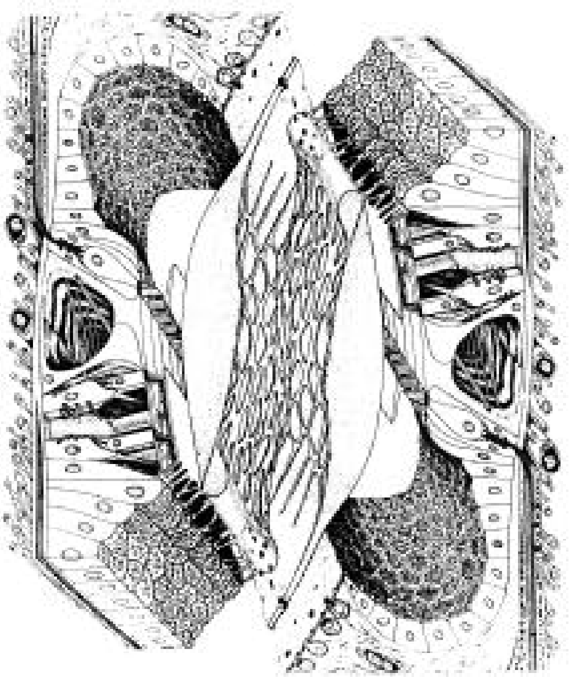

Why do octaves sound the same?

xaq pitkow
spring 2000

Preliminary Qualifying Exam
for Harvard Biophysics

Why do octaves sound the same?

Abstract

The octave, a relation between two tones whose fundamental frequencies stand in the ratio 2:1, is a foundation of tonal musics worldwide: notes separated by an octave are considered harmonically equivalent, and melodies are often sung or played in parallel octaves yet are considered identical melodies. Why do octaves sound the same? I hypothesize that our perceived similarity of octave-related tones derives from other properties of the octave interval — namely spectral fusion, sensitivity to interval tuning, and generalization of response to the common fundamental — which are qualitatively similar to properties of larger intervals in the harmonic series. Consequently, double-octaves (ratio 4:1) should have less similarity (in the above sense) than twelfths (ratio 3:1) have; the octave equivalence present in music is the result of a learned transitivity applied to the natural perceptual similarity. I further hypothesize a mechanism for these properties in which neural circuits in the brainstem inferior colliculus detect coincident firing of neurons tuned to harmonics of a fundamental to compute the periodicity pitch; occasional skipped firings lead to excitation of subharmonically tuned neurons, causing a note to sound like its subharmonic octave. Psychoacoustics experiments to measure musical and nonmusical humans' tuning sensitivity, generalization, and spectral fusion of twelfths and double-octaves should serve to evaluate the first hypothesis. Animals trained to respond to one note generalize this response to other notes (Blackwell and Schlosberg 1943); raising these animals in artificial acoustic environments may alter the patterns of generalization and thus provide a test for learning of octaves. Direct extracellular recording from the brainstem of animals could provide evidence for the proposed neural mechanism of harmonic equivalence.

Contents

Introduction

Harmonic structure of sound

Biologically relevant sounds are often produced by the periodic vibration of a membrane and air column such as our vocal cord and tract (or that of our predator, pet, or prey). This method of creating noise produces vibrations not only with the fundamental frequency of the membrane, but also with higher frequencies or harmonics. The relative strength and phase of these harmonics will tend to covary as the sound source moves, and thus the auditory system can use this information to aid in localization or attention. A single representation of these correlated sensory responses would thus be a useful representation to have, and we have one: when presented simultaneously, harmonically related frequencies fuse into a single percept, *pitch*. Pitch is the subjective analog of physical frequency: the percept ranges from low to high and has a nearly linear, one-to-one correspondence to frequency of pure tones. Pitches are not best described on a linear scale, however: the neurological units of the hearing system are spaced logarithmically. The harmonic series is a set of frequencies with arithmetically constant frequency separation between them, yet we hear the series as notes that get closer and closer together. The natural comparison between two pitches is the ratio between their respective frequencies. These ratios define the *interval* between two pitches. A list of important interval names is found in Table 1.

interval name	ratio	sample notes
unison	1:1	C4–C4
octave	2:1	C3–C4
twelfth / octave+fifth	3:1	C3–G4
double-octave	4:1	C3–C5
two octaves+third	5:1	C3–E5
perfect fifth	3:2	C3–G3
major third	5:4	C4–E4

Table 1: some useful interval names

All non-pitch audible spectral attributes of a sound, including the harmonic spectrum, relative phase relationships between component frequencies, and slow modulations below audible frequencies, are typically subsumed under the term *timbre.* Voices and instruments may produce sounds which share the same pitch, but they are distinguishable based on timbre. A flute has a spectrum with few prominent harmonics; a violin has many; a piano’s are slightly mistuned. These spectral qualities give each instrument its unique sound type. They can all, however, produce the same pitch.

Temporal structures built of pitch intervals *are* melodic and harmonic musics.

Western tonal music tradition emphasizes a particular set of intervals called consonances, meaning, literally, ‘together-sounding.’ Not all musical cultures value this same set of intervals, and untrained observers show no evidence for ‘natural’ categories for musical intervals (Burns 1977). One consonance, however, is absolutely universal across all music cultures. This is the octave.

the Octave: definition and psychophysics

The name, ‘octave,’ has something of a bias, based on the Western system which places this particular tonal relationship eighth in an increasing series of intervals. In a more culture-independent definition, an octave is the relationship between two pitches whose frequencies are related by a ratio of 2:1. Physically, it is the interval between a fundamental and its first harmonic. Vocally, when men and women sing together they typically mirror the melodic pitches one octave apart, and somehow this sounds like the same note. In western music theory, the principle of octave equivalence asserts that notes related by one or more octaves function the same in a harmonic or melodic structure. When listing pitches in increasing order, written systems use a circular scale in which notes separated by octaves are given the same name: DO-re-mi-fa-so-la-ti-DO in the West, SA-re-ga-ma-pa-dha-ni-SA in south Asia. Octaves just seem to sound the same.’[1] It is a great mystery just why this interval is special. Does the cause lie in mathematics? In physics? In neuroscience? Is the octave just an arbitrary, a culturally impressed pitch relation?

Mathematicians and physicists attempted to find a resolution. Circa 500 BCE, Pythagorus explained that simple ratios of the lengths of vibrating strings gave consonance. Much later, von Helmholtz correctly hypothesized from physical principles that the ear takes a Fourier transform of the pressure wave (1885). When two waves are close in frequency, an audible beat pattern of amplitude modulation results. Harmonics of two tones can overlap and produce beat patterns as well. Helmholtz proposed that consonance was actually the minimization of beat patterns, which resulted when the two tones’ frequencies stand in a simple integer ratio. Lipps modified this argument, saying that it applied not to physical vibrations but to mental excitations (1905).

Critically, this type of argument always fails to explain the equivalence of octaves. An octave interval is merely the very first interval which allows greatest alignment of harmonics, 2 being the smallest integer greater than 1. Yet a frequency tripling and a quadrupling also allow alignment of harmonics. If the harmonic-alignment hypothesis is in fact true, then the tripling would actually sound more similar than a quadrupling. A

---

[1] When I refer to an interval sounding similar, I mean that the constituent notes of the interval sound similar.

quadrupling of frequency is two octaves above the fundamental, whereas a tripling is a non-octave interval, also consonant, called a twelfth, or an octave+fifth. If octaves are important for reasons that extend beyond just overlap of harmonics, then the double-octave should sound more similar than the twelfth.

