Auditory perception has always been the stepchild of psychology. The rapid advances of computer technology have, if anything, further increased the hegemony of the visual sense: The prototypical computer combines stunning graphic capabilities with a primitive sound inventory, and it seems useless without a monitor, whereas a loudspeaker can easily be dispensed with. The optic display capabilities of computers are utilized widely in psychological research, their sound-generating capabilities only by a few specialists. Even in those branches of psychology that ostensibly deal with audible things, such as psycholinguistics and psychomusicology (not to speak of their ancient armchair relatives, linguistics and musicology), theory and experimentation are commonly based on visual representations of the objects under study. Books on auditory subjects (such as the three reviewed here) usually have plenty of figures, but no sound sheets.

It comes rather as a shock, therefore, when Stephen Handel opens his book with the confession that “In our culture, I would much prefer to be blind than to be deaf” (p. 1). With this simple but soon convincing statement, he reminds us of the unique importance of the auditory sense to our life experience: More than vision (but rather like the tactile sense, which is much more limited in range), audition keeps us “in touch” with our environment. Moreover, it is the basis of the two most important systems of human communication: speech and music. If television had no sound, it would never have edged out radio as the most popular medium of news and entertainment.

Much of the psychological research on hearing in recent decades has employed simple sounds and sophisticated psychophysical methods; this “psychoacoustics” continues to thrive in a somewhat segregated fashion in many laboratories and in the pages of The Journal of the Acoustical Society of America. Another significant area in auditory research is speech perception, which is even more of a segregated specialty, called “speech science” (or “speech technology”) as soon as some application is in sight, and otherwise largely associated with work done at Haskins Laboratories and the reactions of others to it. Most of the speech perception work has been at the fringe of psychoacoustics, with speech sounds being taken apart into their smallest components until they ceased to be speech and researchers felt on familiar territory again. Much the same can be said about research on music perception, a large part of which has been concerned with pitch, duration, and loudness.

At the same time, a few small rivulets began to flow beside the mainstream of reductionistic auditory research (itself a minor tributary to the St. Lawrence of largely vision-based psychology). James Gibson, in his influential discussions of ecological principles of perception, had relatively little to say about audition (though there is a chapter in Gibson, 1966), but his emphasis on environmental objects and events, and on the perceiver’s attunement to systematically structured physical media, spread seeds which germinated during the 1970s. In the 1980s, Carol Fowler emerged as a champion of an ecological perspective on speech perception (see Fowler, 1986), and James Jenkins wrote a stimulating chapter presaging a science of ecological acoustics (Jenkins,
1985). Albert Bregman's research program on auditory organization had been under way for some time and yielded a steady stream of research reports, with occasional contributions from others (e.g., Kubovy, 1981), applications to speech perception (e.g., Darwin, 1984), and a lively counterpoint of inventive experiments from Richard Warren's laboratory (e.g., Warren, 1984). Music psychology, a relatively obscure enterprise through the 1970s, suddenly gained momentum through publications such as the book edited by Deutsch (1982), Roger Shepard's (1982) influential article on pitch structures, and the extensions of his work by his former student, Carol Krumhansl.1

The three books reviewed here reap the harvest of these developments. One of them, Handel's Listening, is a broad introduction to the perception of auditory events, with special attention to speech and music. Bregman's Auditory Scene Analysis focuses more narrowly on the perceptual organization of simple sound patterns, but treats this topic expansively, with the author's own ideas and research at the center of attention. Krumhansl's Cognitive Foundations of Musical Pitch is even more specialized in that it summarizes the author's research since the late 1970s, with only brief digressions into related literature. It is also considerably more succinct than the other two tomes, and it does not share their overt ecological orientation, being squarely in the tradition of cognitive psychology. What all three books have in common is excellence.2

