least two sets of stubborn facts that suggest otherwise: (i) calculations of the effects of upper partials indicate that intervals cannot explain either the difference between resolved and unresolved chords or the difference between major and minor chords, and (ii) brain-imaging studies show that the brain is doing rather different things when it hears dissonant intervals and when it hears inharmonious chords. Let us examine these issues in some detail.

Harmonies containing three or more simultaneous tones are an important component of "Western" music (and arguably an implicit part of all music traditions due to the effects of upper partials). In spite of the musical importance of multiple pitch combinations, a generally-accepted psychophysical explanation of harmony has not yet been achieved, and most discussions have focused on the harmonic regularities within specific musical genres (e.g., Bharucha 1987; Krumhansl 1993; Regnault et al. 2001). Specifically, while the importance of the acoustics of intervals (including the effects of upper partials) is widely acknowledged, the acoustics of three-tone chords (i.e., harmony) is usually treated as a complex issue of *interval* acoustics and learning ("cognition", "musical traditions") – but without consideration of the role of harmony as a

three-tone acoustical phenomenon. Here, a contrary approach will be outlined in which a general psychophysical model of three-tone harmony is defined, and results from both behavioral and brain-imaging experiments are discussed.

Behavioral experiments

The perception of the dissonance of two tones whose frequencies are near to one another has been thought since Helmholtz (1877) to be due to physiological factors inherent to the human auditory system. In contrast, the degree of consonance of intervals that are not "physiologically-dissonant" is thought to depend at least partly on acquired musical factors. The question arises: in the perception of the "harmoniousness" ("sonority" or "tonalness") of chords consisting of three or more tones, what innate and acquired factors are involved? To answer this question, we have run several behavioral experiments using two- and three-tone sine-wave sounds as stimuli. The objective was to replicate previous findings and to establish a quantitative relationship between interval and chord perception in non-musician subjects.

All subjects for the behavioral experiments were undergraduates (male and female, mean age of 19 years) with normal hearing in both ears. Many had undergone some musical training (principally, the piano) as young children, but none were actively rehearsing musicians. In the first experiment, 26 subjects evaluated two-tone intervals on a six-point scale ranging from very dissonant (1) to very consonant (6), by pressing the appropriate key on the numberpad of a computer keyboard with the right-hand. In previous experiments, we have found no effects of ear or hand usage (data not shown).

The auditory stimuli were computer-generated sine-waves at the frequencies of a 12-tone equal tempered scale (110–440 Hz), sampled at 44100 Hz, stored and replayed as 16-bit sounds on headphones in a quiet room. The sounds were 12 intervals constructed using the (remarkable and remarkably free) Praat™shareware (Boersma & Weenink 2002) and presented 6 times in random order. Subject responses and response times were recorded.

The results of Experiment One were shown in Figure 3-5. There was a significant influence of interval magnitude, and comparisons among the intervals showed several significant differences among neighboring intervals. The decreased consonance from the interval of a fourth to that of the tritone is a trend, whereas the rise in consonance from the tritone to the interval of a fifth was significant. At the end of the curve, both the decrease from an interval of 10 semitones to 11 and the increase from 11 semitones to an octave are highly significant.

All subjects evaluated the semitone and whole-tone intervals to be the most dissonant, but the additional substructure of the dissonance curve suggests that even in the perception of pure sine-wave tones, a mental association with upper partials influences the perception of tones that lack such partials. Even though the stimuli did not contain upper partials, the dips at both 6 and 11 semitones can be explained on the basis of "associated" upper partials (e.g., Kameoka & Kuriyagawa 1969; Sethares 1999). In other words, the entire pattern of perception of interval dissonance might be coherently explained by means of appropriate, post hoc selection of upper partials and their amplitudes (see Figure 3-8). A detailed, quantitative explanation of group effects and individual differences in interval perception based on the variable strength of upper partials is an interesting topic in its own right, but it is of still greater interest to see what role upper partials play in the evaluation of the harmoniousness of three-tone chords.

Chord perception in a 12-tone scale

Previous work has shown that chordal perception can be partially explained on the basis of the constituent musical intervals (Helmholtz 1877; Pierce 1992), but the phenomena of chordal "resolution" and the moods of major and minor harmonies cannot be easily explained solely on the basis of interval effects. Within the framework of the diatonic scales underlying most popular and classical music, certain chordal harmonies are perceived as "stable" or "completed" – with a sense of "finality" or "resolution" that is not felt with other chords. Historical and cultural factors certainly contribute to the sense of restlessness or otherwise of various pitch combinations, but it is likely that there are also properties inherent to the multitone chords themselves that contribute to our perception of tonal stability/instability.

To test this possibility, again pure sine waves (110--440 Hz) from a 12-tone scale were used, this time in combinations of three tones (without higher harmonics). A total of 26 undergraduate subjects (male and female, mean age 20) evaluated the "harmoniousness" of each chord presented binaurally for 1.5 seconds. A total of 108 chords of 6 distinct types were prepared using three tones between A3 and A5: major and minor chords, augmented and diminished chords, and chords containing one dissonant (semitone or whole-tone) interval. The chords were presented twice each in random order and responses between 1 and 6 were recorded.

The results are shown in Figure 3-10. A highly significant effect of chordal structure was found in comparisons of the resolved, unresolved and dissonant

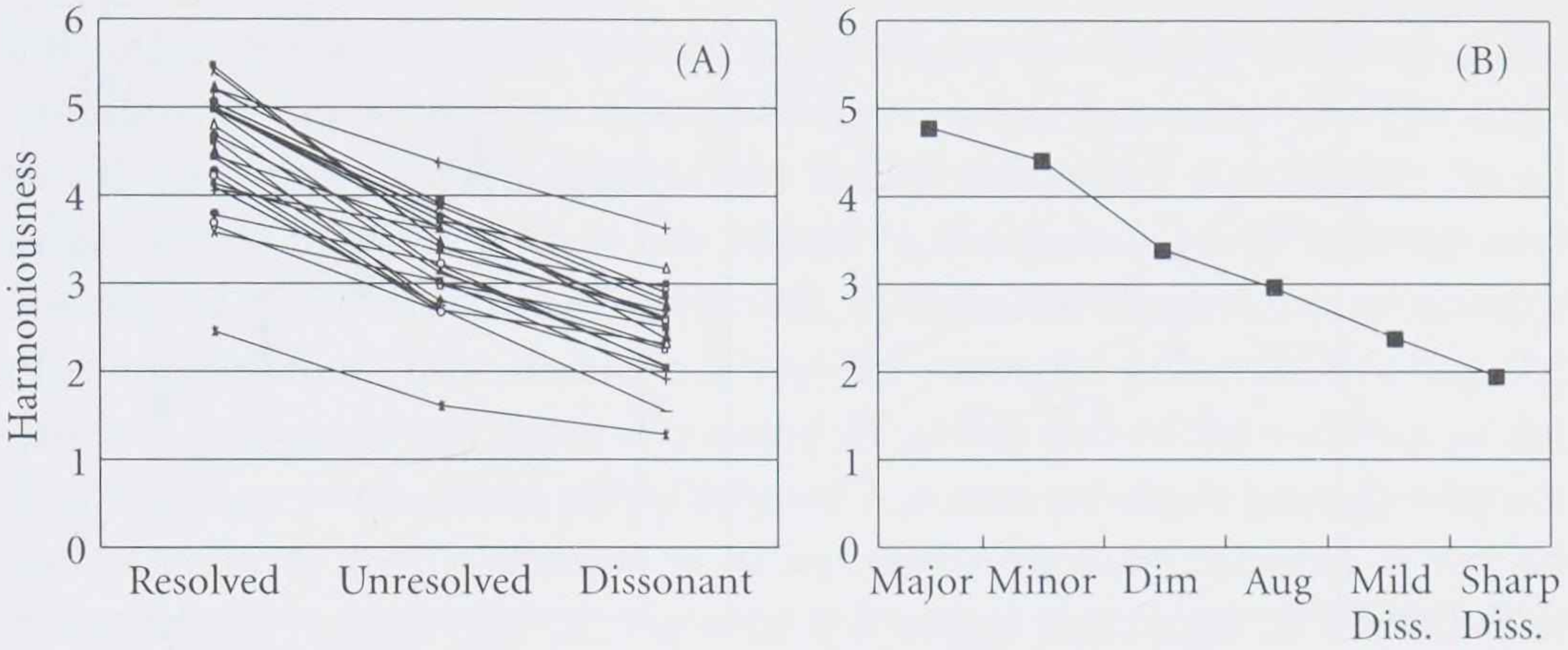


Figure 3-10. The evaluation of chordal harmony using tones from a 12-tone scale. (A) Although there were clear individual differences in the overall level of harmoniousness, these non-musicians had no difficulty in distinguishing among the three types of chords. Lines connect the mean scores of the same subjects. (B) Major chords were perceived as more harmonious than minor chords, diminished chords as more harmonious than augmented chords, and chords containing a whole-tone dissonance were more harmonious than those with a semitone dissonance (Cook 1999a).

chords ($p < 0.001$ in every combination). As expected, the resolved chords were evaluated as the most harmonious, the dissonant chords as least harmonious, and the unresolved chords evaluated in between. This overall pattern was obtained for all 26 subjects, without exception.

Similar results have been reported in the psychology literature. As shown in Figure 3-11, musicians more clearly distinguish among the triads than non-musicians, and root position chords are evaluated as having greater consonance than similar chords in first or second inversion, but the distinction between resolved and unresolved chords appears to be robust.

The results of these experiments are not surprising, but the traditional explanation would be simply the assertion of "cognitive" influences. All subjects – musically trained or untrained – have experienced countless musical stimuli, the vast majority of which are diatonic. The regularities inherent to such music imply that the evaluation of harmoniousness would be as shown in Figures 3-10 and 3-11, simply as a consequence of passive learning. The general idea that learning and cultural experiences influence perception is of course plausible, but before we attribute everything to cognition, there are some further issues to address.

Although there are limitations in distinguishing between cognitive and psychoacoustic effects on the basis of behavioral experiments, the pattern of

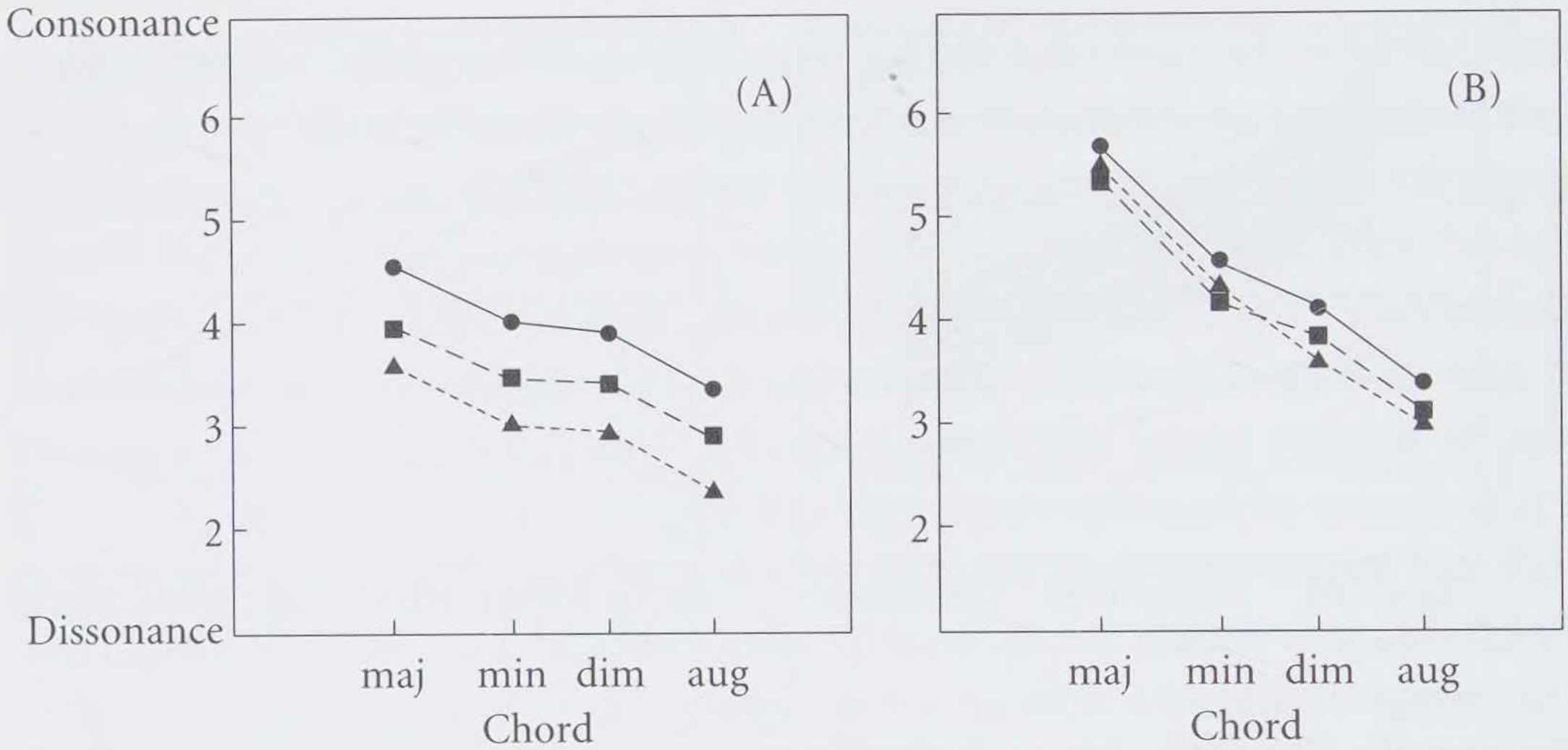


Figure 3-11. Mean ratings for the three different positions of the four triads for untrained (A) and highly-trained (B) subjects, using a 7-level rating scale from "very dissonant" (1) to "very consonant" (7). Circles indicate ratings of the chords in root position, squares show first inversion and triangles show second inversion chord ratings (from Roberts 1986).

experimental results can be examined for psychoacoustic factors using statistical techniques. The first factor studied using multiple regression analysis was the dissonance of small intervals in three-tone chords. That is, did the presence of a dissonant semitone or whole-tone interval in the chord have a distinct influence on subject responses? The second set of factors was the total dissonance of the three intervals in each of the three-tone chords. That is, does the total theoretical dissonance within the chord predict the subject responses? Because a clear influence of upper partials had been found in the interval experiment, theoretical calculations including varying numbers of upper partials (1, 3, and 6) for each chord were made using the model of Sethares (1993, 1999). The empirical interval dissonance curve (Cook 2000b; Figure 3-5) was also used to calculate the total dissonance in the chords. These various "total dissonance" factors were examined in the regression analysis to determine the role of specifically *interval* effects on the overall harmoniousness of the chords. Finally, a third factor related to harmonic tension was examined. Tension was defined as the effect produced in three-tone chords containing two equivalent intervals. In other words, does a chordal structure that entails tension – as distinct from the interval effects – have an effect of the perception of chordal harmony?

Because the auditory nature of each multi-sine-wave stimulus is fully known and easily characterized in terms of its constituent components, mul-

tiple regression analysis can be used to determine the relative influence of the components – in isolation or in combination – on subject responses. The analysis showed that harmoniousness was indeed influenced significantly by all three factors, independently – two related to the dissonance of intervals and one related to the structure of the chords. Specifically, the presence or absence of semitone/whole-tone dissonance was the strongest factor influencing the subjects' responses and alone accounted for about half of the variance in the data. In other words, when subjects heard semitone or whole-tone dissonance within a chord, their evaluation of its harmoniousness plummeted. A second factor related to intervals, i.e., the total theoretical dissonance of the intervals contained in the chord, was also found to be a significant factor. (This factor was significant regardless of the number of upper partials used for the calculation. Moreover, similar analysis using the total dissonance as determined from the empirical dissonance curve in the interval experiment again showed a significant contribution of total interval dissonance to chord perception.) It is noteworthy, however, that, when simple regression was done using only the total theoretical dissonance of the intervals, this factor alone accounted for less than 10% of the variance in the data. That is to say, the total dissonance of the intervals contained within chords plays a statistically significant role in influencing the harmoniousness of the chords, but – contrary to a purely interval-based explanation of harmony, "total interval dissonance" explains only a small part of the empirical data. Finally, and most interestingly, the presence of three-tone chordal "tension" was found to play a significant role. This factor indicates that when subjects heard a three-tone chord containing two intervals of the same magnitude, their evaluation of its harmoniousness decreased.

These results indicate that *the chordal tension of certain three-tone chords is perceived by normal listeners in a manner similar to, but distinct from, the perception of the dissonance of two-tone intervals.* If a chord contains a dissonant interval, that alone will strongly influence subject evaluations of its harmoniousness, but even when the total interval dissonance is mild (calculated in a variety of ways), the chord will be evaluated as rather inharmonious if it has a chordal structure including two equivalent intervals. It is concluded that, in so far as the chordal tension of the unresolved chords cannot be explained as the summation of "two-body" interval effects, this "three-body" chordal effect may itself be a psychological Gestalt: the "instability" or unresolved nature of augmented, diminished and similar chords may indicate the inherent difficulty in finding a subgrouping of the tones. The configuration of tones for which there is a "natural grouping" (and therefore tonal resolution) undoubtedly depends upon factors acquired from the given musical culture (and indeed the loca-

tion of tones at which maximal interval consonance is perceived is also likely to be influenced by musical experience), but the need to resolve chordal tension (i.e., move away from three-tone combinations that are equally spaced) may be a feature as common to the perception of polyphonic music as is the need to resolve interval dissonance. Other cognitive explanations might be pursued here, but stronger evidence can be obtained from brain-imaging.

A functional MRI experiment

On the basis of the behavioral experiments discussed above, it was concluded that the perception of chordal tension is distinct from the perception of interval dissonance. That conclusion was, however, drawn primarily from the statistical significance of factors in the analyses. To test the validity of that conclusion more directly, a brain-imaging experiment was designed (Cook et al. 2001, 2002a, b) to allow conclusions concerning whether or not there are differences in brain activation during the evaluation of the "harmoniousness" of resolved, unresolved and dissonant chords.

Six right-handed subjects (5 males and 1 female between 20 and 50 years of age) participated in the fMRI study. None were musicians, but all were familiar with both Japanese and Western musical styles. Stimuli consisted of six chordal types (major, minor, augmented, diminished, and dissonant chords containing a dissonant interval of one or two semitones) and a control condition consisting of white noise. For each chordal type there were 20 stimuli sampled at 44100 Hz and stored as 16-bit files. The stimuli were three-seconds in duration and consisted of three pure sine waves at frequencies between A3 and A5 in a 12-tone scale.

A block design was used in which 5 chords from the same condition were presented at a rate of 1 every 6-seconds (an inter-stimulus interval of 3 seconds). This was followed by presentation of 5 white-noise control stimuli. Two blocks of each of the 6 chord types were randomly presented for each of the 2 sessions. Randomization was constrained so that all chord types were presented once before a chord type could be repeated. In addition, the same chord type block could not be presented sequentially. Each session was approximately twelve minutes. The task was to judge the harmoniousness of the stimuli. To record subject responses to the chord stimuli, a 5-button glove was fitted to the right hand. Subjects were required to respond with one button press per stimulus during the three second period before the presentation of the next stimulus, with the thumb (1) indicating low harmoniousness and the little finger (5) indicating high harmoniousness. Subjects were not informed of the block de-

sign and responded to each stimulus in each block. For the white-noise control stimuli subjects were required to respond "at random" with a button press. All subjects had practice with the various types of stimuli prior to fMRI scanning.

Brain imaging was performed using a 1.5 Tesla Marconi Magnex Eclipse scanner using an interleaved sequence. First, high-resolution anatomical T2 weighted images were acquired using a fast spin echo sequence. These scans consisted of 50 contiguous axial slices with a $0.75 \times 0.75 \times 3$ mm voxel resolution covering the cortex and cerebellum. Secondly, functional T2* weighted images were acquired using a gradient echo-planar imaging sequence (echo time, 55 ms; repetition time, 6000 ms; flip angle, 90°). A total of 50 contiguous axial slices were acquired with a $3 \times 3 \times 3$ mm voxel resolution.

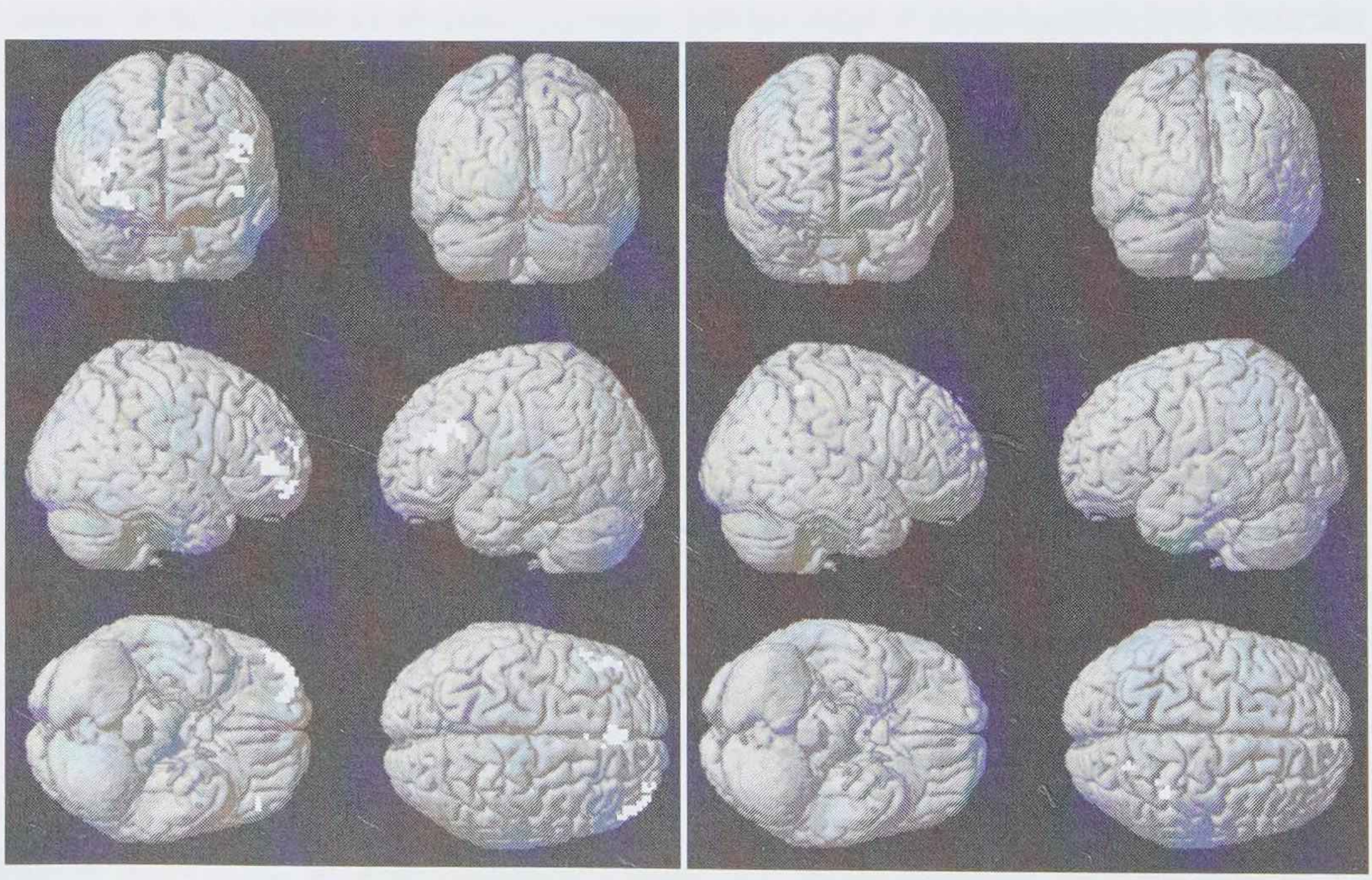
Regional brain activity for the various chord conditions was assessed on a voxel-by-voxel basis. A fixed effect model was employed (the data from all subjects were analyzed together). The data were modeled using a box-car function convolved with the hemodynamic response function. In addition, global normalization and grand mean scaling were carried out. The six chord conditions were paired in the following manner to make three conditions upon which the analysis was carried out: (minor and major = resolved chords), (augmented and diminished = unresolved chords), and (one-semitone dissonant and two-semitone dissonant chords = dissonant chords). Significance for the unresolved and dissonant conditions contrasted with the control white-noise condition was assessed using the conservative random field theory correction for multiple comparisons (corrected $T = 4.84$, $p < 0.05$, $df = 1356$). In order to investigate the brain regions involved relative to other conditions these contrasts were inclusively masked ($p < 0.05$) by the conjunction of the statistical comparison of the condition at hand with the two other conditions. The two analyses of interest were unresolved inclusively masked by the conjunction of (unresolved-dissonant and unresolved-resolved) and dissonant inclusively masked by the conjunction of (dissonant-unresolved and dissonant-resolved).

Behaviorally, the 6 subjects showed the same pattern of chord perception as shown in Figure 3-10. That is, the resolved chords were evaluated as the most harmonious (3.73 ± 0.55), followed by the unresolved chords (3.18 ± 0.39) and then the dissonant chords (2.34 ± 0.49).

Brain activity distinctive of the unresolved (relative to the resolved and dissonant) condition and dissonant (relative to the resolved and unresolved) condition was determined using inclusive masking. Figure 3-12A shows the results for unresolved inclusively masked by the conjunction of (unresolved-dissonant and unresolved-resolved). Distinct activity for the unresolved condition relative to the other conditions is located in the right dorso-frontal area

(Brodmann Area, BA 10 some extending into BA 11 and BA 46), the left dorsolateral prefrontal area BA 46, left inferior prefrontal area BA 47, right cerebellum, and the anterior cingulate BA 32. Figure 3-12B shows the results for dissonant inclusively masked by the conjunction of (dissonant-unresolved and dissonant-resolved). Distinct activity for the dissonant condition relative to the other conditions is located primarily along the right inferior parietal lobule BA 40 and the parietal lobe BA 7.

A comparison of the brain activation during the dissonance and unresolved chord conditions in relation to the resolved chord condition directly indicates the difference in cerebral activity underlying the perception of dissonance versus chordal tension (Figure 3-12). If the perception of the "tension" of unresolved chords were simply a consequence of the dissonance of some of the intervals among the upper partials of the fundamental frequencies, then the pattern of brain activation in response to tension chords would be expected to be similar to the activation in response to dissonant chords. The dissonance of the tension chords would be expected to be weaker than the relatively sharp dissonance of the dissonant chords, but if hearing chordal disharmony is fun-



A B

Figure 3-12. The brain activity distinctive of the unresolved chord condition (A) and the dissonant chord condition (B). All lightly-shaded pixels indicate activity significantly ($p < 0.05$) above that of the other conditions (Cook et al. 2001, 2002a, b).

damentally the same as hearing interval dissonance, the same brain regions should be active. This was found not to be the case.

Clearly, the brain activation during the unresolved chords was distinct from that during dissonant chords in both frontal and parietal areas (at a significance level of $p < 0.05$). In addition to any two-body (interval) effects that are perceived in polytonal music, there are apparently three-body (chordal) effects that are more than the sum of their interval components. Unresolved chordal stimuli show distinctive activity predominantly in right dorso-frontal cortex, left dorso-lateral prefrontal cortex and the anterior cingulate, whereas, dissonant stimuli show distinctive activity predominantly in right parietal cortex.

Previous behavioral and brain-imaging studies[1] have also implicated the right hemisphere, particularly the right frontal lobe, in the perception of patterns of pitch (e.g., Zatorre et al. 1994; Binder et al. 1997). Zatorre has concluded that "frontal lobe mechanisms [are involved] in effecting pitch comparisons, with a particularly important contribution from right-frontal regions" (Zatorre 2001:201).

Harmoniousness is not the same as consonance

The results of the behavioral and brain-imaging experiments are consistent in suggesting that chordal harmoniousness is qualitatively not the same as (or simple summation of) interval consonance. Despite the fact that many discussions of harmony in the music perception literature use the words dissonance and inharmoniousness as synonyms (and tonality, euphony, sonority, harmoniousness and consonance as synonyms), the empirical data discussed above indicate that there are two distinct factors that determine the sonority of three-tone combinations. The first is, to be sure, an underlying interval effect, but the second is an effect that is due to chordal structure itself and cannot be reduced to interval effects. This three-tone factor is properly referred to as harmoniousness or inharmoniousness.

These findings are in line with musical common sense and clearly indicate that a distinction needs to be drawn between two-body and three-body tonal effects. There are, moreover, two theoretical arguments for considering three-tone "chordal tension" to be an important component of a psychophysical theory of harmony. The first can be stated simply as follows. *The summation of interval dissonance (with or without upper partials) cannot explain the perceptual character of various three-tone chords.* The simplest examples illustrating the problematical nature of interval dissonance as an explanatory principle for harmony are the major/minor and augmented/diminished chords (Table 3-1).

Table 3-1. A comparison of the theoretical "harmoniousness" of the 12 major, minor, augmented and diminished chords in various configurations (see also Figure 3-7). The values shown were calculated considering only interval effects, with the chords in root and inverted configurations (with the tones of each chord adjusted to the same mean frequency of 261.6 Hz). The theoretical dissonance (Sethares 1993, 1999) has been calculated with 1, 3 and 5 upper partials, (relative amplitudes decreasing as 1.00, 0.50, 0.33, 0.25, 0.20). The ranking of the chords from the most harmoniousness to the least harmoniousness is shown in brackets for each of the columns. Noteworthy is the *inability* of interval dissonance alone to distinguish between the resolved and unresolved chords. See Appendix 2 for further discussion.

Chord type	Configuration	1	3	5
Major	Root	8.04 [10]	5.42 [9]	4.51 [6]
Major	1st Inversion	4.30 [5]	5.23 [8]	4.65 [10]
Major	2nd Inversion	0.72 [1]	2.59 [1]	2.35 [1]
Minor	Root	8.53 [11]	5.50 [11]	4.71 [11]
Minor	1st Inversion	2.52 [4]	3.83 [4]	3.58 [3]
Minor	2nd Inversion	1.80 [3]	3.72 [3]	3.84 [4]
Diminished	Root	11.32 [12]	5.79 [12]	5.11 [12]
Diminished	1st Inversion	4.39 [6]	5.47 [10]	4.11 [5]
Diminished	2nd Inversion	1.64 [2]	3.10 [2]	3.17 [2]
Augmented	Root	5.59 [7]	5.09 [5]	4.57 [7]
Augmented	1st Inversion	5.59 [7]	5.09 [5]	4.57 [7]
Augmented	2nd Inversion	5.59 [7]	5.09 [5]	4.57 [7]

Whether in root position or in inverted form, the intervals in all such chords are consonant minor-thirds or greater. Summation of their dissonance using models of interval dissonance produces values that do not distinguish among these chord types and yet, perceptually, the difference between resolved and unresolved chords (and between the major and minor chords) is relatively clear, even to the non-musician's ear.

To demonstrate this fact, the theoretical values for the total dissonance of these chords were calculated using Sethares's model of interval dissonance (1999). (See Appendix 2 for further details.) Calculations were done using various numbers of upper partials. As clearly seen in Table 3-1, the number of higher harmonics has an influence on the ranking of the chords, but neither the resolved/unresolved distinction nor the major/minor distinction – the two concepts that lie at the *heart* of traditional harmony theory – can be explained on the basis of interval dissonance. Manipulation of the higher harmonics and the use of other interval dissonance models (e.g., Kameoka & Kuriyagawa 1969) produce slightly different results, but all such models since Helmholtz imply

dissonance curves similar to those shown in Figures 3-8 and 3-9. As a consequence, results such as those in Table 3-1 concerning three-tone harmonies are always obtained. Clearly, a psychophysical model of harmony must correspond more closely with the facts of human perception.

The psychoacoustical model of Terhardt and Parncutt

The values shown in Table 3-1 certainly do not support the idea that the sum total of all dissonances among the upper partials in a chord can explain the most fundamental phenomena of harmony in Western music, but does the problem lie in the hard-nosed acoustical approach? Ever since the work of Helmholtz on precisely this issue, the literature on pitch perception, not to mention musical aesthetics, has been filled with discussions on the subtle differences between physical descriptions and psychological perceptions. There is, to be sure, a rough equivalence between, for example, tone frequencies measured in Hertz and perceived pitch and between amplitudes measured in decibels and the perceived loudness of tones, but the match is not perfect. Would therefore a more sophisticated quantitative psychoacoustical approach that places more emphasis on human perception and less emphasis on the physics of vibrations fare better?

The work of Terhardt and colleagues in the 1970s and the development of that model by Parncutt and colleagues in the 1980s and 1990s provide the most complete expression of what might be characterized as an interval-based psychoacoustical explanation of harmony. Those models are based on a wealth of psychological data and have been described in detail in the terminology of both experimental psychology and music theory. Parncutt's book, *Harmony: A Psychoacoustical Approach* (1989) is the most thorough discussion of all relevant details and is, for good reason, often cited as the definitive modern statement of a psychoacoustical treatment of harmony. For clarity of exposition and thoroughness of both theoretical and empirical work, Parncutt's book is a masterpiece, but it goes no further in explaining three-tone harmony than any of the other interval-based approaches. This fact can be seen from the relevant portion of his Figure 6.2 (reproduced here as Figure 3-13).

Shown in the figure are two related measures of the harmoniousness (the term preferred by Parncutt being "tonalness") of tone combinations. "Pure tonalness" is a theoretical measure of the sonority "dependent on the number and audibility of the pure tone components of a sound" (p. 177), whereas "complex tonalness" is defined as the "audibility of the most audible complex tone sensation of a sound" (p. 169). Both values are of theoretical interest, but

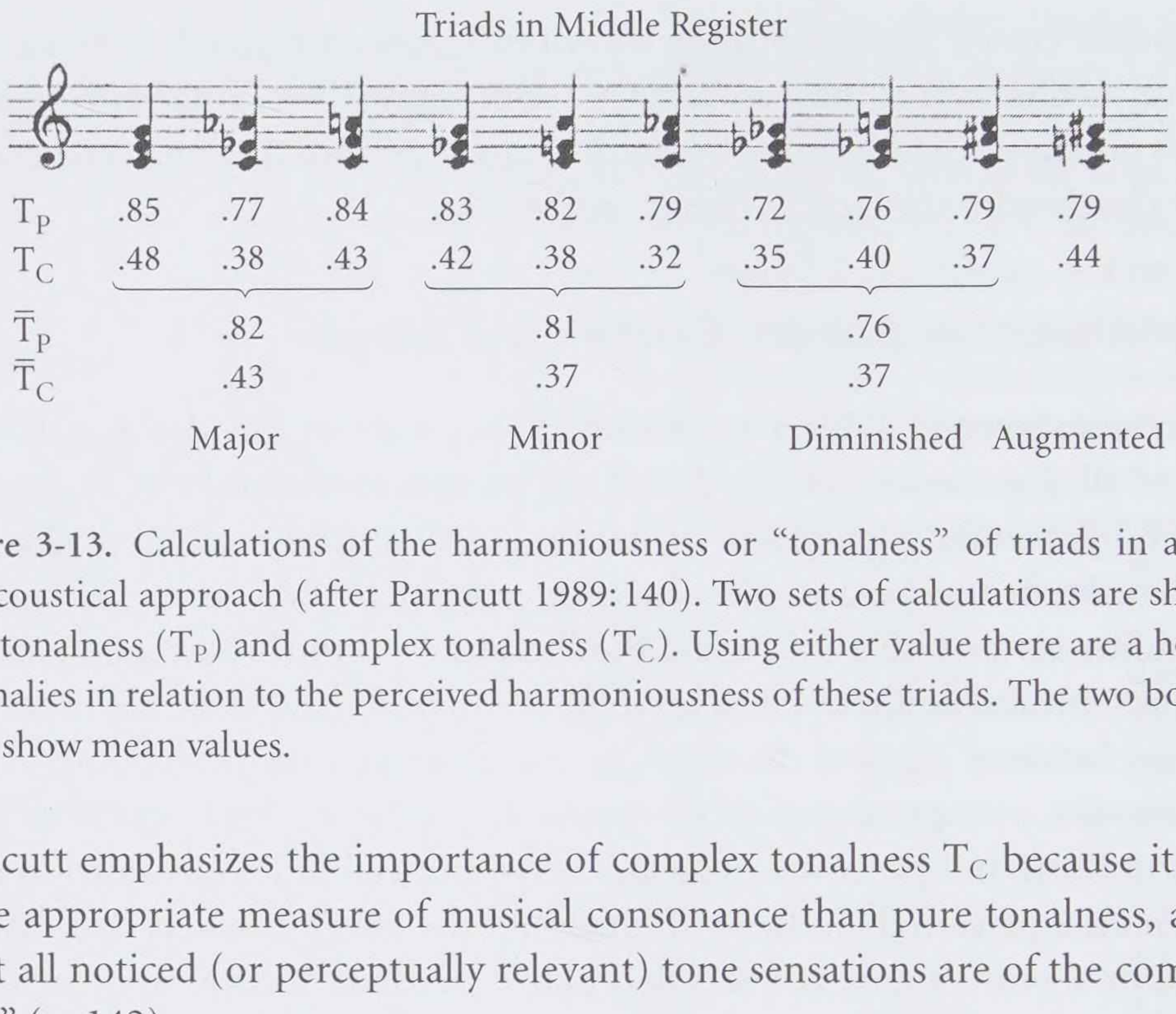


Figure 3-13. Calculations of the harmoniousness or "tonalness" of triads in a psychoacoustical approach (after Parncutt 1989: 140). Two sets of calculations are shown: pure tonalness (T_P) and complex tonalness (T_C). Using either value there are a host of anomalies in relation to the perceived harmoniousness of these triads. The two bottom rows show mean values.

Parncutt emphasizes the importance of complex tonalness T_C because it is "a more appropriate measure of musical consonance than pure tonalness, as almost all noticed (or perceptually relevant) tone sensations are of the complex kind" (p. 142).

What, then, are the psychoacoustical conclusions that one can draw from such theoretical calculations? Most jarring is the finding that, among the 10 triads in Figure 3-13, the augmented chord ($T_C = 0.44$) is the second most sonorous – more sonorous than two of the major chords and all of the minor chords! The average values for the four types of triad indicate the augmented chords to be the most harmonious! Admirably, Parncutt does not evade the problem that this raises: "The dissonance of the augmented triad in music theory is not reflected by its calculated tonalness; it appears to have cultural rather than sensory origins" (p. 141). But that theoretical result alone is an utter disaster for this type of psychoacoustical model, because it shows that *the most striking perceptual phenomenon in all of harmony* – the difference between resolved and unresolved chords – cannot be explained. By calculating average values, the major chords are found to be more sonorous than both the minor and diminished chords, as would be expected, but the model implies that the augmented chords are more sonorous than either the major or the minor chords, and the diminished chords are as sonorous as the minor chords – results that music theorists would not easily accept.

Failing to account for the most basic perceptual facts concerning the harmoniousness of triads (as firmly established in music theory, and quantitatively known through experimental work on musically trained and untrained

listeners, Figures 3-10 and 3-11), the psychoacoustical approach advocated by Terhardt and Parncutt needs to be rescued by "cognitive" factors that, they are forced to conclude, determine how we actually perceive music *in spite of* the contrary influence of psychoacoustical factors. As a mature statement of a psychoacoustical theory of *intervals* (not harmony), such theoretical work is noteworthy, but its contribution to an understanding of pitch phenomena ends precisely where music begins to get interesting.

In light of such calculations, theories of harmony that consider *interval* effects as "psychoacoustic," but regard chord effects as "cognitive" must maintain that the psychoacoustical calculations shown in Figure 3-13 and Table 3-1 are simply overpowered by learned musical associations; in such a view, although single-tone and two-tone perception has a psychophysical basis, three-tone perception is "learned." For a psychoacoustical theory to abandon psychoacoustics already at the three-tone level is simply unsatisfactory. Althougth interval effects are, not surprisingly, important for music perception, the possibility of psychophysical *chordal* effects must be considered before we ignore our ears and blame everything on "cognition."

In the remainder of this chapter (and in Appendix 2), it will be shown that consideration of the three-tone structure of chords allows for a psychophysical understanding of the perceptual difference between resolved and unresolved chords without resorting to inherently unquantifiable "cognitive" factors.

Tension in three-tone chords

The first theoretical argument indicating that three-tone combinations are important in the perception of harmony was essentially the negative finding that interval-based calculations cannot explain the basic phenomena of harmony. The second theoretical argument is more constructive, and is based on the idea that *the resolved/unresolved distinction in harmony theory is as important to the perception of chords as the consonant/dissonant distinction is to the perception of intervals*. Reconstructing the entire theory of harmony central to classical Western music is beyond the scope of the present chapter, but the basic harmonic phenomenon behind such a view can be stated as follows.

Among the triads of diatonic music, there are two types of resolved chord (major and minor) and two types of unresolved chord (augmented and diminished). If the equivalent size of the intervals contained within the chords is taken as the property defining "unresolvedness," then there are several other three-tone combinations that should have a similarly unresolved perceptual quality. These are, collectively, what might be labeled the "tension chords".

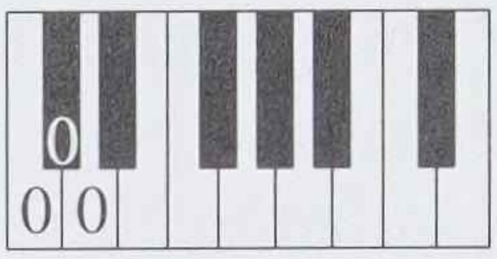
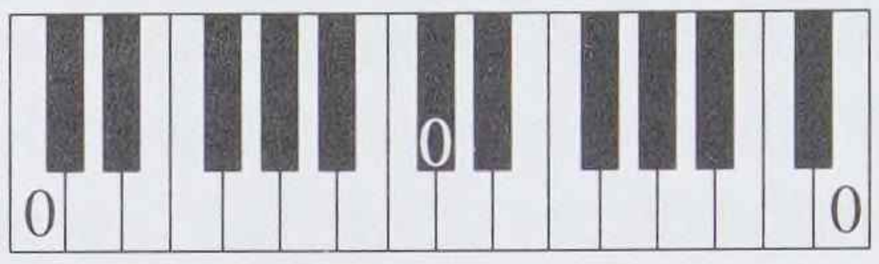
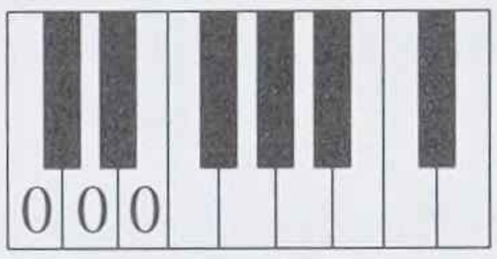
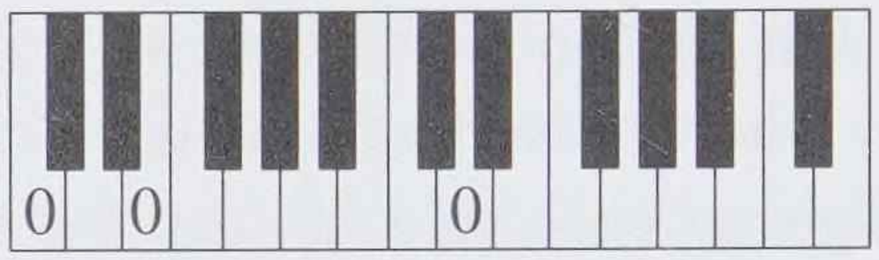
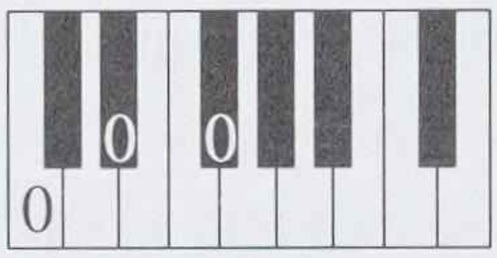
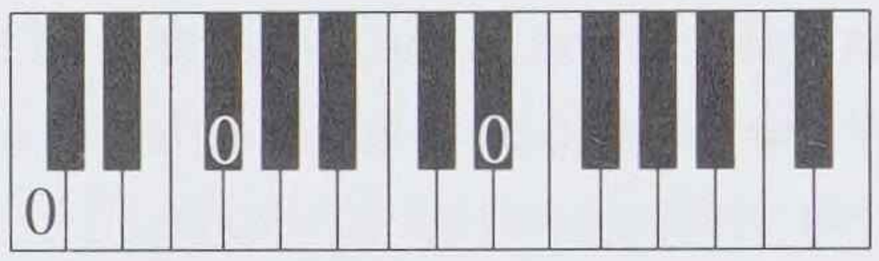
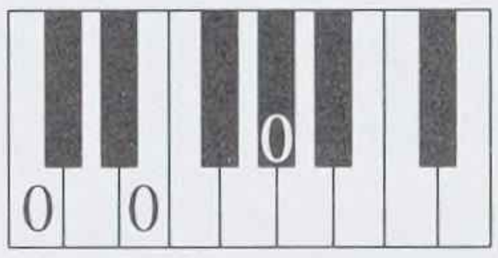
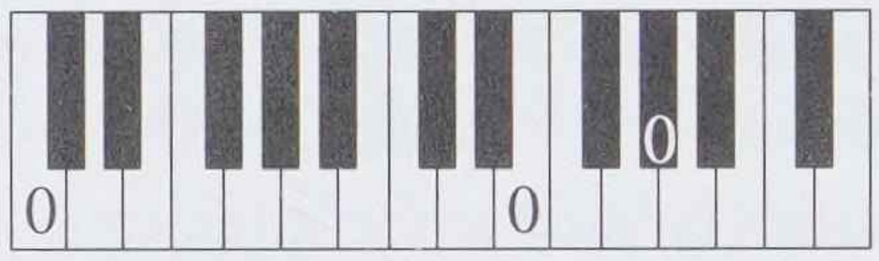
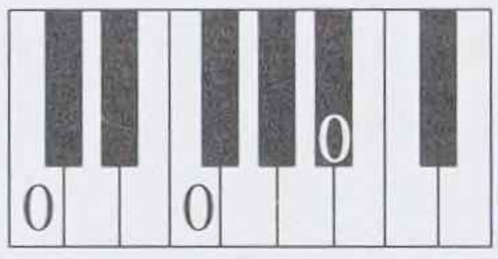
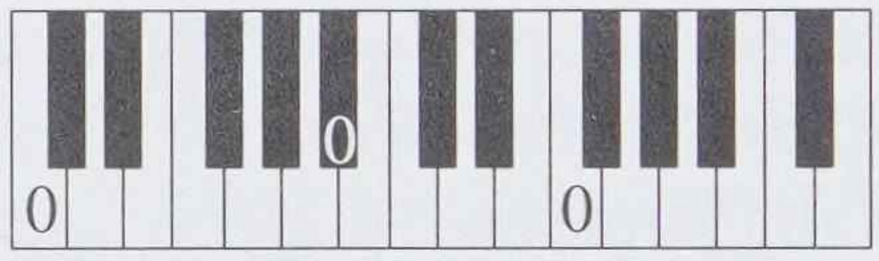
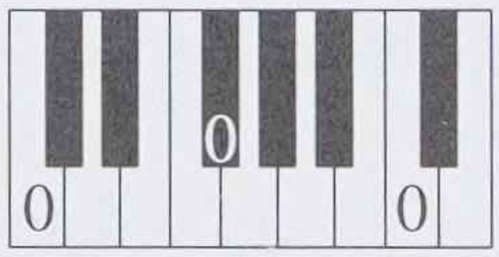
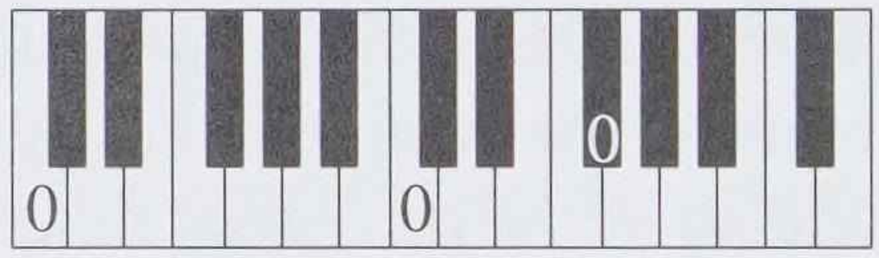
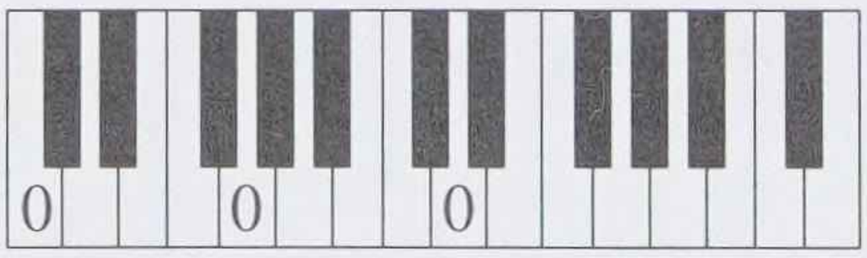
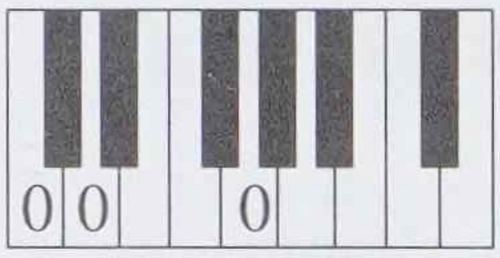
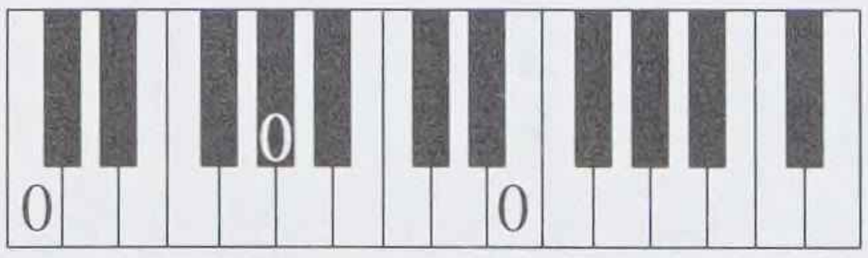
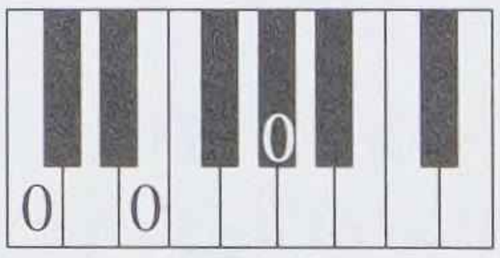
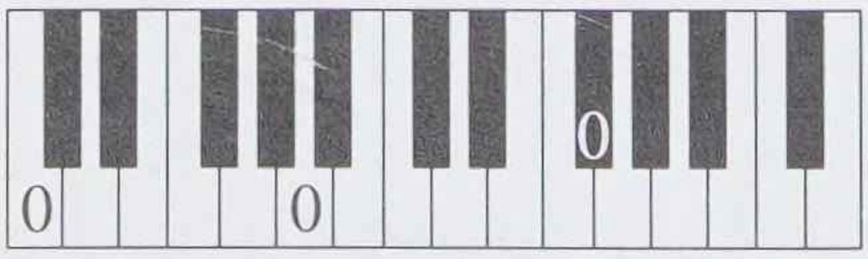
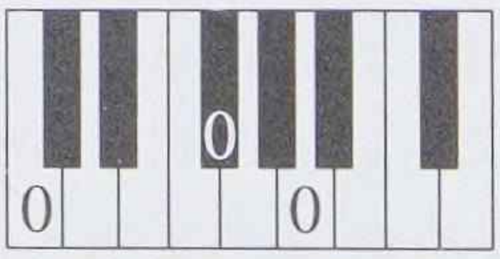
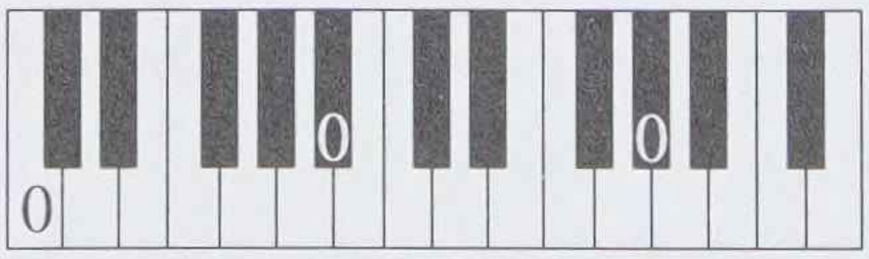
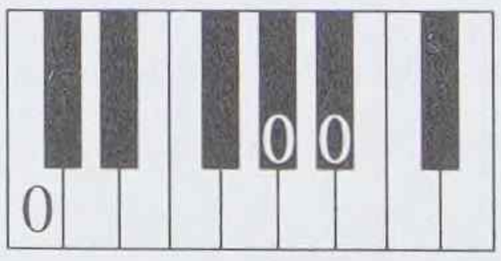
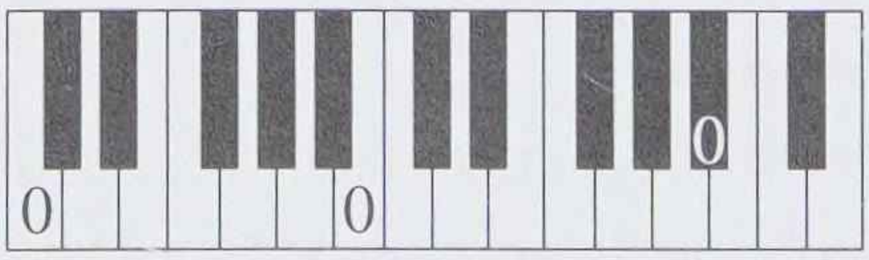
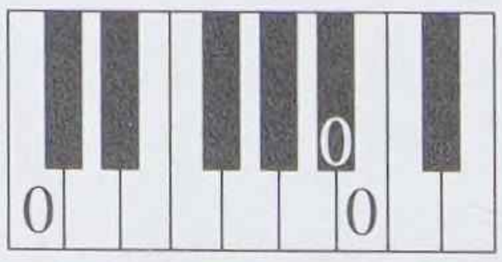
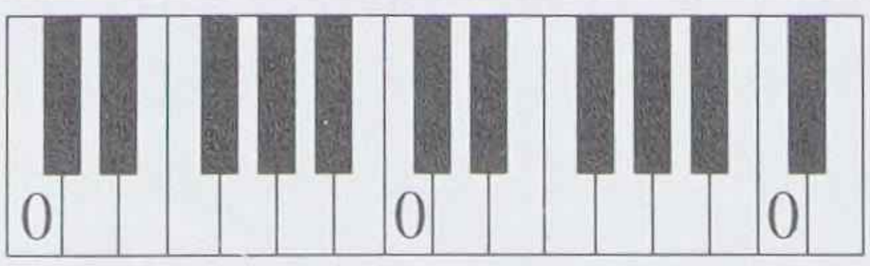
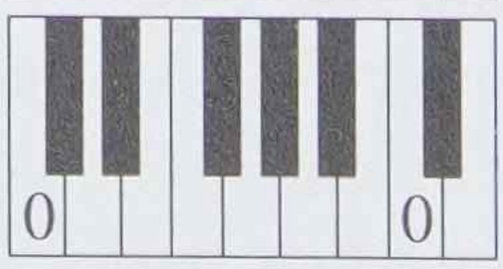
	Root Position	Inverted Form
Tone Cluster		
Tone Cluster		
C-dim		
C-aug		
C7-sus4		
Tritone		
G-sus4		
C-aug		
F#-dim		
A^b-add9		
B(M7)-aug		
C-C'-C"		

Figure 3-14 shows all eleven of the tension chords – three-tone combinations with two equivalent intervals – that begin with C. Although these chords are not recognized as a coherent set in traditional music theory, they have a common structure (in root position, two equivalent intervals) and a common perceptual feature (unresolved character). Beyond the 11th tension chord lies a chord with three tones separated by octaves (C-C'-C") and then the sequence repeats itself, with the same eleven tension chords re-appearing – now spread over several octaves. Among these chords, the "tense" unresolved character of the tone clusters with equivalent intervals of one- or two-semitones is not obvious due to the strong dissonance of the intervals, but when they are spread out over two or three octaves, the stark dissonance of semitone and whole-tone intervals is removed, while the unresolved character remains.

Whether heard in vertical (harmonic) or horizontal (melodic) form, the source of the unresolvedness of these chords is what Meyer has referred to as "intervallic equidistance" (Meyer 1956: 164), which produces the uniformity and tonal ambiguity characteristic of augmented and diminished chords. "If the vertical organization is undifferentiated as to intervallic distance, then there can be no focal point around which organization can take place – out of which shape can arise." Such chords have "no root and no shape and hence no tendency" (p. 166).

The authority of Meyer is important support for the view that equivalent intervals are the source of the clearly-perceptible harmonic tension, but it is nonetheless surprising how little attention has been paid to this phenomenon. Meyer himself has commented that the neglect of this aspect of harmony theory may be due to the lack of a "catchy label," but "intervallic equidistance" certainly states the case, and the phenomenon itself (the unresolved character of the augmented and diminished chords) is of course well-known. Why textbooks on harmony do not address this issue in the first chapter is most likely a consequence of the common view that psychoacoustics is relevant up to and including two-tone intervals, but beyond that only the top-down in-

---

Figure 3-14. The full set of eleven "tension chords" starting at C and the terminating C-octaves (Cook 1999a). The left-hand column shows the tension chords in "root" position, in which the equivalent interval substructure is evident. The right-hand column shows the same chords in various alternative configurations, in which perceptually the unresolved nature is sometimes more evident. Although not recognized as a coherent set in traditional music theory, all of these chords have an unresolved character. Common labels for the chords are noted, but alternatives are possible.

fluences of musical traditions, training and "cognition" need to be considered. In that view, quite aside from their psychoacoustic properties, the diminished and augmented chords in Western music are simply learned as being unresolved – either through active training or passive absorption of popular music. The contrary view is that – aside from both the interval effects and the influences of learning – the three-tone structure itself has a perceptible character. When we hear, for example, an augmented chord followed by the chord produced by raising or lowering any tone of the augmented chord by a semitone, it is not the melodic semitone interval that is the most salient feature of the music and certainly not the regurgitation of years of training that forces us to hear "proper form", but rather it is the higher-level harmonic structure that grabs our attention – the movement away from unresolved "intervallic equidistance" toward harmonic resolution.[2]

The two directions of chordal resolution

Although the unresolved tension chords themselves are characteristically ambiguous with regard to key and, as a consequence, have diverse uses in jazz and some forms of modern classical music, they show a remarkable regularity in relation to the two modes of chordal resolution. As illustrated in Table 3-2, the smallest modification that can be made to a tension chord is the raising or lowering of any one tone by a semitone – a change that sometimes will achieve resolution. When resolution is possible, it can be achieved in only two ways: the fall of one semitone, *inevitably resolving to a major chord*, or the rise of one semitone, *inevitably resolving to a minor chord*. This regularity is illustrated in Figure 3-15 for the case of an augmented chord.

What is noteworthy about the pattern shown in Table 3-2 and Figure 3-15 is that, by setting aside the complex rules of traditional music theory involving keys, scales, chord inversions, the circle of fifths, specific interval effects, etc., and taking *chordal tension* as the reference point for discussing harmony, we are left with an extremely simple regularity: *If resolution of chordal tension can be achieved by a minimal alteration in pitch (a semitone shift), the resolution will be to a major chord when the tone is lowered, and to a minor chord when the tone is raised.* No statement of comparable simplicity can be made within the framework of traditional music theory regarding the major and minor modes.

The regularities of the resolved and unresolved chords shown in Table 3-2 are of course well-known in traditional harmony theory, but are normally stated *in relation to the resolved major triads*. For example, the relations between major and minor chords noted above can be stated in traditional terms,

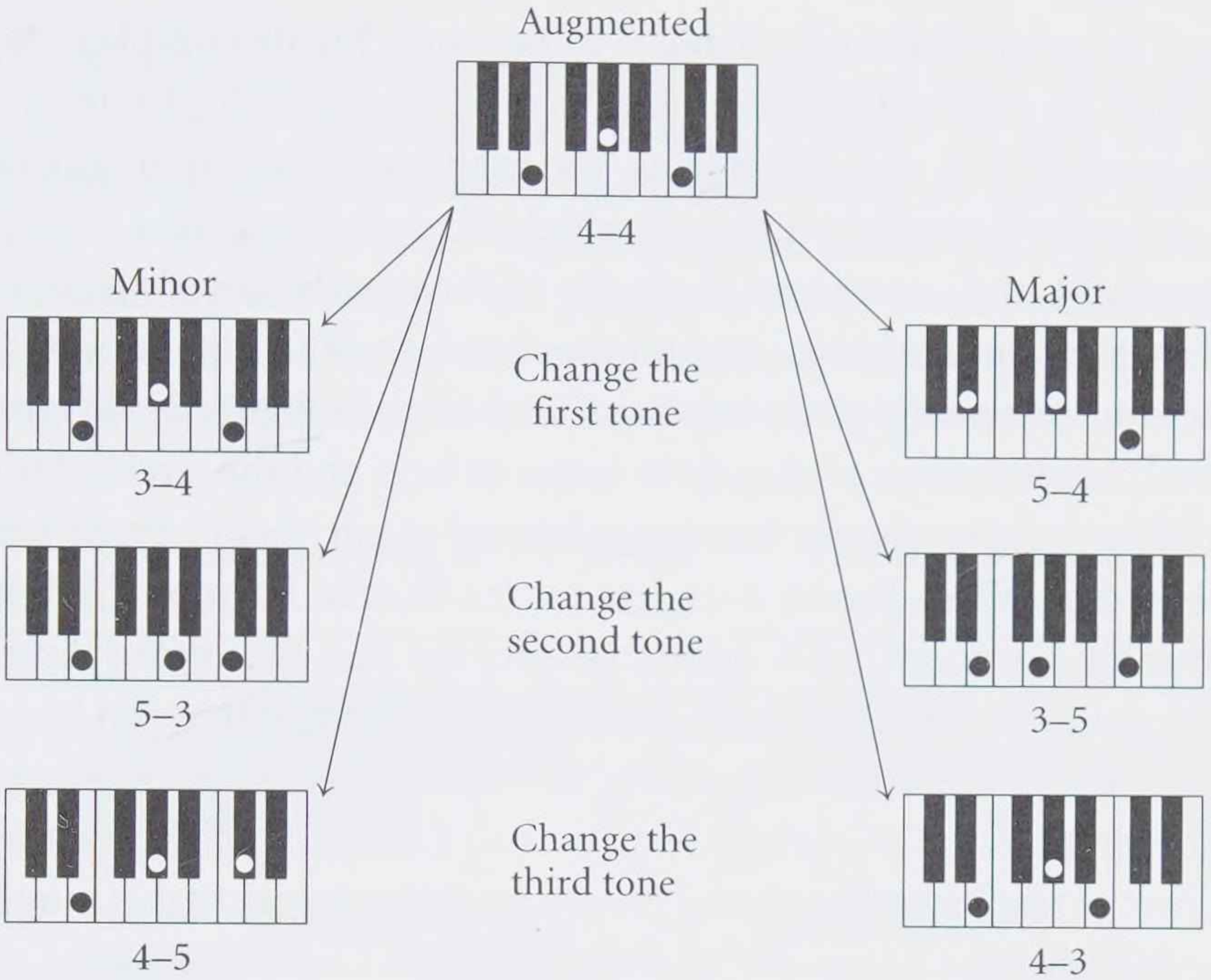


Figure 3-15. Examples of the changes in mode that occur by raising (left) or lowering (right) one tone in a tension chord (E-aug). Starting with an augmented chord, a semitone shift in any of the tones will result in a resolved harmony. The resolution will be to a minor chord (F-min, A-min and C#-min) following a semitone shift *up* and to a major chord (G#-maj, C-maj, E-maj) following a semitone shift *down*. Numbers indicate the interval structure. For other kinds of "tension chords," shifting tones up or down will sometimes result in "dissonant chords," but when a resolved chord is produced it shows the same regularity as seen for the augmented chord above.

as follows: Starting with a major chord in root position or second inversion, a semitone decrease in the interval known as a major third to a minor third or a semitone decrease of the tonic to change the major third to a perfect fourth will change the harmonic mode from major to minor; but if the major chord is in first inversion, the lower tone in the interval of a minor (!) third must be lowered a semitone to a major third or the perfect fourth lowered to a major third to obtain a minor chord. And vice versa for minor chords. Wow! Let's not get into how transitions between resolved and unresolved chords can be stated in traditional harmony theory! Stated either in the terminology of traditional harmony theory or in relation to the tension chords, the same harmonic phenomena can be described, but if harmony theory is seen through the lens of the resolved major triad – as the upper partials story would suggest as the most "natural" approach – then the utter simplicity of the relationship between re-

solved and unresolved chords (Table 3-2) becomes highly complex. Arguably, the insistence on taking the major mode as the centerpiece for harmony theory is as unnecessary and unjustifiable as, for example, studying linguistics with English syntax as the standard against which all else is compared.

Perceptually, the unresolved character of the tension chords relative to the resolved character of the major and minor chords is clear, but, because the pattern of resolutions of any given tension chord (e.g., in Figure 3-15) cuts across the normal classifications of chords in terms of keys and the circle of fifths, the concept of the tension chords stands traditional music theory on its head. The E-augmented chord in Figure 3-15 has no traditional harmonic relationship with F-minor, A-minor or C#-minor, despite the fact that those three minor chords are the three "nearest" resolutions of E-augmented that can be achieved by a semitone increase in tone. Similarly, the nearest resolution of E-augmented to a major chord is to G#-major, C-major and E-major, but this fact is not an integral part of traditional harmony theory. As "unnatural" as this may be from the perspective of the harmonic traditions of 16th Century classical music, it is the pattern that necessarily emerges by emphasizing the perceptual importance of harmonic *tension*, rather than major key *resolution*. The complexities of conventional harmony theory notwithstanding, underlying those venerable traditions is a psychoacoustical simplicity: Without exception, if harmonic resolution can be achieved by a semitone shift in an unresolved tension chord, then the resolution will be major with a semitone decrease and minor with a semitone increase.

As convincing as the argument that harmony theory must be discussed in relation to the *resolved major triad* has seemed to many generations of music theorists (Rameau 1722; Helmholtz 1877; Schoenberg 1911; Tramo et al. 2001), it leaves unsolved the problem of how to explain minor chord resolution, since the interval of a minor third does not appear among the higher harmonics. Complicated mathematical manipulations (e.g., Riemann 1877) have been advocated to explain away the anomalous resolved character of minor chords, but such arguments have, for good reason, left most music theorists unconvinced. The resolution of both major and minor chords is perceptually "obvious" to both trained and untrained listeners, and even to those without a PhD in mathematics – suggesting that the (implicitly major chord) emphasis on the upper partials when discussing three-tone harmony may be unwarranted.

In contrast, by shifting the discussion away from (major) chord *resolution* and taking the concept of three-tone *tension* as the perceptual centerpiece for discussing harmony, a psychophysical explanation becomes possible for both the resolved/unresolved distinction and the major/minor distinction. It

Table 3-2. The relationship between the tension chords and major and minor resolution (Cook et al. 2001, 2002b). Blocks A, B and C represent the same sets of chords and their transformations. The tension chords are shown in the middle column. Lowering any of the three tones of the tension chord can lead to major chord resolution (first three columns) and raising them can lead to minor chord resolution (last three columns). An arrow pointing up (down) means an upward (downward) shift of one semitone of the corresponding tone of the tension chord. An open circle means no change.

	Chord after lowering of one tone			Initial chord	Chord after raising of one tone		
	○○↓	○↓○	↓○○	○○○	○○↑	○↑○	↑○○
A. Interval structure	3–2	2–4	4–3	3–3	3–4	4–2	2–3
	4–3	3–5	5–4	4–4	4–5	5–3	3–4
	5–4	4–6	6–5	5–5	5–6	6–4	4–5
	6–5	5–7	7–6	6–6	6–7	7–5	5–6
	7–6	6–8	8–7	7–7	7–8	8–6	6–7
	2–4	1–6	3–5	(~2–5)	2–6	3–4	1–5
	8–7	7–9	9–8	8–8	8–9	9–7	7–8
	4–3	3–5	5–4	(~4–4)	4–5	5–3	3–4
	9–8	8–10	10–9	9–9	9–10	10–8	8–9
	6–2	5–4	7–3	(~6–3)	6–4	7–2	5–3
B. Examples	C-D#-F	C-D-F#	B-D#-F#	C-D#-F#	C-D#-G	C-E-F#	C#-D#-F#
	C-E-G	C-D#-G#	B-E-G#	C-E-G#	C-E-A	C-F-G#	C#-E-G#
	C-F-A	C-E-A#	B-F-A#	C-F-A#	C-F-B	C-F#-A#	C#-F-A#
	C-F#-B	C-F-C	B-F#-C	C-F#-C	C-F#-C#	C-G-C	C#-F#-C
	C-G-C#	C-F#-D	B-G-D	C-G-D	C-G-D#	C-G#-D	C#-G-D
	C-D-F#	C-C#-G	B-D-G	(~C-D-G)	C-D-G#	C-D#-G	C#-D-G
	C-G#-D#	C-G-E	B-G#-E	C-G#-E	C-G#-F	C-A-E	C#-G#-E
	C-E-G	C-D#-G#	B-E-G#	(~C-E-G#)	C-E-A	C-F-G#	C#-E-G#
	C-A-F	C-G#-F#	B-A-F#	C-A-F#	C-A-G	C-A#-F#	C#-A-F#
	C-F#-G#	C-F-A	B-F#-A	(~C-F#-A)	C-F#-A#	C-G-A	C#-F#-A
C. Labels from Music Theory	–	D7	B	C dim	Cm	–	D#m7
	C	G#	E	C aug	Am	Fm	C#m
	F	C7	–	F sus4	–	–	A#m
	–	–	–	tritone	–	–	–
	–	D7	G	G sus4	Cm	–	–
	D7	–	G	G sus4	–	Cm	–
	G#	C	E	C aug	Fm	Am	C#m
	C	G#	E	C aug	Am	Fm	C#m
	F	G#7	–	F# dim	–	–	F#m
	G#7	F	–	F# dim	–	–	F#m

Major chords | Tension chords | Minor chords

is worth noting in this context that the exact location of the pitches at which chordal resolution is considered to be most "satisfactory" is highly variable, depending on tuning systems, musical genre, and instruments – and is a contentious issue among musicians, but the salience of unresolved harmonies is not disputed. By placing the agreed-upon, unresolved tension chords at the heart of a psychophysical theory of chord perception (Table 3-2), it is seen that the instability of three-tone tension can be resolved by pitch movement in either of two (and only two) directions, corresponding to major and minor chords. Provided only that pitch alterations made on a tension chord do not result in interval dissonance (i.e., do not introduce lower-level factors that prevent resolution), there are in fact only two classes of chord that a semitone pitch alteration can bring about – major, following a semitone decrease, and minor, following a semitone increase.

Schoenberg (1911:96) predicted that the major and minor modes would eventually disappear in the same way that the other church modes have gone out of fashion, but it is arguably the case that the dichotomy of major and minor harmony has a basis that is inherently tied to the phenomenon of intervallic equidistance – and is therefore not something that is subject to musical fashions, styles and traditions, but rather has a psychophysical basis. If indeed the unresolved, unstable character of the tension chords is due to the equal spacing of intervals (Meyer 1956) and therefore has a Gestalt perceptual reality, then a dichotomy of possible resolutions is implied by the fact that there are only two possible directions in which to make changes in auditory frequency. The fact that chordal resolutions in the two directions of decreasing and increasing frequency correspond to major and minor harmonies indicates that chordal tension is as essential to the discussion of three-tone harmony as dissonance is to the discussion of two-tone consonance.

The pattern seen in Table 3-2 can be further generalized by considering the effects of multiple semitone changes to the tension chords. In Table 3-3A, it is seen that, again starting from the tension chords, major chord resolution is obtained by any combination of tone shifts that produces a *net* semitone decrease and, contrarily, minor resolution is obtained by a *net* semitone increase in pitch. Not shown in the Table is the fact that, regardless of the magnitude of the changes in each of the individual tones, if resolution is attainable, a net decrease (increase) of one semitone will produce major (minor) resolution.

Finally, in Table 3-3B it is seen that this pattern of major/minor mode resolutions is reversed when the net pitch change is by a whole-tone rather than a semitone. Starting with a tension chord, a net shift of two semitones up (down) will result in major (minor) resolution. A net change of three semitones in a

Table 3-3. The relationship between the tension chords and major and minor resolution due to the rise or fall of any of the pitches. Section A shows the chordal changes following a net change of one or three semitones. Section B shows the chordal changes following a net change of zero or two semitones. Symbols as in Table 3-2.

A. Interval structure

3-down	Net semitone decrease			Initial chord	Net semitone increase			3-up
↓↓↓	↑↓↓	↓↑↓	↓↓↑	○○○	↓↑↑	↑↓↑	↑↑↓	↑↑↑
3–3	1–3	5–1	3–5	3–3	5–3	1–5	3–1	3–3
4–4	2–4	6–2	4–6	4–4	6–4	2–6	4–2	4–4
5–5	3–5	7–3	5–7	5–5	7–5	3–7	5–3	5–5
6–6	4–6	8–4	6–8	6–6	8–6	4–8	6–4	6–6
7–7	5–7	9–5	7–9	7–7	9–7	5–9	7–5	7–7
2–5	0–5	4–3	2–7	(~2–5)	4–5	0–7	2–3	2–5
8–8	6–8	10–6	8–10	8–8	10–8	6–10	8–6	8–8
4–4	2–4	6–2	4–6	(~4–4)	6–4	2–6	4–2	4–4
9–9	7–9	11–7	9–11	9–9	11–9	7–11	9–7	9–9
6–3	4–3	8–1	6–5	(~6–3)	8–3	4–5	6–1	6–3
Tension chords	Major chords			Tension chords	Minor chords			Tension chords

↓ downward semitone shift

↑ upward semitone shift

○ no change

Table 3-3. *(continued)*

	Net wholetone decrease				No net change		Initial chord		No net change		Net wholetone increase		
	○ ↓↓	↓↓ ○	↓ ○ ↓	○ ↑↓	○ ↓↑	↓ ○ ↑	○ ○ ○	↑ ○ ↓	↓↑ ○	↑↓ ○	↑ ○ ↑	↑↑ ○	○ ↑↑
	2–3	3–4	4–2	4–1	2–5	4–4	3–3	2–2	5–2	1–4	2–4	4–3	2–4
	3–4	4–5	5–3	5–2	3–6	5–5	4–4	3–3	6–3	2–5	3–5	5–4	4–3
	4–5	5–6	6–4	6–3	4–7	6–6	5–5	4–4	7–4	3–6	4–6	6–5	5–4
B. Interval structure	5–6	6–7	7–5	7–4	5–8	7–7	6–6	5–5	8–5	4–7	5–7	7–6	6–5
	6–7	7–8	8–6	8–5	6–9	8–8	7–7	6–6	9–6	5–8	6–8	8–7	7–6
	1–5	2–6	3–4	3–3	1–7	3–6	(∼2–5)	1–4	4–4	0–6	1–6	3–5	2–4
	7–8	8–9	9–7	9–6	7–10	9–9	8–8	7–7	10–7	6–9	7–9	9–8	8–7
	3–4	4–5	5–3	5–2	3–6	5–5	(∼4–4)	3–3	6–3	2–5	3–5	5–4	4–3
	8–9	9–10	10–8	10–7	8–11	10–10	9–9	8–8	11–8	7–10	8–10	10–9	9–8
	5–3	6–4	7–2	7–1	5–5	7–4	(∼6–3)	5–2	8–2	4–4	2–7	7–3	6–2
	Minor chords			Tension chords							Major chords		

tension chord will never result in resolution. This entire pattern due to 1, 2 or 3 semitone shifts up or down then reverses and repeats itself for 4, 5 and 6 semitone changes. Clearly, when multiple semitone steps of all three tones are brought into consideration, the number of possible resolutions (and their musical meaning) increases rapidly, but a remarkable simplicity in terms of major/minor mode remains. *Starting with unresolved chordal tension*, the characteristic mood of major mode harmonics is obtained by making the smallest decrease in pitch allowed within a diatonic scale, and minor mode obtained by the smallest increase. In other words, however one might wish to characterize the psychology of major and minor harmonies, there is an underlying lawfulness that relies on the tonal *direction* of movement away from unresolved chordal tension. Why would a minimal *decrease* in pitch that leads from tension to resolution produce the characteristically "bright" mood of major harmonies, while an *increase* in pitch produces the "dark" mood of minor harmonies?

D. Major and minor modes

Statements about the affective content of the major and minor modes have been made for centuries. Rameau (1722:164) remarked that the major mode is suitable for "mirth, rejoicing, grandeur and magnificence," while the minor mode is characteristically used to express "sweetness, tenderness, plaints" and a "mournful" disposition. Cooke (1959:50) summarized the major and minor modes as "joy, confidence, love, serenity and triumph" versus "sorrow, fear, hate, disquiet and despair" – a dichotomy that he likened to pleasure and pain, but which, historically, has its roots in secular versus spiritual music. Swain (1997:146) commented on the historical evolution of the major mode for the expression of "happy, humorous, or otherwise light emotions" and the minor mode for "the darker ones," but did not pursue the causes or consistency of this distinction. Even Schoenberg (1911:96) commented on the similarity of the major and minor modes to male and female – although this was prelude to declaration that these two modes would eventually disappear under the influence of chromatic scales. In a discussion of 20th Century harmony, Perischetti (1961:35) acknowledged a distinction between "dark" and "bright" scales, with the scales containing the minor third from the tonic being on the dark side, but he did not pursue their psychology further. Harrison (1994) developed an incisive theory of chromatic music that takes the major-minor duality as its starting point – "the Adam and Eve of harmonic function" (p. 17). This contrast between major and minor harmonies has in fact been the basis for much work

in diatonic music theory, often referred to as "dualism" (Riemann 1877), but, under the influence of modern ethnomusicology, such categorical statements on the affect of musical mode have become rare. In any case, the central problem in discussing the "universality" of the affect of major and minor modes is that chords consisting of three (or more) simultaneous tones have a long and pervasive history in Western music, but are non-existent or arguably imported from diatonic traditions in the music of other cultures.

Within the framework of classical Western music, Meyer (1956:222) has made a coherent argument for considering the minor mode to be more ambiguous than major, solely on the basis of the larger number of possible transitions that can follow a minor chord. He rejects the extreme view (e.g., Guernesey 1928) that the psychology of consonance/dissonance, resolved/unresolved and major/minor chords is entirely culturally-determined, but, in developing an argument that is essentially music theoretic, Meyer's view implies that the capability to distinguish the affect of major and minor harmonies is the consequence of a musical sophistication that is clearly well beyond that of the typical listener. As valid as his argument is in terms of traditional music theory, it is likely that there is a more direct psychophysical explanation of the two types of harmonic resolution.

So, despite the prevalence of the view that the major and minor modes reflect happy and sad moods, Sloboda's challenge (1985:61) remains unanswered: "as psychologists, we need to ask what psychological mechanisms allow these meanings to be comprehended by the listener." The claim made here is that the focus on *resolution* that has dominated all previous discussions of harmony unnecessarily introduces the confounding influences of different musical traditions, styles, instruments, tuning systems and, indeed, all the cognitive complexities of music theory in general. In contrast, by focusing on what is a more fundamental phenomenon, i.e., harmonic tension, the psychological need to move *away from* tension (not the *arrival at* resolution) can be addressed on a psychophysical basis without getting embroiled in debates concerning the preferred points of resolution. (The related argument with regard to interval perception is that the salient perceptual phenomenon of two-tone combinations is that of the *dissonance* of small intervals, not the relative *consonance* of various larger intervals.[3]) Given a definition of chordal tension as any three-tone chord with two equivalent intervals (in root or inverted position), there are inherently only two possible pitch directions in which to make tonal changes that will produce unequal intervals and thus achieve harmonic resolution. Regardless of the type of the initial tension chord (i.e., regardless of the size of the equivalent intervals of the tension chord in root position and regard-

less of their possible inversion), these two directions correspond precisely with the major and minor modes of resolution known from music theory. This remarkable pattern is implicit in traditional music theory, but it is hidden from view if the focus is on resolution, rather than on the most salient aspect of 3-tone chords harmonic tension.

What still requires an explanation, however, is why minor is "dark" and major is "bright," and not vice versa. The traditional "cognitive" answer is simply that it is an ingrained musical tradition from which anyone raised in the modern world cannot escape. Why minor is "sad" and major is "happy" is, in that view, a matter of Saussurian chance or perhaps a consequence of the major third being "more natural" (and therefore "happier"?) because it is found among the upper partials of the tonic, whereas the minor third is not. That line of argumentation is clearly speculative, but it is thoroughly unsatisfactory in so far as it implies that there is no inherent, intrinsic, psychoacoustic difference in the mood of major versus minor chords. Music theoreticians may have conceptual motivations for thinking that major and minor moods are nothing more than frozen accidents, but the theoretician's argument is perceptually far from convincing if pairs of major and minor chords are played in succession. To most people, it is clear that a minor chord (e.g., CD#G) followed by a related major chord (e.g., CEG) brings quite a different mood from the reverse sequence. To deny that is to deny the senses – truly a nonsensical argument. Theoreticians will wave their arms and say it is simply not true – and that all experimental support using naïve subjects is nothing more than a consequence of cultural indoctrination! But, whatever the source of this widespread "bias," it is certainly an empirical fact and needs to be explained.