Can one test such a hypothesis by simply asking subjects to compare the similarity of each harmonic to its fundamental? Such a procedure is quickly revealed to be too naïve. Kallman (1982) attempted just such a study, and found that similarity decreased monotonically with interval size: octaves were no more similar to the target than any other nearby interval. One of his musical subject did rate octaves as more similar than neighboring intervals, but this is uninterpretable: for this subject, do octaves sound the same because musical training taught they did, or was there a genuine perceptual similarity? Allen (1967) found precisely the same dichotomy between musically trained and untrained subjects. To investigate the octave, one must use measures of similarity that do not rely on subjective judgments.

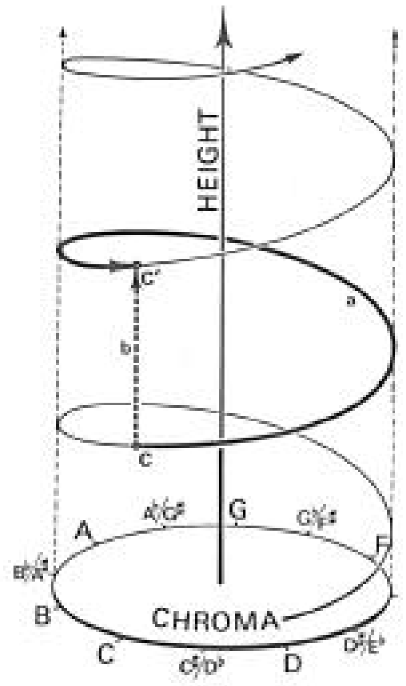

Figure 1: Pitch has been represented as a curve in two dimensions called pitch height and pitch chroma. Height ranges from low, like a thundering rumble, to high, like a shrill birdcall. Chroma represents the musical function of a pitch; in essence, chroma is pitch that has been normalized to a single octave, by assunming all-octave equivalence (from Shepard 1964).

Using better psychophysical methods, several studies have demonstrated some perceptual similarity between two octave-related tones that are not present in tones whose interval differs from an octave by a semitone or more. People with absolute pitch sometimes place notes in the wrong octave with the right note name (Bachem 1954). Pure-tone pairs are sometimes confused with their *inversions*, counterparts in which one note is transposed in the direction of the other by an octave (Plomp 1973). Deutsch (1973) showed that when subjects are asked to compare two tones separated by several intervening pitches, the intervening pitches interfere strongly with memory of the target pitches if they match the target, even if it is an octave above or, more weakly, an octave below the target, but not otherwise. Shepard (1964) demonstrated that a harmonic complex containing all frequencies standing in a ratio of 2^n to a given reference frequency — all octaves of a given pitch — fuse into a single organ-like pitch with ambiguous tone height; and that pairs of these tones are perceived non-transitively. That is, one can choose tones A, B, and C such that when ordering the pitches, subjects will place A>B, B>C, but A<C, indicating that a monotonic scale is inadequate for

representing pitch. His Shepard tone experiments support instead a helical model for pitch, shown in Figure 1, in which notes separated by an octave are equivalent in the 'chroma' dimension. Fusion of multiple frequencies into a single pitch is a property of the auditory system more generally, discussed below; fusion of octave-separated harmonics presented to separate ears, and lack of such fusion for other nearby intervals (octave ± 1 or 2 semitones), was considered evidence for internal octave templates by Demany et al (1987). Demany (1984) presented evidence that 3-month-old infants are less surprised by tonal sequences in which one note is replaced by its octave than those in which that note is replaced by a seventh or a ninth. Blackwell and Schlosberg (1943) showed that rats trained to respond to one frequency have stronger generalization to an octave subharmonic than to other lower inharmonic frequencies. Skin galvanometry measurements by Humphreys (1939) revealed that after mild shock conditioning against one frequency humans have subconscious generalization to octaves, shown by greater skin conductance response to an octave than to a slightly smaller interval. It might seem fair to conclude that, based on perceptual studies of animals, conscious and unconscious human faculties, absolute and relative pitch, and sequential and simultaneous presentation of pitches, there really is something special about the octave.

It is of course important to stress that octaves are not truly 'the same,' but rather have some ill-defined similarity which may depend on context. Similarity judgments are influenced primarily by pitch height in musically naïve subjects (Kallman 1982). Deutsch showed that transposing some notes in a melody by an octave prevents tune recognition, though once the tune is named subjects can still affirm the tune's identity (1982). Risset generated Shepard-tone-like sounds, with stretched octaves in lieu of pure octaves, that when transposed up by an octave actually sound like they decrease in pitch (1986). Both Deutsch (1982) and Risset (1986) thus conclude that for certain tasks, proximity of tones or, alternately, melodic contour determine perception of pitches more strongly than do octave transpositions. No other animals other than Blackwell's rats have yet been shown to have generalization responses to octaves, despite at least one attempt with songbirds (Cynx).

Octaves are not even subjectively fixed intervals. When any given human subject is asked to tune an oscillator to an octave above a given tone, the frequency ratio upon which they decide is usually greater than 2:1. The precise amount of stretch is consistent within one session of trials, but varies significantly from day to day (Ward 1964). Even for trained musicians, stretched octaves and scales are preferred in melodic lines, at least when such intonation is possible as on fretless instruments such as violins and trombones (Terhardt 1975). The amount of stretch is dependent on both the reference frequency and intensity (Sundberg 1972). The subjective octave was actually shown to be additive: tuning to an octave above a reference, and then tuning an octave above that, yields the same frequency as tuning two octaves above the first reference (Ward 1964). This is in fact the only report in the literature that addresses multiple-octave perception. Most psychoacoustic literature explores relationships among pitches that vary within one octave, or perhaps a tone higher or lower; scientists seem to not have been interested in these larger multi-octave intervals. Yet our perception of these tones is extremely relevant to discovering the nature of the octave.

The first harmonic of a given fundamental is an octave above it. Most natural sounds have a strong first harmonic; in the bat this harmonic is far stronger than the fundamental (Suga 1976). Is the harmonic structure of sound somehow responsible for octave similarity?

This general question leads to a much more specific one: do higher harmonics evoke the same kind of similarity percept, or are octaves really qualitatively different? The harmonic series consists of the following intervals from the fundamental: unison, octave, twelfth, double-octave, double-octave+major third... If octaves are special, the double-octave should be more similar, or be more mistakable, or have a lower discrimination threshold, than the twelfth. If not, any aforemention psychophysical measure should show a monotonic progression in harmonic number.

None of the existing studies on the nature of the octave actually address its uniqueness. Every conclusion reached supposedly about the octave is not actually a conclusion about the octave but about the first harmonic. Some of the studies investigate the importance of smaller intervals such as the fifth, which is related to the second harmonic (an octave below a twelfth) but the two may give different results; and the double-octave has not been thoroughly investigated. Thus, no one knows whether infants are less surprised by twelfths than by elevenths or thirteenths, or whether rats will generalize to twelfths and double-octaves. Shepard-like tones composed of stacked non-octave intervals do still give rise to circular illusions; no one knows whether stacked twelfths would fuse and give rise to ambiguous pitch ordering. Would skin galvanometry reveal a subconscious conditioning generalization to any harmonic, or just octaves?

It is thus useful to notate a distinction between two types of octave equivalence. *All-octave equivalence* is the perception that all octave-related notes, e.g. double-octaves, triple-octaves, sound the same. *Nearest-octave equivalence* is the perception that single octaves sound more similar than multiple-octaves do, and that non-octave harmonics sound more similar than even slightly higher octave harmonics. My first hypothesis is, with this terminology, that only nearest-octave equivalence is present in the musically untrained; training or musical exposure creates all-octave equivalence by transitively associating octaves.[2]

It also is interesting to note that the fundamental frequencies of most musical notes and both male and female speech lie below our auditory fovea, our most frequency-sensitive region; however, the low-order harmonics of music and speech lie directly within it (Braun 1999). Furthermore, these low-order harmonics have been found to be most influential in determining the fundamental of a stimulus (Ritsma 1963). Indeed, the spatial resolution of the tonotopic map of the cochlea is inadequate to explain our ability to discriminate pure frequencies below 1 kHz (Langner 1997). Thus, for speech and music, the brain computes pitch by using supplemental information from temporal structure and/or higher-order frequency components.

