10 The Climate of Auditory Imagery and Music*

Diana Deutsch
University of California, San Diego

John R. Pierce
Stanford University

THE TRADITION OF IMAGERY

> From where do I take my ideas? That I cannot say with certainty. They come uncalled, directly and indirectly. I could grasp them by my hands in the freedom of nature, on walks in the silence of the night or in the early morning through moods which turn into tones which sound, blow, storm, until the notes are standing before me. (Cited in Dorian, 1947, p. 8)

These words were written by Beethoven in 1824, three years before his death. The astonishing creations of this great composer which were written decades following the onset of his deafness are perhaps the most striking examples of imagery, in any modality.

Indeed, the procedures generally employed by the masters of composition indicate that feedback, whether auditory, or in the form of symbolic description, is not a necessary part of the compositional process, and might even interfere with it. Composers have frequently asserted that they work over an entire piece mentally, elaborate and polish it in considerable detail, before they set anything down in writing. Again from Beethoven:

> I carry my thoughts a long time, often very long before I write them down. Therein my memory remains loyal to me, since I am sure not to forget a theme even after years, once I have conceived it. Some things I change, reject, try all over again until I am satisfied. (pp. 251–252)

---

*This chapter is based on numerous discussions between the authors when one of us (DD) was a Visiting Scholar at the Center for Computer Research in Music and Acoustics (CCRMA) at Stanford, in the Spring of 1989. The authorship is listed alphabetically, but should be considered as equal.

In: D. Reisberg (Ed.) Auditory Imagery, 1992, Erlbaum

Other composers have given similar accounts. As Gluck wrote:

> Once I am clear about the composition of the whole and about the characterization of the main parts, I consider the opera finished, although I have not yet written one single note. (p. 254)

And again:

> First I go through each individual act, later through the whole work. The plan of the composition I sketch in my mind while sitting in the parterre (of the theatre). Such a preparation usually takes one entire year. (p. 254)

And from Weber:

> Before I approach the execution of the detail, I figure out the great plan of the tonal picture through determining its main colors and its individual parts; namely, I outline for myself the exact sequence of the keys . . . and I strictly weigh the use of the instruments. (p. 255)

This process occurred with such precision that, as his son Max wrote:

> Without any intermediary stage, the whole score would flow out of his pen from the flute to the double bass like an etching. (p. 255)

Mozart also adopted this procedure as his favorite mode of composing, regarding the final writing of the score as an irksome process. As he wrote in 1780 "Composed is everything, written not yet a single note" (Dorian, 1947, p. 256).

Indeed, many composers were not only able to write scores without preliminary sketches, but they also wrote their music directly for the different instrument parts without reference to a full score. For this reason, much music has come down to us in the form of individual parts alone. Composers frequently intended to conduct their pieces themselves, and full scores would for them have been redundant. It was therefore frequently left to others to assemble together the different parts of an orchestral piece (Dorian, 1947).

The process of musical composition not only furnishes us with examples of highly specific and elaborate long term auditory memory, but also gives us insight into how such imagery might be formed. Musical ideas were frequently described as sudden inspirations, often occuring dramatically, for example in dreams. Thus Wagner in his autobiography describes how the Prelude for his *Rheingold* was conceived:

> I fell into a kind of somnolent state, in which I suddenly felt as though I were sinking in swiftly flowing water. The rushing sound formed itself in my brain into a musical sound, the chord of E-flat major, which continually reechoed in broken

> forms; these broken chords seemed to be melodic passages of increasing motion, yet the pure triad of E-flat major never changed, but seemed by its continuance to impart infinite significance to the element in which I was sinking, I awoke in sudden terror from my doze, feeling as though the waves were rushing high above my head. I at once recognized that the orchestral overture to *Das Rheingold*, which must have long lain latent within me, though it had been unable to find definite form, had at last been revealed to me. (p. 55)

Other composers, such as Berlioz, Chopin, Bruckner, and Cesar Franck, have also described the emergence of musical images in their dreams, and as sudden inspirations.

So far we have been considering musical composition, which constitutes an exceptional example of imagery. However, we should note that the successful performance of a piece also requires a high degree of imagery, and even naive listeners must use this faculty to a considerable extent.

EARLIER CONFLICTS

Given such considerations, it is remarkable that for the last few decades, auditory imagery has been neglected in most of the literature on psychoacoustics, and its existence has essentially been denied by many experimental psychologists. This situation is historically anomalous: in previous times the approach to hearing in terms of low-level or peripheral factors has coexisted with considerations of higher-level processing.

Aristoxenus, perhaps the leading music theorist in ancient Greece, argued strongly for a high-level approach. As he wrote in his Harmonics:

> And we must bear in mind that musical cognition implies the simultaneous cognition of a permanent and of a changeable element, and that this applies without limitation or qualification to every branch of music. (Cited in Macran, 1974, pp. 189–190)

And later:

> It is plain that the apprehension of a melody consists in noting with both ear and intellect every distinction as it arises in successive sounds—successive, for melody, like all branches of music, consists in a successive production. For the apprehension of music depends on these two faculties, sense-perception and memory; for we must perceive the sound that is present and remember that which is past. In no other way can we follow the phenomena of music. (p. 193)

The epoch of the scientific revolution was a very fertile one for the study of hearing; most of those responsible for the striking advances in mathematics, astronomy and mechanics of this period also made important contributions to music. Amongst these were Mersenne, Galileo, Kepler, Huygens, and Descartes. It was at this time that the relationships between pitch and rate of vibration in strings, pipes and bells were determined; the phenomenon of beats was discovered; as was the overtone series; and in addition issues such as tuning and temperament, and consonance and dissonance hotly debated.

On this last issue, both low-level and high-level factors were proposed. For example, Galileo espoused a low-level approach, a sort of rhythmic theory of consonance. Consonance resulted from a distinct pattern of two tones beating on the eardrum, and dissonance from irregular beats.