Surprisingly, a plausible answer comes from a field within linguistics known as sound-symbolism, rather than from music psychology. Quite unrelated to music theory, Ohala (1984, 1994; and earlier, Morton 1977) noted that there is a cross-species "sound symbolism" of animal calls: an increase in vocal pitch is indicative of inferior social status and submission, whereas a decrease in vocal pitch is indicative of superior social status and dominance. They related this phenomenon to the well-known cross-cultural tendency in human languages to use rising F_0 in yes/no questions, and falling F_0 in statements (for further discussion, see Bolinger 1978; Cruttendon 1981, 1997; for details of changes in F_0 in a study comparing matched statements and questions, see Eady & Cooper 1986). The "social message" of such pitch changes has been characterized as "deference, politeness, submission and lack of confidence" for the questioning individual using a rising pitch intonation, and "assertiveness,

authority, aggression, and confidence" for the asserting individual using falling pitch[4] (e.g., Bolinger 1978).

The unambiguous regularities of harmony theory (Tables 3-2 and 3-3) suggest a similar psychology for musical pitch changes: minor mode as inherently associated with a *resolved* inferior social status and a submissive mentality, and major mode as associated with a *resolved* superior social status and an established dominance.[5] This psychological dichotomy is reflected in Cooke's (1959) characterization of the affective quality of commonly-used melodic sequences in Western classical music. As summarized by Kaminska and Woolf (2000), specific melodic lines in a minor key typically express "inescapable doom," "pain and acceptance," "pure tragedy," "yielding to grief," "never-ending despair" and "assertion of sorrow." Specific major key melodies, in contrast, express "joy" and "triumph." Their experimental work with musically untrained subjects did not reveal an ability to differentiate among various forms of grief and joy in response to the melodic lines, but their results strongly supported a differentiation between sorrow and joy and between submissiveness and assertiveness for minor and major key melodies, respectively.

A similar conclusion has been reached in studies of speech intonation, e.g., Scherer's (1974) finding of pleasantness due to falling pitch contour and surprise or fear with rising pitch in the voice. The implied evolutionary argument is therefore that, through the eons of mammalian and primate evolution, a rather simple auditory signal emerged to indicate social dominance. The "sound symbolism" of rising pitch is social weakness and of falling pitch is social strength. When music emerged in human societies using fixed musical tones (scales), tonal combinations that indicated a *decrease* in auditory frequency relative to a state of harmonic tension were necessarily – instinctively – associated with the pleasures of dominance, whereas *increases* in frequency were associated with the plaints of social subordinance.

Hypotheses about the origins of the major/minor duality necessarily remain speculations, and definitive resolution of the debate between innate and learned effects would probably require a cross-cultural study of musical traditions that do not use diatonic harmonies. However those issues may eventually be resolved, it is sufficient for the present chapter simply to note that, in musical cultures where diatonic chords are used, the dark/bright dichotomy for minor/major harmonies is perhaps as near to "universal" as any musical phenomenon. Whatever its origins, it is certain that a minimal pitch increase from a state of harmonic tension will *never* result in a major chord and, likewise, a minimal decrease from a state of tension will *never* result in a minor chord. That is an indisputable fact of traditional diatonic music.

It is essential to point out that the distinction between major and minor harmonies is itself distinct from that between resolved and unresolved harmonies. Brown (2000) has recently conflated these issues by arguing that ascending pitch contours in both speech and music "convey a feeling of tension and uncertainty, whereas ... descending contours convey a feeling of resolution" (p. 289), but the harmonic context has a deciding influence here. Clearly, in conventional music theory, harmonic resolution can be reached by either increases or decreases in pitch of various interval sizes, depending solely on the initial pitches, and similarly for moving from a state of harmonic resolution to one of unresolved tension or uncertainty. The effects of context are absolutely critical (see the Appendix for further discussion). Suffice it to note here Cooke's (1959) insistence that minor mode melodies are *not* "uncertainly sad," but are indeed *resolvedly* mournful – settled and resigned in their enduring, unhappy state. Major and minor chords "have one thing in common – their fixedness" (p. 63) – and are therefore distinct from the unresolved triads, which are inherently unstable, ambiguous with regard to musical key, and demanding of change.

A proper understanding of the basic phenomena of harmony is essential for the development of a general psychophysical theory of the perception of pitch combinations, so let us state the conclusions thus far concerning the psychology of pitch combinations as a series of unadorned assertions.

1. Among the many pitch intervals within the normal hearing range, only an interval of one or two semitones is notably dissonant and inherently "unpleasant." The relative consonance/dissonance of larger intervals is probably due to the effects of upper partials. The perception of dissonance is a feature of early sensory processing in the human auditory system and not a cultural artifact.
2. A strongly dissonant interval of 1–2 semitones can be transformed into a consonance either by narrowing the interval (to unison) or widening it further (to an interval of at least three semitones).
3. There are two varieties of three-tone combinations that do not contain interval dissonance: resolved chords (major and minor, both of which are inherently "harmonious") and unresolved ("inharmonious") tension chords. In musical traditions that use three-tone combinations, the distinction between resolved and unresolved chords appears to be a musical constant – suggesting that perception of three-tone "tension" is also a Gestalt that is inherent to the human auditory system (or possibly learned at a very early age). The relative contributions of learning and innate factors remain to

be determined, but the fact that non-musicians and children find no difficulty in distinguishing the moods of major and minor harmonies and in distinguishing between resolved and unresolved chords argues strongly for considering the perception of harmonic tension and its two modes of resolution to be a universal human capability rather than simply a result of passive absorption of certain musical traditions.

4. An unresolved tension chord can be resolved either by raising or lowering one or more tones in the chord (provided only that interval dissonance is not created).
5. If resolution is obtainable through a minimal (semitone) *increase* of pitch, it will be achieved by transforming the mood of tension into the mood of a resolved minor chord; if obtainable through a semitone *decrease* in pitch, it will be transformation into the mood of a resolved major chord.

These facts about pitch phenomena within a 12-tone scale indicate that, starting from a state of harmonic ambiguity, tension or uncertainty, there is something affectively "negative/dark/sad" or minor-like about small pitch increases, and something affectively "positive/bright/happy" or major-like about small pitch decreases. This relationship between pitch phenomena in diatonic music and emotion requires an explanation; the sound-symbolism hypothesis advocated by Ohala (1984, 1994) may provide such an explanation on an evolutionary basis.

E. Conclusions

This musical interlude is in danger of becoming a symphony in its own right, so let us reiterate only the three essential conclusions that will have implications for a discussion of voice intonation and a theory of emotions.

The first point is that the harmoniousness of three-tone combinations is truly a three-tone phenomena – not simply the summation of interval effects. The idea that chords are more than "many intervals" is little more than common sense to most musicians, but there remain a few hard-core reductionists who want to believe that all higher-order pitch phenomena can be explained on the basis of intervals. Both theoretical calculations (Table 3-1; Figure 3-13) and brain-imaging results (Figure 3-12) indicate that such a belief is misguided. As a consequence, whether in simultaneous harmonic combinations or in sequential melodic combinations, a quantitative treatment of harmony requires

consideration of not only the effects of pitch intervals (relative dissonance) but also pitch chords (relative tension).

Secondly, harmonic "stability" (i.e., a major or minor mode) can be established when (at least) three tones over an interval of (at least) a musical fifth have been intoned. When a narrower range of pitches has been used, a major or minor key may be *suggested*, but it cannot be *established* until a resolved chord (played either melodically or harmonically, but requiring an interval of at least a fifth) has been sounded. This is an empirical fact of diatonic music (see the Appendices for further discussion).

Thirdly, there is a well-known dichotomy of musical mode – i.e., two forms of harmonic stability. When other factors are held constant, a minor harmony sounds "sad," "dark," "pessimistic" or "submissive" relative to the "happy," "bright," "optimistic" or "dominant" mood of a major harmony. This psychological effect might find an explanation in Ohala's sound symbolism hypothesis – provided that we consider what a major or minor chord is *in relation to a state of harmonic ambiguity or tension*.

Both in the animal kingdom and in human society, moments of victory (implying dominance and happiness) or defeat (implying submission and sadness) are experienced in small ways on a daily basis. Just as assuredly as we know when we have scored a social "victory" or once again made an embarrassing faux-pas, we often know without theoretical analysis or mathematical calculation when pitch combinations ring major or minor. This musical perceptual ability is undoubtedly strengthened through musical education and passive absorption of the dominant music played in the all-pervasive electronic media, but an understanding of harmony is not unique to the trained musician: the ability to distinguish between resolved and unresolved chords and between major and minor chords is a cognitive skill common to all normal adults.

Similarly, once a major or minor mode has been established in a piece of music, it is intuitively obvious when a note is in key or out of key. Musical training can sharpen the ear, but advanced musicianship is not essential for hearing when tones do and do not fit within musical scales. What this musical ability implies is that, once a musical key is established, the use of tones relative to that key has an inherent, intuitively known affective meaning. Of course, if asked what the "meaning" of this or that musical pitch is, most people will have little to say beyond: "well, it sounds okay" or "it sounds out of place." But that level of insight is already quite remarkable and indicates an ear for musical key (which itself is based on the perception of resolved and unresolved three-tone combinations), that is, an ability to detect what tones are appropriate or inappropriate within the context of neighboring tones.

The fact that the major and minor modes are easily heard in music is psychologically significant, principally because the "ring" of major and minor harmonies is the one inescapable, historically-undeniable link between pitch phenomena and human emotions. Music theorists cringe on hearing simplistic statements about subtle and complex musical phenomena, but it is testimony to the innate musicality of most people that they normally hear that: "Major is happy. Minor is sad." Yes, there are ways of reversing that pattern with rhythmic manipulations – and all such associations can be temporarily destroyed in chromatic music, but the exceptions prove the rule here. For reasons probably as deep as evolutionary theory, certain pitch relations are widely, perhaps universally, understood to have a meaning that musicians have spontaneously exploited for centuries – singing their lost-love laments in a minor key and marching off to war in a major key. Modern composers have rebelled against these modal clichés – and discovered many interesting pitch phenomena in the process – but the dichotomous emotions evoked by major and minor harmonies show no signs of disappearing either in the popular mind or in popular music.

In summary, it can be concluded that there is an inherent mood in certain three-tone combinations that is perceptible to normal adults. Its musical basis is a complex and interesting story, but its significance for human psychology lies in the fact that it provides a link between quantifiable acoustical phenomena and the qualitative psychological phenomena of the emotions.

Notes

1. A fascinating, but still unresolved issue concerns the differential roles of the left and right frontal regions in the perception of musical mode. In the literature on psychopathology, there are indications of LH and RH hyperactivity being associated with, respectively, the positive and negative moods of mania and depression. In the music perception literature, Schmidt and Trainor (2001) have reported EEG indications of greater LH activity in the perception of major chords and greater RH activity in the perception of minor chords in normal subjects.

2. The "uneasiness" of the unresolved chords with intervallic equidistance is probably due to the difficulty of Gestalt-like sensory grouping. The middle tone in such chords is as easily grouped with the lower tone as with the upper tone – and either grouping will, like the alternating perceptions of the Necker cube, directly induce the contrary interpretation. The "correct" organization of the tones is therefore inherently uncertain – and that is perhaps the source of the unresolvedness. Because the Gestalt law of proximity cannot be applied to a tension chord, we expect or want a slight tonal adjustment from the ambiguous organization to a 2-to-1 or 1-to-2 grouping.

3. These two low-level psychological effects are the likely foundations on which the two most important cognitive effects in the perception of music stand: i.e., anchoring and chordal relations. In brief, anchoring is concerned with the resolution of two-tone dissonance within the constraints of diatonic scales. Normally, this involves successive tones that differ by a semitone or whole-tone, where the inherent dissonance is resolved by "anchoring" to a tone of the diatonic scale, most often the dominant or tonic, following a non-diatonic tone. Chordal relations involve a higher-level resolution, where the aptness of a sequence of chords depends on the musical mode and key. These two phenomena are the starting point for the discussion of the cognition of music [see, Bharucha 1984, 1996; Krumhansl 2001; Povel & Jansen 2001; and Tan et al. 1981], but already go beyond the psychophysical argument of the present chapter.

4. A related developmental hypothesis has been suggested by Levelt (1999:112). He notes that "a high register universally expresses vulnerability, helplessness, or special deference. The origin of that impression may be the child's very high speech register". Whether evolutionary or developmental, the basic idea is that high or rising pitch has an intrinsic biological meaning – a significance that we may be able to adorn, qualify or elaborate on in music or speech intonation, but cannot ignore.

5. The high-pitched whimpering and whining of animals in defeat is familiar, but there are also linguistic phrases that express the same phenomenon. In Japanese, the phrase "ne wo ageru" is used to indicate giving up, throwing in the towel or admitting defeat. Literally, it means "to raise one's sound [tone]" (as distinct from raising the volume of one's voice).

Chapter 4

The coding of human emotions

Synopsis

A bimodal pitch-based theory of emotions is outlined in which emotional valence is either positive or negative. The idealized form of valence is found in the major and minor modes of diatonic music, but a more subtle effect can be found in normal human speech. The relevant literature on intonation is reviewed and a general theory of intonation is discussed.

The fundamental argument of this chapter is that the encoding and decoding of human emotions is related to the encoding and decoding of the pitch phenomena discussed in Chapter 3. This does *not* mean that normal speech is truly musical in a diatonic sense, nor that emotions are solely auditory phenomena. Rather, it means that, for ancient evolutionary reasons, the mechanisms for the cortical storage and transfer, i.e., the "carriers," of emotion have characteristics similar to those of musical pitch. In this view, diatonic musical phenomena are idealized, simplified versions of the pitch phenomena of speech and, for that reason, have an affective clarity that is much greater than that found in normal speech, but the principles of pitch combinations are similar in both realms. Why the brain has evolved a fundamental representation of emotion related to auditory – rather than, say, visual – phenomena is undoubtedly a consequence of the importance that animal calls have had for mammalian species over eons of evolutionary time. Well before early primates were capable of close-up visual analysis of each other's grimaces and gestures, our tree-climbing ancestors were communicating long-distance by chirping, purring, squealing and growling at one another – using auditory signals to establish and maintain dominance relations, to signal dangers and opportunities, and to find suitable mates for reproduction. Once a secondary cortical representation of emotions had evolved, access to it through other modalities would have become possible and the direct, monopolistic link with audition necessarily weakened.

For the pursuit of questions concerned with evolutionary origins, analysis of the vocalizations of other species would be of central interest (e.g., Hauser 1996), but in the present chapter the focus will be on the characteristically human emotions – specifically, how it is that, through the voice, people can understand each other's emotional states. Naturally, any emotional response will

also entail a variety of effects in modalities other than the auditory – the entire repertoire of somatic effects elicited by the autonomic nervous system, as well as facial expressions and bodily gestures, so that the questions of emotion could be approached from other directions. The auditory approach is, however, more direct primarily because the auditory component has an older communicative history than the visual.

To explain the relationship between pitch and emotion, some basic terminology must be defined. First of all, a distinction will be made here between feelings and emotions. The term "emotion" will be used strictly to denote the affect human beings feel in relation to other human beings, whereas the term "feeling" will be used to denote a broader category that includes the entire set of uncontrolled, non-verbal, autonomic nervous system responses to (normally) external stimuli. If the wind causes a door to slam shut, a hairy spider suddenly appears on my plate, or I take a ride on a roller-coaster, strong feelings will be experienced. Without any thought or conscious deliberation, my body will react with autonomic responses, including a racing heart, profuse sweating, irregular breathing, shaking knees, and so on. Those feelings may be powerful, but they are rather contentless in terms of human relations. If, however, the slamming door or the ugly spider or the roller-coaster ride somehow involve people, as causal agents, victims or cohorts, then, in addition to the autonomic feelings, I will experience human emotions of pity, joy, anger, sympathy, worry, sadness, etc. with regard to those people. Although many discussions of emotion include the feelings experienced in response to physical events, any theory that includes both anonymous feelings and human emotions is unlikely to be more than a catalog of the autonomic responses that different kinds of stimuli can induce. The present discussion of emotion will therefore focus on human emotions in a social context.

Next, it is important to make a distinction between an emotional state and the specific life events that gave rise to the emotion. Raw emotions do not occur in the abstract; they are psychological responses to events involving people – social events that have a duration, location and causality of their own. Those social events constitute the real-world "contents" of emotions, and are unique in every individual's life. The psychological states that life events engender, however, can be categorized with a commonality among people – generally speaking, across cultures, across races and across generations. It is this common, finite set of definable human emotions that is the subject matter of the present chapter. How many "fundamental emotions" there might be and what kinds of life events can cause them are topics too large to be addressed here, but a related, more tractable topic can be delineated: the expression of emotional

valence through voice intonation. Essentially, the idea is that certain pitch combinations are used (usually, quite unwittingly by the speaker) in the expression of positive and negative emotional states and those pitch combinations are understood (also normally without conscious attention) by listeners as indicative of the speaker's emotional states. Even in this limited auditory context, the complexity of pitch combinations is roughly equivalent to the complexity of musical phenomena, so that only the basic framework can be outlined here (see Appendix 1 for further discussion within a musical context).

The linguist will be quick to ask: "So, just what is the intonational contour that is 'correct' for expressing emotion X through sentence Y?" Some answers to such questions will be offered below, but the possibility that coherent answers are indeed possible becomes apparent if we invert the question: "What intonation is *incorrect* in a particular circumstance?" If I announce to you in a cheery, light-hearted, bouncy tone of voice that "My grandfather died last night," you are likely to ask me if I have lost my mind, if I have myself fallen ill and if I understand what I am saying. The anomaly will be inescapable for you, and you will likely pursue the contradiction to find out which is the actual case: an inherently unhappy death in the family, as denoted in the literal sentence meaning, or an inherently happy frame of mind, as implied by the tone of voice. You will undoubtedly not go through a conscious process of pitch analysis to reach your conclusion, but you will know "in your gut" that something is amiss if the literal linguistic message is about death, and yet the tone of voice conveys a happy emotional state, or vice versa. When words and intonation do not match, the fact that emotional messages can be transmitted through pitch is obvious.

Most often, of course, there will be no conflict between the literal message and the emotional prosody; both the verbal message and the tone of voice will tell the same story – and the conversation will inevitably continue in the one direction indicated by both linguistic and paralinguistic information. At other times, there will be a mismatch between the literal message and the tone of voice, but there will be no uncertainty regarding the "true" meaning. If I say in a sarcastic tone of voice "Oh, yeah, I really want to go out to dinner with you," you will know that the meaning is "no", even if the literal answer was "yes". The tone of voice will have trumped the literal meaning. At other times, the literal language will be inherently ambiguous and the use of intonation will be decisive in determining its meaning:

> *Negation*: "I could go to town right now" (with elevated pitch on "could"), but, although that is a possibility, I won't.

Affirmation: "I could go to town right now" (elevated pitch on "right now"), and I'll be darned if I don't.
Continued ambiguity: "I could go to town right now" (with falling and rising pitch on "now"), or I could go later; I just can't decide.

The question that needs to be answered is: Precisely what are the factors that allow normal people most of the time to figure out what the "real meaning" of conversational speech is amidst a barrage of changing pitches? There is, of course, the realm of literal speech where intonation can be safely ignored. "Where is the train station?" will have only one likely interpretation – void of sarcasm, and lacking either positive or negative affect. But once we have arrived, settled in, and are no longer anonymous tourists in need of literal information – then our conversational exchanges will attain greater interpersonal, emotional significance and intonational cues will become relevant.

It is the thesis of this chapter that, despite the complexity of pitch phenomena (in music or speech) and despite sometimes important individual and cultural differences, there are a few threads of common understanding of pitch phenomena – cross-cultural universals of paralinguistic understanding. Moreover, those common threads are related to the psychophysical phenomena of (interval) dissonance/consonance and (chordal) harmony/disharmony. Let it be clearly noted that, as is well known to musicians, any pitch phenomenon can be strengthened, weakened or negated by the musical context within which it occurs – and the same holds true for speech intonation. In a musical piece where nearly all of the chordal harmonies are neatly resolved major or minor chords, a phrase that ends with a slightly-jazzy minor seventh chord will sound jarring, dissonant and unsettled. But that same chord will sound like a return to home-sweet-home after a chromatic sequence of dissonant intervals and tone clusters. Similarly, musical tempo is known to have a strong effect on mood perception, so that an accelerating tempo in a minor key or a slow and decelerating tempo in a major key can elicit an emotion contrary to what would be expected from the pitch combinations alone. Moreover, in discussing the pitch changes in the voice, the "digital" fixed pitches of many musical instruments are no longer found, so that all pitch phenomena are relative. The analysis of voice intonation is therefore more complex than the analysis of diatonic music, but the psychophysical model of pitch "instability" (and its resolution in the direction of either major or minor harmonies), as outlined in Appendix 2, can be put to use in the examination of voice pitch.

So, bearing in mind that intonation phenomena are complex and unrepresentative counter-examples are easy to produce, let us not slip into nihilis-

tic despair or into an equally unproductive cultural relativism with regard to deciphering the regularities in human intonation. Many studies have shown that emotions can be communicated through prosody, particularly voice pitch (reviewed in Scherer 1986, 1992). Intra- and cross-cultural studies using synthetically altered speech indicate that the positive or negative affect of speech is transmitted, even when semantic understanding is impossible[1] (reviewed in Frick 1985; Bergman et al. 1988).

In a review of the diverse literature on vocal affect expression, Scherer (1986) has summarized the principal factors that convey emotion, as shown in Tables 4-1 and 4-2. Precisely how much affective information is transmitted by pitch alone, by tempo alone, by intensity alone, and in what combinations

Table 4-1. The principal acoustic parameters used in the analysis of voice prosody (from Scherer 1986). Note that most are concerned with pitch.

Parameter	Description
F_0 perturbation	Slight variations in the duration of glottal cycles
F_0 mean	Fundamental frequency (vibration rate of vocal folds as averaged over a speech utterance)
F_0 range	Difference between highest and lowest Fo in an utterance
F_0 variability	Measure of dispersion (e.g., standard deviation of F_0)
F_0 contour	Fundamental frequency values plotted over time (intonation)
Fl mean	Frequency of the first (lowest) formant (significant energy concentration in the spectrum) averaged over an utterance
F2 mean	Mean frequency of the second formant
Formant bandwidth	Width of the spectral band containing significant formant energy
Formant precision	Degree to which formant frequencies attain values prescribed by phonological system of a language
Intensity mean	Energy values for a speech sound wave averaged over an utterance
Intensity range	Difference between highest and lowest intensity valuesin an utterance
Intensity variability	Measure of dispersion of intensity values in an utterance (e.g., standard deviation)
Frequency range	Difference between F_0 and highest pointin the frequency spectrum where there is still speech energy
High-frequency energy	Relative proportion of energy in the upper frequency region (e.g., > 1 kHz)
Spectral noise	Aperiodic energy components in the spectrum
Speech rate	Number of speech segments per time unit

Note. F_0 = fundamental frequency. F1 = first formant. F2 = second formant.

Table 4-2. Summary of results of vocal indicators of emotional states (from Scherer 1986).

Emotion	Pitch			Loudness	Tempo
	Level	Range	Variability		
Happiness/joy	High	?	Large	Loud	Fast
Confidence	High	?	?	Loud	Fast
Anger	High	Wide	Large	Loud	Fast
Fear	High	Wide	Large	?	Fast
Indifference	Low	Narrow	Small	?	Fast
Contempt	Low	Wide	?	Loud	Slow
Boredom	Low	Narrow	?	Soft	Slow
Grief/sadness	Low	Narrow	Small	Soft	Slow
Evaluation	?	?	?	Loud	?
Activation	High	Wide	?	Loud	Fast
Potency	?	?	?	Loud	?

for which emotions remain open questions, but there is wide agreement on the idea that prosody is a "second channel" through which people understand one another – often in spite of the contrary linguistic message in the "first" channel. For decades, linguists have focused on the fundamental frequency of the voice (F_0) – not only because it is intuitively the most important carrier of information, but also because most of the physical energy in the voice lies there. Technical developments have meant, however, that many other acoustical features can also be measured and correlated with emotions. As seen in Table 4-2, some rough distinctions can be made among various emotions on the basis of voice acoustics, but there is no simple answer concerning which acoustic factor(s) convey information concerning which emotion(s). In other words, the link between pitch characteristics and specific emotions is quite approximate. Unlike a typical human listener, the psychoacoustician will have problems distinguishing between joy and anger or between boredom and grief solely on the basis of voice pitch characteristics. Clearly, something is yet missing.

Gregory et al. (1997, 2001) have addressed specifically the relative importance of various nonverbal channels in the analysis of dyadic social communications. They have shown that removal of the fundamental frequency from the voice (while retaining the comprehensibility of speech) prevents the normal communication of social status: "It is apparent that the F_0 band plays an important role in transmission of social status and dominance information and that elimination of F_0 leads to lessened perceived quality of conversation" (Gregory et al. 2001). They conclude that, despite the abundance of other nonverbal

channels, the fundamental frequency of the voice is more important than other auditory and visual signals in establishing social relations.

The general ease with which people judge whether or not others "mean what they say" and the use of intonation in both active and passive ways is strong indication that, by and large, we are not robotic "word-processors" relying solely on the decoding of literal word-strings, but rather that we often rely on and act on the basis of the information obtained on hearing another's tone of voice – the most important component of which is fundamental frequency. Here, individual differences are large – both in the sending and receiving of intonational meaning – and the developmental learning curve is perhaps as slow as the acquisition of any psychological skill. But regardless of our individual successes and failures in judging the meaning of the tone of voice of others and regardless of where we are on the learning curve, deciphering the information contained in the tone of voice is an active concern for most people.

The linguist will immediately counter with, "Yes, OK, intonation is important, but every linguistic culture and indeed every subculture and maybe every family has its own 'intonational system' that makes the search for cross-cultural universals futile." In response, I maintain only that there is a small number of intonational universals. Of course there are individual and cultural differences, and even a fully-acculturated, finely-tuned musician's ear may not be able to distinguish between a friend's joy, happiness and elation solely on the basis of pitch. At a much cruder level of discrimination, however, I will argue that there are four universals of intonational phenomena that all normal people understand in a similar way: (i) interval dissonance, (ii) the movement away from dissonance (toward unison or toward consonance), (iii) harmonic tension, and (iv) the movement away from tension (toward major or minor resolution).

It is only with regard to the fourth universal concerning the major and minor modes in traditional diatonic music that the link with human emotions becomes obvious, and that is because there is a popular conception that major harmonies are "positive, happy, bright, cheerful and optimistic" and minor harmonies are "negative, sad, dark, gloomy and pessimistic." That statement requires some qualification and the meaning of major and minor must be reformulated within the framework of the psychophysical model of Appendix 2, but the central argument remains that pitch *combinations* are important for sending and receiving affective information. Just as no sensible music psychologist would attempt to describe the affective content of music in terms of mean pitch, pitch range or other first order statistic concerning the tones in the melody, I maintain that linguists must face the more difficult issue of pitch combinations – musical intervals and, most crucially, musical chords.

The positive/negative valence of harmonic mode has a known relation to pitch dynamics that has been formalized in the diverse styles of diatonic music over the past 1000 years. Within specific, definable harmonic contexts, pitch sequences tell us about the major key/minor key of diatonic music and bring with them an unambiguous mood. When we move away from the fixed diatonic scales, the entire issue becomes more complex, but some related phenomena can be detected in normal speech – provided only that we focus attention on the tonal relations of two- and three-pitch combinations, and do not naively (from a musical perspective) expect to find affect lying solely in the rising or falling direction of pitch melody.

In a word, we hear emotion in the speech of others because they express themselves using pitch combinations that have intrinsic tonal "meaning" that we have experienced in a somewhat different form in the simplified world of diatonic music. There are interesting evolutionary questions about which came first – "meaningful" pitch intonation in speech or the affective modes of singing, and how the pitch phenomena in one realm have influenced those in the other realm (see, Wallin, Merker, & Brown 2000). I will not pursue those questions, but note only that, whatever the evolutionary origins of musical and speech pitch phenomena may be, both can be empirically studied and provide insights into the nature of the human emotions.

A. Intonation and music theory

Traditional linguistic theory through the first half of the 20th Century concentrated on deciphering the nature of the syntactic rules used to bind the semantic units of speech into regular structures that allow for meaningful communication. However, because verbal discourse often maintains semantic coherence despite the fact that syntactic rules have been violated or abbreviated, modern linguistics has broadened its focus to include various "paralinguistic" aspects of spoken language. These include hand gestures and facial expressions, but the main paralinguistic issue in linguistics is the prosody of speech. The study of prosody has been motivated by three factors: (i) Intonational patterns in normal speech can affect the meaning of speech. (ii) Syntax and first-order semantics are only part of the act of communication, as witnessed by the difficulties in constructing computer systems that are able to generate or to understand the higher-order semantic and emotional aspects of human language. (iii) Brain damage can lead to deficits in language that do not affect basic denotative language skills, but do affect communicative skills, specif-

ically the production and understanding of emotional, pragmatic, prosodic and metaphoric speech. Conversely, other types of brain damage can greatly impair language input and/or output, and yet leave certain kinds of paralinguistic (prosodic) understanding relatively unimpaired. The dissociation between, broadly speaking, denotation and connotation is strong indication that at least a part of the meaning in normal human communications is carried in the prosody of speech.

Prosody itself is normally divided into three distinct parts: (i) the rhythm or tempo of speech, (ii) changes in the intensity or loudness of the voice, and (iii) changes in pitch (i.e., intonation). These are three of the four lowest-level factors studied in music theory in deciphering the "meaning of music" (see, for example, Scruton 1997). The fourth factor in music is timbre, or pitch quality. Timbre in speech is related to the so-called formants, known to play an important role in distinguishing among vowels. As central as they are for vowel structure, they can be safely ignored in a first-order examination of pitch in speech, primarily because fundamental frequency is largely unaffected by formant structure and most of the energy in vowels lies in the F_0. It is relevant to note that most regional accents (as distinct from dialects) differ from one another most clearly in terms of the pronunciation of vowels, i.e., systematic changes in formant structure do not make the language unintelligible to native speakers of the same language (provided that consonant structure remains intact). Indeed, in Hebrew, Urdu and related written languages, specifying vowels is entirely unnecessary, and in that respect formant structure is considered simply irrelevant. In any case, the present focus on F_0 means that the entire issue of higher frequency formants will not be addressed.

Within the tradition of Western music theory, rhythm, intensity and timbre are usually considered as factors that modulate the emphasis on certain pitches within a musical composition, and are consequently taken as secondary relative to the important issues of pitch (Schenker 1954; Cooke 1959; Scruton 1997). Pitch itself is studied in two main aspects: the serial changes referred to as melody, and the relations among simultaneous pitches, referred to as harmony. In speech, only melody is relevant, but, in either case, the degree of consonance and/or dissonance among different pitches *and* the harmonic configuration of three or more pitches are the fundamental issues.

In the present chapter, the focus is on the intonational (pitch) aspects of prosody in relation to music theory, and their emotional meaning. While a complete theory of intonation would necessarily include the modulating effects of tempo and intensity, those aspects have already received a relatively large amount of attention (e.g., Lerdahl & Jackendoff 1983). Pitch has of course

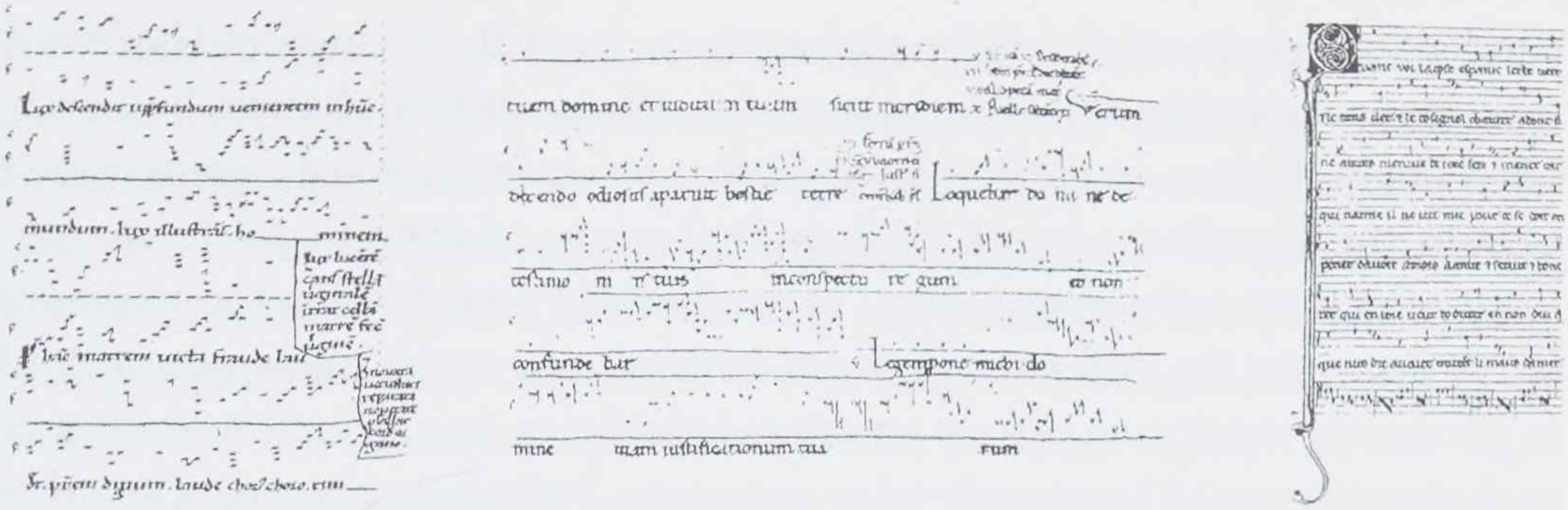
also been studied, but almost never at a level where pitch changes as small as a semitone are brought into consideration (e.g., Frick 1985, "the musical scale is overly precise") and, with a few exceptions (e.g., Fonagy & Magdics 1963; Fonagy 1998, 2001), completely ignoring pitch combinations in a musical sense.

Intonation theory is a relative new-comer in linguistic theory and remains a less coherent body of work than traditional syntactic theory for both conceptual and methodological reasons:

i. Theoretically, there is no over-riding model of intonation (or prosody or paralinguistics) that is comparable to the ambitious universality of the central paradigm of modern linguistic theory, Chomsky's transformational grammar (e.g., Chomsky 1965). Intonation theorists tacitly acknowledge the extreme context- and language/culture-dependence of pitch regularities in speech. In other words, there is, as yet, no "universal" theory of intonation.
ii. Methodologically, intonation theory lacks a widely-accepted, much less a universal, notational system. Many different schemes have been advocated, but none has become dominant – leading to inevitable problems in relating the results of different researchers using different notational systems to one another (Cruttendon 2001).

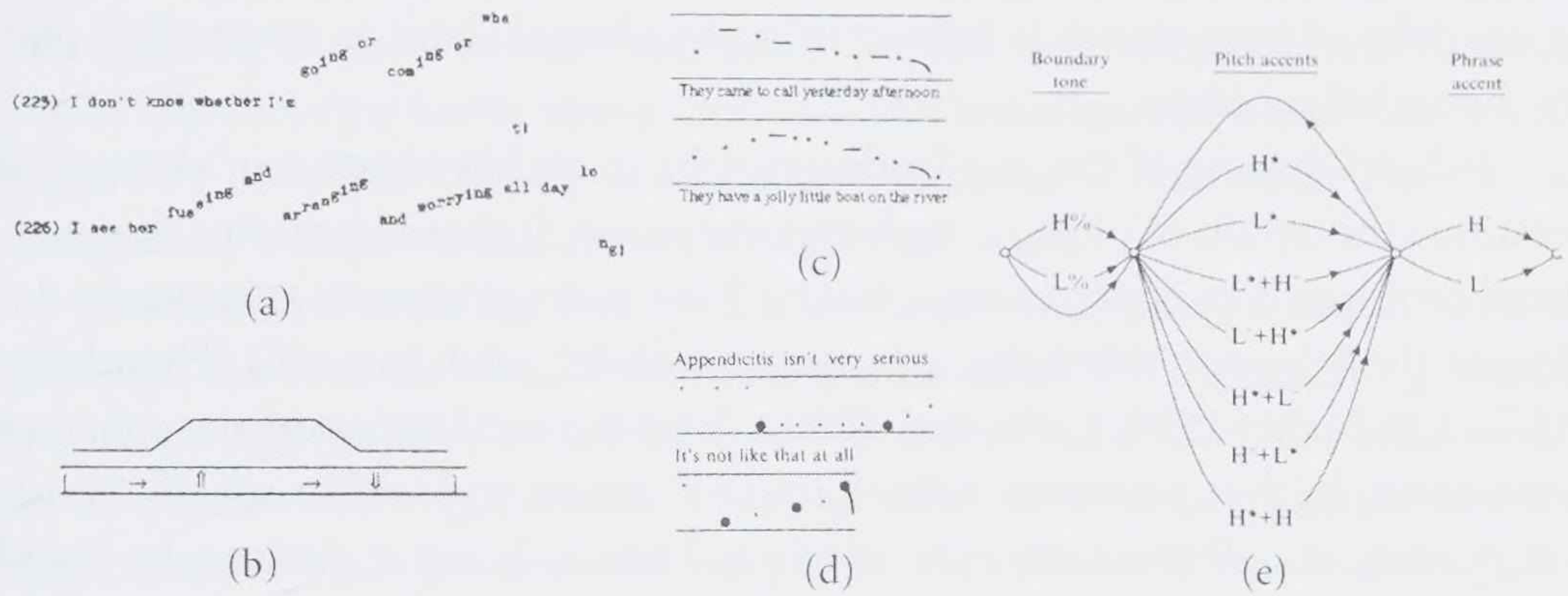
It may be a little bit unfriendly to say it out loud, but the ambiguity of "modern" intonational notation systems is reminiscent of musical notation prior to the emergence of the conventions of modern musical tablature. Up until the 9th Century, a variety of techniques were devised to indicate to the chanters of Christian church prayers what pitch direction and later what approximate interval of pitch changes should be sung (Figure 4-1). Only in the 12th Century did there emerge indication of the duration of specified pitches. And not until the widespread acceptance of the modern tablature in the 13th Century did the notational scheme attain sufficient pitch and tempo specificity that the written notation reliably indicated the acoustic event.

While modern intonation theory may now have an empirical grounding comparable to that of other parts of linguistic theory, the notational techniques currently used (Figure 4-2) contain the same ambiguities as the notational schemes used in music theory a millennium ago! The continued usage of such presumably intuitive, but imprecise notational systems can be explained by the fact that most linguists have not been concerned with the small changes in pitch corresponding to, for example, three versus four semitones, but have been interested in more macroscopic phenomena – for example, the fact that there is rising versus falling pitch, i.e., intonational contours. Because there is a myr-



(a) (b) (c)

Figure 4-1. Early versions of the musical tablature. (a) Gregorian chant tablature from the 9th Century where only the direction of the intervals was explicit. (b) 10th Century tablature where approximate magnitude of pitch intervals was indicated. (c) 12th Century tablature in which the duration of individual notes was also indicated (Grout 1973).



(a) (b) (c) (d) (e)

Figure 4-2. Several current notational systems used for studying intonation: (a) the "scrolling typewriter" system used by Bolinger (1989), (b) the "four line" technique of Gårding (1983), (c) the "close-copy stylization" used by Collier and 't Hart (1981), (d) the "interlinear-tonetic transcription" favored by Cruttendon (1997), and (e) the "finite-state grammar" of Pierrehumbert (1981).

iad of possible descriptions of such contours, numerous notational schemes have been devised – unfortunately, without the discrete structure of musical scales and therefore without the universality of musical notation specified to the semitone level of precision.[2] It is, however, precisely the tacit assumption that a comprehensive theory of speech intonation might be developed *without* attention to the details of pitch that needs to be questioned.

To anticipate the argument developed in this chapter, I maintain that, in the same way that the emotional significance of a piece of music is *highly* de-

pendent on the specificity of the pitches that arise in the melody and harmony (precise to the semitone level), the meaning of intonational changes in speech, often lies in the precise musical intervals that occur in the voice (again, precise to at least the semitone level). To exclude the issue of diatonic pitch from intonation theory is thus equivalent, in music theory, to denying the relevance of, for example, the difference between intervals of a major and a minor third in producing music of a certain emotional character. In music composition or analysis, that would be considered as nonsense. Unfortunately, many linguists have tacitly assumed that the magnitude of the pitch changes in normal intonation are essentially irrelevant or relevant only with regard to the direction of pitch changes – an assumption tantamount to discarding most of the pitch information that is manifestly present in the acoustic signal and that manifestly carries essential emotional significance in music. In the realm of music, pitch is *known* to carry affective meaning that is understood by normal (musically untrained) listeners, and, in the realm of human speech, such information is potentially of importance in transmitting emotional states as an implicit part of the verbal message.

In brief defense of the implication that linguists have been too "sloppy" to examine the details of pitch in studying intonation, it should be noted that a related problem is known in music theory. That is, despite many attempts to delineate the inherent "meaning" of specific, isolated pitch intervals (Helmholtz 1877: 279; Maher 1980; Costa et al. 2000), there is no indication of any coherent consensus. Within narrowly-defined musical genres and within specific musical phrases, the affective meaning of selected intervals can sometimes be stated unambiguously (e.g., Cooke 1959; Kaminska & Woolf 2000). However – and despite the fact that both empirical and theoretical discussions of the meaning of intervals make for fascinating reading! – the musical significance of intervals is context dependent, so that the effort to nail down once-and-for-all the one-and-only meaning of specific diatonic intervals appears to be futile. Linguists are to be commended for not following the musicians down that particular path, but, as discussed more fully in Appendix 1, the insight that some music theorists have pursued (and linguists have not) is that certain three-tone combinations clearly *do* have inherent meaning in so far as they specify a harmonic mode. In other words, two-tone *intervals* are, affectively, "sublexical" units that lack any inherent, intrinsic emotional meaning and that can therefore be used in all manner of emotional music. In contrast, certain three-tone *triads* in music have intrinsic affective implications – i.e., indicating harmonic modality. The significance of three-tone combinations for both music theory and intonation theory lies in the fact that, since certain chords have intrinsic

emotional valence, all pitches and all intervals have affective significance *in relation to those chords*. (The reason why certain three-tone combinations possess inherent meanings that intervals do not is that chords carry within them sufficient tonal "context": the upper and lower tones of a three-tone chord provide a framework within which the third tone can establish a fixed harmony. The meaning of intervals, in contrast, is always subject to change due solely to the location of a third tone; see Appendix 1 for further discussion.)

The hypothesis that will be developed in this chapter is that the coding of emotions in speech involves pitch phenomena related to the major and minor modes. Diatonic music theory is complex and far from non-controversial, and there is no reason to believe that intonation theory will be any simpler, but the essential point is that both intonation and music are concerned, first and foremost, with pitch combinations – accented by rhythmic changes and rises-and-falls in intensity, but with pitch being the main carrier of information. Rather than re-invent concepts, notational schemes and analytical techniques that replicate those in music theory, it is simpler to translate current ideas in intonation theory into the well-established methodologies of music theory and to develop intonation theory in a direction such that it can handle the precise changes in pitch that music theory already deals with.

Both musical melodies and speech are quasi-continuous series of pitches. Although the analysis of scalar music is greatly simplified by assuming that all pitches are approximations of the tones of the 12-tone scale, the total number of pitches and pitch combinations in even short musical pieces can be huge. Because every combination of pitches has potential emotional significance, the total volume of such information over the course of only a few seconds, much less several minutes, is already too large for meaningful study. In musical analysis, there are two principal techniques for dealing with this obvious overload. The first is to undertake a note-by-note analysis of short musical phrases that intuitively have great significance – generally passages that arise at the crescendos and conclusions of larger musical segments (Cooke 1959). The second technique is to contract a musical score into its principal musical components, and then analyze the harmonic significance of the abbreviated version (Schenker 1954).

With regard to the first technique, the justification for analyzing small segments of music or speech lies fundamentally in the concept of the "psychological present" (Poeppel 1997, 2001). In a variety of experimental contexts, it has been shown that sensory information is integrated over short time periods that are subjectively felt as occurring in the "now". The duration of the "now" depends on the modality, behavior and context, but most experimental results

indicate a period of 2–4 seconds, i.e., *a time interval over which recall is accurate without the need for rehearsal.* For this reason, the summation of pitches over a comparable time period in both music and speech is thought to be the essential unit of affective summation. The best example of the psychological "now" in language concerns poetry. The rhyming of consecutive or alternating lines can be heard and appreciated, but not if the rhymes are, temporally, too far from one another. Tempo and voice intensity can be manipulated to prolong the psychological now, but, in poetry read out loud, generally effective rhymes occur within a few seconds of one another.

The literature on the "psychological now" is convincing with regard to the reality of an integration of sensory phenomena over a period of several seconds, but is this true of pitch combinations in particular? Recently, Tillmann and Bigand (2001) have examined the effects of the perception of harmony using chord sequences from classical diatonic music versus the same chords in scrambled sequences. They found that "the effect of harmonic relatedness was unaffected by the scrambling of the sequences." This was true for two different scrambling procedures that were done pair-wise in the chord sequence or in groups of four chords. The focus of that study was in fact on the effects of harmonically related versus unrelated chords, but the result of interest in the present context was that all of the pitch information over the 4 second period of stimulus presentation was evaluated as having similar harmonic effect regardless of the detailed serial structure.

With regard to the second technique (contraction of a complex musical piece into its underlying "essential" harmonic structure), the analytic techniques of Schenker (1954) and his followers are of interest. The principal difficulty in borrowing those techniques for use in intonation theory, however, is that it requires a battery of assumptions concerning how the larger pitch structure of speech can be contracted. Even within music theory, the Schenkerian analytic method is more clearly applicable to harmonic pieces within the so-called repertoire of common practice music from the 16th–18th Centuries, and becomes problematic for other musical styles. To use analogous contraction techniques for the analysis of intonation in speech would require a set of assumptions reliant upon linguistic theory about the relative importance of tones at different grammatical locations within real conversations. The work of Chafe (1994) provides a start concerning what might constitute "intonational units," but many technical problems remains to be solved before unambiguous intonational segments can be abstracted for pitch analysis.[3]

In the intonation literature, there are in fact a few interesting examples of the specifically diatonic musical analysis of speech intonation. Notably, Fonagy

and Magdics (1963) and Fonagy (2001) argued that the use of specific intervals in music and in the intonation of various European languages have specific affective significance and Clynes (1982) provided evidence that anger was often expressed with an interval of a minor sixth. Such works are of fundamental interest, but suffer from their focus on isolated intervals. As vitally important as intervals are as "sublexical" units, the "lexicon" of intonation is more likely to be three-tone harmonic structures (for the musical reasons discussed in Chapter 3 and Appendix 1).

B. The pitch range of normal speech

For many decades, linguists have commented that intonation in normal, "non-emotional" conversational speech varies over approximately a musical fifth (seven semitones, C to G, "Do" to "So," or, in terms of frequency, ranging over half an octave). Such informal comments do not, however, carry much scientific weight, particularly since there are well known age and gender factors influencing pitch production. To address this basic issue, Fitzsimons et al. (2001) had 10 adult male and 10 adult female normal subjects read 10 short sentences twice, first as a declarative sentence and then as an interrogative sentence. The mean pitch range adjusted to semitones per subject is shown in Figure 4-3.

Several points are worth noting. The mean pitch for men and women was of course different, but, when converted to semitones, it is seen that, on average, both men and women used pitch intervals of at least 7 semitones for the 20 sen-

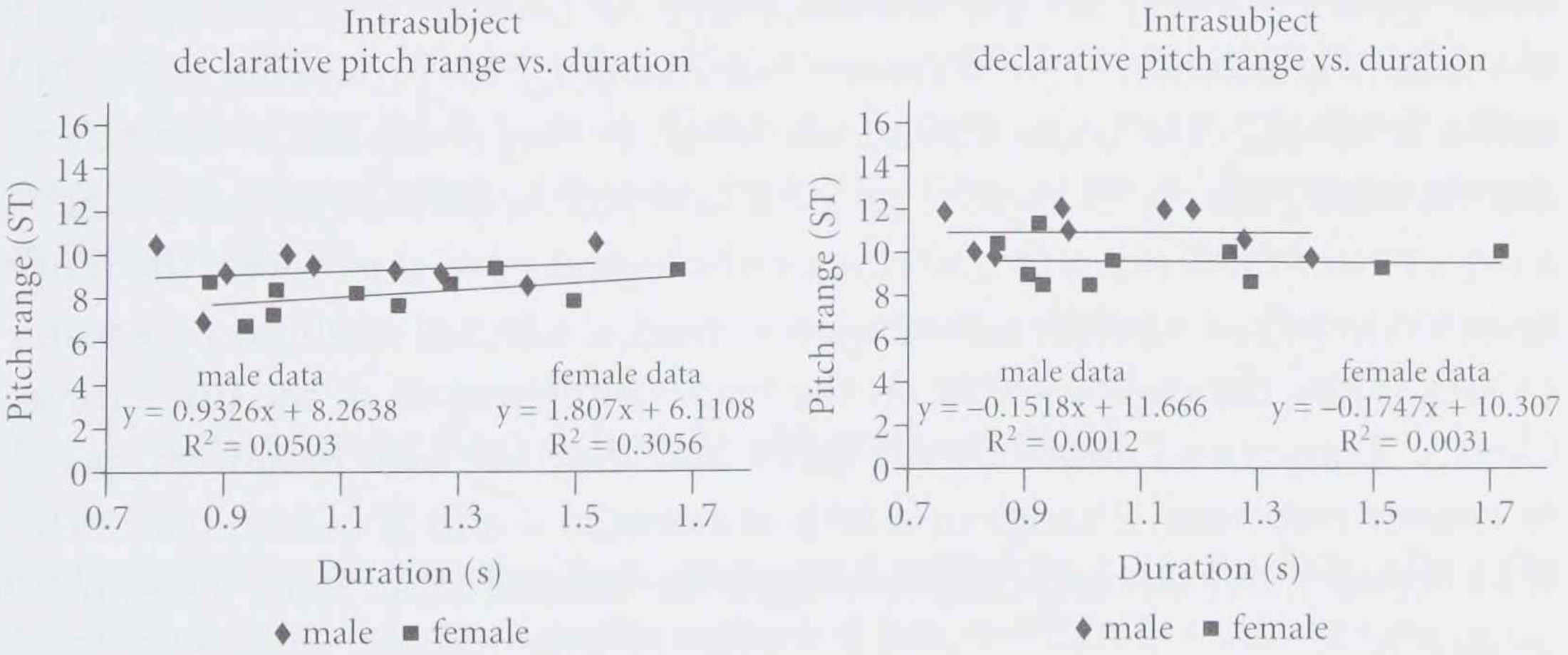


Figure 4-3. The range of intonation produced by 20 subjects during the voicing of 10 declarative (a) and interrogative (b) sentences (Fitzsimons et al. 2001).

tences (Figure 4-3). Examination of the range used per sentence showed that there were a few that did not require 7 semitones, but even for these extremely short sentences containing little emotional content and spoken in a laboratory setting, most were enunciated with a pitch range of 7–12 semitones. Although the human ear is sensitive to extremely small pitch changes (much smaller than a quarter-tone), a relatively large range of pitch is nevertheless routinely employed. Clearly, if the goal of normal linguistic communication were simply to transmit intelligible strings of phonemes with the occasional emphasis placed along the way, there would be no need to vary pitch by more than a few Hertz. In fact, both men and women speak far more "melodically" than is necessary.

In emotional speech that is spontaneous and not the reading aloud of a set text in a psychological laboratory, the pitch range increases still further (Howell & Kadi-Hanifi 1991), but even in the artificial, inherently context-free speech samples used in experiments, frequency changes far in excess of the detectable minimum are employed. Why would that be so? A semitone rise or fall anywhere within the range of human vocalization is easily detected by normal listeners and is often sufficient to indicate emphasis. Why would people go to the trouble of using such a wide pitch range, if it is merely rising or falling pitch that provides the essential intonational information?

Hypothetically, (i) the approximately Do-to-So range of normal conversational pitch could be this large simply to allow for degrees of emphasis – with larger intervals playing the same role of stress that increases in voice intensity can produce. Alternatively, (ii) the magnitude of various pitch intervals might convey information dependent on their magnitude. The former view is the tacit view of most linguists, but there is one crucial fact from traditional music theory that is of obvious relevance here – a musical range corresponding to a fifth (7 semitones) is the *necessary* minimal range for producing major and minor triads (i.e., the major and minor chords in root position). In addition to any affective message of specific melodic *intervals* (sequences of two pitches), a sequence of three tones (within the psychological now) that covers the range from Do to So *can* have the unambiguous musical affect of establishing a major or minor key, depending solely on the pitch configuration of the three tones. Clearly, if major and minor chords (keys, melodies) are used in human speech to convey emotions, a range of a fifth is essential – and is indeed used. This is a necessary, but not sufficient condition for musical affect to be transmitted in speech.

Any interval less than a fifth can of course produce a variety of musical effects, depending on the larger musical context, but musical intervals have an inherent ambiguity with regard to harmonic mode that the major and minor

triadic chords do not have. Within traditional music theory, an isolated interval (two tones anywhere on the keyboard) cannot be unambiguously "major" or "minor". That is to say, the modal meaning of an *interval* depends on the musical context and is affected by the preceding or succeeding tones. For example, two successive tones may constitute an interval of a so-called minor third (C-D#), but the minor third is not truly minor on its own. The location of a subsequent tone can transform the interval of a minor third into a major chord (C-D#-G#) or confirm it as a minor chord (C-D#-G) or leave it ambiguous as to modality (C-D#-F#). In either the major or the minor case, the three-note chord has a perceptible finality and stability that is not felt when hearing a two-note interval lacking a larger musical context. In other words, the major or minor character of any interval of two notes is inherently ambiguous, whereas that of a resolved triad is not. As Cooke (1959) has emphasized, the major and minor chords themselves have a "fixedness" and established feeling that is not subject to contextual reversal. Of course, a resolved major or minor chord can be followed by other pitches, intervals and chords that move in a direction consistent with or contrary to the resolved chord, but the harmony itself cannot be incorporated within a contrarily resolved harmony (unlike an isolated interval that *can* be absorbed within either a major or minor chord).

Normally, a "sad" or "melancholy" feeling is attributed to the minor chord and a "happy" or "joyous" feeling to the major chord (but see the discussion in Chapter 3). However we may wish to characterize the psychology of the "ring" of certain harmonies, chords are apparently more than a sum of their (interval) parts. The establishment of mode requires, at a minimum, the creation of a harmonic "context" that extends over at least a pitch range of seven semitones; the addition of (at least) a third tone can then produce a characteristic harmonic affect. If transmitting affective messages through pitch is one of the things that people do in speech, then a pitch range of seven semitones is a necessary, but not sufficient, criterion for producing "musical" harmonic modes in normal speech.

C. Continuous or discontinuous pitches in speech?

That normal speech varies over a range of sufficient magnitude to allow for major and minor harmonies is an empirical fact, but clearly we do not normally speak like opera singers and the duration of individual pitches in speech is usually quite short relative to those used in music, in general, and opera, in particular. Nevertheless, because we can easily perceive pitch durations on the

order of a few tens of milliseconds, the enunciation of brief pitches of sufficient intensity or sufficiently often may yet suffice to allow us to be cognizant of the "music of language" in each other's voices. So, are there other indications that people select specific pitches and pitch intervals in normal speech, rather than simply slide up and down along a continuous range of pitches in an essentially atonal manner?

A more decisive indication of the musicality of speech can be obtained by summing the pitches used in speech, as shown in the following experiment. A group of 24 20-year old undergraduates (8 female, 16 male) read aloud 12 sentences in the Japanese language. Many had taken piano lessons as children, but none were active musicians. Four sentences were concerned with affectively happy events; four concerned unhappy events; and four were explicitly ambiguous. Prior to reading aloud these sentences of 1–2 second duration, the subjects read a short (3–4 line) scenario that set the scene for the "emotional" sentences. The emotional sentences were constructed to convey stereotypical positive and negative emotions of the kind that a typical 20-year-old might actually experience (e.g., "I won a computer in the lottery" and "My grandfather died last night"), and situations that were ambiguous with regard to affect ("I am not sure if I should go or not."). A typical recording (with pauses removed) is shown in Figure 4-4.

Pitch and intensity values were obtained at intervals of 10 milliseconds (using the remarkable freeware, Praat™, by Boersma and Weenink 2002). The product of frequency and intensity was calculated, and assigned to the nearest semitone in an equitempered scale (A = 220 Hz) – resulting in essentially a semitone power spectrum. This normalization procedure resulted in a semitone histogram for each sentence, examples of which are shown in Figure 4-5. For each sentence for each subject, a "tonic" tone was then determined from the most frequently used semitone in the lower half of the histogram. Summation of all the data adjusted to the tonic tone for the 24 subjects for the three categories of sentences is shown in the histograms of Figure 4-6 through 4-8.

The most remarkable aspect of these results is already apparent in Figure 4-5: most speech utterances have notable interval or chordal structure. Clearly, a description of the pitch characteristics in terms of mean pitch, pitch range or pitch variability would be fundamentally misleading in so far as the pitch distribution is bimodal or trimodal. In fact, a minority of utterances were unimodal (10%) or showed four or more pitch peaks (11%), but most were either bimodal (45%) or trimodal (34%). The interval structure is apparent in the raw pitch data, but subsequent cluster analysis often revealed that individual peaks consisted of two overlapping Gaussians.

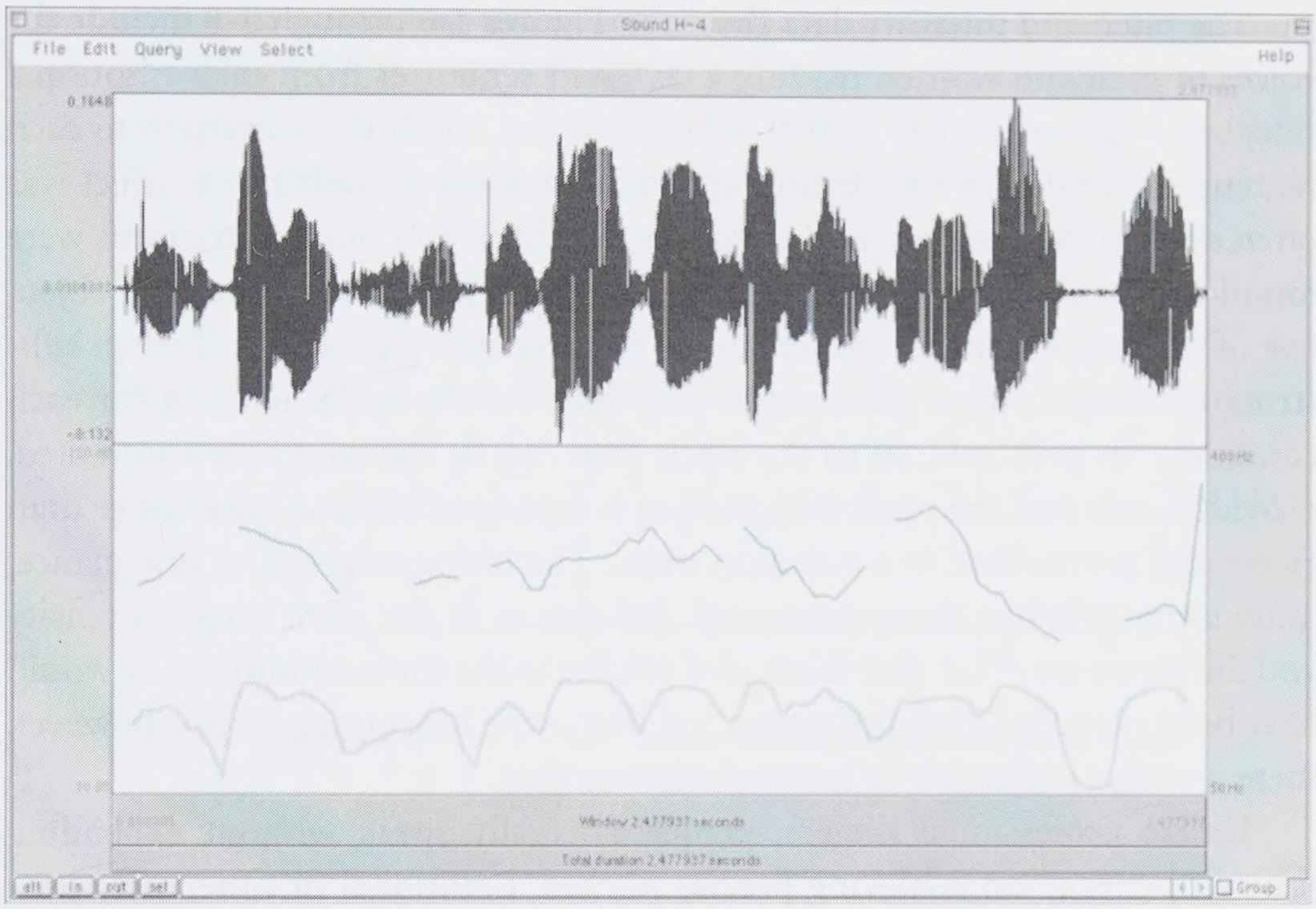


Figure 4-4. The edit window of the Praat™ software, showing the sound wave (above), the changes in pitch (middle) and the changes in intensity (below).

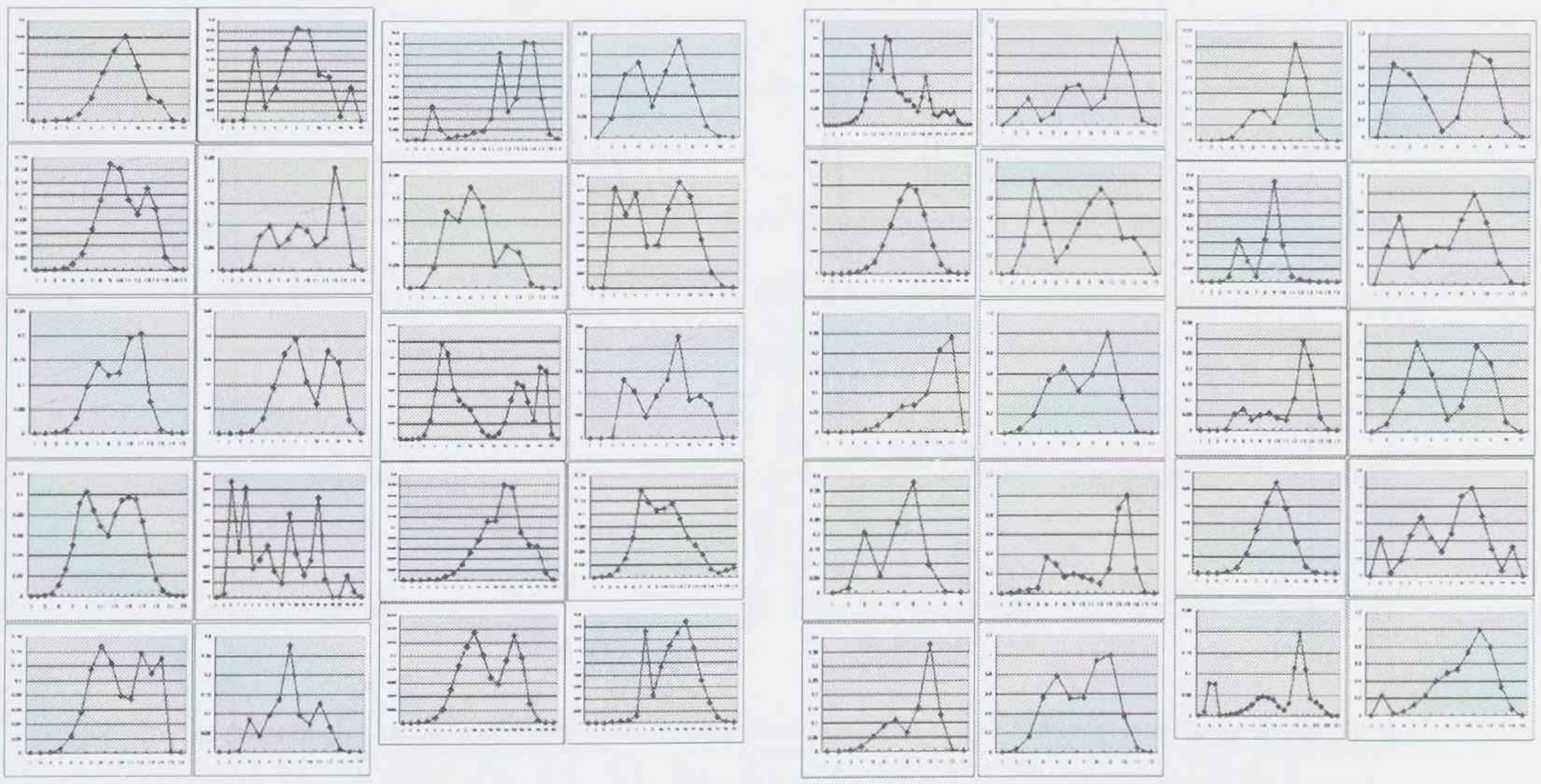


Figure 4-5. Typical examples of the pitch structure of short (~2.0 second) "emotional" sentences. Amplitudes are relative and pitch frequencies have been normalized to an equitempered 12-tone scale. On the left are 20 "sad" sentences, on the right are 20 "happy" sentences. The two-tone interval or three-tone "chordal" structure of most utterances is noteworthy.

The pitch and intensity data shown in Figures 4-6 through 4-8 include the voices of men and women reading 4 different sentences from each emotional category – approximately 18,000 data points per condition. Listening to each sentence individually, the slightly upbeat, downbeat or ambiguous affect was often evident, but, as seen in the Figures, very different pitch histograms were found for the different sentences within the same condition. Since the objective of the experiment was to search for macroscopic regularities of pitch substructure due to affect, the analysis was done on the entire data set for each condition. In principle, all of the pitch data can be meaningfully summated, provided only that the pitch information is expressed in semitones rather than Hertz and normalized to a common tonic. The histograms are, at first glance, quite unremarkable: there were small differences in the pitch range, intensity and the skewness of the distribution of pitches in the three conditions, as would have been expected from the intonation literature. Most importantly, however, there was also indication of interval substructure.

Unlike professional singers, people normally speak without expending time and energy on achieving precise pitches. Emotional or otherwise, people are after all talking, and therefore focusing attention on the meaning of words rather than the tone of voice. A rather relaxed, approximate control of

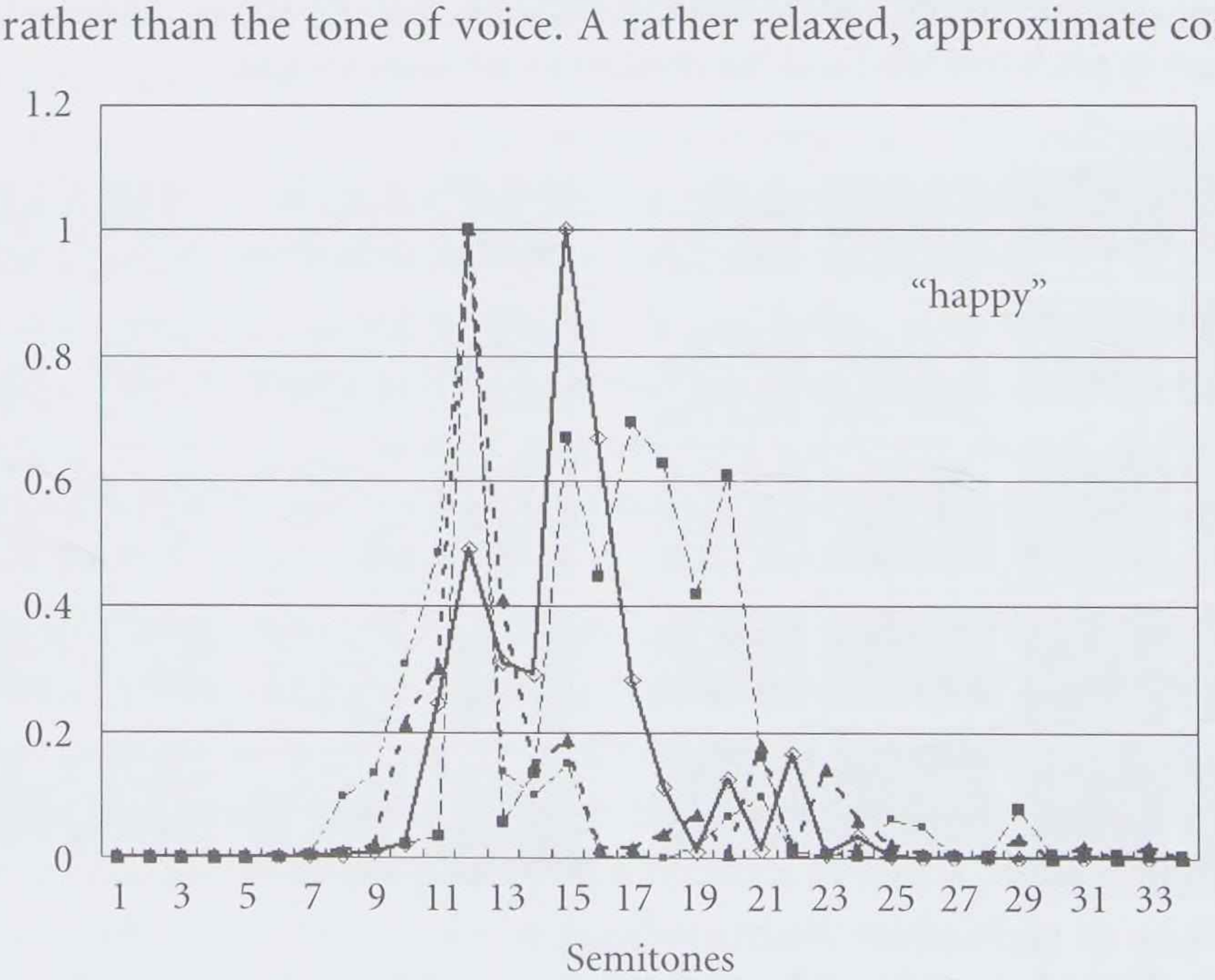


Figure 4-6. The pitch histograms for the four "happy" sentences. Amplitudes are relative and pitches are in semitone steps normalized to preserve the interval between 12 and 15 semitones (thanks to K. Takami, M. Hanazono and T. Kawashima for data collection and analysis).

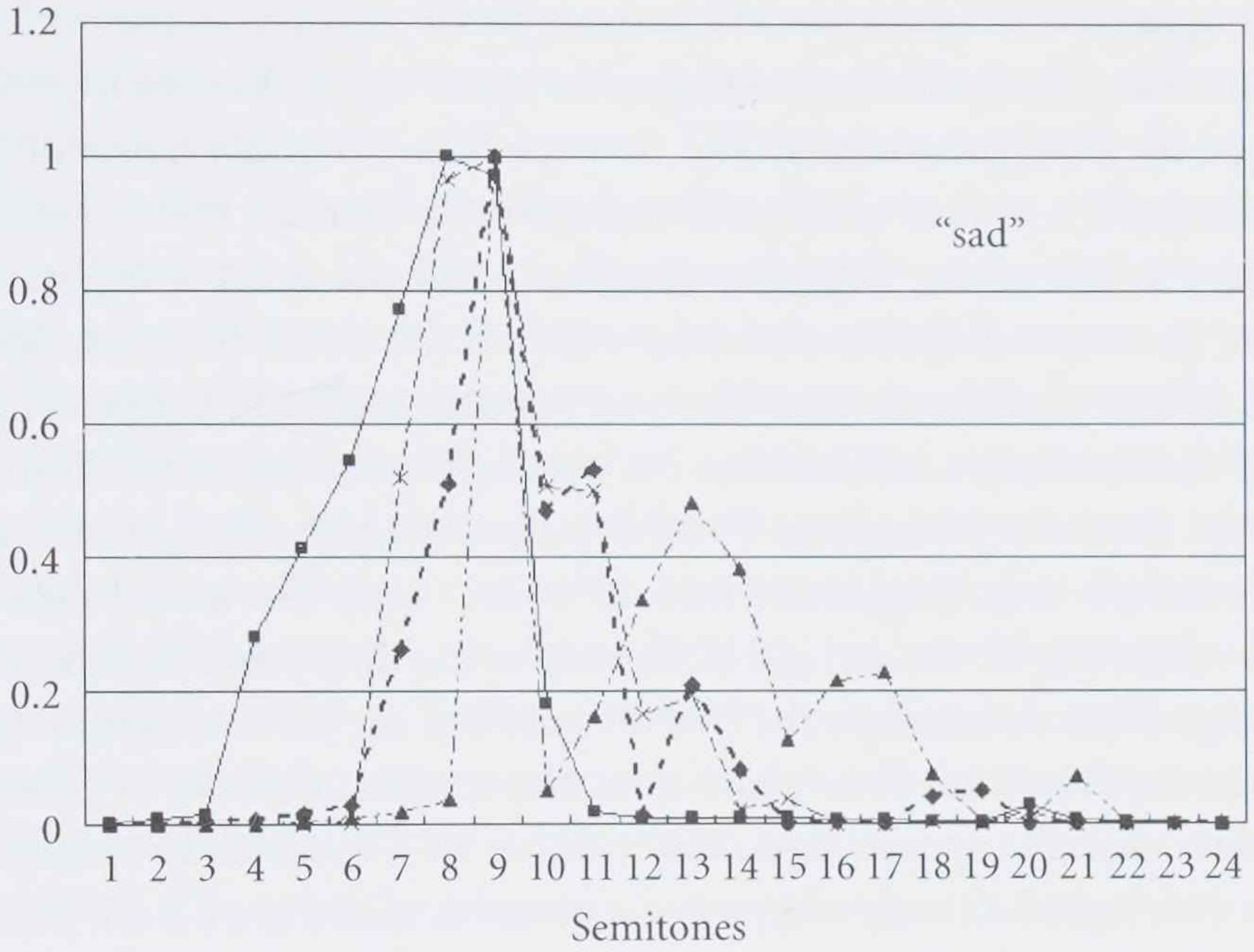


Figure 4-7. The pitch histograms for the 4 "sad" sentences (thanks to Takami et al. for data analysis).

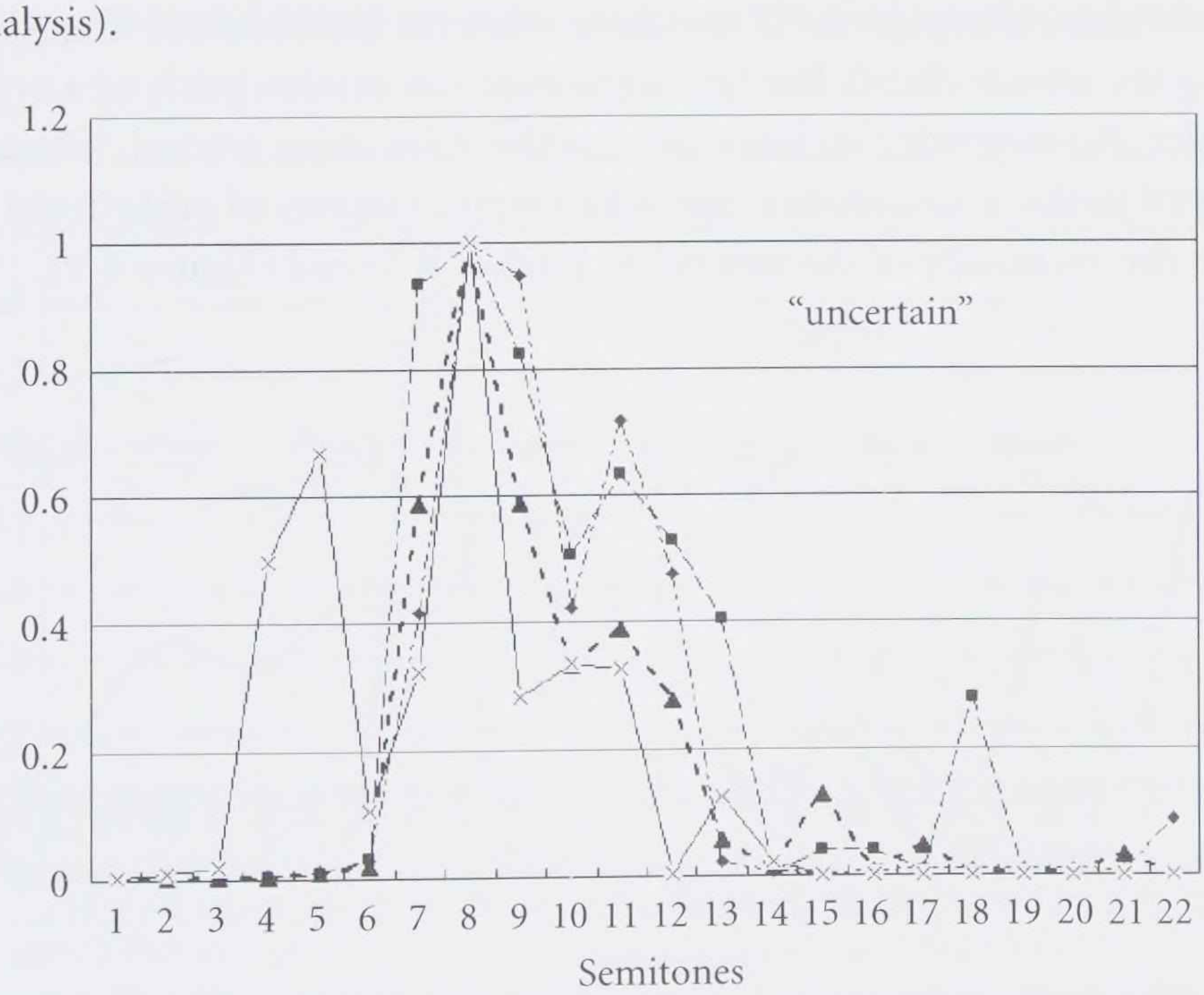


Figure 4-8. The pitch histograms for the 4 sentences with "uncertain" emotional valence (thanks to Takami et al. for data analysis).

the pitch in normal speech would therefore lead to broad peaks of pitch in a pitch histogram. If it is assumed that those pitch peaks nonetheless have a normal Gaussian distribution around specific scalar tones, then each pitch can be described by three parameters. The first specifies the pitch frequency (the location along the x-axis at which the maximum is found); the second is its intensity (the height of the bell-shaped curve); and the third is the standard deviation of the normal distribution (the width of the bell at half its height).

The analytical problem is therefore how to optimize the fit between the experimental data and a set of Gaussian curves corresponding to the dominant pitches in the semitone histogram. For this purpose, a technique known as radial basis functions (e.g., Poggio & Girosi 1990) was used. The actual algorithm was that developed by Bouman (2002). By employing maximum likelihood criteria, the algorithm determines the best fit between the sum of several Gaussians and the experimental data, while using the smallest number of Gaussians to fit the data. For the speech data, the number of Gaussians is usually two, three or four – even when only one peak is apparent (as in Figure 4-5). The general idea applied to speech is easily understood from the radial basis functions of a minor chord (Figure 4-9). A spectrogram of the chord when played on a piano would show sharp peaks at the three relevant fundamental frequencies comprising the minor chord, but the somewhat less certain pitch of a typical human voice shows greater variance around the three main pitches. Instead of three isolated peaks, a mountain range with varying degrees of peakedness that depend on the variability of the individual pitches is found (Figure 4-9).