If pitch is perceived through analysis of harmonic or temporal structure, and octaves might be caused by harmonic structure, then it makes sense to see what studies have been done on the perception of harmonics. Most work on the subject applies to the perception of harmonic complexes — simultaneous tones composed of frequencies all of which are harmonics of a single fundamental.

Pitch perception

Psychophysics shows us just how good we are at resolving harmonic spectra into a single pitch: if we artificially stimulate with a few upper harmonics of some base frequency, or mask the fundamental with noise (Licklider 1954), our brains recreate a missing fundamental which we perceive as a pitch. Importantly, note that there is no resolved spectral energy at the frequency corresponding to this pitch.

Over time, various experiments discovered that more and more limited spectra could give rise to a perception of the missing fundamental. First, a complete harmonic complex without the fundamental (Schouten 1962) then a set of five (de Boer 1976), then three (Ritsma 1962), then two harmonics (Houtsma 1971) were all found to create a low pitch corresponding to the missing fundamental. Amazingly, even a single harmonic was shown to elicit a pitch at a subharmonic, though only if noise were present and if attention were directed to the expected pitch region (Houtgast 1976).

---

[2] Transitivity of octaves means if A is an octave of B and B is an octave of C then A is an octave of C.

There is no direct cochlear response to the low pitch perceived from harmonic stimuli missing the fundamental: periodicity pitch does not arise from nonlinear distortion of the basilar membrane (Galbraith 321), yet can be heard even in the presence of masking noise at that frequency (Smith 1978) and is perceived even when only two harmonics are presented, one to each ear (Houtsma 1972). Low pitch perception emerges only when information from different auditory nerve fibers is brought together centrally.

We are not only able to detect a missing fundamental whose spectral components are resolved as separate peaks in the peripheral auditory system, but we can also detect a missing fundamental in higher harmonics which are unresolved (Houtsma 1990). While a stimulus with a missing fundamental has by construction no spectral energy at that frequency (say f_0), nonetheless it has the period $1/f_0$ (Figure 2). Here, only temporal information can be available, because these frequencies are not isolated into separate channels by the inner ear. Auditory nerve fibers are known to phase-lock with acoustic waveforms, firing stochastically at peaks of vibration. Cariani and Delgutte (1996) demonstrated that the perceived pitch for both resolved and unresolved harmonic stimuli was predicted by pooling the temporal information of all auditory nerve firings and identifying the most common interspike interval one could predict; one could even explain the origin of the stretched perceptual octave as due to the refractory period and of neural firings (McKinney). However, certain pitch-evoking click series caused serious problems for this model. The model predicted our poor perception of pitch from nonsequentially-correlated clicks when only sequential interspike intervals were used in the computation (Kärnbach 1998), whereas it predicted intensity invariance of pitch correctly only when higher-order sequences were considered (Cariani 1996). Thus, the simple temporal or autocorrelation model is insufficient to explain certain pitch percepts. However, Kärnbach's click trains were admitted to have only a weak pitch sensation that could not even be used to play a recognizable melody; so Cariani and Delgutte's model remains quite successful for normal pitch stimuli.

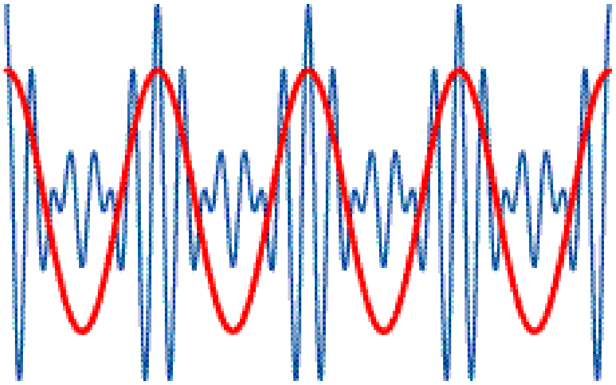

Figure 2: a wave composed only of harmonics 5–7 (blue) has the same periodicity as their fundamental (red) but no spectral energy at the fundamental frequency.

Nonetheless, stimuli with resolved and unresolved harmonics seem to be perceived by different mechanisms (Steinschneider 1998). Shackleton and Carlyon (1994) found comparison amongst tones of unresolved harmonics, or comparison amongst tones of resolved harmonics, was easier than comparison between resolved and unresolved tones.

White and Plack (1998) showed that the dependence of pitch identification on stimulus duration was different for these two categories of tones: pitch determination from unresolved harmonics improved with continuous stimulus duration up to 80 ms, whereas that from resolved harmonics was established within about 20 ms. One implication is that resolved harmonics generate useful information in spatially separated neural channels, whereas unresolved harmonics yield information only through temporal structure of pooled responses which must be interpreted for a longer time.

The conclusion is that both place- and time-coding are both at work in pitch perception, to differing degrees based on the resolvability of harmonic structure. For a wide range of spectral types, our brain tries to compute the single subharmonic pitch whose harmonics best match the perceived acoustic spectrum.

A possible neural mechanism of pitch perception

This is interesting in its own right, but if we can understand neurophysiologically how we perceive a missing fundamental, we may be stumbling onto the very same mechanism that makes us perceive octaves as the same. After all, the spectrum of a note is just like the spectrum of a note an octave lower which is missing its fundamental and even harmonics. If a circuit exists which extracts a missing fundamental from spectral information, then that same mechanism might be partly excited by its octave. Recent psychophysical, encephalographic data, brainstem single-unit recordings, and anatomical evidence suggests a specific neural circuit in the inferior colliculus may generate the missing fundamental (Braun 1999).

In psychoacoustic studies, the percept of two pure tones depends on their frequency separation. For a difference of <20 Hz, we perceive beat patterns due to interference. As the frequencies separate by more than 20 Hz, we cease to perceive individual fluctuations but hear instead a sensation of roughness. As the separation increases still further, beyond a threshold separation that varies approximately logarithmically with center frequency, the two tones resolve into individual components. This demonstrates the existence of a sensory feature called the critical bandwidth (CB), representing the simultaneous-tone resolving capacity of our auditory system. Frequencies within the critical band, typically one quarter-octave in width, fuse into a single percept, while frequencies separated by more than one critical band are resolved (Rasch 1982).

This poorly resolved frequency response is not present in cochlear vibrations or in auditory nerve responses (Pickles 1979), but is found in noise-masking curves of single-neuron recordings from the inferior colliculus (Ehret 1985). Neurons in the inferior colliculus have are excited maximally by a particular frequency, and detailed three-dimensional mappings of these characteristic frequencies reveal distinct layers of closely tuned neurons separated by discrete jumps in tuning (Schreiner 1997).

Critical bands are hypothesized to arise from this structural arrangement (Schreiner 1997); anatomical evidence supports this notion. Inhibitory effects of stimuli within a critical bandwidth were found to exist at a cellular level and were localized to individual lamina, measured by activity-induced labeling of tissue in the inferior colliculus (Webster 1985) and extracellular recording upon suppression of GABA inhibition (Palombi 1996). This jives with morphological studies which show extensive, oriented dendritic connections within a lamina and between both neighboring and second-nearest laminae (Oliver 1991).