> The Offence [the Dissonances] give, proceeds, I believe, from the discordant and jarring Pulsations of two different Tones, which, without any Proportion, strike the Drum of the Ear: And the Dissonances will be extreme harsh, in case the Times of the Vibrations are incommensurable [. . .] Those Pairs of Sounds shall be Consonances and will be heard with Pleasure, which strike the *Timpanum* in some Order, which Order requires, in the first Place, that the Percussions made in the same Time be commensurable in Number, that the Cartilage of the *Timpanum* or Drum may not be subject to a perpetual Torment of bending itself two different Ways, in submission to the ever disagreeing Percussion. (Cited in Cohen, 1984, p. 90)

Mersenne also adopted a low-level approach, but placed the locus of the interactions a little further along the auditory pathway:

> . . . the external air excites the air inside the ear, and it impresses a state of motion upon the auditory nerve that resembles the one it received; and the mind that is present in each part of the body, and consequently in the said nerve, perceives at once the movement of the organs of the ear, and thereby judges the qualities of the motion of the sound, and of the external objects that produce it. Now one could imagine that the mind is like an indivisible and intellectual point, to which all sense impressions taper, like all lines of a circle towards their center, or like all the threads of the web of the spider that spins and weaves them . . . (p. 110)

Kepler, in contrast, argued that low-level factors were inadequate to explain the phenomena:

> If the velocity of one string has the ability to move another, proportionate string that to the eye appears not to be moved, then would not the equal velocities of two strings have the ability pleasantly to titillate the ear, since in a certain sense it is excited uniformly by both strings, and since the two strokes of both tones or vibrations coincide every moment? No, I say, this is too simple a way to settle the matter [. . .] For please, what relationship could there be between the titillation of the sense of hearing, which is a corporal thing, and the incredible delight that we perceive deep within our soul through the harmonic consonances? If the delight comes from the titillation, would not then the main part of this delight be played by

the organ that sustains the titillation? In my *Dioptrics* it seemed best to me to define any sense organ in such a way that the sense perception that brings forth pleasure or grief is not completed until the species of the organ that is destined for the perception in question, as it is affected from outside, has reached inwards, through the guidance of the spirits, the tribunal of the general sense. Hence I now ask, what part in this delight in hearing consonant voices and notes the ear has? Do we not sometimes feel pain in our ears when we listen to music and, because of the horrible blare, shut them with our hands, although we continue to perceive the consonances and our heart jumps? Add to this that the explanation taken from motion is applied in the first place to the unison; but sweetness does not primarily lie in the unison; but in the other consonances and the combination thereof. Much [more] might be adduced in order to destroy this alleged explanation of the sweetness that comes from the consonances; but for the moment I prefer to desist from a more detailed disquisition. (p. 31)

Descartes also had misgivings about invoking a low-level explanation of consonance alone, and proposed what today is described as the difference between sensory consonance and musical consonance:

For it should be observed that all these calculations [on coinciding strokes] serve only for showing which consonances are the simplest, or, if you prefer, the sweetest and the most perfect ones, but not on that account the most agreeable [. . .] . But in order to determine what is most agreeable, one should consider the capacity of the listener, that changes like taste, according to the person in question; thus some will prefer to hear one single melody, others part-music, etc; just as the one prefers what is sweet, and the other what is somewhat bitter or acid, etc. Concerning the sweetness of the consonances, two things should be distinguished: namely, what renders them simpler and more accordant, and what renders them more agreeable to the ear. Now, as to what renders them more agreeable, that depends on the places where they are employed: and there are places where even dimished fifths and other dissonances are more agreeable than consonances, so that one could not determine absolutely that one consonance is more agreeable than another. One could, indeed say, however, that, normally speaking, the thirds and sixths are more agreeable than the fourth; that in cheerful songs major thirds and sixths are more agreeable than minor ones, and the opposite in sad [songs], etc., in that there are more occasions where they can be employed agreeably. But one can say absolutely which consonances are the most simple and the most accordant ones; for that depends only on how often their sounds unite, and how closely they approach the nature of the unison; so that one can say absolutely that the fourth is more accordant than the major third, while ordinarily it is not so agreeable, just as cassia is definitely sweeter than olives, but not so agreeable to our taste.[1] (p. 169)

One can indeed sympathize with the desire to invoke partially understood peripheral structures as explanatory devices, and to avoid resorting to more

[1]We may note that Descartes wrote after thirds had been admitted to music.

abstract types of explanation; however this approach is not without its pitfalls. Such pitfalls are illustrated in a report by Everard Home, in his 1799 Croonian lecture to the Royal Society of London on the "Structure and Uses of the Membrana Tympani of the Ear." Having described his researches on the tympanic membrane in elephants (reasoning, as Bekesy did in our time, that the animal's size makes it easier to observe its anatomical features) he compared the ear drum to a monochord "of which the membrana tympani is the string; the tensor muscle the screw, giving the necessary tension to make the string perform its proper scale of vibrations; and the radiated muscle acting upon the membrane like the moveable bridge." He then proceeded to assert that accuracy of music perception is determined by the degree of perfection in the activity of these muscles:

> The difference between a musical ear and one which is too imperfect to distinguish the different notes in music, will appear to arise entirely from the greater or less nicety with which the muscle of the malleus renders the membrane capable of being truly adjusted. If the tension be perfect, all the variations produced by the action of the radiated muscle will be equally correct, and the ear truly musical; but, if the first adjustment is imperfect, although the actions of the radiated muscle may still produce infinite variations, none of them will be correct: the effect . . . will be similar to that produced by playing upon a musical instrument which is not in tune. (Cited in Miller & Cohen, 1987, p. 30)

THE TURN OF THE CENTURY

Let us consider next the approach to hearing taken by those in the latter part of the last century and the beginning of this one. Here again we find that representation of complex structures was not neglected, and imagery was considered alongside phenomena that allowed for peripheral explanations.

The work that was perhaps the most influential for twentieth century psychoacoustics was Helmholtz's *On the Sensations of Tone* (1885/1954). Helmholtz not only dealt with issues that were consistent with peripheral explanations, such as the pitch of complex tones, beats, and combination tones, but he also gave detailed consideration to higher-level musical phenomena such as scales, modes, root progressions of chords, systems of key relationships, and the concept of tonality. Our ability to learn, remember, recall and call up played a considerable part in his thought. Amongst his several discussions of imagery he wrote:

> The effort felt in singing the leading note does not lie in the larynx, but in the difficulty we feel in fixing the voice upon it by mere volition while another tone is already in our mind, to which we desire to pass . . . (pp. 286–287)

and later:

> Supposing that I have been used to hearing Fifths taken at all possible pitches, and have recognized them by aural sensation as having a very close melodic relationship, I should know the magnitude of this interval by experience for every tone in the scale, and should retain the knowledge thus acquired by the actions of a man's memory of sensations, even of those for which he has no verbal expression . . .
>
> And just in the same way I shall be able to recognise, as previously known, other melodic passages or whole melodies which are executed in simple tones, and even if I hear a melody for the first time in this way, whistled with the mouth or chimed by a clock, or struck on a glass harmonicon, I should be able to complete it by imagining how it would sound if executed on a real musical instrument, as the voice or a violin.
>
> A practiced musician is able to form a conception of a melody by merely reading the notes. If we give the prime tones of these notes on a glass harmonicon, we give a firmer basis to the conception by really exciting a large portion of the impression on the senses which the melody would have produced if sung. (pp. 289–290)

The Gestalt psychologists, who were primarily concerned with issues of high-level representation, made frequent reference to music in their discussions. At that time, Mach (1898/1943) raised the question of "Whether there is anything similar to the symmetry of figures in the province of sounds".