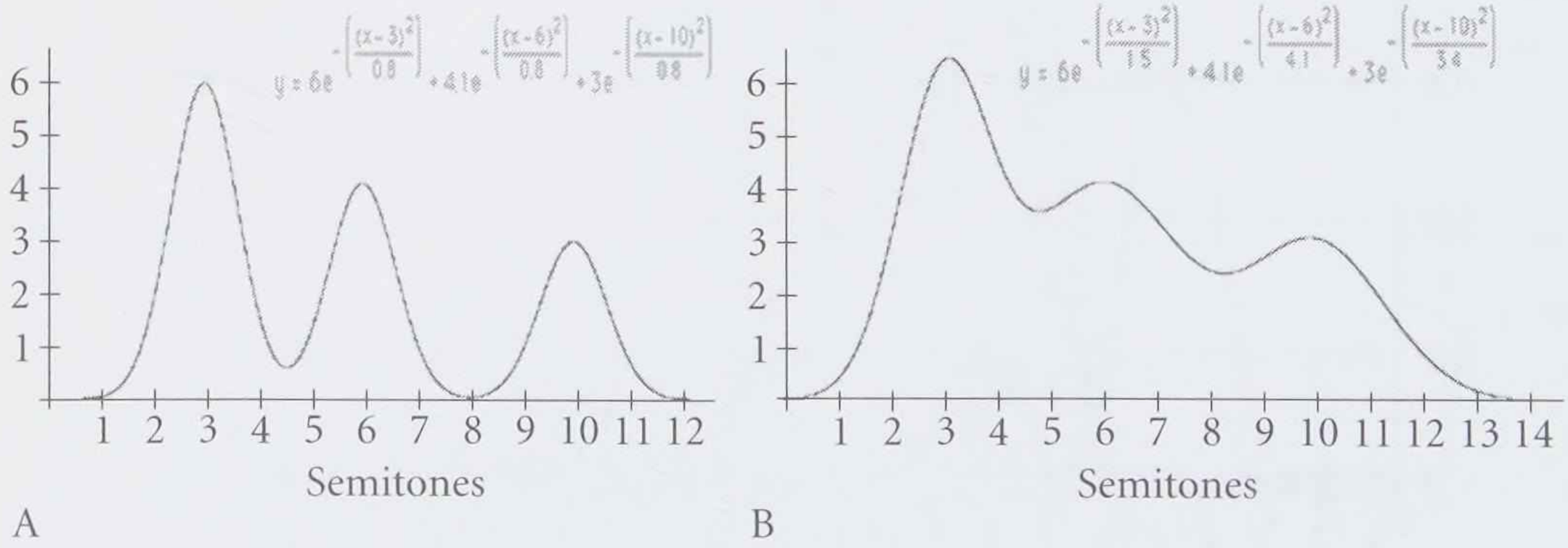


Figure 4-9. The semitone structure of a minor chord with (A) relatively well-defined or (B) poorly-defined pitches and varying amplitudes. The *pitch* substructure is identical in both (A) and (B) – with peaks at 3, 6, and 10 semitones, while the widths of the Gaussians differ.

Using this radial basis function technique for analysis of the intonation data, it is simply a matter of fitting a number of such Gaussians to the empirical pitch curve, and drawing conclusions about the pitch relations in the speech from the parameters of the curve. The experimental data shown in Figures 4-6 through 4-8 have been fitted with such Gaussians and are shown in Figures 4-10 through 4-12. What is found is that the pitch curves for all three conditions are well-fitted with 3–4 Gaussians, the pitch locations of which are different in each of the three emotional conditions. Specifically, in the "happy" sentence condition the three loudest pitches are located at the tonic, three semitones above the tonic, and eight semitones above the tonic (Figure 4-10). Musically, this corresponds to a major chord in first inversion. If, however, the fourth peak is also brought into consideration, the chord becomes a minor seventh – which is more difficult to interpret as indicative of a happy mood! In the "sad" sentence the three strongest pitches are located at the tonic, and at three semitones below and four semitones above the tonic. This corresponds to a minor chord. The small fourth peak changes the chord into a minor seventh with an added eleventh – still minor, but more ambiguous in character. Finally, in the affectively "uncertain" sentence, the pitches are at the tonic and three semitones above and below the tonic, i.e., an unresolved diminished chord.

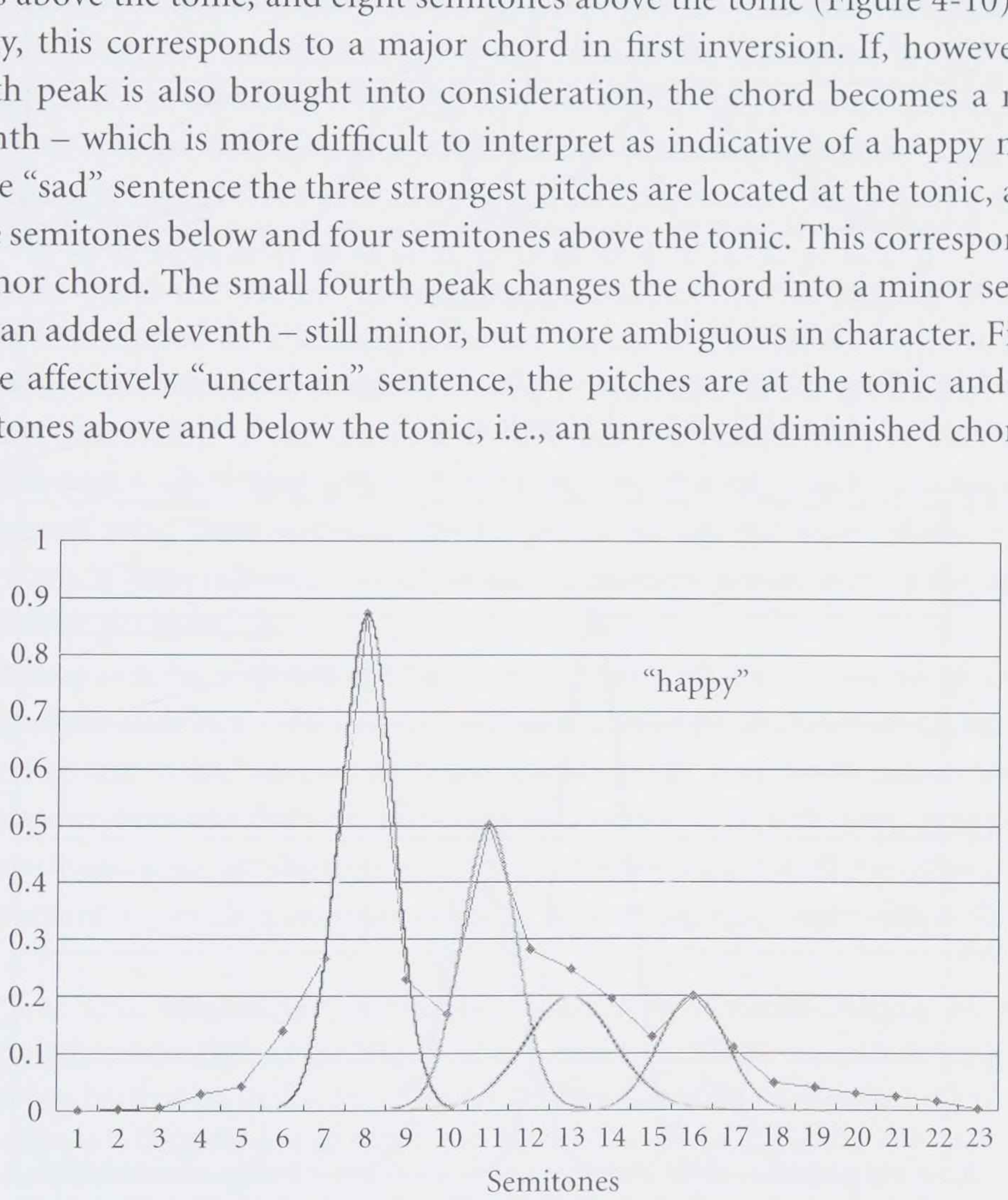


Figure 4-10. The pitch histogram for the "happy" sentences fitted with four Gaussians. They are spaced at intervals of 3, 2 and 3 semitones. The upper curve shows the experimental data.

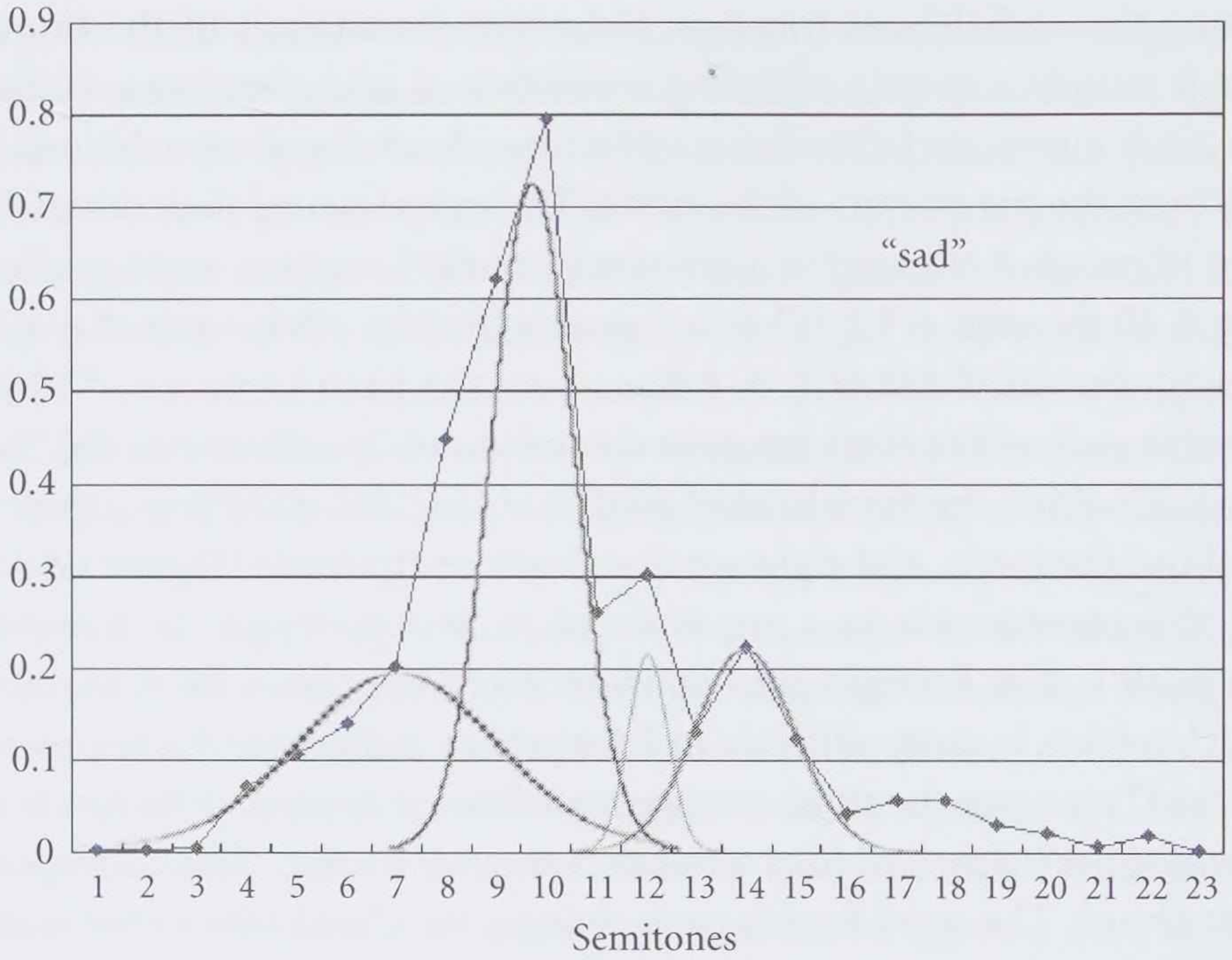


Figure 4-11. The pitch histogram for the "sad" sentences fitted with three Gaussians. They are spaced at intervals of 2 and 3 semitones.

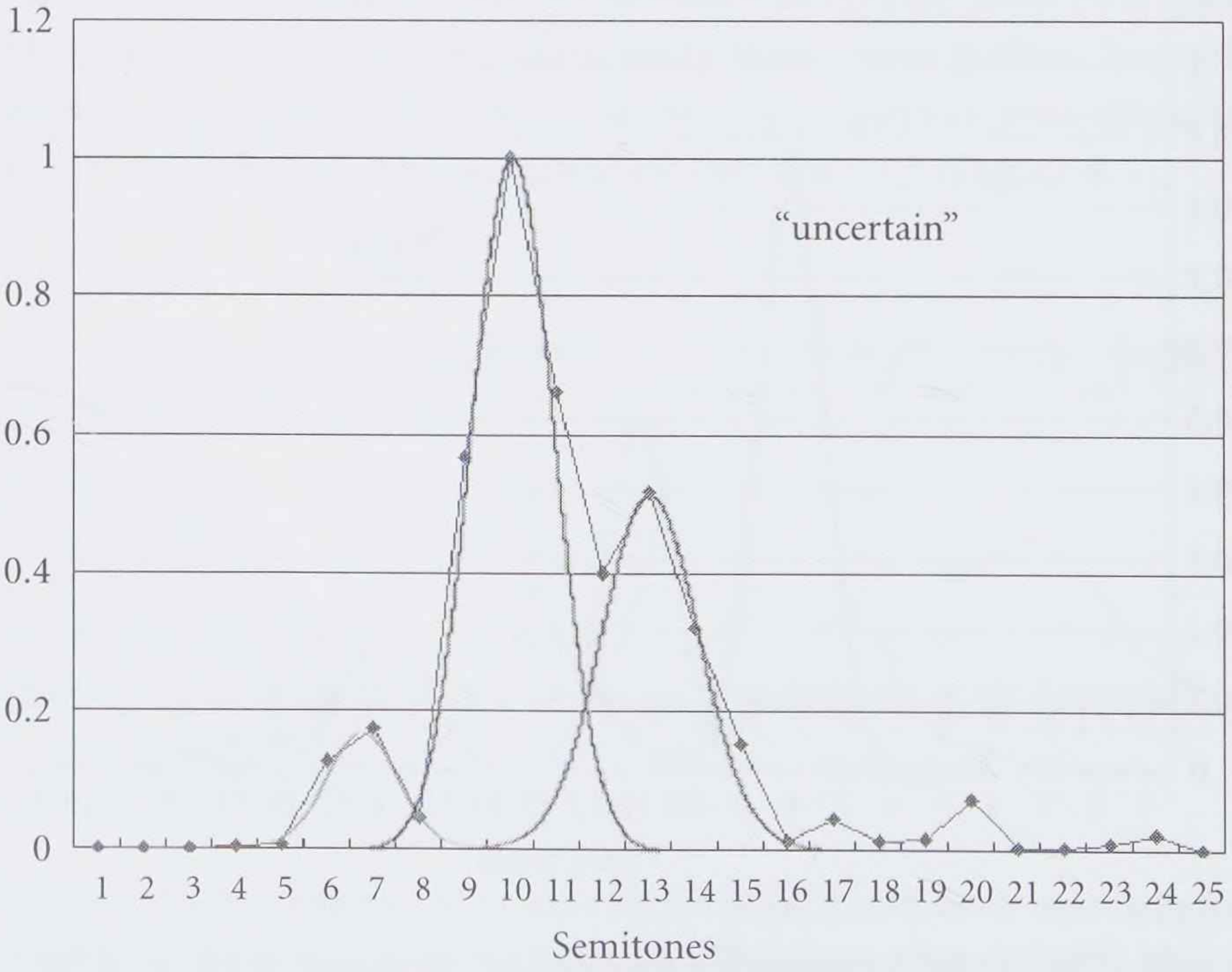


Figure 4-12. The pitch histogram for the "ambiguous" sentences fitted with three Gaussians. They are spaced at intervals of 2 and 2 semitones.

Despite the fact that clear triads were only infrequently found in the individual data (Table 4-3), the group data revealed harmonic phenomena that are clearly suggestive of the expression of mood through musical mode. The generality of this phenomenon remains to be examined, and the patterns within the individual data require further study, but it is certain that the harmonic structure of emotional speech can be studied in a quantitative manner by searching for the radial basis functions that underlie what otherwise appear to be complex pitch-intensity curves. In conclusion, it can be said that the group results are in general agreement with the musical phenomena reviewed in Chapter 3. Although it is extremely unlikely that normal speech is diatonic at the individual level, the non-diatonic model of pitch "instability" discussed in Appendix 2 may provide a means for evaluating the "instability" or "tension" in the pitches used in a spoken sentence, and moreover to determine which of the two directions of resolution are indicated by the size of intervals relative to a state of chordal tension.

In so far as the use of "intervallic equidistance" in the melody of speech can be considered as a kind of affective "tension" (Meyer 1956), any movement of pitches away from tension should indicate a trend toward positive or negative resolution of the tension. Even when the pitches do not correspond to a diatonic scale (using any particular tuning system), chordal tension and movement away from tension can be perceived on the basis of the relative size of pitch intervals without introducing musical concepts that are specific to diatonic phenomena.

Strong conclusions cannot be drawn from only one experiment using a new analytic technique. The experimental evidence presented above is certainly suggestive, but raises almost as many questions as it answers concerning the relationship between diatonic harmony and non-scalar pitch combinations. At the very least, however, these results are clear indication that the *relative interval substructure* in speech intonation should be investigated – not within the rigid

Table 4-3. Triad substructure in the three affective conditions. Diatonic harmonies were found in only 18% of the 288 recorded sentences. There was a clear tendency to use tension harmonies in the emotionally-ambiguous condition, but major and minor triads were not frequent in any of the conditions (thanks to K. Sakurai for analysis).

	Triads			Other pitch combinations
	Major	Tension	Minor	
Happy	5	8	8	75
Ambiguous	4	12	2	78
Sad	2	4	7	83

framework of the fixed scales of diatonic music, but taking important hints from music theory and music psychology concerning the emotive force of pitch combinations.[4] The relationship between pitch substructure and emotion will be addressed once again in Chapter 5.

D. Methodological questions

We have run a variety of related experiments to determine the degree of "musicality" in normal speech, and have learned the hard way why intonation theory is so complex! Nonetheless, three consistent findings are worth noting: (i) a pitch range of 7 or more semitones is normally used in whole sentences; (ii) the distribution of pitches is not simply a normal (or skewed normal) distribution around the speakers preferred tonic, but shows distinct interval and chordal substructure; and (iii) by assuming that the intonation of the normal speaking voice is the sum of several discrete pitches each with a normal distribution, certain regularities of harmonic substructure can be seen. Tentatively, I conclude that the dimension of positive versus negative affect may be detectable solely in terms of pitch substructure, but subcategories of positive and negative affect appear to require consideration of pitch amplitude and/or the sequential ordering of pitches. The "semitone histogram" technique outlined above completely ignores the time series effects of changing pitches, so that any information that is inherently temporal is excluded from the analysis. Although there is some justification for ignoring the time domain when considering all pitch events within the psychological "now," it is specifically the temporal changes that have been the focus of most previous intonation studies on pitch contours. As in musical phenomena, it is unlikely that simply an alteration in the sequence of pitches that conveys positive or negative affect could transform a minor mood into major, or vice versa. In other words, for distinguishing between positive and negative affect, the time domain may not be crucial, but the ascending pitches that signal an interrogative versus a declarative sentence clearly rely on the temporal sequence. For this reason, the quasi-Schenkerian harmonic analysis discussed above will undoubtedly require some introduction of the temporal changes in pitch to distinguish among subtypes of positive or negative emotional states.

Most of the remaining methodological questions are essentially empirical in nature – and can be resolved in the laboratory. Two of the main problems that we have faced but not adequately resolved are (i) the transformation of speech pitches to a 12-tone scale, and (ii) the normalization of different speak-

ers to the same scale for the analysis of group data. Using a 12-tone equal-tempered scale with a fixed pitch of middle C (261.6 Hz), all pitches can be assigned to the nearest semitone for diatonic analysis. By adjusting the central frequency of each pitch in the diatonic scale to more closely match the speaker's chosen frequencies, small changes in the semitone histograms are induced, but their significance remains uncertain. Similarly, by taking the speaker's "tonic" pitch to be that pitch most frequently used in each spoken sentence, group data analysis is possible, but other normalization techniques are possible and would again lead to somewhat different results. Specifically, from a musical perspective it might be more valid to normalize to the speakers' "dominant" pitches above the tonic, and then study the interval substructure.

Some linguists already despair of finding any intonational regularities common to any one language, much less common to all of humanity. While it is not possible to declare with certainty whether that pessimism is the realism of experience or a premature abandonment of the task, it is certain that the psychoacoustics of intervals and harmony (Chapter 3), i.e., two- and three-tone combinations, have *not* previously been examined in intonation studies. From a music theoretic perspective, it is clear that musical mode *cannot* be defined by first-order statistics of F_0 (mean, range, standard deviation, etc.). To define harmonic mode, pitch combinations need to be studied – and the same probably holds true for voice intonation.

The experiments we have run thus far indicate that the intonation of normal speech contains more "musicality" than would be expected by most linguists, but true melodies and triadic harmonies in conversation are apparently infrequent. Nevertheless, when, in normal conversation, we want to add emotional emphasis or drive a point home, we naturally use sustained pitches and larger pitch intervals. It is as if truly "musical" pitch is used sparingly in speech until it becomes appropriate to supplement the denotative meaning of language with an affective "punch". Study of the "musicality" of highly emotional speech in more ecological settings is therefore a priority.

If indeed the expression and understanding of emotion occurs essentially through pitch, and specifically through pitch phenomena related to those used to establish major and minor modes (dominant and subordinate moods), then the pitch phenomena that (i) establish, (ii) reinforce, (iii) modify, and (iv) destroy the major and minor modes should produce corresponding changes in mood. As such, this hypothesis concerning the pitch phenomena underlying human emotions acknowledges only two fundamental affective directions. However those two moods might be characterized – happy and sad, positive and negative, pleasant and unpleasant, favorable and unfavorable, optimistic

and pessimistic, etc. – they are likely to have their evolutionary origins in animal calls used to establish dominance relations.

From a purely theoretical point of view, it would be of interest to develop a more elaborate pitch-based theory of the emotions, in which various pitch combinations on a non-diatonic analog scale correspond to each of several "basic" emotions, but let us remain on the relatively solid ground of a model that distinguishes only between harmonic stability and instability, and identifies only two forms of stability corresponding to positive and negative affect. How could such a theory of emotions be implemented in the brain?

Notes

1. As Levelt (1999:108) has commented: "Although emotion can be expressed at all levels, from the semantic to the prosodic, the most immediate expression of emotion in speech is through pitch movement." Abercrombie (1967:103) has also noted: "Fluctuation in the pitch of the voice is probably the most important of the features of voice dynamics. It owes its importance partly to its outstanding role as a bearer of affective indices. These indices, together with affective indices conveyed by register differences, make up what is meant by 'tone of voice', and the flow of conversation much depends on them."

2. The unsatisfactory disarray of intonation theory has often been noted, e.g., "In our opinion, some explicit scheme for describing prosody must replace the vague, intuitive, and theoretically unmotivated descriptions psychologists [!] have often used in the past" (Cutler & Clifton 1999). The problem remains, however, that all linguists who have come to this realization have immediately put forth their own idiosyncratic scheme in an unrealistic anticipation that the psychologists would see the error of their ways. A partial exception to this rule has been the quasi-musical system advocated by Lerdahl and Jackendoff (1983). Their insufficient treatment of pitch phenomena greatly limits the utility of their system, but they have pointed out the value of musical concepts of rhythm for describing speech prosody.

3. Jackendoff (1999, 2002) has argued that linguistics needs to examine a level of organization lying between lexical units and whole sentences, "hierarchical phrase structure. Such mid-sized semantic units may constitute the units of a lexicon of affect. Using the technique for marking prosodic structure advocated by Lerdahl and Jackendoff (1983), it may be possible to extract only the pitches relevant to the affective message – effectively discarding the "passing notes" in the linguistic melody. In principle, such selection of pitches in speech would clean up the pitch structure of a sentence, making for less ambiguous pitch analysis (see Section C). Ultimately, some such technique would be desirable, but since it demands assumptions concerning "relevant" linguistic structures, it will not be pursued here.

4. The fact that the relative sizes of pitch intervals, rather than their absolute sizes, are important in speech has long been recognized. For example, Abercrombie (1967:107) states: "The intervals between the points in the pattern are absolute and constant in the patterns of musical melody, but they are relative and variable in the patterns of speech melody. Thus a

pattern in speech melody can be either compressed or expanded in the dimension of pitch and still remain the same pattern, although in one case the intervals are smaller and in the other larger." The problem Abercrombie points to here is absolutely fatal if only pitch intervals are considered, because any interval can be expanded or compressed to have any possible affect. However, study of the relative size of multiple intervals, i.e., chords, allows for quantitative statements about harmonic patterns that do not rely on the "absolutes" and "constants" of musical melody.

Chapter 5

The brain code

Synopsis

The cortical representation of pitch and pitch combinations on tonotopic maps is discussed. Together with earlier conclusions concerning voice intonation, harmony and the central dogma, a theoretical argument concerning the code by which the brain stores and transmits emotions is developed.

The central dogma outlined in Chapter 2 is unabashedly a "dogma" – a set of assumptions with diverse empirical support, but, on its own, the dogma is little more than one of several ways of thinking about the human brain. There is no shortage of other theories about the relationship between the brain and behavior that focus on other neural structures, specific genes, a particular neurotransmitter or even a special ion. Each such theory is arguably a "dogma" in the sense that it marshals the evidence to provide a coherent "story," but also demands the acceptance of certain starting assumptions concerning causality and the setting of priorities. Logical analysis of the dogma itself is unlikely to provide an unambiguous conclusion concerning its validity or invalidity – in much the same way that the analysis of religious dogma is unlikely to "prove" or "disprove" a religion.

The inherent difficulty in pronouncing unambiguously for or against any given dogma means that the validity of the dogma must be evaluated on the grounds of what it implies concerning lower level mechanisms. This has certainly been the case in the field of molecular biology. Originally (in 1957), the central dogma was little more than a "plausible story," but the subsequent details of the genetic code were found to be consistent with the dogma, so that both theoretical arguments have become fully established and guide virtually all thinking in cell biology today. In so far as the psychological dogma advocated here can be seen as a more-or-less coherent "story," it is fair to ask what it implies for lower level (neuronal) mechanisms – i.e., the brain code, and what specific predictions can be made. Unlike the relatively "macroscopic" central dogma itself, for which diverse, but not definitive, arguments are available, the details of the brain code are, by their very nature, "microscopic" and explicit

enough that conclusions concerning physiological processes can be drawn and subjected to empirical testing.

Having taken a lengthy detour through the realms of music theory and intonation phenomena, let us now return to the question of the "brain code." How do the pitch phenomena of speech (and music) relate to the central dogma and to an understanding of how information is stored and relayed within the brain? What is known empirically and has already been discussed in the previous chapters can be summarized as follows:

1. The left hemisphere is dominant for language – both the literal understanding of the speech of others and the control of the organs of speech for verbal expression. When people speak, they speak through the left frontal cortex – roughly, Broca's area and the associated motor cortex. And when people understand speech, they understand grammatical structures through Wernicke's area and the associated left temporo-parietal cortex.
2. It is also known, however, that damage to the right hemisphere or the corpus callosum results in the loss of the normal production and understanding of the prosody of speech – indicating that the right hemisphere is also involved in language processing at a "paralinguistic" level. The fact that callosal damage affects intonation in a manner similar to right hemisphere damage shows that, although the actual processing of pitch phenomena occurs predominantly in the right hemisphere, information is transmitted to the left hemisphere across the corpus callosum for the generation of intonation in normal speech.
3. From the psychological literature on music perception, it is known that the right hemisphere is superior, relative to the left hemisphere, at the encoding and recognition of musical intervals and harmonic modes. Although rhythm and tempo are of great importance to real music (and are more capably handled by the left hemisphere), the pitch aspects of music – melody and harmony – are handled predominantly by the right hemisphere. The emotional effects of music are also known to be strongly influenced by the usage of certain pitch intervals, chords and keys (with tempo, timbre and intensity playing supporting roles). Since pitch information is processed predominantly by the right hemisphere, it follows that, for the verbal expression of affect through the pitch changes of music or speech, the affective state of the right hemisphere must be represented on the cerebral cortex in a form that can be transferred to the left hemisphere.
4. The affective valence of music is largely determined by major and minor modes. In the harmony model discussed in Chapter 3, the positive and neg-

ative valence of mode in diatonic music was shown to have a simple structure in relation to a state of harmonic tension. Specifically, if three pitches intoned within the psychological "now" produce two intervals of the same magnitude, unresolved harmonic "tension" is perceived. From such a state of tension, a decrease in pitch resolves toward a major mode, and an increase in pitch resolves toward a minor mode. Probably for evolutionary reasons related to animal calls, most people hear major mode harmonies as having positive affect, whereas minor mode has negative affect. The use of certain combinations of pitches in emotional speech is suggestive of a similar phenomenon occurring in speech intonation.

To those four basic findings, the present chapter will add two more concerning the cerebral cortex and the corpus callosum. In abbreviated form, they can be stated as follows:

5. Auditory information is processed and stored cortically in so-called tonotopic maps – on which auditory frequency is one of the dimensions of organization.
6. The area-to-area, homotopic, topographical connectivity between cortical areas across the corpus callosum implies that a pattern of neuronal activity on a 2-D tonotopic map can be relayed contralaterally in several ways that depend on details of callosal physiology.

With the addition of these two relatively uncontroversial facts about the human brain, we have the necessary building-blocks for a brain code: (1) the anatomical pathways, (2) the physiological mechanisms, and (3) the psychological content of what goes on between the cerebral hemispheres of the human brain during music and language processing. In other words, we know enough about brain mechanisms that the central dogma argument need not remain an abstract one concerning disembodied cognition that, in some unspecified way, occurs in neuronal "black-boxes". The point – the entire basis for calling it the "brain code" – is that, for certain forms of characteristically-human information-processing, we know the basic building-blocks that underlie how the brain performs (linguistic) cognition and processes (intonational) emotion.

Reduced to its skeletal structure, the argument concerning the roles of the cerebral hemispheres in language functions can be easily stated. Words are arbitrary phonetic units that are associated through learning experiences with concrete or abstract objects, conditions and actions. How auditory sounds were first used as signs, and later became abstract symbols remains lost in prehistory

(Deacon 1997), but it is widely believed today that the sound of a word has little bearing on its meaning. Following Saussure, it seems likely that the relationship between phonetic sounds and meaning are essentially conventional – frozen accidents, the source of which is rarely traceable and never worth tracing. Each culture and phonetic system has its own peculiarities and some inherent onomatopoeia where sound-symbolism becomes important, but there is little, or perhaps nothing that is truly universal in terms of the relationship between the phonetic structure of words and their meanings.

Although each human being has his or her own slightly idiosyncratic usage and connotations for words as a function of personal experience, a word lexicon is learned in a social setting so that any given word has a more-or-less agreed upon conventional meaning. Grammar has a similar conventional basis. The transformational grammar advocated by Chomsky (1965), or some variation on it, may unite the structural principles of all human languages at an abstract level, but, however that debate may be resolved, there is general agreement that a conventional lexicon plus a conventional grammar allow for meaningful communication between human beings. Brain implementations remain uncertain, but there is in principle nothing mysterious about how lexicons, word associations and simple grammars can be used together to produce language or how those processes might be instantiated using neurons. Indeed, a simplified vocabulary and grammar have been taught to monkeys and parrots – who are then able to communicate with people at a rudimentary, but nonetheless symbolic level. And of course machine translation systems have demonstrated that language-processing can be achieved in a fully mechanical way by feeding words, word associations, basic semantics and grammatical rules into a computer. The electronic system can then deduce the meaning of the written or spoken sentence and express it again in a different language. While there are limitations to such artificial systems, they have already achieved a level of practicality that was unimaginable a few decades ago. The speed of learning and the flexibility of real brains in using language remain beyond the reach of artificial systems, but the "mystery" of language has largely disappeared. Teaching tricks to animals and creating computers with crude language recognition and production skills does not mean that our understanding of language is complete, but monkey tricks and talking robots do indeed show that modern science has achieved a level of basic understanding of what language is and how it is used.[1]

The normal mode of human communication – and the one that is learned first in life – is of course spoken language. Communication through speech is, however, more than the transmission of strings of symbols; it is heavily influenced by the actual acoustical delivery, that is, the tone of voice used while

speaking. Stated conversely, the information conveyed by the literal meaning of words strung together in syntactically normal sequences is only part of the message. Sometimes, literal understanding is sufficient and transfer of the literal symbolic meaning is the entire message, but other times we are also aware of or perhaps solely interested in conveying/interpreting an emotional message: "This is my feeling about X" or "This is my feeling about you" or "This is how I feel". And in these cases, the lexical meanings, the sequence of words and the syntactic regularity of the sentences may be less important than the tone of voice.

If, in line with the central dogma outlined earlier, there is hemispheric cooperation during language that involves conventional syntax/semantics in the LH and the sound-symbolism of intonation in the RH, what form does the cooperation take at the neuronal level? Three issues must be examined in some detail to answer that question: the known cortical mechanisms of information storage, the callosal mechanisms of information transfer between the hemispheres, and the nature of the information involved.

A. Cortical maps

It is known that the patterns of activation on the sensory organs that detect external stimuli maintain their topographical structure as they are brought to the central nervous system – and are there mapped and re-mapped numerous times (Stein & Meredith 1993). The most studied and best understood of these maps are the retinotopic maps of the visual system. The retina itself is a small two-dimensional surface. When light strikes it and excites retinal neurons, the pattern of excitation is carried in its entirety to the primary visual cortex (see Figure 5-1). From there it is copied at least 30 times per hemisphere, with each retinotopic cortical map presumably carrying out slightly different information processing on the raw retinal image (Posner & Raichle 1994).

In comparison with vision, the sense of touch has a much larger two-dimensional sensory organ – the entire skin surface, but it too is known to have a two dimensional representation on the surface of the somatosensory cortex just posterior of the central sulcus (Figure 5-2). Somatosensory information is mapped separately several times for the perception of touch, warmth and pressure, and the existence of secondary and tertiary maps in parietal cortex is known. The somatosensory maps cover the bulk of the cortical surface of the parietal lobe (in both hemispheres), just as the visual maps cover most of the bilateral occipital cortex.

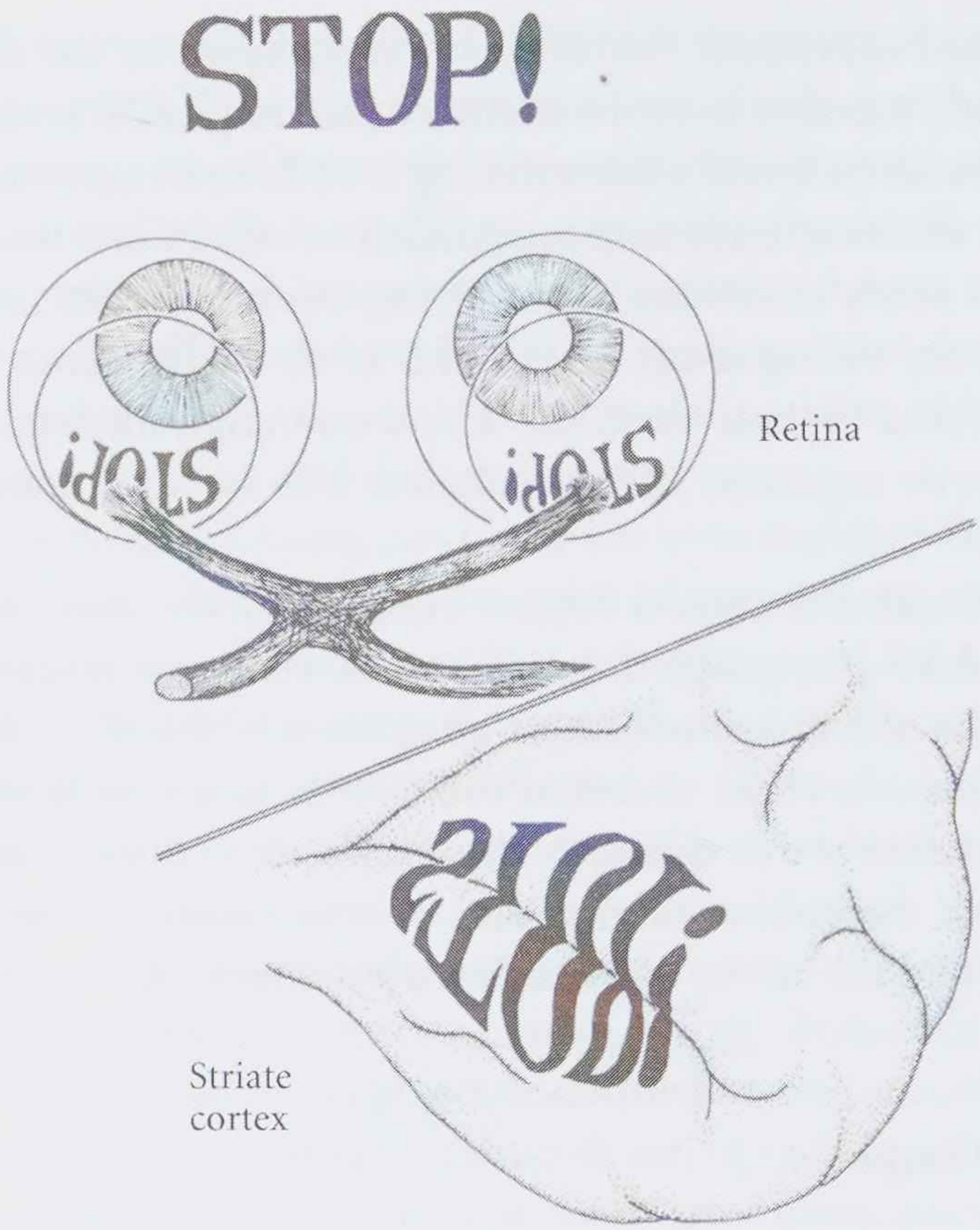


Figure 5-1. Retinotopic mapping. While there occurs some deformation of the visual image as it is transported from the retina to the neocortex, the general topography remains intact (from Cook 1986).

Of importance to the discussion that follows, the somatosensory map is known to have a motor analog. That is, immediately anterior of the central sulcus, there is a somatomotor map that is a virtual mirror-image of the somatosensory map (Figure 5-2). The somatomotor map is the last cortical way-station in the processing of motor commands that are sent from the cortex down the pyramidal tract to the somatic musculature for the production of behavior. The somatosensory map and the somatomotor map are directly connected by a set of parallel arcuate fibers.

Auditory stimuli have cortical representations that are curiously analogous to those of both touch and vision. Sounds create oscillations of the air, causing changes in the air pressure within the ear. Depending on the frequency of those oscillations, different regions of the cochlear membrane will then be stimulated. The cochlear membrane itself has an intrinsic organization such that high-frequency sounds excite one end of the membrane, and low-frequency

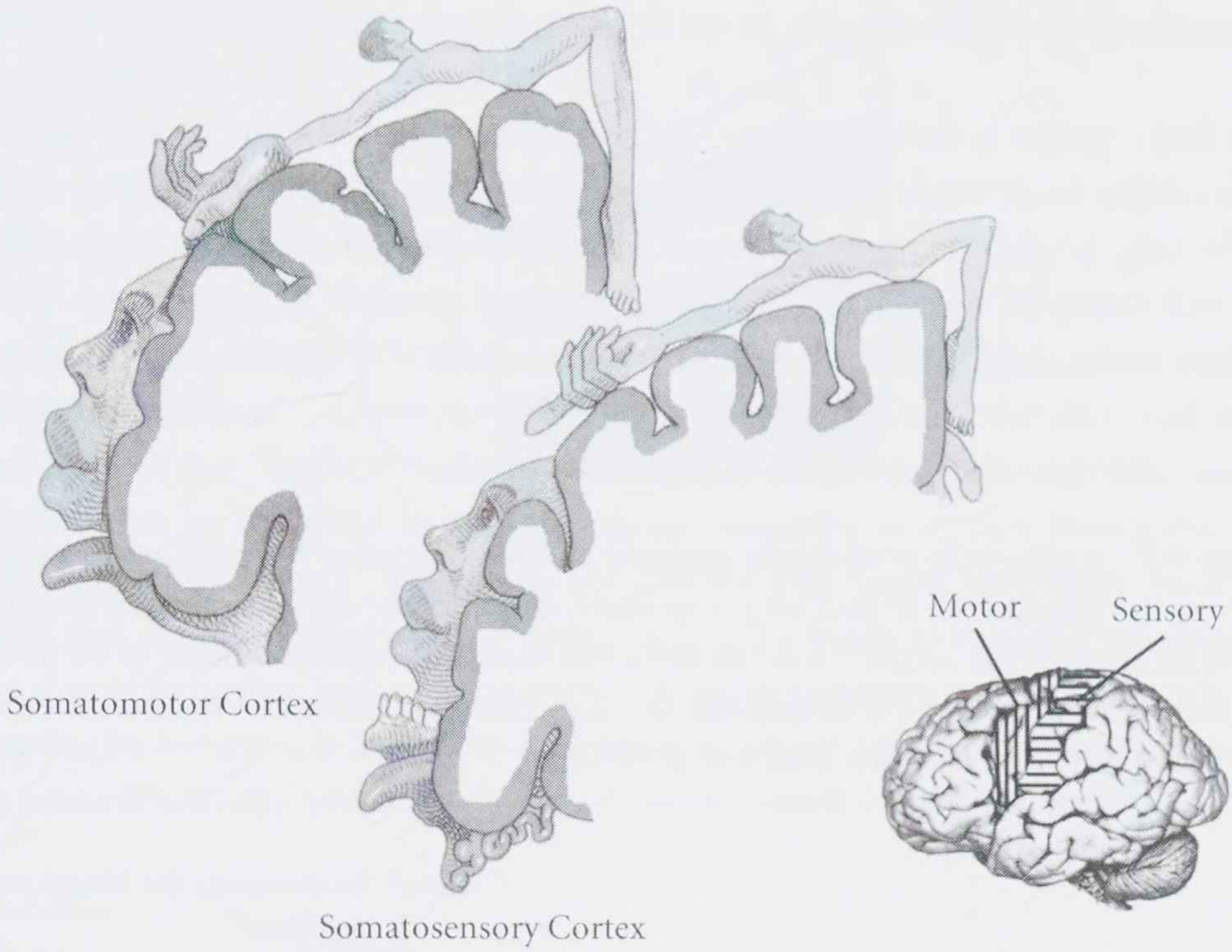


Figure 5-2. Somatomotor and somatosensory maps of the left hemisphere (after Posner & Raichle 1994: 14). The motor and sensory maps are separated from one another by the central sulcus, but have similar topography.

sounds the other end, with an orderly mapping of frequency from one end to the other. That organization means that a sound containing a particular set of auditory frequencies will produce a related, two-dimensional pattern of cochlear membrane excitation. It is that 2-D pattern that is transmitted from the ear to the cerebral cortex (Figure 5-3).

On the neocortical surface of the temporal lobe, the tonotopic representation of auditory stimuli is mapped and re-mapped many times – each time maintaining one dimension that corresponds to the frequency of the sounds, as originally encoded at the cochlear membrane. Realistic neural network models of tonotopic cortical mapping have been developed to process auditory signals using similar topographical structures (Figure 5-4).

Unlike the skin-to-neocortex and retina-to-neocortex mapping, the necessity of two-dimensional spatial maps for the cortical representation of the temporal phenomena of sound is not intuitively obvious. Nonetheless, it is empirically known that, even for auditory processing, the cortex creates 2-D representations of sound stimuli. As in other sensory modalities, multiple maps of

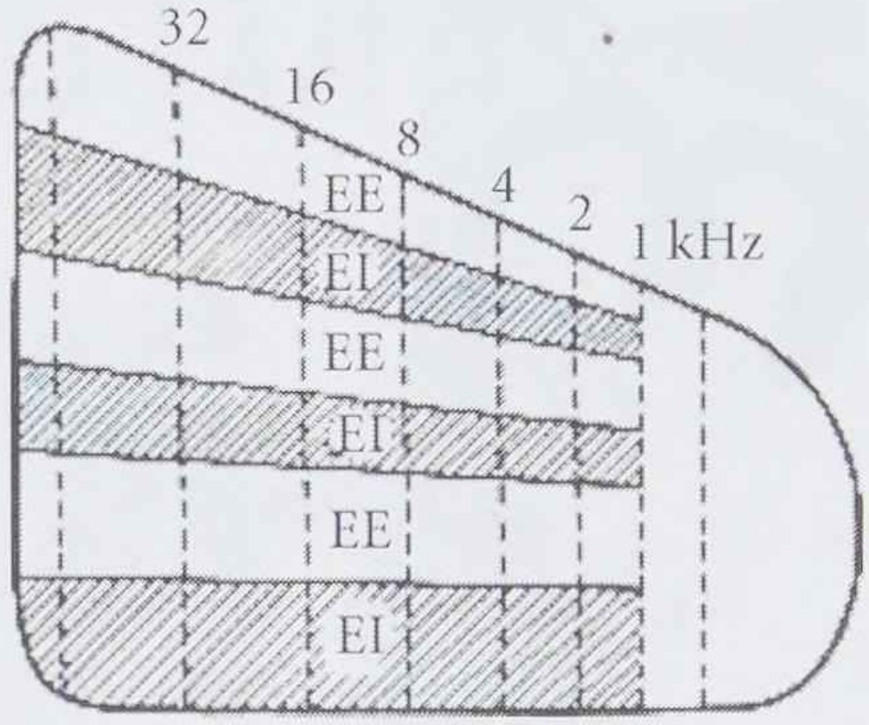


Figure 5-3. A tonotopic map on the primary auditory cortex of the cat. One dimension of the map corresponds to auditory frequency, and another dimension corresponds to lateral location of the sound (neurons activated from both ears are found in EE stripes; neurons activated by the contralateral ear and inhibited from the ipsilateral ear are found in the EI stripes) (after Creutzfeldt 1995).

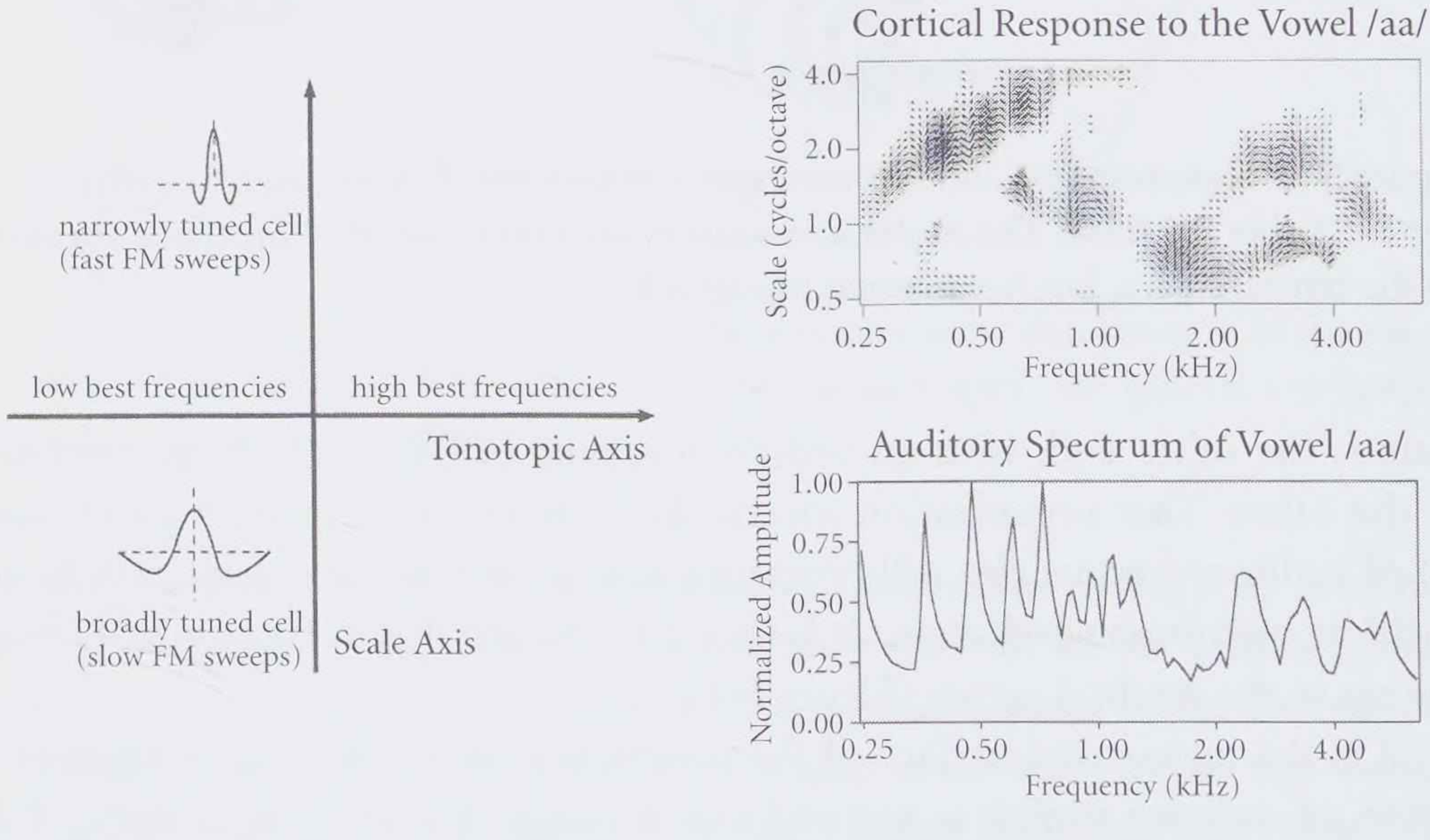


Figure 5-4. On the left are shown the two main axes of topographical mapping of auditory stimuli: auditory frequency and tuning width. On the right is shown the tonotopic activation on a simulated cortical map in response to a vowel /aa/ sound. Individual neurons on the cochlear membrane will respond most strongly to sounds of a particular "best" frequency and particular range of frequencies around that "best" frequency (modified from Shamma 1995). The lower right plot shows the frequency spectrum of the vowel sound, whereas the upper right plot shows the pattern of activity on the tonotopic map. Peaks of auditory amplitude have corresponding regions of cortical activation.

auditory information are constructed and cover a large portion of the temporal lobes bilaterally.

The reality of 2-D topographic "maps" on the cerebral cortex (and elsewhere in the brain) has been known for decades, but there is an unfortunate tendency to consider the cortical maps as nothing more than a part of the "early" processing of sensory information prior to "higher" abstract-level information processing. While non-topographical cortical organization and associations without 2-D structure are also likely to occur, the fact that more than 50% of the "highest" information-processing organ, the cerebral cortex, is known to be involved in topographical mapping is strong indication that some kinds of "higher" processes are carried out topographically. Auditory, visual and somatosensory information enters the brain as two-dimensional arrays motor behavior is generated from 2-D somatomotor maps, and much that happens in-between is also two-dimensional. Other kinds of processing occur, but there is nothing particularly "lower-level" about cortical maps.[2]

B. Homotopic callosal connectivity

The anatomy of the human corpus callosum has been intensively studied (Reeves 1985) and it is known that primary sensory and motor cortical regions characteristically lack callosal fibers, while most of the other parts of human cerebral cortex send and receive commissural fibers (Figure 5-5). The absence of callosal fibers in early sensory cortex is presumably due to the fact that there is no advantage in mixing the sensory stimuli from both the left and right sensory fields at an early stage of processing. In contrast, at higher-level cortical regions, a comparison of the sensory data obtained from the left and right can be a valuable means determining the nature and location of the stimulus. This is known to be the case both for depth perception in vision, where the slightly different images at the left and right retinas allow for an estimation of the distance of the object from the observer, relative to the distance between the eyes. Similarly, the slight differences in the intensity of auditory stimulation reaching the ears allows one to estimate the direction from which the sound was generated. Direct connections between the neocortices of the left and right hemispheres do exist, however, for virtually all cortical regions outside of primary sensory and motor cortex with the total number of callosal fibers estimated as exceeding 200 million. It is a remarkable anatomical fact about the human brain that there are sufficient numbers of callosal fibers that, on average, approximately one afferent and one efferent callosal fiber travel be-

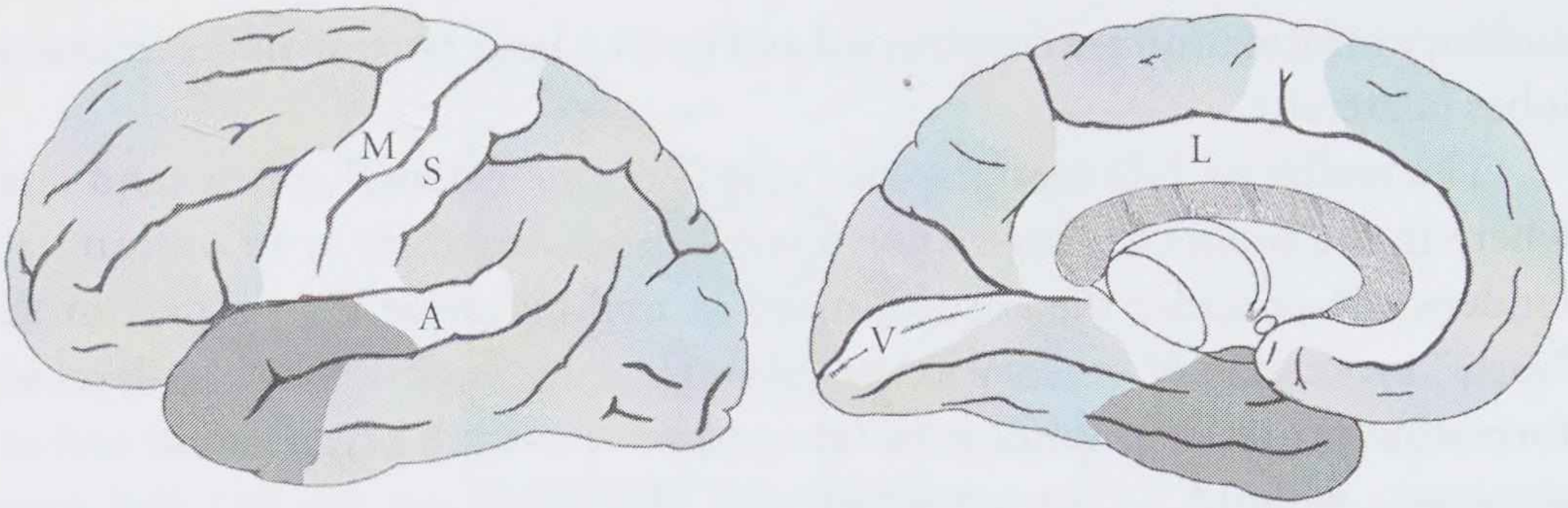


Figure 5-5. Most neocortical regions send and receive commissural fibers crossing either in the corpus callosum (light gray) or in the anterior commissure (dark gray). Cortical regions lacking commissural fibers (white) are confined to primary motor (M) and sensory (S, A, V) regions, and portions of limbic paleocortex (L).

tween every cortical column in the neocortex and its approximate analog in the contralateral hemisphere (Cook 1986). It is also known that symmetrical brain regions have a larger percentage of callosal fibers than asymmetrical regions (Rosen et al. 1989) – suggesting that it is specifically the communication between homotopic regions that is the main business of the corpus callosum. As illustrated in Figure 5-6, most callosal fibers terminate in homotopic regions of the contralateral hemisphere. The homotopicity applies not only to the general region of cortex (Figure 5-6B), but also, in the case of primates, to the layer of neocortex (Figure 5-6C). In so far as the cortical column is thought to be the functional unit of cortical information processing (Mountcastle 1997), that huge number of commissural fibers implies that whenever a cortical region is actively involved in neuronal information-processing of some kind, it is also communicating with a homotopic region of the contralateral hemisphere on a column-to-column basis.

Whatever role that callosal fibers may play in relatively early sensory processes, it is known that the bulk of callosal fibers originate and terminate in posterior *association* cortex and the broad expanses of mixed *modality* frontal cortex. In other words, while some fibers of the corpus callosum are involved in certain forms of sensory processing (stereopsis and sound localization), the majority of callosal fibers do something else. The nature of the information transmitted remains quite uncertain for most cortical regions, but the size of the corpus callosum suggests that it is not unimportant. The corpus callosum is the largest nerve tract in the human brain – containing 3-fold the number of fibers in the pyramidal tract (but in other species pyramidal fibers outnumber callosal fibers). Although the brains of all mammalian

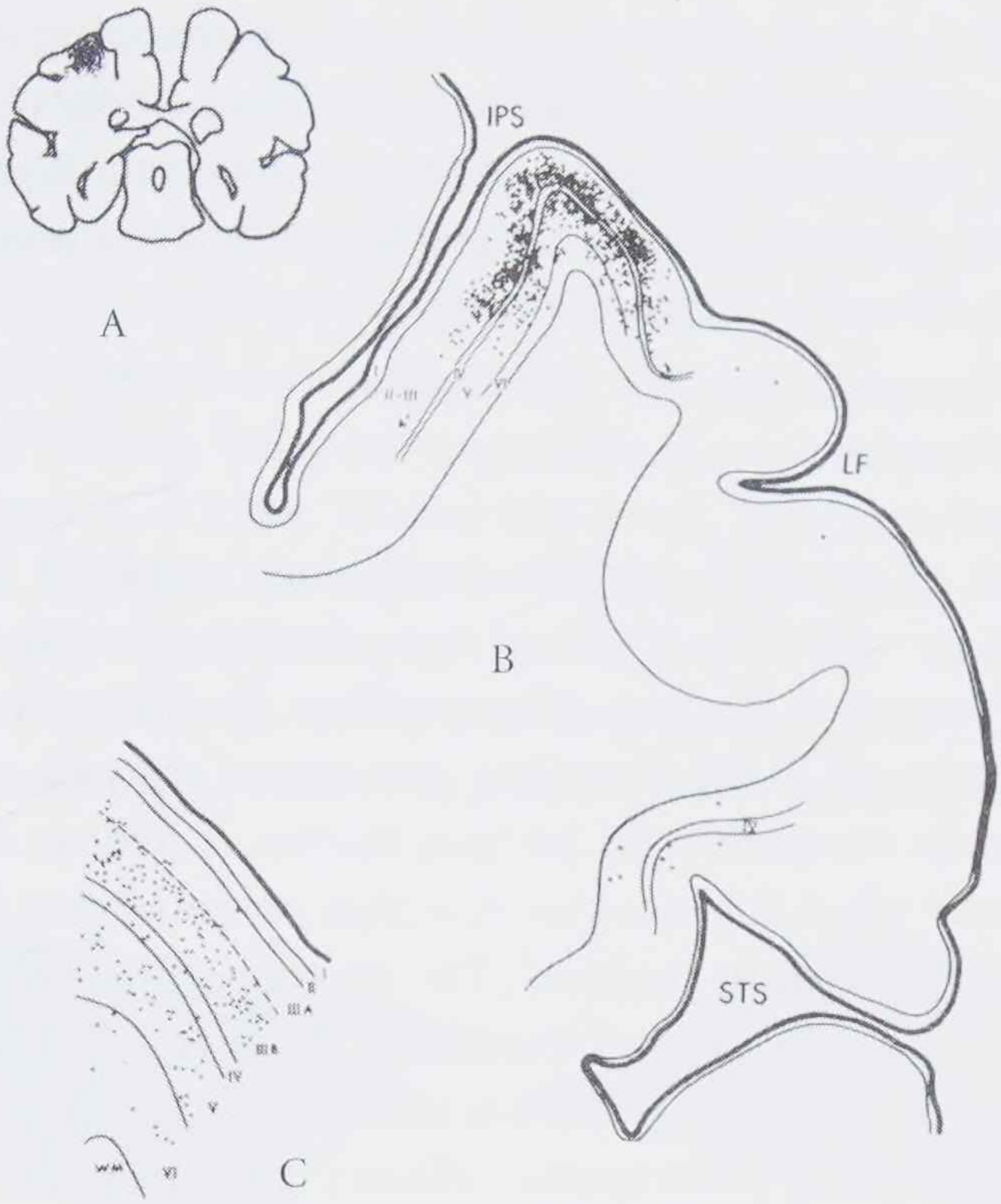


Figure 5-6. Homotopic callosal connectivity demonstrated by tracing the degeneration of axonal fibers (B) following the unilateral injection of a poison at a small intracortical region (A). Such experiments reveal not only the gross connectivity of the corpus callosum, but also the microscopic layering of fiber termination (C) (Hedreen & Yin 1981).

species are similar in design and structure, the relative size of various structures is generally a reliable guide to the importance for the given species. Moles whose visual systems are undeveloped have correspondingly small visual cortices. Bats, with highly developed auditory sensors have relatively large auditory cortices. Rabbits with whiskers providing useful touch information have large whisker-related somatosensory areas. Monkeys with acute vision and sophisticated visually-guided behaviors have large visual cortices. With regard to the relative size of the corpus callosum, it shows progressive growth up the evolutionary ladder, but even in the cat, horse or dog, its size is modest and only a small percentage of neocortex is callosally connected. In the monkey and gorilla, callosal fibers are seen to connect the bulk of non-primary sensory cortex and in man there are only small patches of cortex that are *not* connected. Most callosal fibers and the densest regions of callosal termination are seen in association cortex.[3]

Given what is known about (i) topographical mapping of both sensory and motor information at the cortical level, (ii) the anatomy of the corpus callosum, and (iii) the neuropsychology of the cerebral hemispheres in language functions, neural network simulations can be used to examine what kinds of communication can occur across the corpus callosum.

Simulations

From the point of view of cognitive psychology, neurons are rather simple things. To be sure, the topic of consciousness makes the argument more involved (see Part II), but strictly within the framework of neuronal "information-processing" the functions of neurons are well-understood and easily simulated in computers. As far as is known, a neuron has significance for cognition only when it generates an action potential that leads to the release of neurotransmitters at synapses. The strength of the synaptic effect and the frequency of firing can vary continuously from small to large values, but the synaptic polarity is normally fixed as either excitatory or inhibitory for all the axonal synapses of a given neuron (Dale's law). The complexity of neuronal "information-processing" therefore lies in (i) the spatial configuration of neurons (numbers of neurons, numbers of synapses and pattern of connectivity), (ii) the polarity of the synapses, and (iii) the temporal structure of neuronal firing.

At a relatively macroscopic level, it is known that the dimensions along which information is organized in early sensory and late motor cortex give rise to *topographical mapping* of large portions of the neocortex. Despite some indication of semantic organization outside of primary cortical regions obtained by direct stimulation at the surface of the brain (Penfield 1959), the modal simplicity of the topography of primary sensory and primary motor cortex is not found in association cortex, and individual differences may be large. Nonetheless, assuming only that association cortex – like all other known cortical areas – has some form of meaningful two-dimensional organization, it can be concluded that the interaction between homologous regions in the LH and RH will be influenced primarily by two factors: the anatomical connectivity between them, and the nature of the synaptic effects. The main types of hemispheric interaction are easily summarized and simulated in artificial neural nets (e.g., Reggia et al. 1998; Cook 1999b). If a relatively fine-grained topographic connectivity between cortical regions is assumed, then the transfer of information (either excitatory or inhibitory) from one hemisphere to the other is theoretically possible. In contrast, if callosal effects are relatively dif-

fuse, the possibilities for the transmission of detailed topographical information decrease, but the corpus callosum might then act to alter the hemispheric balance of arousal and attention. Idealized patterns of cortical activation in callosally-connected bilateral portions of cortex can be shown as in Figures 5-7 and 5-8.

Reggia and colleagues have undertaken the most thorough simulations of bilateral cortical activity, using self-organizing neural nets and homotopic callosal fibers (Reggia et al. 1998; Shkuro et al. 2000; Levitan & Reggia 2000; Reggia et al. 2001a, b). Unlike simulations designed specifically to examine the effects of inhibition (Cook 1984a, b; Cook & Beech 1991; Cook 1999b), those of Reggia et al. have had the strength of demonstrating the effects of systematic variation in network parameters on the symmetry/asymmetry of cortical

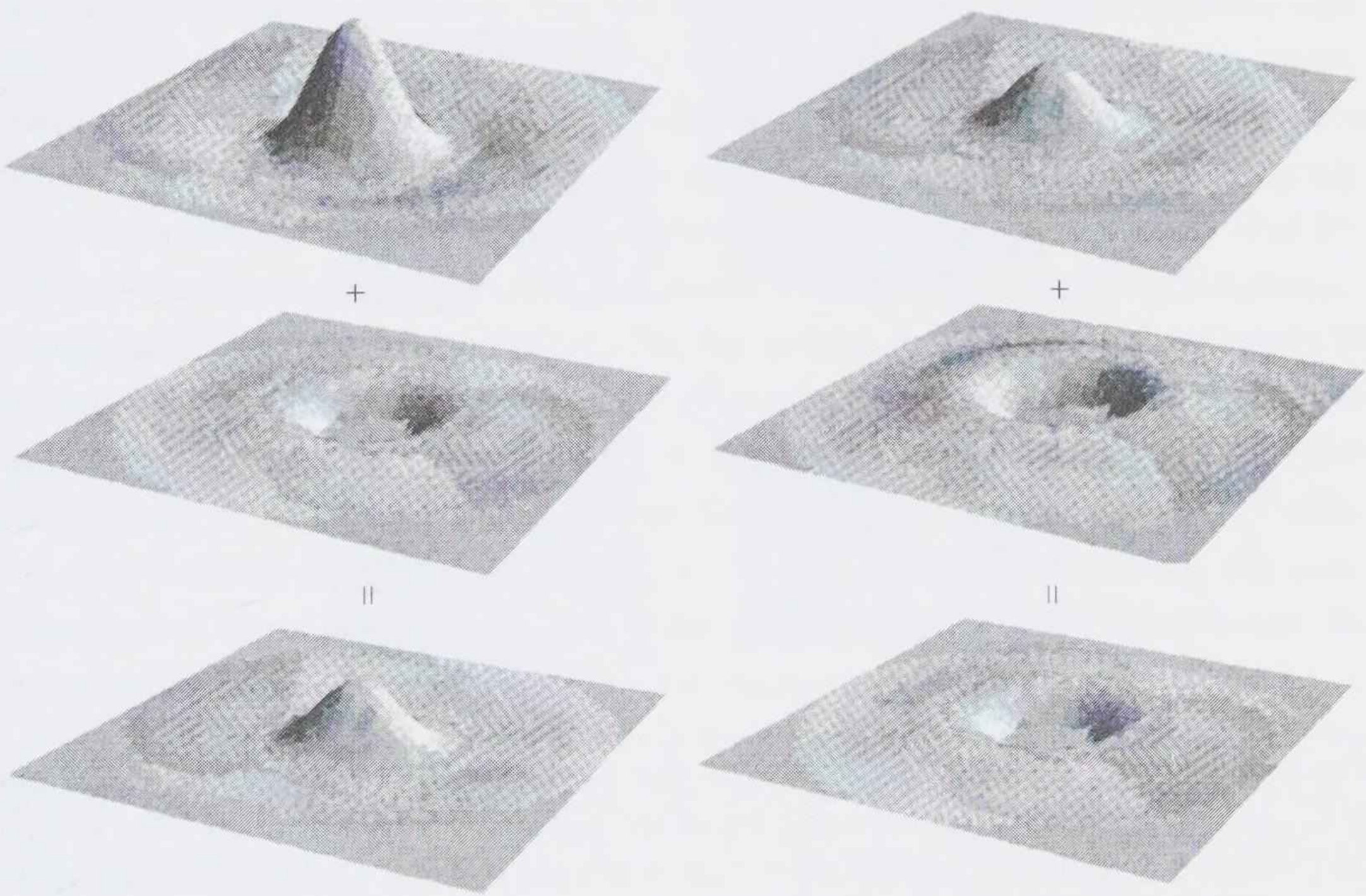


Figure 5-7. Bilateral "Mexican hat" activation of cortex with inhibitory callosal connections. The top row shows activation at identical sites in the left and right hemispheres, with stronger activation in the left. The middle row shows the strength of mutual callosal inhibition. The strength of inhibition is proportional to the original activation, so that the right hemisphere is more strongly inhibited from the left, than vice versa. The bottom row shows the final level of activation in the two hemispheres – essentially the same focus in the left hemisphere, but a reversal in the right hemisphere. The net effect is complementary patterns of cortical activity – a focus of activity with the surround inhibited on the left and the converse contralaterally.

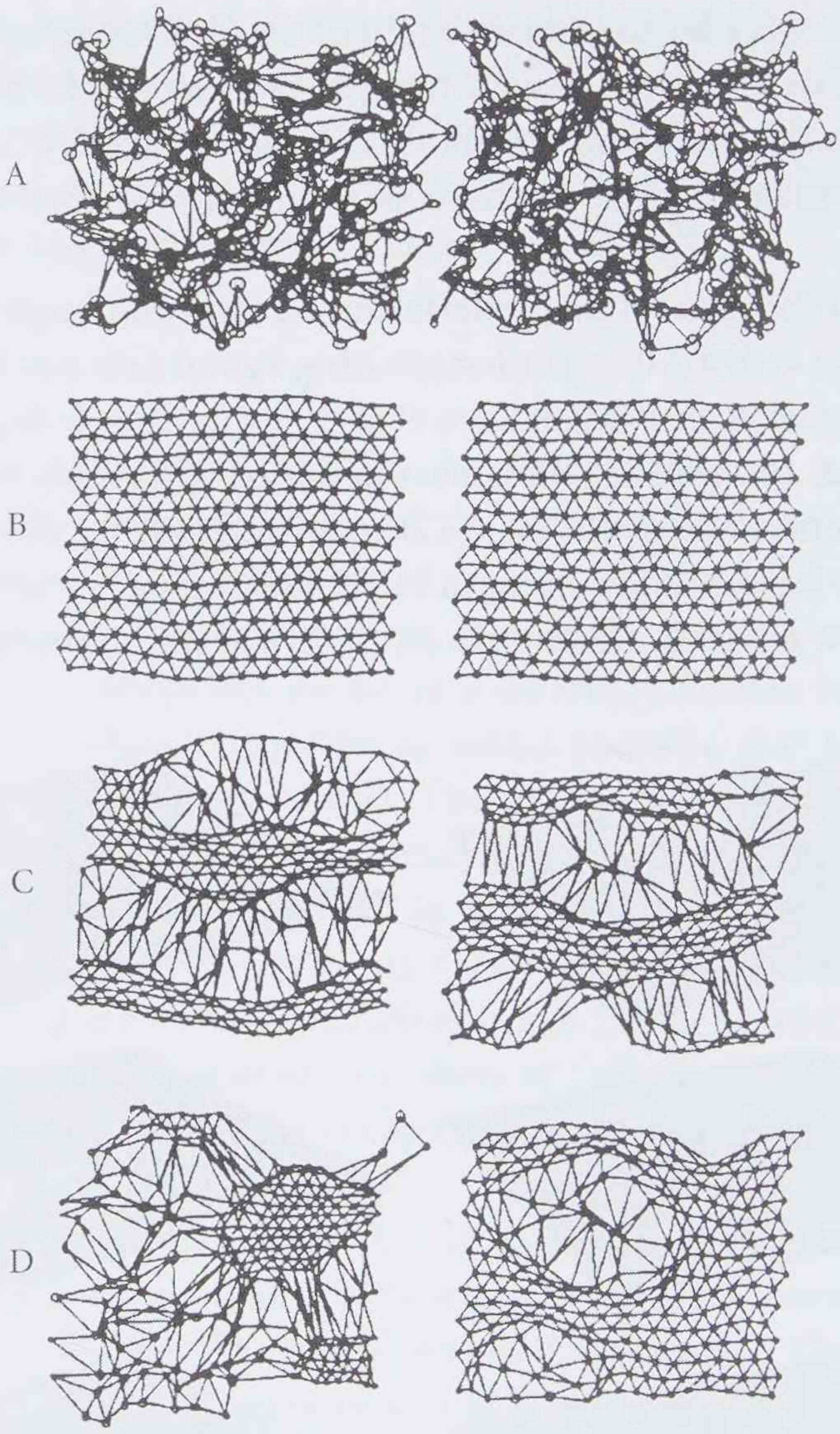


Figure 5-8. Patterns of bilateral activation due to various callosal parameters (Levitan & Reggia 2000). The neural networks are of the self-organizing Kohonen type, so that well-ordered nets show a regular mesh-like structure, whereas unorganized nets show random structure. (A) shows unorganized nets; (B) shows highly organized symmetrical nets; (C) and (D) show nets with bilaterally complementary organization. Wherever the left hemisphere shows clear organization, the right hemisphere does not, and vice versa.

activation. The upshot of their work is that there are several ways of producing hemispheric complementarity – i.e., asymmetrical information-processing, depending primarily on the balance of excitatory and inhibitory effects, mediated either by callosal or brainstem fibers. Definitive demonstration of the

effects of specifically the human corpus callosum remains unachieved, but, for the present purposes, it is sufficient to note that the constraints normally imposed on neural network simulations in light of neuron physiology produce a range of effects that include: (i) complete hemispheric independence of information processing, (ii) symmetrical or identical processing in both hemispheres, and (iii) asymmetrical or complementary hemispheric activity. These are the three main possibilities that must be examined in relation to the known functional asymmetries of the human brain.

Despite the fact that almost every animal nervous system has approximately bilateral symmetry and the problem of bilateral coordination is perhaps the first behavioral issue that must be solved by every mobile organism, there is a surprising lack of simulations specifically on the bilaterality issue. (The exceptions are the bilateral aspects of stereoscopic vision, sound localization and bipedal locomotion, but the bilaterality of the cerebral cortex is not generally acknowledged by neural network theorists!) Particularly for mammalian brains, where the number of direct cortico-cortical fibers connecting the two hemispheres suddenly increases and, in the primate, where commissural connections become a major factor in brain anatomy, it is likely that the cortical mechanism of dual control is more than merely a small variation on a "more basic" vertical organization. On the contrary, resolution of the issues of horizontal organization – at a variety of motor, sensory and, most outstandingly, cognitive levels – has clearly been central to evolutionary success. For researchers in artificial intelligence to design "physiologically realistic" neural nets that are not bilateral is thus patently absurd and, similarly, the construction of humanoid robots with only one unpaired, highest-level center for information-processing is arguably an exercise in brain damage.

Attentional models of hemispheric functions have been advanced by Kinsbourne (1970, 1982), Heilman and Van Dan Abell (1979) and Guiard (1980), and more recently by Banich (1998) and Liederman (1998). These attempts to explain laterality effects on the basis of asymmetrical "arousal" or "attention" are psychologically plausible, but suffer from terminological problems that either obscure genuine insights or mask conceptual confusions. Commonly-used phrases such as "the dynamic allocation of attentional resources" have no obvious physiological meaning and thus lack the kind of specificity needed to bridge the gap between (quantifiable) psychological phenomena and (quantifiable) neuronal mechanisms. Interestingly, recent advances in explaining arousal, attention and awareness on the basis of the synchronization of neuronal firing mean that the psychological concepts of the attention theories, such as "resources," "spotlights" and "bottlenecks," might eventually be translatable

into the language of neuronal activity (Singer 1993; Crick 1994). The importance of the synchronization hypothesis for the issue of human laterality is that it has the potential to replace a host of plausible, but inherently fuzzy descriptions of hemispheric relations with an explicit neuronal mechanism that allows distant (including bihemispheric) cortical modules to collaborate without requiring a cortical region at which all cognitive "results" are accumulated.

C. Intonatopic maps

Every new concept should lead to at least one decent neologism! Here, the word "intonatopic" is introduced to mean the topological organization of auditory information related to the production and perception of intonation in speech. In so far as auditory information-processing at the cortical level is known to have tonotopic organization, it is likely that the processing of specifically voice intonation involves the analysis of patterns of excitation in similar maps. For that process alone, the term "tonotopic" would suffice, but the processing of intonation involves auditory frequencies over the relative narrow range of voice pitch F_0 (60–400 Hz), has both sensory and motor aspects and, in line with the musical modality arguments in Chapters 3 and 4, demands a rather narrow focus on three-tone combinations. For these reasons, two forms of affect-related tonotopic maps can be expected to exist – i.e., intonatopic sensory mapping in the right hemisphere and intonatopic motor mapping for voice production in the left hemisphere.

If the laterality of voice intonation briefly discussed in Chapter 2 and the pitch phenomena of music and speech discussed in Chapters 3 and 4 are generally valid, then a theory of hemispheric communication across the corpus callosum between intonatopic maps is implied – and has clear implications concerning the brain code. Specifically:

1. Musical mode is due to relative interval sizes in three-tone chords.
2. The affective meaning of voice intonation is also a function of multiple pitch intervals.
3. The "positive/negative" valence of voice intonation is due to the patterns of tone combinations on intonatopic maps.
4. The pattern of activation on intonatopic maps in the right hemisphere can be transmitted to the left hemisphere via homotopic callosal fibers to an intonatopic map which explicitly controls the speech organs for the production of pitch. And, most importantly,

5. The motivational force of emotion is due to the need to resolve unresolved harmonic tension.

Implication (5) leads directly into the world of the relationship between language and emotion (Kövecses 2000) and will not be pursued here. Suffice it to say that when emotional states are expressed through language, the subsequent linguistic response (in a monologue or conversation) will be motivated by the tonal implications in the RH intonatopic map. In a word, harmonic tension in the voice will not be left standing, but will be followed by resolution in the form of the establishment of dominance or subordinance of one of the players in the narrative.

That voice intonation leads to some form of tonotopic mapping in the RH is almost certain, given what is known about auditory cortex and the effects of brain damage (Chapter 2), but alone the neuropsychological findings do not provide sufficient detail to elaborate a "brain code". What is more contentious is the idea that the affective valence of the pitch information in the voice obtains its meaning from chordal structure that is related to known diatonic harmonic phenomena. A more cautious statement of the nature of the affect produced by three or more tones was attempted in Chapter 3, where the fundamental dimension of tonal affect was stated simply as "stability/instability" – with two forms of stability in the directions of major and minor harmonies. Although the general case can be stated most precisely in that way, it is only when we have related the pitch combinations to positive and negative affect that we have the possibility of a cortical theory of emotions. Therefore, while acknowledging that further qualification might well be needed – especially when considering primitive cultures and young children, it is worth exploring what implications a harmonic theory of emotions has for the bilateral activation of intonatopic maps.

What specific patterns of cortical activation would be expected from a tonotopic theory of emotion? In so far as the right hemisphere is a more sensitive detector of pitch and pitch combinations, it should be dominant for the detection and representation of emotional states. Moreover, in so far as harmonic mode is determined by three tones – two of which, the high and low tones, set the tonal context, and a third mid-lying tone that specifies the tension or positive/negative affect of the harmony – three basic patterns of activation in tonotopic maps might be expected to correspond to three distinct affective states (Figure 5-9).