AUDITORY EVENTS
Handel's book begins with a detailed but very readable introduction to the physics of sound production, presented without mathematical formulas but with many illustrations. This is followed by a chapter on sound propagation in the environment, by two chapters dealing specifically with sound production by musical instruments and by the human vocal tract, and by a chapter summarizing acoustic (and, very briefly, perceptual) commonalities between speech and music. The remaining chapters, which constitute roughly two thirds of the book, deal with issues of perception. The first of these chapters is on auditory stream segregation and lucidly overviews a topic treated in much more detail in Bregman's book. Chapter 8, "Identification of Speakers, Instruments, and Environmental Events", is particularly valuable in pointing out the common aspects of these important activities, which have been given less research attention than they deserve. Chapter 9 deals primarily with categorical perception and context effects, with the focus on speech. Under the unfortunately misleading title, "Grammars of Music and Language", the next chapter deals almost entirely with music, particularly pitch structures, anticipating Krumhansl's more detailed treatment. (Was an earlier section on linguistic grammar deleted in the last minute?) The following chapter on rhythm is more balanced and provides a very useful discussion of music in juxtaposition with prosodic aspects of speech. The final chapter, somewhat surprisingly, is on auditory physiology, but summarizes what is known about the auditory processing of complex sounds and speech, so that it ties in well with the general thrust of the book. A brief epilogue points out two aspects that were neglected in the book: the role of the listener's expectations and knowledge, and a characterization of the experience of listening.

Handel's book contains a wealth of information, presented accurately, in simple prose, with numerous instructive examples. It brings together, often for the first time, topics that have been treated in articles scattered through the research literature, and it provides a coherent perspective on them. The writing is modest, thoughtful, and balanced; there is no dogma or strident criticism, nor any oversimplification of complex issues. Handel always shows a healthy respect for the complexity of natural phenomena, and he inculcates the same attitude upon the receptive reader. As Albert Bregman says on the book jacket, "Listening is obviously the work of a master teacher."

AUDITORY SCENE ANALYSIS
Bregman's own book, Auditory Scene Analysis, is narrower in scope than Handel's but probes the topic in much greater depth. At 773 pages surely one of the heftiest monographs ever published in psychology, it rests heavily on Bregman's own research since the late 1960s and on the contributions of a few other scientists working in the same area. Its leisurely, narrative style at times gives it the quality of a historical or philosophical treatise. In a very real sense, Bregman serves as the historian of his own ideas and research. One rarely gets such an intimate view of a scientist's mind at work, nor such a comprehensive picture of personal observations, experimental explorations, and alternative interpretations. Bregman invites the reader to join him on his intellectual journey, and I, at least, found the book difficult to put down. If Handel's book shows a master teacher at work, then this is the product of a master thinker.

The term "auditory scene analysis" was coined by Bregman to refer to the process of organizing
complex auditory input into internally coherent “streams” or auditory objects. He distinguishes two classes of such processes: “primitive” and “schema-based” stream segregation. The book deals primarily with primitive processes, which do not depend on a listener’s domain-specific knowledge. Bregman believes (although he acknowledges that further research is needed) that primitive scene analysis segregates auditory events before they are interpreted with reference to learned “schemata,” as in listening to speech or music.

Following an introductory chapter, more than half of the book is taken up by Chapters 2 and 3, which deal with sequential (temporal) and simultaneous (spectral) integration/segregation, respectively. Chapter 2 introduces the now well-known phenomenon of auditory stream segregation and the seminal work of van Noorden (1975)—surely the most cited unpublished dissertation in the field—and proceeds to discuss exhaustively what is known about the various factors that influence the perceptual grouping of acoustic elements. Chapter 3 discusses the factors that cause simultaneous tones to fuse into a single percept or to be perceived as separate pitches or timbres. It covers a good deal of more traditional psychoacoustics (such as pitch perception, binaural fusion, masking, etc.), but Bregman never strays very far from his own research and brings in the findings of others primarily to illuminate or supplement the story of his enterprise.