> Now, although in all the preceding examples I have transposed steps upward into equal and similar steps downward, that is, as we may justly say, have played for every movement the movement which is symmetrical to it, yet the ear notices little or nothing of symmetry. The transposition from a major to a minor key is the sole indication of symmetry remaining. The symmetry is there for the mind but is wanting for sensation. No symmetry exists for the ear, because a reversal of musical sounds conditions no repetition of sensations. If we had an ear for height and an ear for depth, just as we have an eye for the right and an eye for the left, we should also find that symmetrical sound-structures existed for our auditory organs. (p. 103)

Mach was canny in differentiating what is pattern to the eye and what is pattern to the ear.

DOCTRINES IN MUSIC AND SCIENCE

Figure 10.1 illustrates the potential symmetry of which Mach writes. It is interesting to compare Mach's conclusion with the following statement by Schoenberg (1975) who, although not a psychologist, propounded what he thought

FIG. 10.1. Musical example given by Mach, presented in both original and inverted form. This was used to illustrate Mach's conclusion that passages related by inversion are not perceptually equivalent. Adapted from Mach (1898/1943), Fig. 26, p. 102.

should be or must be true of musical images as a justification for his method of twelve-tone composition:

> THE TWO-OR-MORE DIMENSIONAL SPACE IN WHICH MUSICAL IDEAS ARE PRESENTED IS A UNIT . . . The elements of a musical idea are partly incorporated in the horizontal plane as successive sounds, and partly in the vertical plane as simultaneous sounds . . . *The unity of musical space demands an absolute and unitary perception.* In this space . . . there is no absolute down, no right or left, forward or backward . . . To the imaginative and creative faculty, relations in the material sphere are as independent from directions or planes as material objects are, in their sphere, to our perceptive faculties. Just as our mind always recognizes, for instance, a knife, a bottle or a watch, regardless of its position, and can reproduce it in the imagination in every possible position, even so a musical creator's mind can operate subconsciously with a row of tones, regardless of their direction, regardless of the way in which a mirror might show the mutual relations, which remain a given quantity. (pp. 220–223)

His theory of abstract representation of tones series is illustrated in Fig. 10.2. In referring to the illustration he writes "The employment of these mirror forms corresponds to the principle of the absolute and unitary perception of musical space" (p. 225). For Schoenberg, musical space must be just like the space of vision.

In sharp contrast to the concerns of composers, the general approach of 20th-century psychoacoustics has been to disregard issues of higher-level representation, and to focus instead on highly circumscribed phenomena which could, it is assumed, be explained in terms of simple neural activity, preferably in terms of the actions of the peripheral hearing apparatus. A huge amount of data has amassed concerning, for example, absolute thresholds, difference thresholds for frequency and loudness, the masking of one tone by another, the lateralization of sinusoids and noise bands, adaptation and fatigue. Similarly, auditory physiologists have focused on the response of the auditory system to narrowly defined stimuli, and have addressed a few circumscribed problems; for example they have examined the response of the basilar membrane to clicks and to sinusoids, plotted tuning curves for neural units in various parts of the auditory system, described relationships between stimulus amplitude and firing rate, mapped out tonotopic organization in central auditory structures; and so on.

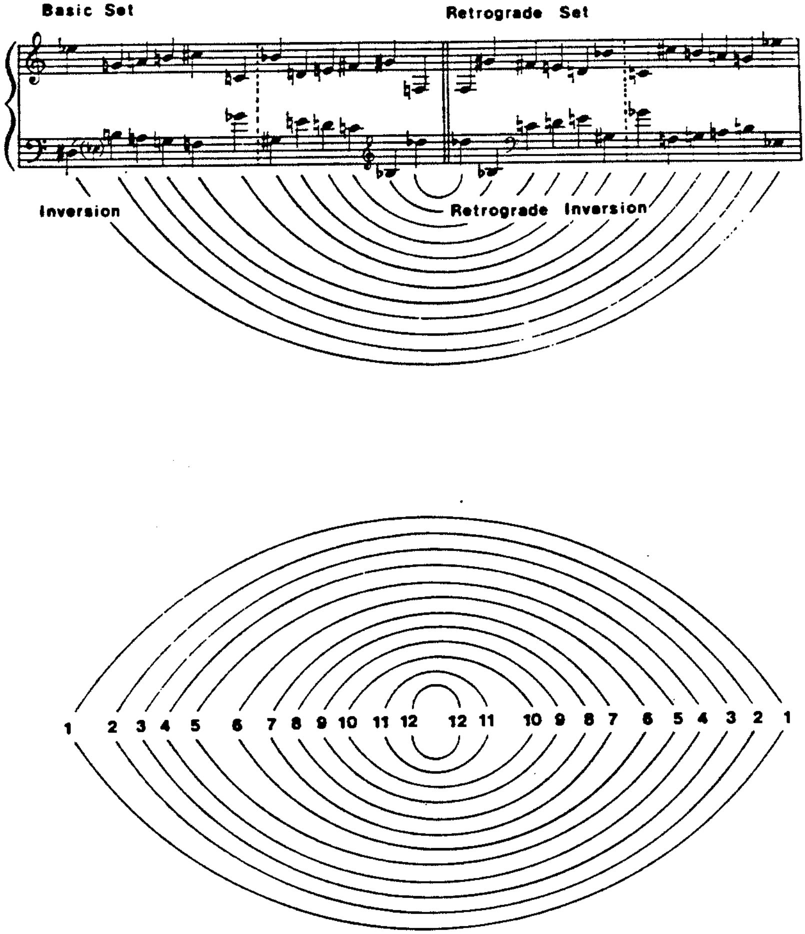

FIG. 10.2. Musical example given by Schoenberg. This was used to illustrate Schoenberg' assumption that passages related by inversion and retrogression are perceptually equivalent. Adapted from Schoenberg (1975), Example 4, pp. 224–225.

This reductionistic approach has gathered much particular information, some of which has value. Unfortunately, it has been coupled with the doctrine that higher-level issues should be excluded from consideration until low-level functioning is entirely understood. The problem here is that one cannot, logically, come to any satisfactory conclusion by following this approach. More specifically, one cannot exclude central or complex factors by focusing only on low-level stimuli or on explanations in terms of peripheral mechanisms; the perceptual characteristics at issue could in principle be taking place anywhere in the processing chain. On the contrary, it is *peripheral* or *simple* explanations which can logically be excluded, by demonstrating that effects occur as a result of central interactions, or as a result of some complex process. Thus, even assuming that one's sole aim is to understand the action of peripheral mechanisms, it is necessary to generate stimulus patterns which enable the determination that observed effects are indeed peripheral rather than central in origin, or that they are based on simple rather than on complex operations. Thus issues involving central structures and complex operations cannot, logically, be avoided.