The tonotopic organization shown in Figure 5-9 is depicted in relation to absolute pitch values (in Hertz), but one further abstraction of the auditory

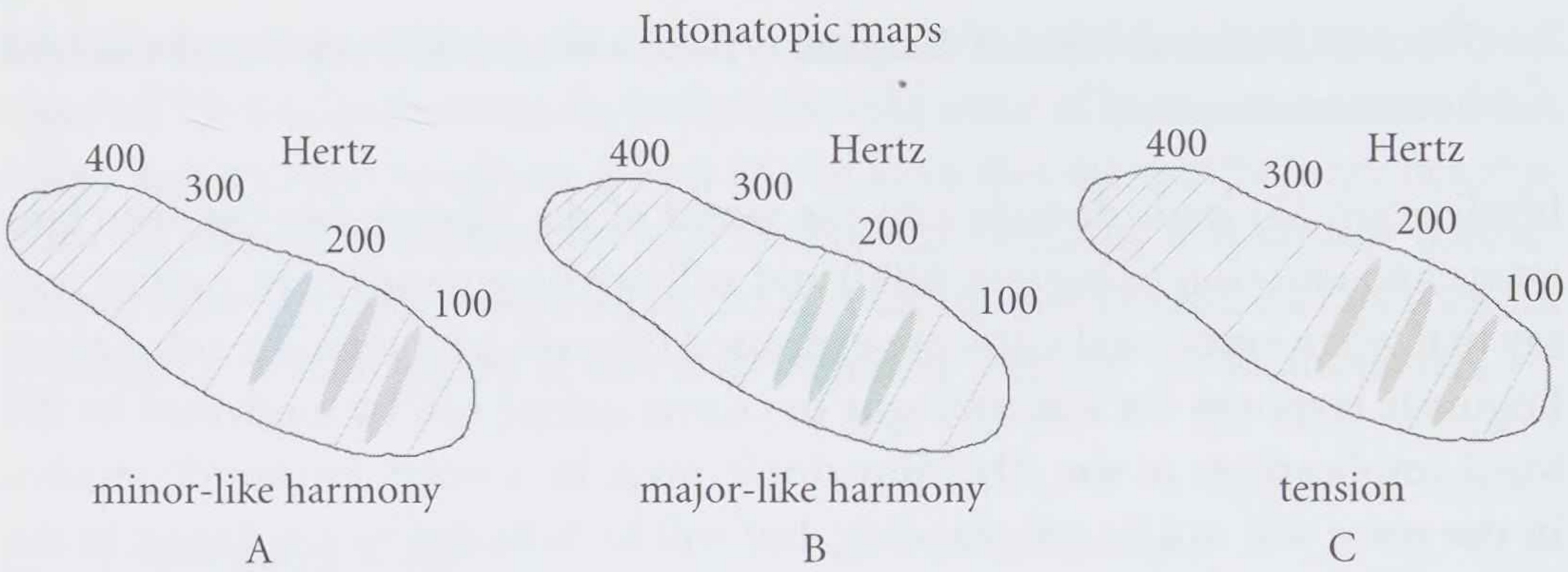


Figure 5-9. Three patterns of activation on intonatopic maps. The pattern in (A) corresponds to a smaller interval at low frequencies and a larger interval at higher frequencies (i.e., a minor-like mode). (B) corresponds to a major-like mode, and (C) corresponds to harmonic tension (equivalent intervals).

information is also worth considering. As Bharucha has shown within a musical context, the phenomenon of octave equivalence can be simulated in a self-organizing neural network (Bharucha 1987; Bharucha & Mencl 1996). In effect, due to the ubiquitous effects of the upper partials of all musical tones, the network organizes auditory stimuli within a tonotopic map that extends over just one octave – with any given tone activating a region of the map corresponding to the relevant pitch, regardless of the actual auditory frequency of the acoustical stimulus. By representing all pitches within one normalized octave, the basic modes of harmony (~emotion) require the activation of only three sites on the map (Figure 5-10). Here, the frequencies of the tones in Hertz is replaced by a relative scale in which auditory signals are summated within the

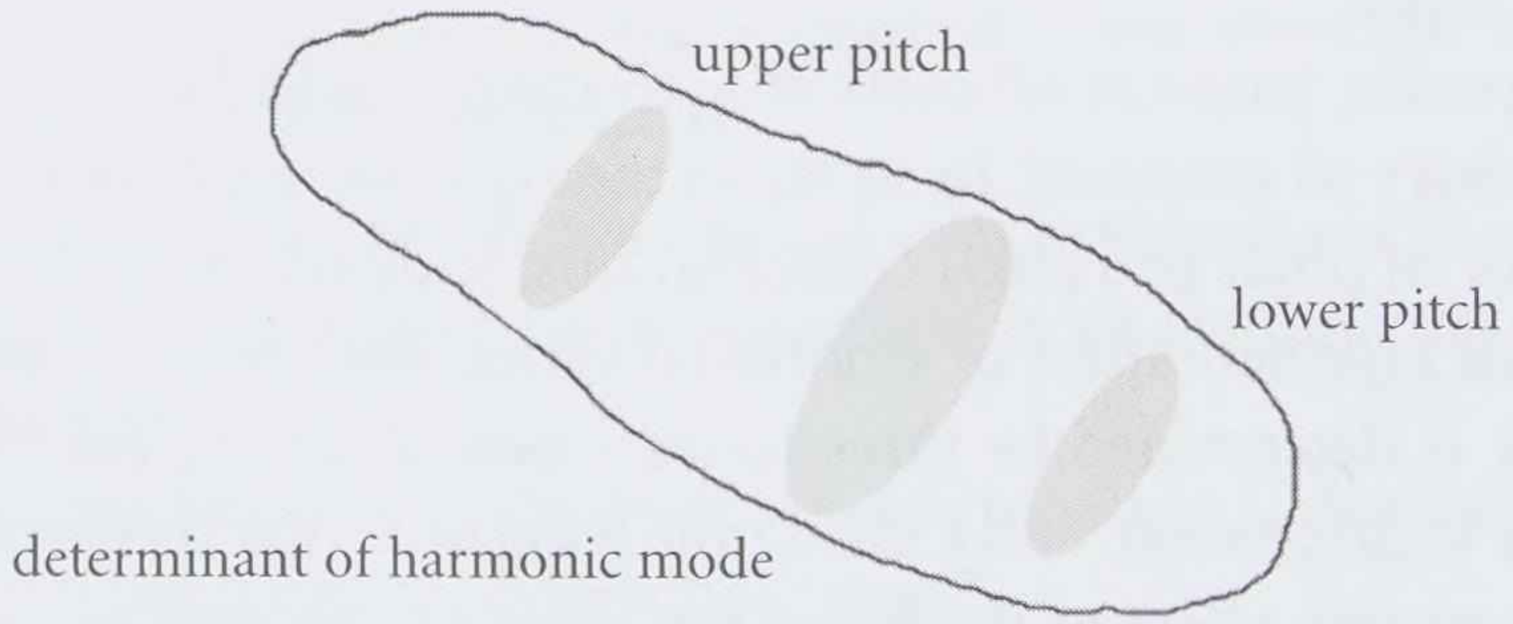


Figure 5-10. The representation of emotion on a minimalist intonatonic map. The upper and lower pitches define the pitch range – i.e., the emotional context, and the midlying pitch determines the harmonic mode simply as a function of its relative distance from the other pitches.

psychological now and represented simply as higher and lower tones that fix the auditory context, and a mid-lying tone that determines the affect of the tonal combinations.

While most theories of emotion implicitly or explicitly assume that the brain correlates of emotion are due to limbic system effects – essentially, the kind of brain activity that human beings have in common with rabbits, toads and alligators, a tonotopic theory of emotion is based on the assumption that there is a class of human emotions that are neocortical. There is of course also an autonomic *response* that is a function of the limbic-hypothalamic-somatic axis, but the pitch-based representation of emotions implies that there are characteristic emotional states that are *not* part of the neurological *response* to life events, but are "cerebral" – a form of neocortical activation that we can recall, recreate and manipulate as a part of our normal thought processes without necessarily calling on behavioral or autonomic responses. It is important to note that those thought processes are not simply the "informational" replay of recalled facts. When we ponder a long-gone life event, there occur only faint whispers of the autonomic effects that were once intensely experienced, but the thoughts nonetheless have emotional content just as clearly as they have cognitive content.

That emotions have a tonotopic representation follows from the considerations above and in previous chapters. The main theoretical conclusion is that there are three distinct affective states: two patterns of harmony with explicit, resolved affective meaning (corresponding to the "idealized" diatonic forms known as major and minor chords) and an in-between state that is affective in being an explicit state of uncertainty wavering between the two poles of positive or negative resolution (corresponding to the tension chords). Of course, the magnitude of the emotional state can be influenced by many factors, but there are inherently three and only three qualitatively distinct states along the dimension running from positive to negative emotion. The affect of a resolvedly positive or negative outcome is self-explanatory, but it is worth noting that a state of tension between the positive and negative poles is *not* a state without emotional import. On the contrary, the uncertainty with regard to the direction of resolution is itself a powerful feeling – with the magnitude of the tension being greater if the possible positive and negative outcomes are extreme.

The theoretical argument up to this point is relatively straight-forward, given what is known about musical mode, the importance of F_0 in voice intonation and the reality of the tonotopic mapping of auditory information. What is less certain, however, is the nature of the information transfer to and from the left hemisphere. Specifically, assuming that there is an intonatopic map in

the right hemisphere, and a related, homotopic intonatopic map in the left hemisphere that actually drives the production of voice intonation, what is the nature of the callosal communication between them? Three physiologically-plausible answers must be considered.

i. Excitatory transfer. If the callosal transfer of a RH pattern of activation on an intonatopic map to a contralateral LH intonatopic map is excitatory, then the major/minor/tension mood of the right hemisphere will be transferred in its entirety to the left hemisphere, and the left hemisphere will then use those activated pitch locations when it generates voice intonation. This is the simplest, most parsimonious hypothesis of hemispheric relations (Figure 5-11A), but there are both theoretical and empirical grounds for considering other mechanisms, as well.

ii. Inhibitory transfer. If the callosal effect is inhibitory, then the pattern of excitation corresponding to the harmonic "mood" of the RH will necessarily be *transformed* as it is transmitted from the right to the left. The nature of such transformations is complex in a musical sense, but can in principle be delineated (Figure 5-11B). This pattern of activation corresponds to the active suppression of "emotionality," and yet, through the process of inhibition, the pattern of activation in the LH is different from, but directly related to the chordal patterns of musical modality.

iii. Synchronization. Alternatively, interhemispheric communication could entail neither form of explicit information transfer, but rather the synchronization of neuronal firing in homologous maps in the LH and RH. In effect, the LH intonatopic motor map that directly controls voice intonation would be activated in so far as it is synchronized with the harmonic mood patterns in the RH. Only to the extent that both hemispheres are synchronized would a definite affective tone of voice be employed by the speaker. This synchronization mechanism is by its very nature the most flexible, by allowing the "linking" of any pattern of RH affective mapping with any arbitrarily-constructed intonation map. In principle, any form of cortical representation of emotion could be synchronized with any pattern of intonation. That such linking is not entirely arbitrary is indicated by the non-arbitrariness of the major/minor distinction, the universality of rising and falling pitches in questions and statements, and other indications of "sound-symbolism", but a combination of fixed structures in the RH and individual differences in synchronization patterns is a theoretical possibility.

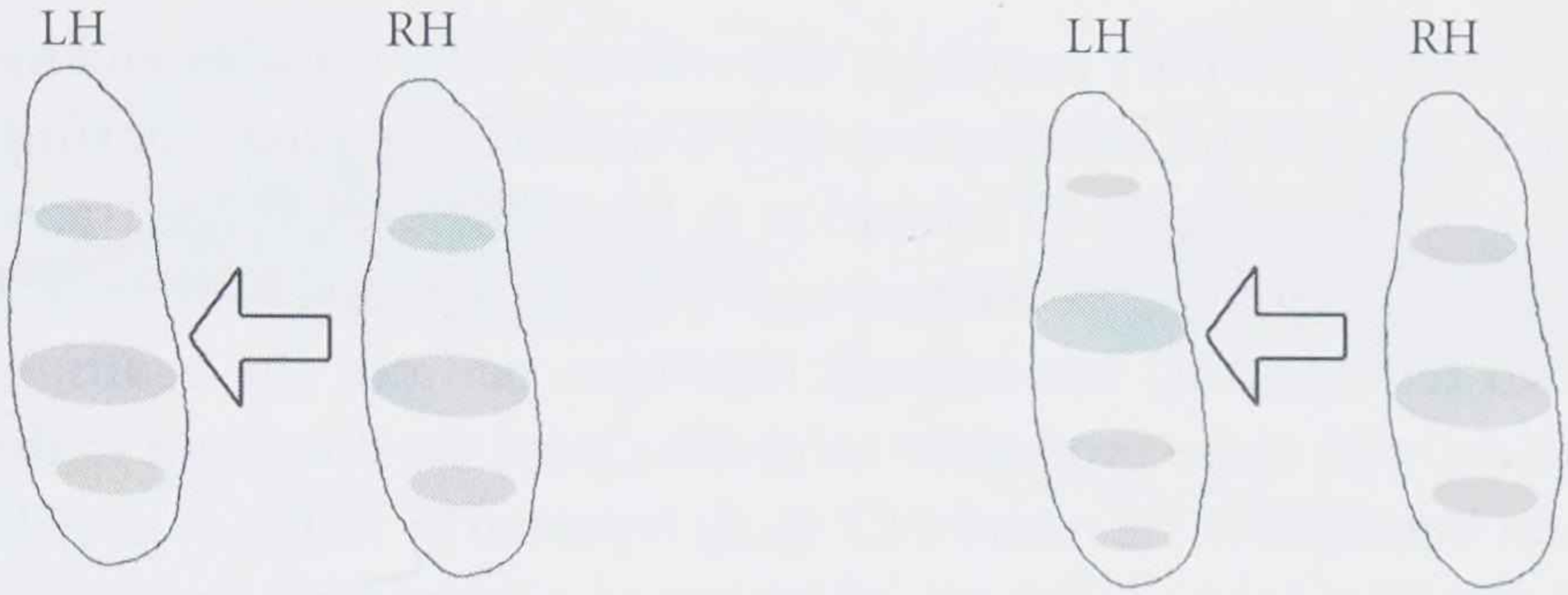


Figure 5-11. Excitatory and inhibitory of communications between homotopic intonatonic maps in the left and right hemispheres. An excitatory transfer from right to left implies a literal reproduction of the pattern of activation, whereas an inhibitory transfer implies a complementary representation.

The differences among these three basic patterns can be stated as (i) literal transfer – where the tonotopic pattern generated in the RH is used in that form for the generation of the tone of voice by the LH, (ii) complementary transfer – where the tonotopic pattern corresponding to a definite emotional valence in the RH is the causal force determining the LH voice production, but it is not a "literal" re-presentation of the emotional tone of the RH, and (iii) the quasi-arbitrary linking of two patterns – where, presumably through a learning process, the pattern of affect representation in the RH is activated together with a learned pattern in the LH, but with no inherent tonal causality linking the two. These three patterns imply three different relations between emotions and intonational phenomena, the reality of which must be weighed empirically.

Unrelated to the specific pitch content of the patterns of topographical activation, it can be said that the choice among these three possibilities is a long-standing problem with regard to all types of interhemispheric cooperation. As I have argued elsewhere (Cook 1986, 2002a, b), the excitatory mechanism is the most straight-forward and indeed most corticocortical communications are known to be excitatory (similar to the excitatory relay from retina to thalamus to V1 to V2, etc.), but excitatory interhemispheric transfer offers the fewest possibilities for generating new information, and does not contribute in any way to an explanation of hemispheric functional differences. On the contrary, mirror-image excitatory transfer would imply that the hemispheres would be more similar with increases in the number of callosal fibers. From an evolutionary perspective, it appears that the exact opposite is the actual case.

In contrast, inhibitory transfer is theoretically more complex and more interesting (Cook 1999b), but there are only a handful of studies reporting physiological evidence for callosal inhibition in human beings. Using transcranial magnetic brain stimulation and measuring motor activity, Netz et al. (1995) concluded that "mutual transcallosal inhibition between the hemispheres is asymmetric, with more prominent inhibition from the dominant to the nondominant hemisphere." A recent PET study reported by Karbe et al. (1998) indicated that "disynaptic collateral inhibition of interfering neuronal information is particularly important for orchestration of functionally asymmetrical regions in the speech-relevant cortex" (p. 1161). A more complex model that includes simultaneous subcortical interactions, in addition to callosal effects, has been advocated by Reggia et al. (2001).

The third alternative of bilateral hemispheric synchronization is known to occur in certain experimental preparations, but its plausibility as a general mechanism of interhemispheric cooperation remains uncertain. Thus far, the evaluation of all such models through a comparison with experimental data has been severely hampered by the uncertainty concerning the nature of the cortical information that is involved. To return to early sensory cortex and study the callosal mechanisms of stereoscopic vision or sound localization necessarily implies a retreat from higher cognitive functions, and is therefore inappropriate. For this reason, the study of intonation and cortical information stored in a tonotopic fashion is perhaps the most promising approach to examine high-level, language-related cooperation between the cerebral hemispheres.

Without being able to choose among the three mechanisms outlined above, we are not in a position to draw final conclusions, but a small number of theoretical possibilities have been identified. Arguably, the components of the highest-level "brain code" involved in language processes are in hand. The basic elements include: (i) a RH-based cortical representation of emotions, that has (ii) tonotopic organization and perhaps (iii) the regularities and affective qualities of chordal harmony. The known bilateral symmetry of the cerebral hemispheres and the homotopic nature of the corpus callosum provides (iv) a channel of interhemispheric communication and (v) several plausible candidates for the neuron-level mechanisms of information transfer between the cerebral hemispheres. The theoretical patterns of bilateral activation on homotopic intonatopic maps are illustrated in Figure 5-12 for the two most obvious cases of interhemispheric cooperation. In the same way that the nature of nucleotide-nucleotide bonding was the conceptual insight that led eventually to the deciphering of the genetic code, the finite set of mechanisms for

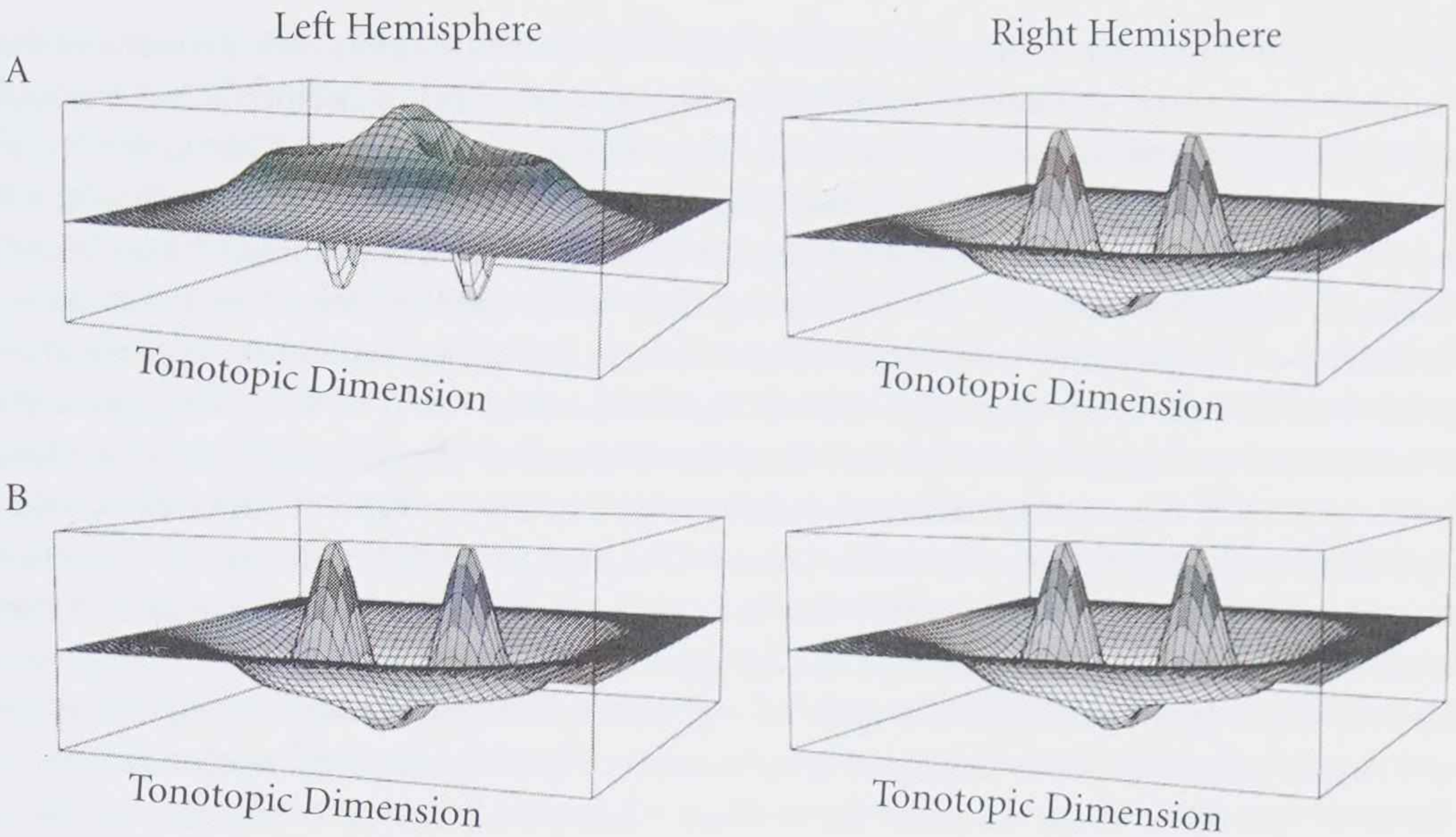


Figure 5-12. Patterns of cortical excitation implied by inhibitory (A) and excitatory (B) communication between intonatopic maps. In case (A), a dominant pitch interval represented by the two peaks is perceived in the right hemisphere; the inversion implied by callosal inhibition produces a single peak of activity in the left hemisphere, that would influence the tone of voice in speech. In case (B), excitatory transfer produces two peaks in the left hemisphere.

cortico-cortical communications outlined above (particularly, excitatory and inhibitory hemispheric "binding") constitutes the core of the brain code.

Notes

1. The principal strength of the Chomskyian view continues to be the insistence that there is something unusual about the human brain that allows us to learn language. The formulation of this idea solely in the linguist's terminology of "universal grammar" or a "language gene" is probably inappropriate, in so far as the other characteristically-human talents of tool-usage and music may be results of the same mental revolution that allows language. Rather than place all our bets on one of these three branches, it is more sensible to search for their underlying commonality – and then spell out how the one-step from "monkey-to-man" can take three distinct forms.
As Deacon (1997) and Jackendoff (2002) have noted, what lies behind human language is not simply symbolization – taking signs or icons to have symbolic meaning, but rather the conceptual revolution implied by syntax. Linkage between X and Y in a Pavlovian sense simply does not add up to language, because we need higher-order structuring to give the lexical associations the power to delineate patterns of causality through language. This issue

can be pursued in a purely linguistic context (notably, determining the core elements of the "universal grammar" – possibly Jackendoff's "hierarchical phrase structure"), but I would argue that the entire discussion should be taken to a more abstract level. That is, the "two-body" problem of one-to-one associations that linguists have shown to be insufficient for explaining language should be supplemented by consideration of appropriate "three-body" issues. In a linguistic context, the underlying three-body problem might be the Subject-Verb-Object (agent-action-recipient) structure that, in various permutations, appears to be a universal feature of all languages. That is, in-between the world of the associative links between icon and meaning (which the monkey can do) and the full grammar of modern language (which the monkey cannot do), there is necessarily a notion of "three-body linkage" that is at the heart of all syntactic structures. The linguist might want to call it "universal grammar," but the important point is that it lies at a pre-linguistic psychological level that needs to be approached in terms that can be applied to linguistic issues, but is not derived from or dependent on them! In Chapter 3 I argue that music psychology has already faced and largely solved the three-body problem in terms of traditional harmony theory. This is the tonal "primitive" that lies at the heart of the concepts of musical mode, chordal resolution and the affect of major and minor harmonies. In a musical sense, it is "simple," has a well-understood regularity as embodied in diatonic music, and from which the full complexity of Western musical phenomena is derived. But three-tone harmony is nonetheless a more complex issue than two-tone intervals, and cannot be reduced to the topic of musical intervals.

2. On the contrary, cortical maps are the brain's way of recreating certain external events on a 2-D surface with greatly simplified dynamics.

The most basic form of biological "dissonance", "conflict" and other two-body problems is somatosensory (in the widest sense including touch, smell and taste – i.e., the direct contact between the biological organism and biologically-relevant molecular structures). This is the world of direct physical assault, pain and dissolution of the intact biological system. Auditory (and visual) phenomena have such two-body meanings *only* in an analogical or metaphorical sense. That is, the "overlap," "contact" and "collision" of tones, and the concept that some tones fit well or poorly with other tones have no physical reality; they are a consequence of higher cognitive activity – that is, interpretation of auditory phenomena *as if* they had real biological significance. As a metaphorical set of signs that are translated to have significance concerning biological reality, meaning can be attributed to auditory events. But on their own, as auditory phenomena, they have almost no biological significance ("almost" because a sufficiently loud sound could be painful and have biological significance). Generally, sounds have meaning only in an indirect sense – and language further distances the sound from the meaning in so far as both semantics and syntax are conventional.

For this reason, all of the relationships among tones – consonant and dissonant intervals, harmonious chords, pleasant and unpleasant melodies – necessarily borrow from the world of real biological meaning by establishing parallels between patterns of sounds and more fundamental somatosensory events. Sounds themselves do not have biological significance! They have meaning only as metaphors. So, in order to know what sound relations are analogous to, it is essential to inquire first about the world of "important" somatosensory biology. This is the world of physical space (territory) – which, in concrete terms, comes down to two issues: sex and food. Those are the things that count in the biological world! Sights and

sounds are only hints, indications, reflections and echoes passing through the air and telling us about the biological reality of reproduction and territorial conflict.

In order to understand audition, to understand the meaning of sounds in speech and music, it is necessary to begin at a low level and translate what the primitives of sound signify within the context of biology. While our goal is to explain the meaning of music and language, those higher-level, abstract phenomena are necessarily built upon a primitive ability of the (sound-sensitive and light-sensitive) organism to translate the phenomena of (acoustical of electromagnetic) vibrations into the "real" phenomena of biological existence. It cannot be said too often! Sound has no biological meaning! Sound is noise! Sound is nothing more than changing air pressure on a tiny hypersensitive membrane in the ear. If there were no biological follow-up to the lion's roar, the snake's rattle or the lover's giggle, sound would be meaningless, leaves rustling in the breeze, inconsequential. Since, however, sounds are strongly correlated with certain kinds of biologically-significant events, we, as a biological species, have learned to take sounds seriously – to interpret them as if they were directly significant. The non-biological character of sound implies that the power of sound (language, music, etc.) does *not* lie in the power of mental associations; our attributing importance to auditory stimuli is not because we have learned (as individuals or as a species) that the lion's roar is followed by our becoming a meal or that an increasing vocal pitch is associated with "submission" weakness, deference, or interrogative sentences. On the contrary, those are the kinds of generalizations we may ultimately deduce as pondering academics, but not as organisms engaged in the struggle of life. If we relied on mental associations between sound and meaning, long before we have drawn conclusions and initiated behavior, biological reality would have already sealed our fate. What we know – in our bones, instinctively, without reflection, without linguistic or any other kind of cognitive operation is that physical, somatosensory punishment – pain (usually PAIN!!!) – is not good. Pain is intolerable and we instinctively strike out and change the situation from one of pain to one of painless biological integrity. In response to somatosensory pain, we are instinctual animals, and there is simply nothing comparable in the world of audition (and vision), where all knowledge is obtained through the hard-won association of inconsequential sights and sounds with consequential biology.

Territoriality has two forms – protecting personal turf in order to obtain necessary nutrition and in order to reproduce sexually. Our instinctual responses to territorial threats of either kind cannot be rationally explained. In a post hoc fashion, one may attempt to justify actions on the basis of nutritional need or sexual desire, but whether such rationalizations are coherent or otherwise, the animal form, mammalian form, primate form and even the Homo sapiens form is such that we have instincts – and, as a consequence, gravitate toward sources of nutrition and we are attentive to the biology of the opposite sex. As biological entities, we are instinctively attracted to the opposite sex, and instinctively reacting to impingement on our personal territory. And in this regard we are only modestly different from the dog, gorilla, pig or toad! At our core, we are animals. And yet, we are more than somatosensory animals because we have highly-developed cognitive mechanisms for analogizing about the biological world. Unlike the toad and others, human beings have the capability for metaphorical thinking. Eons before the emergence of language or music, animals evolved means to capture sounds – and to associate auditory ephemerals with the biological events that truly mattered. The association between certain sound patterns and the kinds of

biological events that are related to food and sex has aided survival. But in addition to that set of associations – in principle, identical among åll animal species – the human brain has developed an analogical system – not direct signs linked to their signifiers, but symbols. Direct associations between certain sound patterns and certain biological meanings undoubtedly grounds the entire system, but the fact that auditory stimuli are themselves not directly biologically relevant means that whatever pattern of auditory stimuli that we learn is linked to biologically significant events system, it is itself one step removed from the instinctual stimulus-response reflexes that guide our biological beings. (A trivial example concerns the relative urgency of responding to an auditory threat – "I will kill you" – versus a pin-prick to the fingertip! We are capable of completely ignoring the words, but are unable to suppress a response to the trivial sting.) Sounds may signify something biological, but they themselves are not biological. In the extreme form of human language, they become fully Saussurian: arbitrary sound combinations that, through the complex process of language learning, are indicative of real world events, but are only weakly tied to that reality.

The significance of the "unreality" of sounds is that a theory of mind cannot be constructed solely from the world of direct associations. More important than the Pavlovian associations are the regularities in the auditory patterns themselves, for they constitute a world of sensory events that are a universe running in parallel with and linked to the biological reality. That, I maintain, is the source of the ability to understand metaphors and analogies. With our enlarged brains and our lethargical jungle living (no fear of starvation today!), we have evolved into sex and territory machines. We are not plants – fixated in space and "sexual" only in the sense of passive germination and producing fruit. On the contrary, we are territorial and aggressively sexy.

Once an animal organism can detect and produce pitches, then the realm of auditory analogy is possible. The auditory analogy begins with the two-tone problem. How can one tone overlap, approach or intersect with another? What possible relationship can one tone have with another? Of course, in a biological sense, nothing can happen between tones. Tones are noise – whether single, multiple or overlapping, they are immaterial. So what we, as a species, have done is to analogize tones. Two tones in auditory space are taken as representative of the two-body problem. And the only two-body problems we understand as *biological* organisms are sex and territoriality. The auditory metaphor is therefore necessarily the problem of representing the basic dynamics of the somatosensory world of sex and territory into an analogical auditory world: we take what we know – biological existence – and we apply those principles of interaction between bodies to *imagine* something similar in tones. It is a fiction – a total fiction, an utter fiction! But it is specifically the fictional, abstracted character of the analogy that makes it of interest.

Two tones. What can be done with two tones? We can take those two tones and treat them as if they were two material entities: we can think metaphorically. If tonal space were as complex as biological reality, then thinking in language or musical terms would be as complex and demanding as experiencing – actually living out – the thought. There would be no advantage in thinking and planning through language. Fortunately, the essential simplicity of the world of tones is due to the fact that auditory frequency is one dimensional: a tone is either higher or lower than a second tone, so the relations among pairs of tones are, if not simple, clearly finite. The "real" world is reduced from 3-D to 1-D, and in this finite world

of tonality, it is possible to create and think and talk about simplified representations of our biological reality.

Before examining the complex issue of language, let us consider the other wonder of mentation, music. We start with two tones, two "bodies". On a stimulus-sensitive membrane – and more specifically a membrane that has an inherent topological topography – two tones that are of different frequencies will produce oscillations in the membrane at distinct locations, and changes in the auditory frequencies of either or both tones will mean movement of those foci of excitation toward or away from each other. From there it is merely a matter of sending that sensory data to the brain where various associations can be made on related topographical maps. In the abstracted world of association cortex, the two-tones become two-entities – self and other, sister and brother, X and Y – and all of the dynamics of two physical bodies can be analogized in that auditory space. As a consequence of tonotopic encoding, any animal species with an auditory system will necessarily have an adequate understanding of the two-body problem. Aggression for territory is primarily one-on-one: self against other. Lust for sexual satisfaction is also a two-body interaction. But the three-body problem is suddenly complex – self and lover and other, self and enemy and brother, X, Y and Z. The dynamics of three-bodies can again be analogized in auditory space – the movement of three tones on a tonotopic map. Territorial issues among three entities are more complex than those among two entities, because of the possibilities for pairings among the three entities. The two-gender aspect of sex implies that the two-body problem in sex is more complicated than that in territoriality. Lust itself remains a two-body problem, but sexuality is inherently a three-body problem: two-body attractions within the context of unchosen attractions. The presence/absence of lust – and all that flow from that – is not problematical: it is animal and fully known. But two-body attraction normally occurs within a social setting, and the attraction to her is necessarily experienced together with less attraction to the other her. The complexities of territorial and sexual strategies are too complex to pursue here, but the important point is that our cognition about those realms occurs in the analogized world of tonal space.

3. Human brain-imaging techniques may answer some basic questions here in the near future, but it is relevant to note the caution that must be taken when extrapolating from the anatomical and physiological results on the corpus callosum of other species (Bianki 1988). Not only are the number of callosal fibers markedly different among mammalian species, but the layer of termination also differs (Jacobson & Marcus 1970). Callosal fibers that terminate in the superficial layers of the neocortex are more likely to have a diffuse excitatory, arousing effect, but callosal fibers in the primate terminate contralaterally in the same layer from which they were sent (Figure 5-6). The anatomy of the human corpus callosum is undoubtedly similar to that of other primates, but it should not be forgotten that it is specifically the functional left-right asymmetry of the human brain that differs so clearly from ape brains. Clearly, uniquely-human characteristics of the corpus callosum could play a role here.

Part II

Consciousness and cognition

In the first two chapters, the issues underlying a basic understanding of *human* neuropsychology were outlined. That first-order description of the brain pathways utilized in characteristically-human behavior might be labeled as "central dogma," and involves the asymmetrical use of the cerebral hemispheres. Similar to its dogmatic cousin in molecular biology, the central dogma in psychology is not by any means a complete theory of the brain, but it *is* a succinct summary of well-established facts concerning the lateral asymmetry of brain activity related to quintessentially human behavior, specifically, tool-usage and language. The basic argument is that motor execution of characteristically human behavior is performed predominantly by the left hemisphere, while the right hemisphere contributes affective and contextual information to the left hemisphere via the corpus callosum. Many qualifications of this dogma (for left-handers, for the effects of brain-damage, for individual differences, and so on) are necessary, but normally one hemisphere is dominant – primarily for reasons relating to the achievement of unified motor control, while the other hemisphere is concerned with other processes related to, above all else, affect. Precisely because it is *not* involved in the sequentialization of motor behavior – i.e., the problems of the temporal ordering of muscular contraction, the right hemisphere seems to have a better appreciation of what the overall desired final-state is – no matter what sequence of behavior leads to the ultimate goal and without concern for the practical problems of motor execution. It is *not* preoccupied with the topic of *how* to achieve a goal-state, but it *is* concerned with what that state should be. While the left hemisphere is "dominant" in the sense that it actually carries out the command sequence for control of the musculature, the right hemisphere plays an important role in "masterminding" behavior by focusing on emotion.

Given the known specialization of the right hemisphere for intonation and emotion, the tonotopic organization of pitch on the cerebral cortex implies that the pattern of cortical activity corresponding to a specific pattern of pitches can be transmitted to the left hemisphere and utilized in the motor control required for vocal intonation. In Chapter 3, it was shown that the distinction between major and minor musical chords has a simple regularity in relation to harmonic tension. By generalizing that diatonic pattern to non-scalar pitch combinations, the framework for a pitch-based theory of emotions was outlined (Chapter 4). In so far as certain pitch combinations (e.g.,

major and minor chords) have "universal" emotional meaning, it can be said that the affective state of the RH is transmitted to the LH as a pattern of neuronal activity in the form of a tonotopic "intonation map" and expressed vocally by the LH. Several candidate neuronal mechanisms concerning the interhemispheric transfer of topographically-organized information are known and constitute a small set of physiologically-plausible processes by which information can be stored and transmitted to other brain regions (Chapter 5). While the scientific argument is as yet far from proven, the major pieces concerning cortical mapping, interhemispheric transfer, and the perception and the production of auditory affect have been empirically established. Those patterns of corticocortical communication are the physiological pieces that constitute the "brain code."

The arguments outlined in Part I are incomplete on many counts, but – even granting the theoretical possibility of such a "code" – many psychologists, and perhaps all those working on the problems of consciousness, would nonetheless object to an assertion that the establishment of a brain code is a *sufficient* understanding of the brain. Even if, for the sake of argument, one were to accept that the central dogma provides a rough summary of the major pathways of information flow in the human brain, and even if one were to accept that the cortical representation of emotions takes the form of activation of tonotopic maps, still questions remain: Why do we not only "think" about, for example, the pitch changes in the voice, but also "feel" them? It is relatively easy to understand how cognition and the control of behavior might occur due to the firing of neurons, but why do people subjectively experience emotions that are distinct from cognition?

And those are important questions.

In the chapters that follow, I argue that the key to these and related issues concerning subjectivity does not lie in the realm of neuropsychology; in other words, these issues are unlikely to be resolved at the "whole-brain" level and require that we rethink the entire issue of subjectivity starting from the neuron and working gradually to larger neural systems. Thus far all arguments concerning the brain code have been at the relatively macroscopic level of cortical mapping and the sequential flow of patterns of activation from one brain region to another. That, I maintain, is the proper level at which to address questions concerning the higher-level cognition that leads to characteristically-human behavior. In effect, much of the discussion of the brain code has been concerned with differences between the cortical activity of human beings and that of apes, with particular emphasis on the significance of the functional cerebral asymmetry that we, but apparently few other species, have evolved.

Behaviorally, it is certain that human beings are more than slightly oversized monkeys primarily because of talents with·language, tools and music – and all three involve functional asymmetry of the brain. Conversely, the biological commonality among many mammalian species, including man, is reflected in the extreme similarity of neuronal physiology and neurotransmitter biochemistry. There are few if any meaningful differences among mammalian species at the cellular level, so that the apparent "awareness" that we normally acknowledge to exist in human beings and certain other mammals is likely to be a function of the similarity in underlying neuron biology.

In the following two chapters, I argue that, unlike the lateralized cerebral functions that distinguish human beings from other animal species, the issues in consciousness studies should begin with the known *similarities* among all animal brains. Topics concerning specifically *human* consciousness will again demand consideration of hemispheric specialization (Chapter 8), but we share with many mammalian species a rudimentary level of animal "awareness" that kicks in well before sophisticated language functions and the complexities of the asymmetrical use of the cerebral hemispheres. As a consequence, questions about animal awareness must be addressed at a level where our commonality, rather than our human uniqueness, is central.

CHAPTER 6

Synapses and action potentials

Synopsis

Basic neurophysiology provides key insights into how, at the single-cell level, the "information-processing" underlying cognition and the "direct experience" of the environment underlying subjective awareness can occur. The nature of information-processing through excitatory/inhibitory synaptic effects is already well-understood, is easily simulated in computers, and is no longer controversial. Synaptic communication is thus widely understood to be the functional unit of cognition. A more controversial argument is developed here: that the neuron's other fundamental mode of contact with the biochemical environment, i.e., ion-exchange across the semipermeable cellular membrane, is the single-cell functional unit of sensitivity to the external world. Although the cell membrane normally maintains an ionic potential gradient between the cellular "self" and an external "non-self," periodically that barrier is dropped and there occurs a massive transmembrane flow of ionic charges during the action potential. The action potential of course plays an important role in triggering neurotransmitter release and consequent neuronal "information-processing," but the exchange of ions during the action potential itself is not "informational" – it is a means of directly "feeling" the electrostatic state of the environment simply by allowing the free diffusion of ions into and out of the cell.

Unanimity regarding definitions, much less mechanisms, has not been reached with regard to either cognition or consciousness, but the debate has reached some level of consensus concerning cognition. That is to say, there is wide agreement in the scientific community that *cognition is a function of neuronal information-processing via synaptic activity*. Such a statement may seem quite unremarkable, but its significance becomes apparent from the fact that far fewer of the same hard-headed, materialist scientists who are willing to summarize cognition so boldly would be willing to make a similar assertion with regard to consciousness. While the likelihood of "neuronal correlates of consciousness" is often acknowledged, the subjective quality of consciousness is thought by many to present new and unprecedented problems that simply might not be resolvable solely in terms of neuronal circuitry.

From the debate concerning the origins of subjectivity, several schools of thought – from quantum mechanical arguments to agent/environment interaction approaches – have emerged. Most of those arguments have some relevance to the consciousness debate, but it is surprising how few hypotheses concerning specifically the problem of subjectivity have been developed with neuron physiology (as distinct from brain circuitry) as the focus of the argument. While many appear to believe that a particularly cleverly-designed circuit might produce subjective awareness and while others argue that the topic of subjectivity can be resolved at a strictly philosophical, psychological or even quantum mechanical level, here I argue that the key to understanding the emergent property of subjectivity exhibited by multicellular nervous systems lies at the single-neuron level. The argument that I develop in this chapter is that there are cellular-level "protophenomena" (MacLennan 1996) whose explication allows for an understanding of the whole-brain phenomena of central interest in consciousness studies. To the physiologist, the idea that cell physiology may be relevant is again perhaps unremarkable, but consideration of neurophysiology has thus far *not*, according to most commentators on consciousness, provided answers to the problem of specifically subjectivity. While I agree that hypotheses about neuronal circuitry leave the subjectivity issue untouched, I advocate a return to the basic science of cellular life rather than pursuing still more sophisticated philosophical analysis for clues about how cells interact with their local environments. Once a proper understanding of neuronal functions is obtained *at the single-cell level*, then it is a relatively easy matter to extrapolate to the functions of many-neuron systems – and gradually build-up neuronal circuitry models that show various forms of "sensitivity" to specific types of stimuli, awareness of the external world, and ultimately "consciousness" of the role of self in a wider social context. However it is that such larger neuronal networks allow organisms gradually more complex interactions with the wider world, the starting point is the neuron.

The essential idea to be developed here is that physiological mechanisms that are already well-characterized at the neuron-level can be understood as providing the "unit" structures and processes that underlie the phenomena of mind. Alone, the cellular-level argument clearly will not suffice to answer many questions about whole-brain psychology. Moreover, there are many important issues of brain circuitry that are arguably the "main business" of consciousness studies and will be dealt with only briefly here, but there are certain phenomena at the single-cell level that are potentially relevant and that require our attention before we proceed to the more subtle, more complex issues addressed in consciousness studies.

Let me begin with a simple analogy.

A large-scale phenomenon such as democracy can be *empirically* studied only by examining the "low-level" mechanisms that make it possible. Our main concern may well be the "high-level" concepts of social justice and self-determination, but such social phenomena depend fundamentally on the details of the underlying democratic procedures undertaken by individual citizens (filling out ballot sheets, counting the votes, examining the locations and accessibility of voting stations, studying the frequency of citizen participation in referenda and the election of politicians, etc.). Stated negatively, the concept of democracy is *not* fruitfully approached by discussion of anecdotal impressions of the "atmosphere" of democratic or undemocratic states. Similarly, the processes that lead to both cognition and consciousness must be studied initially at the "local" cellular level. It is of course *not* the case that the discussion of mind ends with an understanding of the neuron, but it is essential that relevant facts at the cellular level be brought into the discussion of brain functions at the outset. In this view, just as there is no need to deny the reality of, for example, the particular "feel" of an authoritarian regime or the "feel" of an open democratic society – which are, after all, the kinds of subjective impressions that commonly motivate discussions about different types of social systems, there is no need to insist that the "subjective feeling" of consciousness is simply a delusion (Dennett 1991). On the contrary, the "moods" inherent to societies with different political systems are tangible and the "feeling" of consciousness is subjectively as real as any perception (Chalmers 1996; Searle 1997; Velmans 2000). The fact that we "feel conscious" – in addition to undertaking "neuronal information-processing" – can be denied only if the entire discussion of subjective consciousness is to be denied. But, just as political scientists do not debate the relative merits of the "atmosphere" of democracies versus autocracies, very little mileage (if lots of verbiage) can be squeezed from linguistic elaborations on the "feeling" of subjectivity. In the same way that there is very little content to a conversation that remains at the level of "I like Cuba" or "I like Switzerland," even an elaborate description of the "subjective feeling of consciousness" does not advance our understanding if mechanisms are not discussed. Therefore, rather than pursue the whole-brain issue solely at the macroscopic level, a more fruitful approach is first to study the underlying microscopic "protophenomena." Once the cellular-level mechanisms are understood, then a coherent extrapolation from microscopic to macroscopic becomes possible, and the large-scale phenomena can be discussed without invoking dualism, mysticism, complementarism or other "isms" that lack an empirical foundation. This view implies that we accept Dennett's (1991) hard-headed rejection of philosophical

dualism, but also accept Chalmer's (1996) persistent claim that there is indeed a subjectivity issue that needs explaining.

The present chapter is an attempt to return to neuron physiology in order to clarify what single-cell physiology can reveal about the properties of complex brains. The discussion will eventually be placed within the framework of recent philosophical discussions, because it is ultimately the philosopher's "hard problem" that needs to be solved, but the relevant arguments, elaborated on below, are concerned primarily with cell physiology. In order to show that this general approach of elucidating cellular mechanisms before proceeding to whole-brain problems is indeed sensible, the first step will be to address issues in neuronal information-processing. That is, before addressing the "hard problem" of consciousness, the "easy problem" of cognition needs to be reviewed.

A. What needs to be explained?

The easy problem of cognition

As complex as the neurophysiology of information-processing in multimillion neuronal networks may actually be, cognition is often considered among the so-called easy problems of the human mind for the simple reason that basic logical functions and memory mechanisms are theoretically well-understood and can be implemented in computers using simulated neurons. In principle and in technological fact, large numbers of such operations can be built into inorganic, robotic systems that are capable of certain kinds of admittedly primitive, but seemingly intelligent behavior. Highly-simplified examples of cognition, in the sense of sensorimotor integration of an autonomous agent interacting with its environment (Pfeifer & Scheier 1999), can be readily produced. Even the largest, most complex and most expensive artificial systems do not rival the functions of primitive living brains, so that no practitioner of modern AI would claim that anything approaching the full complexity of ant (much less, human) cognition has been implemented. Nevertheless, most researchers would strike optimistic notes concerning the possibility of someday building sophisticated silicon machines that "cognate" and flexibly control robot behavior in intelligent ways. In other words, the consensus is that there is no deep, fundamental mystery about how neurons are capable of storing and manipulating sensory information for the goal-directed control of motor behavior. Although there is still great uncertainty about *consciousness*, most neurosci-

entists and virtually everyone in the field of artificial intelligence believe that the essential mechanisms of *cognition* are known: primitive cognitive functions (e.g., edge detection) can be understood as the consequence of the activity of small networks of neurons, and less primitive functions (e.g., object recognition) can be understood as the consequence of the hierarchical arrangement of such small networks. Although the temporal structure of neuronal interactions adds a new level of complexity (Shastri & Ajjanagadde 1993; Maass & Bishop 1999), the simplicity of synaptic effects and the all-or-none nature of neuronal firing lead directly to coherent hypotheses about cognition. This has been fully appreciated since the mid-20th Century and the basic facts of neuron physiology were in fact the principal fuel that fed, initially, the cybernetic revolution (Wiener 1948), the brain science revolution (Hebb 1949) – and, subsequently, the revolutions in artificial intelligence and neural networks.

As illustrated in Figure 6-1, the logical functions that are understood to be the basis for cognition can be implemented using extremely small numbers of neurons. Provided only that there are sufficient numbers of such circuits and that they are suitably connected, it is a truism to say that *any* logical operation that can be explicitly defined can be implemented using these simple components.

As a consequence of the fact that (i) neurophysiology has provided a solid foundation for understanding the functions of real neurons and the fact that (ii) computer science has demonstrated how highly-simplified, simulated neurons can perform logical tasks, most scientists make the "leap of faith" and believe that cognition – including human cognition – is due to the operations of many neurons acting in a coordinated manner. While many details remain to be clarified concerning both the circuitry and the temporal patterning of neuronal firing (Chapter 7), there appears to be no obstacle to explaining complex cognitive functions on the basis of quite simple "non-cognitive" neuronal units. In other words, the bedrock belief in modern neuroscience is that the physical substrate of cognitive psychology *is* neuronal circuitry and therefore that all non-neuronal (physical, chemical, biological and social) influences on cognition have their effects directly or indirectly through neurons.

To state this conclusion in another way, it can be said that, although there is a huge distance between known (empirically verifiable or technologically implementable) neural network systems and the "real" cognition of human beings, few would argue that there is an unbridgeable, conceptual "gap" between the logical operations of small, knowable neural networks and the full-blown cognitive operations of mammalian nervous systems. In moving from small deterministic systems to large stochastic systems, the mechanisms of cognition

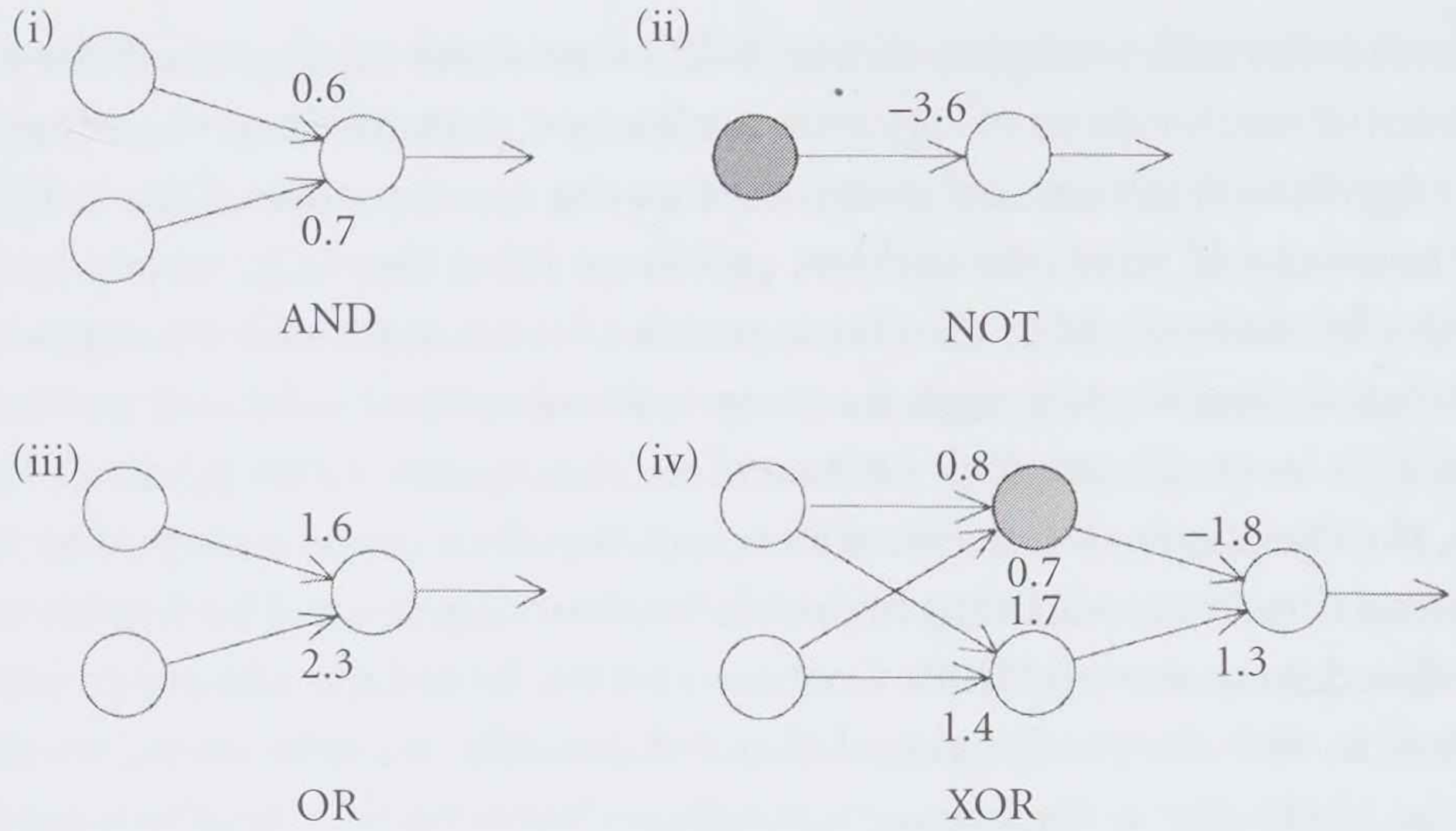


Figure 6-1. Possible neural implementations of a few basic logical functions (using neurons with firing thresholds of 1.0, and excitatory/inhibitory synaptic weights, as illustrated): (i) is the AND (&) function, (ii) is the NOT (∼) function, (iii) is the inclusive-OR (^) function, and (iv) is the exclusive-OR (XOR) function. According to Penrose (1989:522), "*all* logic gates can be constructed from just '∼' and '&' (or even just from the one operation '∼(A&B)' alone)." In fact, the neuronal circuits of even the smallest invertebrates with *known* numbers of neurons and *known* connectivity are not this simple (Kandel & Schwartz 1991), but the gist of the neurocognitive revolution has been the gradual realization that the full complexity of animal cognition is a consequence of such logic circuits in huge numbers and in complex temporal sequencing. This is the "astonishing hypothesis" that Crick (1994:3) correctly emphasizes: "This hypothesis [that mind is nothing more than the firing of neurons] is so alien to the ideas of most people alive today that it can truly be called astonishing." This type of synaptic communication is sufficient to explain the neuronal information-processing underlying cognition, but the subjectivity of awareness is a phenomenon that *cannot* be explained by synaptic events.

change from formal logic to fuzzy logic, but this is the same difference seen when moving from any small-scale mechanism to the related large-scale phenomenon. To return to the analogy with social systems, it is not unusual for social commentators to note the merits and demerits of a society solely on the basis of rather general impressions of the mood of a society, because that is the level that we most frequently experience foreign cultures. But, for the political scientist, the relevant issues are the mechanisms, universality and frequency of citizen participation in social control – and those phenomena can be measured and quantified at a "local" level by examining voting procedures, patterns of voter turn-out, and so on. By extrapolating to society at-large, general state-

ments about social systems can then be made, but the crux of the quantitative argument lies at the level of explicable small-scale mechanisms – not hunches, general impressions or feelings that are inherently subjective and not open to quantitative examination.

With regard to the jump from few-neuron systems to many-neuron systems, there inevitably arises a new complexity inherent to large numbers and probabilistic events, but materialist neuroscience allows for no conceptual mysteries. Difficult and complex problems may abound, but unquantifiable ghosts and ethereal forces belong to the speculations of previous eras. Already by the middle of the 20th Century, the outlines of cognition by means of basic neuronal operations were understood in principle and cognitive neuroscience of the 21st Century is well on its way to elucidating the actual mechanisms of both animal and human *cognition*. Progress is rapid and the conceptual foundations are solid.

The hard problem of subjective awareness

But if neurons are fundamentally such simple "information-processing" entities and the problems of cognition are soluble and "easy," why is consciousness not similarly understood? Why is there simply no consensus about basic mechanisms in the field of consciousness studies? Why is there widespread confusion about the possibility of silicon-based computer systems becoming aware? And why do philosophers insist that consciousness presents more difficult problems than does cognition?

The chief obstacle to achieving as basic an understanding of consciousness as has already been achieved for cognition is that the terms "awareness" and "consciousness" include – some would say, mean – "subjective feeling" (Chalmers 1996; Searle 1997; Shear 1998). Subjectivity has the unusual property of being knowable only to the individual experiencing it. As a consequence, the usual scientific methods for making objectively verifiable measurements on an external reality cannot be applied and there is, in fact, no known method of measuring directly the "internal" perspective of consciousness. As important as it would be to have a list of the physiological events that are necessary and sufficient for consciousness to arise (the "neuronal correlates of consciousness"), it is specifically the internal, subjective aspect that motivates the study of consciousness – and it is quite uncertain that any list of mechanisms will tell us why there is an internal feeling. What we would like to understand on the basis of scientific principles is why indeed there is subjectivity – not simply what the physiological correlates or the related behavior is. Whereas most psychologists

would argue that the essence of *cognition* is neuronal information-processing (and not the *feeling* of information-processing!), the essence of consciousness is specifically the feeling itself, with other aspects – word associations, autonomic nervous system responses, etc. – being related, but nonetheless secondary issues. This is to say, there seems to be an inherent difference between the objective, material mechanisms at the heart of cognition and the subjective feelings of the conscious state. Reconciling these two seemingly orthogonal dimensions is the conceptual task implied by the hard problem.

The explanatory gap

Most researchers (within the general framework of a scientific study of consciousness) agree that consciousness is somehow a natural phenomenon emergent from a material universe. Without this common starting point, no further discussion is possible; the evolutionary-materialist premise is therefore the point at which consciousness studies parts company with the mystics. But even with a common understanding of the material universe, the subjectivity of consciousness appears to be without precedent in the natural world and offers no appropriate objective "handle" to discuss it along with topics such as atomic weights, DNA sequences and blood pressure. This dilemma of having subjectivity all alone in a category by itself has led to many attempts to "objectify the subjective". An ever-growing literature in fields ranging from quantum mechanics to cosmology attests to the depth of the problem and the scope of the search, but nothing approaching consensus has been achieved: the hard problem has remained hard, unsolved, uncracked.

One of the many descriptions of the problem of subjective feeling is in terms of the so-called "explanatory gap." Declaration of a gap is recognition that even a complete, exhaustive, fully-scientific (physical, chemical, biological) explanation of brain phenomena would still suffer from a discrepancy between the known, objective mechanisms of neuron physiology and the equally-known, but non-quantifiable subjective aspects of psychology (Levine 1983, 2000). Some philosophers have doggedly pursued this line of thought, and have ended up arguing that the gap is simply unbridgeable: although we are capable of seeing that there is a problem, the human mind is incapable of obtaining a sufficiently objective perspective on itself (McGinn 1990). In that view, whatever successes may be achieved in identifying the neurological correlates of consciousness, there will always and necessarily remain aspects of the nature of consciousness that will remain opaque to human understanding. In essence, such a stance is tantamount to abandoning explanation, and need not concern

us further. In the end, if the science of consciousness fails to make progress on the hard problem, the pessimists will have been proven correct, but today many avenues of scientific inquiry remain to be explored. While we may salute the devil's advocate for explicating the defeatist point-of-view, more constructive results may yet be had from not abandoning the task.

Among those still hopeful of an explanation, it is axiomatic that there must be neuronal correlates of consciousness: as far as we know, mind does not drift free of a material, specifically biological, and more precisely a neuronal substrate (Searle 1997). But, although there may be general agreement that mind and brain are somehow locked together, it has remained thoroughly unclear how any particular set of material "correlates" could lead to subjectivity. What we therefore need in a fundamental theory of consciousness is some indication that the objectively-verifiable material mechanisms of the science of mind should indeed have a subjective "internal" dimension.

In so far as current explanations of consciousness do not provide links between subjective processes and objective brain mechanisms, an explanatory gap remains. To fully close that gap, (i) more precise specification of neuronal mechanisms (a scientific task) and (ii) the development of a self-consistent terminology concerning the relevant psychological phenomena (a philosophical task) are needed. Progress has certainly been made in both of those realms. Non-invasive techniques have succeeded in demonstrating brain responses to extremely subtle psychological phenomena (e.g., the sudden perception of the face of Jesus in a collection of splotches of ink [Landis et al. 1984] or the changing perspective on the Necker cube [Inui et al. 2000]). Philosophical progress is harder to quantify, but the focus on the phenomena of consciousness has arguably led to a more consistent use of the terminology in consciousness studies. (For example, the use of the word "consciousness" to mean exclusively "self-consciousness" is certainly on the wane.) Neither path of study is likely to reach final conclusions any time soon, but I maintain that it is already possible to draw a link between the objective and subjective worlds, i.e., to build a conceptual bridge across the explanatory gap.

Responses to the explanatory gap

At the risk of unfairly caricaturizing the various responses to the problem of the explanatory gap, it is convenient to summarize some of the prominent ideas concerning the problem of subjectivity along the lines of the five academic fields involved in the debate. The first is:

1. *The "engineer's" response: "What's the problem?!"*
To some, the idea of the explanatory gap is simply a misunderstanding, and requires no explanation. The engineer's tendency to underestimate the importance of philosophical debate is well-known, but forgivable in so far as it is accompanied by the tendency to go ahead and build clever machines: "On the basis of what is currently known, let us see what kinds of artificial intelligence, simulated cognition, and perhaps someday conscious behavior can be built." As constructive as that approach may be, the engineer's response to the explanatory gap is implicitly a denial that there is a problem. The philosopher, Dennett (1991), the engineer, Aleksander (1996), and the physiologist O'Regan (O'Regan & Noe 2001), adopt this view. It is of course an inherently frustrating task to debate with someone who insists he has no inner life – no feeling beyond neuronal "information-processing," but it is, in any case, a fact that most researchers in consciousness studies acknowledge both the existence of the explanatory gap and the need to close it. Particularly with the recent development of semi-intelligent robotic systems, examples of primitive "cognition without feeling" have become common, and the "gap" is no longer merely a theoretical construct. For this reason, the engineer's toys are among the clearest demonstrations that thinking is not the same thing as – and does not necessarily imply – feeling. The fact that even the engineers have no reluctance in flipping the off-switch of their robotic creations at the end of the work day clearly illustrates their true understanding about this issue. In any case, the counter-intuitive insistence that there is no explanatory gap – and that, in principle, the consciousness of a carbon-based system is no different from that of a silicon-based system – should be considered as a technological challenge, rather than as a principled philosophical argument. The challenge to build more intelligent silicon systems is fundamentally constructive, but the argument does not explain why the objective/subjective duality has exercised the minds of philosophers for centuries. The philosopher's focus on the two seemingly irreconcilable dimensions of the mind/brain problem is not a consequence of other-worldly mysticism or blind ignorance of the engineer's techniques, but a realization that mind and brain cannot be easily described in each other's terms. Denying the problem does not solve it.

The second response is:

2. *The "philosopher's" assertion that: "It is an unsolvable, eternal paradox!"*
At the far extreme are some philosophers who not only acknowledge the reality of the explanatory gap and discuss it as an example of the fascination and difficulty of philosophical inquiry, but have grown so accustomed to this para-

dox as a philosophical party trick that they have abandoned all hope of ever solving it (McGinn 1990). On the one hand, the academic world is indebted to those who have helped to define the gap and brought it full-force to the attention of neuroscientists (e.g., Levine 1983, 2000; Chalmers 1996), but we should nonetheless remain wary of delight in a paradox. The failure of well–intentioned philosophical analysis to resolve the issue without invoking some form of dualism should be seen as indication of the limitations of verbal gymnastics – and the need for introducing new, presumably scientific, evidence into the debate. Philosophers may have the admirable intellectual sensitivity to *detect* and *define* problems, but they may need assistance in *resolving* those problems in light of the empirical facts of modern neuroscience.

The next response is:

3. *The "physicist's" pronouncement that: "We can solve this – and every other – paradox by introducing the concepts of quantum theory!"*

A small, but vocal minority of physicists (e.g., Penrose 1989; Stapp 1993; Walker 2000) has tried to resolve various problems in biology and psychology by rephrasing old problems ("What is life?" "What is creativity?" "What is parapsychology?") in terms of the concepts of quantum physics. There can be no doubt that quantum theory at the atomic level presents an unusual view on physical reality, but it is unclear whether or not shifting the consciousness problem to the quantum level actually solves anything. The possibility of reformulating psychological problems in the mathematical formulae of the uncertainty principle is an obvious attraction, but simply rephrasing the paradoxes of consciousness in a different terminology is not a solution, unless specifically the connection between the objective and subjective realms is clarified. Particularly in light of the fact that the interpretation of quantum mechanics itself has remained a hotly debated topic strictly within the realm of atomic physics (e.g., Bell 1987; Bohm 1952; Bunge 1959; with, incidentally, many of the founders of quantum theory [Einstein, Planck, Schrödinger] rejecting the indeterminist philosophy favored by Bohr and Heisenberg), any quantum mechanical "solution" that simply replaces the objective/subjective conundrum with a different contentious puzzle cannot be considered as a contribution to the understanding of consciousness. As of today, the possible contribution of quantum theory to consciousness issues must be considered an open question – as yet, an unredeemed promissory note.

A fourth response is:

4. *The "psychologist's" promise that: "All we need is a better brain model!"*
The psychologist's answer is inherently constructive, and at least partially correct. Certainly, the brain structures and wiring diagram involved in the different forms of mental activity need to be identified before the general topics of animal awareness and human consciousness will be adequately understood. Among a large number of "brain-circuitry" suggestions, four are particularly noteworthy: (i) distinct circuits for (brainstem) arousal and (cortical) attention will be needed (e.g., Crick 1994); (ii) special reverberatory loops are required to prolong the "psychological now" (e.g., Edelman & Tononi 2000); (iii) special cerebral modules – "self-consciousness circuitry" – are needed to place a concept of "self" within a larger conceptual framework (e.g., Taylor 1999), and (iv) a critical threshold in the number of neurons must be crossed to obtain a sufficient complexity of mental associations to allow for awareness. All four points may be relevant to the emergence of certain types of human or animal mental states, but none of these suggestions indicates how specifically the *subjectivity* of consciousness can be introduced. Increasing the number or changing the configuration of the black boxes in a brain-circuitry model does not alone explain the internal perspective and declaration that subjectivity may simply arise from "complexity," or from a new neuronal component begs the question. Although the main work that remains for neuroscience to achieve an understanding of human brain functions involves issues of brain modules and cortical circuitry, it is not apparent how the problem of subjectivity itself can be resolved at that level. The very nature of the explanatory gap suggests that conceptual pieces of the puzzle are yet missing.

Finally, there is:

5. *The "physiologist's" response: "Examine the neuron more closely!"*
The physiologist's intrusion into psychological matters is not always welcomed because, whatever strengths reductionist science may have, higher-level psychological phenomena often require descriptions in a terminology that is foreign to cell physiology. Nonetheless, I suggest that the physiologist's approach to consciousness is the most appropriate and can contribute an important element to the debate on the explanatory gap. Thus far, two concrete suggestions have been advanced – one by Pribram (1991) and the other by Engel, Singer and colleagues (1999). Pribram emphasized the continuous, analog-character of consciousness and argued that dendritic encoding was a better candidate than the digital events at axonal synapses. Engel et al. have argued for more than a decade for the relevance of neuronal synchronization for cognitive binding (see below), and have recently extended the synchronization argument to

the phenomena of consciousness. While both points are important, the contention concerning a relationship between temporal neuronal firing and subjectivity is made by fiat rather than explanation. It may well be the case that consciousness is continuous and temporally coordinated, but neither point alone indicates why there is also a subjective dimension. To understand where subjectivity arises, it is necessary to explore the implications of cell physiology in greater detail.

B. Cell physiology

The first and most important fact about all living, organic systems is that they must maintain themselves in a state of relative biochemical stability to ensure long-term survival. Although the kinds of environments within which single-cell organisms can thrive are extremely diverse, the range of allowed biochemical changes conducive to life processes is greatly restricted in the case of multicellular animal organisms (Gross 1998). Specifically, the cytoplasmic interior must be maintained at a rather neutral pH – avoiding the drastic biochemical effects brought about by strongly acidic or alkaline solutions. It must also be kept at a temperature that is neither too warm that proteins will be denatured and made dysfunctional nor too cold that the watery interior will freeze and normal biochemical reactions prevented. Only by maintaining so-called "physiological" conditions, can the typical cell conserve its genetic information within a fluid, functional biochemical milieu for the benefit of future generations. The chief weapon used by all living cells to achieve such autonomy and biochemical balance is the protective cellular membrane (Figure 6-2). The membrane is a semipermeable, bilipid 2-D surface that separates the biochemical functions of the cellular "self" from an ever-changing, potentially dangerous external world. Within the cell membrane, the cytoplasm is maintained at a pH and state of aqueous fluidity that allows the cellular apparatus to function normally – maintaining an entropy-defying metabolism to further its own unlikely existence and replicating its genetic material to spread the patterns of its living form to corners unknown.

The second most important fact about the living cell is that, despite its autonomy and separateness from the external world made possible by the cellular membrane, the cell must nonetheless interact with the wider world. That is, the normally "hermetically sealed" cell must make sufficient contact with its local environment that it can absorb nutrients and interact in a controlled manner without entirely losing its biological integrity. There are many known

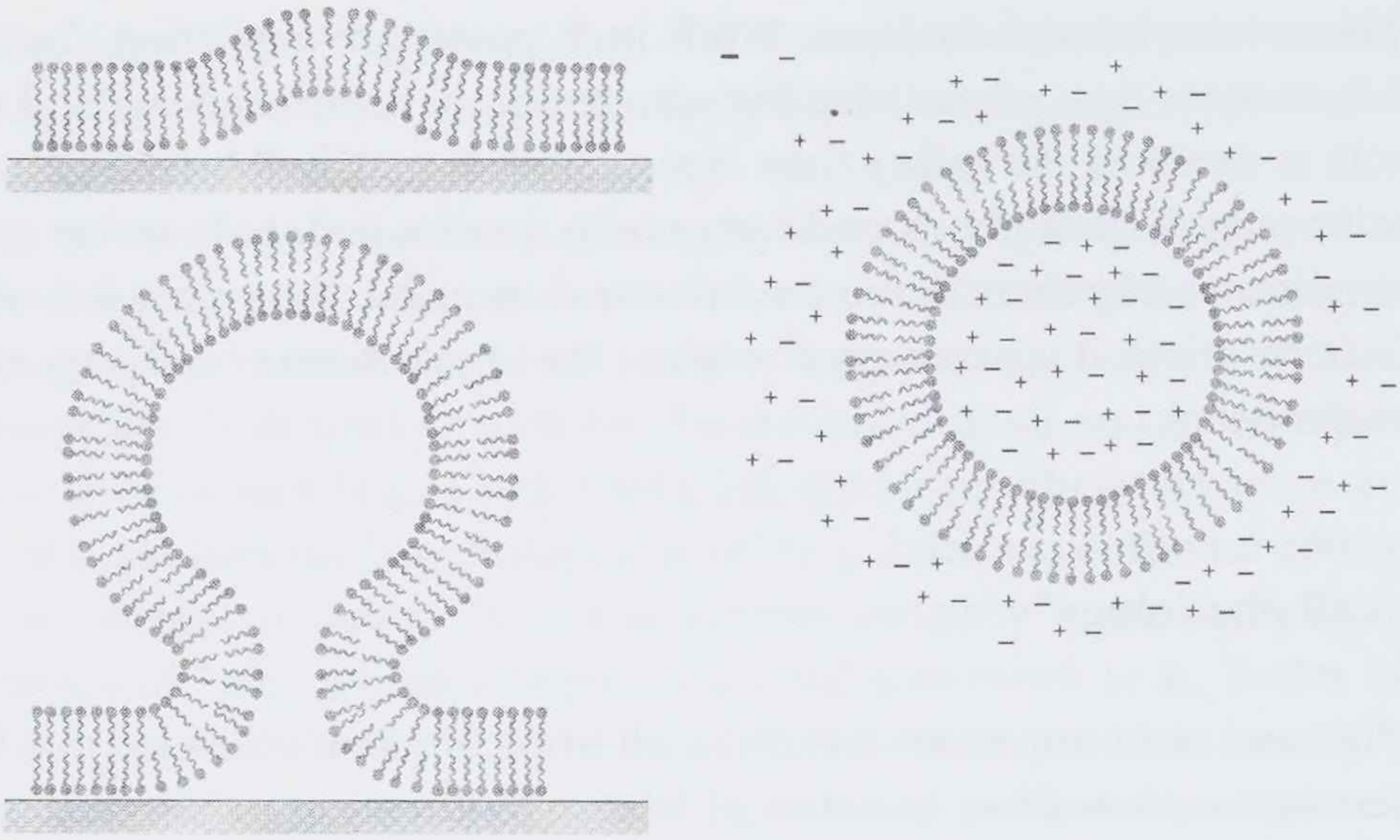


Figure 6-2. The formation of the bilipid cell membrane is a consequence of hydrophobic and hydrophilic molecular interactions. Due to the oil-in-water reluctance of (hydrophobic) lipid molecules to face polar water molecules, the lipids aggregate and tend to form 2-D surfaces that spontaneously wrap upon themselves to form hollow vacuoles containing an intracellular fluid.

mechanisms for importing and exporting molecules that are succinctly summarized by the phrase "semipermeable membrane". By allowing certain substances to cross the bilipid layers of the cell membrane, the cell can ingest nutritional, energy-supplying molecules and discharge the biochemical detritus of the metabolic process. Partial permeability is thus essential to the continued metabolic viability of all living cells (Figure 6-3).

The importing and exporting of materials across the cell membrane is a universal cellular process, but, on the cellular time-scale, it is rather lethargic and energy-consuming. While the slow pace of engulfing foreign objects, digesting them and expelling useless materials is sufficient for obtaining nutrition from the environment, transport mechanisms do not allow rapid interactions with the external world. For mere survival, transport mechanisms provide sufficient contact for most bacterial and plant systems, but for a more dynamic interaction with the world, a more rapid, more direct mode of contact would be valuable and has in fact evolved in the nerve (and muscle) cells of animal organisms. This mechanism of rapid interaction involves the related phenomena of (i) the transmembrane diffusion of ions across the axonal membrane during the action potential and (ii) the release/uptake of neurotransmitters at synapses. Both can be considered as specialized mechanisms for rapid contact

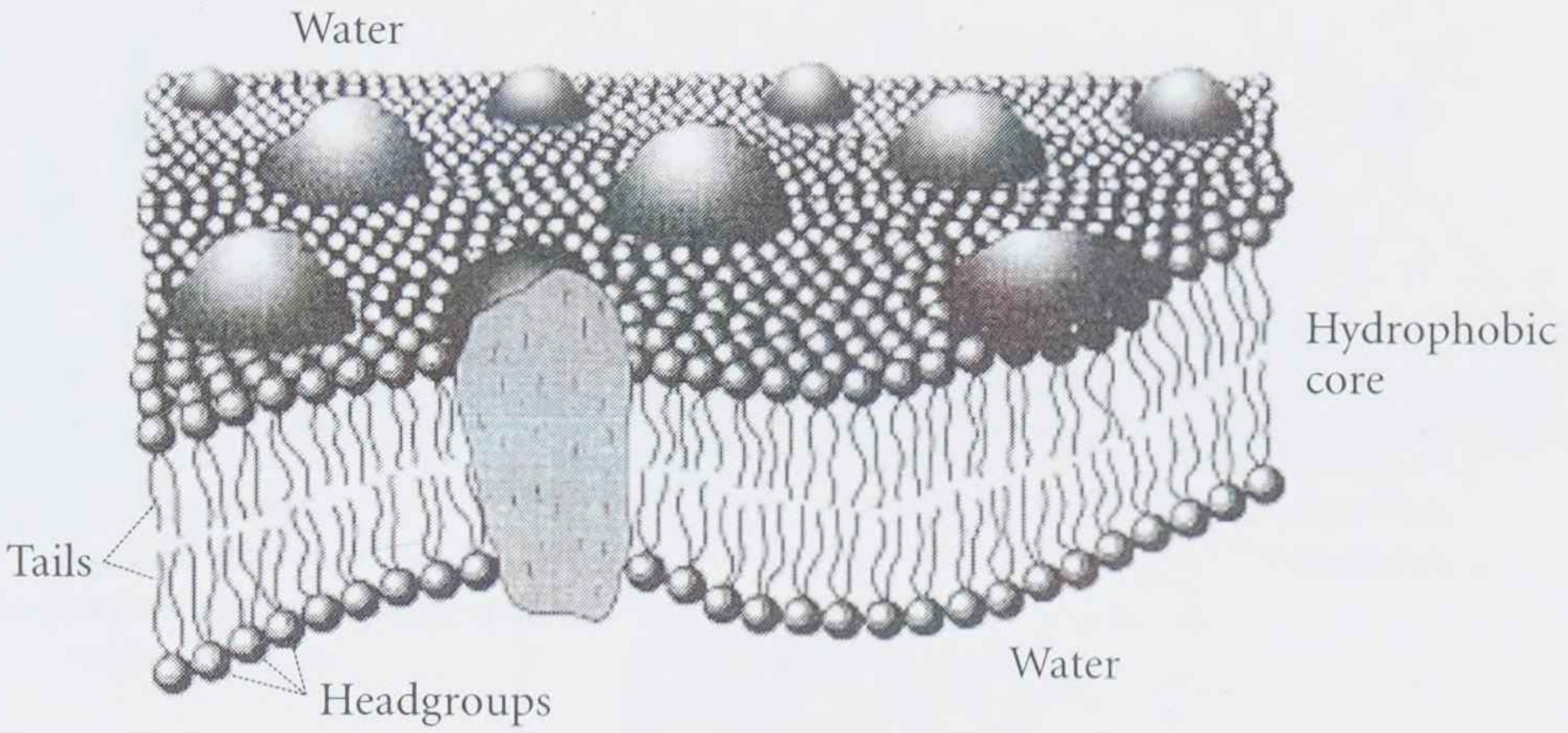


Figure 6-3. The bilipid cellular membrane with numerous channels and pores for contact with the external world (after Loewenstein 1999:214). Although the membrane separates the cellular self from a biochemically-dangerous external environment, the channels allow for selective transmembrane communication and contact with that environment.

with the extracellular environment, but they are fundamentally different types of cellular behavior that lead to different types of organism-level phenomena. Let us examine these processes in some detail.

All living cells are protected by semipermeable membranes, but the functions of the neuronal membrane are unusual. Not only must it maintain a barrier between the intra- and extracellular worlds, but, because neurons are not autonomous organisms, they have some properties that are atypical of cells in general. Among these, the most important are (i) the release and uptake of neurotransmitters at synapses and (ii) the action-potential that triggers synaptic activity. Suffice it to say here that, as a consequence of the arrival of the action potential at the axonal terminal, a complex set of physiological reactions results in the release of normally one type of neurotransmitter, that is in turn capable of influencing the excitability of neighboring, post-synaptic neurons (Figure 6-4).

Subsequent to transmitter release into the synapse and uptake by the post-synaptic membrane, the transmitters are re-released from that membrane and re-absorbed into the neuron from which they originally came. These processes require time and energy, but eventually the neuron is again in a condition to respond to a subsequent action potential with release of the same neurotransmitters.

As important as neurotransmitter release is – and as essential as it is for all discussions of neuronal information-processing, there is a second, related, but

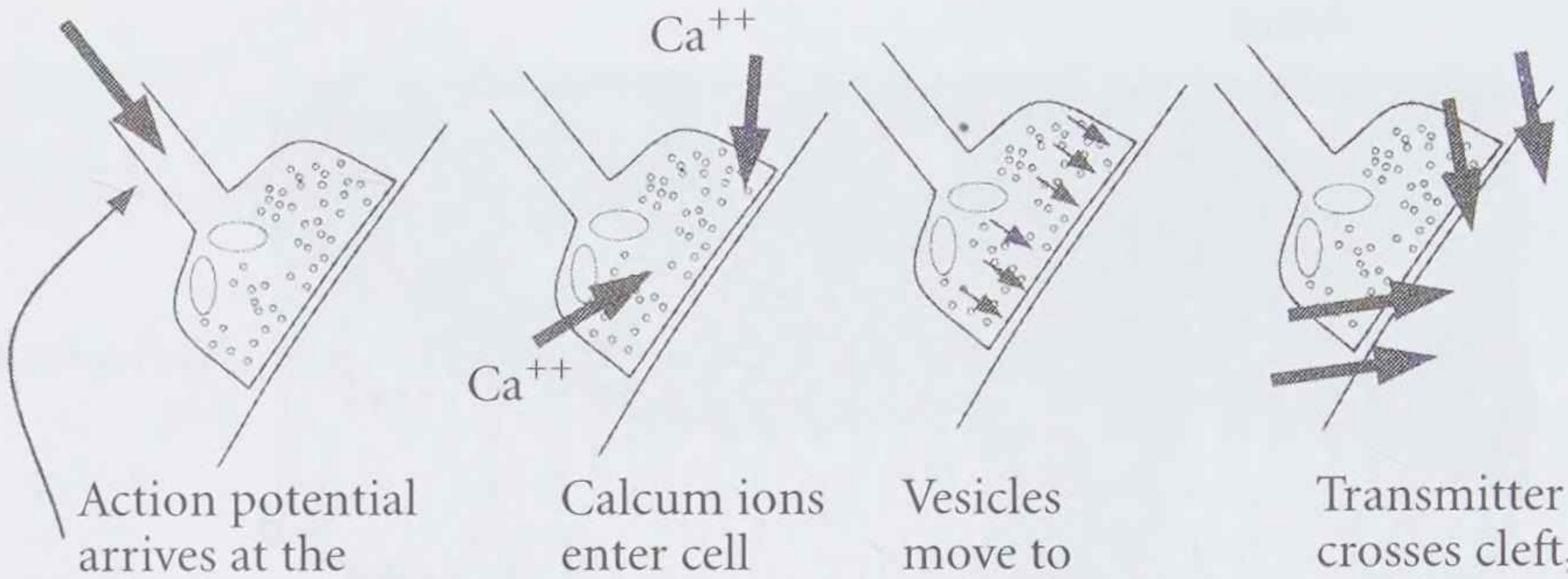


Figure 6-4. Neurotransmitter release (after Anderson 1995:39). The action potential induces the release of neurotransmitters that have effects on other nerve cells.

distinct neuronal process for contact with the extracellular world. This other fundamental process is the transmission of an impulse from the neuronal cell-body down the axonal cable to the synapse, i.e., neuronal conduction via the action potential. The mechanism of generation of the action potential is also well-understood: when ionic concentrations within the intracellular fluid cross a threshold value – the intracellular fluid becoming slightly less alkaline than its normal "resting" state, the neuron initiates a process at the axon hillock whereby external, positively-charged sodium ions are briefly allowed to flow into the cell, followed immediately by the outflow of positive potassium ions. In essence, the action potential is an interval of a few milliseconds during which ions flow from regions of high to low concentrations. Thereafter, the ion channels are shut and the relatively slow process begins of re-establishing the resting potential by transporting ions one-by-one across the membrane. Through this moment of free diffusion of ions during the action potential, *the neuron directly experiences the electrotonic state of the extracellular fluid*, and adjusts its own internal state in response. The causal factors driving the dynamics are the relative concentrations of positive and negative charges – chemistry, not biology. In essence, when ion concentrations change such that the difference between the intra- and extracellular pH is reduced, the biological neuron responds by allowing a brief return to the physicochemical processes of charged-particle diffusion – a moment of statistical thermodynamics that the living cell normally prohibits. By allowing a momentary relaxation of biological control, the chemical diffusion stimulates the neuron to modify its intracellular pH in direct response to changes in the biochemical environment (Figure 6-5).