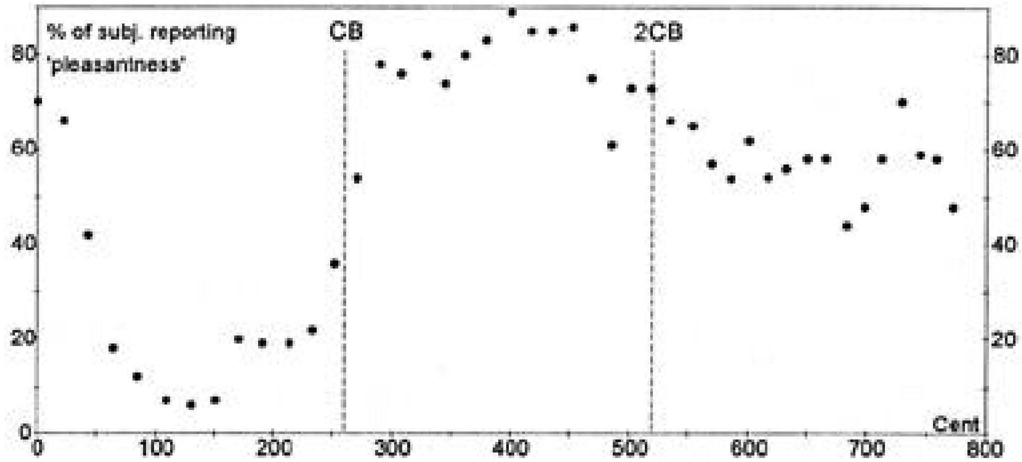

Figure 3: Compiled data of psychophysical percepts as a function of interval size (in cents, which is a logarithmic measure of frequency ratio r: cents = 1200 log$_2$*r). Note here the vertical lines demarcating pitches' trends toward unpleasant for intervals less than one critical band (CB = 260 cents), neutral for those larger than two critical bands (2CB), and pleasant for those between one and two critical bands (from Braun 1999).*

Braun (1999) emphasizes the finding of some specific next-neighbor axonal connections in Malmierca (1995), and claims this as a possible anatomical substrate of a previously unnoticed double-critical bandwidth (2CB) which he observes in pooled psychophysical data. The double-critical bandwidth is, simply, twice the bandwidth of the critical band. Whereas tone pairs within the CB sound dissonant, tones whose frequencies stand in a relation between the CB and 2CB appear consonant; and tones pairs whose frequencies are separated by more than 2CB elicit neutral reports (Figure 3). A similar weak trend may exist in the number of spontaneous otoacoustic emission (SOAE) pairs as a function of frequency ratio (Braun 1997). Otoacoustic emissions are physical vibrations of the basilar membrane, induced by contractions of the Outer Hair Cells which are thought to actively regular the nonlinear properties of the basilar membrane based on descending input from the brainstem (Kim 1986, Long 1997).

The connection between the single- and double-critical bandwidth and the computation of a missing fundamental arises when one notes the frequency separation of the psychophysically most important harmonics, partials 3-5, in perceiving the fundamental. The difference between each subsequent frequency is of course constant when measured in frequency; but when measured in terms of frequency ratio, the spacing between harmonics grows smaller with increasing harmonic number. The critical bandwidth is approximately proportional to center frequency, so as harmonic number grows, neighboring harmonics eventually fall within a single critical bandwidth and cease to be resolved (Figure 4). The fundamental and second harmonic fall outside of the 2CB, and thus may be considered less necessary to pitch extraction by the brain.

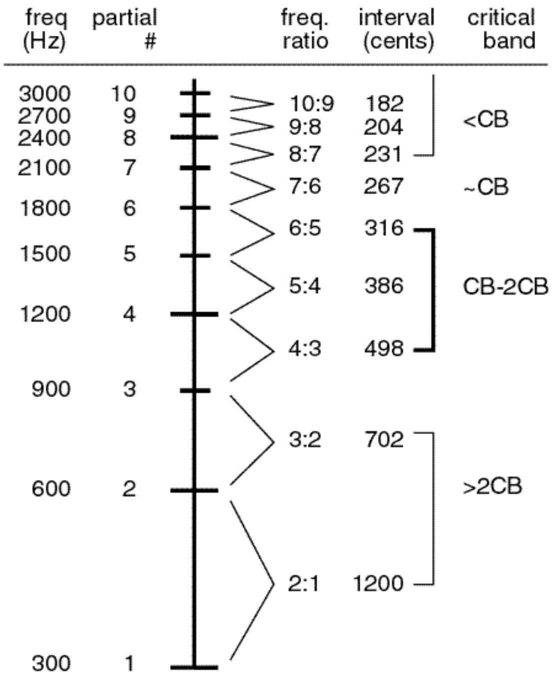

Figure 4: Harmonics 3–5 fall between one and two critical bands (from Braun 1999).

For harmonics 3-5, whose ratios fall between CB and 2CB, the consonance observed in pairwise psychophysical experiments may reflect a physiological coalescence of the information from their respective laminae. Thus, the circuitry of the inferior colliculus may be designed to extract the fundamental from the dominant harmonics.

The temporal model of pitch perception may also be relevant in this computation. Neurons are synchronized to the acoustic waveform: they tend to fire with highest probability at a particular phase in the stimulus period. If such phaselocked neurons corresponding to two consecutive partials converge, the greatest coincidence of spikes will occur with the period of the fundamental. A coincidence detector may have already been found in the whole-cell patch-clamp recordings of Covey et. al. (1996). They measured post-stimulus currents in several neurons of the bat inferior colliculus, and found subthreshold oscillations over a wide range of frequencies. Inputs whose frequencies are commensurate with the intrinsic oscillations could excite spiking output. Schwarz et al. (1993) also found oscillating cells in the rostromedial area of the inferior colliculus, which had preferred frequencies that varied from cell to cell to cover four octaves despite the lack of any phaselocked input due to a surgical removal of the entire inner ear.

Such neurons might serve as efficient coincidence detectors of harmonically related responses, and it may thus be ideally suited to compute and transmit the fundamental frequency of a harmonically complex stimulus (Langner 1997, Braun 1999). Pure tones could not have their frequencies calculated by this mechanism, but frequency-sensitive neurons that respond directly to the fundamental could be used in tandem with the periodicity information (Langner 1997).