The reader who has been persistent enough to plow through these two enormous but fascinating chapters, each a small book in itself, is faced with five additional, shorter chapters. Chapter 4, “Schema-Based Integration and Segregation” is shorter because Bregman’s goal is to distinguish and separate knowledge-guided processes from primitive auditory scene analysis, and to keep the focus on the latter. Perhaps the most important theoretical argument of the book is that primitive scene analysis is independent of acquired knowledge, though what has been divided by scene analysis can sometimes be recombined into a higher-level (schema-based) unit. In the popular jargon of contemporary cognitive science (which Bregman studiously avoids), primitive scene analysis is modular and noninteractive. Chapters 5 and 6 deal with auditory organization in music and speech perception, respectively. Again, these discussions focus on the role of primitive scene analysis, not on the perceptual consequences of the categories and structures specific to each system. Thus they address such basic questions as “What makes a melody hang together?” and “What makes the different voices in polyphonic music distinct?”, or in speech, “Why are the sounds of speech perceived as a coherent stream?” and “How do we separate several simultaneous voices from each other?” The parallel nature of these questions in music and in speech reflects the universality of auditory scene analysis. Music- and speech-specific knowledge is considered a nuisance factor from the perspective of this book, which treats music and speech as pure sound. This may disappoint some musicologists and linguists among the readers, but Bregman should not be blamed for saying little about topics that his book is not about; rather, the rigor of his approach must be admired, for there is a continuous temptation to elevate (or, rather, reduce) knowledge-based processes to the status of auditory primitives, particularly in the case of speech.

Chapter 7 presents a relevant case study. Under the heading of “The Principle of Exclusive Allocation in Scene Analysis”, Bregman discusses the phenomenon of duplex perception, an instance in which the principle is violated (i.e., the same sound appears to be heard as part of two different streams). In fact, Alvin Liberman and his collaborators at Haskins Laboratories claim that speech schemas (Bregman’s term) override and even “pre-empt” auditory scene analysis (see, e.g., Liberman & Mattingly, 1989). Bregman discusses evidence to the contrary. Still, the issue remains somewhat unresolved at the end of the chapter, which is the most difficult and the least definitive in the book. The last chapter, “Summary and Conclusions: What We Do and Do Not Know about Auditory Scene Analysis” condenses the book’s contents onto 65 pages, much for the benefit of readers who just want to get the gist of it. Here and throughout the volume, Bregman’s honesty in acknowledging unresolved questions and missing empirical evidence is exemplary. There are many leads for future research to be done, and Bregman’s accomplishment is made all the more impressive by his careful delineation of its current limits. This book will stand as an important milestone in the history of 20th century psychology, as well as an inspiring human document.³

**PITCH STRUCTURES**

With Krumhansl’s monograph we enter a different world, yet one that dovetails nicely with Bregman’s and especially with Handel’s introduction. Krumhansl is concerned with some of the
schema-guided processes in music perception that exceeded the scope of Bregman's book, specifically the relationships among the pitches of the Western tonal system. The monograph is a natural outgrowth of Krumhansl's exceptionally systematic and coherent research program, which is almost unique in the burgeoning field of music psychology.

Lucid and organized throughout, Krumhansl's writing lacks the old-fashioned charm of Bregman's meandering thoughts. Instead, there is a crystalline quality to her orderly designs and structural representations. Clearly, her most distinctive achievement is in the domain of sophisticated quantitative analysis. As one of Roger Shepard's most brilliant students in the 1970s, she absorbed the multidimensional techniques pioneered by her mentor and proceeded to apply them to musical problems in an imaginative and revealing way. Despite the formal complexity of these analyses, she makes the results always easy to grasp, with the help of many illustrations which are an essential part of the methodology. Having accorded Handel and Bregman master status, without intending to stereotype them in any way, I regard this book as the work of a master analyst.

Only a very brief summary of the contents can be given here. To convey the full flavor and elegance of Krumhansl's research, a much longer precis would be necessary, which will appear elsewhere (Repp, in press). Krumhansl's primary experimental technique is a probe task in which a musical context (a melodic fragment, sequence of chords, or excerpt from a composition) is followed by a probe tone or chord, whose adequacy as a continuation of the preceding context the listener is to judge. Probe elements are sampled exhaustively from a fixed set (such as the 12 tones of a scale), and a profile of average ratings across these elements is obtained. The autocorrelation matrix of this profile, which represents the similarities of the rating profiles for the same probe elements in the context of all different keys, is subjected to multidimensional scaling, which results in a spatial representation of keys similar to such maps constructed intuitively by musicologists. The ingenious part of the methodology begins when the key rating profiles established in the initial experiments are used as diagnostic tools for determining the perceived key following some arbitrary context. Thus Krumhansl devises a key-finding algorithm based on the frequency distribution of the most recent notes and their correlations with all possible key profiles; by presenting probe tones after some musical excerpt, she determines which (if any) key is dominant at that point by finding the prototypical key profile that most closely resembles the obtained rating profile; and in the most advanced application of these procedures, she traces the listener's changing sense of key through a modulating sequence of chords by obtaining a probe tone profile after each chord, correlating each of these profiles with the standard key profiles, and finally mapping these relationships into a multidimensional space, where they trace a modulatory path among the various keys (represented as points in the space). Highwire acts such as these are complemented by results from simpler memory tasks and other studies in the literature.