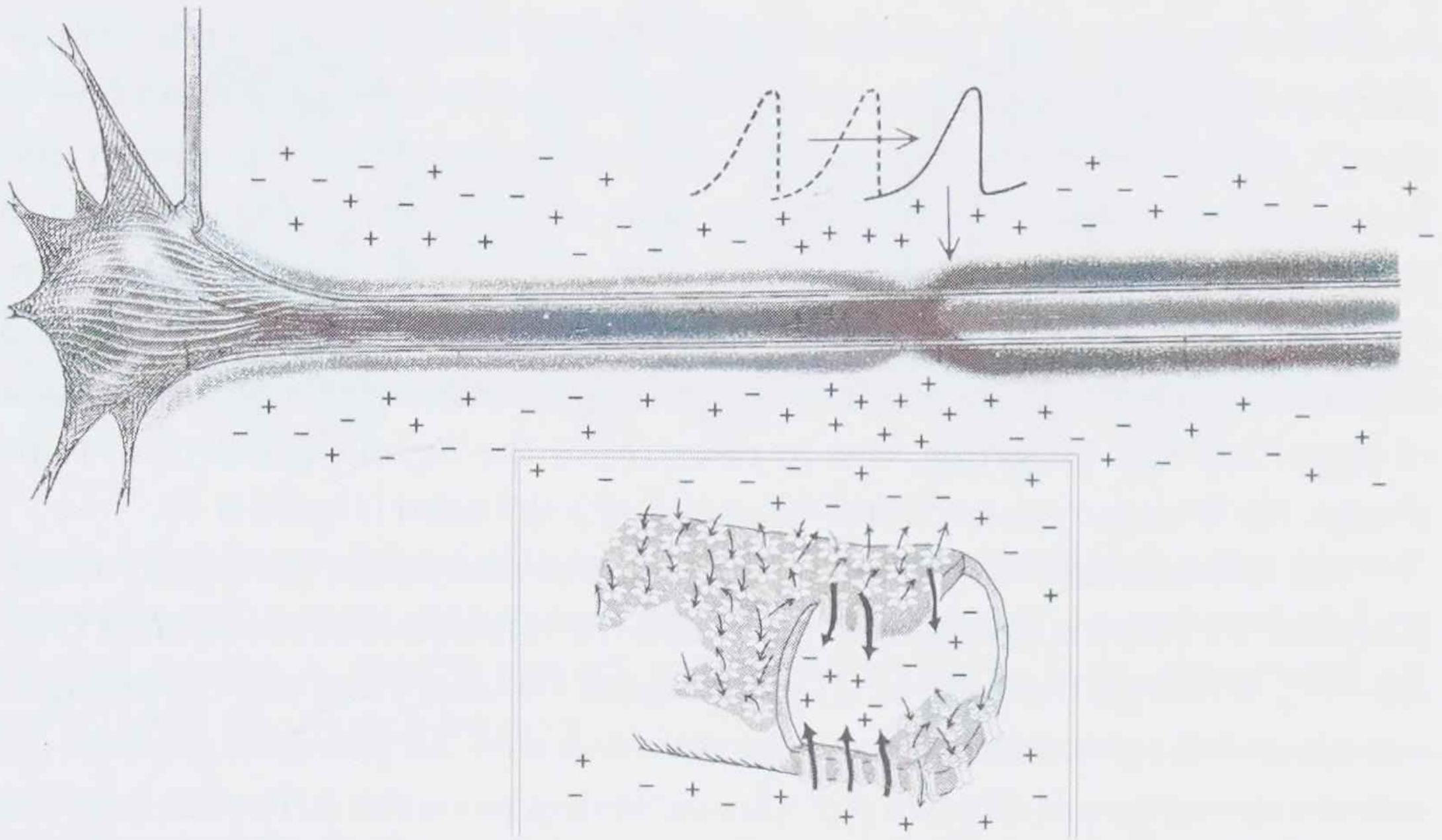


Figure 6-5. From a cognitive perspective, the directionality of the action potential is most important, because it triggers synaptic activity, but the action potential itself is a form of communication between the neuron and its environment. The inset illustrates the free diffusion of charge from the extra- to the intracellular environment.

What is most interesting about the neuron's response to the ~10 mV decrease in the resting potential is that it does not respond by directly increasing the intracellular negative charges – as might be expected from a simple homeostasis mechanism. The small decrease in negative ion concentration is indication that the electrotonic state of the intracellular environment has become more similar to the extracellular environment. Instead of clamping down on all cross-membrane traffic or actively pumping negative ions into the cell, the neuron opens the floodgates – thereby initiating a massive influx of sodium. That stimulus in turn triggers a reverse flow of potassium out of the cell, and then the active pumping of various ions until the normal alkaline state is re-established.

Textbooks typically treat the action potential as the neuron's surprisingly clumsy way of igniting the synapse in response to sufficient dendritic stimulation. Instead of simply sending free electrons – as in a copper wire – down the axon, the neuron has developed a mechanism whereby the impulse radiates out from the cell body through a sequential opening of the cell membrane. Starting at the axonal hillock, ion pores that are normally closed to the free diffusion of charged ions are opened, and the ionic charges inside and outside of the cellular membrane are free to follow the physicochemical push and pull of various ion gradients. As a consequence of this "low-level" physical process where ions tend

to move away from high concentrations of similar ions, the momentary transparency of the cellular membrane induces an uncharacteristically-rapid flow of charge. At the outset, the bulk of the ionic flow is that of positive sodium ions flowing into the cell. After 1–2 milliseconds of such ionic diffusion, the ion channels close and this process repeats itself at the next step along the axon. Potassium channels then briefly open for an outflow of positive charge, and eventually all channels are closed and the neuron returns to the slower process of importing and exporting ions to re-establish the normal gradient of ionic charge, the impulse proceeds step by step down the axon (Figure 6-6).

The free diffusion of ions across the neuronal membrane during the action potential is a process that occurs only when the ion-channels are opened (Figure 6-7). Although there are other mechanisms for importing and exporting of ions *against* the potential gradient, ion diffusion *with* the potential gradient is a massive movement of charged particles occurring when the difference between intra- and extracellular potential gradients has exceeded a threshold value.

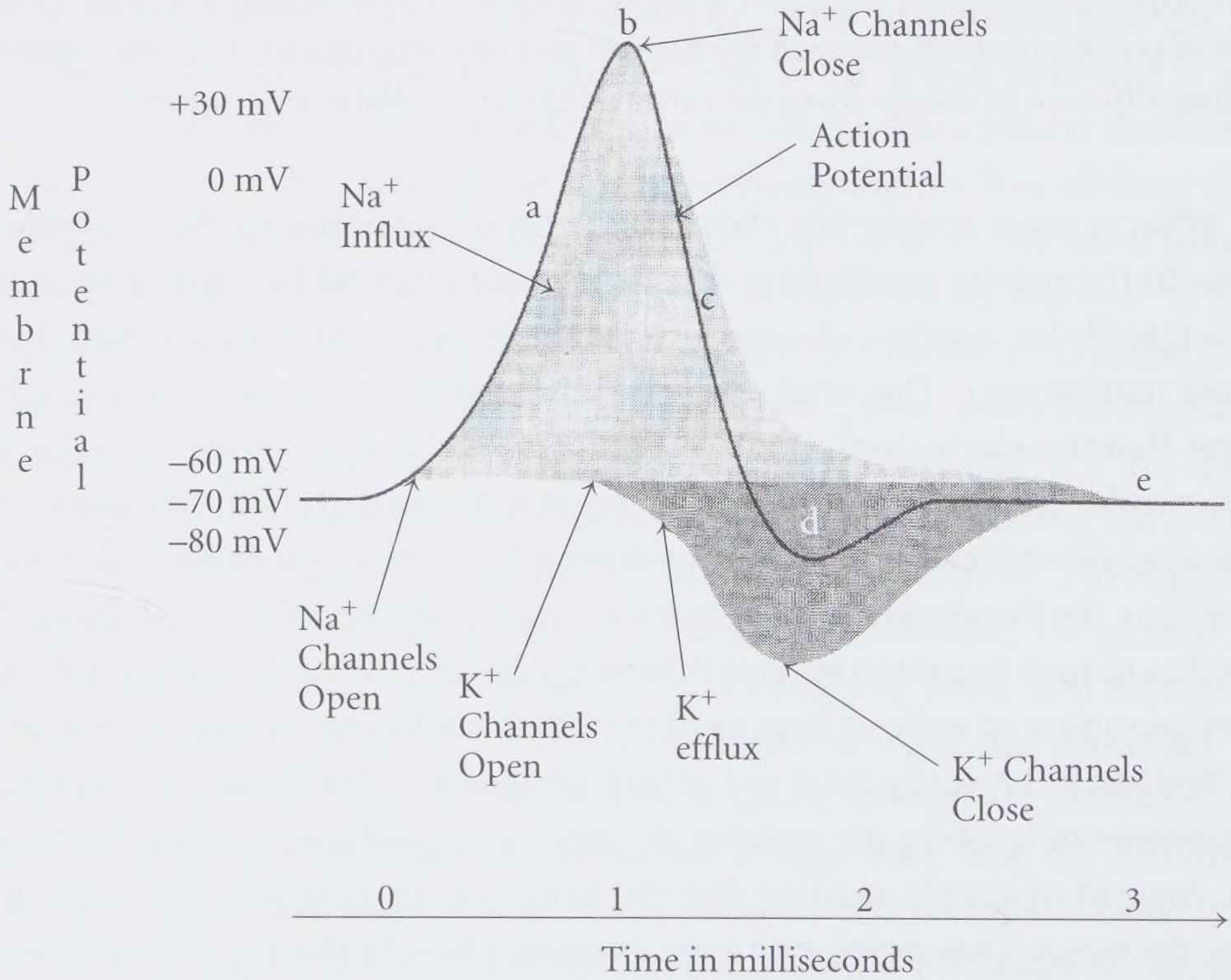


Figure 6-6. Details of the action potential (after Grilly 2002: 59). Sodium channels are opened due to a small *decrease* in the resting potential from –80 to –70 milliVolts. The influx of positive charges triggers the opening of potassium channels and the closing of the sodium channels. Eventually, the resting potential is re-established.

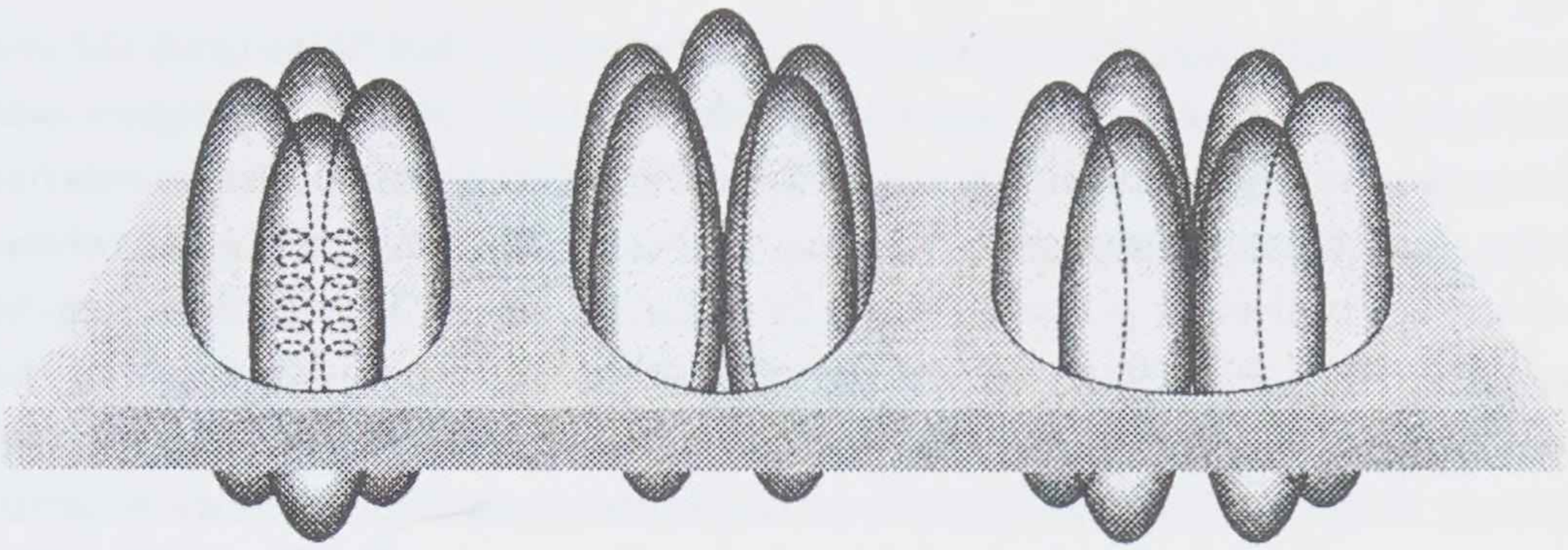


Figure 6-7. The structure of some ion channels (after Loewenstein 1999: 288). Depending on the radius of the channel and the nature of the charges lining its interior, large or small, positive or negative ions will pass freely between the intra- and extracellular worlds when the channels are open.

The magnitude of ion exchange is truly astounding – $\sim 10^8$ ions/channel/second (Koch 1999), a deluge of charged particles that is many orders of magnitude greater than the active transport mechanisms. In fact, since the pores remain open for only 1–2 milliseconds, the net movement of charge is not huge, but relative to the normal dynamics of cross-membrane traffic, the action potential can be characterized as a wild moment of "direct contact" with the external biochemical world. The essential point is that, in addition to an "impulse transmission" role (that leads to synaptic transmission and neuronal information-processing), the action potential itself is a brief moment of "biologically-uncharacteristic" opening of the cellular system to direct input from the environment.

Ultimately, the synaptic region is stimulated to release its neurotransmitters, but the membrane-mediated physiological process is notably different from an electron-conductance mechanism that the engineer might design. Not only is there "impulse conduction" from dendrites to cell body to axon to synapse, but there is also a material exchange in an orthogonal direction between the cell and its environment – and *that orthogonal process has no correlate in the copper wire or in a computer simulation*. For strictly "information-processing" purposes, the difference between carbon- and silicon-based systems is perhaps trivial, but there are important differences between brains and computers – not at the atomic level, i.e., organic versus inorganic chemistry, but at the "functional unit" level, i.e., *living* cells versus *non-living* circuitry. The difference is crucial not in terms of the information-processing capabilities of the two systems, but in terms of whether or not there is any real contact between the cognating system and its immediate environment.

The action potential, a wave of channel-opening that flows from the cell body down the length of the axon, proceeds toward the terminal synapses, and there causes neurotransmitter release. Synaptic activity is of course the mechanism by which a neuron has influence on other neurons, but neurotransmitter release is a molecular process distinct from ion exchange. Specifically, once the action potential has flowed from the cell body to the terminal region of the axon, the massive ion inflow results in the so-called exocytosis of vesicles containing neurotransmitters. The neurotransmitter molecules are released from the vesicles into the synaptic cleft and then have a direct effect on the ion permeability of the post-synaptic membrane.

Depending on the type of molecule contained in the synaptic vesicles, synaptic activity will have either of two effects on the post-synaptic membrane of its neighboring neurons – either increasing or decreasing the intracellular ionic concentrations through secondary biochemical effects. Much is known about the molecular mechanisms, but, from a psychological perspective, what is important is that the synaptic communication relays the excitation of one neuron to other neurons by increasing/decreasing intracellular charge. Although one might be tempted to dismiss these cellular processes as rather uninteresting nuts-and-bolts physiology, such fundamental mechanisms of contact between the neuron and its external environment are truly the essence of what neurons do – and suggest cellular-level definitions of the "protophenomena" that underlie the more interesting organism-level phenomena of animal psychology.

The "story" of neuron-to-neuron communications and the many subtle variations on that theme rightfully takes center-stage in any discussion of how the brain works. As briefly mentioned above, the entire "book" of cognition can be convincingly constructed out of these relatively simple non-cognitive neuronal events – and there is probably not a neuroscientist alive who does not believe that synaptic communication is the building-block of higher-level cognition. But it bears repeating that synaptic communication is one of only two modes of contact between the neuron and its world – and both modes have implications for "whole-brain" psychological phenomena.

Primarily through these two processes, the neuron is able to maintain an intracellular pH conducive to normal cellular functions, but also to have effects on other neurons. Both to the individual neuron and to the organism within which it dwells, gross changes in the concentrations of positive and negative ions indicate an "abnormal" biochemical state – and therefore circumstances that could be fatal. In order for the *organism* to take action to bring about a return to a normal biochemical state, it must (i) sense where in the organ-

ism biochemical imbalances have occurred and (ii) make cognitive deductions about appropriate behavior. For both of these functions, neurons are essential. In other words, the neuron is the functional unit for both sensing the external world (ion exchange) and initiating motor behavior that will lead to changes in that external world (neurotransmitter release).

In this view, the release of neurotransmitters – which is the end result of impulse conduction – is only one half of the neuron's interaction with its external world. In an informational sense, the release of inhibitory or excitatory molecules and the subsequent effect on other neurons is all-important, but for the individual neuron, the ion-exchanges occurring down the length of the axonal processes are more drastic, more energetically-demanding, and more direct interactions with the environment (Figure 6-8).

It is not for us to make anthropomorphic guesses about what the release of neurotransmitters or the action potential itself "feels like" for the neuron, but it is a fact that the neuron periodically opens up to its local environment and allows the thermodynamic flow of ions into and out of the cell. It is of course true that the "opening up" of the neuron to environmental molecules is restricted to

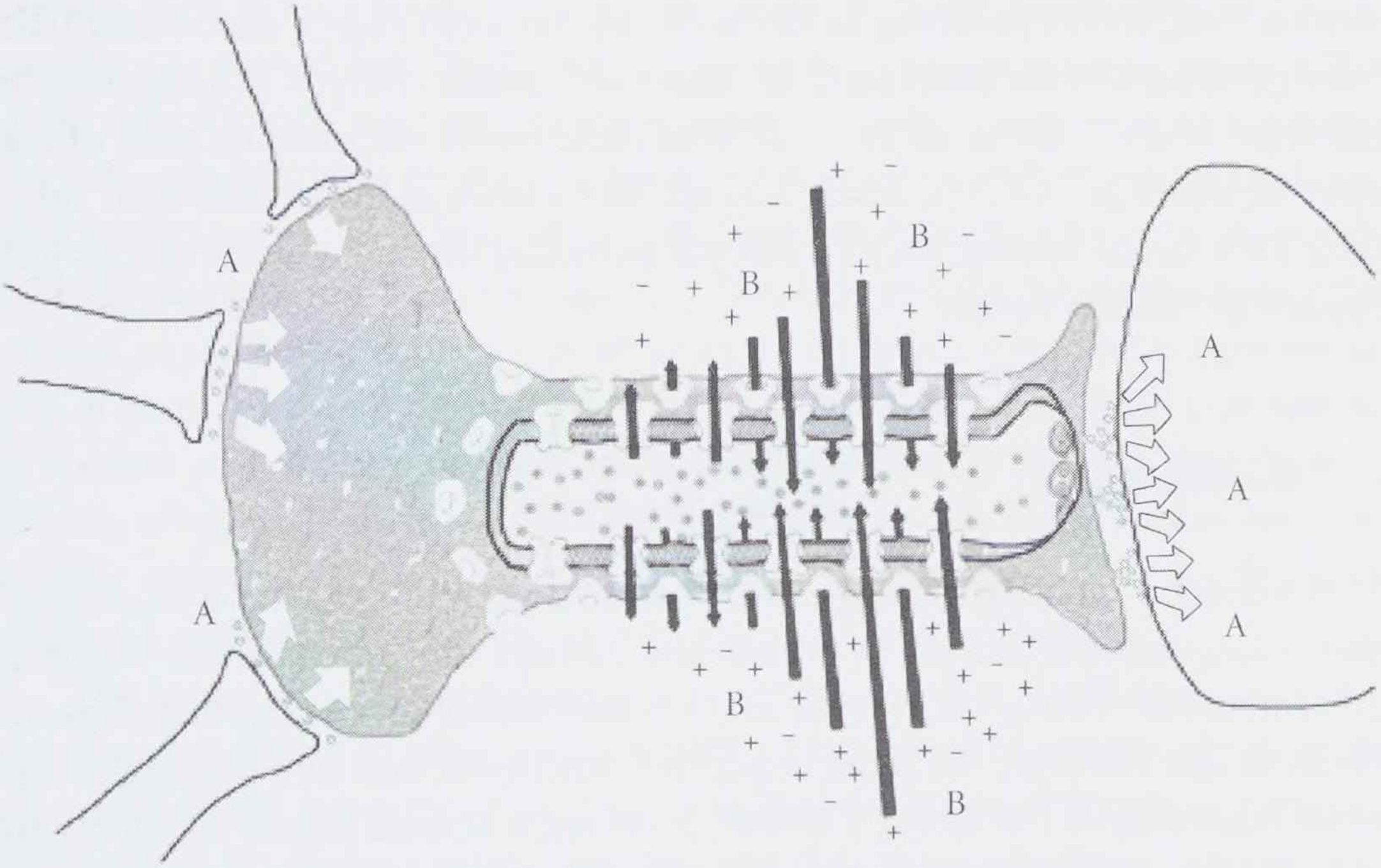


Figure 6-8. The two modes of contact between a neuron and its extracellular environment. Synaptic communications (A) are central to the phenomena of cognition, whereas the deluge of ions from the extracellular fluid into the cell during the action potential is arguably the single-cell protophenomenon of subjectivity – i.e., directly "feeling" the (ionic state of the) external world (Cook 1999c, 2000a, 2002a, b).

a very narrow range of charged atoms (principally, sodium, calcium, potassium and chlorine ions). The opening of ion channels is thus a controlled "exchange" that is far more regulated than simply allowing a molecular free-for-all through gaping holes in the cellular membrane. Nevertheless, the opening up of specific types of ion channels is, for the cell, sufficient to learn of general biochemical changes in the extracellular fluid. In analogy with the whole-organism perception of external stimuli, channel opening in the single neuron is similar to the narrow bands of sensory channels that are available to the animal organism. The band of visible light and the band of audible frequencies constitute only a tiny fraction of available environmental information, but by having visual and auditory mechanisms that are tuned to even a small spectrum of external physical events, those limited windows onto the external world that are provided by the organism's eyes and ears suffice to keep it informed of relevant environment changes. By the same token, the range of "information" obtained by the cell through the inflow of extracellular ions is extremely limited, but nonetheless constitutes vital indication of the neuron's current biochemical surroundings. It is worth recalling that the ions in physiological solutions are of course nothing more than concentrations of positive and negative charge.[1] At a fundamental level, there is nothing in the cell's environment except electromagnetic fields produced by the movement of these electrostatic charges. What the living cell must know – above all else – is what the current electrostatic state of the external world is – i.e., the acidic and alkaline nature of the extracellular solution. This can be known only by directly sampling the ion concentrations just beyond its cellular border.

C. Implications

At the heart of neurophysiology is the phenomenon of synaptic communication – i.e., the mechanism by which one neuron can influence the firing of other neurons. This is the essence of the information-processing that neurons perform and therefore the most important functional unit underlying the *cognitive* functions of the brain. There is, however, a second means by which the individual neuron interacts with its surrounding milieu – and that is the controlled flow of ions across the entire expanse of the neuronal membrane. The inflow and outflow of materials (particularly, charged ions) across the cellular membrane is an essential part of the mechanism of neuronal conduction and therefore the necessary precursor to the release of neurotransmitters at the axonal synapse. Synaptic transmission and ion diffusion are, however, distinct

phenomena at the cellular level: they involve different molecular mechanisms, and contribute to different types of organism-level, psychological phenomena. On the one hand, neurotransmitter release is the process by which neurons have direct effects on one another and how they bring about the motor responses of various effector mechanisms. While today we may have only piecemeal glimpses of how the totality of cognition can be accomplished on the basis of such neuronal communications, the underlying physiological mechanisms are known, and we have a good understanding of how analogous processes, as implemented in artificial neural networks, can achieve rudimentary cognitive functions. Many difficult questions remain unanswered, but there are, in principle, no remaining mysteries surrounding neuronal information-processing through synaptic transmission.

On the other hand, the physiology of the action potential is also well understood. Although the emphasis is normally placed on the role of the action potential in allowing for neuronal conduction, the relevant membrane dynamics are known to be a process through which the ionic concentration of the intracellular medium is maintained within physiological limits through periodic "direct contact" with the extracellular medium. Unlike neurotransmitter release, there is no informational significance – from the perspective of the organism – to the ion flow itself, except in so far as it triggers synaptic activity. But quite distinct from the interneuronal communications occurring at synapses, the action potential is an *intraneuronal* event – concerned only with the ionic concentrations of the lone neuron itself. Instead of the discrete event of synaptic transmission, ion flow is an ongoing "analog" process taking place over the surface of the entire neuron – peaking at the moment of the action potential, but occurring at a lower rate at all times. For the neuron, these two modes of contact with the extracellular world are fundamental: the information processing of synapses and the diffusion of charged ions across the semipermeable cellular membrane. Both are vital processes, but they are of a very different character with different significance to the individual neuron.

The topic of this chapter has been a fairly simple fact about neuronal life – the neuron's ability to make contact with its biochemical surroundings while maintaining the integrity of its own biological functions. Through a presumably long and complex evolutionary process, this balance between internal stability and external contact has been achieved through the emergence of the semi-permeable cellular membrane. On the one hand, maintaining a distinction between the molecular constituents of the cellular "self" and the outer world through the workings of a closed membrane is perhaps the most important function of any living cell, but, at the same time, allowing some form

of molecular exchange with that world is also essential for survival. Neurons in animal organisms have retained the fundamental "closed membrane" structure typical of all living cells, but they have also evolved quite unusual mechanisms for direct contact with the external biochemical world. The "dual functions" of the neuron – the action potential and neurotransmitter release – reflect two distinct modes of interaction with that world. While some form of self/other interaction is necessary for all living systems to survive in suitable environments, only animal organisms have evolved these mechanisms of rapid-fire, direct exchange.

What has *not* been addressed in this chapter is how these actions of the individual neuron can be coordinated with other neurons. On its own, each neuron can have only a small effect on other neurons through neurotransmitter release, and can sense only a very local biochemical environment through moments of increased membrane permeability. In order to move from the level of "neuronal unit" functions to the phenomena of organism-level cognition and consciousness, mechanisms for coordinating these primitive neuronal functions need to be considered.

Notes

1. Because the action potential is central to the issue of neurotransmitter release, the ion dynamics that underlie it have been intensely studied, but the flow of sodium and potassium ions is only one aspect of the maintenance of an appropriate balance of electrostatic charge. This topic is dealt with in detail in a remarkable monograph, *pH and Brain Function*, by Kaila and Ransom (1998). That book alone serves as an antidote to the synapse-centered view of the neuron that lies at the heart of cognitive neuroscience. Among the many insights provided there, perhaps the most important is the idea that the glial cells (that outnumber neurons at a ratio of 10:1 in the mammalian CNS) work to control the excesses of ion concentrations that arise during neuronal functions. The neuron normally has a slightly alkaline pH (7.4) that must be actively maintained by extrusion of positive ions, particular hydrogen ions. During sustained excitatory or inhibitory neuronal activity, the intracellular pH falls (acidosis), while that of surrounding glial cells rises (alkalosis) correspondingly (Ballanyi & Kaila 1998). "Glial cells … may be H+ caretakers, which not only regulate their cytoplasmic pH, but also help to control the acid-base content of the neuronal environment" (Deitmer 1998). Noteworthy is the fact that extreme neuronal acidosis leads to impairments of consciousness (Kauppinen & Williams 1998) – clearly indicating the importance of the charge state of neurons in maintaining normal sensitivity.

Chapter 7

Synchronization

Synopsis

Individual neurons interact with their immediate environments by means of exchanges of materials across the neuronal membrane, but to proceed from unit neuronal functions to truly psychological functions, the coordinated actions of many neurons must be considered. Unlike most other cell types, neurons are capable of acting in concert to amplify both their capacity for information-processing and their capacity for sensing the extracellular world. Amplification is achieved through the synchronization of action potentials such that many neurons detect their biochemical states in unison and deliver their information-processing results in a time-locked fashion. In this view, the synchronized "information-processing" and the synchronized "sensitivity" of many neurons acting together result in, respectively, the "cognition" and "awareness" of the animal organism.

Various suggestions for dealing with the explanatory gap in the consciousness literature have already been mentioned and dismissed as unsatisfactory in so far as they provide no indication of the "internalness" of subjectivity. Even the fanciest of circuit diagrams, recursive algorithm or cerebral module for emulating "self" end up being typical neuronal structures – and therefore offer no possible resolution of the explanatory gap. In contrast, by returning to the single-cell level, it was shown that there are two fundamental modes of neuronal contact with the external world – neurotransmitter release and cross-membrane ion exchange. These two mechanisms are known to be, respectively, (i) how the neuron has effects on other neurons and (ii) how the neuron senses the electrotonic state of its biochemical surroundings. It is therefore arguably the case that we already have some scientific knowledge concerning how a (primitive unicellular) living system can "feel" the world outside itself by quite literally allowing a small part of that external world to rush in. If ion influx through the semipermeable membrane and neurotransmitter release are the predominant means by which the neuron senses the external world and has effects on it, it nonetheless remains to see how these primitive neuronal functions might lead to the macroscopic phenomena of cognition and awareness. Addressing that

problem is the constructive argument of the present chapter, but before that discussion, some dirty linen needs to be hung out.

A. Scandal

There is a long-standing, unacknowledged "scandal" in the field known as consciousness studies. It can be simply stated, but not easily explained. The scandal is that, although the phenomenon of "awareness" is known for a certainty to arise in animal organisms with nervous systems, hypotheses concerning causal mechanisms have included everything *except* the neuron. There have been a few notable exceptions to the general neglect of neuronal properties and, to be fair, much of the theoretical work in consciousness studies has rightly been concerned with the configuration of neural systems (*multineuron* neural networks), but there has been scandalously little focus on the properties of neurons themselves. Simply to state the problem is to suggest that there may be some naked emperors loose in the academy, but it needs to be said yet again: As every healthy 10-year-old fully understands, *animals* exhibit signs of awareness, but *plants* and *minerals* do not. If we are interested in more subtle questions concerning, for example, the differences between sleeping brains and awake brains, or between infant and adult brains, or in clarifying the types of awareness exhibited by different animal species, then the configuration of multicellular neural *systems* rather than single-cell neuron physiology is undoubtedly the right level at which to start. But for first-order issues concerning why there is a psychology of animals, but not of other material entities, surely the neuron should be on stage front-and-center.

There are perhaps not many tips to be taken from 10-year-olds on the topic of consciousness, but the fact of the matter is that few scholars have focused on neurophysiology, while many have been led astray by, for example, fantasies about robot "awareness". Unfortunately, the fictitious AI arguments do not reflect a sophisticated understanding of CPU design, cache memory or recursive programming, but rather a Hollywood-style inattention to basic science. Eventually, we may want to address questions about what behavioral capabilities inorganic "brains" can in principle achieve, but the first order of business lies in the life sciences. Since awareness is present in animals with nervous systems, but absent elsewhere, and since it is known that the types of neurons found in all animal species are generally indistinguishable from those in other species, the subtle difference between, for example, monkey and human awareness is not likely to be due to cell types or minor differences in neurotransmitter bio-

chemistry. First-and-foremost, the fact that all mammals share a fundamental neurobiology clearly indicates that some of the core issues in consciousness studies should be addressed at the neuronal level.

So, before the second-order issues about types of animal awareness are addressed, it is important to answer basic questions: Why do plants react to external stimuli, but are not "aware"? What is it about neurons that makes animal organisms so much more "alive" than plant organisms? And why do even the most powerful computer systems lack the faintest glimmer of spontaneous mental life? Perhaps life-like robots can someday be constructed, and perhaps some broader definition of consciousness would encompass the phenomena of plant life – or, for that matter, the fundamental connectedness of all physical matter. But before we set out on such a cosmic journey, it is important that we have a clear idea about what is meant by "awareness" in those systems that are demonstrably, unambiguously and universally acknowledged to have it. If we set aside questions about the relative mental activity of rats, dogs, monkeys and men, and concentrate on their common characteristics, it is clear that the common denominator for all such organisms that show signs of "awareness" is the presence of neurons. How far down the evolutionary line of the animal kingdom we might still want to use the word "awareness" is debatable, but neurons are a necessary (if perhaps not sufficient) prerequisite to anything currently existing that we might describe as mental life.

To begin with, we must ask why the neuron has thus far been so ignored. The most obvious reason is that the nerve cell is, in many respects, a rather typical biological cell. If we were to somehow attribute awareness to all living cells, then we would be obliged to conclude that plants and bacteria are also aware. Although there is a certain logical consistency to that kind of panpsychism, clearly it leads away from a definition of awareness in which human (or mammalian) awareness is a relatively well-developed form and that of multi-billion cellular systems such as Sequoias and coral reefs is relatively undeveloped. It is therefore (rightly) concluded that the light of "consciousness" is not to be found emanating from every living cell. For this reason alone, many scholars may have concluded that cell physiology is not the key to the questions of consciousness.

A second reason why the neuron has been neglected is that, for virtually everyone except the neurophysiologist, the neuron is a rather simple, all-or-nothing on/off mechanism; on its own, it is only a bit in a byte in a bucketful of neurons packed into our heads! Considered as an on/off switch, the neuron is a decidedly uninteresting object – and, just like every boring nanocorner of a computer chip, the functional units of information-processing become in-

teresting only when put into large complex systems. The immediate jump to "complex neural systems" is therefore tempting. Moreover, it is certainly the case that, whatever the eventual resolution of specifically the subjectivity problem may be, most issues concerning the human psyche are concerned with multineuron thought processes. Scholarly debate at the level of brain mechanisms does not require and generally is greatly hampered by the introduction of tangential arguments about, for example, the meaning of electrostatic charge in physical theory, the nature of chemical bonding, the mechanisms of genetic replication, and so on. No one would deny that there are fascinating and important issues at those and other levels of material organization, but to make any progress at all it is essential to draw boundaries within which coherent arguments can be developed. So, do we really want to consider cellular matters that seem so far from the main topic?

A third reason why the neuron is not on center stage in the consciousness drama is the two-sided coin of career choices: neurophysiology is one of the "hard sciences" – more akin to chemistry than psychotherapy. As a consequence, the physiologist is, by nature and training, rather disinclined to speculate while there are still interesting measurements to make. Most psychologists, on the other hand, would rather not spend time in the laboratory when there is still some hardcore speculation to complete! Even if we have some sympathy for both sides of this cultural divide, the situation is such that the typical psychologist without training in cell biology has the idea that the neuron is, for most purposes, reasonably considered to be a simple logic gate, rather than a complex biological entity with a variety of properties characteristic of living systems. In contrast, the typical physiologist with a good understanding of the complexities of the neuron is interested in cellular mechanisms per se or their medical implications, but is unlikely to be in search of philosophical meaning at the cellular level.

Whatever reasons scholars have had for ignoring neuron physiology, the scandal of consciousness studies is the fact that the general case for considering neuronal "protophenomena" to underlie the "real" phenomena of consciousness has already been eloquently made by MacLennan (1996) [and reprinted in Shear's (1998) volume on the hard-problem]. And yet, despite the lucidity of MacLennan's argument, that essay has never been cited in the forest of verbiage on the topic of subjectivity and the entire line of argumentation concerning how to interpret the facts of neurophysiology has remained undeveloped.

The essence of MacLennan's (1996) argument for considering the existence of "protophenomena" that underlie subjectivity is as follows. To begin with, he notes the widely accepted view that:

the subjective is not reducible to the objective according to the usual pattern in science. Nevertheless it is important to strive for some kind of reduction of the more complex to the simpler or better understood. This can be accomplished by an analysis of the structure of consciousness, which allows a phenomenological subjective-to-subjective reduction that parallels a neurological objective-to-objective reduction. It is to be expected that progress in each reduction will facilitate the other, in turn.

Taking hints from what is known about the neurophysiology of cognition, he then argues that:

> The receptive field of a neuron in a topographic map refers to the stimuli to which it responds; for example, a neuron in a somatosensory map might respond to pressure on a particular patch of skin, or a neuron in the visual cortex to light on a particular patch of the retina. In such a case we can identify the subjective experience corresponding to activity in this neuron, namely, the feeling of pressure in that patch of skin, or the sensation of light on that patch of retina. I call such an element of experience a … *protophenomenon*. Further, we can see how, to a first approximation, the objective neurological processes corresponding to tactile or visual sensation can be reduced to a large number of receptive fields of this kind. This suggests phenomenological subjective-to-subjective reductions (phenomena to protophenomena) paralleling neurological objective-to-objective reductions (e.g., topographic maps to their neurons, and sensory surfaces of skin to receptive fields).

The importance of the temporal synchronization of protophenomena in producing conscious states is implied:

> Normally a change in the intensity of a single protophenomenon will not be perceived since it will not usually lead to a macroscopic change in conscious state… Normally the intensity of a large number of protophenomena must change *in a coherent way* for a phenomenon to appear in consciousness, that is, for there to be a macroscopic change in conscious state. On this view, the state of the phenomenal world, that is, the content of consciousness, is identical with the intensities of all the protophenomena. The appearance of coherent or stable phenomena can be identified with cohesive or coherent patterns of intensity among the protophenomena (emphasis added).

So, what precisely is the neuronal essence of the protophenomena? MacLennan (1996) remained agnostic – calling them "activity sites," probably involving the pre- or post-synaptic membrane potential – but it is in this general realm that further consideration is required.

B. The neuron's two modes of contact with the external world

Scandal or otherwise, it is clear that, with a few notable exceptions (Eccles 1990; Pribram 1991; MacLennan 1996), the properties of neurons have not been a major point of discussion in consciousness studies. Without necessarily being able to defend an explicit argument about the relevant neuronal wiring, the view that awareness is a property that emerges full-blown without a neuron-level analog from large brain systems is tacitly held by many. The unfortunate consequence of proceeding directly to the "complex network" level without first getting the "unit processes" into proper focus has meant that problems that are perhaps easily resolved at the cellular level are left to fester into nightmares of entangled issues where cellular, organismic and philosophical problems are jumbled together. Let us therefore return once more to the single-cell level to see if a self-consistent view of the "big" problems of consciousness studies can then be achieved.

As already discussed in Chapter 6, at the cellular level, the essence of "sensitivity" is the ability of the living cell to engage in a regulated "openness" to its biochemical environment. This takes the form of the controlled exchange of substances across a semipermeable membrane; the import and export of neurotransmitters and the sudden, rapid exchange of ions during the action potential of the neuron. For the living cell, the partial opening of its protective membrane is a "dangerous" form of activity – literally allowing the inflow of foreign substances and the outflow of the molecular constituents of the cellular self, but such permeability is the cell's most fundamental means of interacting with its local environment. A cell that is hermetically sealed will of course eventually die for a lack of energy resources, but it immediately loses its sensitivity if its membrane is no longer permeable – if it can no longer directly "know" what the electrotonic state of its biochemical environment is. By this definition, all living creatures (plant, animal and bacterial, but excluding the metabolically-inert virus) have some form of (cellular level) sensitivity as a consequence of their semipermeable membranes. In essence, "sensing" the environment is a necessary part of maintaining the cellular "self" within certain biochemical parameters. By continually sampling the biochemistry external to the cellular self, the cell can make the internal adjustments needed to keep the cellular machinery in working order. But the commonality among animals, plants and bacteria ends there. In addition to cellular-level self/other interactions, animals with nervous systems are capable of coordinated neuronal responses to environmental stimuli and, as a consequence, have the ability to know and respond to the external world more directly and rapidly. They have a capacity for greatly

amplified cellular sensitivity, i.e., synchronized bursts of receptivity and reactivity to their environments through the coordination of many cells working in concert.

I suggest that such amplified sensitivity, typical of the more highly evolved animal species, can be considered as the smallest unit of "awareness." It is, in essence, nothing more than neuronal "openness" – particularly that during the action potential, but it takes on qualities that are not seen at the single-cell level due to the temporal coordination of the action potentials of many neurons. In other words, awareness is a product of animal nervous systems – the result of the synchronous firing of neurons engaged in sensory, motor and/or cognitive functions, but the functional unit is the cellular sensitivity exhibited during the action potential. The magnitude of animal awareness is therefore a function of the number of neurons involved in synchronous activity, their level of activity and the precise timing of their firing relative to one another. As a consequence, larger brains and waking, metabolically active brains will in general have the potential for greater awareness – the simultaneous, momentary "feeling" of the biochemical surroundings of many nerve cells. The kind of feeling that an animal organism will experience is necessarily a function of the set of simultaneously active neurons: Just as individual neurons do not perform "cognition," but rather execute simple excitatory or inhibitory response functions, the summation of which is the information-processing of whole brains known as "cognition," an individual neuron does not "feel" anything beyond its own surrounding biochemistry, but, when acting in concert with many neurons, the summation of the sensations of many neurons produces a total "feeling" for the current state of the "external world" – a feeling of awareness that depends directly on what subset of neurons within the nervous system are synchronized. It is specifically the simultaneity of a large number of cells opening up to the environment that produces a "feeling" in the organism as a whole, and it is the specific locations and patterns of connectivity that will determine the psychological content of the feeling. With regard to both sensitivity and information-processing, the single-neuron event is well understood, but the related whole-brain events of awareness and cognition can be understood only by extrapolation.

The lowly paramecium provides convincing evidence attesting to the fact that "sensitivity" and "reactivity" do not require a nervous system of any kind (Figure 7-1). If that single-cell organism is prodded mechanically at its rostral end, the cell membrane reacts with pore opening, the in-flow of ions, a change in intracellular pH, and consequently a reversal in the flagellating hairs on the cellular surface: the paramecium backs up. If it is prodded at its caudal end, a

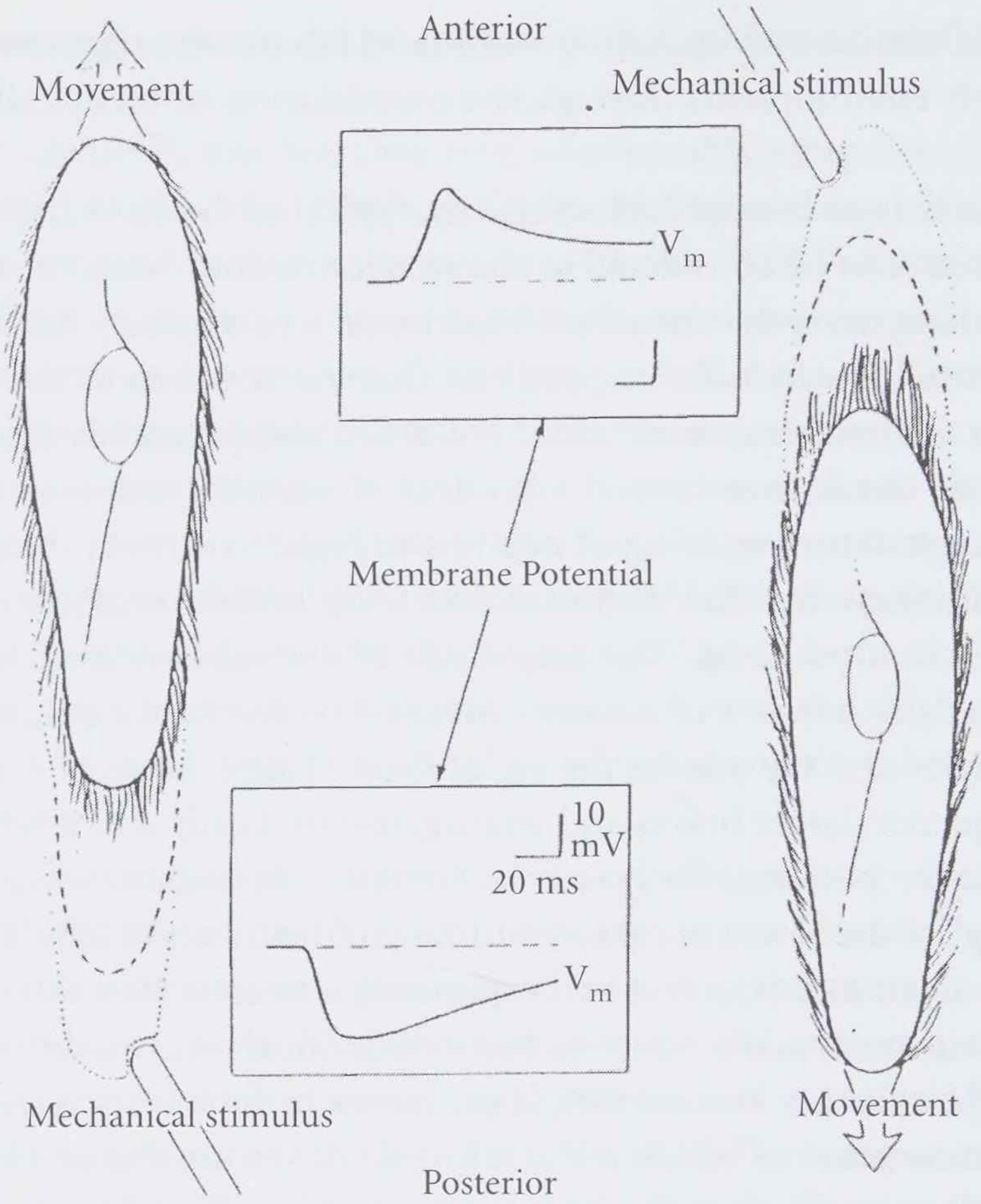


Figure 7-1. Although entirely lacking a nervous system, the single-cell paramecium can respond with appropriate behavior following mechanical stimulation (after Ogura & Machemer 1980). The cellular mechanisms are closely related to those of neurons – involving the controlled flow of ions across a semipermeable membrane.

similar membrane mechanism is elicited, pushing the organism ahead. Grand words such as awareness and consciousness should be avoided here, but clearly the paramecium is able to make meaningful responses to environmental stimuli simply as a consequence of the mechanistic effects elicited by changes in intracellular ion concentrations.

What then is different about animals with nervous systems?

C. The temporal coordination of neuronal firing

Synchronization and cognitive binding

Throughout the 1990s, many arguments were presented concerning the importance of temporal correlations in neuronal firing (synchronization) for the "binding" together of object features during cognition. The basic idea is that the processing of different types of information simultaneously at various spatial locations on the cerebral cortex can be unified temporally through the coordinated timing of neuronal action potentials. When neurons fire more-or-less in unison, the information implied by the firing of those neurons is processed simultaneously, and perceived as being related. When, however, neuronal firing is temporally unrelated, i.e., desynchronized, the lack of simultaneity implies that the information of the unsynchronized neurons is perceived as being concerned with unrelated stimuli. The theoretical importance of synchronization lies in the fact that, by allowing for a type of cognitive integration *without* requiring a brain region to which all neuronal impulses must flow, the synchronization of neuronal firing may solve one of the oldest enigmas in the philosophy of mind, that is, the unity of thought processes. The unity of mind is a problem that has, paradoxically, been aggravated by modern neuroscience, because there is strong evidence indicating that different features of sensory stimuli are processed at diverse locations in the brain. The perceptual mechanisms and locations of the taste of an apple, its aroma, its redness, its texture and its roundness are to some extent understood, but they occur at brain sites so distant from one another that topographical unification is unlikely. If, however, these diverse sites can be temporally unified, then the spatial diversity of feature processing does not imply the disunity of mind.

The cognitive binding issue is most frequently discussed in terms of the unification of diverse sensory attributes, but the same synchronization mechanism may be at work in all forms of higher cognition. Shastri and Ajjanagadde (1993) have illustrated the binding as shown in Figures 7-2 and 7-3. Given a setting in which several objects, agents and actions are involved, it is imperative for the organism not only to recognize the individual objects/agents/actions, but also to perceive their causal relationships – often, in the human case, with the aid of symbolic language. As shown in Figure 7-2, the cognitive problem that the brain must solve is to determine which of several objects/agents/actions are causally involved in the (for some reason) salient event of "giving the book" that is at the center of attention. Any suitable linguistic grammar might be used to represent the observed scene, but the overwhelm-

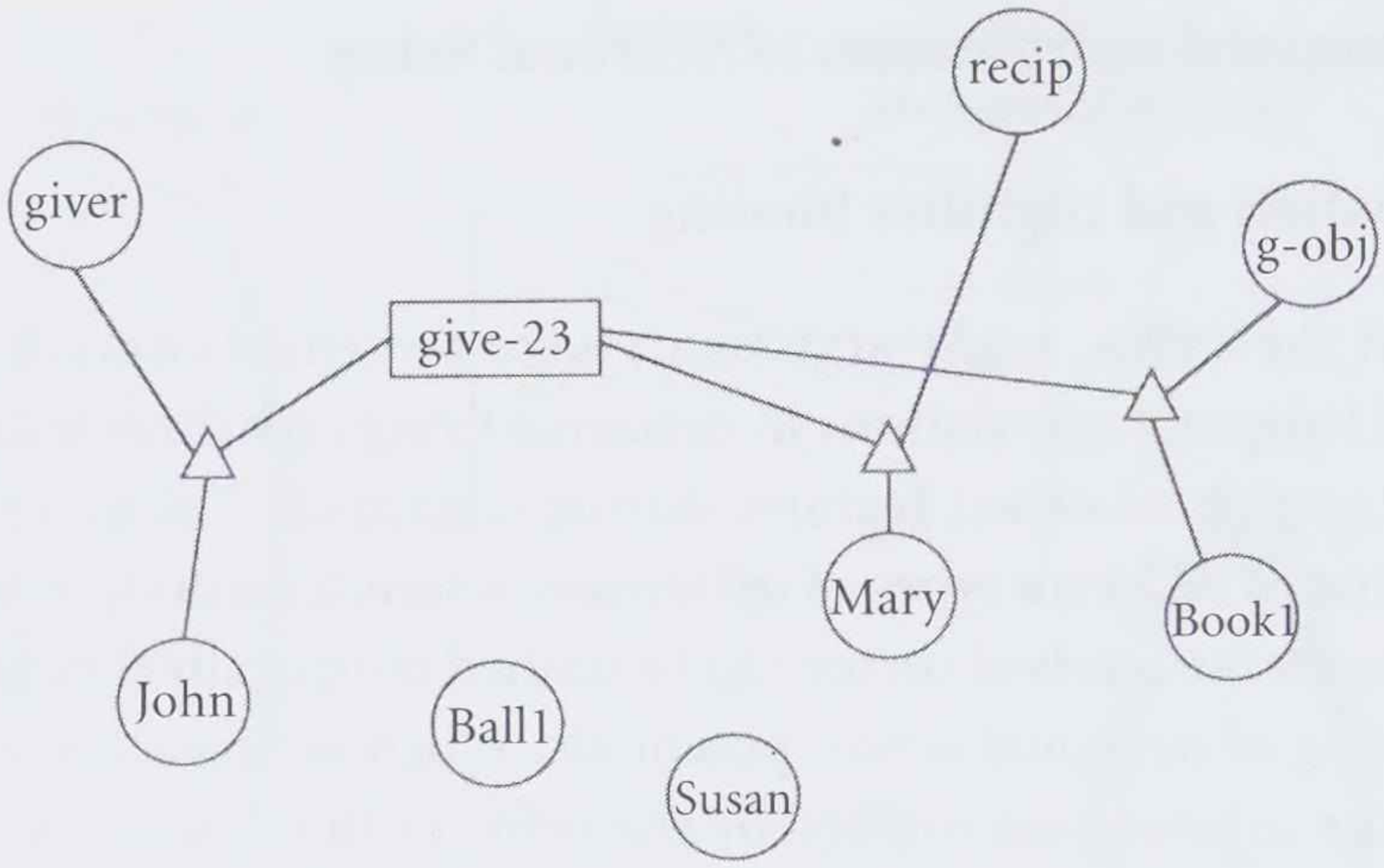


Figure 7-2. A typical "parsing" of an event involving several agents and objects, and excluding others. The brain's problem is how to sort out the causal relations in "real time" (Shastri & Ajjanagadde 1993).

ingly complex problem that the brain faces is not simply representing the event, but representing the event amidst a barrage of other stimuli that are of limited relevance.

The synchronization model of Shastri and Ajjanagadde indicates that the hierarchical segmentation of a complex event can be achieved without brute force logical deduction concerning the relationships in the event simply through the appropriate timing of activity of neuronal modules dealing with the parts and relations within the scene (Figure 7-3).

Many unanswered questions about the neuronal mechanisms and cognitive implications of synchronization remain, but (i) progress has been made in defining the role of synchronization in arousal and attention (Abeles 1982; von der Marlsburg 1981; Crick & Koch 1990; Crick 1994; Singer & Gray 1995; Engel & Singer 2001), (ii) the relevant frequencies of oscillation have been studied in various animal species (Engel et al. 1999; Engel & Singer 2001; Gray et al. 1989; Roelfsema et al. 1996), and (iii) possible neuronal mechanisms have been explored in artificial neural nets (Shastri & Ajjanagadde 1993; Schillen & Konig 1994; Arnoldi & Brauer 1996). Although interesting ideas about the temporal aspect of neuronal activity have been discussed for several decades (e.g., Abeles 1982), the recent development of multiple single-cell recording techniques and of detailed computer simulations has, at long last, brought the issue of the timing of neuronal activity into mainstream neuroscience. It is too early to draw firm conclusions, but the concept of synchronization is perhaps the single most important addition to the basic findings of neurophysiology that were already

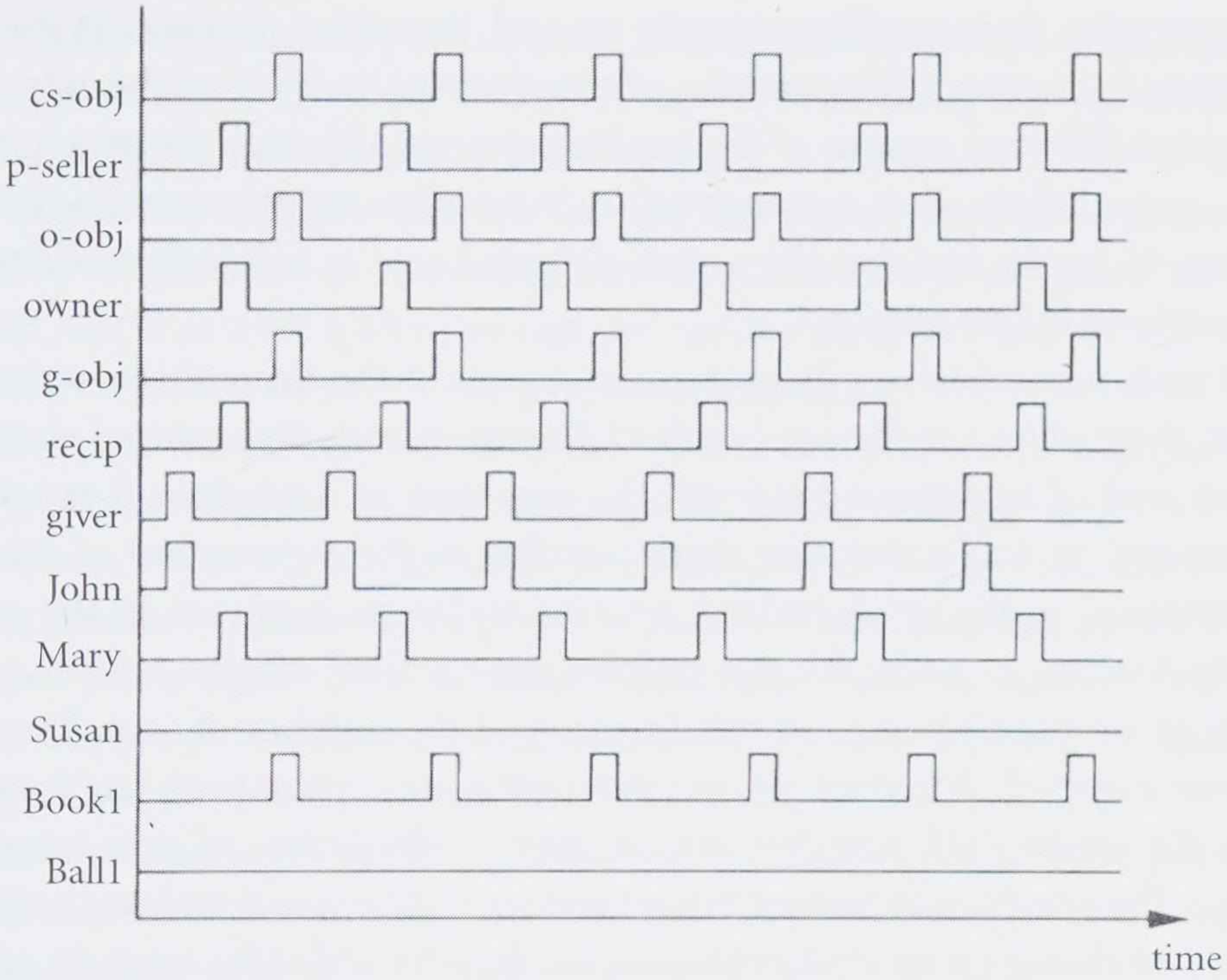


Figure 7-3. Hierarchical synchronization is hypothesized to bind the relevant pieces together in a coherent manner (Shastri & Ajjanagadde 1993).

established by the middle of the 20th Century. Nobel Prizes may well be deserved for the empirical demonstration of the "cognitive binding" argument alone, but before the celebrations begin, let us consider what synchronization may mean for the issue of subjectivity.

Synchronization and awareness

The phenomenon of synchronization that has been hypothesized to play an important role in cognitive binding is essentially the temporal coordination of multiple action potentials, such that many neurons simultaneously allow the inflow of ions and simultaneously re-equilibrate their ionic concentrations. While the emphasis with regard to the cognitive binding due to synchronization is normally placed on the coordinated information-processing resulting from the simultaneity of synaptic events (i.e., cognition), it is in fact the temporal coordination of *the action potentials* that is experimentally measured. That is to say, when two or more neurons are synchronized, the brief opening of ion channels in their axonal membranes occurs simultaneously during the action potentials. The cells quite literally "sense" their respective external worlds at

the same time: their membranes open up and, through a moment of chemical diffusion, the neurons take samples of the electrostatic state of the extracellular fluid in different regions of the nervous system. Although the synchrony of action potentials is what is measured, approximate synaptic synchrony is also guaranteed by the fact that the action potential acts as a deterministic trigger to neurotransmitter release.

If each neuron senses its own environment at the time of the action potential, then when a sufficient number of neurons are synchronized above the chance level of background activity, the organism as a whole will experience a characteristic subjective state corresponding to the summation of the local "sensitivities" of the synchronized neurons at diverse locations in the central nervous system. As such, this is a simple "mass action" effect – with the magnitude of awareness being a direct function of the number of (synchronized) neurons involved. For example, on perceiving an approaching truck rapidly filling the visual field, together with auditory stimulation of ever-increasing volume, the synchronization of many neurons in the visual and auditory cortices (recording the sensory stimulus itself), the amygdala (the emotion of fear) and the hypothalamus (the autonomic response) will summate to a predominant feeling appropriate to the situation. In this hypothesis, quite aside from the unity or coherency of the related *cognition* (conclusions concerning the nature of the stimulus and decisions concerning the proper course of action), the non-cognitive "feeling" – the subjective impression that one is in direct contact with an external reality – is *not* a consequence of "subjective state circuitry" or some other unknown, unprecedented fiat of neural connectivity, but simply the summation of the simultaneous receptivity – literally, ion influx – of many neurons. Why would the simultaneous firing of certain neurons in visual and auditory cortex, amygdala and hypothalamus lead to the distinct feeling of terror related to the impending collision with a truck? Because, on the basis of all previous experience, the currently-activated sensory neurons have a set of established connections that imply certain inescapable associations. An object that fills the visual field at such acceleration and a crescendo of noise over the same time interval indicate rapid approach and violent contact. The logical deduction that an AI program would lead to might be identical to that in the living brain: "Impending collision! Threat to biological integrity! Evasive action recommended!" But the subjective feeling of the living system will not be a function of the algorithmic encoding of the state of urgency of the logical conclusions, but a more quantitative issue of numbers of neurons involved. In other words, the strength of the feeling will necessarily be a function of the number of activated, synchronized neurons.

From an evolutionary perspective it is likely that neuronal conduction (i.e., action potentials) occurs *solely* for the purpose of allowing the neuron to perform synaptic communications and ultimately to have some influence on organism behavior and survival. The fact that the action potential is instrumental in allowing neuron-to-neuron communications should not, however, cloud the fact that it is itself a form of neuron-environment interaction that differs from the cognitive effects of synaptic communication. The flow of the nerve impulse from the cell body toward the synapse involves a direct, physical exchange between the cell and its environment. Unlike the flow of electrons in a copper wire, where no contact with the external world is implied – the biological mechanism of transmission *prior* to synaptic communication entails the exchange of materials between the neuron and its biochemical surroundings and, as a consequence, there is both neuronal information-processing (~cognition) and neuronal sampling of the extracellular world (~awareness).

Synchronization of the firing of neurons produces a coordinated pattern of the inflow of ions to many neurons at diverse locations throughout the nervous system. Clearly, there are two distinct levels at which the issue of subjectivity must be considered. First of all, there is the single-cell phenomenon of the direct influx of extracellular ions by physicochemical diffusion. This is the basic mechanism by which a living cell "re-evaluates" its own electrotonic state in relation to that of the environment. At the next level, there is the coordination of such ion-inflow by many neurons. The synchronization of action potentials implies the simultaneity of the "feeling" that is analogous to, and inherently linked with the cognitive binding of synchronized neurons. The difference between the cognitive binding of different features into a unified percept and the coordinated "feeling" of those same neurons lies only in what aspect of the neuronal process we focus on: the cognition resulting from the complex summation of excitatory and inhibitory synaptic activity or the feeling resulting from the opening of diverse neurons to their local environments. In this view, subjective awareness is "unitary" in the same sense that a volume of gas will have a single temperature, despite the fact that the molecular constituents of the gas will each be in distinct microstates, none of which suffices to explain the temperature of the gas as a whole. In contrast, the unity or disunity of cognition is a consequence of the progression of serial information-processing toward a cognitive result that drives behavior.

Singer and colleagues have emphasized the importance of synchronization for many years, but their focus has been on cognitive binding and its role in cognitive unity. More recently, they have argued that synchronization also allows for the unity of subjective experience (Engel et al. 1999), but simple decla-

ration that cognitive binding equals subjectivity does not suffice to explain why there is an "internal perspective" that we know subjectively (!) to be distinct from the processes of cognition. My suggestion is quite simply that the subjectivity occurring during cognitive binding emerges from a cellular mechanism that is distinct from synaptic processes. Once the neuron itself is understood to have a cellular-level primitive capability of "experiencing the environment" – a process which is predominantly due to ion inflow during the action potential – then it is possible to understand why many neurons acting in concert might produce an organism-level feeling that is also felt to be "experiencing the environment" and typically referred to as subjective awareness.

The unhappy inference is, however, that the organism-level feeling of "experiencing the wider world" – specifically, the feeling of direct connectedness with an external reality is an illusion.[1] The fact that we all have such feelings of varying magnitude and frequency is not due to any form of "deep" empathic understanding of an external reality, but more mundanely to the fact that our neurons are sensitive living systems. At the single-cell level, there is true receptivity and local material exchanges with the extracellular world, but at the organism-level there is only the summation of many "locally-sensitive" cells. Here, we would perhaps like to believe that the whole is more than the sum of the parts, but – in light of what is known about the biochemical nature of neuronal "sensitivity" – we are obliged to conclude that the whole-brain "feeling" is nothing more than multineuronal openness to the intracranial cerebrospinal soup! Whether we are observing the sunset, actively falling in love, or feeling the "oneness of the cosmos," the simultaneous receptivity of neuronal action potentials is the causal factor underlying the qualia – the "feelingness" of perception. Unfortunately, what each and every one of our neurons actually experiences is a very local bath of biochemicals – nothing more. It knows nothing about the "environment" in the sense that the organism *cognitively* understands it. Because of each individual neuron's anatomical location within the brain, the sensitivity of each neuron is locked into certain kinds of cognitive events – visual, auditory, motor, neurosecretory, etc. – and its individual sensitivity is similarly local. Individually, the activity of any one neuron is associated, on the basis of previous experience, with certain kinds of perceptions and behaviors, and its firing will synchronize with some probability to the firing of other neurons, but its experience is strictly cellular. To return to the analogy with social systems, the individual voter might be said to have an experience of "protodemocracy" in pulling the lever in a voting booth, but the true meaning of the word "democracy" is concerned with a society-level phenomenon that the individual *cannot* directly experience. By definition, "democracy" is sociopolitical.

Without the participation of individual voters, democratic phenomena cannot occur, but the lone voter is but a small piece in the puzzle, a small cog in a huge machine – and *cannot* internalize the social phenomenon of democracy except as a highly-abstracted symbolic concept. In the same way, it is an incorrect confusion of ideas related to different levels of material organization to assert that the individual neuron is "conscious." It is *not* conscious: "consciousness" is the term used to describe a property of complex nervous systems, but the sensitivity of the individual neuron is nonetheless the essential protophenomenon – the unit-function that underlies the whole-brain phenomenon.

For the information-processing of many-neuron systems, the timing of the processing is not crucial. Any effects that occur within the time-limits of the summation of dendritic inputs will produce similar informational results. However, for the summation of the "feeling" of many neurons, the timing is crucial. The influx of ions during the action potential must be timed such that neurons open up to their respective environments simultaneously, so that the organism as a whole has a single, dominant feeling. If temporally spread out, the unity of the feeling would be lost and the motivation for action of the organism as a whole would be dissipated in the separate activities of uncoordinated neuronal events. To return to the looming truck example, it is imperative that the limbic (fear) component of the response to the auditory/visual input is activated together with the external stimulus in order that the appropriate behavioral response will be taken. If the visual information were processed at a leisurely pace without limbic/brainstem activation, cognitive conclusions concerning the impending collision might be unchanged, but clearly the cognition, planning and decision-making would be too late if not linked to the feeling of immediate terror.

In brief, for the strength of the affective result, temporal coordination of neuronal firing is important. In contrast, the simultaneity of information-processing is not crucial for cognition. The roundness, shininess and redness of the apple will be cognitively bound due to the driving sensory stimulus, but cognitive implications can occur at various speeds "off line." For synaptic events, the crucial factor is only whether or not the information-processing results in an effect on effector mechanisms – other neurons, striate or smooth muscle cells, or endocrine tissue. If an effect is obtained, the goal-directed information-processing can be considered a success, whereas there is no goal-directed effect of the ion-influx of individual neurons: there is only affect – the diffuse feeling of some number of neurons sampling their diverse local environments.

Subjectivity

The philosophical argument that this focus on neurophysiology implies is therefore that the momentary opening of the cell membrane at the time of the action potential is the single-cell protophenomenon (MacLennan 1996) underlying "subjectivity" – literally, the opening up of the cell to the surrounding biochemical solution and a brief, controlled breakdown of the barrier between cellular "self" and the external world. The synchronization of the action potentials of many neurons produces a pattern of "transparency" of the nervous system as a whole – arguably a simple by-product of the coordination of neuronal firing that is *needed* for feature "binding" in cognition. The normal ebb-and-flow in the strength of subjective feeling is thus real, and a direct consequence of the variable number of neurons participating in synchronous firing. The only thing "illusory" about subjective awareness is that we tend to mistakenly think that we feel the "outer world" directly – a world that, in fact, exists millimeters or miles outside the neuronal membranes. In contrast, it is *not* an illusion that individual nerve cells are momentarily in direct contact with their extracellular environments, with each cell briefly sampling the electrostatic state of the surrounding ion solution. The origins of the so-called hard-problem are thus apparent: our neurons are individually in direct contact with their cellular-level "external worlds" and produce a multineuron feeling of subjectivity that is *real*. However, although the direct "feeling of the environment" at the cellular level is real, brain-level subjectivity does not have a corresponding "direct experience" of an extracorporeal external world, except in the quite uninspiring sense of many neurons subjectively experiencing the intracranial extracellular soup at the same time. Our "feelings" for external events and our empathy for people whose experiences we cannot literally share are as strong as our synchronized neurons are numerous, but our psychological internalization of external phenomena is entirely *cognitive*.

Already at the cellular level, there is an explanatory gap between computable cognition and incomputable awareness – between quantifiable synaptic communications that lead to behavior and diffuse electrotonic effects across semipermeable membranes that result in an unquantifiable "feeling" with no direct behavioral implications. The gap is real, but not inexplicable, and the problem is perhaps not so hard, if it can be understood as a direct consequence of the two modes of neuronal interaction with the "external," biochemical world.

Synaptic transmission is widely believed to be the functional unit of cognition: in the act of neurotransmitter release and the exertion of inhibitory

or excitatory effects on other neurons, a neuron participates in small-network logical functions, the sum of which quite simply is cognition, but a single neuron does not "think" anything beyond its own input/output response function. Similarly, at the moment of the action potential, the neuron exchanges ions with its environment and participates in the organism's overall feeling of the external world, but a single neuron is not "aware" of anything other than its own relative membrane transparency. Of course, the action potential is the causal trigger leading to neurotransmitter release at the synapse, but ion flow and neurotransmitter release are distinct phenomena and contribute to two different types of organism-level psychology: in addition to whatever small role any lone neuron may play in the cognition of the whole organism, it also makes a small contribution to the overall subjective feeling of openness of the organism through contact with its extracellular environment.

What this synchronization-based interpretation of subjectivity achieves is the separation of the core topics of awareness and cognition in terms of proto-phenomena *at the neuronal level*. Synaptic communications lead ultimately to the cognition of multicellular nervous systems, and the transmembrane flow of ions leads ultimately to the direct feeling of an "external" world, i.e., subjective awareness. Questions about the relationship between animal awareness and human consciousness, and questions about animal cognition and human, symbolic cogitation are not addressed by this synchronization argument, but by identifying the relevant cellular mechanisms the construction of strictly *neuronal* theories of both awareness and cognition may be possible – and obviate the need for introducing other-worldly philosophical or quantum mechanical postulates.

D. Cellular-level and organism-level definitions

The term "sensitivity" is used here simply to indicate the cellular-level phenomenon that results from the transmembrane receptivity to electrostatic changes in the external world, not any higher-level psychology. In other words, the physiological mechanism of membrane "openness" results in cellular-level sensitivity to the biochemical environment – and this is the protophenomenon underlying higher (organism-level) psychological phenomena. By this definition, single cells have "sensitivity," whereas all living creatures (comprised of many living cells) have some form of "proto-awareness" as a consequence of the collective sensitivities made possible by many living cells. The proto-awareness of most plants, however, is the simple summation of individual cell sensitivity,

with no coordination among them. Animal nervous systems, in contrast, are capable of coordinated neuronal responses to environmental stimuli and have the capacity for amplification of proto-awareness, i.e., synchronized bursts of receptivity and reactivity to their environments that might be referred to as awareness. The magnitude of awareness is thus a function of the number of neurons simultaneously activated. (An interesting implication of this hypothesis, suggested by MacLennan, is that larger neurons "feel" more strongly than smaller neurons, as a consequence of the greater ion traffic across a wider expanse of neuronal membrane. The reduced ion traffic caused by the myelin sheath in myelinated neurons would reverse this effect, but in any case a correlation between the volume of ion-flow and the strength of subjective feeling is predicted.)

The neuron learns about its biochemical situation by periodically opening its membrane and allowing the flow of ions into and out of the cell, but "feeling" and "thinking" are organism-level phenomena that the lone neuron does not experience. There are necessarily "proto-feeling" and "proto-thinking" phenomena at the cellular level, but only the protophenomena can be rigorously defined at the cellular level; the resultant organism-level phenomena require a different terminology. To return to the analogy with social systems, the protophenomenon of democratic control in society is voting, but the individual voter has no experience of "democratic fairness" while in the voting booth. There, he typically experiences only a yes/no choice for or against the status quo, while the "fairness" and "majority rule" characteristic of democratic decision-making arises only in the summation of the effects of many voters acting together on voting day. Although clearly related to the individual voter's activity itself, democratic justice is a society-level phenomenon that cannot be adequately described solely at the level of individual voter psychology/behavior.

At the single-cell level, the term "sensitivity" suffices to indicate the ability to react to environmental stimuli and the paramecium is a good example of such biological sensitivity. Even when the organism has a nervous system, reactivity that is undertaken solely in response to one type of external irritation is also normally described as "sensitivity." For example, not only do most insects appear to respond directly and "blindly" to direct stimulation, we commonly use the sensitivity label for animal and human responses in the context of reactions to a finite stimulus: a dog that is sensitive to a high-pitched whistle, an ex-smoker sensitive to tobacco smoke, a boy sensitive to poison ivy, a girl sensitive to remarks about her curly hair, and so on. In such contexts, we would

never say "aware" or "conscious" of the stimuli; it is a direct stimulus-response pair that indicates sensitivity.

In contrast, the term "awareness" is typically used in the context of meaning both sensitive to a specific stimulus and understanding the larger context of that stimulus. The dog that is "staying" before pouncing on a morsel of food is certainly sensitive to the aroma of the food, but is also aware of the larger context from training concerning the trainer's demands. We have no hesitancy in declaring that the dog is aware of the need *not* to react to the stimulus of food. It is aware of the larger context – not simply sensitive to the stimulus.

But is the dog conscious? We might sometimes be tempted to give it the benefit of the doubt, but the normal test of consciousness that we apply to *people* is whether they can spell out the consequences of their actions. The dog is certainly aware of the trainer's presence and, on the basis of previous experience, wary of the possible smack on the nose that uncalled for advance toward the food will bring, but we remain ignorant of (and skeptical regarding) the depth of the awareness in so far as the dog cannot indicate that it has concern for the good condition of its nose if it should seize the food. A chimpanzee, on the other hand, might well indicate by pointing that it does not wish physical punishment to a certain body part – and thereby communicate to us its understanding of consequences. Even without a verbal report, we might then be willing to use the "conscious" label. If the chimp is indeed able to communicate that it is not only aware of the fact that a wrong action will lead to something unfavorable, but capable of indicating what that unfavorable outcome might actually be, then the "awareness" of danger is supplemented by a "consciousness" of lawful cause-and-effect.

In the case of training a child, we would verbally inquire to the child about what would happen if the conditions set up by the adult are not met. We would normally request a verbal delineation of the possible unfavorable consequences of taking another cookie out of the cookie jar before we would be satisfied that the child is conscious of the full meaning of the situation. The child is not simply "aware" and "wary" of the fearsome power of the adult: he is certainly aware of some danger if behavior is incorrect, but, in addition to animal awareness, if the child can verbally explain consequences, we are then normally willing to say that he is conscious of his own actions.

By these definitions, sensitivity refers to the receptivity of individual cells. Without a coordinating mechanism, the quantitative increase in sensitivity due to an increase in the number of cells does not lead to higher psychological properties; the cell-conglomerate may be "very sensitive," but it cannot be aware unless its many-cells act together – are synchronized in relation to the coordi-

nated activity of other neurons. By its very nature, the focus of sensitivity needs to be contrasted with a larger context. In other words, "foreground" objects are in the foreground *because* they stand out from a background; it is the presence of the background that makes the foreground more than a lone stimulus. In this view, the reptilian brain that can respond with the coordinated activity of a large number of motor neurons arguably lacks "awareness" in so far as the motor response cannot be suppressed. But the mammal capable of suppressing the motor response arguably has the neural mechanisms necessary for awareness – an ability to see the forest while responding to the trees. Finally, human consciousness denotes a capacity not only to suppress motor responses – and thus take the delaying action that can allow one to plan and make decisions, but also to simulate the decision mechanisms before or after the event. In order for a psychological process to be simulated, however, there must be mechanisms available for re-presenting the sensory and motor events. In other words, language or some other mechanism for symbolic representation is a prerequisite to consciousness and is thus a mental capacity unique to human beings.

It is a curious and wonderful fact that the French have one word, "conscience," to mean both consciousness and conscience. If the meaning of consciousness, as distinct from awareness and sensitivity, lies in the ability to understand consequences, then the psychological link between consciousness and conscience is clear. In English, conscience means to have a retrospective ability to trace back consequences once a decision has already been made (and action taken). A "guilty conscience" therefore implies a sense that the current situation can be traced back to a specific decision made by the individual. It was a conscious, remembered decision to set out in one direction, not the other, that led to the current state. To have consciousness means to have the potential for conscience – and that is why we are more forgiving of naughty young children and animals – who typically show no remorse suggestive of conscience – than we are of equally remorseless criminals that demonstrate clever forethought and scheming purposeful decisions – i.e., consciousness afore, but no retrospective conscience. A certain class of criminal has forward-thinking consciousness in abundance, but does not demonstrate even rudimentary conscience, while young children and animals (and arguably a certain class of adult criminal) simply lack the purposeful planning – and understanding of cause-and-effect – of consciousness, and necessarily also lack conscience – a sense of remorse. Without consciousness, there can be no conscience.

E. Three explanatory gaps!

Having outlined what I believe to be a sensible distinction between (synaptic) information-processing and (ion-influx) sensitivity at the neuronal level, and how it relates to whole-organism issues, it is now possible to return to the problem of the explanatory gap, outlined in Chapter 6. I maintain that there are three distinct "explanatory gaps" in modern neuroscience, although only one has been considered to be a major philosophical problem. First of all, there is the gap between "information-processing" at the single-neuron level and "cognition" at the whole-brain level. Secondly, there is a gap between what is understood concerning how a neuron can directly "sense" the electrostatic charge of the extracellular environment and how organisms have an "awareness" of the external world at the whole-brain level. Thirdly, there is a gap between whole-brain cognition and whole-brain awareness. The third of these gaps is commonly referred to as "the" explanatory gap – and refers to the seemingly unbridgeable distance between the well-understood, objective, scientific, material mechanisms underlying neuronal information-processing and the subjectivity of organism-level feeling. This "gap" is between seemingly orthogonal dimensions and has been the focus of much philosophical discussion, but I suggest that the gap can be bridged by an appropriate understanding of the relationship between single-cell and whole-organism phenomena.

1. The gap in understanding cognition

The way in which individual neurons can "process information" is well-understood and can be implemented in quasi-intelligent computer systems, but the jump to "real cognition" and human-level thinking requires an extrapolation from a small to a large system. Such extrapolation indeed seems reasonable to most neuroscientists and most would agree in principle that cognitive functions are "just" the summation of very simple logical functions made possible by synaptic connections. In fact, however, there is a huge gap between, on the one hand, our understanding of how small systems comprised of simple "non-cognitive units" with weighted inhibitory and excitatory connections can perform logical operations and how, on the other hand, we – as highly-complex, multimillion-neuron systems – actually perform operations that we normally refer to as "cognition".

The explanatory gap in cognition is not often addressed, but is arguably as real – and of the same general character – as other gaps that require a "leap of faith" from well-understood unit-mechanisms to whole systems. While there

may be some reluctance to make the bold assertion that cognition is "nothing but" coordinated synaptic activity, that is the general conclusion to which modern science is headed – i.e., Crick's "astonishing hypothesis." Rigorous empirical demonstration of such a view will remain difficult in any nervous system with more than a few dozen neurons, but the demonstration of complex logical functions using simplified neuron-like logic gates in computers has already proven the argument. If neurons can function as logic gates, and many logic gates in computers can exhibit cognition, then it is not an unreasonable leap of faith to believe that complex brains are cognitive because of neuronal *synaptic* activity.

2. The gap in understanding consciousness

The above argument concerning *cognition* via synaptic activity is entirely straight-forward. For cognition, there is an end-product of the cognitive activity, normally followed by behavior that can be measured. But, for the question of awareness, any "resultant behavior" is, by definition, not we are looking for. The argument regarding *awareness* is simply that there is an analogous extrapolation to be considered with regard to the receptivity of individual neurons and the apparent receptivity of multi-million neuron systems. There is no doubt that, principally through the mechanisms of the action potential, the neuron briefly opens its membrane to the free diffusion of ions – leading to equilibration of the ion concentrations inside and outside the neuronal membrane. Quite literally, the neuron feels its environment by letting some of the external world pour in! When that process is coordinated among many neurons simultaneously, there is inevitably a coordinated feeling of experiencing many extracellular environments by many neurons. In other words, the simultaneous sensitivities of many neurons sum to a grand total that we, at the whole-brain level, experience as animal awareness. My contention is that the single-neuron mechanism of transmembrane "sensitivity" that we understand with the full certainty of modern science can be extrapolated to the whole-brain level with some confidence.

3. The gap between awareness and cognition

Finally, there is the "classical" explanatory gap, as defined by Levine (1983, 2000). This gap is the discrepancy between a "complete" scientific description of the material events that underlie all mental events and the psychological description of what awareness feels like. As accurate and complete as

both descriptions may be, the gap is acknowledgement of our uncertainty with regard to the relationship between the two domains. The phenomena themselves and the vocabulary used to describe them appear to be incommensurate, but I maintain that the main problem lies in trying to resolve the problem at the "whole-brain" level (Figure 7-4). At the cellular level, the gap (between synaptic communications and the action potential) is understood and is simply not a mystery. The findings of modern-day neurophysiology already provide plausible links between well-understood cellular phenomena and less well-understood, high-level "cognition" and "awareness" of the organism as a whole. At the organismic level, we cannot directly experience the internal "feeling" of the sensitivity of individual cells, but we have more insight into cellular sensitivity as a primitive mechanism of opening up to the local environment than we do to other hypothetical mechanisms – quantum mechanical, biochemical or molecular processes – for which we have only a rather abstract intellectual understanding, and no hint of any possible "internal" aspect.