We attribute the reluctance to consider high-level processing in hearing to a number of factors. The first has been the difficulty involved in generating complex stimuli with sufficient stimulus control; this severely inhibited researchers from embarking on experiments using such stimuli. Fortunately, over the last few decades this problem has been overcome. With new advances in sound generation and analysis by computer, stemming primarily from work begun around 1960 by Max Mathews (see Mathews, 1969), researchers now have available the tools with which to address any issues concerning sound perception, however complex, and they can do so with both versatility and precision.

A second factor that led researchers to focusing on low-level parameters was the enthusiasm resulting from new understandings of the workings of the inner ear. It is understandable that researchers would attempt to attribute as much as possible to the characteristics of newly discovered mechanisms. We are reminded here of Galileo's enthusiasm for the tympanic membrane as the organ responsible for the perception of consonance and dissonance, ignoring the rather obvious points made by Kepler and by Descartes. On the other hand, we see clearly in Home's Croonian lecture on the muscles controlling this membrane the pitfalls inherent in being overly enthusiastic about some known peripheral structure as an explanatory device.

A third factor (and this is perhaps the most important one) is a lack of awareness of musical issues in our culture. In the case of vision, specialized knowledge is not required in order to pinpoint some of the analyses that the nervous system would need to perform in order for useful perception to occur. Obvious operations include, for example, the recognition of objects when these are presented in different orientations, or at different distances from the observer. However, in the case of hearing, analogous operations do not readily present themselves, except to the musically initiated. As we have seen, in previous times scientists were musically knowledgeable, and this enabled them to consider, as Helmholtz and Mach did, high-level along with low-level factors in their explanations of the hearing process.

In this context we can also consider the approach taken by experimental psychologists to auditory memory. The framework that has generally been adopted for work on human memory is a three-stage model consisting of a sensory register, a short term store, and a long term store. The first detailed version of this model was proposed by Broadbent (1958), who hypothesized that information presented to the observer is first retained in a large capacity sensory system, where it decays very rapidly. The material is saved from obliteration only through a process of verbal encoding, which enables it to enter short term memory. Information is there saved from obliteration by verbal rehearsal, until it is transferred to long term memory. Several later versions of the three-stage model have been proposed (see, for example, Neisser, 1967).

Despite the general implausibility of the notion of an extremely short term store for auditory memory, certain experimental results have been cited in its favor (for example, Pollack, 1972; Treisman & Rostron, 1972). However, the results of these experiments are also consistent with alternative explanations in terms of interference effects that occur under highly specialized stimulus conditions (see Deutsch, 1975a, for an extended discussion). The incontrovertible fact remains that most people can remember quite complex music over entire lifetimes. This cannot be reconciled with the view that auditory memory decays within a second or two unless it is recoded into verbal form. Again, we attribute the widespread acceptance of this view, which must surely be incorrect, to the lack of consideration of musical phenomena.

A fourth factor has arisen through the advanced computer technology that, through the work of Max Mathews, has so expanded the complexity and variety of sounds that can be generated and used in experiments on music and musical imagery. With the computer has come work on artificial intelligence and work on neural networks and other varieties of computational neural science.

Artificial intelligence, or AI, purports to gain insight into human function by programming computers to perform a variety of "human" tasks. Computational neural science, or CNS, purports to gain insight into human behavior by proposing computer models of the functioning of the nervous system. Both deal with high-level as well as low-level behavior.

What have AI and CNS told us about human capabilities and human behavior? Both deal with the behavior of complicated computer programs rather than with the behavior of actual organisms. It is not clear that, in general, workers in AI and CNS are aware of what is actually known about human behavior and its neurophysiological substrate—or particularly care.

SEARCH FOR SYNTHESIS

We have discussed some of the conflicts and hazards in the current relation among science and technology on the one hand and music and its imagery on the other. The interrelations involve conflicting doctrines various in source and scope. Yet, our time is a time of unprecedented resources for exploration and enlightenment.

Computer technology has supplied us with resources for generating complex musical sounds, and for generating quasi-musical sounds in which various aspects of quality and perception have been intensified or eliminated. Computer technology has also made it far easier to devise and conduct experiments and to gather data.

Audio technology, including the computer generation of sounds and their recording and distribution on audio cassettes, and lately, on compact disks, have opened new potentialities of illustration and demonstration in the study of musical perception and imagery.

In the field of vision, illustrations in books and the projection of patterns fixed or in motion have played an essential role in discovering, confirming, and demonstrating various visual phenomena. Today the illustration of auditory and musical phenomena through recordings, or through direct digital sound generation, can have a similar value in auditory perception.

Our argument from here forward is that science and technology have provided us with new and varied opportunities for studying music and musical imagery. In exploiting these profitably, we must choose wisely among fields of endeavor that are widely different in procedures, standards, modes of presentation, and relevance to musical imagery.

Among other things, we must keep in mind that it is clear from current psychological studies that the human being is no *tabula rasa*. Some of our capabilities are built in. Some capabilities must be learned. But, we can learn only skills that are within our scope. Some distinctions, ready reactions to some sort of formal order, are beyond our inherent capabilities.

Along with the nature and variety of the physical phenomena essential to musical instruments, human capabilities of perceiving and distinguishing must be the stuff of which music is made. Fortunately, some realization of this is becoming apparent among psychoacousticians and cognitive psychologists, if not in the fields of AI and CNS.

It is easy to say that we should be cautious of wrongheaded simplification in psychoacoustic demonstrations and experimentation. It is also easy to say that we should get at the root of important matters, rather than carefully gathering huge amounts of narrow data that give us no deep insight. Our hope and belief is that various different and competing *aspects* of perception can be sorted out and studied in a meaningful, insightful way.

In any simplification important things are bound to be left out. We can only recommend caution in interpreting and applying to music and its perception the results of experiments and demonstrations made with simplified, quasi-musical sounds in a non-musical or barely musical environment.

THE ROLE OF NEUROPHYSIOLOGICAL STUDIES

We have advocated a holistic approach and inveighed against a specialization that loses meaning in detail. Nonetheless, detail can be important. Between the instrument and the ear, music travels as sound waves. Within the ear vibrations of air excite vibrations of little bones, and these excite vibrations in the fluid of the cochlea and the basilar membrane. The basilar membrane performs a rough frequency analysis, and different regions, corresponding to different small ranges of frequency, excite different nerve endings of the auditory nerve. The nerve impulses on the auditory pathways up to the auditory cortex are processed in various interconnecting and bypassed way stations. When we speculate about mechanisms of perception and imagery, we should try to keep our ideas consistent with what is known about the auditory system.