The central chapter of the book is Chapter 6, which defines three principles of tonal stability, and their effects on the perceived relations between tones. Tonal stability is the central concept of the book, and indeed of traditional music theory; it refers to the fact that, in tonal music, there is a hierarchy of pitches, such that one pitch (the key-defining pitch or tonic) is most preferred or most important or most representative—in other words, most stable—at any given time, a second pitch (the dominant) is preferred next, and so on. Krumhansl's three principles, then, are (in simplified language): A stable tone seems more similar to itself (e.g., is remembered better) than an unstable tone; two tones seem more similar to each other if either of them is stable; and two tones of unequal stability seem more similar when the unstable tone precedes the stable one. A variety of evidence supports these statements, which parallel predictions made with regard to prototypicality in many other areas of cognitive psychology. Krumhansl's work indeed falls squarely into mainstream cognitive psychology and should contribute significantly to making music psychology seem part of this larger enterprise.

THREE FIELDS ON THE MOVE

It is perhaps appropriate to conclude this review with some musings on the current state of three fields of research that are addressed by the books reviewed (ecological acoustics, speech perception, and music perception), and on the influence the books might have on research in the 1990s. As it happens, the three fields named are at rather different stages of development: one nascent, one burgeoning, and one temporarily stagnant. These impressions are subjective, of course, and depend in large measure on where I draw the boundaries of these domains of inquiry.
**Ecological acoustics.** Under this rubric I would consider studies that deal with the analysis and perception of information in complex sounds other than the message elements of speech or music—information that helps us identify individuals, objects, and events in our environment. (Ultimately, of course, the “ecologically valid” study of speech and music as gestural events must be included, too.) Under this definition, Bregman’s work is merely a prolegomenon to an ecological acoustics, though an essential one. Handel’s Chapter 8 (“Identification of Speakers, Instruments, and Environmental Events”) is most pertinent, inasmuch as speaker and instrument identification are not really linguistic or musical activities. The literature on human speaker identification is relatively small (much smaller than that on automatic speaker recognition), and most of it originates in Europe. Research on instrument identification is almost nonexistent. The related topic of the acoustic expression and perception of emotion in speech and music is likewise under-researched, with Klaus Scherer’s work on speech standing as a single beacon in the desert (see Scherer, 1986). Handel discusses Warren and Verbrugge’s (1984) study of breaking and bouncing events, which is a prototype for ecological acoustics research built on Gibsonian premises, but little has happened since except for a few isolated studies on seemingly exotic topics including “chilling” sounds (Halpern, Blake, & Hillenbrand, 1986), hand clapping (Repp, 1987), and the sounds of kitchen pans struck with mallets (Freed, 1990). However, those who doubt the potential significance of studies in ecological acoustics may be converted by reading Tom Johnson’s (1984) still unpublished dissertation on doctors’ perception of human heart beats, a brilliant foray into real-world relevance. The message of all these studies is that we hear not just sounds but, through their structure, environmental happenings and organisms in action. Hopefully, Handel’s book will stimulate more research on how we use our ears to perceive actions and events—how we hear the world.