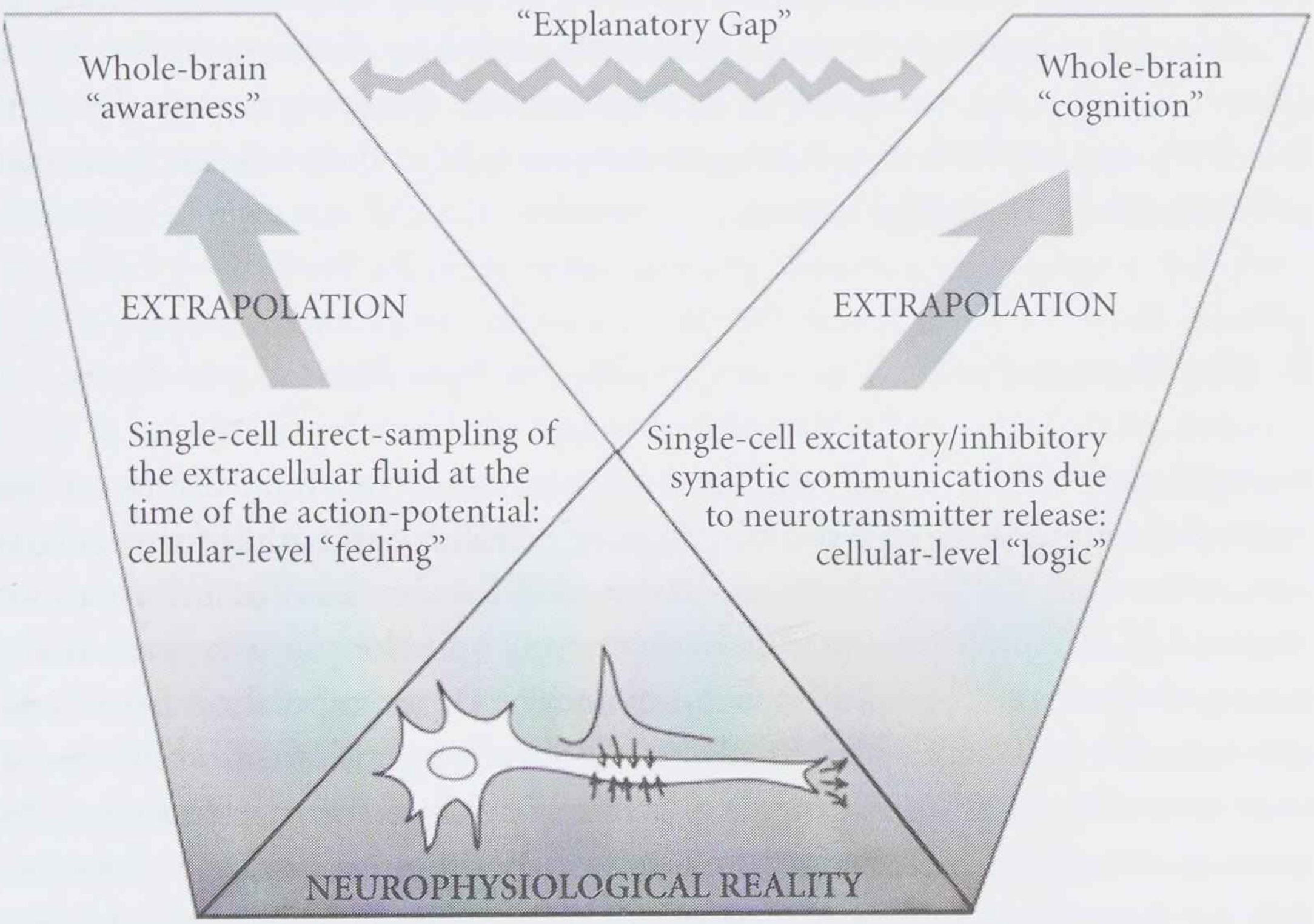


Figure 7-4. The traditional "explanatory gap" is the seemingly unbridgeable distance between the whole-brain phenomena of awareness and the whole-brain phenomena of cognition. The origins of the gap can be understood, however, by consideration of the single-neuron functions that underlie both awareness and cognition.

The so-called "explanatory gap" has been considered to be one of the big mysteries among the many puzzles of consciousness. At the whole-brain level, it is unresolvable, but, I maintain that it can be understood if we return to the single-neuron level and give consideration to how a neuron interacts with its extracellular world. Related "gaps" between the single-cell and whole-organism levels exists for both cognition and awareness, but those gaps are real only in the sense that extrapolation from well-understood microscopic phenomena to the sum total of many such effects occurring simultaneously in large systems *always* produces a qualitative difference between the mechanistic cause-and-effect at the microscopic level and the stochastic effect at the macroscopic level. The reality of such a gap does not mean, however, that we do not understand the macroscopic level, but the simplicity and one-to-one causation at the microscopic level is replaced by a stochastic causality that is just as real, but requires a different form of understanding. The qualitative change requires a new terminology when referring to the two levels of material organization, but the conceptual links are entirely clear. Similarly, the now-classic "gap" between the known mechanisms of cognition and the strangely unquantifiable feelings of subjective awareness cannot be flippantly closed by declaring either realm irrelevant, but it can be closed in an intellectually satisfying manner by consideration of how both cognition and awareness have their origins in certain primitive properties of the neuron.

F. Conclusions

It is no longer controversial to maintain that synaptic communications are the essence of cognition: simple combinations of excitatory and inhibitory neurons can perform all the logical operations needed for any kind of information-processing. Although we experience subjectively only the complex, probabilistic cognition of multi-million neuron networks, the extrapolation from well-understood small neuronal systems to the "real" cognition of whole-brains does not strike us as unreasonable – precisely because the mechanism of the proto-phenomenon is known. As a consequence, although it may never be possible to verify the activity of thousands or millions of neurons involved in any specific cognitive event, this is not seen as a crisis in cognitive psychology. On the contrary, the fact that we have a firm understanding of the logical operations of few-neuron systems suffices as the basis for an understanding of potentially all cognitive processes as "simply" their temporal coordination and summation.

With regard to awareness, there is a similar "gap" between the feelings that we actually experience and well-understood single-neuron events, but again extrapolation from the single-cell level may be justifiable. The subjective impression that we directly experience the external world – i.e., that we have feelings that are distinct from cognitive processes – is a whole-brain process that has a corresponding protophenomenon at the single-cell level – the transitory opening of the neuronal membrane during the action potential. The semipermeable membrane quite literally allows a small part of the environment to enter into the cell and results in the readjustment of the electrostatic state of the intracellular fluid to a state more similar to that of the extracellular fluid. It is a mechanism by which the cell directly learns about its environment and becomes more like it through the diffusion of charged ions across the cellular membrane.

No attempt has been made to develop this argument to explain the many forms of human consciousness, but it is clear that consideration of the complexities of neuronal circuitry will be required to make meaningful distinctions between normal consciousness, abnormal states of consciousness, subconscious phenomena, and the unconscious state (sleep, drug-induced loss of consciousness and coma). The contention in the present chapter has simply been concerned with how neuron physiology may provide the basis for an explanation of even the difficult philosophical issue in consciousness studies, i.e., subjectivity – thereby allowing consciousness studies to move beyond the objective/subjective quandary and to focus on the neural circuitry issues. Making the connection between objective phenomena and the mechanisms of subjectivity is important primarily in that it could draw energies away from some of the more bizarre attempts to account for subjectivity through dualist philosophy or outright mysticism. If it is possible to conclude that we have identified the single-cell correlate of cognition (synaptic events) and the single-cell correlate of awareness (opening of the cell membrane during the action potential), then theoretical and experimental efforts can be fruitfully spent on the complex issues of brain circuitry – and routed away from the fruitless search for subjectivity in dualistic philosophy. Clarifying how specific brain circuits and their temporal coordination can produce different states of consciousness and cognition is the main business of human neuroscience – and much work remains to be done.

Although there inevitably remains a "gap" between our understanding of cellular-level "openness" and the subjectivity of whole organisms, the gap is not a conceptual mystery. On the contrary, it is identical in magnitude to the gap between the logical functions of small neuronal networks and the high-level

cognition of whole brains. We have a scientific understanding of the functional units and can extrapolate to the operations of large and complex systems without introducing mysticism, religion, inconclusive arguments based on a controversial interpretation of quantum mechanics or other forms of dualism that are antithetical to the scientific endeavor.

Notes

1. Having stated the obvious "realist" conclusion concerning the utter lack of direct internalization of the wider world by neurons, it is of interest to speculate on possibilities for synchronization of small neuronal modules with the cyclic regularities of external events. In so far as the patterns of synchronization within one brain are temporally-linked with the behavioral patterns of other people or events, the feeling of "oneness" and the significance of "chance" coincidences may have a neuronal basis. This, at any rate, is what I believe may be the sane backdoor into the phenomenon of synchronicity in a Jungian sense.

CHAPTER 8

A bilateral neural network simulation

Synopsis

A technique for the bilateral activation of neural nets that leads to a functional asymmetry of two simulated "cerebral hemispheres" is described. The simulation is designed to perform object recognition, while exhibiting characteristics typical of human cognition – specifically, a dual focus of attention, corresponding to the "nucleus" and "fringe." Sensory neural nets self-organize and the system is then taught arbitrary symbolic labels for a small number of similar sensory stimuli. Mutual homotopic inhibition across the "corpus callosum" produces functional cerebral asymmetries, i.e., complementary activation of homologous maps within a common focus of attention – a "nucleus" in the left hemisphere and a "fringe" in the right hemisphere. An object is recognized as corresponding to a known label when the total activation of both hemispheres (nucleus plus fringe) is strongest for that label. The functional dualities of the cerebral hemispheres are discussed in light of the nucleus/fringe asymmetry.

A distinction was drawn in Chapters 6 and 7 between consciousness and cognition in terms of basic properties of the neuron. In line with that view, it is unrealistic to believe that a silicon-based computer could become conscious in the normal biological sense of that word, because it is specifically the exchange of materials across the cell membrane that gives the neuron and neuron-systems their property of sensitivity to an external world. For the neuron, "receptivity to the surrounding world" is not a metaphor, but rather a factual statement about how it learns of external changes. For existing silicon systems, "receptivity" is metaphorical and all knowledge of the external world is hard-won through data processing. This fundamental difference between the living and the non-living worlds implies deep and probably lasting differences in the quality of "minds" of men and machines.

Nevertheless, the information-processing that neurons are capable of can certainly be simulated (but not the feeling of the information-processing!), so that, in principle, any form of cognition that can be explicitly defined can be simulated in a computer. Thus far, many sophisticated simulations have been developed world-wide (e.g., Edelman 1992), but the issue of the bilaterality of the nervous system, particularly with regard to language, has been thoroughly

neglected. To begin to fill this gap, a rather simple simulation of object recognition will be outlined in the present chapter, in which the bilaterality issue is considered. The computational capabilities of the simulation are in fact modest, but represent one way in which the functional asymmetry of the human brain might be addressed in an artificial system.[1]

A. Bilateral cognition

The *starting* point for the present simulation is a set of subjective psychological features that are widely acknowledged to be part of the normal human mind. No list of properties will be without controversy, but, as Galin (1996) has argued, a viable definition must include, at the very least, those properties that people frequently indicate are central to their own subjective experience. To leave those properties out of a "scientific" description is truly to leave out the very topic that motivates our interest in human psychology. As a final theory of mind, the working definition below will not suffice, but it is useful for illustrating the kinds of issues that need to be included in computer simulations of human cognition.

Underlying the simulation is the assumption that the "higher" forms of cognition are not of a fundamentally different character from "lower" level processes such as visual perception or motor control. On the contrary, given what is known about neurons, it is extremely likely that the types of neuronal processes known to contribute to sensory perception and motor output are also involved in all forms of cognition as well. This is indeed the essence of Crick's "astonishing hypothesis" (1994): there is nothing "higher" than neuronal circuitry! Therefore, unless we have clear indication that other factors (such as glial metabolism, electromagnetic fields or quantum effects) have *empirically verifiable* effects on cognition – and therefore that neuronal activity is *not* the sole substrate underlying all mental phenomena, we are logically obliged to address both "lower" level phenomena, such as perception, and "higher" level phenomena, such as human cognition using language, in computer simulations. There is no way around that argument except a quasi-mystical declaration that human language somehow transcends the material foundations of all other types of neuronal events.

Starting assumptions

The present chapter is an attempt to show how a computer simulation can exhibit certain properties that are often related to the human mind. The philosophical assumptions that underlie the simulation are as follows:

i. Characteristically, the human mind exhibits "coherence," "unity" or "singularity" (Baars 1997; Dennett 1991; Velmans 1991). Stated negatively, the mind has a "limited capacity." Despite the continual barrage of sensory stimuli and the huge array of memories that might at any time be brought to mind, we typically have *one* topic "in mind" at any given moment, and information extraneous to that topic is temporarily excluded. This is most clearly demonstrated in the mutually exclusive perceptions of the Necker cube and the either-or nature of awareness during binocular rivalry, but it is equally true during other forms of cognition: when we ponder an abstract problem, we necessarily limit our thoughts to that one topic. While it is of course possible to do two or more things "at once," introspection shows that the mind is focused on only one of several simultaneous activities, with the others somehow on a psychological back-burner.
ii. Despite the unity of conscious attention, there is nonetheless a familiar duality of mind. Given only one focus of attention at any one time, we are nonetheless often introspectively aware of two simultaneously active types of thinking – two perspectives on that one topic. The dual nature of such cognition was described by James (1890) as a "nucleus" and a "fringe," labeled by Galin (1996) as "feature-awareness" and "nonfeature-awareness," referred to as "text" and "context" by Ornstein (1997), and summarized by Mangan (1993) as "articulated" and "unarticulated" experience. James (1890:245) emphasized the duality of the mind, insisting that human beings normally have both a focus of attention and a peripheral or "fringe" awareness: "The significance, the value, of the image [in mind] is all in this halo or penumbra that surrounds and escorts it."

The simultaneity of "unity of attention" and "duality of mind" is a source of confusion, and most authors have focused on one or the other topic. As illustration of how these two topics are not necessarily contradictory, consider a situation where one is drinking coffee and awaiting a telephone call with music playing in the background and a book in hand. At any instant attention will be focused on either the flavors in the drink, the sounds of the music, the words in the book or the thought of the awaited call – one, not more, of several possibilities at a time. Other topics might spring to mind through mechanisms that

are not apparent from introspection and there may be brief transitional states where one topic is fading from and another topic ascending into consciousness, but it is quite unusual to merge the senses and think simultaneously of, for example, the coffee and the music or the book and the awaited telephone call at precisely the same moment. The normal situation would be an abrupt shifting of attention between such themes. That reflects the "singularity" of the mind. If, however, we were to focus attention on, for example, the music, it is possible to listen to the specific auditory features of the music and, simultaneously, to savor the musical mood within which those sounds organically, cohesively fit. Similarly, when we hear speech or read written text, we understand the words literally and, simultaneously, appreciate the wider implications of the language. In that sense, we have only "one topic" in mind – one focus of attention, but it is possible to consider two aspects of that topic. Often the duality is such that one aspect is concerned directly with sensory features and the other aspect is of a more indirect, implicative or contextual nature. In the realm of language, the duality of mind often takes the form of denotative and connotative understanding. The essential difference between "two different topics" and "two perspectives on the same topic" lies in the intrinsic associations between them. There is no inherent cognitive link between the coffee and the book, or the music and the awaited phone call – other than the fact that we are now engaged in all those activities, whereas the literal and contextual meanings of the text, or the melodic progression and the overall impression of the music feed on each other and relate to one another. On a moment-by-moment basis, attention to the implications of the text (chordal harmonics) follows directly from the literal text (melody), and leads back to it. In contrast, there are no inherent, reverberating cognitive paths between the coffee, the music, the book and the telephone call.

Both the singularity and the duality of mind should be included in computer simulations if they are to match with our subjective understanding of mind. Other properties might also be considered, but the focus/context duality of mind and the fact that we can actively think about only one topic at a time are certainly among the most familiar and frequently noted subjective properties of the human mind. Although the duality is less frequently discussed in the consciousness literature than the singularity of attention, some insightful accounts can be recommended (Galin 1994, 1996; James 1890; Mangan 1993). Here the focus is on the duality of consciousness in relation to the duality of the cerebral hemispheres. Galin (1996), in particular, has warned against a simple reduction of James's idea of nuclear and fringe awareness to the various dichotomies found in the laterality literature. Nevertheless, there are enough

similarities between certain aspects of hemispheric specialization and the nucleus/fringe idea to warrant further study. By formalizing those ideas within the framework of a simulation, the relationship between nuclear and fringe cognition can be made explicit, the nucleus/fringe duality can be related to various descriptions of hemispheric functions, and psychological mechanisms can be explored in ways not possible in a purely philosophical discussion.

The duality of mind

One of the enduring puzzles of human neuropsychology concerns the differences in the higher cognitive functions of the cerebral hemispheres (e.g., Corballis 1991; Gazzaniga 1995; Jaynes 1976; Kinsbourne 1978; Landis 1987; Ornstein 1997; Schiffer 1998; and Sperry 1966; reviewed in Springer & Deutsch 1998). A great many functional asymmetries have been reported, but the central problem concerns the asymmetry of language – not only because of the importance of language in human life, but also because the strongest laterality effects in both normal subjects and clinical patients involve linguistic processing. Although subtle disturbances of language can arise following right hemisphere (RH) damage (see Chapter 2), the main clinical finding is catastrophic disruption of speech and/or speech understanding following left hemisphere (LH) damage. The accepted wisdom is that inherent structural differences between the hemispheres must be responsible for the functional specialization, and indeed many small anatomical and physiological asymmetries of the human cerebral hemispheres are now known (e.g., Geschwind & Galaburda 1984; Galaburda 1995). However, the postulation of unique "language processing circuitry" in the LH begs the main question: Why would two essentially complete "brains" – with a full set of sensory, associative and motor components, and neuronal circuitry that is inherently plastic – develop in such a way that one "brain" is competent for language, and the other incompetent? Particularly in light of the fact that LH excision early in life results in the development of only mildly subnormal language abilities in the RH (Gott 1973), it is difficult to understand why the RH – which, throughout the individual's life, experiences every language event experienced by the LH – would not have a roughly comparable linguistic ability.

One possible resolution of this problem, as outlined in Chapters 1 and 2, is to consider (i) the benefits of unilateral control of the midline organs of speech, and (ii) the maintenance or accentuation of hemispheric differences through callosal inhibition. The traditional idea underlying a callosal inhibition model of hemispheric asymmetry is that a mono- or multisynaptic in-

hibitory effect would prevent the storage of identical information bilaterally, and therefore promote the development of distinct patterns of neuronal activity in the two cerebral hemispheres (e.g., Cook 1986; Denenberg 1980; Doty, Overman & Negrao 1979; Kinsbourne 1982; Landis 1987; Moscovitch 1973; Wigan 1844/1985). Although the theoretical arguments remain fundamentally unresolved and experimental evidence is inconclusive about underlying mechanisms, what follows is the description of a more complete system that builds on the idea of inhibitory callosal connections between informationally identical, bilateral neural nets. The network consists of 80 individual cortical maps learning in either supervised or unsupervised ways. The system utilizes an excitatory corpus callosum to relay sensory information between the hemispheres, and an inhibitory corpus callosum to connect areas of association cortex. The excitatory connections assure that both hemispheres receive all sensory and verbal information, while the mutually inhibitory interactions in association cortex result in a functional asymmetry that can be described as "nuclear" and "fringe" information-processing. These complementary modes are then put to use in a simple object recognition.

In any simulation using a bilateral (two hemisphere) architecture, the issue of the unity or duality of mind must be addressed. The topic of dual consciousness has already been thoroughly debated within the framework of the so-called split-brain research (e.g., Puccetti 1981; Schiffer 1998) and the preponderance of behavioral evidence seems to favor the view that, in the split-brain condition, the hemispheres can exhibit two rather independent modes of thought (Gazzaniga 1995; Sperry 1966). In the normal state with an intact corpus callosum, however, behavioral duality is rare or absent, suggesting that interhemispheric neural connections somehow force a mental unity on two nerve centers that, at least potentially, could diverge. This unity is a consequence of two factors:

i. Outside of artificial laboratory situations, the organism as a whole is situated in only one physical state. Despite sensory hemifield differences, all environmental stimuli will directly or indirectly enter both cerebral hemispheres. Insofar as cognition and motor behavior are driven by sensory stimuli, independence of the cortical activity of the hemispheres would be unusual. Moreover, when the organism undertakes motor behavior, it must be able to act in a unified manner. Whether the coordination of the organs of speech or coordination of the limbs for locomotion or tool-use, goal-directed behavior requires non-contradictory motor actions that involve the entire organism.

ii. Whether inhibitory or excitatory, the numerous fibers of the corpus callosum function to "yoke" the two hemispheres together. The nature of the yoking remains debatable, but physiological measures of cerebral cortical activity (EEG, glucose or oxygen consumption, blood flow, etc.) indicate that homologous cortical regions on the left and right are normally activated together (Chapter 1). There are normally massive *bilateral* increases in neuronal activity in relevant brain regions during task performance. This is not to deny that there are functional asymmetries or cerebral dominance, but the roughly *symmetrically bilateral* activation of the cortex is the normal state.

The unity of sensory experience and the continuous interaction of the hemispheres via the cerebral commissures mean that the potential "duality" of mind is greatly attenuated and that both hemispheres in the waking organism normally attend to the same topic. For this reason, the issue of "two different consciousnesses in one head" that may arise in the split-brain condition will not be addressed here. Our main concern is with the nature of the linked or complementary cognition found in the normal, callosally-connected brain.

B. The simulation

The goal of the present simulation is to implement highly-simplified versions of the above-mentioned properties of the human mind and then to study how the system performs in an object recognition task.[2] Specifically, after the system has undergone unsupervised (nonverbal) and supervised (verbal) learning phases for obtaining appropriate verbal labels for visual objects, it is presented with noisy stimuli for recognition. The underlying idea is that a "two-hemisphere" system with a division of labor for "nuclear" and "fringe" information-processing should outperform a "one-hemisphere" system or a "two-identical-hemispheres" system.

The neural network structure is shown in Figure 8-1. There are four levels of processing which correspond roughly to (visual) sensory cortex, (auditory) language cortex, association cortex and frontal cortex. The correspondence with specific cortical regions of the human brain is however weak; the simulation is designed primarily to show how functional asymmetries may arise from structurally and informationally similar hemispheres through the actions of an inhibitory corpus callosum, not to present arguments about the functions of specific cortical regions.

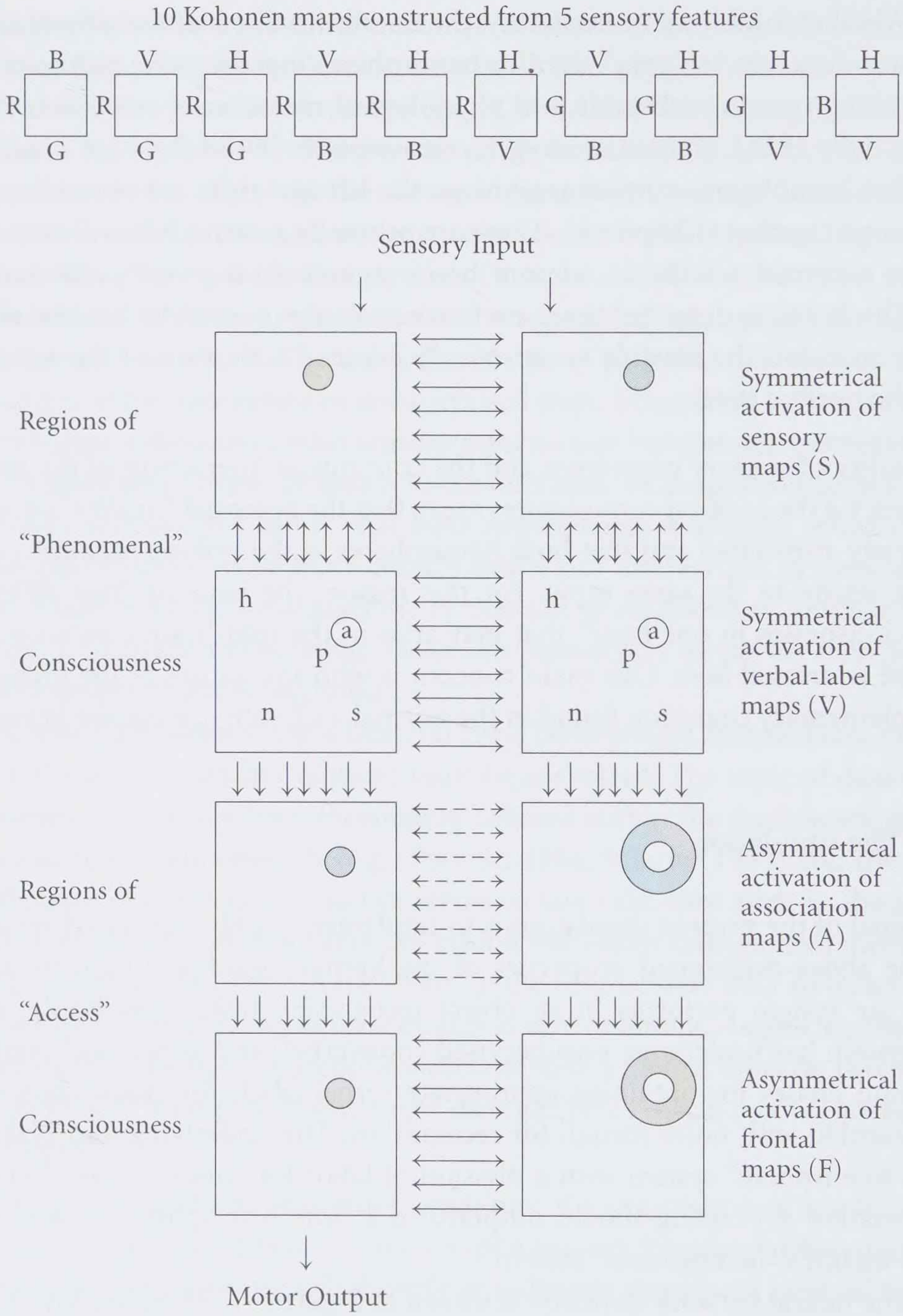


Figure 8-1. The entire simulation consists of four levels of paired maps in the left and right hemispheres. At each level (sensory [S], verbal [V], association [A], and frontal [F]), there are 10 maps [20×20 "neurons"] in each hemisphere. Each map handles three sensory features – giving a total of 32,000 cortical neurons. Interhemispheric connections include 24,000 excitatory (solid) and 8,000 inhibitory (dashed) fibers; all 32,000 intrahemispheric fibers between maps are excitatory. The complete processing in one set of related maps handling the same three features is shown. After the completion of

For implementation, a variety of practical issues concerning the flow of sensory information must be dealt with, and are discussed below. The strategies used here are not unique, but are thought to be simplified, physiologically-plausible mechanisms through which the psychological assumptions concerning the mechanisms of cognition can be realized.

Cortical maps are bilateral

One of the main motivations behind the simulation is to study how two structurally similar cerebral hemispheres might interact to produce functional asymmetries. For this reason, all of the neural structures in the simulation are paired (left and right) nets. To begin with, it is assumed that the bilateral nets are qualitatively and quantitatively identical – in recognition of the fact that both hemispheres experience essentially the same world. This assumption is, strictly speaking, not justifiable for brain regions involved in visual depth perception, auditory localization, etc., but such issues are not relevant to the current object recognition task. A *quantitative* asymmetry related to language is later introduced as a means of inducing hemispheric specialization.

It is noteworthy that "different modes of information processing" are not imposed on the simulated hemispheres. On the contrary, rather than explicitly implementing differences based upon preconceived notions about digital/analog, verbal/nonverbal, or analytic/holistic (etc.) dualities, differential effects are generated on the basis of quasi-physiological mechanisms of neuronal information-processing. Psychological dichotomies, such as analytic/holistic, are perhaps appropriate descriptors for summarizing the subsequent functional asymmetries, but *a priori* mechanisms should not be imbedded in the simulation at the outset by the designers; rather, such functional differences should emerge from neuronal processes. To impose high-level, information-processing asymmetries would be to give the hemispheres differential capa-

---

both nonverbal and verbal learning, a sensory stimulus will activate bilaterally symmetrical regions in "sensory cortex" (black dots). The label that lies closest to the region of sensory excitation is activated bilaterally in "verbal cortex" (the circled label). Stronger LH activation and mutual callosal inhibition result in "nuclear" and "fringe" activation in "association" and "frontal" maps (the paired dot and doughnut). Cortical activation of the maps is symmetrical until the third stage, where a quantitative asymmetry is transformed into a qualitative asymmetry, resulting in different, but related regions of cortical activation in the two hemispheres.

bilities that are as useful or as limiting as are the starting assumptions, but such a "designer-imposed duality" is unlikely to tell us anything about self-organizing systems. For this reason, it is preferable to produce functional asymmetries using general mechanisms of corticocortical communication, and later to discuss descriptive dualities that follow from the differential activation of the cortical networks.

The processing of five visual features

In the simulation that follows, the information processing of each hemisphere is generated separately from the competitive activation of its multiple neural nets; the hemisphere which shows greater overall activation becomes "dominant" and has greater influence on the other hemisphere and on behavior. Their combined effects determine the one "topic" of conscious attention. The simulation is thus essentially a proposal concerning mechanisms for activating multiple cortical maps bilaterally, summing their effects separately in each hemisphere, and using that bilateral activity to control motor behavior.

The system is extremely primitive in terms of both sensory input and motor output, but it is a "complete" system with continuous sensory, cognitive and motor processing. It processes five visual sensory features [redness, greenness, blueness, vertical edges and horizontal edges (R, G, B, V, H)], and performs object recognition on the basis of those features. The sensory features of which human beings are normally conscious are of course more numerous, multimodal and of a "higher" character, but it should be possible, in principle, to discuss the basic mechanisms of consciousness (if not the richness and complexity) regardless of the content of sensory stimulation or the nature of the system's motor behavior. The sensory features used here are such that they can be easily extracted from a video camera image in a robotic implementation, but they could be replaced by more complex features for simulations of greater psychological interest. For example, if "geons" were the visual primitives, or if abstract features appropriate for categorizing animal species, edible plants or common tools were used, considerably more preprocessing of the raw visual stimuli would be required, but the subsequent topological organization of sensory maps would be analogous to that obtained using primitive RGBVH features.

Input of these sensory features is hard-wired, but the cortical organization of sensory information is a consequence of a learning process that produces 2-D Kohonen-style (1995) topological maps (see below), each of which can handle three of the five sensory features. More features and multidimen-

sional processing of sensory information could have been assumed, but the purpose has been to keep the system simple, while aiming for the principles of "higher" cortical processing. By restricting the sensory system to topological maps that process only three sensory features each, learning is rapid and does not involve a lengthy error-correction procedure. These characteristics facilitate real-time implementations and are essential for using the neural nets in a truly autonomous agent (Pfeifer & Scheier 1999).

Lower brainstem and thalamic activity is undoubtedly necessary to maintain, a waking state and a focus of attention, respectively (Steriade, Jones & Llinas 1990; Jones 1985). As important as both issues are for an understanding of the physiology of consciousness, these functions have not been implemented here. That is, the system is always "alert" when running, and "attentional focus" on environmental stimuli is assumed, i.e., the system does not need to distinguish among several objects simultaneously present in the visual field. This is again an extreme simplification of the actual biological situation, and thus avoids a host of early sensory processing issues. As shown by Schillen and König (1994), synchronization of neuronal responses is a powerful means for distinguishing among multiple objects and is a likely brain mechanism underlying visual attention. The precise mechanism of excluding some sensory information from attention and "binding" other information into a single percept is, however, not essential for illustrating the phenomenology of attention itself, once the attentional spotlight has focused on one of several possible stimuli. Therefore, for the purposes of the present simulation it is assumed that the system always sees a stimulus that corresponds to one object at the "focus of attention" and that simultaneously activated neurons in separate cortical maps have a temporal synchrony that binds the features of the activated portions of such maps together. The inclusion of mechanisms for synchronization (Singer 1995; Roelfsema et al. 1996) and the effects of the intralaminar and reticular thalamic nuclei (Bogen 1995; Crick & Koch 1990) would increase the realism of the simulation, but the interaction among cortical areas for cognition can be simulated without implementation of these (and many other) neuronal details.

Two-dimensional cortical maps

Other aspects of brain anatomy and physiology, however, have direct bearing on the nature of corticocortical communications and must be an explicit part of the simulation. Specifically, the storage of much sensory (visual, auditory and somatosensory) information is known to occur at the cortical level in two-dimensional maps for various modalities and submodalities (Chapter 5). A

central contention of the model is that the (synchronous) activation of certain of these cortical maps represents the psychological contents of (that is, the information processing associated with) consciousness. Precisely which cortical maps correlate with consciousness will change over time and will differ among animal species, but the main argument is that conscious neuronal activity will at all times correspond to the localized activation of certain cortical regions, i.e., the synchronized activity of *cortical* neurons. In the human case, early visual processing of, for example, edges or moving edges is apparently *not* accessible to consciousness. The "edginess" of a visual image or its optical flow pattern is not what we consciously have "in mind." On the contrary, given a variety of such early visual processes occurring in the human visual system, the conscious mind is normally concerned with higher-level aspects of visual stimuli – e.g., the identity of objects based on such lower-level visual features. Early visual processes such as edge extraction and optical flow thus correspond to *preconscious* processing of visual information for utilization in cortical maps that are organized on the basis of higher-level visual features and that *are* accessible to consciousness. In the cat or frog, these early processes may already be accessible to the cat or frog "consciousness," while various kinds of higher-level visual processing may be absent. Whatever individual or species differences in the nature of subjective awareness that may exist, in the present simulation, the earliest stage of visual processing (the sensory maps using RGBVH features) is sufficient for undertaking the object recognition task. This type of cortical "representation" of sensory features is therefore the information processing associated with the system's "phenomenal consciousness" (Block 1995) of sensory stimuli. Depending on the tasks required of the agent, the contribution of higher level cortical maps may not be necessary, but the general principle is that, at some level, the sensory information that guides motor behavior should be the central theme "in mind" for the system – both as awareness of relevant sensory information and as "access consciousness" of the cognitive elements that act as the precursors to behavior.

Two illustrations of the representation of sensory stimuli in Kohonen maps are shown in Figure 8-2. Given five sensory features (RGBVH, $n = 5$), a total of ten such maps, each processing three features ($f = 3$), can be produced $[n!/(n–f)!f!]$. While this number is far smaller than the number of sensory features available to living organisms, it is sufficient to illustrate the principles of multi-feature sensory processing and interareal communications. The values of the sensory features are all randomly-generated numbers normalized to 0.0~1.0 and homogeneously distributed. Whenever a site on such a map is ac-

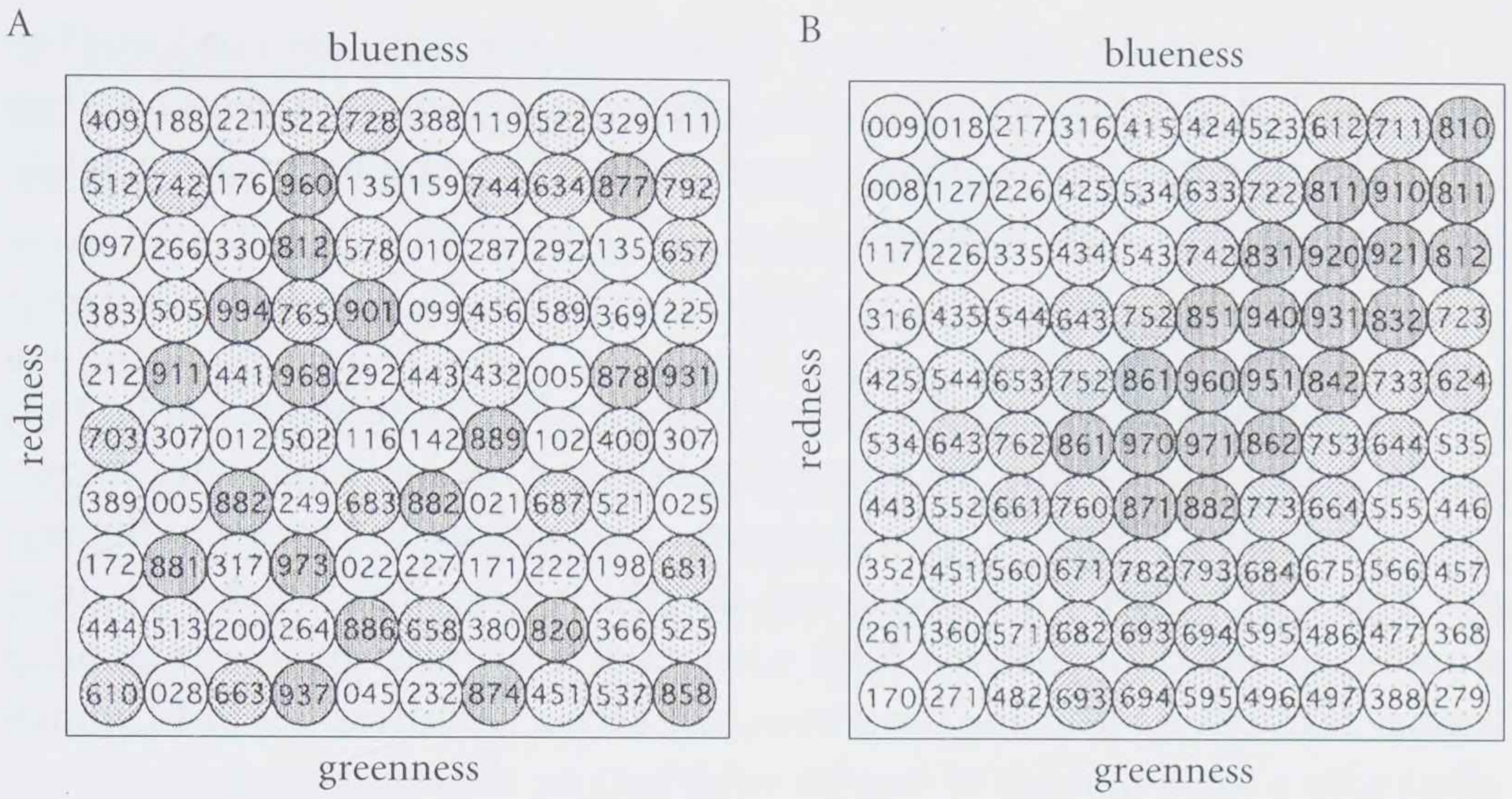


Figure 8-2. Unorganized (A) and organized (B) Kohonen maps formed in response to stimuli with varying intensities of primitive sensory features, say, redness, blueness and greenness. The strengths of the three features to which individual neurons respond are shown as three numbers in each "neuron". The first number indicates the strength of redness to which the neuron responds, the second, the strength of blueness, and the third, the strength of greenness. Each neuron thus responds to a specific combination of color information. On the left (A) is shown the map prior to learning: the topological organization of the values is random (the shading corresponds to the intensity of redness to which the neuron responds). On the right (B) is shown the map after learning: the strength of stimulus features to which each neuron responds is not markedly different from that of its nearest neighbors, indicating the establishment of 2-D topology. Note that the topological organization of the map after learning is different for the three different features. The shading of the neurons corresponds to the strength of only one feature (redness). The size of the sensory maps has been reduced to 100 units for illustration of the topology.

tivated, it corresponds to awareness of a correlation among certain intensities of the three features handled in the map.

At any given time, stimuli will activate localized regions on the maps depending on the fit between the current sensory features and the existing topology of the map. During learning, interneuronal effects within each map operate such that the neuron that most closely matches the stimulus features is activated and influences its near-neighbors also to respond to the current stimulus. The maps gradually self-organize, such that similar stimuli will produce similar activations of a given map and dissimilar stimuli will produce dissimilar activations.

Prior to learning, the neuronal responses in the maps are random (Figure 8-2A), but with sensory experience, topological organization develops (Figure 8-2B). The standard Kohonen learning algorithm employs a kernel neighborhood function, *h*, defined as:

$$h_{ci} = a(t)\exp(-||\mathrm{r}_c - \mathrm{r}_i||^2 / 2s^2(t)) \qquad (8\text{-}1)$$

where *t* indicates the discrete time, function *a* is a "learning factor", and *c* and *i* are the map coordinates of the best-fitting unit and the input stimulus, respectively (Kohonen 1995:79). s is a parameter that defines the width of the kernel. Initially, the kernel *h* is large, such that the neuron that matches an input stimulus influences a relatively large number of nearby neurons to respond to similar features. As learning progresses, the kernel size decreases until stimuli affect only a small number of nearest neighbors on the sensory maps.

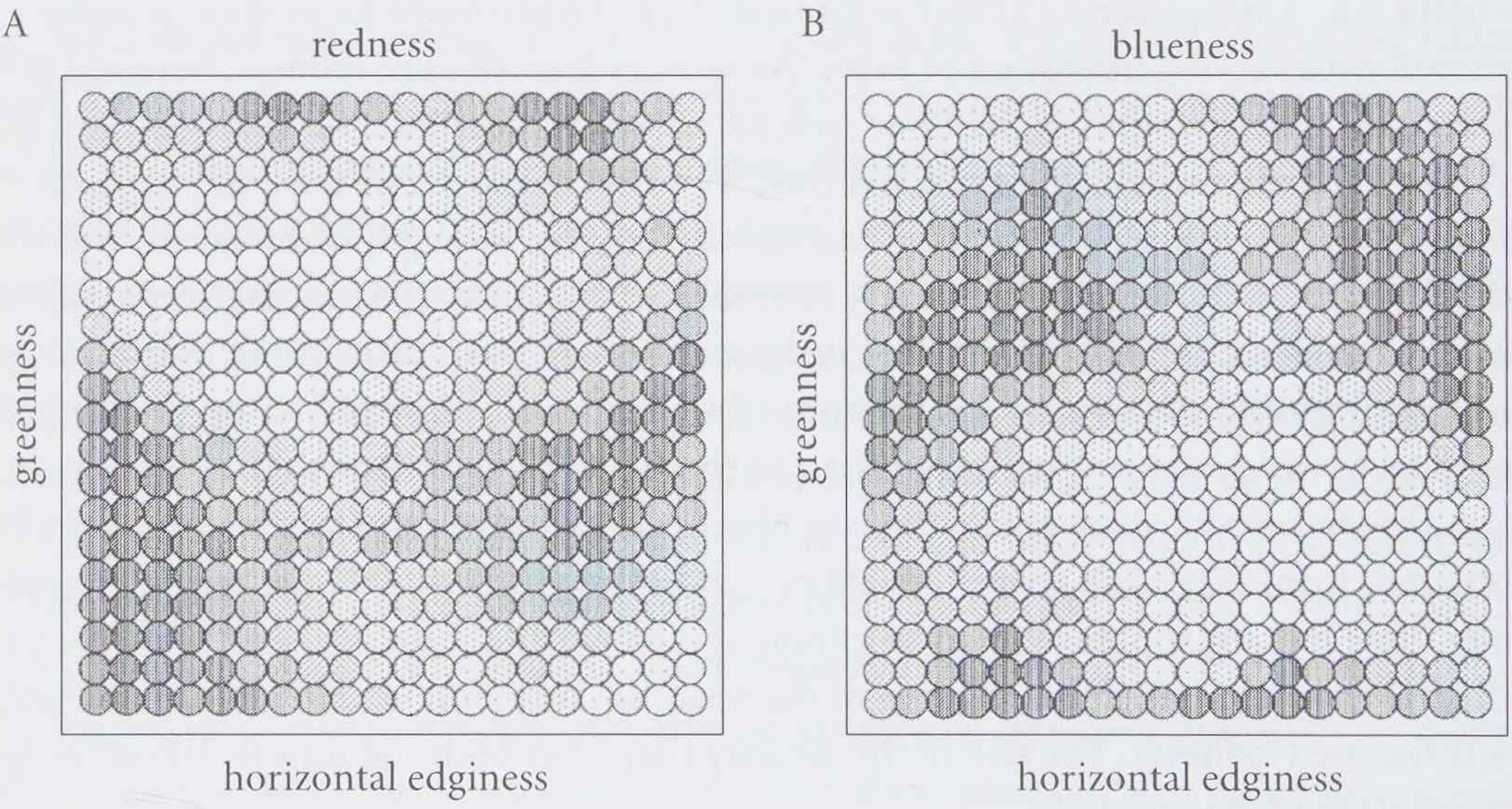


Figure 8-3. Each of the 10 sensory maps in the simulation contains 400 neurons, each of which responds to a unique combination of the 3 relevant features per map. In the simulation the color intensity was calculated from the number of red, blue or green pixels in the visual field. The horizontal (vertical) "edginess" was calculated from the number of horizontal (vertical) lines (three contiguous pixels of the same color). All sensory features were then normalized to 0.0~1.0 and used for the unsupervised Kohonen-style learning. The shading of the neurons indicates the intensity of greenness to which the neuron responds. The gradual changes in shading across the maps indicate that learning has occurred and the sensory maps have "self-organized." Note that the topology of the responsiveness to green in the sensory maps (A and B) differs, as a consequence of the random initial values, the differences in the features handled by each map, and the learning process.

Because each sensory map handles a different set of features and is influenced by the random numbers that determine neuronal responses prior to learning, the 2-D topology of the ten sensory maps (each of which processes a unique combination of three sensory features) will not have any obvious relationship with one another (Figure 8-3). As a consequence, the location of activation on any map in response to an environmental stimulus is different from the locations of activation in other maps. The orthogonality of the sensory maps has significance once verbal labels have been given to various sites (see below). When topological changes to the sensory maps are no longer brought about by sensory stimulation, the system's "infancy" has ended, and the verbal learning of "childhood" can begin.

Supervised learning of verbal labels for sensory events

After a period of nonverbal learning, the sensory maps become self-organized and further stimulation produces minimal changes in the map topology. The rigidity of trained Kohonen maps can be a drawback in certain applications, but the gradually decreasing reorganization of the sensory maps can be used to determine when the system has become sufficiently familiar with its local environment. Once nonverbal sensory learning has come to a halt, it is then possible to teach verbal labels for a smaller number of such stimuli. In practice, randomly produced stimuli (the number of which is ~5–10% of the number of neurons [20 × 20] in each sensory map) are presented and the neuron in each sensory map that shows the closest match to the stimulus is found and given a verbal label on separate verbal maps (supervised learning). The nature of the label is entirely arbitrary and is seen simply as a place-holder corresponding to a certain combination of the three relevant sensory features (Figure 8-4). Verbal-label learning is a one-shot process, after the completion of which sensory features can be activated directly via the verbal labels. A more sophisticated verbal learning process would entail repeated presentation of a given stimulus as seen from different perspectives, with the label ultimately being assigned to a set of prototypical feature strengths. The current one-shot learning process is chosen solely for simplicity.

Learning occurs separately for each verbal map in association with a sensory map. Because each of the sensory maps shows unique topology that depends on its sensory features, the random settings at initialization and the learning process, identical labels applied to each of the verbal maps are spatially *unrelated* to one another (see, for example, Figure 8-3). The "childhood" stage of verbal learning ends with the labeling of the maps with 20~40 labels.

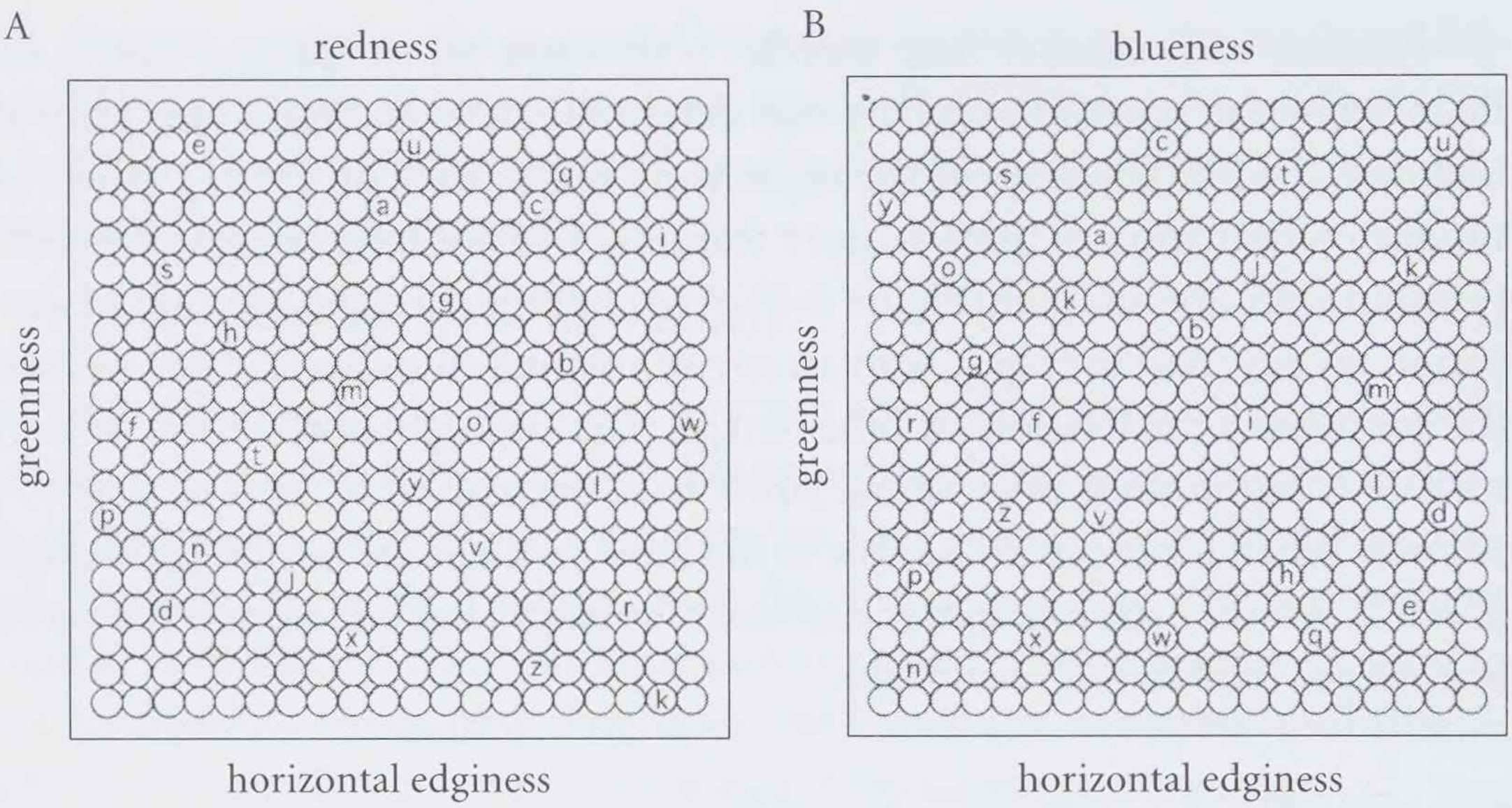


Figure 8-4. Two verbal maps to which 26 labels (arbitrary symbols, a–z) have been taught. The topological organizations of these maps (not illustrated in the figure) are identical to the maps in Figure 8-2. Noteworthy is the fact that the locations of the labels are orthogonal on the two maps. The site of maximal activation in response to a given stimulus will differ among sensory maps, and therefore differ among the verbal maps.

Map activation using verbal or nonverbal stimuli

After the system has undergone both unsupervised (sensory) and supervised (verbal) learning, further stimulation leads to activation of both kinds of maps due to reentrant connectivity among them (Edelman 1992). When a nonverbal stimulus is given to the "adult" system, the appropriate label will be found according to the similarity between its features and those on the sensory maps. A label is found for a sensory stimulus whenever the match between the new nonverbal stimulus and the previously learned verbal labels meets some criterion value for goodness-of-fit. In other words, if a sensory stimulus is sufficiently close to any of the previously learned verbal labels on the various verbal maps, the system responds by recognizing the stimulus as an exemplar of X. If not sufficiently close, no recognition occurs and the system awaits further sensory stimulation. The amplitude of activation on the verbal maps, s, is calculated from the mean distance, d, between the site of the sensory stimulus and the nearest label in each verbal map:

$$\text{if } s > t_{VERBAL}, \quad s = 1 / \exp(d) \quad \text{else } s = 0 \tag{8-2}$$

where d is the Euclidean distance between the peak of activation due to the current stimulus and the various labels on the verbal maps. By setting a threshold

value, t_{VERBAL}, above which s is seen as a "match" between labels and stimuli, the level of precision or "looseness" of the system when undertaking its object recognition task is fixed. When a low threshold value is used, the system will "recognize" most nonverbal stimuli as being exemplars of specific labels; when a high threshold value is used, the system will frequently fail to recognize nonverbal stimuli and will await further nonverbal stimulation before deciding on a label.

Using these definitions, if a nonverbal stimulus approximately matches a previously learned verbal stimulus on most of the sensory maps, the stimulus will be recognized. If, however, the new stimulus matches conflicting labels or no labels on such maps, it will not be seen as a recognizable object and motor behavior will not ensue (see Figure 8-5 for examples). (Various algorithms such as Kohonen's LVQ method [1995] and Grossberg's ART technique [1987] have been suggested for manipulating the location of the "labels" [prototypes or categories] in response to the fit with a specific sensory stimulus. Such techniques are of interest for improving the network's overall performance, but are thought not to be relevant in the present context.) Conversely, verbal stimuli can be shown to the system, and they in turn will activate the sensory maps. In this way, symbolic information generates sensory map activity corresponding to the perception of various sensory features.

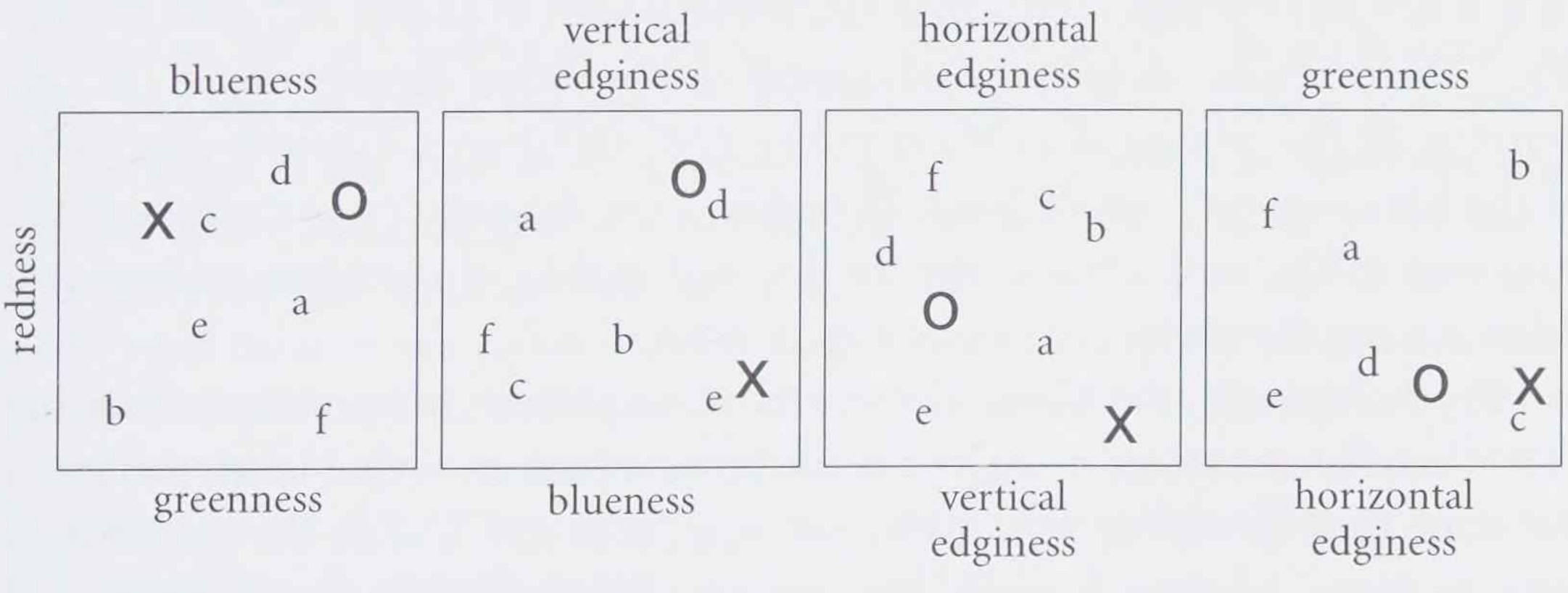


Figure 8-5. The mapping of sensory features onto the verbal maps determines whether or not an appropriate verbal label can be found for a new stimulus. Here are shown 4 of the 10 verbal maps, on which only 6 labels are shown. The position of the nonverbal stimulus whose sensory features place it at the site shown by "O" on the 4 maps is sufficiently close to the label "d" that the system correctly recognizes the new stimulus as an exemplar of "d". In contrast, the stimulus indicated by "X" has features similar to "c" on two maps, but clearly diverges on the other two maps. It does not reach a criterion level for labeling as "c". The success or failure of recognition is determined by the distance of the stimulus to the labels in all 10 verbal maps.

The role of language

The special role that "language" (i.e., the verbal labels) plays in specifically human consciousness is a key aspect of the present model. That is, the existence of verbal maps is essential, not only for the logical operations that might be done using the symbols (none of which are implemented here), but more fundamentally for the use of the labeled sites on the maps in determining the focus of activation of cortical maps. On the sensory maps that have undergone learning, activation takes the form of a simple Gaussian with a peak at the location that corresponds most closely to the specific combination of sensory features. All that is required on each of the sensory maps is that a "winner" is found – that is, production of a peak of activation at one site per map. In that respect, recognition is solely a process of matching the current stimulus with previous experience. The use of verbal maps, on the other hand, separates the stimulus from the sensory maps – and creates a distance between the two.

The verbal maps are assumed to have a more complex pattern of activation, i.e., a summation of excitatory and inhibitory Gaussians (a Mexican hat). The Gaussian of excitation has its peak at the labeled site lying closest to the peak in the corresponding sensory map; the overlapping Gaussian of inhibition is specified such that the second-nearest-neighbor labeled site receives maximal inhibition. This is to say that the process of choosing a label for a sensory stimulus is declaration that: "This stimulus corresponds to object 'A' " ***and***, implicitly, "This stimulus does *not* correspond to the rather similar object 'B'." By denying 'B', the relative sizes of excitatory and inhibitory Gaussians (the shape of the Mexican hat) are defined. Depending on the sensory features and the structure of the verbal maps, the heights and widths of the Mexican hats will differ among the verbal maps (see Figure 8-6).

Psychologically, the Mexican hat activation pattern is equivalent to saying that when we recognize a sensory stimulus and find its verbal label, we are at the same time declaring what labels are inappropriate. Unlike the response of sensory maps, labeling is essentially the act of distinguishing among a set of vaguely similar phenomena – for which we must inhibit those labels that are similar to, but less appropriate than, the best-fit label. For example, if we look at a table with dishes on it, we may see plates, platters, saucers, cups, bowls, glasses, mugs, etc. Focusing on one such dish, if we declare, "That's a bowl!", we are *also* implicitly asserting that "It's not a cup [because it is too big]" and "It's not a plate [because it's too deep]". We do not need to say "It's not a fork" or "It's not an elephant," because there are only very weak similarities between

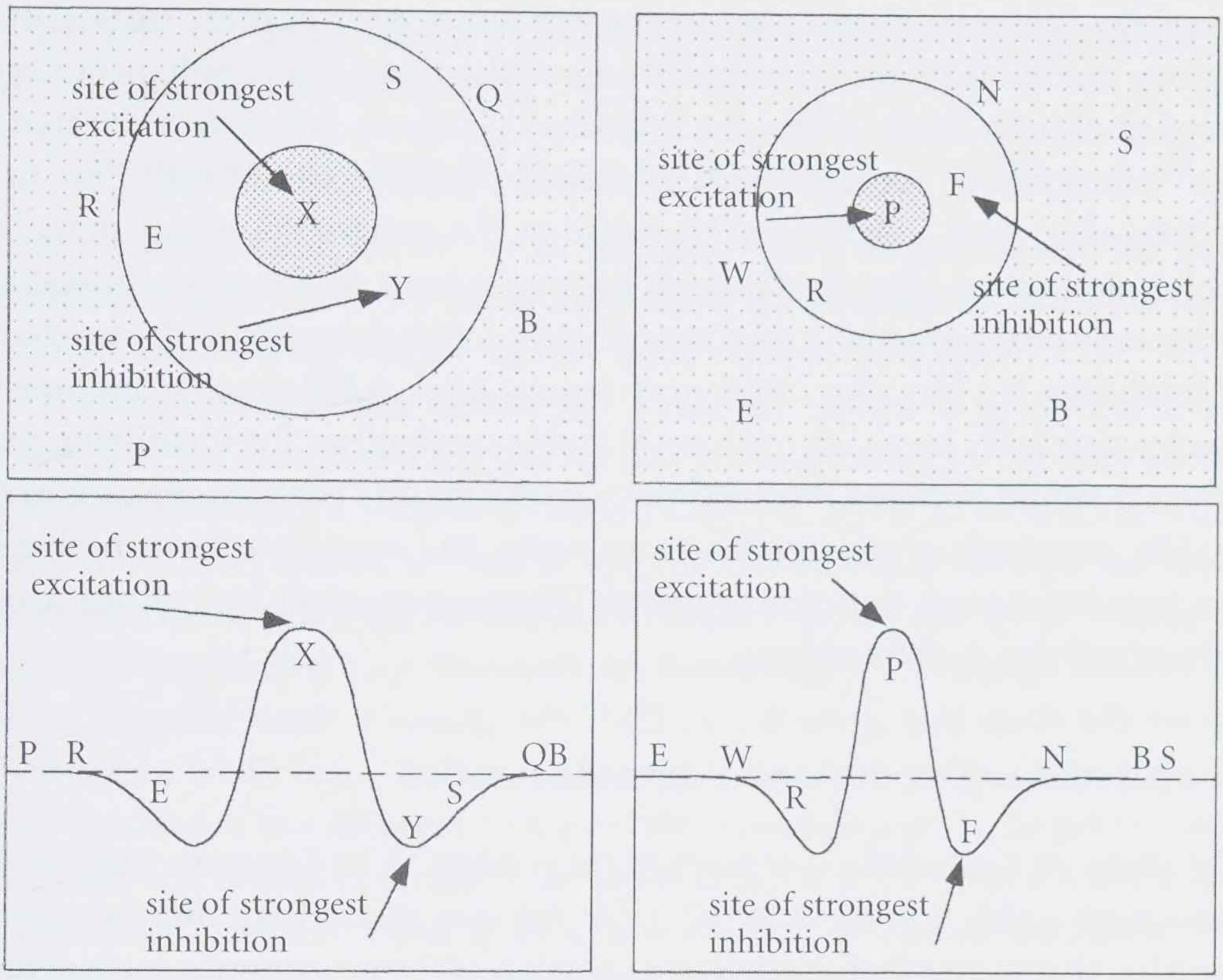


Figure 8-6. Two examples of the activation of the verbal maps such that there occurs inhibition of the labeled site that lies closest to the activated labeled site. This produces a "Mexican hat" of a width that depends on the characteristics of the particular verbal map. The top row shows the sites (X and P) and the regions (shaded) of activation in 2-D maps, as seen from above. Surround regions of inhibition are unshaded, while regions with a background level of activity are stippled. The radii of the regions of activation and inhibition are determined by the distance between the labeled site (X and P) and the next-nearest labeled site (Y and F, respectively). The bottom row shows the same maps from the side. Here, it is seen that the amplitudes of maximal activation and inhibition are identical in both cases, but the widths of the hats and the number of suppressed labels differ. These differences in the shape of the Mexican hat are a function of the locations of the labels on the various verbal maps.

the sensory features of a bowl and a fork or elephant. What we need to inhibit are the labels corresponding to the close neighbors of a bowl.

As a consequence of the nature of competitive 2-D maps, only one region can be activated at any given time. The peak of activation is necessarily singular; the region of inhibition is, however, defined in terms of the distance of the next-closest label from the peak – thus leading to a penumbra or surround of inhibition within which several labels may be suppressed. With regard to ter-

minology, "conscious" processes are, by definition, taken as those cortical map regions that exceed some threshold of neuronal activity. The simultaneous suppression of next-closest labels, the inhibitory Gaussian, is therefore referred to as a "subconscious" process. Such inhibitory processes contribute to the final pattern of cortical activity, and therefore to the current conscious state, but inhibition is not a conscious process because it leads to *decreased* cortical activity; only peaks of activity correspond to conscious processes. If the surround of inhibition is a function of the next-nearest label, it follows that any animal species that lacks language (or a comparable mechanism for identifying prototypes) would perform recognition by matching the peak on sensory maps against previously experienced sensory peaks, and would *not* have inhibitory mechanisms at work during recognition. They consequently lack subconscious mental activity, as defined here.

The parameters that determine cortical activation

The shape of activation on sensory maps is taken to be Gaussian with a half-maximum width arbitrarily set at 1/5th the size of the map dimensions. In practice, that means that a rather large portion of the neurons of a cortical map will be activated to some degree in response to a stimulus. The neuron with the best fit with stimulus features will be activated maximally (normalized to 1.0), and its neighboring neurons will be activated more weakly. Distant neurons will be unaffected. Activation of the sensory maps is defined as:

$$\text{SensoryMap}_{\text{LH}} = \text{SensoryMap}_{\text{RH}} = a \,/\, \exp(b) \qquad (8\text{-}3)$$

where a is a factor dependent upon the dimensions of the sensory maps and b is the Euclidean distance between the site of the current stimulus on the map and each of the neuron sites.

The site of maximal activation on the verbal maps is the labeled site that most closely matches the sensory stimulus. If the match is exact, the maximal activation on the verbal map is 1.0, with correspondingly weaker activation of neighboring neurons that match the perceptual features less well. The width of the Mexican hat is determined by the next-closest labeled site; maximal inhibition is assumed to occur for the second-closest labeled site. In order that the maximum of excitation is not entirely canceled by the overlapping inhibition in the Mexican hat function, an arbitrary amplitude ratio of 3:2 is used

for the excitatory and inhibitory Gaussians, respectively, and the amplitude of activation per map is calculated as:

$$\text{VerbalMap}_{LH} = \text{VerbalMap}_{RH} = 3.0 / \exp(r) - 2.0 / \exp(fr) \qquad (8\text{-}4)$$

where r is the distance on the verbal map from the activated label to the next closest labeled site and f is a factor influencing the strength of inhibition such that more closely lying labeled sites are more strongly inhibited.

The key feature of this calculation is that the trough of inhibition is determined by the second-nearest labeled site, but the resulting inhibitory effects necessarily extend to a wide area surrounding the center of activation – not simply to the label nearest to the activated site, but outwards from the activated site to perhaps several neighboring labeled sites. It is this process of inhibition on a discretely labeled map that produces interesting psychological effects (see below). In some cases, a nearby label will be inhibited, and there will be no other nearby labels (e.g., "It is a rhinoceros; it is not a hippopotamus [and all other species are too different to come into consideration]."). In other cases, inhibition will extend to several other labels (e.g., "It is a coyote; it is not a dog or a hyena or a jackal or a wolf or a fox [and all other species are too different to come into consideration]."). In either case, the surround of inhibition provides a kind of contrast enhancement that is not found in the sensory maps.

What the Gaussian activation of the sensory maps and the Mexican hat activation of the verbal maps produce are two effects. The first is simple recognition of a previously-labeled object that most-closely matches with the current sensory stimulus. The second is rejection of the second most-closely matching labeled object per map and, as a by-product, rejection of other somewhat less-closely matching objects.

Laterality effects via callosal inhibition

In the cortical processing discussed thus far, bilaterally symmetrical maps with excitatory callosal connections were implemented. These callosal fibers are in fact redundant in so far as the sensory input is identical to both hemispheres, but in a slightly more realistic simulation with visual hemifields, etc., the corpus callosum connecting regions of secondary sensory cortex would allow for the excitatory transfer of sensory information, such that both hemispheres contain information about the entire sensory field. In either case, the system is bilaterally symmetrical up to this point.

For information processing in association cortex, however, callosal effects are assumed to be inhibitory. As shown below, laterality effects are thereby

achieved without explicitly imposing qualitatively different "modes of hemispheric information processing." Specifically, the left and right association maps are initially constructed via ipsilateral excitatory transfer of the topology of the sensory maps and the labels of the verbal maps, such that the maps have the same dimensions and organization as the sensory maps and the same labels as the verbal maps. A quantitative asymmetry is then introduced by giving one of the bilaterally-paired association nets a stronger verbal activation for any given stimulus. The association maps in the left and right hemispheres differ *only* in regard to the overall level of activation – with one of the hemispheres being "dominant" (assumed to be the LH) in showing a stronger response on the association maps to stimuli with recognizable verbal labels. Such asymmetry of activation in language areas of the human brain is thought likely to be due to the necessity of having one of the two hemispheres control the midline organs of speech (e.g., Passingham 1981) – that is, to exert motor dominance over musculature that could simultaneously receive contradictory commands from the other hemisphere. Empirically, while motor nerves to the midline organs of speech come from both hemispheres, speech dominance of one hemisphere is the normal state in some 95% of people (Annett 1985). As discussed in Chapter 2, a failure to achieve unilateral motor dominance is associated with stuttering, and stuttering patients who momentarily overcome the stuttering show greater LH dominance during fluent speech (Fox et al. 1996). Such issues go well beyond the present simulation, but are indicative of the need for unambiguous unilateral dominance in at least some language behavior. In the simulation asymmetries are introduced in the form of greater left than right hemispheric activation due to language activity. The locations of cortical activation are nevertheless identical in the sensory and verbal maps of both hemispheres.

Association areas of the human cortex are known to have the densest concentration of callosal fibers – more than enough fibers to connect every cortical column afferently and efferently with its contralateral homologue (Cook 1986; Creutzfeldt 1995). If homologous regions with homotopic connections were activated identically, functional asymmetries could of course not emerge, but homotopic callosal inhibition acting between organizationally identical, but asymmetrically-activated maps can be shown to result in a qualitative asymmetry (Figure 8-7). The basic mechanism is a point-to-point (bidirectional) inhibitory effect on top of a bilaterally symmetrical activation between neurons in paired maps. The inhibition is proportional to the initial activation of each map. Assuming that cortical neurons normally fire at some background level, then excitatory and inhibitory effects will increase or decrease this firing rate. As illustrated in Figure 6, the post-callosal firing rate of a neuron can thus

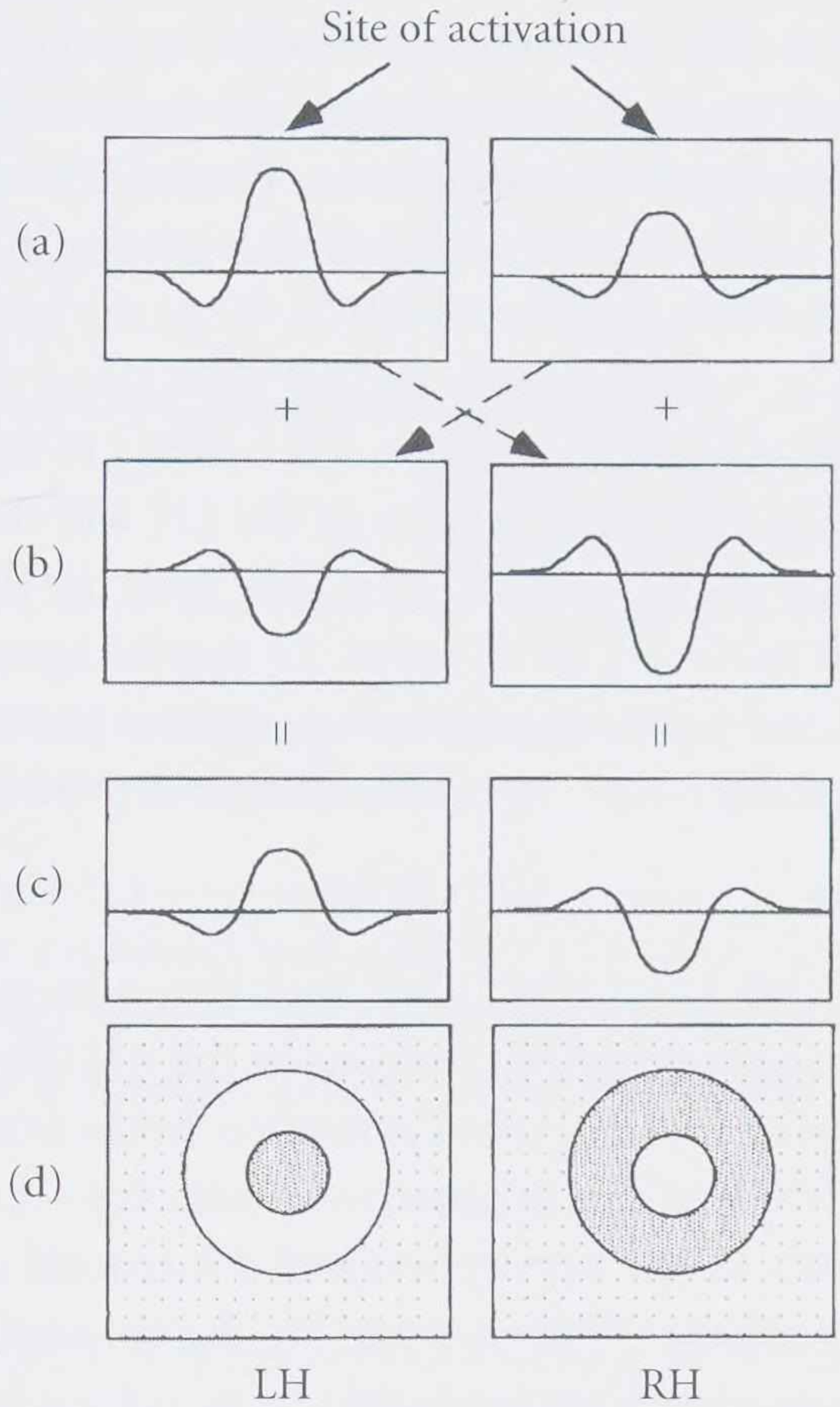


Figure 8-7. The effects of callosal inhibition. (a) The activation of a site on an association map is assumed to be larger in the language-dominant (left) hemisphere. (b) Both hemispheres send inhibitory callosal fibers to contralateral homotopic regions; the strength of callosal inhibition corresponds to the strength of cortical activation, thus leading to stronger inhibition of the RH by the LH than vice versa. (c) The result is a reduction in the amplitude of the "focus" in the LH at the original site of activation and a "fringe" of activation surrounding the original (now inhibited) site in the RH (see Rodel et al. 1992, for further discussion). (d) The pattern of activation, as seen from above. The LH association maps have a "nucleus" of activation; the RH association maps have a "fringe" of activation. It is noteworthy that the amplitude and sign of the peaks in these Mexican hats are a function of the intra- and interhemispheric transfer functions; suitable adjustments can produce various types of nuclei and fringes in the two hemispheres.

be calculated as the summation of an initial (ipsilateral) Mexican hat activation and the callosal (contralateral) effects, taken as an inversion of the contralateral activation.

Given appropriate callosal inhibitory strength, a strongly activated neuron (cortical column) will produce strong inhibition contralaterally and a weakly excited neuron will produce weak inhibition. Contrarily, an inhibited neuron (in the trough of the Mexican Hat) will result in a relative increase in activation contralaterally (Cook 1986; Rodel et al. 1992). Since the foci in LH association nets are, by definition, more strongly excited than those in homologous RH association nets, callosal inhibition of the RH by the LH will be stronger than vice versa. The Mexican hat of activation in the LH will therefore be reduced in amplitude after homotopic callosal inhibition from the RH, whereas the RH Mexican hat of activation will be reversed, or turned upside-down. The end result is that the LH and RH have *qualitatively* distinct patterns of activation of organizationally identical maps. The callosal effect is produced as follows:

$$\text{AssociationMap}_{\text{LH(RH)}} = \text{VerbalMap}_{\text{LH(RH)}} - C * \text{VerbalMap}_{\text{RH(LH)}} \quad (8\text{-}5)$$

where the callosal factor C is set at a level such that a Mexican hat remains in the LH and an inverted Mexican hat remains in the RH (Figure 8-7).

Given a quantitative asymmetry of activation in the bilateral maps, the inhibitory callosal effects lead to a qualitative asymmetry – i.e., a single focus or "nucleus" of activation in one hemisphere and a surround or "fringe" of activity in the other hemisphere. While supporting the concept of dual awareness, Galin (1996) has previously objected to this terminology [originally proposed by William James (1890)] because of the implicit unimportance of the "fringe" relative to the "nucleus," but the *spatial* implications of "nucleus" or "focus," on the one hand, and "fringe," "halo" or "penumbra," on the other, are precisely what is required in the present simulation. It is worth emphasizing that the "psychological importance" of any cortical region is a function of its amplitude (frequency and synchrony of neuronal firing) in relation to other regions. For this reason, the terms "fringe" and "nucleus" are used not to imply anything about their relative importance, but simply to specify the spatial pattern of activation – regardless of the amplitude which can be large or small for either the fringe or the nucleus. As Kinsbourne (1997) has noted, "'Fringe' information is no less well specified than information that is in focal attention. It differs in the nature of the contents, not in 'how conscious it is'."

Under the various assumptions discussed above, callosal inhibition produces a single peak of cortical activity (per association map) in one hemisphere and *multiple* fringe peaks in association maps in the other hemisphere (Figure 8-8). Depending on the level of activation of the focus and fringe, the total activity of the association maps in response to a stimulus may or may not cross a threshold for participation in the object recognition. Depending on the rel-

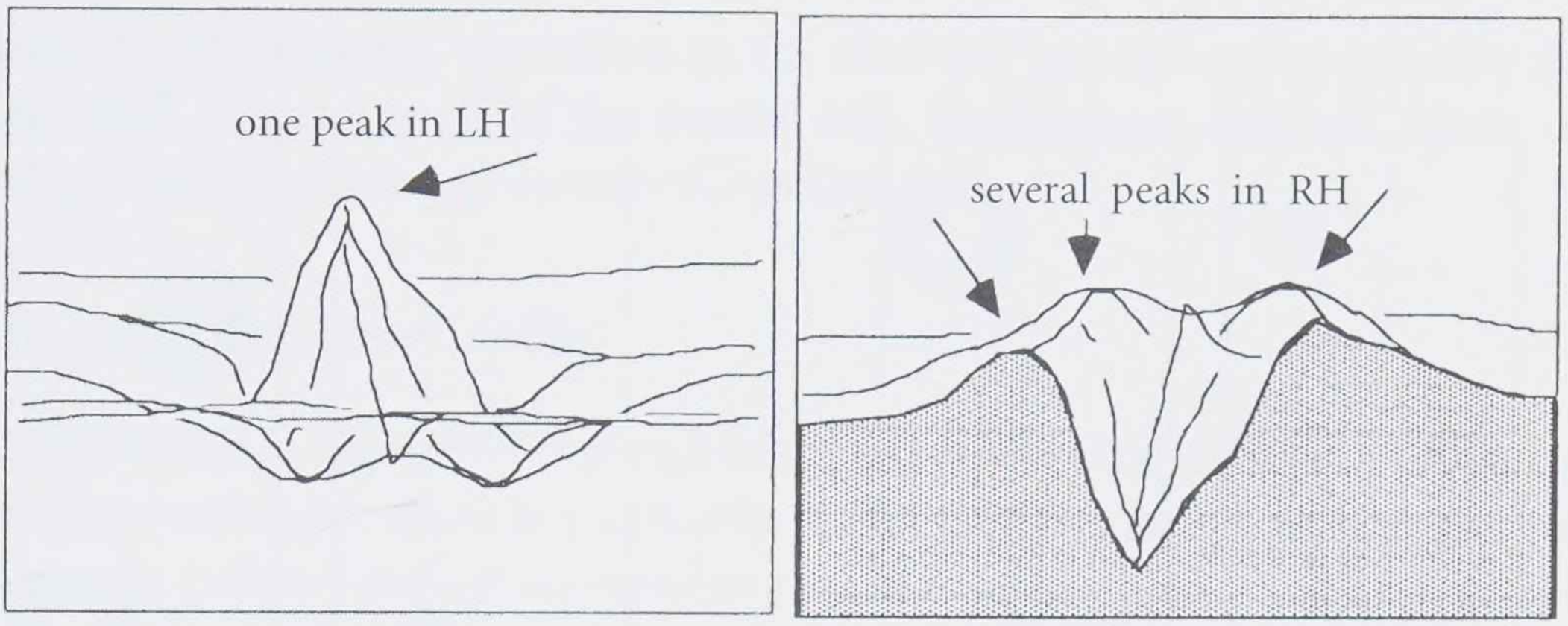


Figure 8-8. Illustration of a single peak of activation on the left and multiple peaks on the right, as a consequence of callosal inhibition.

ative strength of the excitatory and inhibitory processes, either the peak or the fringe could represent greater cortical activity, but *bilateral activation* that is asymmetrical at the millimeter level would be the normal state.