In the past 20 or so years a great deal has been learned about the structure and functions of auditory pathways. A volume such as *Auditory function: Neurobiological bases of hearing* (Edelman, Gall, & Cowan, 1988) will convince any reader that it is not easy to understand the current status of auditory neurophysiology. It is clear, however, that important advances have been made over the background given in recent books on hearing. Rhode (1970, 1971) found that in live animals the frequency analysis performed by the basilar membrane is much finer grained than deduced by von Bekesy (1960) through studying cadaver material; nevertheless, fairly current textbooks on hearing still show von Bekesy's drawings. It is found from animal studies (Schreiner & Langer, 1988) that in general only low time rates of excitation are preserved up to the auditory cortex—at most a few tens of changes a second will be followed. Time response is faster lower down in the neural pathways of the auditory system, but it is only as far as the first interaction of the two ears that response is followed to a small fraction of a millisecond, and that time differences of around 10 microseconds can be distinguished.

Some aspects of animal hearing have been worked out inconsiderable detail. This is particularly true of the auditory system of barn owls (Konishi, Takahashi, Sullivan, Wagner, & Carr, 1988; Knudsen, 1988). The details of binaural time and loudness comparisons for various frequency ranges, and the neural pathways that, through inhibition, overcome ambiguities in source location have been explicated. The bringing together in one part of the brain of visual and auditory clues to object location has been demonstrated.

Where does the relevance of such knowledge to musical imagery lie? It lies chiefly in restricting our range of sensible speculation. It does not tell us what to believe about human perception, but it can indicate things that we should not believe.

It is very difficult to use what little we know about auditory pathways in the study of music perception. Certainly, however, we should not hypothesize functions or processes that seem grossly at odds with what we know of the anatomy and physiology of hearing.

THE THREAT OF MODELS

As we have noted, the capabilities of human hearing are astonishing. Complicated anatomical structures and physiological processes must lie behind these capabilities. We are certain concerning some of these, as the rough frequency analysis performed in the cochlea, and the importance of both relative intensity and relative time of arrival of sounds in the two ears in judgment of the location of a source of sound. Many details of the processes of perception are not known, even when we sense important aspects.

The psychoacoustic literature is full of models, including models of pitch perception. Some models have been worked out in analytical detail beyond present sure neurophysiological or psychophysical verification. Autocorrelation models of pitch perception (Licklider, 1951) call for functions that have not been found in the auditory system, and involve amounts of delay of waveform or its features that seem neurophysiologically dubious.

One model for the perception of virtual pitch (Terhardt, 1974) calls for the internal creation of a subharmonic frequency that early work (Fletcher, 1924) indicates to be perceptually unimportant when present. In the work cited, Fletcher himself felt impelled to hypothesize the generation of such a fundamental frequency internally through nonlinearities, presumably because Helmholtz had erroneously concluded that pitch is perceived only through the fundamental frequency.

As we have noted, psychoacoustic speculators have turned to electronic or computer simulated "neural networks," or have made various computational models in seeking explanation of auditory perception. Such work is commonly pursued without any adequate demonstration that the networks or functions hypothesized actually correspond to processes found in behavioral and neurophysiological explorations of animal hearing.

An engineer may wish to make a machine that will do something previously done by people or animals. Often there will be several possible avenues open, as in the case of flying. We should not argue from the success of the airplane that birds fly by spinning a propeller. Some psychoacousticians write as if the making-up of an unverified and sometimes unverifiable model lends creditability to a proposed explanation, even when there is no evidence for that the processes hypothesized have been found in animals, or are plausible on the basis of what is known. In our view, most models are a distraction from thought and an inspiration to certitude for the uncertain.

CAUTIONS ON THE USE OF TRAINED LISTENERS

We have all heard of Pavlov's dogs which, through training, salivated at the sound of a bell. This does not mean that the sound of a bell and the smell or sight of food are similar sensations. It means that through training the same "objective" behavioral response can be evoked by different stimuli.

Is it fanciful that the evocation of the same responses through different sensations could lead to erroneous conclusions in psychoacoustic experiments? We don't think so. Let us note that a tune can be recognized through listening to it, or through a musician examining a score. And, that a tune can be recalled either as a sequence or sounds that just "come to" the person who sings or whistles the tune. Or, as a sequence of note names and durations.

There seems to be no danger here. But suppose that we are studying the perception of pitch. A tune with a very characteristic rhythmic pattern may be recognized by the rhythmic pattern alone. So, in experiments aimed at the study of pitch we should avoid tunes with rhythmic patterns. But a familiar tune can be recognized by its melodic contour, even when it is played with stretched or otherwise distorted intervals. Here we are on shaky ground, but the ground can be shakier.

The rate of periodic bursts of high-frequency tones can be distinguished as faster or slower. But, up to a few hundred tonebursts a second, the match of different patterns of tonebursts is on tonebursts per second, not on fundamental frequency or harmonic spacing, as it is at higher rates. And, when the match is on rate rather than fundamental frequency, changing the rate in a ratio of 5:4 does not result in the perception of a musical interval of a major third (Pierce, 1990).

Subsequent informal experiments have shown that for a subject with absolute pitch tonebursts with rates up to a few hundred a second periodic tonebursts sound like a buzz with no pitch. No identification occurs with a note of the scale or position on the keyboard. Yet, faster or slower can be distinguished.

Because faster or slower can be distinguished, it seems plausible that with sufficient training a subject could be taught to respond to higher rates with the term "higher pitch," and to lower rates with the terms "lower pitch," and perhaps even to respond with position on a keyboard, just as in the presentation of a note on a staff. Here is a chance for a real confusion.

In psychoacoustic experiments, much is made of well-trained subjects. Yet, objective responses to stimuli may depend on training. Thus, the similarity that musicians find in isolated chords and their inversions is not found in data on subjects without musical training (Mathews, Reeves, Pierce, & Roberts, 1988).

In the study of musical perception, experimental outcomes will in some cases depend on musical talent and training. In part, the responses of subjects without formal musical training will depend on exposure to and degree of care in listening to music. This makes it difficult to distinguish inherent abilities from those resulting from exposure to music.

In particular studies, prolonged training of subjects could lead to responses more characteristic of the training than of anything else. And, to "the right," "the expected," or "the requested" response as to sensation or image.

EXPERIMENTS VERSUS DEMONSTRATIONS

In connection with visual perception, we have seen wonderful demonstrations that leave the participant utterly convinced. These include apparent motion of flashing lights, errors of relative size of figures in perspective drawings, depth perception of random-dot stereograms—and a host of others. Such demonstrations establish a general fact concerning human response beyond all question, though details and quantitative laws may have to be worked out through time-consuming experiments.

Convincing demonstrations work when the salient fact to be demonstrated can be disentangled from other effects that would make it less clear and obvious. Alas, the psychoacoustic literature is full of wonderfully detailed experiments in which any clear point is hidden by obscuring processes, and any conclusion must be teased out by detailed analysis.