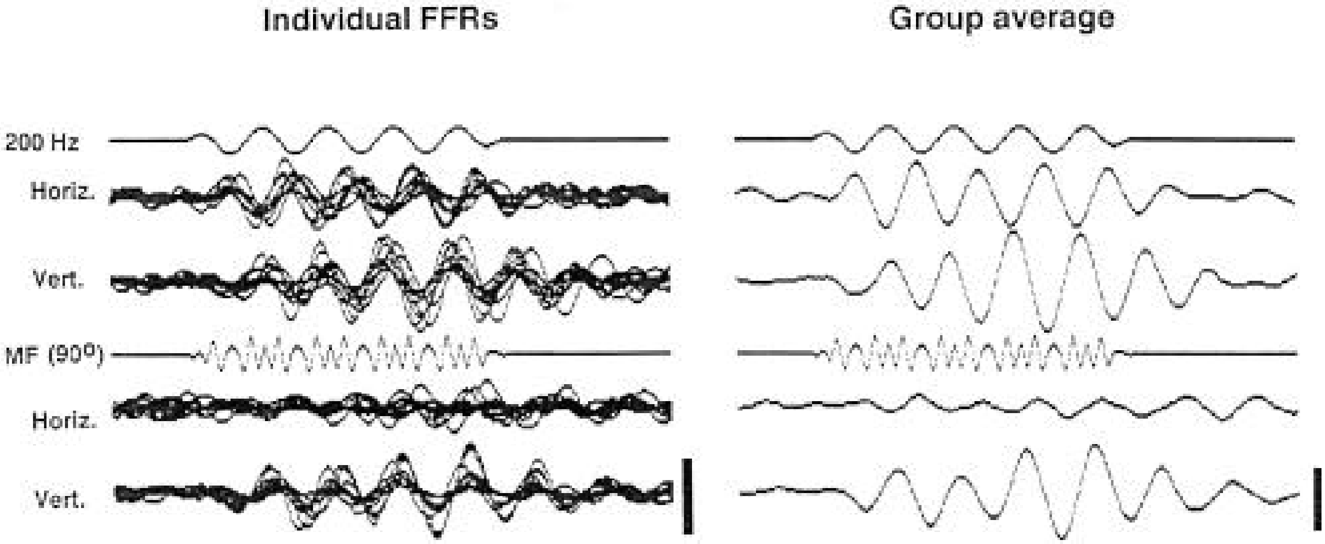

Figure 5:Electroencephalogram (EEG) of Frequency Following Response (FFR) to pure tone and missing fundamental stimuli. From top to bottom, traces are: pure-tone stimulus waveform, horizontal and vetical current dipole responses, a missing fundamental (MF) stimulus of three harmonics with one component 90° out of phase with the others (for minimum amplitude modulation), and horizontal and vertical current dipole responses. EEGs from horizontal current dipole correspond to auditory nerve fibers; those from vertical dipoles correspond to brainstem activity. Vertical bars represent 0.5 µV for individual and 0.25 µV for group FFRs; horizonal dimension shows 50 msec of time. Note that the vertical current dipole response to the missing fundamental is larger than the horizontal response (from Galbraith 1994).

Noninvasive studies of human brains have discerned a signal strongly correlated to acoustic stimuli. Electrodes that measure current dipoles in the head have detected current oscillations with the same waveform as pure tone stimuli, the so-called Frequency Following Response (FFR), associated with the synchronized firing of action potentials. Galbraith (1994) discovered that FFRs would arise even when the stimulus lacked the fundamental frequency; but importantly, this evoked response was measured only for the vertical current dipole associated with brainstem activity, and not for the horizontal dipole associated with the auditory nerve fibers (Figure 5). This is evidence that the missing fundamental is indeed computed in the brainstem, not the auditory cortex, and is not present in the peripheral auditory sense.

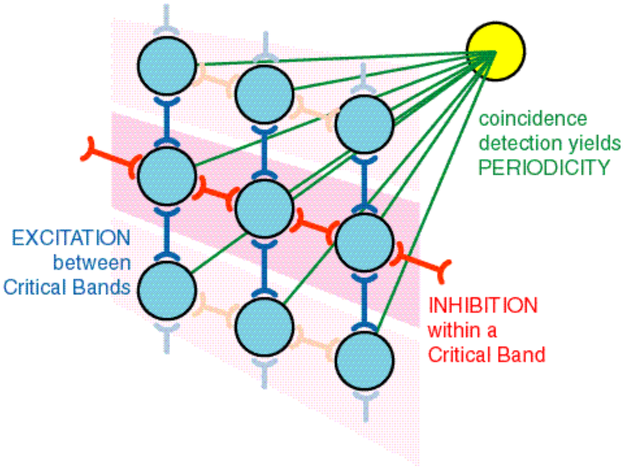

Figure 6: Neural networks for calculating the missing fundamental or periodicity of a stimulus. Frequency-tuned neurons(cyan circles) in the inferior colliculus are arranged in bands of similar tunings (light magenta). Connections are inhibitory (red) within a layer, and excitatory (blue) between neighboring layers. A coincidence detector (yellow) receives input (green) from nearby neurons and then fires with periocity of the stimulus waveform.

Hypothesis

Coincidence detection of stochastically doubled periodicity in inferior colliculus fundamental extraction excites neurons sensitive to a subharmonic octave.

If different frequency channels converge in the inferior colliculus and temporal coincidence detection there calculates the periodicity, then the stochastic nature of neural firings could affect the calculation. If several neurons all ‘miss a beat’ simultaneously, then the implied periodicity would actually be double its true value, and its pitch would seem an octave lower. If such a mistake were to be given as input to a neuron which is most sensitive to that lower frequency, then it would respond with a low level of excitation. A given pitch would thereby also excite a subharmonic an octave below; it would truly resemble another pitch one octave lower.

If extraction of the fundamental is realized as hypothesized by Braun, there would be both an innate and a learned component. The rough laminar structure of the inferior colliculus is developmentally programmed but fine-tuning of the auditory system in general requires appropriate auditory input (Friauf 1993). This suggests that appropriate teaching soon after birth might allow subtle but measurable reconfiguration of the spectral pattern recognition developed.

Is it necessary to suppose this fundamental extraction circuit to make this leap? After all, the first harmonic of a natural fundamental is an octave, and this is much more strongly correlated with the fundamental than any randomly occurring subharmonic octave mistakes. Terhardt’s influential theory of virtual pitch (1973) is relevant here. He hypothesized that through pattern recognition the brain learns to associate harmonically related frequencies as a single percept, the virtual pitch. Other harmonic complexes could excite subpatterns, and this would give percepts of consonance. For example, the spectral composition of a harmonic series missing its fundamental would closely match the excitation pattern of the complete spectrum, and thus elicit the pitch of that missing fundamental. However, when the first upper harmonic is present, so is the second —

which is not an octave — and a whole slew of other non-octave intervals. There is extra information in the sounded tone which unambiguously distinguishes it from an octave above. In contrast, there is an absence of information which might distinguish it from a lower *sub*harmonic note: as far as the brain is concerned, this information might or might not exist. The brain will use information it has, but cannot use information it does not have.

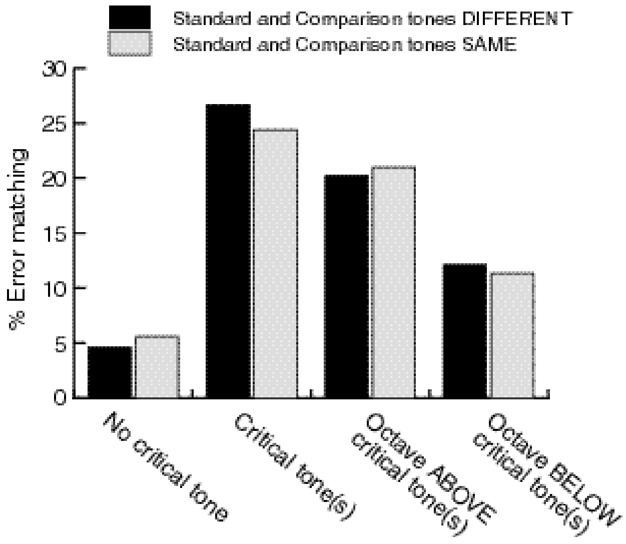

Figure 7: Octave generalization demonstrated by interference with pitch memory. Two tones are presented, subjects are asked whether they are the same or different, and the fraction of errors is recorded. Other 'interfering' tones are presented between the standard (target) and comparison tones, lowering the fraction of correct responses. When an interfering tone matches the target tone, subjects make more mistakes in pitch memory; the same holds for the octaves of the target tone, but more so for an octave above than for an octave below (from Deutsch 1973).