Action

The differential activation of the hemispheres is put to work by summing the positive activation of each hemisphere plus the effects of a facilitatory callosal transfer in frontal maps – the fourth and final level of cortical information processing:

$$\text{Frontal}_{\text{LH(RH)}} = \text{Assoc}_{\text{LH(RH)}} + \text{Assoc}_{\text{RH(LH)}} \qquad (8\text{-}6)$$

The sites of activation in each map are taken as simple Gaussians without inhibition. In effect, the combination of activated LH nuclei and activated RH fringes leads to a competition among various LH labels; the choice of a verbal label in the object recognition task (and subsequent motor execution) is then guided by the dominant LH. Examples are shown in Figure 8-9. When the RH fringes lie at various regions and are inconsistent among the frontal maps, the RH contribution to LH activation (via callosal transfer) is small, and the activation of the LH "focus" may not reach a criterion threshold for initiating motor behavior. In such a case, the system as a whole will *not* act on the basis of the current sensory stimulus, but will await further stimulation until the threshold for motor behavior is crossed. The summation of nucleus and fringe effects that also occurs in the RH is a computational "epiphenomenon" with currently no bearing on cognition and subsequent motor activity. The pattern of RH ac-

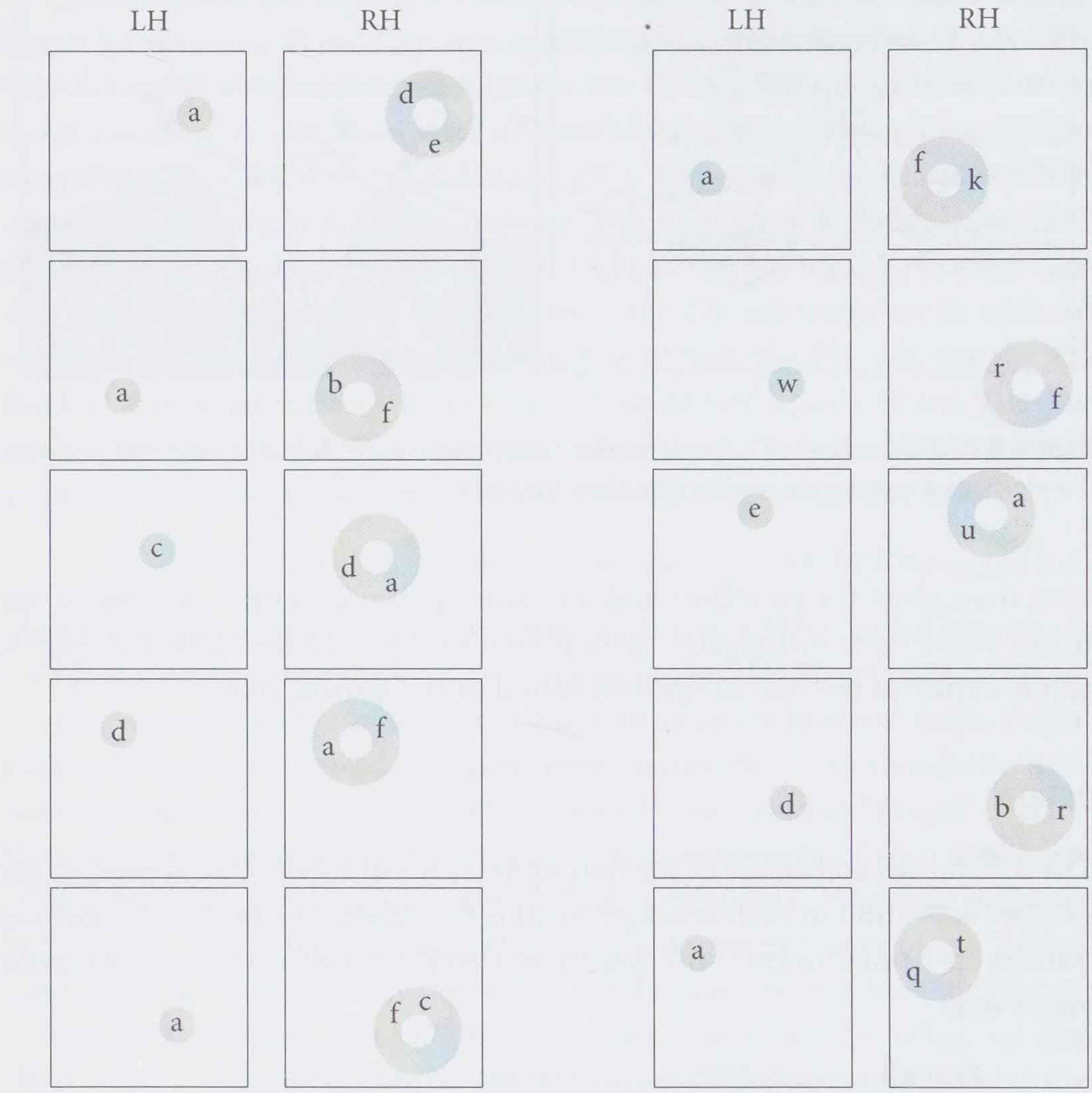


Figure 8-9. Two examples of the summation of LH and RH regions of activation. On the left is shown an example of RH fringe "support" of the LH nucleus: 5 of the 10 pairs of bilateral frontal maps, with only 6 of the 26 labels shown. The labels falling within activated regions (shaded) are included in the summation. In (A), the LH indicates 'a' as the correct label in 3 of the 5 maps; the RH indicates 'a' in 2 maps – giving more indications of 'a' than any other label. In this case, the threshold for recognition was crossed. In (B), an example of a RH "veto" of the LH choice of a label is shown: the LH indicates 'a' as the correct label in 2 of the 5 maps, and the RH indicates 'a' in only 1 map – giving a total vote for 'a' that is insufficient to meet the minimal criteria. No label selection is made for such a stimulus. Recent human brain-imaging studies have indicated an unanticipated involvement of the right prefrontal cortex during memory retrieval processes (see Chapter 2).

tivation will, however, contribute to the affective "mood" and specifically the subjective consciousness of the system with, for example, indirect effects on subsequent processing of sensory information.

The entire simulation system

As illustrated in Figure 8-1, for each of the ten sensory maps (where the feature topology emerges), there is a verbal map (where labels are given), an association map (where cerebral asymmetries via callosal inhibition are introduced), and a frontal map (where the relative strength of hemispheric activation is computed and object recognition occurs). The system as a whole is engaged in a cycle of activities – sensory perception, verbal-labeling, associative manipulation, and frontal summation for determining motor behavior. When the threshold for motor behavior is not exceeded, the system returns to a sensory level to await further stimulation and begin another cycle (Figure 8-10). Since competing tasks have not been implemented, the singularity of mind in the simulation refers only to the fact that activation of all cortical maps is limited to one region (focus or fringe). In a more complex simulation in which the mechanisms of neuronal synchronization have been implemented, the agent would be capable of multiple tasks and a competitive interaction among tasks would bring about a switching from one task to another at the completion of any sensory-verbal-associative-frontal cycle. Depending on the biological salience of competing tasks, the focus of attention would then persist as long as the synchronization of one neuronal subsystem involves more neurons than other subsystems. The essential idea behind the bilateral architecture is that, unlike a "unilateral brain," a "bilateral brain" has a principled method by which the "limited capacity" of LH focal attention can be increased, but without introducing conflicting, competing or simply irrelevant information. For the object recognition task, when the fit between external stimulus and previously learned features is close, there is no difficulty for the stand-alone LH to undertake motor behavior; in such a case, the LH finds the correct label for the stimulus, regardless of the presence or absence of fringe information. When the fit is not exact, however – as is normally the case in a real-world situation, interhemispheric communication leads to "contextual" fringe activity which can influence the focal activity of the motor dominant LH. The relative activation of the nucleus and fringe indicates whether or not the "best-fit" in the LH should be relied upon for motor behavior. If the RH activation is supportive, the RH serves to guide the LH in its motor output. If it is not supportive, the

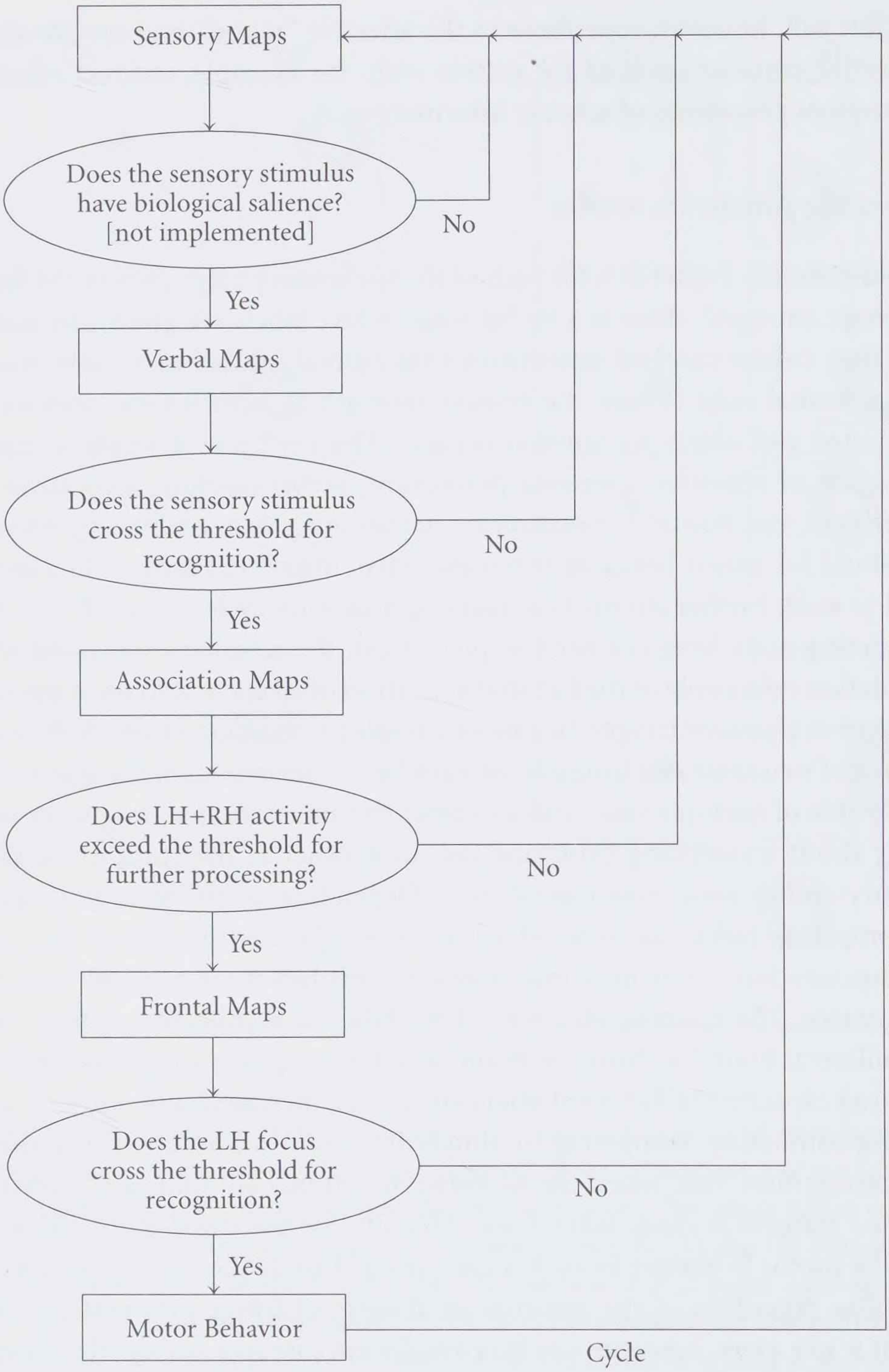


Figure 8-10. The flow of information among sensory, verbal, association and frontal areas in the fully-trained "adult" system. The first decision fork is not implemented; as a consequence, all sensory stimuli are assumed to have salience for the system's only possible activity, object recognition. At each of the four subsequent decision forks, the flow of information-processing can be aborted, at which time the system awaits further sensory input.

RH can effectively veto a weak LH nucleus of activation, and send the system into another cycle in search of more definitive sensory information.

Motor behavior is always controlled by one of the two hemispheres, by default the LH. The fringe in the RH acts as the feature "context" on each of the cortical maps, and can influence LH behavior, but the RH does not itself act independently. When the total activation of any given label in the LH frontal maps exceeds a threshold value, the LH acts immediately. However, when the contextual fringes of the RH combine with a weak LH focus to give activation above or below the threshold value, the LH is encouraged to or prevented from acting. In this way, the RH has an influence on LH motor activity that can be as great or greater than that of the LH, but the RH itself does not execute motor functions. Speculatively, it is the combination of two properties of the Mexican hat – nuclear vs. fringe activation (Figure 8-7) and single vs. multiple peaks (Figure 8-8) – that may underlie the most important functional differences of the human cerebral hemispheres. The coherence and singularity of LH activation is conducive to utilization in serial motor behavior, whereas the less-focused, inherently ambiguous, multiple activation of the RH maps is necessarily less useful in the guidance of unified, goal-directed motor behavior. The pattern of RH activation nonetheless corresponds to relevant contextual information and can be put to use *in relation to* LH activity, rather than on its own.

The entire system can be summarized in terms of its three temporal stages. The first is the system's "infancy," during which sensory maps are constructed. In this stage, the network must be allowed sufficient time to experience the full range of sensory stimuli in its ecological niche in order to construct the relevant topological maps. The second stage is the system's "childhood," during which verbal labels are attached to maps that are homologous to the sensory maps. Here, the network is taught labels for recurring, ecologically-meaningful stimuli in its environment. The third stage is the system's "adulthood" – an endless loop during which stimuli repeatedly activate the sensory, verbal, association and frontal maps.

The simulation was designed explicitly to perform object recognition on the basis of the closeness of the fit between a current sensory stimulus and information concerning similar objects previously stored in topological maps. When the fit between the sensory stimulus and labeled sites on the maps is close, the network performs well solely on the basis of LH activity. When the fit is not close, the network makes use of the contextual "fringe" information in the RH. As shown in Figure 8-11, there is a clear improvement in the system's object recognition capability when information from both hemispheres

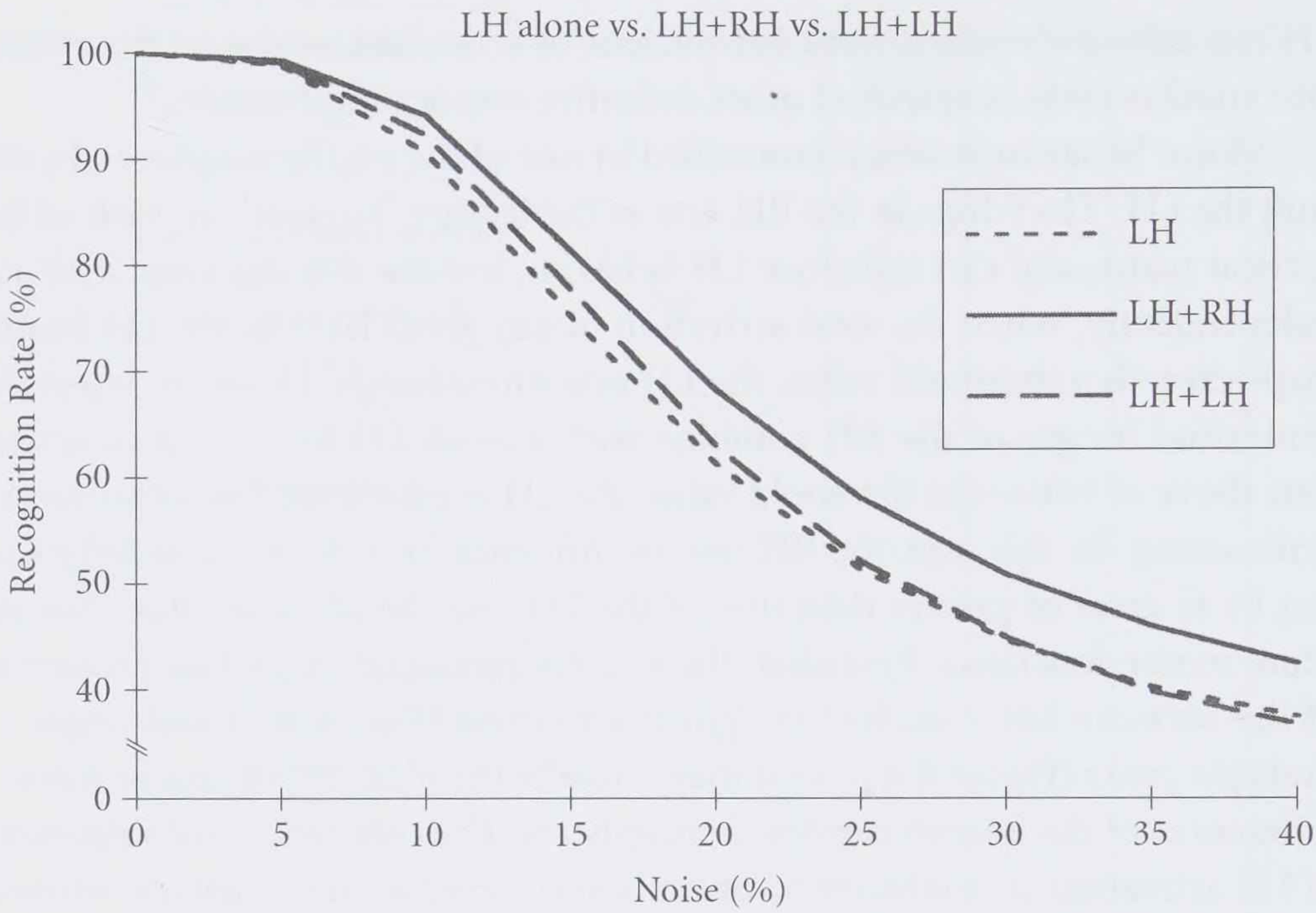


Figure 8-11. Comparison of the object recognition capability of a "one hemisphere system" (LH alone), a "complementary, callosally-connected two hemisphere system" (LH + RH), and a "two similar hemispheres system" (LH + LH). Correct object recognition is plotted as a function of noise (each point is the mean of 10 trials each consisting of 100 runs using the same map topology and all 20 of the labeled stimuli). Random noise (0–40%) was added to each of the values of the features for each of the labeled stimuli, and then the system was tested for recognition. (Different noise was used for the two similar hemispheres in LH + LH.) Correct hits decreased as noise increased for all three systems, but the LH + RH system showed significantly better recognition of noisy stimuli relative to either LH alone or LH + LH, as a consequence of the complementary RH fringe activity. (Thanks to S. Doi for simulation work.)

(LH + RH) is utilized, relatively to one hemisphere alone (LH alone). This effect is a consequence of the RH fringe augmenting the LH focus above the threshold value for recognition for those cases in which the LH focus indicates a particular object rather weakly. When the noisiness of the sensory stimuli is increased, the accuracy of the LH object recognition decreases sharply, and RH contextual information leads to improvements.

The advantage of the bilateral system over the unilateral LH system is consistent, and statistically highly significant (pairwise t-tests for noise levels, $10\% < n < 40\%$, $p < 0.001$). Because the bilateral system has twice as many neurons as the unilateral system, the improved performance by the former is not sur-

prising. A more relevant comparison is therefore between the bilateral system (LH + RH) and two hemispheres not connected by an inhibitory corpus callosum (LH + LH) to which similar levels of noise have been added. It is seen that the callosally-connected system (LH + RH) performs significantly better than the unconnected system in which one hemisphere functions merely as a back-up (pairwise t-tests for noise levels, $10\% < n < 40\%$, $p < 0.001$). These results were obtained without fine-tuning of the LH + RH system, i.e., without manipulating the strength of LH and RH contributions relative to the amount of noise.

C. Conclusions

A major caveat concerning the "bilateral cognition" simulated here should now be stated: there is nothing inherently "bilateral" about these patterns of cortical map activation in the computer simulation. They could be arranged in a fully "unilateral" architecture that implements the same functions; moreover, in the human brain it is likely that, following brain-damage or simply as a consequence of individual differences in development, the plasticity of the nervous system sometimes allows for similar complementary functions in a "unilateral" brain. What is essential for realization of the "nucleus" and "fringe" functions, however, are certain structural features that are easily and naturally incorporated into a bilateral approach. The most important feature is homotopic, bidirectional connectivity – or point-to-point mutual connections between cortical areas of comparable size (Edelman 1992). Such interhemispheric connectivity is of course found at many regions of association cortex, principally, across the corpus callosum, but bidirectional "reentry" connections are also found among many cortical areas within a single hemisphere (Creutzfeldt 1995). Similarly, the simultaneous activation of cortical regions is essential for the interhemispheric Mexican hat dynamics outlined here. In theory, appropriate timing might be achieved between any two (ipsi- or contralateral) regions of cortex, but the simultaneous activation of homotopic cortical regions is well-known experimentally (following either sensory stimulation or direct excitation of the brainstem). For these reasons, the nucleus/fringe duality is most naturally – if not uniquely – implemented in the framework of a bilateral simulation.

Of particular interest is the psychological characterization of nuclear/fringe patterns generated in the simulation. Since each of the cortical maps has a topological organization such that neighboring units represent closely related sensory features, it follows that the fringe generated for any given peak is the

natural semantic "context" of the peak. Because the cognitive processing in the association cortex of the simulation goes no further than recognition of the appropriate symbolic label for a sensory stimulus, the contextual processing of the fringe is no more than the activation of features close to, but not identical with those of the nuclear peak. Although this type of "context" is rather simple, it is easy to imagine that a more complex form of LH nuclear processing would generate correspondingly more complex forms of RH context. The "focal" and "contextual" functions of the left and right cerebral hemispheres in a variety of verbal tasks are a well-documented aspect of human neuropsychology (Cook 1986; Gardner et al. 1983; Springer & Deutsch 1998). While debate continues concerning the nature of such dichotomies, the simulation shows that a straight-forward, quasi-physiological neuronal mechanism can generate focus/context dualities.

Finally, in light of the simulation, let us consider the functional dualities that have been suggested as summaries of the information processing of the cerebral hemispheres. There have indeed been many such suggestions. Of interest is the suggestion of Bogen (1969) that the LH is "propositional" and the RH is "appositional." In the object recognition simulation, the LH "proposes" specific labels for the perceived objects and the RH "apposes" (or, more conventionally, "juxtaposes") several alternative labels, as a direct (transcallosal) result of the LH proposition. Also of relevance is Semmes's (1968) suggestion of "focal" and "diffuse" activation of the LH and RH, respectively. In the simulation, for every instance of focal LH activation, there is a simultaneous instance of relatively diffuse RH activation. Similarly, the high and low spatial frequency hypothesis of Ivry and Robertson (1998) can be seen as a restatement of the focal pattern of LH activation and the broader pattern of RH activation. (Interestingly, one of the more mysterious features this hypothesis – that the hemispheric preferences for higher and lower frequencies are relative, rather than absolute – is an implicit feature of the callosal inhibition hypothesis: at whatever scale of nuclear activation, fringe activation will be relatively more diffuse.)

Human cognition is of course far more complex than anything that can be produced in a ~32,000 – neuron simulation. Moreover, it is likely that the information processing of the human brain has a unified character because of the temporal coordination of brainstem arousal and thalamic attentional circuitry – i.e., due to the synchrony of neuronal firing in various brain regions including, but not limited to, neocortex (Chapter 7). Nonetheless, while the size of the nets, the nature of the neuronal coordination, and the mechanisms of excitation and inhibition in the human brain differ from those used

in the simulation, the simulation demonstrates that certain of the well-known, high-level, information-processing characteristics of the human mind can be reproduced using multiple, competitive, bilateral cortical maps with appropriate callosal connections. Specifically, the singularity of the focus of attention is a consequence of the competitive nature of the cortical maps. Each sensory and verbal map in the LH can have but one focus of activation and the summation of the activation of the maps produces a dominant singular "theme." The duality of mind, functional asymmetries (nucleus and fringe), dominance and cerebral "specialization" are a consequence of inhibitory callosal effects between bilaterally-paired association and frontal maps.

Among current attempts to define more clearly the subjective contents of the human mind, Mangan (1993) has revived James's concept of "fringe" awareness. He has summarized James's description of the fringe as including the following six features:

i. The "vague" impressions of the fringe are a "constant presence" in consciousness.
ii. They occupy the "periphery surrounding the clear-cut contents" of consciousness.
iii. They represent "contextual information" related to the focus of consciousness.
iv. The impressions of the fringe are "elusive". Even when strongly felt, they are vague, indefinite and do not "stabilize in the focus of attention". [As Galin (1996) has emphasized, the fringe is by its very nature "peripheral" awareness, and it is *not* information that is not yet in proper focus.]
v. There is an "alternation" of consciousness, such that the fringe briefly but frequently comes to the fore and is dominant over the nucleus of awareness.
vi. Fringe awareness can be important as an evaluative indication of the "rightness or wrongness" of the activity of nuclear consciousness. It guides the activity of the focus.

All of these properties are also found in the present simulation of the nucleus/fringe:

i. A fringe is always generated in the RH association maps at the time when the nucleus is generated in the LH.
ii. The fringe is literally a halo of RH cortical activity surrounding the site corresponding to the LH focus.
iii. The informational content of the fringe is, by the nature of the topological organization of the sensory maps, that information which is most closely

related to the nuclear information. It is "contextual" in the sense that for every activated LH label, there is one or more activated RH labels, each of which has sensory features similar (but not identical) to those of the LH focus. Of course, "context" in the sense of the "web of relations" that defines a concept is not found in this simple simulation, but, given the primitive nature of the sensory and motor processing of the system, it is perhaps appropriate to call the near labels with some similarity of sensory features the natural "context" of the chosen label.

iv. The content of the fringe in the simulation is never the information used for the direct control of motor behavior. If fringe information is to be the focus of consciousness, it is activated as a LH nucleus and a different fringe is generated in the RH. That is, in the act of "grasping" some part of the elusive information of the fringe by focusing nuclear attention on it, it becomes the nucleus and then no longer has the multiplicity and "vagueness" of fringe activity.

v. Both nucleus and fringe are generated in response to sensory stimulation. Moment-by-moment, the strength of fringe activation can be greater or lesser than nuclear activation. When greater, the psychological content of the RH fringe is "dominant" and has a strong influence on LH activity. When the LH focus is stronger, the LH is dominant and can undertake the recognition task regardless of RH activity.

vi. Finally, when the RH fringe is strong, its contextual information influences the amplitude of the LH focus – thus bringing the LH focus to cross a threshold value for initiating behavior or, contrarily, acting to reduce the amplitude of the LH focus. In this respect, it supports or vetoes LH motor activity.

Following James, several researchers have emphasized the need to discuss both the nuclear and the fringe aspects of consciousness – and *not* to force all issues of consciousness into the framework of the "spotlight of attention." More precise delineation of the various modes of the attentional spotlight itself is required, but it may be a mistake to consider all non-spotlight processes as being simply weak or fleeting "would-be spotlight" effects, rather than processes of a different, and inherently unfocused nature. If we follow James's advice and bring the "vague" issues of the fringe into the discussion of consciousness, there is still a need to define more clearly the various phenomena that are now referred to as "fringe" (Galin 1996) and to make necessary distinctions with "nonconscious contexts" (Baars 1988), implicit processing (Schachter 1987) and other "non-spotlight" phenomena of the mind.

In the present simulation, four levels of cortical processing have been implemented, each of which corresponds to a form of consciousness in the human brain. The first two, the sensory and verbal maps, correspond to Block's (1995) phenomenal consciousness, and are regions where awareness of sensory stimuli (verbal or nonverbal) is possible. The latter two, the association and frontal maps, correspond to access consciousness, and are those regions where there is awareness of candidate verbal labels in the object recognition task. Designing a simulation that more faithfully corresponds to cortical regions of the human brain would allow for greater clarity in future discussions of the mind.

It might be argued that the present model deals with such low-level sensory features that the analogy with the human mind is simply too great a leap to be convincing. Clearly, human beings are normally conscious of stimulus features of a more complex nature than the sensory features used here, so that, even if successfully embodying some relevant principles of brain functioning, the simulation is unquestionably primitive. If the difference between human cognition and, say, insect cognition were simply a matter of the complexity of the sensory features used in cortical processing, this objection would be valid indication that the simulation is too primitive. However, while acknowledging the quantitative difference between the sensory features utilized by human beings and those in the simulation, the use of symbolic labels (and the associated mechanism for asymmetrical cortical activation, leading to a focus and its context) means that the simulation contains important qualitative elements of the human mind – most notably, the nucleus and fringe, and their complementary information.

Notes

1. The simulation of this chapter is based on a previously reported work (Cook 1999b), but the terminology is somewhat different. Specifically, when that paper was written, I still considered consciousness to be an issue of "information-processing" – somehow a higher-level form of cognition. As discussed in previous chapters, I now consider that view to be fundamentally mistaken. The present chapter is thus an attempt to straighten out some conceptual and terminological confusions, but, more importantly, to focus on the mechanisms of bilateral *cognition*. If, as I argue in Chapters 6 and 7, the subjective quality of consciousness is the whole-brain correlate of neuronal transmembrane "sensitivity" to the external world, then there is in principle no possibility for recreating the subjectivity of consciousness in silicon systems that do *not* exchange materials with their environments. Since the release of neurotransmitters is triggered by an action potential, the neuronal phenomena of synaptic communication and transmembrane sensitivity will always occur together; similarly, due

to the large-scale synchronization of neurons, cognition and awareness will normally occur together, but the terms refer to either the information-processing or the subjective feeling aspects of neuronal activity.

I remain skeptical about the production of consciousness in non-biological systems for the one reason that silicon systems do not embody anything even vaguely similar to the openness and receptivity to the physical environment that biological cells clearly do: Living cells *feel* as they process information, non-living systems only process information. This subjectivity issue may, however, be irrelevant with regard to possibilities for simulating cognition. Provided only that the biological urgency that living systems can experience through subjective feeling is in some way implemented in the simulation of cognition, I cannot see any limitation to the level of cognition that might be reproduced in non-living systems. This is not to say that an artificial system with a well-constructed simulation of feelings actually feels! In the same way that we do not take a fellow human being seriously who pretends to sympathize with us or pretends to be angry or pretends to be emotionally involved for some ulterior reason, hopefully we as a species will never lose a healthy skepticism concerning "emotional" robots, no matter how pathetic their sobs or friendly their grins!

2. For technical reasons, the entire simulation is pitched at a level that is easier to implement than the issues of hemispheric cooperation during spoken language. Reworking the simulation to include the phenomena of intonation would be of interest, but the principles of callosal interaction, hemispheric specialization, etc. can be illustrated even in the realm of visual object recognition.

Chapter 9

Conclusions

From a neuropsychological perspective, there is no doubt that the topic of cerebral lateralization is important for most high-level issues in human psychology (Chapters 1 and 2). Others have said it more eloquently and more convincingly, and yet it needs to be said again: There is something truly unusual about the left-right dimension in the human brain. The specialization of the cerebral hemispheres underlies the three kinds of mental activity that pre-occupy us for most of our waking lives: language, tool-use and music. To be a human being is to be involved in these kinds of activities – and, by inference, to use our brains asymmetrically. Other animal species – even our closest evolutionary cousins – show only faint glimmers of cerebral laterality, and their linguistic, tool-usage and musical talents are correspondingly impoverished.

Starting from that one absolutely firm conclusion, what is known about the cognitive differences of the cerebral hemispheres and the nature of neuronal activity has led us inextricably down the path of the "brain code" of Chapters 3, 4 and 5. Here, the evidence on the perception of emotions, voice intonation and musical harmony suggests that the right hemisphere is particularly capable at detecting affective states from an analysis of pitch combinations. An isolated pitch or even a pitch interval does not have unambiguous emotional content, but the psychology of music perception tells us that configurations of three harmonious pitches can have inherent affective meaning. Although the "sound-symbolism" of certain chords does not allow the conveyance of complex symbolic messages, such as those made possible through learned syntactical structures and lexical associations, pitch combinations can nonetheless convey simpler messages concerning an emotionally-favorable or unfavorable situation. The meaning of pitch combinations has been formalized in a variety of musical cultures throughout the world that are consistent in using pitch to convey emotions. The easy access to the emotions that music provides is undoubtedly the main reason why it is a central part of all known cultures and is, generally speaking, pleasurable: music is a means to experience emotionality in the abstract without any concrete referent – and without getting involved

in the complex, troublesome and time-consuming human relations that real emotions necessarily imply!

By focusing on what is well-known within the framework of traditional diatonic music theory, it has been possible to generalize the regularities inherent in harmony theory and apply them to the pitch phenomena in normal speech. By then pursuing the likely mechanisms by which the human brain is able to decipher the meaning of pitch and to use that information in the delivery of emotional speech and music, some conclusions concerning the underlying brain structures and processes have been suggested. Most importantly, since pitch information is known to be represented cortically in the form of two-dimensional tonotopic maps, mechanisms for transmitting and transforming patterns of cortical activity provide the basis for a physiological code through which the brain detects emotions in the right hemisphere, conveys that information to the left hemisphere via the corpus callosum, and expresses the affective state through the speaker's tone of voice.

Details of the code remain to be empirically confirmed, but it is certain that the issues discussed in the first five chapters will remain the focus of experimental and theoretical work. Asymmetrical cerebral activity, cortical maps, synaptic inhibition and excitation, and the timing of neuronal activity simply are – and are likely to remain – the essence of modern human neuroscience. The neuronal phenomena that produce cognition, affect and behavior might be reconfigured in different ways to spell out different brain codes for different species or different types of information-processing, but it is difficult to imagine how the mechanisms of mind could be explained without those neuronal phenomena playing a central role. Stated negatively, a dopamine or chromosomal or hippocampal theory of the human mind will not suffice. Those and many other material structures may be relevant components in a scientific theory of the human mind, but they will attain importance by fitting into the cerebral hemispheric specialization framework (not vice versa).

Yet, even if it can be said that some progress has been made in deciphering the brain code, an irritating question immediately arises whenever there is talk about deciphering how the human brain works. If linguistic cognition and pitch-based emotion are simply a matter of certain processes of cortical activation and information flow within the brain, why couldn't computers with virtually identical mechanisms of information-processing become psychologically fully human? Why would such machines not be conscious "feeling" beings whose power switch is as untouchable as human murder is unthinkable?

Most psychologists probably do not take artificial intelligence seriously enough to worry about such questions, but the logic is forceful. Is current

robot technology not the first wave of creating machines that we will someday be morally incapable of turning off? For better or worse, cognition in machines will continue to evolve and may someday be truly super-human in realms more impressive than board games. Be that as it may, I believe that by focusing on the properties of living neurons, it is evident that there are indeed some "special properties" of biological cells – and those properties imply that the artificial construction of consciousness – specifically the feeling of subjective awareness, will remain beyond technological reach unless the "artificial" systems are constructed of living cells.

The unusual approach to this problem discussed in Chapters 6 and 7 arrives at the conclusion that, whatever the coursing of electrons through a computer chip may feel like to the computer, the subjective feeling of neural systems is due to the fact that they feel their local environments by allowing a small part of the surrounding fluid into the cell. As soon as a computer chip is designed to do something analogous, there may be grounds for a plausible subjectivity argument for robots, but, until then, we should not forget that the only systems with indications of subjectivity are animal brains – for which the subjectivity of neuronal "feeling" is a direct consequence of living cells opening up their membranes and interacting in measurable, material ways with their environments.

The philosophical issues of mind are old and complex enough that the myopic view of resolving the consciousness vs. cognition problem through consideration of the dual nature of neuronal contacts with the outer world could prove to be precisely that – myopic. Nonetheless, the attraction of that view lies in the fact that it offers a self-consistent view on the objective/subjective duality while remaining squarely within the field of neuroscience. Many of the alternative theories of subjectivity imply a focus on material mechanisms that are non-neuronal and therefore lead to counter-intuitive and often silly suggestions about possibilities for robot consciousness or hidden aspects of plant consciousness or imply a full-blown anti-scientific philosophical dualism. In fact, most current work on consciousness remains sufficiently neurophysiological that it will contribute to a coherent theory of the human mind, but the properties of the living neuron – the crucial building block of the only system that we know beyond doubt to be conscious! – will necessarily be central to an understanding of why certain kinds of living systems have attained the seemingly unprecedented property of subjective awareness.

Given the neuropsychological certainties of Chapters 1 and 2, the physiological plausibility of Chapters 3, 4 and 5, and the philosophical common sense of Chapters 6 and 7, I am tempted to conclude that modern neuroscience has

provided us with all of the major pieces to solve the puzzles of the human mind. While the extensive task remains of sharpening the focus and discarding the false leads that were stated or implied in these chapters, there are also inevitable questions about what a science of human psychology might ultimately mean for the evolution of society. At this point, final conclusions are out of the question, but the task of sketching out possible implications is not wasted effort if it stimulates discussion based on the relative certainties of empirical science – with an eye opened and an ear tuned to the psychological common sense that is at the heart of all human cultures.

Appendix 1

Musical emotions

In the establishment of a major or minor mode, individual tones play specific roles in relation to other tones – and those roles can be delineated explicitly within the framework of traditional music theory. As discussed in Chapters 3 and 4, the underlying idea in elaborating a pitch-based theory of emotions is that three-tone major and minor chords constitute the core of all pitch phenomena in diatonic music. (Chords are of course related to scales and keys, but it is the three-tone chords that provide the psychological primitives of the emotional response to music.) This emphasis on mode, known as "dualism," has a long history in Western classical music, and has guided much of the work in traditional music theory from Rameau (1722) through Riemann (1877). Although Schoenberg (1911) explicitly denied the importance and inevitability of major and minor modes, it is not therefore necessarily the case that dualism is irrelevant to chromatic music. On the contrary, many theorists have explicitly or implicitly agreed with Cooke's (1959: 50) description of major and minor as "the twin poles of expression" – in relation to which the subtleties of musical affect are elicited, but in the modern era most composers have been interested in using the two modal extremes in rather complex ways that transcend the simple harmonic relations of, for example, 16th Century classical music. Notably, Harrison (1994) has argued for a revival of dualist theory specifically for the development of an understanding of atonal music. He maintains that even "rampant modal mixture in chromatic music does not alter our perception of this music as fundamentally tonal" (p. 19). Although the emphasis on the qualities of chromatic music is not on resolved harmonies, the extremes of major and minor modality provide essential reference points – without which music becomes a mere succession of tones with no higher structure.

The idea of modal duality is the starting point for explicating a basic theory of musical emotions. Assuming only that there is a perceptible difference between major and minor keys (whether learned, innate or, most likely, a combination of both factors), then the functions of all pitch combinations become

known in relation to the two modal extremes. By assuming that (1) the creation of harmonic tension and (2) its resolution in either of two (and only two) directions is the cornerstone of all "affectively significant" pitch phenomena, melodic sequences and harmonic combinations of pitches then attain certain unambiguous meanings within the framework of the twelve tones that are available for creating harmonic tension and then resolving the tension using major or minor chords. The implications of this approach are wide-ranging; here only the basic framework for a theory of the musical emotions will be outlined.

It is important to emphasize that the entire discussion is essentially orthodox in terms of traditional, diatonic music theory. It does not introduce idiosyncratic assumptions about the meaning of music beyond the one starting postulate of dualist theory. There remain questions with regard to other musical traditions that are of interest to ethnomusicology, but the crucial point is that the relations between tones (in diatonic music) are *known* in relation to their possible functions in major and minor harmonies. Naturally, as more and more complex music is considered, the number of permutations of affective pitch phenomena increases dramatically, so the present discussion is confined to the relations among 2- and 3-tones in the creation of major and minor chords. Clearly, this is an extreme reductionist approach, but it is worth noting that it is not the reduction of the phenomena at one level to that of a qualitatively different level (e.g., the reduction of psychological phenomena to biological factors). On the contrary, the reduction is from the complexity of polyphonic musical phenomena to the musical primitives of simple harmonies.

If one denies the necessity of harmonic resolution – maintaining, as some theorists would, that all of the regularities of Western diatonic music are simply culturally enforced habits, then the following theory of emotions also becomes "diatonic culture-specific," i.e., not universal. The main argument against such cultural determinism, as noted in Chapter 3, is that the "Western" diatonic scales are now de facto nearly universal and have a convincingly solid acoustical explanation in light of the upper partials of vibrating bodies. In either case – whether we assume that diatonic scales are an inevitable and natural outgrowth of the physics of vibrating bodies or, more sinisterly, a manifestation of Western European cultural hegemony, the meaning of 2- and 3-tone combinations in diatonic music is so familiar to all normal adults worldwide that only a cross-cultural study of tribes unexposed to the modern media might reveal the limits of the universality. Be that as it may, the meaning of pitch combinations strictly within the framework of diatonic music is known, and can be stated as follows.

We begin with a first, isolated tone – a pitch lying somewhere in the range of maximal sensitivity for the human ear, say, middle C at 261 Hz. Musically, the first pitch is, by default, the suggested tonic of the major key in which subsequent pitches will be interpreted. In Riemann's description, the first tone is the "thesis" that will normally be followed by other tones in the development of a musical theme. In a discussion of harmony, Schoenberg (1995:311) refers to the steps in establishing harmony as "assertion", "challenge" and "confirmation". (The assumption that the first tone is a proposed tonic for a *major* key is perhaps not necessary, but the positive/negative affect of major/minor harmonies would suggest that the positive affect – major key – will be the default choice.) The first pitch, however, only suggests, and does not, on its own, establish the musical key. Moreover, alone it gives no indication concerning the establishment of a major or minor mode.

The second pitch can be either consistent or inconsistent with the key suggested by the first tone. In Riemann's terminology, the second tone is an "antithesis" in relation to the first tone, but, as discussed below, the tonal features inherent to diatonic scales imply that the second tone can be either strongly antithetical to the first tone or largely supportive of it. That is, if the second pitch is one of the seven tones of the major scale or one of the seven tones of the minor scale built from the first tone, then the second pitch is consistent with the possibility that the subsequent tones will produce music in a key with the first pitch as the tonic. For example, if the first tone is C, then a major scale will use tones DEFGAB and a minor scale will use tones DD#FGG#A#. (Variants of the minor scale use tones from the major scale – A and/or B instead of G# and/or A#, but that is not of importance in the present context.) If, however, the second tone is not from either the major or the minor scale of the proposed tonic, then it is a true antithesis. In the example of C as the proposed tonic, tones C# and F# are *inconsistent* with any subsequent harmonic development in the key of C.

If the second pitch is consistent with the first pitch, it keeps the door open for possible establishment of either a major or minor key utilizing both the first and second pitches. It is noteworthy, however, that regardless of the interval produced, the first two pitches alone cannot *establish* a harmonic mode, i.e., a fixed and unambiguous musical mood. Even when the second tone creates a so-called major or minor interval, the major or minor character of the harmony will be established only upon the addition of a third tone that produces a resolved chord (that is, a minimum of three tones is required to establish mode).

If the second pitch is *inconsistent* with both major and minor scales based on the first tone, its inconsistency can be either weak or strong. Specifically, a tone that lies one semitone above the first pitch is inconsistent with either a major or minor key based on the first pitch as tonic (e.g., C# acts to deny the key of C – refusing the possibility that a major or minor mode will be built in that key). Both C and C# are, however, consistent with the C#-major scale (with C# as the tonic). A second tone of C# therefore effectively replaces C as the suggested tonic (demoting C to the role of so-called "leading" 7th tone in the key of C#-major), but implies (without establishing) that the new key will be major (since the seventh in a C#-minor scale would not be C, but rather B). In contrast, most other tones following C are consistent with the eventual establishment of harmony in either C-major or C-minor. The one notable exception is a second tone that is three whole-tones above or below the first tone – i.e., F#, which together with C produces a tritone interval (the "diabolus in musica"). A tritone interval is thus strongly inconsistent with development in the key suggested by the first tone. When a tritone interval is sounded, either or both of the first two pitches may remain in the establishment of mode, but the key will not have either tone as its tonic. (That is, the key will *not* be C-major, C-minor, F#-major or F#-minor, but the CF# interval is consistent with either the key of A#-minor or G-major). In this respect, the choices of a second tone that create a semitone interval or a tritone interval are the only ways of denying the original tonic.

In effect (and discounting the possibility of simple repetitions of previous tones here and in the subsequent discussion), the second pitch can in principle be any of the remaining 11 tones of the diatonic scale. Affectively, it can be either supportive or unsupportive: the second tone can either act to affirm the assertion of key suggested by the first tone (9 possibilities), or negate it (2 possibilities) (as illustrated in Figure A1-1 for the case of C as tonic). Negation can take two forms: replacement of the initial pitch as the tonic (a semitone interval) or denial of both pitches as the tonic for the establishment of key (a tritone interval).

In sounding a second tone that does *not* deny the first pitch, the affirmation of the tonic suggested by the first pitch takes three possible forms, corresponding to suggestions for proceeding to a major mode, to a minor mode or to either mode (i.e., ambivalent with regard to mode). Specifically, given the proposed tonic pitch, intervals of a third, a sixth and a seventh can be either of a major or a minor character, whereas intervals of a second, fourth and fifth are common to both major and minor scales. To begin with, let us disregard the differences within each of these three categories, and examine what the tonal

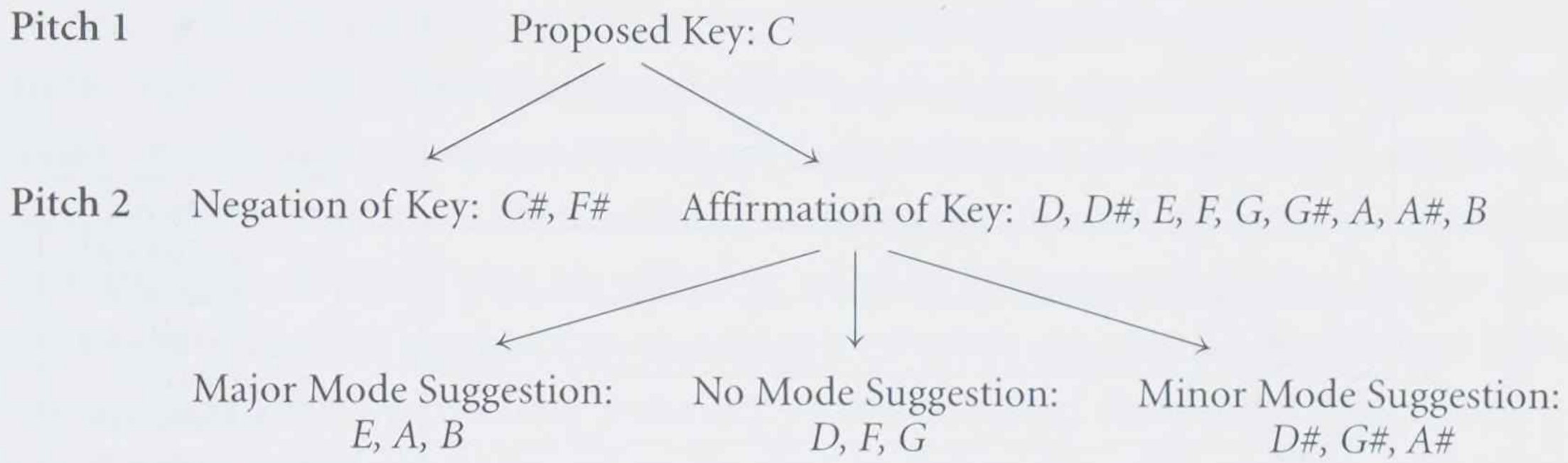


Figure A1-1. The roles of two consecutive pitches in the "suggestion" of musical key and harmonic mode. Letters in italics indicate pitches. Given the initial tone, *C*, there are two subsequent tones (1 and 6 semitones above *C*) that are incompatible with the first tone becoming established as the tonic of the key (i.e., either C-major or C-minor). The other 9 tones are consistent with the eventual establishment of a major or minor mode in the key of *C*.

"meaning" of the categories are. It can be said that a second pitch that does not work against the first pitch, can play two distinct roles: (1) supporting the suggestion of key provided by the first pitch, and adding a suggestion of mode (major or minor), or (2) supporting the suggestion of key and offering no suggestion as to affective mode. This pattern of tone combinations is shown in Figure A1-1.

The classification of pitches shown in Figure A1-1 is a simple consequence of the regularities of traditional diatonic music theory, but aspects of the classification also have some empirical justification in the folk songs of various cultures. Huron (2001) collected 181 songs (for which written notations were available) from 7 different cultures and tabulated the frequency of usage of various intervals. The results are shown in Figure A1-2.

The data summarized in Figure A1-2 are of interest with regard to cultural differences, but several outstanding cross-cultural effects are also evident. The first is the general tendency to use small intervals: in typical melodies from around the world about 75% of the intervals are 5 semitones or less, with fully 55% being intervals of 0–3 semitones. In other words, most melodies consist of gradual increases and decreases in pitch. Within that general trend, however, there are two notable dips that correspond to a reluctance to use one-semitone and six-semitone steps. The frequency of one-semitone steps is about half that which would be expected from the tendency to use small intervals in general and, most remarkably, the tritone interval is never used (in folksongs). Those empirical findings (from inevitably a finite sample) strongly indicate a tendency for sequential pitches to proceed in the direction of harmonic resolution

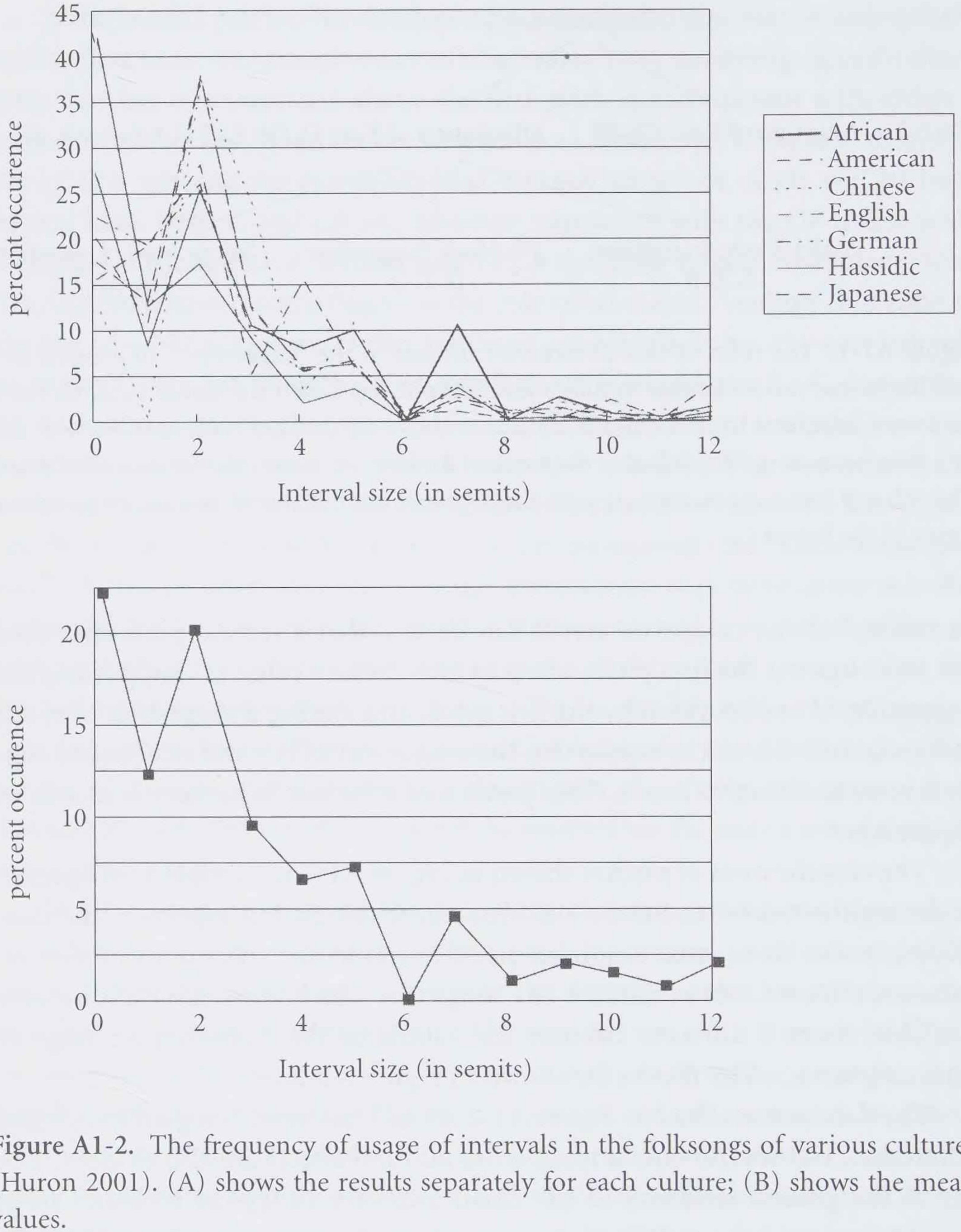


Figure A1-2. The frequency of usage of intervals in the folksongs of various cultures (Huron 2001). (A) shows the results separately for each culture; (B) shows the mean values.

and not to "negate" harmonic development through the use of the only two intervals that effectively deny the key suggested by a preceding tone.

No matter what the first two pitches are, harmonic mode itself cannot be established until (at least) a third pitch has been intoned. The third tone is referred to as the "synthesis" by Riemann, but clearly resolved harmonic "synthesis" is only one of several possibilities. What, specifically, are the effects of the third pitch? Because this is essentially a combinatorial problem, the number

of possibilities suddenly explodes, but there remains an underlying simplicity to the various effects of the third pitch. That is to say, the third pitch can either establish a (major or minor) mode, or fail to do so. If it establishes a resolved, major or minor chord, then both a harmonic mode and an emotional mood are implied. At that point an affective musical statement has been made; the addition of further pitches can strengthen or rebel against the established harmony (and thereby produce a variety of musical and emotional effects), but, unlike the mere suggestion of mode that an interval can imply, the establishment of a resolved chord produces an explicit harmonic affect that cannot be erased or retroactively modified by further tonal additions.

To state this more bluntly, major and minor chords have affective meanings that are unambiguous and "universal". Given that there is a lengthy historical record of the usage of major and minor modes and the fact that a "bright/dark" affective duality holds for both adults and young children, then the only theoretical work needed to elaborate a theory of musical emotions is an examination of the process by which tonal combinations can reach either of those two known modal end-states. Let us do just that.

If the third tone fails to produce a resolved chord, there are two ways in which it can fail. The majority of "harmonic failures" are tone combinations that work effectively as negations ("refutations" according to Schoenberg) – setting up a dissonance with one or both of the previous pitches. In effect, not only is a mode not established, but the negation implies that the process of establishing mode must begin again – starting again from either the first or the second pitch (i.e., re-initialization to some earlier stage in harmonic development and implicitly declaring earlier pitches to be false starts irrelevant to subsequent developments). Musically, the "failure" to establish harmony is often desirable as a means of delaying the resolution until a certain level of tension and uncertainty has been created. In other words, to jump immediately to the modal ending of the story takes all the excitement out of the narrative, and is often avoided, but ultimately most diatonic music and virtually all popular music proceeds – after a variety of false starts, dead ends and contrary hints – toward modal resolution. When a third tone does not proceed directly to the establishment of mode, it can nonetheless have two distinct effects. It can create interval dissonance with either of the two preceding tones or it can establish harmonic "tension".

The possibilities that a third tone brings with regard to ultimate harmonic resolution are outlined in Figure A1-3.

What is essential to note in Figure A1-3 is that the pattern of establishing mode is nothing more than a summary of the possible contributions of

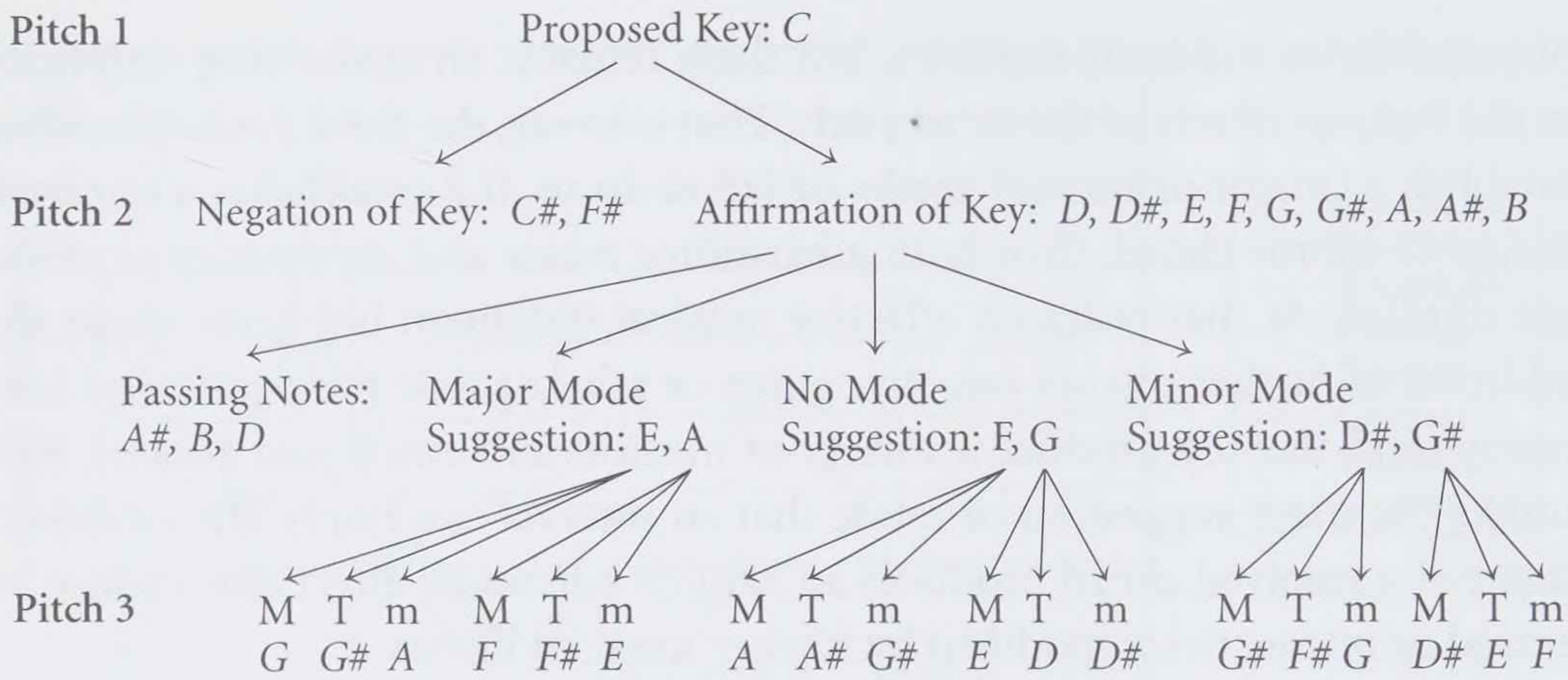


Figure A1-3. The roles of three consecutive pitches in the establishment of harmonic mode. Letters in italics indicate pitches, relative to the initial pitch, *C*. Among the effects of the second pitch, the so-called passing notes remain within the proposed key, but constitute tonal dissonances that cannot lead directly to harmonic resolution; they have implications with regard to mode, but, similar to pitch repetitions, they delay but do not negate the proposed key. Musically, the "passing notes" are incidental to the development of mode, and can be neglected. The effects of the third pitch that leads to a harmonic statement are noted as: "M" (establishment of the major mode), "m" (the minor mode), and "T" (the unresolved tension chords). In addition to the three unambiguous harmonic effects (MTm) that can follow the second pitch, there are 8 other diatonic tones (per second pitch) with various musical implications. A complete theory of musical emotions would require their enumeration, but will be omitted here.

individual pitches to harmonic states, as understood in conventional diatonic music theory. Some tones function to support the previous tones in establishing a self-consistent key and either a major or minor mode, and some do not. This is not an argument about what makes for "good" music or what "should" be done in musical composition; the figure is simply a statement about the known relationships among tones in a diatonic scale (when delivered sequentially as melody). The third pitch can produce a dissonance with a previous pitch or pitches and prevent the arrival at a harmonic statement or, alternatively, allow such a statement and establish a harmonic affect. Among harmonic statements, there are the resolved (major and minor) triads and the unresolved (diminished and augmented) triads.

In summary, Figure A1-3 implies a set of tone combinations that play specific roles in the establishment of major and minor modes. Provided only that the major and minor modes have (for some reason, possibly related to the use of pitch in animal calls, as described in Chapter 3) an emotional valence

for normal human listeners, then it can be concluded that the tonal combinations that lead in three distinct stages toward the production of major or minor chords are the essential auditory steps leading up to a resolved mood of either positive or negative affect. If the establishment of a major mode has an evolutionarily-deep connotation of "happiness" in the sense of denoting social dominance, and a minor mode "unhappiness" in connoting social submission, then it is a straight-forward task to delineate the individual affective tonal steps along the way to such resolved moods.

As such, a rather simple dichotomous theory of musical emotions has been arrived at – a theory in which positive emotions are associated with the feelings inherent to social dominance, and negative emotions associated with subordinance. (An in-between state of harmonic tension is also implied, but it is by definition a state of vacillation or ambiguity between the poles. Harmonic tension lacks the fixedness of major and minor chords, but is nonetheless itself an explicit harmonic state: a condition of uncertainty and anxiety that is unresolved with regard to a happy or sad affect.) Many theories of emotion that include a happy-sad dimension include subcategories of the positive and negative poles (e.g., happy, joyous, elated, ecstatic versus sad, mournful, depressed, bored that differ with regard to activation, ego-involvement, etc.), but the pattern outlined in Figure A1-3 suggests something quite different: that there are a variety of transitional states *prior to* arriving at the resolved positive or negative emotions or the explicit statement of unresolved tension.

Note that the steps shown in the figure are clearly not an exploration of the social dynamics underlying emotions, but simply enumeration of the tonal patterns that lead, step-by-step, to major, minor or tension chords. In principle, it might be possible to elaborate on the nature of the corresponding types of social interactions that lead up to social dominance or subordinance – events that precede the resolution of a tension in either a positive or a negative way. In such a psychological theory, the magnitude of the positive or negative emotions inherent to resolution would be a function of the individual's involvement in the event. If there occurs a whole-hearted struggle implying heavy losses or substantial gains, the emotional magnitude of the win or loss will necessarily be great (and definable with different adjectives that indicate the strength of the emotion). But the qualitative nature – the positive or negative character of the emotion – does not change with quantitative differences. Suffice it to say that, in the present theory of musical emotions, the magnitude of emotions will be strongly context-dependent.

Nothing more needs to be said about the extremes of "resolved positive emotion" and "resolved negative emotion" – other than that they are the psy-

chological effects (strong, weak or middling) that are felt as a result of a (large, small or medium) "victory" or "defeat" (as commonly understood in terms of the typical social dynamics among people). All other aspects of emotion remain rather nebulous and are context-dependent, but victory and defeat are universally known – at one scale or another – and are associated with characteristic emotions. For this reason, further description of the "meaning" of dominance and submission, or positive and negative feelings, is pointless; it is given a priori on the basis of human experience. Simply stated, everyone understands the affective states of "win" and "lose" on the very small scale of personal experience (initially intrafamilial, childhood social interactions and gradually expanded into ever-larger social contexts). The poles are known and are not in need of explication, but there is still a need for a better understanding and a more detailed description of the possible states that lead up to victory or defeat. It is these transitional states that provide the subclassification of a tonal theory of emotions.

If it is assumed as a starting premise that, underlying all social interactions, there are dominance relations at work, and therefore that, among all the buzz and confusion of complex social phenomena there is a bottom line of power, then every social encounter can be evaluated in terms of increasing or decreasing power. As small as any one move on the chessboard of social interactions may be, each move can then be understood as a small step in the direction of greater or lesser social dominance. With "world domination" at one extreme and a decrepit, insipid rejection from all social contact being the other (equally unimaginable) extreme, our daily lives can be seen as normally involving small adjustments of dominance relations in the lukewarm hospitality of familiar people. So, without jumping to the heroic extremes of ecstasy and despair, it is in principle possible to outline what kinds of steps can be taken in either one direction or the other. Again, there is no need to elaborate on the types of situations that lead to the affective extremes of social dominance/submission, but there is a need to delineate what set of states can be generated in the process of establishing major chord "dominance" or minor chord "subordinance".

The underlying contention of this model of emotions is that each of the precursor states (distinguished on the basis of the known pitch relations in music) is a definable emotion. If we experience an emotional state in the context of a world war or defending a loved one, the emotion will be strong and memorable, but even at the much smaller scale of tricking a younger brother into taking out the garbage, receiving praise for a job well done, or taking blame for a minor mishap, in each and every case there is a definable, recognizable, human emotion that we have all experienced and that, on the basis of which,

allows us to empathize with the grander emotions that happen to famous and notorious people on the world scene. In a word, emotional states are universal constants that are reflected in the regularities of harmonic phenomena.

The establishment of a chord requires, at a minimum, three tones. The first tone is, by definition, a proposal to build a resolved harmony with that tone as the tonic. Emotionally, it is a bold assertion – the sounding of a tone in a noisy, but neutral auditory universe. Since no event is thoroughly context-free in an ongoing, existing world, the proposal of a tonic, the opening salvo itself, can be described as implicitly containing some level of daring and boldness; it is a first statement, suggesting that something (musical) can be built where nothing exists.

The next tone is, by definition, sounded in relation to the first tone. Without again going through the full set of possibilities in Figure A1-1, the second tone can be either affirmation of the possibility of a resolved chord based upon the previously sounded tonic, or it can be negation of that possibility.

If the second tone is negation of the tonic, it is, from the perspective of the original tone, a conflicting, troublesome challenge to the boldness of the original tonic assertion. Since the second tone negates the first tone and is itself also a new assertion, it shows a boldness in *not* following the direction of the original tonic. If, however, the second tone is affirmative of the tonic, it is explicit support of the original tonic assertion. In this case, the second tone is affirmative of a possible harmonic resolution in conjunction with the first tone.

Depending on the interval relationship between the original tonic assertion and the second tonal affirmation of key, the proposed chordal resolution implied by the second tone can be any of three possibilities – (i) inherently a suggestion for a minor resolution, (ii) inherently a suggestion for a major resolution, or (iii) inherently ambiguous with regard to subsequent modal development. In all three cases, the second tone "joins with" the first tone in establishing an interval that can be a part of a harmonic statement. However, the possibilities for resolution implied by the second tone put clear constraints on the nature of the subsequent harmonic development. Depending on the nature of the second tone, the suggested resolution may be either major or minor, with the emotional implications that those two types of resolution bring, or it can be a tone that does not indicate a preference between major and minor.

In any of these three cases, the addition of a second tone amounts to nothing more than a *suggestion* for an upcoming resolution – and is not itself definitive. With the exception of the two tones that act to deny the original assertion of a tonic, the second tone (one of nine remaining, non-repetitive possibilities) leaves entirely open the question of whether the third tone will result in major

resolution, minor resolution, or proceed to a non-resolved option. The choice of the second tone does, however, imply a certain modal stance of the second tone in relation to the first tone. That is, there is an affective significance of the second tone in choosing either to be supportive or unsupportive of the key implied by the first tone and, if supportive, then suggestive of a major or minor resolution. In so far as it can be assumed that the original assertion of the first tone is, by default, the proposal of a tonic for the eventual resolution of a major chord, then the suggestion offered by the second tone with regard to possible resolution will be either fundamentally supportive of the major key or rebellious in suggesting a minor key.

More precisely, the choice of a "minor" interval (3 semitones – a minor third, 8 semitones – a minor sixth, or 10 semitones – a minor seventh) implies that resolution would most easily, most naturally (within the implicit context of diatonic music) occur at a minor chord. In contrast, the choice of any of the "major" versions of these same intervals (4, 9 and 11 semitones) implies a major mode resolution. In addition, there are three noncommittal intervals that have no inherent modal force (intervals of 2, 5 or 7 semitones – a second, a fourth or a fifth – all three of which are equally consistent with major or minor chord resolution).

In any one of the three scenarios (suggestive of major, minor or either form of resolution), the second tone is not definitive with regard to the nature of the chordal resolution, but, depending upon the interval created by the second tone, it can leave open the possibility for immediate major or minor resolution (by adding a subsequent third tone), or not. The tones that do not leave the possibility of resolution open require that the third tone "restart" the chord building – either returning to the stage of an opening tonic assertion or – rejecting that – building an interval from the second tone that can then subsequently be resolved. In both of these latter cases, the unsupportive character of the second tone implies that the process of harmonic development must return to an earlier stage. For this reason, such unsupportive second tones are, by their very nature dissenting, denying, rebelling, contrary-minded assertions, relative to the first tone.

In each of these scenarios there is a tonal implication of the relation between the first and second tones. If the possibility of resolution has been left open by the choice of the second tone, then the third tone itself can proceed directly to a major or minor chordal resolution, or – once again – proceed to something else. Emotionally, the significance of major or minor resolution is clear, but the third tone can also be such that one or more additional tones is required in order to construct a resolved harmony. That is, the sounding

of the third tone can be the assertion of harmonic resolution, or it can be delay of same.

The three types of *suggestion* regarding resolution that the second tone can bring (layer two in Figure A1-3) have already been considered – that is, the possibilities tending toward major resolution (left part of layer two), tending toward minor resolution (right part of layer two) or allowing either form of resolution (middle part of layer two). In all three cases, the affective power is mild because the interval created by two pitches alone *does not establish* mode; in other words, two-tone intervals in music are sublexical affective units, whereas the addition of a third tone establishes a lexical affective unit, a harmonic chord that carries intrinsic emotional force.

Starting with the left-hand set of "suggestions" (in Figure A1-3), in response to the original assertion of a tonic, the second tone can be a tone that lies at a major key interval from the tonic – and therefore would most naturally be resolved as a major chord by sounding a subsequent third tone. The most appropriate phrase to describe the evolving affect of three steps headed in the direction of a major chord is confident, unswerving, self-assured happiness (CEG, CAF). That is, such chords are examples of a tonic followed by a tone which forms a major key interval, followed by a third tone that establishes a major chord – resolution of a direct, three-step process that leads to a positive harmonic affect. Let us call that "joy".

A variant on that three-step process of establishing a major chord is the choice of a third tone that turns the suggested major key into a minor chord resolution. This emotion is not simply the negative affect of minor chords in general; it is also the surprise sadness that follows upon two preceding tones, both of which had indicated the likelihood of the positive affect of major chord resolution. For this reason, the negative, dark emotion of such a chord has also an element of a sudden, unexpected, unfortunate turn of events, and might be described as the emotion of de-spiritedness, or "despair" (CEA, CAE).

The third possibility following two potentially major chord tones is to choose not to resolve the chord as either the expected major or as the unexpected minor, but to add a third tone which sets up an unresolved tension triad. The tension chord remains unresolved until further tonal adjustments are made, and is explicitly neither major nor minor. Unlike an *interval* that is ambiguous as to major or minor harmony due to its underdeterminedness, the tension triad is itself a harmony – an explicit statement of relations among three tones, but it is not a resolved chord. A third tone that produces an unresolved harmony means that the positive and negative affect of the major and minor chords is actively avoided, so that the implied emotion is more complex than

ambiguous interval; it also has an element of unexpectedness (similar to the sudden resolution in a minor chord), but remains ambivalent with regard to final outcome, and ultimate harmonic mode. It corresponds to the emotion of hesitancy and the purposeful delay of (the anticipated) major chord resolution until further tones are sounded. It is not simply unresolved, but is a harmony that is an active retreat from the simple and obvious resolution into a major chord (CEG#, CAF#).

A second set of scenarios corresponds to the possibilities following the second tone's suggestion of a *minor* key; these are shown on the right-hand side of layer two in Figure A1-3. Again, the first tone (C) is, by default, a proposal to proceed in a certain key. If the second tone forms a so-called minor interval (D# or G#), it constitutes a mildly provocative, rebellious *suggestion* to proceed in a minor mode. The second tone can then be followed up in three ways. First of all, the interval can be resolved as a minor chord by the appropriate selection of the third tone (CD#G, CG#F). The sequence of events can then be summarized as: proposal of a key that is major (by default), choice of a minor interval (suggesting a minor mode that is inherently supportive of the originally proposed key, but rebellious in suggesting a minor direction), and then resolution to a minor chord. In this case, the third tone's affirmation of the minor "reversal" suggested by the second tone entails abandonment of the original major mode direction and acceptance of the minor mode suggestion of the second tone. This development might be described as the emotion of a fatalistic embracing of sadness (CD#G, CG#F).

In contrast, in response to the suggestion of a minor mode by the second tone, the third tone can be such that it "ignores" that suggestion and insists on a major chord resolution. This is the emotion of defiant optimism – effectively and definitively rejecting the suggestion for minor resolution and plowing on to major resolution regardless of the contrary mode suggested by the second tone (CD#G#, CG#D#).

Finally, the suggestion for minor chord resolution can be mildly rejected by a choice of the third tone that creates a tension chord; in so doing, the issue of chordal resolution and the establishment of mood is delayed by neither accepting the minor chord suggestion nor entirely overriding it. This is the "diplomatic" choice, corresponding to the state in which major chord resolution has not been insisted on, but neither has the suggestion for minor chord resolution been agreed to. In effect, the third tone has not demanded finality immediately with the sounding of the third tone. Instead, by moving to an unresolved chord, the unresolvedness of the situation has been made explicit. This emotion is one of frank uncertainty and anticipation, for it neither backs

away from the initial assertion (succumbing to minor key resolution), nor vetoes the contrary suggestion with a forced resolution in major; instead, it leaves the issue of chordal resolution open to further tonal input (CD#F#, CG#E).

The final set of third tone possibilities are in response to the second tone suggestions that do not imply a preferred mode (the middle part of layer two in Figure A1-3). Second tones with no modal suggestion are inherently deferential to the original proposal of a tonic; they accept the proposal of key implied by the first tone, but they are truly noncommittal with regard to the direction of chordal resolution. In response to the second tone showing neither support nor denial, the third tone has again three options. It can be sounded such that it immediately terminates the sequence and chooses the major chord resolution – essentially acting without the active consent or dissent of the second tone. This corresponds to the emotion of cool, level-headed confidence, but lacks the assuredness of a major chord developed with the active consent of the second tone (CGE, CFA).

Alternatively, the third tone can exhibit a "change of heart" and produce a minor resolution – acting entirely arbitrarily and without reference to other tonal suggestions, by intoning a pitch that turns the original proposition of a major key, followed by a non-committal assent, into an unexpected minor key resolution. (CGD#, CFG#). This corresponds to the emotion of a loss of will and sudden despair, that is, "hopelessness." The negative affect of these sequences is more sudden that that of the minor chords discussed above (CD#G, CG#F) because the "darkness" of the latter triads (CGD#, CFG#) was not anticipated in a preceding minor interval.

The final option is that in which the third tone does not proceed on to (unsupported) major resolution, and does not retreat into (uncalled for) minor resolution, but opts to delay resolution until further tones are sounded. This is the creation of a tension chord (CGD, CFA#), in spite of the fact that resolution would have been possible. This option corresponds to the emotion of lack of conviction, or uncertainty, as the original proposition of a tonic for a major key does not proceed to resolution; without active commitment of the second tone, the easy and natural final step toward major chord resolution is not taken, thus leaving the harmonic state in one of uncertainty.

Further elaboration of the affect (due to differences between diminished and augmented chords, and due to chordal resolution achieved in sequences of four, five or more tones) could be pursued, and the affective meaning of tonal combinations that include dissonances could be similarly enumerated. While the choice of adjectives to describe the affective states may be somewhat idiosyncratic, there are generally common phrases for describing the emotions

that arise in the progression toward happy or unhappy outcomes – and those words can be applied in relation to the progression toward chordal resolution and the positive or negative affect of major and minor chords. Suffice it to say here that, if it is assumed that three-pitch combinations are the essence of harmony (the dualist assumption), then there is a finite number of pitch combinatorial states that stand in explicit relation to such modal resolution. If the major/minor dichotomy has an inherent link with positive/negative affect, then the entire pattern of pitch combinations in Figures A1-1 and A1-3 necessarily has an analog in "affective space" – where each combination of two or three tones plays an explicit role in relation to the sequential process of major and minor chordal resolution.

Sounded simultaneously as harmonic chords, there are three classes of (non-dissonant) chords, major, minor and tension (with several configurations of each). The affect of these three classes can be stated simply as positive, negative and uncertain. Sounded sequentially as melodic chords, the sequence of tonal events can be broken down into several subcategories, depending on the implications that each tone has in relation to the key and mode implied by the preceding tones. Assuming only that major and minor have affective significance for most people, the above subclassification of emotions inevitably follows in light of what is known about diatonic music. Clearly, real music that is more than combinations of 2 or 3 tones will contain such a huge number of such affective implications that it would be of questionable value to summate the affect of its many chords and intervals, but the affective power of music in general and the importance that music has in the lives of many people may be explained precisely in the fact that even rather short musical pieces create, manipulate and rearrange a large number of pitch events that elicit affective responses of varying strength, varying expectedness, and varying valence.

The entire discussion above concerning three tone chords and emotions can be summarized as in Table A1-1. Several points concerning the table are worth making. The first is that the same tones in different sequence lead to somewhat different emotions. Although the same pitches are of course a part of identical major, minor or tension harmonies, when they are played as sequences, the second tone has a meaning relative to the first, and the third tone has a meaning relative to the interval created by the first two tones. Again, the choice of adjectives can be disputed, but the underlying argument is simply that, since the positive or negative emotional valence of resolved major and minor chords is known, the emotional qualities of the three tones that lead to one form of resolution or the other can be delineated. What the second tone implies in relation to the first tone is either (i) an indication of the preferred

Table A1-1. The musical emotions associated with sequential three-tone combinations. "M" denotes the sounding of a major interval or chord, "m" denotes a minor interval or chord, "n" denotes a neutral interval that lacks indication of mode, and "T" denotes a tritone interval or tension chord. For each description, there are 2 or 3 equivalent tone sequences that differ in the magnitude and/or direction of the sequences. Twice as many could be generated by considering the combinations possible when the second tone falls below C.

Tone sequence	Ultimate valence	Affective sequence	Description of the process of establishing or avoiding modality
CEG	Positive	MM	Confident procession from major interval to major chord.
CAF	Positive	MM	
CD#G#	Positive	mM	Defiant rejection of minor interval in establishing major chord.
CG#D#	Positive	mM	
CGE	Positive	nM	Persistent establishment of major chord following a modaly-non-committal interval.
CFA	Positive	nM	
CEA	Negative	Mm	Loss of confidence in abandoning a major interval to establish a minor chord.
CAE	Negative	Mm	
CD#G	Negative	mm	Resignation to a minor chord following a minor interval.
CG#F	Negative	mm	
CFG#	Negative	nm	Active retreat to a minor chord following a neutral interval that could have resolved to major.
CGD#	Negative	nm	
CAD#	Tension	MT	Active retreat to chordal ambiguity following a major interval that could have been resolved as a major chord.
CAF#	Tension	MT	
CEG#	Tension	MT	
CD#F#	Tension	mT	Assertion of chordal ambiguity following a minor interval that was headed toward minor chord resolution.
CD#A	Tension	mT	
CG#E	Tension	mT	
CGD	Tension	nT	Affirmation of chordal ambiguity following a neutral interval that suggested neither major nor minor mode.
CFA#	Tension	nT	
CGA#	Tension	nT	
CF#A	Tension	TT	Emphasis on ambiguity in the establishment of a tension chord following the inherently unresolvable tritone interval.
CF#D#	Tension	TT	
CF#C'	Tension	TT	

(major, M, or minor, m) resolution, or (ii) an indication that neither form of resolution is preferred (modal neutrality, n, or mode-defying tension, T). What the third tone implies as it completes a non-dissonant chord is an unambiguous statement of positive, negative or ambiguous affect emerging from the preceding interval. The final mood is thus a consequence of (i) the role of the third tone in proceeding to major, minor or tension harmony, in light of (ii) the modal direction that has just been provided by the second tone.

The second point concerns the previous work of Cooke (1959), who described the affect of 3, 4, and 5 tone sequences in classical music in a similar way (see also Kaminska & Woolf 2000). His argument was essentially historical – arguing that certain sequences of tones had emotional qualities that had been used in a variety of classical works relating principally to Western mythology and love themes. Although his descriptions are somewhat similar to those in Table A1-1, his tonal justification of the affect was a combination of musical intuition and historical usage. In contrast, the argument of this Appendix is more simply that the *known* roles of individual tones in producing a major, minor or tension harmony leads to a rudimentary, but unambiguous description of what any tone (in relation to other tones) can and cannot do in a modal sense. In that respect, the logical deduction of the affect of three-tone combinations (based upon the dualism of major and minor harmonies) differs from the historical argument provided by Cooke, but in fact the conclusions are similar.

Schoenberg (1911) and other modern composers fought against the "tyranny" of the diatonic scales and their implied "correct" and "incorrect" tonal members. The fact that most normal listeners nonetheless hear the indelible character of the major and minor triads (and their implications for what other notes can or cannot be used in combination with them) is indication of how strong that "tyranny" is. Whether the tyranny is learned or innate, should be resisted or accepted in a musical sense, is a musical "law" or merely a bad habit, is not of relevance here. What is important in the present context is simply that normal listeners in the modern world experience the "appropriateness" or "inappropriateness" of certain tones in a given musical context. In other words, diatonic keys are indeed tyrannical – and spread their tyranny through virtually all popular music, the bulk of traditional classical music, and perhaps most importantly all advertising jingles; as a consequence of that tyranny, we know "instinctively" when something is in or out of tune. Schoenberg and others versed in chromatic scales would agree that the diatonic scales have become ingrained in most listeners, but would argue that this is not inevitably so and is a consequence of historical serendipity. They would maintain that it is both possible and desirable to train oneself *not* to be swayed by the regularities inherent to diatonic scales and the pull of major and minor harmonies. But without such specific training, the diatonic scales used in diverse ways in the diverse musical cultures of the world (with the exception of Schoenberg and friends) tell us quite explicitly when a major or minor key has been established. Whether this is a sign of musical naïveté or indication that we have come to an understanding of the patterns of diatonic music is debatable, but there are indeed patterns inherent to diatonic music.

The acoustic effects of the upper partials mean, at the very least, that the diatonic scales are not unnatural – i.e., not an artificial cultural artifact that the Western world has forced down the unwilling throats of its many listeners and the diverse indigenous musical cultures throughout the world. Diatonicism is one of a small number of possible musical styles in musical cultures that use vibrating objects in combination with one another – perhaps not unique, but certainly not an arbitrary invention enforced by anything other than the physics of vibrating bodies.