We hold that good demonstrations are more convincing than experiments. A good demonstration need not be preceded by special training which just *might* influence what we hear, or how we respond to it. In a good demonstration, that which is demonstrated is right out front; we can't approach a demonstration with a kit of statistical tools. But, a good demonstration is less easily arrived at than a "careful" experiment.

The good auditory demonstrations we know of involve ingeniously simplified stimuli whose perception casts light on other stimuli, such as musical sounds or our reactions to them. However there is some danger here. A powerful demonstration may not be applicable to what we really want to understand.

AMATEURS AT WORK

Computers and synthesizers have put tools of unprecedented power into the hands of talented musicians and, indeed, of talented tinkerers who are unused, as was Galileo, to the laborious protocols of experiment and analysis that have become standard in psychoacoustic literature.

Such "amateurs" are interested chiefly in powerful and demonstrable results. Some have contributed very important synthesis techniques and insights as, Chowning (1973) in fm synthesis, in his vividly convincing evocation of moving sound sources (1971) and in fm voice synthesis (1980, 1989); Risset in various paradoxical sounds and brass and bell sounds (Risset & Wessel, 1982), and Rodet, Potard, & Barriere (1984) in synthesis of the singing voice. Much of such work is to be found in print, other such work has been presented at computer music conferences, and some becomes scuttlebutt among the like minded. Much real insight is not to be found in the standard literature. One of us (JRP) was surprised to find that improvement of consonance by adequate frequency separation of higher harmonics is to be found in U.S. patent 4,117,413 filed by Robert A. Moog in 1977.

A wealth of important or potentially important insights, observations, and techniques may be escaping the scientific literature. It is not escaping those who wish to use it. Much of it is unlikely to be available as standard papers in standard journals. It is available through word of mouth and through publication in new journals with new standards. And, often it is available through a very convincing demonstration on a tape or even compact disc.

ANALYSIS BY SYNTHESIS

Analysis by synthesis is a very powerful tool. In essence, the investigator seeks a (usually simple) algorithm for producing a perceptual effect. In this search he may as legitimately be guided by hunches or black magic or chance observations as by keen experimental or analytical insight. It doesn't matter *how* he arrives at the algorithm. But, he succeeds only if the algorithm works, and works well.

Risset (Risset & Wessel, 1982) sought an algorithm for producing brasslike sounds. He waded through much analytical detail, relevant and irrelevant. What he finally found was that for a sound to be brasslike the components of higher frequency must rise later than those of lower frequency. Guided by this result, Mathews processed bowed tones in such a way that they sounded brassy.

Analysis by synthesis can be subject to abuse. The algorithm arrived at must be patently successful. It must not be "just sort of." It must find what is clearly a chief and sufficient, or almost sufficient ingredient. The origin of some consonants in frequency motions toward a following vowel (Cooper et al., 1952) is convincing, we guess, but we'd be more convinced and happier with examples that sounded more like human speech.

SOME STRONG EFFECTS AND THE DEMONSTRATIONS THEREOF

In this section we call attention to some powerful aspects of music perception, and illustrations of the decisive impact that ingenious demonstrations can have.

The Ratios of Small Integers and Harmonic Effects

The sense of consonance, of harmony, and of progression to a resolution are central to Western music. Since the day of the Greeks, the sense of consonance has been associated with the ratios of small integers: the octave 2:1, fifth 3:2, fourth 4:3, major third 5:4, and minor third 6:5.

The traditional explanation of Helmholtz, as elaborated by Plomp (1966), is that for these intervals, harmonic partials of the two tones will not beat disagreeably, because many either coincide in frequency (or for tempered intervals very nearly coincide), or are well separated in frequency. But, musicians recognize musical intervals and chords when the individual tones are sine waves, for which there is no such beating. Could it be that the ratios of small integers somehow gave us the diatonic scale and its intervals, and that the considerations that

Helmoltz and Plomp put forward are sufficient but not necessary in the appreciation of Western music? How could we possibly find out?

That beats, or the interactions of partials lying within a critical bandwidth, are absolutely essential in music of several voices is conclusively demonstrated by demonstration 31, tracks 58–61 of the IPO NIU ASA compact disc of Houtsma, Rossing, and Wagenaars (1987). The musical material is a four-part Bach chorale. The tones used have 9 partials. There are two versions of the scale: normal and logarithmically stretched to replace the octave by a ratio 2.1 : 1. There are two versions of the frequencies of the partials: harmonic (which goes with the normal, 2 : 1 octave) and stretched logarithmically to conform to the stretched scale.

When the stretched partials are used with the stretched scale the partials of two notes that coincide before stretching will coincide after stretching, and partials that lie well apart in frequency before stretching will lie well apart after stretching. Thus, if relative positions of the partials of the several simultaneous tones are the harmonically important criterion, the unstretched scale and tones will sound all right together, and the stretched scale and stretched tones will sound all right together, while the stretched tones will sound bad with the unstretched scale, and the unstretched tones with the stretched scale. This is what one hears in the demonstration, very strongly. It is clear that even small departures in scale *or* partials undermines the harmonic basis of Western music, while the same departure from scale *and* partials can be acceptable.

Percussive sounds may play a larger part in African and Indonesian music than in western music. In many percussive instruments, overtones are not harmonic in frequency. Do the intervals of the scales in such cultures depart somewhat from diatonic intervals, perhaps to produce intervals with less beating?

Perceptual Fusion

Chowning (1973) showed how complex spectra can be produced efficiently through fm (frequency modulation) of a sinusoidal carrier frequency. When he used this approach in synthesizing the singing voice (1980, 1989) he found that the harmonic frequency components depicting the formant structure of a sung vowel simply did not fuse with the fundamental frequency to produce the sound of a vowel. Rather, the percept was that of a nondescript collection of tones, a collection of tones that could lead to no sensible musical percept, to no image of a singer.

When all frequency components were given a common vibrato, they fused into a sung vowel sound. Chowning found that he could use this technique to attain the effect of three voices singing a major triad together, and, beyond that, of several voices singing together a note of the same pitch. Without the use of vibrato, there was no sensible musical image. With appropriate vibrato, the image of a chorus could be invoked.

The value of a common vibrato in making a complex sound evoke a musical image had been appreciated before Chowning's work. Chowning's contribution is that there can be an absolute necessity for common small erratic deviations in pitch and/or amplitude if a complicated assortment of component frequencies are to be heard as a fused sound coming from one source. Such common erratic variations are of course characteristic of the human voice and of traditional musical instruments.

Octave Equivalence

It is traditional to take into account two aspects of pitch: the note within the octave (called the *pitch class* or *chroma*) as, C, E, G, and the octave within which the note lies (sometimes called *pitch height*), which can be designated C4 (middle C), C5, C6, and so on.