This argument suggests there should exist a measurable asymmetry in octave perception: a note should sound more like its subharmonics than like its harmonics. When directly comparing the 'similarity' of two sounds this can't be measured: does the first sound more like the second or does the second sound more like the first? But for generalization experiments this asymmetry can become apparent. Actually, a close examination of existing data shows just such an effect. Deutsch demonstrated (1972) that if a note interferes with memory of a pitch, its octave above will also interfere almost as much, whereas the octave below will interfere but less than half as much; her data is reproduced in Figure 7. The interpretation is that a note an octave above the interfering tone sounds like the interfering tone itself — which is the real note's subharmonic octave — and thus serves to interfere; a note an octave below the interfering tone sounds less like the interfering tone — it is distinguished by having more harmonics present — and interferes less. Furthermore, a single frequency can elicit a subharmonic percept if noise is present (Houtgast 1976), which the current hypothesis explains as noise tending to increase the random neural firing background and increasing the probability of superthreshold coincidence detection of subharmonic periodicities. Cariani and Delgutte's algorithm (1996) to determine pitch from interspike intervals also shows a very interesting effect when applied to Shepard tones: a mistake yielding the first subharmonic of such a harmonic complex would predict a pitch an octave below; one yielding the second subharmonic would predict a pitch an octave plus a fifth below, and since all octaves are present, this would be manifest as a pitch a fifth below one of the physical frequencies present. Sometimes, their algorithm does choose a pitch that is a fifth below the pitch chroma on which the tone is based, shown in Figure 8. In this case the most common interspike interval was actually the second subharmonic, indicating that random stochastic firings might produce an occasional period-doubling or even tripling in a population of neurons.

In summary, I hypothesize that octave similarity arises because periodicity detection in the inferior colliculus can mistakenly excite subharmonic responses, and that the octave, while naturally the most similar interval amongst harmonics, gives rise to all-octave equivalence only as a result of learning. This has the following testable consequences: all comparisons between a fundamental and its harmonics or subharmonics would exhibit anomalous psychophysical effects (in tuning sensitivity, pitch fusion, generalization) that differ from comparisons between inharmonically related frequencies; certain effects would be stronger for harmonics than for subharmonics; effect strength would decrease with harmonic number; and all-octave generalization would be stronger in trained musicians than in untrained.

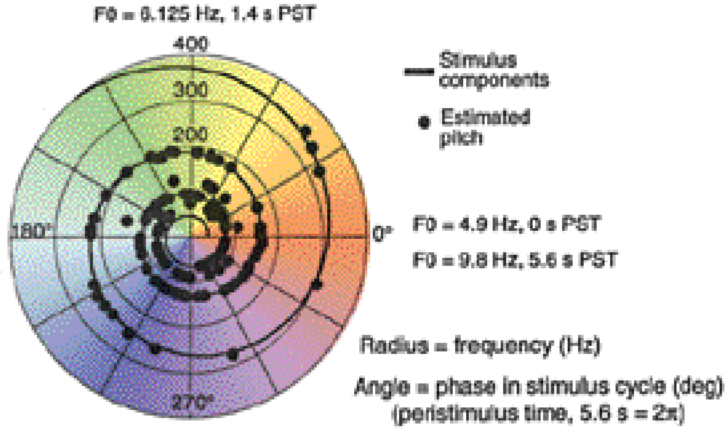

Figure 8: Cariani and Delgutte's algorithm (1996) predicts subharmonic pitch mistakes. Dots represent pitches estimated as the most common interspike interval in pooled recordings from cat auditory nerve fibers, when stimulated by Shepard tone complexes composed of all octaves of a given pitch. The color represents the pitch chroma on which the tone is based. The black spiral line represents frequency components physically present in the stimulus. Most estimated pitches lie on this line. Those that do not lie instead on another spiral (not drawn) which corresponds to the second subharmonic of one of the frequency components, which is not present in the stimulus.

Experiments

Specific Aims

A. *Psychophysics*: Test whether musical and nonmusical human subjects demonstrate all-octave equivalence or near-octave equivalence in tuning sensitivity, pitch fusion, and generalization, thus establishing the relevance of harmonic and subharmonic series in octave perception.

B. *Animal Behavior*: Develop an animal behavioral assay for tuning sensitivity and generalization.

C. *Auditory System Development Modification*: Measure generalization patterns in animals raised in artificial acoustic environments to test whether octave perception is learned from natural structure of the auditory world.

D. *Electrophysiology*: Collect evidence for a possible neural mechanism involved in octave perception.

A. Psychophysics: Human discriminabilities for twelfths and double-octaves

Much work has been done on testing humans for our ability to perceive octaves, as described above. No one to my knowledge has compared our ability to discriminate, tune, or generalize to larger intervals of the harmonic series, however. Most studies have focussed on consonant intervals within a single octave, e.g. the major third, the perfect fifth. Yet for the current investigation, we are asking why the octave is special, rather than assuming it to ask why other consonances are special. So this existing body of data is alas not sufficient. We want to examine whether notes of a double-octave are perceptually more or less similar than those of a twelfth.

I reiterate that measuring the ‘similarity’ of intervals by simply asking humans for subjective judgments along some arbitrary similarity scale alas is uninterpretable data, as the precise meaning of such responses cannot be determined. Non-subjective tasks must be used for such measurements. Not all of these have direct correspondence with our commonsense notions of similarity, but they nonetheless form a dimension along which we can assess the uniqueness of intervals.

Three pyschophysical measurements will be useful. Firstly, one can measure subjects’ ability to choose the closest match to a target interval. This must be done as a comparison task to avoid biasing results towards nameable intervals. Secondly, one can measure the degree of harmonic fusion of two simultaneous pure tones. Thirdly, one can measure subconscious response generalization to punishment-associated stimuli: after one stimulus is negatively reinforced with mild electric shock, similar stimuli may elicit an involuntary skin galvanic response as in a lie-detector test. All three measures are useful in determining the perceptual significance of specific intervals, and all three should be applied to a sampling of consonant and nonconsonant intervals ranging from two and a half or three octaves below a given fundamental to two and a half or three octaves above.

The first psychophysical experiment would ask humans to match one of two simultaneous tone pairs to a target tone pair. One pair would be a target repeat and the other would be a similar tone pair where one note is detuned. The degree of detuning would be varied from large to small in an adaptive procedure. Correct matching would reduce the detuning; incorrect answers would increase it. The threshold detuning would be indicated by a sufficient number of alternations between correct and incorrect answers. This procedure would be repeated for a series of both harmonic and inharmonic intervals at a number of different fundamental frequencies.

The second experiment would replicate Demany and Semal's test for dichotic fusion (1990), except over a larger range of intervals. Two pure tones related by an interval to be tested would be presented to the left and right ears individually. Both of these tones would be frequency modulated at 2 Hz with a 10% frequency change. Thus, the interval between them would be preserved. Then, the phase of the frequency modulation cycle would be changed, so that the two frequencies no longer rose and fell in parallel. In an extreme case, where the modulation phase would differ by one half-cycle, the upper tone would be rising while the lower one was falling (Figure 9). In the first case, when the two tones modulate in phase, listeners report a fusion wherein the two tones become a single pitch; in the second, the two tones cannot fuse and are perceived as distinct. As the phase difference grows increases from zero, a loss of fusion becomes more and more apparent. The ratio fluctuations are to a good approximation a linear function of the modulation phase difference. Therefore, the smallest detectable phase difference becomes a measure of just-noticeable interval fluctuation. Demany and Semal found that fusion occurred best when

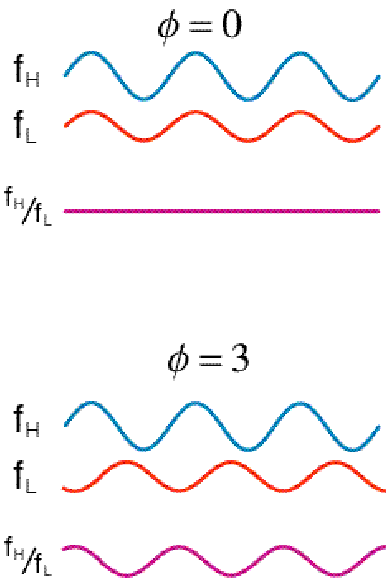

Figure 9: Demany and Semal's stimulus for measuring pitch fusion. Two tones, high and low (blue and red) are modulated in frequency and presented to different ears. As the relative phase (ϕ) between the frequency modulation changes, so does the interval between the two tones (purple).

the fluctuating interval was centered on an octave (Figure 10). For other intervals fusion did not occur; the smallest detectable phase difference was much larger because subjects could not rely on cues of pitch fusion.