In a remarkable book that explores the significance of the upper partials, Sethares (1999) has shown that non-diatonic scales can be produced and musically enjoyed in ways that Schoenbergian intellectualisms cannot be – solely by making appropriate changes to the upper partials of the fundamental frequencies. Using computer technology, the harmonics of vibrating bodies can be adjusted and upper partials that are stretched or shrunken multiples of the fundamental frequency can be easily produced. In effect, what that allows for is non-diatonic music that maintains the consonance of upper partial intervals. In principle, similar modifications could allow for non-diatonic music that has all of the complexity of Western harmonic music, but without the "tyranny" of the 12-tone scale.

However we may regard the acoustical foundations of diatonic music, the regularities that it embodies and that have been elaborated on in not only Western classical music, but in the popular music of many cultures have come to be associated with certain affective states. How those tonal regularities may emerge in the voice intonation of normal speech is the difficult issue raised in Chapters 4 and 5, but the affective content of diatonic music is relatively straight-forward. Assuming only that there is a recognizable affective polarity in the major and minor modes – the roles that individual tones play in establishing such "positive" and "negative" moods is known with the full certainty of traditional harmony theory.

Appendix 2

Calculating harmoniousness

The sonority of isolated intervals (with or without upper partials) can be calculated using dissonance curves (e.g., Kameoka & Kuriyagawa 1969; Sethares 1999) or using more sophisticated psychoacoustical models (e.g., Terhardt et al. 1982a, b; Parncutt 1989). For computational simplicity the following model will be used here:

$$D(x) = \mu_A \cdot \exp(-ax) - \exp(-bx) \qquad \text{(A2-1)}$$

where μ_A is the mean amplitude of the two tones (0.0~1.0), x is the interval size (in semitones) and a, b and c(1.20, 4.00, 3.53) are chosen to give a maximal dissonance value at roughly one quartertone, a dissonance of 1.00 at an interval of one semitone, and smaller values for larger intervals.

As discussed in Chapter 3, at least one further factor is required to account for the basic facts concerning the perception of chords, i.e., a "tension" factor related to the relative size of intervals contained within the chord. For this purpose, a more general model of pitch "instability" that includes both a two-tone component and a three-tone component will be introduced. The total "instability" of multiple tones is then defined as the summation of all two-tone dissonance (D) effects and all three-tone tension (T) effects. Higher order effects remain possible, but, as in theoretical physics, they are likely to contribute relatively little beyond the dominating two- and three-body effects.[1] That is, a model of pitch instability (I) can be stated as:

$$I_{\text{n-tones}} = D_{\text{intervals}} + k\,T_{\text{chords}} \quad (n \geq 3) \qquad \text{(A2-2)}$$

where the first term on the right of Eq. A2-2 corresponds to a model of interval perception (e.g., Eq. A2-1), and the second term is a parametrization of the perception of three-tone chordal tension (Eq. A2-3) (see Figure A2-1). The shape of the dissonance and tension curves and the relative strength of these factors ($k \sim 0.1$) must be determined empirically.

From an experiment on three-tone chord perception using a 24-tone scale, it was found that the harmoniousness of chords containing quartertone intervals (e.g., 3.5 and 4.0 semitones) was approximately midway between the harmoniousness of the resolved and unresolved chords (Figure A2-1B). The parametrization of the chordal tension, T, can then be expressed as:

$$T(x) = \mu_A 2\mu_x \cdot \exp(-(|x1 - x2|/d)^2) \qquad \text{(A2-3)}$$

where μ_A is the mean amplitude of the three tones, c is a constant (0.6), $x1$ and $x2$ are the magnitudes (in semitones) of the two smaller intervals in each three-tone combination of pitches, and μ_x is the mean value of $x1$ and $x2$.

The model defined by Equations A2-1 through A2-3 and illustrated in Figure A2-1 provides a means for calculating the total instability of any combination of pitches, unrelated to the specific musical scale or tuning system. It is explicitly designed to reproduce the musical "common sense" concerning the relative stability of dissonant, unresolved and resolved chords (Figures 3-5, 3-10 and 3-11). In effect, the model allows for a psychophysical understanding of the stability of resolved chords relative to the instability of unresolved chords that is more closely in line with perceptual findings than is possible using only interval effects (compare the results in Table A2-1A and A2-1B).

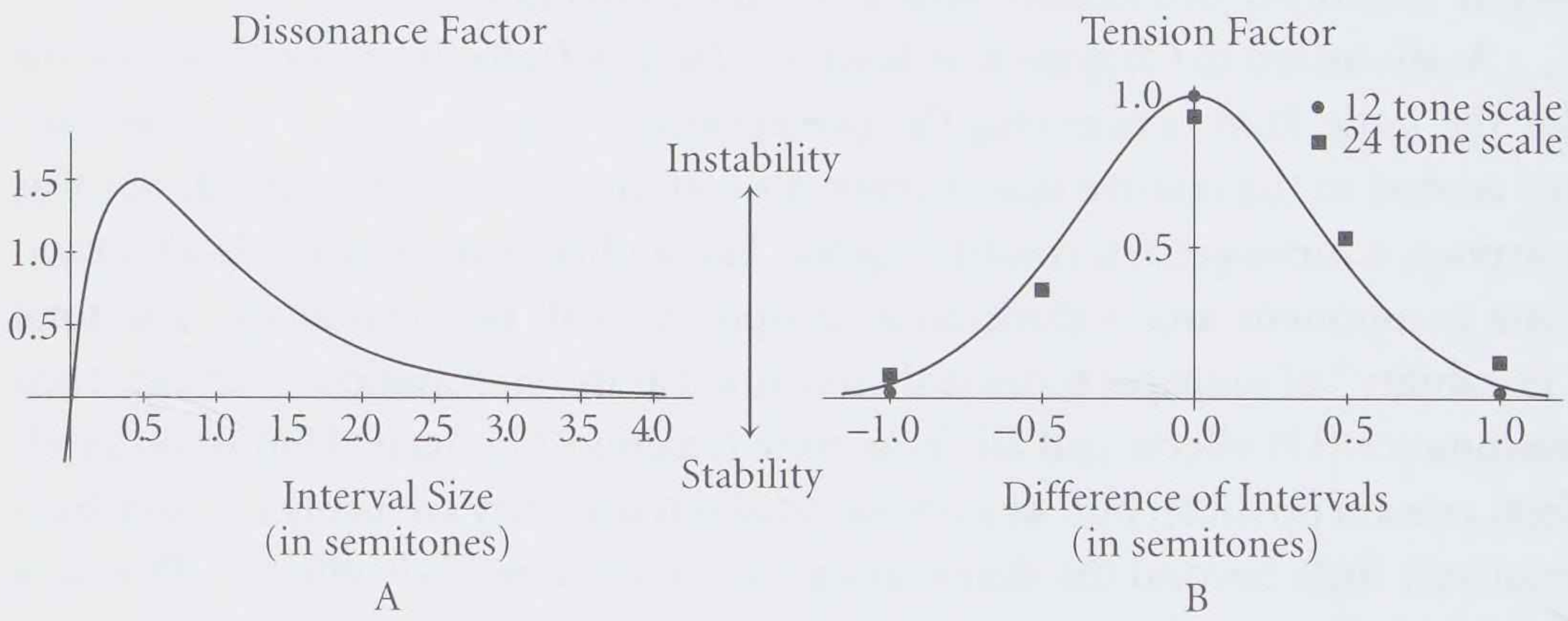


Figure A2-1. The two components contributing to the "instability" of pitch combinations. (A) The interval dissonance factor (Eq. A2-1). (B) The chordal tension factor (Eq. A2-3). The data points are from 12-tone and 24-tone scale chord experiments in which the "harmoniousness" of various three-tone chords were evaluated. Major and minor chords containing tones that are one quartertone flat or sharp (–0.5 and 0.5 on the x-axis in the figure) were evaluated as somewhat less sonorous than the resolved chords, but somewhat more sonorous than the diminished and augmented chords (Cook et al. 2002b). (Thanks to M. Okada for 24-tone scale experimental work.)

Table A2-1. A comparison of the theoretical "harmoniousness" of the 12 major, minor, augmented and diminished triads in various configurations (compare with Figures 3-10 and 3-11). The values shown in (A) were calculated considering only interval effects (Eq. A2-1); similar results would be obtained with other interval-based models. The results in (B) were calculated considering both interval and chordal effects (Eq. A2-2). Calculations were made with 1-3 upper partials (relative amplitudes decreasing as 1.00, 0.50, 0.33, 0.25). The ranking of the chords from the most harmoniousness to the least harmoniousness is shown in brackets for each of the columns. Noteworthy is the *inability* of interval dissonance alone to distinguish between the resolved and unresolved chords (A1, A2 or A3); inclusion of a three-tone chordal factor produces the expected distinction provided that some upper partials are included in the calculations (B2 and B3).

Chord type	Chord configuration	A			B		
		1	2	3	1	2	3
Major	Root	0.13 [10]	0.20 [8]	0.29 [6]	0.21 [7]	0.30 [5]	0.41 [1]
Major	1st Inversion	0.11 [8]	0.18 [6]	0.76 [10]	0.11 [3]	0.22 [1]	0.81 [5]
Major	2nd Inversion	0.04 [1]	0.12 [4]	0.43 [1]	0.15 [5]	0.29 [3]	0.80 [4]
Minor	Root	0.13 [10]	0.20 [8]	0.29 [11]	0.21 [7]	0.35 [6]	0.45 [2]
Minor	1st Inversion	0.04 [1]	0.12 [4]	0.43 [3]	0.15 [5]	0.29 [3]	0.78 [3]
Minor	2nd Inversion	0.11 [8]	0.18 [6]	0.76 [4]	0.11 [3]	0.24 [2]	0.83 [6]
Diminished	Root	0.20 [12]	0.30 [12]	0.79 [12]	1.40 [9]	1.80 [9]	2.49 [9]
Diminished	1st Inversion	0.10 [6]	0.21 [10]	0.80 [5]	0.10 [1]	0.57 [7]	1.16 [7]
Diminished	2nd Inversion	0.10 [6]	0.21 [10]	0.80 [2]	0.10 [1]	0.66 [8]	1.25 [8]
Augmented	Root	0.06 [3]	0.11 [1]	0.61 [7]	1.66 [10]	3.19 [10]	3.96 [10]
Augmented	1st Inversion	0.06 [3]	0.11 [1]	0.61 [7]	1.66 [10]	3.19 [10]	3.96 [10]
Augmented	2nd Inversion	0.06 [3]	0.11 [1]	0.61 [7]	1.66 [10]	3.19 [10]	3.96 [10]

Two points about the model should be emphasized. The first is that, for use of Equation A2-3, the interval structure of a 3-tone chord must be "normalized" to span one octave or less in order to avoid unrealistically large tension values for chords that extend over several octaves. Normalization of the triads is well-known and described with the terms pitch-class and octave equivalence, and has some justification from experiments on the musical effects of intervals (e.g., Costa et al. 2000). Because of the finite size of the 12-tone octave, in effect, all normalized chords will have a pitch range of 11 semitones or less, and the difference in the size of the two smaller intervals in the chord will range from 0 to 10 semitones. All "tension" chords in root position will have an interval difference of 0-semitones; all major and minor chords will have an interval difference of 1–2 semitones; and all "dissonant chords" will contain within them at least one dissonant interval of one- or two-semitones.

The second point is that the μ_A factor in Eq. A2-3 is included expressly for the purpose of obtaining a higher instability value for the augmented than the diminished chords (Tables A2-1 and A2-2). The underlying idea is that the relatively greater instability of the augmented chord is a consequence of the perceptual clarity of the "intervallic equidistance" relative to that heard in the diminished chord or the tone clusters. Despite the fact that the intervals of the augmented chord are large and therefore individually consonant, the three-tone tension effect is perceptually strong. In contrast, the small interval dissonance in the diminished chord or tone clusters may mask the three-tone tension effect. Further experimental work is needed here.

In effect, the normalization procedure reduces all 3-tone combinations to three classes of chord (resolved, tension and dissonant). Calculation of the total instability of any chord will then result in: (i) a large *dissonance* factor for the chords containing dissonant *intervals* (regardless of the stability/instability of the chordal factor), (ii) a large *inharmoniousness* factor for the tension chords (regardless of the consonance/dissonance of the interval factor), and (iii) low dissonance and low inharmoniousness factors for the resolved major and minor chords. The only exceptions are the diminished chords in inverted positions when only the fundamental frequency is considered (Table A2-1, Column B1). When upper partials are included in the calculation, however, the tension factor among various combinations of F0, F1, etc. tones introduces a significant tension component that results in large instability values for these chords as well.

As to the relative sonority of the various chord inversions, the results from non-musicians (Figure 3-11a) indicate clear distinctions, whereas those from musicians (Figure 3-11b) do not. Since it is likely that the musicians in such an experiment gave responses in accordance with their identification of the major/minor mode of the triad, rather than a naïve sonority rating, it may be

Table A2-2. The total "instability" of several types of chords, as calculated using Eq. A2-2 and three upper partials (F0~F2). Note that a clear numerical distinction between the resolved and unresolved chords is obtained.

Chord	Type	Total instability
C -E-G#	Augmented	3.96
C-C#-G	Semitone dissonant	2.97
C-D#-F#	Diminished	2.49
C-D-G	Whole-tone dissonant	1.49
C-D#-G	Minor	0.45
C-E-G	Major	0.41

better to believe the non-musicians. In either case, slight modifications of the dissonance or tension factors and/or the strength of the upper partials would be needed to achieve perfect agreement with experimental data. In any case, the model, as it stands, is an improvement over interval-based models.

The model for pitch instability discussed above provides a straight-forward explanation of the most important regularities of traditional diatonic harmony and is consistent with conventional ideas in music theory. The potential importance of the model, however, lies in its implications for pitch phenomena that are not based on the diatonic scales (Chapter 4).

Notes

1. For example, there may well be a unique affective quality of four-tone chords, but the four-body effect will necessarily be heard simultaneously with all combinations of two-tone and three-tone effects. Assuming that the four-tone effect would have a magnitude comparable to the individual two- and three-tone effects, it would be numerically swamped by the 6 two-tone combinations and 3 three-tone combinations within the four-tone chord – constituting only 10% of the affective quality of the chord. The corollary with regard to intervals and three-tone chords is that the three intervals of a three-tone chord provide 75% of the affective quality of triads – whereas the quality of the three-tone combination itself contains 25% of the total effect, and pushes the combined consonance of the contained intervals toward harmonic tension or toward the positive or negative resolution of major and minor chords.

References

Abeles, M. (1982). Role of the cortical neuron: integrator or coincidence detector? *Israel Journal of Medical Science, 18*, 83–92.

Abercrombie, D. (1967). *Elements of General Phonetics*. Chicago: Aldine Publishing.

Alcock, K. J., Wade, D., Anslow, P., & Passingham, R. E. (2000). Pitch and timing abilities in adult left-hemisphere-dysphasic and right-hemisphere-damaged subjects. *Brain and Language, 75*, 47–65.

Aleksander, I. (1996). *Impossible Minds: My neurons, my consciousness*. London: Imperial College Press.

Alivisatos, B., & Petrides, M. (1997). Functional activation of the human brain during mental rotation. *Neuropsychologia, 35*, 111–118.

Anaki, D., Faust, M., & Kravetz, S. (1998). Cerebral hemispheric asymmetries in processing lexical metaphors. *Neuropsychologia, 36*, 353–362.

Anderson, J. A. (1995). *An Introduction to Neural Networks*. Cambridge, MA: MIT Press.

Annett, M. (1985). *Left, Right, Hand, and Brain: The right shift theory*. London: Erlbaum.

Arnoldi, H.-M. R., & Brauer, W. (1996). Synchronization without oscillatory neurons. *Biological Cybernetics, 74*, 209–223.

Baars, B. A. (1988). *Cognitive Theory of Consciousness*. New York: Cambridge University Press.

Baars, B. A. (1997). *In the Theater of Consciousness: The workspace of the mind*. New York: Oxford University Press.

Ballanyi, K., & Kaila, K. (1998). Activity-evoked changes in intracellular pH. In K. Kaila & B. R. Ransom (Eds.), *pH and Brain Function*, Chapt. 16. New York Wiley-Liss.

Banich, M. T. (1998). The missing link: The role of interhemispheric interaction in attentional processing. *Brain and Cognition, 36*, 128–157.

Beeman, M., & Chiarello, C. (Eds.). (1998). *Right Hemisphere Language Comprehension*. Mahwah, Hillsdale, NJ: Erlbaum.

Beeman, M., Friedman, R. B., Grafman, J., Perez, E., Diamond, S., & Lindsay, M. B. (1994). Summation priming and coarse semantic coding in the right hemisphere. *Journal of Cognitive Neuroscience, 6*, 26–45.

Behrens, S. J. (1989). Characterizing sentence intonation in a right hemisphere-damaged population. *Brain and Language, 37*, 181–200.

Bell, J. S. (1987). *The Speakable and Unspeakable in Quantum Mechanics*. Cambridge: Cambridge University Press.

Bergmann, G., Goldbeck, T., & Scherer, K. R. (1988). Emotionale Einsdruckswirkung von prosodischen Sprechmerkmalen. *Zeitschrift für Experimentelle und Angewandte Psychologie, 35*, 167–200.

Bharucha, J. J. (1984). Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology, 16*, 485–510.

Bharucha, J. J. (1987). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception, 5*, 1–30.

Bharucha, J. J. (1996). Melodic anchoring. *Music Perception, 13*, 383–401.

Bharucha, J. J., & Mencl, W. E. (1996). Two issues in auditory cognition: Self-organization of octave categories and pitch-invariant pattern recognition. *Psychological Science, 7*, 142–149.

Bianki, V. L. (1988). *The Right and Left Hemispheres of the Animal Brain*. New York: Gordon and Breach.

Binder, J., Frost, J., & Hammeke, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *Journal of Neuroscience, 17*, 353–362.

Block, N. (1995). On a confusion about a function of consciousness, *Behavioral and Brain Sciences, 18*, 227–287.

Boersma, P., & Weenink, D. (1992). *Praat: A system for doing phonetics by computer*. http://www.praat.org

Bogen, J. E. (1969). The other side of the brain. *Bulletin of the Los Angeles Neurological Society, 34*, 135–162.

Bogen, J. E. (1995). On the neurophysiology of consciousness: I. An overview. *Consciousness and Cognition, 4*, 15–24.

Bogen, J. E. (1997). Does cognition in the disconnected right hemisphere require right hemisphere possession of language? *Brain and Language, 57*, 12–21.

Bohm, D. (1952). A suggested interpretation of the quantum theory in terms of 'hidden' variables. *Physical Review, 85*, 166–193.

Bolinger, D. L. (1978). Intonation across languages. In J. H. Greenberg, C. A. Ferguson, & E. A. Moravesik (Eds.), *Universals of Human Language: Phonology* (pp. 471–524). Stanford: Stanford University Press.

Bolinger, D. (1989). *Intonation and Its Uses: Melody in grammar and discourse*. Stanford: Stanford University Press.

Bottini, G., et al. (1994). The role of the right hemisphere in the interpretation of figurative aspects of language. A positron emission tomography activation study. *Brain, 117*, 1241–1253.

Bouman, C. A. (2002). Cluster: An unsupervised algorithm for modeling Gaussian mixtures. http://www.ece.purdue.edu/~bouman

Brown, S. (2000). The 'musilanguage' model of music evolution. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The Origins of Music*. Cambridge: MIT Press.

Brownell, H. H., Potter, H. H., Michelow, D., & Gardner, H. (1984). Sensitivity to lexical denotation and connotation in brain damaged patients: A double dissociation. *Brain and Language, 22*, 253–265.

Bulman-Fleming, M. B., & Bryden, M. P. (1994). Simultaneous verbal and affective laterality effects. *Neuropsychologia, 32*, 787–797.

Bunge, M. (1959). *Metascientific Queries*. Springfield, IL: C. C. Thomas.

Burgess, C., & Chiarello, C. (1996). Neurocognitive mechanisms underlying metaphor comprehension and other figurative language. *Metaphor and Symbolic Activity, 11*, 67–84.

Burgess, C., & Lund, K. (1998). Modeling cerebral asymmetries in high-dimensional space. In C. Burgess & C. Chiarello (Eds.), *Right Hemisphere Language Comprehension* (pp. 215–244). Hillsdale, NJ: Erlbaum.

Chafe, W. (1994). *Discourse, Consciousness and Time*. Chicago: University of Chicago Press.

Chalmers, D. (1996). *The Conscious Mind*. Oxford: Oxford University Press.

Chiarello, C., & Maxfield, L. (1996). Varieties of interhemispheric inhibition, or how to keep a good hemisphere down. *Brain and Cognition, 30*, 81–108.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Clynes, M. (1982). *Music, Mind and Brain: The neuropsychology of music*. New York: Plenum Press.

Code, C. (1987). *Language, Aphasia and the Right Hemisphere*. New York: John Wiley.

Collier, R., & 't Hart J. (1981). *Cursus Nederlandse Intonatie*. Leuven: Acco/De Horstink.

Cook, N. D. (1984a). Homotopic callosal inhibition. *Brain and Language, 23*, 116–125.

Cook, N. D. (1984b). Callosal inhibition: the key to the "brain code". *Behavioral Science, 29*, 98–110.

Cook, N. D. (1986). *The Brain Code: Mechanisms of information transfer and the role of the corpus callosum*. London: Methuen.

Cook, N. D. (1989). Toward a central dogma for psychology. *New Ideas in Psychology, 7*, 1–18.

Cook, N. D. (1999a). Explaining harmony: The roles of interval dissonance and chordal tension. In R. J. Zatorre & I. Peretz (Eds.), *The Biological Foundations of Music* (pp. 382–385). New York: New York Academy of Sciences.

Cook, N. D. (1999b). Simulating consciousness in a bilateral neural network: "nuclear" and "fringe" awareness. *Consciousness and Cognition, 8*, 62–93.

Cook, N. D. (1999c). On defining awareness and consciousness: The importance of the neuronal membrane. *Proceedings of the Tokyo-99 Conference on Consciousness*. Singapore: World Scientific.

Cook, N. D. (2000a). Bridging the explanatory gap. *Proceedings of the 26th Annual Meeting of the Psychological Association*. Stockholm.

Cook, N. D. (2000b). Understanding harmony. *Proceedings of the Fifth Annual Meeting of the Society for Music Perception and Cognition*. Keele, England.

Cook, N. D. (2001). Acknowledging and resolving the hard problem without recourse to dualism. *Proceedings of Toward a Science of Consciousness*. Skovde, Sweden.

Cook, N. D. (2002a). Modèles de la communication interhémisphérique. *Revue de Neuro-psychologie, 12*, 51–85.

Cook, N. D. (2002b). Bihemispheric language: How the two hemispheres collaborate in the processing of language. In T. Crow (Ed.), *The Speciation of Modern Homo Sapiens*, Vol. 106, Chapter 9. London: Proceedings of the British Academy.

Cook, N. D., & Beech, A. S. (1991). The cerebral hemispheres and bilateral neural nets. *International Journal of Neuroscience, 51*, 101–110.

Cook, N. D., Callan, D. E., & Callan, A. (2001). Behavioral and fMRI study of harmony. *Proceedings of the Sixth Annual Meeting of the Society for Music Perception and Cognition.* Kingston, Ontario.

Cook, N. D., Callan, D. E., & Callan, A. (2002a). Frontal areas involved in the perception of harmony. *8th International Conference on Functional Mapping of the Human Brain.* Sendai, June 2–6.

Cook, N. D., Callan, D. E., & Callan, A. (2002b). Behavioral and fMRI study of harmony (submitted for publication).

Cook, N. D., Frueh, H., Mehr, A., Regard, M., & Landis, T. (1994). Hemispheric cooperation in visuospatial rotations: Evidence for a manipulation role for the left hemisphere and a reference role for the right hemisphere. *Brain and Cognition, 25*, 240–249.

Cooke, D. (1959). *The language of music*. Oxford: Oxford University Press.

Corballis, M. C. (1991). *The Lopsided Ape: Evolution of the generative mind*. New York: Oxford University Press.

Costa, M., Bitti, P. E. R., & Bonfiglioli, L. (2000). Psychological connotations of harmonic musical intervals. *Psychology of Music, 28*, 4–22.

Creutzfeldt, O. D. (1995). *The Cerebral Cortex*. New York: Oxford University Press.

Crick, F. (1994). *The Astonishing Hypothesis: The scientific search for the soul*. New York: Scribners.

Crick, F., & Koch, C. (1990). Toward a neurobiological theory of consciousness. *Seminars in Neuroscience, 2*, 263–275.

Crow, T. J. (1997). Schizophrenia as failure of hemispheric dominance for language. *Trends in Neurosciences, 20*, 339–343.

Crow, T. J. (1998). Why cerebral asymmetry is the key to the origin of Homo sapiens: How to find the gene or eliminate the theory. *Cahiers de Psychologie Cognitive, 17*, 1237–1277.

Crow, T. J. (2002). *The Speciation of Modern Homo Sapiens*. London: Proceedings of the British Academy.

Cruttendon, A. (1981). Falls and rises: meanings and universals. *Journal of Linguistics, 17*, 77–91.

Cruttendon, A. (1997). *Intonation*. New York: Cambridge University Press.

Cruttendon, A. (2001). Mancunian intonation and intonational representation. *Phonetica, 58*, 53–80.

Cutler, A., & Clifton C. (1999). Comprehending spoken language: Blueprint of the listener. In C. M. Brown & P. Hagoort (Eds.), *The Neurocognition of Language* (pp. 123–166). Oxford: Oxford University Press.

Deacon, T. W. (1997). *The Symbolic Species*. New York: Norton.

Deitmer, J. W. (1998). pH regulation in mammalian neurons. In K. Kaila & B. R. Ransom (Eds.), *pH and Brain Function*, Chapt. 12. New York Wiley-Liss.

Denenberg, V. H. (1980). General systems theory, brain organization and early experience. *American Journal of Physiology: Regulatory, Integrative and Comparative Physiology, 238*, 213–223.

Dennett, D. (1991). *Consciousness Explained*. New York: Little, Brown.

D'Esposito, M., Detre, J. A., Aguirre, G. K., Stallcup, M., Alsop, D. C., Tippet, L. J., & Farah, M. J. (1997). A functional MRI study of mental image generation. *Neuropsychologia, 35*, 725–730.

Doty, R. W., Overman, W. H., & Negrao, N. (1979). Role of forebrain commissures in hemispheric specialization and memory in macaques. In I. S. Russell (Ed.), *Structure and Function of the Cerebral Commissures* (pp. 333–342). New York: Macmillan.

Eady, S. J., & Cooper, W. E. (1986). Speech intonation and focus location in matched statements and questions. *Journal of the Acoustical Society of America, 80*, 402–415.

Eccles (1990). A unitary hypothesis of mind-brain interaction in the cerebral cortex. *Proceedings of the Royal Society London, B240*, 433–451.

Edelman, G. (1992). *Bright Air, Brilliant Fire: On the matter of the mind*. New York: Basic Books.

Edelman, G. M., & Tononi, G. (2000). *A Universe of Consciousness: How matter becomes imagination*. New York: Basic.

Efron, R. (1990). *The Decline and Fall of Hemispheric Specialization*. Hillsdale, NJ: Erlbaum.

Engel, A. K., & Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends in Cognitive Science, 5*, 16–25.

Engel, A. K., Fries, P., Konig, P., Brecht, M., & Singer, W. (1999). Temporal binding, binocular rivalry, and consciousness. *Consciousness and Cognition, 8*, 230–250.

Fitzsimons, M., Sheahan, N., & Staunton, H. (2001). Gender and the integration of acoustic dimensions of prosody: Implications for clinical studies. *Brain and Language, 78*, 94–108.

Fletcher, P. C., Shallice, T., & Dolan, R. J. (1998a). The functional roles of prefrontal cortex in episodic memory. I. Encoding. *Brain, 121*, 1239–1248.

Fletcher, P. C., Shallice, T., Frith, C. D., Frackowiak, R. S. J., & Dolan, R. J. (1998b). The functional roles of prefrontal cortex in episodic memory. II. Retrieval. *Brain, 121*, 1249–1256.

Fonagy, I. (1998). Intonation in Hungarian. In D. Hirst & A. Di Cristo (Eds.), *Intonation Systems* (pp. 328–344). Cambridge: Cambridge University Press.

Fonagy, I. (2001). *Languages within Language: An evolutive approach*. Amsterdam: John Benjamins.

Fonagy, I., & Magdics, K. (1963). Emotional patterns in intonation and music. *Zeitschrift für Phonetic, 16*, 293–326.

Fox, P. T., Ingham, R. J., Ingham, J. C., Hirsch, T. B., Downs, J. H., Martin, C., Jerabek, P., Glass, T., & Lancaster, J. L. (1996). A PET study of the neural systems of stuttering. *Nature, 382*, 158–162.

Frick, R. W. (1985). Communicating emotion: The role of prosodic features. *Psychological Bulletin, 97*, 412–429.

Gainotti, G., Caltagirone, C., & Miceli, G. (1983). Selective impairment of semantic-lexical discrimination in right-brain-damaged patients. In E. Perecman (Ed.), *Cognitive Processing in the Right Hemisphere* (pp. 149–167). New York: Academic Press.

Galaburda, A. M. (1995). Anatomic basis of cerebral dominance. In R. J. Davidson & K. Hugdahl (Eds.), *Brain Asymmetry* (pp. 51–73). Cambridge, MA: MIT Press.

Galin, D. (1994). The structure of awareness: contemporary applications of William James' forgotten concept of "the fringe". *Journal of Mind and Behavior, 15*, 375–400.

Galin, D. (1996). The structure of subjective experience. In S. R. Hameroff, A. W. Kaszniak, & A. C. Scott (Eds.), *Toward a Science of Consciousness*. Cambridge, MA: MIT Press.

Garding, E. (1983). A generative model of intonation. In A. Cutler & D. R. Ladd (Eds.), *Prosody Models and Measurements* (pp. 11–21). Berlin: Springer.

Gardner, H., Brownell, H., Wapner, W., & Michelow, D. (1983). Missing the point: the role of the right hemisphere in the processing of complex linguistic materials. In E. Perecman (Ed.), *Cognitive Processing in the Right Hemisphere* (pp. 169–192). New York: Academic Press.

Gazzaniga, M. S. (1995). Consciousness and the cerebral hemispheres. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences*. Cambridge, MA: MIT Press.

Geschwind, N. (1982). Disorders of attention: a frontier in neuropsychology. *Philosophical Transactions of the Royal Society London, B298*, 173–185.

Geschwind, N., & Galaburda, A. M. (Eds.). (1984). *Cerebral Dominance: The biological foundations*. Cambridge, MA: Harvard University Press.

Gibbs, R. W. (1994). *The Poetics of Mind: Figurative thought, language and understanding*. New York: Cambridge University Press.

Gill, H. S., O'Boyle, M. W., & Hathaway, J. (1998). Cortical distribution of EEG activity for component processes during mental rotation. *Cortex, 34*, 707–718.

Gott, P. S. (1973). Cognitive ability following right and left hemispherectomy. *Cortex, 9*, 266–274.

Gray, C. M., Konig, P., Engel, A. K., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature, 338*, 334–337.

Greenfield, P. M. (1991). Language, tools and brain: The ontogeny and phylogeny of hierarchically organized sequential behavior. *Behavioral and Brain Sciences, 14*, 531–595.

Gregory, S. W., Dagan, K., & Webster, S. (1997). Evaluating the relation of vocal accomodation in conversation partners' fundamental frequencies to perceptions of communication quality. *Journal of Nonverbal Behavior, 21*, 23–43.

Gregory, S. W., Green, B. E., Carrothers, R. M. Dagan, K. A., & Webster, S. W. (2001). Verifying the primacy of voice fundamentral frequency in social status accomodation. *Language and Communication, 21*, 37–60.

Grilly, D. M. (2002). *Drugs and Human Behavior*. Boston: Allyn and Bacon.

Grimshaw, G. M. (1998). Integration and interference in the cerebral hemispheres: Relations with hemispheric specialization. *Brain and Cognition, 36*, 108–127.

Gross, M. (1998). *Living on the Edge: Amazing creatures thriving in extreme environments*. New York: Plenum.

Grossberg, S. (1987). Competitive learning: From interactive activation to adaptive resonance. *Cognitive Science, 11*, 23–63.

Grout, D. J. (1973). *A History of Western Music*. New York: Norton.

Guernesey, M. (1928). The role of consonance and dissonance in music. *American Journal of Psychology, 40*, 173–204.

Guiard, Y. (1980). Cerebral hemispheres and selective attention. *Acta Psychologica, 46*, 41–61.

Harrington, A. (1987). *Medicine, Mind, and the Double Brain: A Study in Nineteenth-Century Thought*. Princeton: Princeton University Press.

Harrison, D. (1994). *Harmonic Function in Chromatic Music*. Chicago: University of Chicago Press.

Hauser, M. D. (1996). *The Evolution of Communication*. Cambridge: MIT Press.

Hebb, D. O. (1949). *The Organization of Behavior*. New York: Wiley.

Hedreen, J. C., & Yin, T. C. T. (1981). Homotopic and heterotopic callosal afferents of caudal inferior parietal lobule in Macaca mulatta. *Journal of Comparative Neurology, 197*, 605–621.

Heilman, K. M., & Van Dan Abell, T. (1979). Right hemispheric dominance for mediating cerebral activation. *Neuropsychologia, 17*, 315–321.

Hellige, J. B. (1993). *Hemispheric asymmetry: What's right and what's left*. Cambridge, MA: Harvard University Press.

Helmholtz, H. (1877/1954). *On the Sensations of Tone*. New York: Dover.

Heun, R., Klose, U., Jessen, F., Erb, M., Papassotiropoulos, A., Lotze, M., & Grodd, W. (1999). Functional MRI of cerebral activation during encoding and retrieval of words. *Human Brain Mapping, 8*, 157–169.

Hevner, K. (1936). Experimental studies of the elements of expression in music. *American Journal of Psychology, 48*, 246–268.

Howell, P., & Kadi-Hanifi, K. (1991). Comparison of prosodic properties between read and spontaneous speech. *Speech Communication, 10*, 163–169.

Hugdahl, K., Bronnick, K., Kyllinsbeak, S., Law, I., Gade, A., & Paulson, O. B. (1999). Brain activation during dichotic presentations of consonant-vowel and musical instrument stimuli: A ^{15}O-PET study. *Neuropsychologia, 37*, 431–440.

Huron, D. (1991). Tonal consonance versus tonal fusion in polyphonic sonorities. *Music Perception, 9*, 135–154.

Huron, D. (2001). Tone and voice: A derivation of the rules of voice-leading from perceptual principles. *Music Perception, 19*, 1–64.

Iacoboni, M., & Zaidel, E. (1996). Hemispheric independence in word recognition: Evidence from unilateral and bilateral presentations. *Brain and Language, 53*, 121–140.

Inui, T., Tanaka, S., Okada, T., Nishizawa, S., Katayama, M., & Konishi, J. (2000). Neural substrates for depth perception of the Necker cube; a functional magnetic resonance imaging study in human subjects. *Neuroscience Letters, 282*, 145–148.

Ivry, R. B., & Robertson, L. C. (1998). *The Two Sides of Perception*. Cambridge, MA: MIT Press.

Jackendoff, R. (1999). The representational structure of the language faculty and their interactions. In C. M. Brown & P. Hagoort (Eds.), *The Neurocognition of Language* (pp. 37–79). Oxford: Oxford University Press.

Jackendoff, R. (2002). *Foundations of Language: Brain, meaning grammar, evolution*. Oxford: Oxford University Press.

Jacobson, S., & Marcus, E. M. (1970). The laminar distribution of the corpus callosum: a comparative study in the rat, cat, rhesus monkey and chimpanzee. *Brain Research, 24*, 513–520.

James, W. (1890/1950). *The Principles of Psychology*. New York: Dover.

Jaynes, J. (1976). *The Origins of Consciousness in the Breakdown of the Bicameral Mind*. Boston: Houghton-Mifflin.

Jones, E. G. (1985). *The Thalamus*. New York: Plenum Press.

Jourdain, R. (1997). *Music, the Brain and Ecstasy*. New York: Avon.

Kaestner, G. (1909). Untersuchungen ueber den Gefuehlseindurck unanalysierter Zweiklaenge. *Psychologishe Studien, 4*, 473–504.

Kaila, K., & Ransom, B. R. (Eds.). (1998). *pH and Brain Function*. New York Wiley-Liss.

Kameoka, A., & Kuriyagawa, M. (1969). Consonance theory. *Journal of the Acoustical Society of America* (Parts I and II), *45*, 1452–1459; *45*, 1460–1469.

Kaminska, Z., & Woolf, J. (2000). Melodic line and emotion: Cooke's theory revisited. *Psychology of Music, 28*, 133–153.

Kandel, E. C., & Schwartz, J. H. (1991). *Principles of Neural Science*. New York: Elsevier.

Karbe, H., Herholz, K., Halber, M., & Heiss, W. D. (1998). Collateral inhibition of transcallosal activity facilitates functional brain asymmetry. *Journal of Cerebral Blood Flow and Metabolism, 18*, 1157–1161.

Kashiwagi, A., Kashiwagi, T., Nishikawa, T., & Okuda, J. (1989). Hemispheric asymmetry of processing temporal aspects of repetitive movement in two patients with infarction involving the corpus callosum. *Neuropsychologia, 27*, 799–809.

Kauppinen, R. A., & Williams, S. R. (1998). Use of NMR spectroscopy in monitoring cerebral pH and metabolism during systemic and focal acid-base disturbances. In K. Kaila & B. R. Ransom (Eds.), *pH and Brain Function*, Chapt. 30. New York Wiley-Liss.

Kiefer, M., Weisbrod, M., Kern, I., Maier, S., & Spitzer, M. (1998). Right hemisphere activation during indirect semantic priming: Evidence from event-related potentials. *Brain and Language, 64*, 377–408.

Kiehl, K. A., Liddle, P. F., Smith, A. M., Mendrek, A., Forster, B. B., & Hare, R. D. (1999). Neural pathways involved in the processing of concrete and abstract words. *Human Brain Mapping, 7*, 225–233.

Kim, S. G., Ashe, J., Hendrich, K., Ellermann, J. M., Merle, H., Ugurbil, K., Georgopoulos, A. P. (1993). Functional magnetic resonance imaging of motor cortex: Hemispheric asymmetry and handedness. *Science, 261*, 615–617.

Kinsbourne, M. (1970). The cerebral basis of lateral asymmetries in attention. *Acta Psychologica, 33*, 193–201.

Kinsbourne, M. (1978). *Asymmetrical Function of the Brain*. New York: Academic Press.

Kinsbourne, M. (1982). Hemispheric specialization and the growth of human understanding. *American Psychologist, 37*, 411–420.

Kinsbourne, M. (1997). What qualifies a representation for a role in consciousness? In J. D. Cohen, S. & J. W. Schooler (Eds.), *Scientific Approaches to Consciousness*. Hillsdale, NJ: Erlbaum.

Kitayama, S. (1991). Impairment of perception by positive and negative affect. *Cognition and Emotion, 5*, 255–274.

Klouda, G. V., Robin, D. A., Graff-Radford, N. R., & Cooper, W. E. (1988). The role of callosal connections in speech prosody. *Brain and Language, 35*, 154–171.

Koch, C. (1999). *Biophysics of Computation: Information processing in single neurons*. New York: Oxford University Press.

Kohonen, T. (1995). *Self-organizing Maps*. New York: Springer.

Kosslyn, S. M. (1994). *Image and Brain*. Cambridge, MA: Harvard University Press.

Kövecses, Z. (2000). Metaphor and Emotion: Language, culture, and body in human feeling. New York: Cambridge University Press.

Krumhansl, C. L. (1993). *Cognitive foundations of musical pitch*. Oxford: Oxford University Press.

Krumhansl, C. L. (2001). Rhythm and pitch in music cognition. *Psychological Bulletin, 126*, 159–179.

Ladd, D. R. (1996). *Intonational Phonology*. Cambridge: Cambridge University Press.

Landis, T. (1987). Right hemisphere reading: a clinico-experimental approach to the dual brain interaction. Habilitationsschrift. Zurich: Zurich University.

Landis, T., Lehmann, D., Mita, T., & Skrandies, W. (1984). Evoked potential correlates of figure and ground. *International Journal of Psychophysiology, 1*, 345–348.

Lee, A. C. H., Robbins, T. W., Pickard, J. D., & Owen, A. M. (2000). Asymmetric activation during episodic memory: the effects of stimulus type on encoding and retrieval. *Neuropsychologia, 38*, 677–692.

Leonard, C. L., Waters, G. S., & Caplan, D. (1997). The use of contextual information by right brain-damaged individuals in the resolution of ambiguous pronouns. *Brain and Language, 57*, 309–342.

Lerdahl, F., & Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.

Levelt, W. J. M. (1999). Producing spoken language: A blueprint of the speaker. In C. M. Brown & P. Hagoort (Eds.), *The Neurocognition of Language* (pp. 83–122). Oxford: Oxford University Press.

Levine, J. (1983). Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly, 64*, 354–361.

Levine, J. (2000). *Purple Haze*. Cambirdge, MA: MIT Press.

Levitan, S., & Reggia, J. A. (2000). A computational model of lateralization and asymmetries in cortical maps. *Neural Computation, 12*, 2037–2062.

Liederman, J. (1998). The dynamics of interhemispheric interaction and hemispheric control. *Brain and Cognition, 36*, 193–208.

Loewenstein, W. R. (1999). *The Touchstone of Life*. New York: Penguin.

Maass, W., & Bishop, C. M. (Eds.). (1999). *Pulsed Neural Networks*. Cambridge, MA: MIT Press.

MacLennan, B. (1996). The elements of consciousness and their neurodynamical correlates. *Journal of Consciousness Studies, 3*, 409–424.

Maher, T. F. (1980). A rigorous test of the proposition that musical intervals have different psychological effects. *American Journal of Psychology, 98*, 309–327.

Mangan, B. (1993). Taking phenomenology seriously: The "fringe" and its implications for cognitive research. *Consciousness and Cognition, 2*, 89–108.

Mathiak, K., Hertrich, I., Lutzenberger, W., & Ackermann, H. (2002). Functional cerebral asymmetries of pitch processing during dichotic stimulus application: a whole-head magnetoencephalography study. *Neuropsychologia, 40*, 585–593.

McGinn, C. (1990). *The Problem of Consciousness*. Oxford: Blackwell.

McNeely, H. E., & Parlow, S. E. (2001). Complementarity of linguistic and prosodic processes in the intact brain. *Brain and Language, 79*, 473–481.

Mehta, Z., Newcombe, F., & Damasio, H. (1987). A left hemisphere contribution to visuospatial processing. *Cortex, 23*, 447–461.

Meyer, L. B. (1956). *Emotion and Meaning in Music*. Chicago: University of Chicago Press.

Morton, E. W. (1977). On the occurrence and significance of motivation – structural rules in some bird and mammal sounds. *American Naturalist, 111*, 855–869.

Moscovitch, M. (1973). Language and the cerebral hemispheres. In P. Pliner et al. (Eds.), *Communication and affect: Language and thought* (pp. 89–126). New York: Academic Press.

Mountcastle, V. B. (1997). The columnar organization of the neocortex. *Brain, 120*, 701–722.

Netz, J., Ziemann, U., & Homberg, V. (1995). Hemispheric asymmetry of transcallosal inhibition in man. *Experimental Brain Research, 104*, 527–533.

Nichelli, P., Grafman, J., Pietrini, P., Clark, K., Lee, K. Y., & Miletich, R. (1995). Where the brain appreciates the moral of a story. *Neuroreport, 6*, 2309–2313.

Nottebohm, F. (1994). The song circuits of the avian brain as a model system in which to study vocal learning, communication and manipulation. *Discussions in Neuroscience, 10*, 72–80.

Nyberg, L. (1998). Mapping episodic memory. *Behavioral Brain Research, 90*, 107–114.

Nygaard, L. C., & Queen, J. S. (1999). Communicating emotion: Linking affective prosody and word meaning (preprint).

Ogura, A., & Machemer, H. J. (1980). Distribution of mechanoreceptor channels in the *Paramecium* surface membrane. *Journal of Comparative Physiology, A 135*, 233–242.

Ohala, J. J. (1984). An ethological perspective on common cross-language utilization of F_0 in voice. *Phonetica, 41*, 1–16.

Ohala, J. J. (1994). The frequency code underlies the sound-symbolic use of voice pitch. In L. Hinton, J. Nichols, & J. J. Ohala (Eds.), *Sound Symbolism* (pp. 325–347). New York: Cambridge University Press.

O'Regan, J. K., & Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Brain and Behavioral Sciences, 24*, 1–31.

Ornstein, R. E. (1997). *Right Mind*. New York: Harcourt-Brace-Jovanovich.

Parncutt, R. (1989). *Harmony: A psychoacoustical approach*. New York: Springer.

Passingham, R. E. (1981). Broca's area and the origin of human vocal skills. *Philosophical Transactions of the Royal Society London, B292*, 167–175.

Pell, M. D. (1999). Fundamental frequency encoding of linguistic and emotional prosody by right hemisphere-damaged speakers. *Brain and Language, 69*, 161–192.

Penrose, R. (1989). *The Emperor's New Mind*. Oxford: Oxford University Press.

Perischetti, V. (1961). *Twentieth-Century Harmony*. New York: Norton.

Pfeifer, R., & Scheier, C. (1999). *Understanding Intelligence*. Cambridge, MA: MIT Press.

Pierce, J. R. (1992). *The Science of Musical Sound*. New York: Freeman.

Pierrehumbert, J. (1981). Synthesizing intonation. *Journal of the Acoustical Society of America, 70*, 985–995.

Plomp, R., & Levelt, W. J. M. (1965). Total consonance and critical bandwidth. *Journal of the Acoustical Society of America, 38*, 548–560.

Poeppel, E. (1997). A hierarchical model of temporal perception. *Trends in the Cognitive Sciences, 1*, 56–61.

Poeppel, D. (2001). Pure word deafness and the bilateral processing of the speech code. *Cognitive Science, 25*, 679–693.

Poggio, T., & Girosi, F. (1990). Networks for approximation and learning. *Proceedings IEEE, 78*, 1481–1497.

Posner, M. I., & Raichle, M. E. (1994). *Images of Mind*. New York: Scientific American.

Povel, D. J., & Jansen, E. (2001). Perceptual mechanisms in music processing. *Music Perception, 19*, 169–198.

Pribram, K. H. (1991). *Brain and Perception: Holonomy and structure in figural processing*. Hillsdale, NJ: Erlbaum.

Price, C. J., Wise, R. J. S., Warburton, E., Moore, C. J., Patterson, K., Howard, et al. (1996). Hearing and saying: The functional neuroanatomy of auditory word processing. *Brain, 119*, 919–931.

Puccetti, R. (1981). The case for mental duality: Evidence from split brain data and other considerations. *Behavioral and Brain Sciences, 4*, 93–123.

Querné, L., Eustache, F., & Faure, S. (2000). Interhemispheric inhibition, intrahemispheric activation, and lexical capacities of the right hemisphere: A tachistoscopic, divided visual-field study in normal subjects. *Brain and Language, 74*, 171–190.

Rameau, J-P. (1722/1971). *Treatise on Harmony*. New York: Dover.

Reeves, A. G. (Ed.). (1985). *Epilepsy and the Corpus Callosum*. New York: Plenum.

Regard, M., Wieser, H. G., Landis, T., & Cook, N. D. (1994). The dynamics of cerebral dominance during unilateral limbic seizures. *Brain, 117*, 91–104.

Reggia, J. A., Goodall, S., & Shkuro, Y. (1998). Computational studies of lateralization of phoneme sequence generation. *Neural Computation, 10*, 1277–1297.

Reggia, J. A., Goodall, S., & Levitan, S. (2001a). Cortical map asymmetries in the context of transcallosal excitatory influences. *Neuroreport, 12*, 1609–1614.

Reggia, J. A., Goodall, S. M., Shkuro, Y., & Glezer, M. (2001b). The callosal dilemma: Explaining diaschisis in the context of hemispheric rivalry via a neural network model. *Neurological Research, 23*, 465–471.

Regnault, P., Bigand, E., & Besson, M. (2001). Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context. *Journal of Cognitive Neuroscience, 13*, 241–255.

Riemann, H. (1877). *Musikalische Syntaxis: Grundriss einer harmonischen Satzbildungslehre*. Leipzig: Breitkopf und Härtel.

Roberts, L. A. (1986). Consonant judgements of musical chords by musicians and untrained listeners. *Acustica, 62*, 163–171.

Rodel, M., Cook, N. D., Regard, M., & Landis, T. (1992). Hemispheric dissociation in judging semantic relations. *Brain and Language, 43*, 448–159.

Roederer, J. G. (1995). *The Physics and Psychophysics of Music: An introduction*. New York: Springer.

Roelfsema, P. R., Engel, A. K., & Singer, W. (1996). The role of neuronal synchronization in response selection: a biologically plausible theory of structured representation in the visual cortex. *Journal of Cognitive Neuroscience, 8*, 603–625.

Rosen, G. D., Sherman, G. F., & Galaburda, A. M. (1989). Interhemispheric connections differ between symmetrical and asymmetrical brain regions. *Neuroscience, 33*, 525–533.

Ross, E. D., Harney, J. H., deLacoste-Utamsing, C., & Purdy, P. D. (1981). How the brain integrates affective and propositional language into a unified behavioral function. *Archives of Neurology, 38*, 745–748.

Ross, E. D., Thompson, R. D., & Yenkosky, J. (1997). Lateralization of affective prosody in brain and the callosal integration of hemispheric language functions. *Brain and Language, 56*, 27–54.

Ross, M. K., Reeves, A. G., & Roberts, D. W. (1984). Post-commissurotomy mutism. *Annals of Neurology, 16*, 114–123.

Rushworth, M. F. S., Nixon, P. D., Wade, D. T., Renowden, S., & Passingham, R. E. (1998). The left hemisphere and the selection of learned actions. *Neuropsychologia, 36*, 11–24.

Schachter, D. S. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory and Cognition, 13*, 501–518.

Schenker, H. (1954). *Harmony*. Chicago: University of Chicago Press.

Scherer, K. R. (1974). Acoustic concomitants of emotional dimensions: Judging affect from synthesized tone sequences. In S. Weitz (Ed.), *Non-verbal Communication* (pp. 105–111). New York: Oxford University Press.

Scherer, K. R. (1986). Vocal affect expression. *Psychological Bulletin, 99*, 143–165.

Scherer, K. R. (1992). Vocal affect expression as symptom symbol and appeal. In H. Papousek, U. Jürgens, & M. Papousek (Eds.), *Nonverbal Vocal Communication* (pp. 43–60). New York: Cambridge University Press.

Schiffer, F. (1998). *Of Two Minds: The revolutionary science of dual-brain psychology*. New York: Free Press.

Schillen, T. B., & Konig, P. (1994). Binding by temporal structure in multiple feature domains of an oscillatory neuronal network. *Biological Cybernetic, 70*, 397–405.

Schirmer, A., Alter, K., Kotz, S. A., & Friederici, A. D. (2001). Lateralization of prosody during language production: A lesion study. *Brain and Language, 76*, 1–17.

Schmidt, L. A., & Trainor, L. J. (2001). Frontal brain electrical activity (EEG) distinguishes valence and intensity of musical emotions. *Cognition and Emotion, 15*, 487–500.

Schneiderman, E. I., Murasugi, K. G., & Saddy, J. D. (1992). Story arrangement ability in right-brain damaged patients. *Brain and Language, 43*, 107–120.

Schoenberg, A. (1911/1978). *Theory of harmony*. London: Faber & Faber.

Schoenberg, A. (1995). *The Musical Idea.* New York: Columbia University Press.

Scruton, R. (1997). *The Aesthetics of Music*. Oxford: Clarendon Press.

Searle, J. R. (1997). *The Mystery of Consciousness*. New York: New York Review Books.

Sethares, W. A. (1993). Local consonance and the relationship between timbre and scale. *Journal of the Acoustical Society of America, 94*, 1218–1228.

Sethares, W. A. (1998). Consonance-based spectral mappings. *Computer Music Journal, 22*, 56–72.

Sethares, W. A. (1999). *Tuning, Timbre, Spectrum, Scale*. New York: Springer.

Shallice, T., Fletcher, P., Frith, C. D., Grasby, P., Frackowiak, R. S. J., & Dolan, R. J. (1994). Brain regions associated with acquisition and retrieval of verbal episodic memory. *Nature, 368*, 633–635.

Shamma, S. A. (1995). Auditory cortex. In M. A. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks* (pp. 110–115). Cambridge: The MIT Press.

Shastri, L., & Ajjanagadde, V. (1993). From simple associations to systematic reasoning: A connectionist representation of rules, variables and dynamic bindings using temporal synchrony. *Behavioral and Brain Sciences, 16*, 417–494.

Shear, J. (1998). *Explaining Consciousness: The hard problem*. Cambridge, MA: MIT Press.

Shkuro, Y., Glezer, M., & Reggia, J. A. (2000). Interhemispheric effects of simulated lesions in a neural model of single-word reading. *Brain and Language, 72*, 343–374.

Singer, W. (1993). Synchronization of cortical activity and its putative role in information processing and learning. *Annual Review of Physiology, 55*, 349–375.

Singer, W. (1995). Development and plasticity of cortical processing architectures. *Science, 270*, 758–768.

Singer, W., & Gray, W. C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience, 18*, 555–586.

Sloboda, J. A. (1985). *The Musical Mind*. New York: Oxford University Press.

Sperry, R. W. (1966). Brain bisection and consciousness. In J. Eccles (Ed.), *Brain and Conscious Experience*. Berlin: Springer.

Sperry, R. W. (1968). Hemisphere deconnection and unity in conscious awareness. *American Psychologist, 23*, 723–733.

Sperry, R. W., Gazzaniga, M. S., & Bogen, J. E. (1969). Interhemispheric relationships: the neocortical commissures; syndromes of hemisphere disconnection. In P. J. Vinken & G. W. Bruyn (Eds.), *Handbook of Clinical Neurology*, Vol. 4 (pp. 273–290). Amsterdam: North-Holland.

Springer, S. P., & Deutsch, G. (1998). *Left Brain, Right Brain* (5th ed.). New York: Freeman.

Stapp, H. P. (1993). *Mind, Matter and Quantum Mechanics*. New York: Springer.

Stein, B. E., & Meredith, M. A. (1993). *The Merging of the Senses*. Cambridge, MA: MIT Press.

Steriade, M., Jones, E. G., & Llinas, R. R. (1990). *Thalamic Oscillations and Signaling*. New York: Wiley.

Swain, J. (1997). *Musical Languages*. New York: Norton.

Tallal, P., et al. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science, 271*, 81–84.

Tan, N., Aiello, R., & Bever, T. G. (1981). Harmonic structure as a determinant of melodic organization. *Memory and Cognition, 9*, 533–539.

Taylor, J. G. (1999). *The Race for Consciousness*. Cambridge, MA: MIT Press.

Taylor, K. I., Brugger, P., Weniger, D., & Regard, M. (1999). Qualitative hemispheric differences in semantic category matching. *Brain and Language, 70*, 119–131.

TenHouten, W. D., Hoppe, K. D., & Bogen, J. E. (1985). Alexithymia and the split-brain I–III. *Psychotherapy and Psychosomatics, 43*, 202–208; *44*, 1–5, 89–94.

Terhardt, E. (1974). Pitch, consonance and harmony. *Journal of the Acoustical Society of America, 55*, 1061–1069.

Terhardt, E., Stoll, G., & Seewann, M. (1982a). Pitch of complex signals according to virtual-pitch theory. *Journal of the Acoustical Society of America, 71*, 671–678.

Terhardt, E., Stoll, G., & Seewann, M. (1982b). Algorithm for extraction of pitch and pitch salience from complex tonal signals. *Journal of the Acoustical Society of America, 71*, 679–688.

Tervaniemi, M., Kujala, A., Alho, K., Virtanen, J., Ilmoniemi, R. J., & Naatanen, R. (1999). Functional specialization of the human auditory cortex in processing phonetic and musical sounds: A magnetoencephalographic (MEG) study. *Neuroimage, 9*, 330–336.

Tillmann, B., & Bigand, E. (2001). Global context effect in normal and scrambled musical sequences. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 1185–1196.

Trainor, L. J., & Trehub, S. E. (1992). The development of referential meaning in music. *Music Perception, 9*, 455–470.

Tramo, M. J., Cariani, P. A., Delgutte, B., & Braida, L. D. (2001). Neurobiological foundations for the theory of harmony in western tonal music. In R. J. Zatorre & I. Peretz (Eds.), *The Biological Foundations of Music* (pp. 92–116). New York: New York Academy of Science.

Tulving, E., Kapur, S., Markowitsch, H. J., Craik, F. I. M, Habib, R., & Houle, S. (1994). Neuroanatomical correlates of retrieval in episodic memory: auditory sentence recognition. *Proceedings of the National Academy of Science, 91*, 2012–2015.

Van Lancker, D., & Kempler, D. (1987). Comprehension of familiar phrases by left but not by right hemisphere damaged patients. *Brain and Language, 32*, 265–277.

Van Lancker, D., & Sidtis, J. J. (1992). Identification of affective-prosodic stimuli by left and right hemisphere damaged subjects: All errors are not created equal. *Journal of Speech and Hearing Research, 35*, 963–970.

Velmans, M. (1991). Is human information processing conscious? *Behavioral and Brain Sciences, 14*, 651–726.

Velmans, M. (2000). *Understanding Consciousness*. London: Routledge.

von der Marlsburg, C. (1981/1994). The correlation theory of brain function (Internal Report 81–2). Max Planck Institute for Biophysical Chemistry. In E. Domany et al. (Eds.), *Models of Neural Networks*, Vol. II (pp. 95–119). Berlin: Springer.

Walker, E. H. (2000). *The Physics of Consciousness*. Cambridge, MA: Perseus.

Wallin, N. L., Merker, B., & Brown, S. (2000). *The Origins of Music*. Cambridge, MA: MIT Press.

Wapner, W., Hamby, S., & Gardner, H. (1984). The role of the right hemisphere in the apprehension of complex linguistic materials. *Brain and Language, 14*, 15–33.

Wedin, L. A. (1972). A multidimensional study of perceptual-emotional qualities in music. *Scandinavian Journal of Psychology, 13*, 241–257.

Weiskrantz, L. (1980). The problem of hemispheric specialization in animals. *Pontificiae Academiae Scientarium Scripta Varia, 45*, 573–592.

Weniger, D. (1984). Dysprosody as part of the aphasic language disorder. *Advances in Neurology, 42*, 41–50.

Wiener, N. (1948). *Cybernetics; or control and communication in the animal and the machine.* New York: Wiley.

Wigan, A. L. (1844). *The Duality of the Mind*. London: Longman (Reprinted Malibu: J. Simon Publ., 1985).

Winner, E., Brownell, H., Happé, F., Blum, A., & Pincus, D. (1998). Distinguishing lies from jokes. *Brain and Language, 62*, 89–106.

Winner, E., & Gardner, H. (1977). The comprehension of metaphor in brain-damaged patients. *Brain, 100*, 719–727.

Wurm, L. H., Vakoch, D. A., Strasser, M. R., Calin-Jageman, R., & Ross, S. E. (2001). Speech perception and vocal expression of emotion. *Cognition and Emotion, 15*, 831–852.

Zaidel, E. (1985). Language in the right brain. In D. F. Benson & E. Zaidel, *The Dual Brain: Hemispheric specializations in humans* (pp. 205–231). New York: Guilford Press.

Zatorre, R. J. (1997). Hemispheric specialization of human auditory processing: Perception of speech and musical sounds. In S. Christman (Ed.), *Cerebral Asymmetries in Sensory and Perceptual Processing*. Amsterdam: Elsevier.

Zatorre, R. J. (2001). Neural specializations for tonal processing. In R. J. Zatorre & I. Peretz (Eds.), *The Biological Foundations of Music* (pp. 193–210). New York: New York Academy of Science.

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends in Cognitive Science, 6*, 37–46.

Zatorre, R. J., Evans, A. C., & Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience, 14*, 1908–1919.

Index

12-tone scale 62, 63, 66, 88, 105, 118, 263
24-tone scale 266

A
Abeles, M. 188
Alcock, K. J. 18, 19, 29
Aleksander, I. 164
Alivisatos, B. 15
Anaki, D. 31
Anderson 170
Annett, M. 26, 226
anterior commissure 132
Arnoldi, H. M. 188
auditory system 12, 48, 49, 61, 87
augmented chord 63, 72, 76, 77, 259, 266, 268
awareness 137, 154, 156, 160, 161, 166, 179–181, 184–186, 189–192, 194–203, 207, 208, 216, 217, 228, 237, 239, 240, 243

B
Baars, B. A. 207, 238
Banich, M. T. 137
Beeman, M. 27, 31
Behrens, S. J. 28, 29
Bell, J. S. 114, 165
Bharucha, J. J. 60, 140
Bianki, V. L. 149
bilaterality 137, 205, 206
Binder, J. 69
Boersma, P. 61, 110
Bogen, J. E. 33, 36, 215, 236
Bohm, D. 165
Bolinger, D. L. 85, 86, 103
Bottini, G. 31
brain code vii–x, 43, 123–125, 138, 139, 144, 145, 153, 241, 242
Brown, S. 87, 100
Brownell, H. H. 30
Bulman-Fleming, M. B. 29
Bunge, M. 165
Burgess, C. 31, 44

C
cell membrane 167, 168, 171, 185, 194, 203, 205
central dogma viii–x, 40, 43, 45, 47, 123–125, 127, 152, 153
cerebral cortex 15, 22, 34, 37, 124, 125, 129, 131, 137, 152, 187
Chafe, W. 106
Chalmers, D. 161, 165
Chomsky, N. 102, 126
Clynes, M. 107
complementarity 15, 20, 33, 136
conscience 198
consonance 52, 53, 57–59, 61, 63, 66, 69, 72, 80, 84, 87, 99, 101, 263
context viii, x, 7, 8, 29, 30, 32, 33, 35, 41, 43, 47, 80, 87, 89, 94–96, 100, 102, 104–106, 108, 109, 139–141, 152, 156, 196, 197, 207, 208, 221, 231, 233, 234, 236–239, 247, 253–256, 262
Cook, N. D. 17, 28, 40, 44, 53, 63, 64, 66, 68, 75, 79, 128, 132, 134, 135, 143, 144, 175, 197, 210, 226, 228, 236, 239, 266

Cooke, D. 83, 86, 87, 101, 104, 105, 109, 245, 262
Corballis, M. C. 209
corpus callosum vii, 10, 14, 28, 35–37, 39–41, 124, 125, 131–135, 137, 138, 144, 152, 210, 211, 225, 235, 242
Costa, M. 53, 104
Creutzfeldt, O. 37, 130, 226, 235
Crick, F. viii, 138, 160, 166, 188, 200, 206, 215
Crow, T. A. 12, 23
Cruttendon, A. 85, 102, 103

D
D'Esposito, M. 15, 16
Deacon, T. W. 126, 145
Denenberg, V. H. 210
Dennett, D. 164, 207
diatonic music 73, 84, 88, 89, 93, 96, 99, 100, 105, 106, 125, 242, 245, 246, 249, 251, 252, 256, 260, 262, 263
diminished chord 56, 62, 63, 70, 72, 75
dominance 3, 8, 14, 15, 25, 26, 38, 41, 43, 44, 58, 85, 86, 89, 93, 98, 120, 139, 211, 226, 237, 253, 254
Doty, R. W. 210

E
Eady, S. J. 85
Edelman, G. 166, 205, 220, 235
Efron, R. 9
emotions 3, 60, 83, 88, 90, 93, 94, 97–100, 105, 108, 110, 119, 120, 139, 141, 143, 144, 152, 153, 241, 242, 245, 246, 251–255, 259–261
Engel, A. K. 166, 188, 191
excitation 127, 129, 138, 142, 145, 174, 213, 222, 224, 235, 236, 242
explanatory gap 162–164, 166, 179, 194, 199, 200, 202

F
feeling viii, x, 2, 3, 60, 87, 94, 109, 127, 141, 161, 162, 164, 175, 183, 185, 190–196, 199, 201–205, 240, 242, 243, 253, 254
Fitzsimons, M. 107
Fletcher, P. C. 37, 38
fMRI 30, 37, 66, 67
folk songs 249
Fonagy, I. 102, 106
Fox, P. T. 26, 225, 226
Frick, R. W. 97, 102
fringe 207–211, 213, 227–231, 233–239
frontal cortex 16, 38, 124, 132, 211

G
Gainotti, G. 30
Galaburda, A. M. 209
Galin, D. 206–208, 228, 237, 238
Gardner, H. 31, 32, 236
Gazzaniga, M. S. 10, 209, 210
Gårding, E. 103
Geschwind, N. 32, 209
Gibbs, R. W. 33
Gill, H. S. 15, 16
Gott, P. S. 209
Gray, C. M. 132, 188
Greenfield, P. M. 40
Grimshaw, G. M. 29
Gross, M. 16, 133, 167, 174
Grossberg, S. 221
Grout, D. J. 59, 103
Guernesey, M. 84
Guiard, Y. 137

H
handedness 13, 17, 26, 27, 40, 41, 44
harmony 19, 20, 56, 59–61, 63–65, 69–73, 75–78, 80, 83, 84, 86–89, 96, 101, 104–106, 109, 119, 124, 139–141, 144, 241, 242, 247, 248, 251, 255–258, 260–263
Harrington, A. 9

Harrison, D. 83, 245
Hauser, M. D. 93
Hebb, D. O. 159
Hedreen, J. C. 133
Heilman, K. M. 137
Hellige, J. B. 9
Helmholtz, H. 61, 62, 70, 71, 78, 104
HERA 37–40, 44
Heun, R. 37
Hevner, K. 60
homotopic connections 44, 226
Howell, P. 108
Hugdahl, K. 19–21
humor 32, 83
Huron, D. A. 59, 249

I
Iacoboni, M. 31, 36
image generation 15, 16
inhibition 135, 142, 144, 145, 209, 213, 222–229, 231, 236, 242
instability 65, 80, 96, 120, 265, 266, 268, 269
ions 168, 171–177, 184–186, 189, 191, 193, 195, 196, 200, 203
Ivry, R. B. 236

J
James, W. 207, 208, 228, 237, 238
Jaynes, J. 209
Jones, E. G. 11, 215
Jourdain, R. 59

K
Kaestner, G. 53
Kameoka, A. 51, 53, 57, 62, 70, 265
Kaminska, Z. 86, 104, 262
Kandel, E. C. 160
Karbe, H. 144
Kashiwagi, A. 41
Kiefer, M. 31
Kiehl, K. A. 30
Kim, S. G. 13, 14
Kinsbourne, M. 33, 137, 209, 210, 228
Klouda, G. V. 29, 34
Koch, C. 173, 188, 215
Kohonen, T. 136, 214, 216–219, 221
Kosslyn, S. 15, 27
Kuriyagawa, M. 51, 53, 57, 62, 70, 265

L
Ladd, D. R. 133
Landis, T. 33, 163, 209, 210
laterality 8–11, 13, 15, 25, 31, 33, 44, 47, 137, 138, 208, 209, 225, 241
Lee, A. C. H. 38
left hemisphere 6–8, 13–15, 18–21, 28, 40–43, 47, 124, 129, 135, 136, 138, 141, 142, 145, 152, 209, 242
Leonard, C. L. 33
Levelt, W. J. M. 49, 50
Levine, J. 162, 165, 200
Liederman, J. 137
limbic system 141
Loewenstein, W. R. 169, 173
logic gates 160, 200

M
Maass, W. 159
MacLennan, B. 156, 182–184, 194, 196
major mode 78, 83, 86, 125, 248, 252, 253, 256, 258
Mangan, B. 207, 208, 237
McGinn, C. 162, 165
Mehta, Z. 15
melody perception 124
mental rotation 15–17
metaphor 30, 31, 33–35, 49, 101, 205
Mexican hat 135, 222–225, 227, 228, 233, 235
Meyer, L. B. 75, 80, 84
minor mode 76, 80, 83–87, 89, 90, 99, 105, 119, 124, 125, 245, 247–249, 251–253, 258, 261, 263

Morton, E. W. 85
Moscovitch, M. 210
Mountcastle, V. B. 132

N
Netz, J. 144
neural networks 136, 159, 177, 180
neurotransmitters 2, 134, 168–170, 173–176, 184, 239
nucleus viii, 22, 207, 209, 227–231, 233, 235, 237–239
Nyberg, L. 37
Nygaard, L. C. 29

O
O'Regan, J. K. 164
object recognition 159, 206, 210, 211, 213, 214, 216, 221, 228, 229, 231–234, 236, 239, 240
Ogura, A. 186
Ohala, J. J. 85, 88, 89
Ornstein, R. 207, 209

P
paramecium 185, 186, 196
Parncutt, R. 53, 58, 59, 71–73, 265
Passingham, R. E. 25, 226
Pell, M. D. 29
Penfield, W. 134
Penrose, R. 160, 165
Perischetti, V. 83
Pfeifer, R. 158, 215
phonemes 19, 21, 27, 108
phrases 27, 31, 40, 104, 105, 137
Pierce, J. R. 49, 51, 59, 62
pitch perception 29, 71
Plomp, R. 49, 50
Poeppel, E. 105
Poggio, T. 114
Posner, M. I. 127, 129
Praat™ software 110, 111
Pribram, K. H. 166, 184
Price, K. H. 7
prosody 28, 29, 41, 48, 56, 95, 97, 98, 100–102, 124
psychopathology 12
Puccetti, R. 210

Q
Querné, L. 44

R
radial basis functions 114
Rameau, J.-P. 78, 83, 245
Reeves, A. G. 131
Reggia, J. A. 134–136, 144
Regnault, P. 60
resolved chords 63, 67, 85, 87, 266
retinotopic cortex 127
rhythm perception 29
Riemann, H. 78, 84, 245, 247, 250
right hemisphere vii, 6, 8, 9, 13–15, 18–21, 25–27, 29, 35, 41–43, 47, 48, 69, 124, 131, 135, 136, 138, 139, 142, 143, 145, 152, 209, 212, 226, 242
Roberts, L. A. 64, 236
Rodel, M. 31, 227, 228
Roederer, J. G. 59
Roelfsema, P. R. 188, 215
Rosen, G. D. 132
Ross, E. D. 29, 34, 35
Rushworth, M. F. S. 14, 40

S
Schachter, D. S. 238
Schenker, H. 101, 105, 106
Scherer, K. R. 86, 97, 98
Schiffer, F. A. 11, 209, 210
Schillen, T. B. 188, 215
Schirmer, A. 29
Schneiderman, E. I. 32
Schoenberg, A. 78, 80, 83, 245, 262, 263
Scruton, R. 59, 101
Searle, J. R. 161, 163
Semmes, J. E. 236

sensitivity 156, 165, 184, 185, 195–199, 201, 205, 239, 247
Sethares, W. A. 51, 58, 62, 64, 70, 263, 265
Shallice, T. 38
Shamma, S. A. 130
Shastri, L. 159, 187–189
Shear, J. 161, 182
Shkuro, Y. 135
simulation 134, 135, 137, 173, 188, 205–216, 218, 225, 226, 228, 231, 233–240
Singer, W. 109, 112, 138, 166, 188, 191, 215
Sloboda, J. A. 84
Sperry, R. W. 10, 34, 209, 210
Springer, S. P. 209, 236
Stapp, H. P. 165
Stein, B. E. 127
Steriade, M. 11, 215
stuttering 26, 27, 226
Swain, J. 83
synapse 134, 166, 168, 169, 171, 173, 174, 176, 177, 191, 195
synchronization 137, 138, 142, 144, 166, 183, 187–189, 191, 194, 195, 204, 215, 231, 240

T
Tallal, P. 29
Taylor, J. G. 166
Taylor, K. I. 30
TenHouten, W. D. 34
Terhardt, E. 53, 71, 73, 265
Tervaniemi, M. 19, 21
Tillmann, B. 106
tonotopic cortex 129
Trainor, L. J. 60
Tramo, M. J. 78
triads 55, 63, 64, 72, 73, 76, 87, 104, 108, 117, 252, 259, 262
tritone 52, 58, 61, 248, 249, 261
Tulving, E. 37

U
unresolved chords 55, 60, 63, 65, 67–70, 72, 73, 76–78, 87–89, 266–268
upper partials 51–54, 57, 58, 60, 62, 64, 65, 68–71, 77, 78, 85, 87, 140, 246, 263, 265, 267–269

V
Van Lancker, D. 29, 31
Velmans, M. 207
visual system 127, 133, 216
von der Marlsburg, C. 188

W
Walker, E. H. 165
Wallin, N. L. 100
Wapner, W. 32
Wedin, L. A. 60
Weenink, D. 61, 110
Weiskrantz, L. 9
Weniger, D. 28
Wiener, N. 159
Wigan, A. L. 210
Winner, E. 31, 32, 222
words viii, 5–7, 25, 27, 29–31, 34–36, 38, 39, 42, 60, 62, 64, 65, 69, 83, 95, 98, 102, 104, 109, 112, 118, 125–127, 132, 153, 158, 159, 175, 185, 186, 195, 207, 208, 220, 249, 251, 257, 260, 262
Wurm, L. H. 29

Z
Zaidel, E. 31, 36, 42
Zatorre, R. J. 20, 27, 29, 69

In the series ADVANCES IN CONSCIOUSNESS RESEARCH (AiCR) the following titles have been published thus far or are scheduled for publication:

1. GLOBUS, Gordon G.: *The Postmodern Brain.* 1995.
2. ELLIS, Ralph D.: *Questioning Consciousness. The interplay of imagery, cognition, and emotion in the human brain.* 1995.
3. JIBU, Mari and Kunio YASUE: *Quantum Brain Dynamics and Consciousness. An introduction.* 1995.
4. HARDCASTLE, Valerie Gray: *Locating Consciousness.* 1995.
5. STUBENBERG, Leopold: *Consciousness and Qualia.* 1998.
6. GENNARO, Rocco J.: *Consciousness and Self-Consciousness. A defense of the higher-order thought theory of consciousness.* 1996.
7. MAC CORMAC, Earl and Maxim I. STAMENOV (eds): *Fractals of Brain, Fractals of Mind. In search of a symmetry bond.* 1996.
8. GROSSENBACHER, Peter G. (ed.): *Finding Consciousness in the Brain. A neurocognitive approach.* 2001.
9. Ó NUALLÁIN, Seán, Paul MC KEVITT and Eoghan MAC AOGÁIN (eds): *Two Sciences of Mind. Readings in cognitive science and consciousness.* 1997.
10. NEWTON, Natika: *Foundations of Understanding.* 1996.
11. PYLKKÖ, Pauli: *The Aconceptual Mind. Heideggerian themes in holistic naturalism.* 1998.
12. STAMENOV, Maxim I. (ed.): *Language Structure, Discourse and the Access to Consciousness.* 1997.
13. VELMANS, Max (ed.): *Investigating Phenomenal Consciousness. Methodologies and Maps.* 2000.
14. SHEETS-JOHNSTONE, Maxine: *The Primacy of Movement.* 1999.
15. CHALLIS, Bradford H. and Boris M. VELICHKOVSKY (eds.): *Stratification in Cognition and Consciousness.* 1999.
16. ELLIS, Ralph D. and Natika NEWTON (eds.): *The Caldron of Consciousness. Motivation, affect and self-organization – An anthology.* 2000.
17. HUTTO, Daniel D.: *The Presence of Mind.* 1999.
18. PALMER, Gary B. and Debra J. OCCHI (eds.): *Languages of Sentiment. Cultural constructions of emotional substrates.* 1999.
19. DAUTENHAHN, Kerstin (ed.): *Human Cognition and Social Agent Technology.* 2000.
20. KUNZENDORF, Robert G. and Benjamin WALLACE (eds.): *Individual Differences in Conscious Experience.* 2000.
21. HUTTO, Daniel D.: *Beyond Physicalism.* 2000.
22. ROSSETTI, Yves and Antti REVONSUO (eds.): *Beyond Dissociation. Interaction between dissociated implicit and explicit processing.* 2000.
23. ZAHAVI, Dan (ed.): *Exploring the Self. Philosophical and psychopathological perspectives on self-experience.* 2000.
24. ROVEE-COLLIER, Carolyn, Harlene HAYNE and Michael COLOMBO: *The Development of Implicit and Explicit Memory.* 2000.
25. BACHMANN, Talis: *Microgenetic Approach to the Conscious Mind.* 2000.
26. Ó NUALLÁIN, Seán (ed.): *Spatial Cognition. Selected papers from Mind III, Annual Conference of the Cognitive Science Society of Ireland, 1998.* 2000.
27. McMILLAN, John and Grant R. GILLETT: *Consciousness and Intentionality.* 2001.

28. ZACHAR, Peter: *Psychological Concepts and Biological Psychiatry. A philosophical analysis.* 2000.
29. VAN LOOCKE, Philip (ed.): *The Physical Nature of Consciousness.* 2001.
30. BROOK, Andrew and Richard C. DeVIDI (eds.): *Self-reference and Self-awareness.* 2001.
31. RAKOVER, Sam S. and Baruch CAHLON: *Face Recognition. Cognitive and computational processes.* 2001.
32. VITIELLO, Giuseppe: *My Double Unveiled. The dissipative quantum model of the brain.* 2001.
33. YASUE, Kunio, Mari JIBU and Tarcisio DELLA SENTA (eds.): *No Matter, Never Mind. Proceedings of Toward a Science of Consciousness: Fundamental Approaches, Tokyo, 1999.* 2002.
34. FETZER, James H.(ed.): *Consciousness Evolving.* 2002.
35. Mc KEVITT, Paul, Seán Ó NUALLÁIN and Conn MULVIHILL (eds.): *Language, Vision, and Music. Selected papers from the 8th International Workshop on the Cognitive Science of Natural Language Processing, Galway, 1999.* n.y.p.
36. PERRY, Elaine, Heather ASHTON and Allan YOUNG (eds.): *Neurochemistry of Consciousness. Neurotransmitters in mind.* 2002.
37. PYLKKÄNEN, Paavo and Tere VADÉN (eds.): *Dimensions of Conscious Experience.* 2001.
38. SALZARULO, Piero and Gianluca FICCA (eds.): *Awakening and Sleep-Wake Cycle Across Development.* 2002.
39. BARTSCH, Renate: *Consciousness Emerging. The dynamics of perception, imagination, action, memory, thought, and language.* 2002.
40. MANDLER, George: *Consciousness Recovered. Psychological functions and origins of conscious thought.* 2002.
41. ALBERTAZZI, Liliana (ed.): *Unfolding Perceptual Continua.* 2002.
42. STAMENOV, Maxim I. and Vittorio GALLESE (eds.): *Mirror Neurons and the Evolution of Brain and Language.* n.y.p.
43. DEPRAZ, Natalie, Francisco VARELA and Pierre VERMERSCH.: *On Becoming Aware.* n.y.p.
44. MOORE, Simon and Mike OAKSFORD (eds.): *Emotional Cognition. From brain to behaviour.* 2002.
45. DOKIC, Jerome and Joelle PROUST: *Simulation and Knowledge of Action.* n.y.p.
46. MATHEAS, Michael and Phoebe SENGERS (ed.): *Narrative Intelligence.* n.y.p.
47. COOK, Norman D.: *Tone of Voice and Mind. The connections between intonation, emotion, cognition and consciousness.* 2002.
48. JIMÉNEZ, Luis: *Attention and Implicit Learning.* n.y.p.
49. OSAKA, Naoyuki (ed.): *Neural Basis of Consciousness.* n.y.p.

3 9346 00844809 6

NO LONGER PROPERTY OF
FALVEY MEMORIAL LIBRARY

"Tone of Voice and Mind" is a synthesis of findings from neurophysiology (how neurons produce subjective feeling), neuropsychology (how the human cerebral hemispheres undertake complementary information-processing), intonation studies (how the emotions are encoded in the tone of voice), and music perception (how human beings hear and feel harmony). The focus is on the psychological characteristics that distinguish us from other primate species. At a neuronal level, we are just another mammalian species, but the functional specialization of the human cerebral hemispheres has resulted in three outstanding, uniquely-human talents: language, tool-usage and music. To understand how the human brain coordinates those behaviors is to understand who we are.

Without taking hemispheric asymmetry into account, there is simply no way forward in unraveling the functioning of the human brain. Cook takes the argument a stage further in showing that the mysterious "non-dominant" hemisphere contains a 2-D representation of the affective charge of the linguistic message that balances the syntactic structures in the dominant hemisphere. These developments in the theory of bi-hemispheric specialisation will be followed closely by all those with an interest in the evolution of the hominid brain.

Timothy J. Crow MD, Professor of Psychiatry, University of Oxford

Norman Cook has done it again! Some fifteen years after his puzzling and controversial book on "the brain code", he presents his matured views on basic brain functioning in a brilliant writing on consciousness and how the brain might code emotions in musical terms. A book full of insights in a search for the underlying rules that guide the brain's behavior.

Theodor Landis MD, Chairman of Neurology, University of Geneva

This fascinating book is remarkable both in the breadth of the scientific knowledge it brings to bear on understanding the mind, and in the presentation of a "central dogma" for neuropsychology. Informative and provocative, it leaves one with a new sense of what it is to be human.

James A. Reggia MD, Professor of Computer Science and Neurology, University of Maryland

John Benjamins Publishing Company

ISBN 90 272 5174 6 (Eur) / 1 58811 276 4 (US)