Shepard (1964) produced well-known tones in which one goes around the pitch class circle, seemingly without changing pitch height.[2] These tones consist of octave-related partials whose amplitudes are given by a rising, then falling curve fixed with respect to frequency. As these octave-related partials are moved up together in small steps along the log frequency continuum (and so clockwise along the pitch class circle) new partials enter at the low-frequency end at levels below the threshold of perception, and exit at the high-frequency end, again at levels below the threshold of perception. An impression of an ever-increasing pitch can thus be attained. For this particular pattern, then, pitch height appears virtually constant. Physically, the C at the beginning of an octave's span of such movement results in the identical tone as the C at the end of the octave's span of such movement.

These quasi-musical tones exhibit stepwise changes of pitch class from note to note. But the final note at the beginning of what would be expected to be the next higher octave is instead identical to the initial note. This is octave equivalence with a vengeance.

What about octave equivalence for musical sounds? Musically, how much are the notes in succeeding octaves really alike? If they are *really* alike, shouldn't they be interchangeable?

This has been settled with a simple demonstration (Deutsch, 1972). A very well known melody ("Yankee Doodle") was chosen in which successive notes have the same length. This melody was played with notes of the correct pitch

[2]This effect has since been shown to result from pitch class or spectral proximity between temporally adjacent tones. When tone pairs composed of such octave-related partials are presented, such that the tones within a pair are diametrically opposed along the pitch class circle (C–F#, D–G#, and so on) so that the principle of proximity cannot be invoked, an orderly relationship appears between the perceived height of a tone and its region along the pitch class circle: Tones lying in one region are heard as higher and those lying in the opposite region are heard as lower (Deutsch, Kuyper, & Fisher, 1987). Interestingly, the form of relationship between pitch class and perceived height here varies substantially across listeners.

class, but the octaves in which the notes lay were chosen randomly over a three-octave range. Only very, very few listeners succeed in identifying the melody under these conditions.[3]

What, then, of octave equivalence? There is clearly "something in" this concept, but the above experiment shows that Cs, Ds, and Es in different octaves just aren't necessarily interchangeable. Indeed, accuracy of melodic contour may be more important than accuracy of pitch class in evoking a known melody.

Streaming

It is remarkable how a skilled player can bring out the separate parts of Bach's works for the solo violin. Experiments concerning and demonstrations of streaming, so well discussed in Bregman's (1990) book can illustrate this only in part. But, demonstrations of streaming are simple and striking.

The simplest is the alternate, periodic playing of notes of a lower and higher pitch. When the sequence of notes is slow we hear the notes as one stream, jumping up and down in pitch. At a high enough rate we hear two streams, repeated notes of higher pitch and repeated notes of lower pitch. We lose any clear sense of the time relation of the notes in the two streams.

The phenomenon of streaming is important in getting some sense of the organization we impose on what we hear, or how we envision sounds. Proximity in pitch will link sensations together when they are close enough in time—perhaps, close enough to indicate a common source. Difference in pitch tends to differentiate sensations as to the envisioned source.

Demonstrations of streaming can be made in which the upper and lower voices do not repeat in pitch, but constitute simple melodic patterns without large jumps in pitch. These, too, are fascinating and enlightening.

It is a long way from simple streaming experiments to the skilled player creating an image of two or more sound sources in playing a Bach composition for the solo violin. But, the simplicity and inevitability of demonstrations of streaming are impressive.

The Deutsch Octave and Scale Illusions

In listening attentively we commonly listen to a sound of interest in one ear or from one direction and ignore or suppress sounds in the other ear or from another direction. This human ability is as well known as the phrase "whisper in your ear."

It would and does seem remarkable that for ingenious musical sound patterns this commonplace ability can be overridden at a high level, and the sense of

---

[3]We may note that if listeners are given clues which enable them to hypothesise the correct melody, identification becomes much easier: Listeners are able to achieve such identification by matching each incoming note to their auditory image of the untransformed version, and so confirm that each note is indeed of the correct pitch class.

sound source rearranged so as to give a sense of a "musically simpler" presentation to the two ears.

Deutsch (1974) has shown that such high-level rearrangement of sense or image of source is possible by devising binaural presentations of successive notes of different pitches which demonstrate the listener's creation of a simple sound image from a "complicated" sequence of stimuli.

In the *octave illusion*, tones are presented simultaneously to the two ears by means of headphones. The stimuli to the right and left ears and what is heard by the right and left ears are:

Stimulus to right ear	G5	G4	G5	G4	G5	G4	G5
Stimulus to left ear	G4	G5	G4	G5	G4	G5	G4
Heard by right ear	G5		G5		G5		G5
Heard by left ear		G4		G4		G4	

Somehow, one of the notes is assigned to the right ear and the other to the left ear. In this case, G5 in the right ear predominates over G4 in the left ear, but G4 in the right ear is heard in the left ear while simultaneously G5 in the left ear is not heard. The ear to which the higher note is assigned differs somewhat among individuals. For right-handed subjects the note of higher pitch is usually assigned to the right ear.

In the *scale illusion* (Deutsch, 1975b) the pitches of the notes played to the right and left ear jump up and down a great deal, but the listener produces the image of two sources, in each of which the pitch changes from note to note by no more than one step.

Stimulus to right ear	C′	D	A	F	F	A	D	C′
Stimulus to left ear	C	B	E	G	G	E	B	C
Heard by right ear	C′	B	A	G	G	A	B	C′
Heard by left ear	C	D	E	F	F	E	D	C

In this case, regardless of which ear individual notes go to, notes which form a descending and ascending scale are assigned to one ear, and notes which form an ascending and descending scale are assigned to the other ear. Again, for most right-handed subjects the higher notes are heard as in the right ear, and the lower notes as in the left ear—the reverse is true for some subjects.

These illusions show that a listener may derive a plausible musical organization or image of sound sources that seems quite at variance with the stimuli presented to the two ears.

CONCLUDING REMARKS

We have noted the unhappy climate of auditory imagery and music. Until recently, at least, psychoacousticians and cognitive psychologists have not held musicians, their art, and their technology in the esteem in which they were held through the 17th century. Even in the best of days, many ideas concerning music perception and musical imagery have been doctrinaire with little relevant foundation. Alas, musicians of the serial and post serial period appear to have put arbitrary ideas about music and perception ahead of actual musical experience. This has been bad for psychophysics and psychology, and perhaps also for music.

Happily, a new generation of musicians, skilled in the digital production and analysis of sound, have made great advances in the evocation of musical images of sound source, location, and motion. This has led to real advances in our understanding of music perception. Many or most of these advances have been motivated by a search for musically useful sounds, sound relations, sound sequences, soundscapes, if you will, rather than from an interest in psychoacoustic phenomena as such. However, compelling demonstrations have shown the scope and reality of such advances.

There is every reason to believe that new musical technology will be used to produce new and compelling effects and insights. We hope that compelling demonstrations and illustrations may be as fruitful in the psychophysics and psychology of musical sound as illustrations have long been in the field of visual perception.