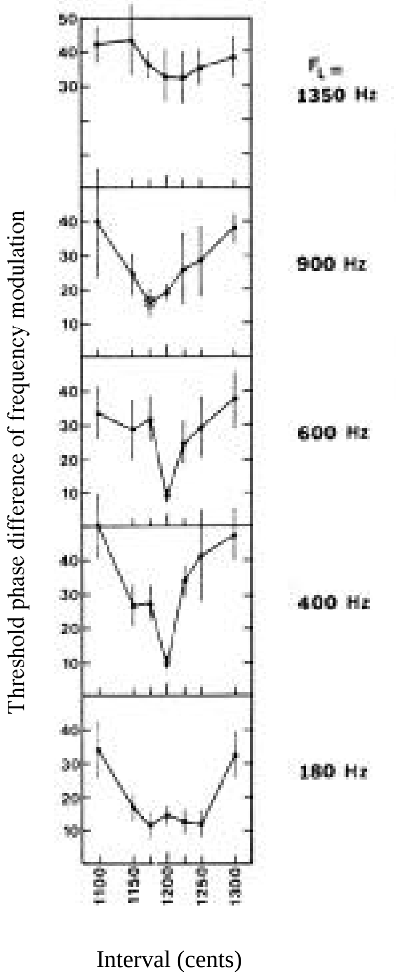

Figure 10: Improved pitch fusion for octave-related tones, measured by just-noticeable phase difference between high and low frequency modulated tones. Frequency to right is that of lower frequency (from Demany and Semal 1990).

The third experiment would involve subjects volunteering for a test involving unpleasant but not painful mild shock conditioning. Shock would be administered along with one target pure tone presentation. Conditioning leads to a measurable involuntary change of skin conductance upon presentation of this same tone. If a slightly detuned tone is presented instead, if the detuning is unnoticeable, then the same galvanic skin response will result. Response was shown to decrease with detuning, except that octave detuning resulted in anomalously large response (Humphreys 1939), attributed to a perceptual generalization. This anomaly should be measured for the key intervals of a double-octave and a twelfth, to see whether humans generalize to all octaves or to harmonics.

In all of the above tests, both the harmonically relevant intervals and nearby intervals should be tested. I expect all of the measures to reveal that harmonic-related intervals are perceived as different from nearby non-harmonic intervals, which should be a verification of previous psychophysical studies of consonant intervals. These experiments will test the verity of my hypothesis that our perceptual system has the property of near-octave equivalence rather than all-octave equivalence, by measuring different perceptual attributes of twelfths compared to double-octaves, namely discrimination thresholds, pitch fusion, and subconscious generalization.

Furthermore, because I hypothesize that double-octave specialness is due to musical training, by segregating the test subjects into three groups based on musical experience, we should see a distinct correlation between group and upper-octave perception. Tone deaf people should have no particular facility for discriminating harmonics or octaves, the normal but musically naïve should exhibit some near-octave equivalence, while the musically educated should exhibit all-octave equivalence.

B. Animal Behavior: Test whether double-octaves or twelfths are more discriminable by rats.

To test the mechanism underlying octave perception, we need an animal model. First, we must find an animal that actually perceives octaves.

How can we test whether animals can detect octaves? We can't just ask them if two notes sound the same. Blackwell and Schlosberg (1943) found that a rat will make subharmonic octave mistakes in behavioral tests relying on pure tone discrimination. However, their conclusion, that rats perceive octaves, is not the only interpretation of these results. Another interpretation could be that rats are perceiving the first subharmonic, and not actually identifying it as an octave. In this interpretation, the number of mistakes would decrease with subharmonic or harmonic number. In particular, a frequency ratio of 3:1 would be more often mistaken than one of 4:1. The hypothesis, that octaves are truly special amongst harmonics, would argue the reverse. Thus, these two comparisons would distinguish between the hypotheses. Since the two intervals have a fairly large separation by frequency ratio, it might be better to consider also 7:1, 8:1, and 9:1.

Yet it seems quite likely that the quality of data will decay with the substantial change in pitch height: rats will make discrimination to a great extent on the absolute frequency and slight effects of harmonic overlap will be imperceptible. Thus a slightly new assay is necessary to keep each rat listening for a narrow band of pitch height. I propose training rats to match the various harmonic *combinations* and determine their learning rate and eventual frequency discriminability. Different rats would be trained to respond to a pure tone pair of a fundamental, f, and a harmonic, nf. The octave hypothesis would hold that rats learn octaves best; the harmonic series hypothesis would predict a diminishing performance with harmonic number. Other rats should be trained to identify an inharmonic pair, and both hypotheses would predict that harmonic pairs are learned better than inharmonic pairs.

With these new stimuli, I would attempt a variation on Blackwell and Schlosberg's experiment, which surprisingly (or not) has not been replicated since 1943. The apparatus is as follows. A rat is placed in a box with two chambers connected by an L-shaped tube which can be electrified for foot shocking. Each chamber will have an electronically controllable door, food dispenser, and photodiode for detection of the rat's position. A hi-fi speaker system will deliver pure tone stimuli. The experiment will consist of an acclimation period, a training period, and a testing period. During the acclimation period, the target combination tone will be presented, of frequencies f and nf, after which the chamber doors will be opened so the rat is given opportunity to leave the current chamber and traverse the unelectrified L-tube to find the food reward in the opposite chamber. The photodiode will sense the entry of the rat into the chamber and the doors will be closed. Time of departure and arrival will be digitally recorded. During the training phase, the punishment is introduced. Just as in the acclimation period, target combination tone stimuli will be presented and the doors opened so the rat may obtain the food reward. This ‘go’ stimulus will be randomly interspersed with ‘no-go’ events: the rat’s current chamber door will open without any tone stimulus, and the L-tube will be electrified for footshock. The far chamber door will remain closed to prevent the rat from canceling the punishment with reward. Finally, the testing phase will resemble the training phase except the ‘no-go’ silence will be replaced with a new stimulus, with the same fundamental f but a detuned harmonic. The amount of detuning on no-go trials will start quite large and decrease from day to day as the rat learns.

Raw data will consist of the tones produced and the times of rat movements through the tubes. This will be analyzed to find the latency of response and the fraction of correct assessments of go versus no-go. Various groups of rats will be trained to learn discriminate different note pairs, and I expect that the performance of each group should be related to the target frequency ratio, and to a lesser degree on the particular frequencies that make up the interval.

A few controls would establish a baseline for behavioral comparisons. A few rats should be given normal tone stimuli but with randomized reward/punishment, to establish the behavior associated with no learning. A further group should be presented with two different modes, auditory and visual, for the go and no-go stimuli, to establish optimal learning. A final group of rats should have the electric shock replaced with an animatronic cat, to simulate a more natural aversion response.

I hypothesize that the specialness of the octave arises from purely harmonic structure, and thus this would demand that in this task, rats will be less able to detect well-tuned double- or triple-octaves (frequency ratios 4:1 or 8:1) than they can twelfths or twenty-oneths (3:1 or 7:1) respectively; yet they will be more able to detect well-tuned harmonics than they can inharmonic pairs.