**SPEECH PERCEPTION**

Research on speech perception began at Haskins Laboratories in the early 1950s and largely remained dependent on the technology available there for the next two decades. Then, with computers getting smaller and cheaper, and with software replacing hardware synthesizers, other laboratories got into the business. Much of that research, however, remained methodologically and conceptually dependent on the Haskins research: Nearly everyone tried to support, refute, or extend the claims of the Haskins researchers. (Among the few significant exceptions, Richard Warren’s consistently original—though psychoacoustically tinged—contributions are especially noteworthy; see, e.g., Warren, 1982, 1984.) The 1970s and early 1980s were fertile years for speech perception research, with several popular paradigms being milked dry and lively arguments going back and forth. Now these activities seem to have slowed down, very much in proportion to the decline of speech perception research at Haskins Laboratories, where most of the effort is nowadays directed to speech production. The older generation of speech perception researchers has reached retirement age, and many of the younger (now middle generation) protagonists of the 1970s have turned to different topics or tend to publish less, with only a few die-hards continuing to suck on the dry teats of their superannuated paradigms. There seems to be a general lack of intellectual ferment in the field.

Was speech perception research (as defined rather narrowly here) just a historic episode? Did Bregman in his chapter on duplex perception capture the last, already somewhat peripheral controversy? Perhaps, as far as the dominant and unifying (or, rather, constructively divisive) role of Haskins Laboratories is concerned. It may take a while before new ideas develop and strong voices emerge to put them forward. Handel’s book will make only a minor contribution here; what is needed is a coherent body of research that makes a point, comparable to what Krumhansl and Bregman have to offer. The ideas of the most innovative theorist in recent years, Carol Fowler, hold much promise but have not yet resulted in a critical mass of empirical findings. Meanwhile, of course, there is a rich matrix of related research activities in experimental phonetics, speech science, speech technology, and psycholinguistics, which are not experiencing a similar recession and may provide hotbeds for new directions in speech perception research.

**MUSIC PERCEPTION**

Music psychology, and particularly research on music perception, is on the rise. One of the prime movers is Diana Deutsch who has earned the socio-scientific triple crown by editing the first modern collection of articles on the subject (Deutsch, 1982), founding the journal *Music Perception* in 1983, and by recently establishing the Society for Music Perception and Cognition, in addition to be-
ing a fertile and original researcher at the psychoacoustic end of the music spectrum. A number of other significant books have appeared in recent years, of which Sloboda's (1985) is the most original, and music-related conferences abound. There is a rapidly increasing pool of talented young investigators, each of whom quickly seems to find a niche in the vast territory offered by musical questions and phenomena. Interdisciplinary conferences bring psychologists together with music technologists, musicologists, composers, and performers—some sceptical, to be sure, but all eager to exchange ideas and explore new avenues. Electronic instruments, sophisticated software, and MIDI systems offer new and exciting possibilities for research and practice. As an added special touch, a shared love for music unites scientists of very different theoretical persuasions: The fact that music gives aesthetic pleasure and spiritual sustenance is never far from their minds and frequently invades their discussions, whereas speech researchers, for example, rarely think of drama or poetry in connection with their work.

Krumhansl's book rides the crest of a wave to which she herself contributed significantly. The book serves mainly to bring her work to the attention of those who have not been following her progress in the journals, and it is admirably suited for that purpose. The research itself, of course, has led and influenced the field for some time, and also has aroused some controversy (Butler, 1989; Krumhansl, 1990; Butler, 1990)—a healthy sign of a science's vitality (cf. Hull, 1988). Unlike Bregman's life work, which has the quality of a fortress under construction, with open doors but numerous escape routes, all thoroughly explored in advance of any possible attack, Krumhansl's work, with its carefully planned design, its built-in dependencies among experiments, and its strong reliance on one particular methodology, appears much more vulnerable and transparent, more like a contemporary office building in a historic neighborhood. It remains to be seen whether her constructs and methods can withstand critical onslaught. Meanwhile, her book is required reading for anyone interested in the contemporary psychology of music, as indeed are the other two volumes reviewed here. In concert, this admirable trio should convince anyone that auditory perception is worthy of much more attention by psychologists than it has received in the past.

REFERENCES


FOOTNOTES

1 Naturally, these are just selected highlights which happened to leave a strong impression on me.

2 What is missing from my shelf is a research monograph on speech perception from a cognitive or ecological perspective.

3 A useful supplement to the book would have been a sound sheet or CD illustrating the auditory phenomena discussed in the book. Some years ago, Bregman produced a cassette with such demonstrations and distributed copies to colleagues; I understand that copies can still be obtained by sending $5.00 to him at McGill University.