REFERENCES

Aristoxenus. (1974). *The harmonics of Aristoxenus* (H. S. Macran, Ed. and Trans.). New York: Verlag.

Bekesy, G. von. (1980). *Experiments in hearing*. New York: Robert E. Krieger Publishing Company. (Original work published in 1960)

Bregman, A. S. (1990). *Auditory scene analysis*. Cambridge, MA: MIT Press.

Broadbent, D. E. (1958). *Perception and communication*. New York: Pergamon.

Chowning, J. M. (1977). The simulation of moving sound sources. *Computer Music Journal, 1*, 48–52. (Reprinted from *Journal of the Audio Engineering Society*, 1971, *19*, 2–6.

Chowning, J. M. (1977). The synthesis of complex audio spectra by means of frequency modulation. *Computer Music Journal, 1*, 46–54. (Reprinted from *Journal of the Audio Engineering Society*, 1973, *21*, 526–534.

Chowning, J. M. (1980). Computer synthesis of the singing voice. In J. Sundberg (Ed.), *Sound generation in winds, strings, and computers* (pp. 4–13). Stockholm: Royal Swedish Academy of Music.

Chowning, J. M. (1989). Frequency modulation synthesis of the singing voice. In M. V. Mathews & J. R. Pierce (Eds.), *Current directions in computer music* (pp. 57–63). Cambridge, MA: MIT Press.

Cohen, H. F. (1984). *Quantifying music*. Dordrecht: Reidel.

Cooper, F. S., Delattre, P. C., Liberman, A. M., Borst, J. M., & Gerstman, L. J. (1952). Some experiments on the perception of synthetic speech sounds. *Journal of the Acoustical Society of America, 24*, 597–606.

Deutsch, D. (1972). Octave generalization and tune recognition. *Perception and Psychophysics, 11*, 411–412.

Deutsch, D. (1974). An auditory illusion. *Nature, 1251*, 307–309.

Deutsch, D. (1975a). Auditory memory. *Canadian Journal of Psychology, 29*, 87–105.

Deutsch, D. (1975b). Two-channel listening to musical scales. *Journal of the Acoustical Society of America, 57*, 1156–1160.

Deutsch, D. (Ed.). (1982). *The psychology of music*. New York: Academic Press.

Deutsch, D., Kuyper, W. L., & Fisher, Y. (1987). The tritone paradox: Its presence and form of distribution in a general population. *Music Perception, 5*, 79–92.

Dorian, F. (1947). *The musical workshop*. London: Secker & Warburg.

Edelman, G. J., Gall, E. W., & Cowan, W. M. (Eds.). (1988). *Auditory function: Neurobiological bases of hearing*. New York: Wiley.

Fletcher, H. (1924). The physical criterion for determining the pitch of a tone. *Physical Review, 23*, 427–437.

Helmholtz, H. von. (1954). *On the sensations of tone*. (A. J. Ellis, Trans.). New York: Dover. (Original work published in 1885)

Houtsma, A. J. M., Rossing, T. D., & Wagenaars, W. M. (1987). *Auditory demonstrations* [Compact Disc, Philips 1126-061].

Knudsen, E. I. (1988). Experience shapes sound localization and auditory unit properties in the barn owl. In G. M. Edelman, W. E. Gall, & W. M. Cowan (Eds.), *Auditory Function* (pp. 721–745). New York: Wiley.

Konishi, M., Takahashi, T. T., Sullivan, W. E., Wagner, & Carr, C. (1988). Neurophysiological and anatomical substrates of sound localization in the owl. In G. M. Edelman, W. E. Gall, & W. M. Cowan (Eds.), *Auditory function* (pp. 721–745). New York: Wiley.

Licklider, J. C. R. (1951). A duplex theory of pitch perception. *Experientia, 7*, 128–133.

Mach, E. (1943). On symmetry. In *Popular Scientific Lectures*. T. J. McCormack (Trans.). La Salle: The Open Court Publishing Company. (Originally published in 1898)

Mathews, M. V. (1969). *The technology of computer music*. Cambridge, MA: MIT Press.

Mathews, M. V., Pierce, J. R., Reeves, A., & Roberts, L. (1988). Theoretical and experimental explorations of the Bohlen-Pierce scale. *Journal of the Acoustical Society of America, 84*, 1214–1222.

Miller, L., & Cohen, A. (1987). *Music in the Royal Society of London;* 1660–1806. Detroit: Detroit studies in music bibligraphy, No. 56.

Moog, R. A. (1977). Amplifier with multiplier. U. W. Patent 4, 117,413, filed June 21, 1977.

Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.

Pierce, J. R. (1990). Rate, place, and pitch with tonebursts. *Music Perception, 7*, 205–212.

Plomp, R. (1966). *Experiments in Tone Perception*. Eindhoven, The Netherlands: Institute for Perception.

Pollack, I. (1972). Memory for auditory waveform. *Journal of the Acoustical Society of America, 51*, 1209–1215.

Risset, J.-C., & Wessel, D. L. (1982). Exploration of timbre by analysis and synthesis. In D. Deutsch (Ed.), *The psychology of music*. New York: Academic Press.

Rhode, W. S. (1970). *Measurement of the amplitude and phase of vibration of the basilar membrane using the Moessbauer Effect*. Unpublished dissertation. University of Wisconsin, Madison.

Rhode, W. S. (1971). Observations of the vibration of the basilar membrane in the squirrel monkey using the Moessbauer technique. *Journal of the Acoustical Society of America, 49*, 1218–1231.

Rodet, X., Potard, Y., & Barriere, J.-B. (1984). The CHANT project: From the synthesis of the singing voice to synthesis in general. *Computer Music Journal, 8*, 15–31.

Rossing, T. D. (1990). *The Science of Sound*, 2nd ed. New York: Addison Wesley.

Schoenberg, A. (1975). Composition with twelve tones (1). in *Style and Idea: Selected writings of Arnold Schoenberg* (pp. 214–244). (Edited by L. Stein, with translations by L. Black.) New York: St. Martin's Press.

Schreiner, C. E., & Langer, G. (1988). Coding of temporal patterns in the central auditory nervous system. In G. M. Edelman, W. E. Gall, & W. M. Cowan (Eds.), *Auditory function*. New York: Wiley.

Shepard, R. N. (1964). Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America, 36,* 2345–2353.

Sundberg, J. (1987). *The science of the singing voice*. DeKalb: Northern Illinois University Press.

Terhardt, E. (1974). Pitch, consonance, and harmony. *Journal of the Acoustical Society of America, 55,* 1061–1069.

Treisman, M., & Rostron, A. B. (1972). Brief auditory storage: a modification of Sperling's paradigm applied to audition. *Acta Psychologica, 36,* 161–170.