C. Auditory System Development Modification: Are octaves learned or hardwired?

If rats are able to detect octaves or harmonics with particular accuracy, then we will be able to ask whether they derive this ability from genetically hardwired circuits or whether it is learned. Learning would be the likely response, considering that many of the important natural sounds they hear are generated by vibrating membranes and air columns which produce correlated harmonic spectra. It would be nice to test how they learn if the natural environment taught inharmonic correlations or lack spectral correlations altogether; alas it is impossible to change the laws of physics. A study by Liberman (1977) found that the auditory system of cats was slightly altered by raising them in a low-noise chamber; one might try to alter the rats' auditory system with such silence, or with noise, or with masking tones, or with notched noise. More interestingly, one could use active noise control to interfere with the effects of the natural physical processes. Feedback circuits linking a microphone to a noise-canceling speaker or earphones can reduce broadband noise by 15 dB or more and can diminish pure tones by 30 dB (Figure 11). Thus — as an auditory analog of Stryker's 1978 experiment in which the vision of kittens was restricted to stripe patterns presented via goggles — one could use destructive interference to hide the natural structure of the acoustic environment and replace those normal sounds with distorted ones. For example, a frequency analyzer could identify and then subdue all harmonics and strengthen their fundamentals, or shift harmonics by a constant frequency, or augment only the second harmonic.

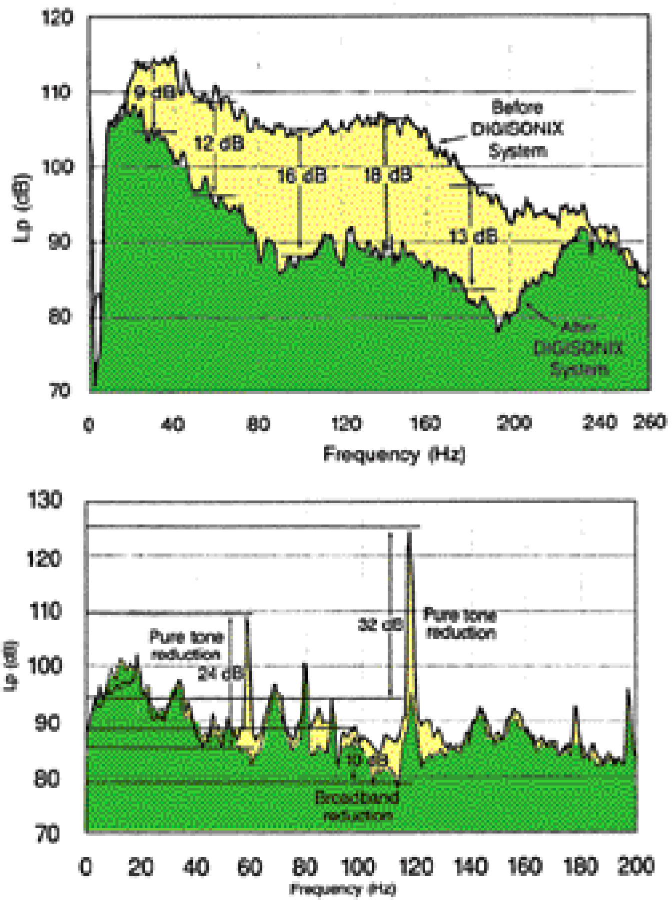

Figure 11: Active noise cancelation efficiency. Traces show power spectrum before (yellow) and after (green) negative feedback, for broadband noise (upper graph) and individual tones (lower graph) (from Eargle 1995).

Rats' auditory systems are not fully wired at birth (Friauf 1993), and so one could raise rats in a noise-cancellation chamber to allow improved control over the auditory stimuli they receive during postnatal development. The type of stimuli should determine their later behavioral responses in the assay described above. If harmonics are suppressed, behavioral tests should show dramatically reduced octave generalization since the relevant covarying information from harmonic structure would have been absent and never learned; only pure tone neurons would be strengthened, and connections amongst the laminae of the inferior colliculus would likely be disrupted. If harmonics were shifted during postnatal development, the behavioral generalization that normally would occur for octaves would occur for shifted octaves. If the fundamental and first harmonic were strengthened and the other harmonics reduced, then octave generalization should be stronger than average, but upper harmonic generalization should be weaker.

Some data suggests (Demany and Armand 1984) that 3-month-old infants possess octave templates; the study concludes that octave templates may be hardwired and not learned. Three months of learning may be sufficient to establish the harmonic patterns recognition circuits that I hypothesize exist.

D. Electrophysiology: Neuronal mechanism of harmonic perception

If rats do indeed perceive octaves as Blackwell and Schlosberg showed, then I may attempt to explain the results of the behavioral studies in a neurophysiological context. As described above, the laminar structure of the inferior colliculus may be perfectly suited to extract the fundamental. The circuits that are tuned to extract a particular fundamental may occasionally make octave mistakes, randomly and stochastically; and it might even make larger subharmonic mistakes, with a probability that would depend on the harmonic number. Indeed, such a model could quantitatively predict the accuracy of tone-matching upper harmonics by correlating this to the likelihood that a subharmonic mode would arise by chance due to the properties of neurons that give a stochastic input, of phaselocked firing probablity, to the coincidence detector.

A search in the inferior colliculus might expose such neurons. In the bat, neurons with intrinsic subthreshold oscillations were found. The frequency of oscillations should, in Braun's model, be correlated with that neuron's spectral receptive field, not to pure tones but to combination tones. Two-lobed single-tone receptive fields preferring harmonically related frequencies have been measured in the auditory cortex of the mustache bat (Suga 1976) and cat (Evans 1974). Better, two-tone receptive fields that are not simply the trivial addition of two single-tone receptive fields have been measured by Nelken *et al* (1997) in the cat dorsal cochlear nucleus, an earlier region in the auditory pathway.

To isolate these receptive fields, one could record extracellularly from the inferior colliculus while presenting a series of stimuli consisting of two pure tones of equal loudness, one whose frequency sweeps logarithmically across the animal's audible range and the other whose frequency is constant for each stimulus but increases over the series. This gives an input matrix in which every frequency is played at some time with every other frequency. Each frequency pair is actually presented twice within the series: once the upper component would be sweeping, and once the lower would be. This provides an internal control against the effects of sweeping the frequency slowly compared to presenting a constant frequency. The whole series would be presented several times, to increase the number of samples. The average firing rate of the neurons would constitute the neural response. To refine the receptive field one could reapply the frequency pairs that elicited the greatest responses, but with more stimulus presentations, finer frequency steps, and a slower frequency ramp. Plotting the responses against the input frequency pair would yield the two-tone receptive field matrix. The diagonal of this matrix would correspond to the single-tone receptive field. Off-diagonal structure — ones that are not horizontal or vertical bands aligned with peaks along the diagonal — would correspond to nonlinear interactions between two frequencies. One ought to examine the temporal structure of the response at these frequency pairings, and should detect phase locking at a frequency of the missing fundamental.

Implications, Interpretations and Renovations

A. Psychophysics

What is the connection between ‘similar sounding’ and the three proposed psychophysical measures: increased discriminability, pitch fusion, and generalization? If special circuits exists to identify harmonics within sounds, then harmonic frequency complexes will have more processing dedicated to their perception than inharmonic frequency complexes will have. Consequently, the brain will have a greater discriminability for harmonic complexes. If the similarity of two tones is a result of this same circuitry, then measurements of generalization and discriminability are correlated. Correspondingly, specialized harmonic-detection circuitry could be a common cause both pitch fusion and similar-sounding effects. Both these measures should exercise the circuitry responsible for harmonic detection. The third psychophysical experiment, asking whether humans have a subconscious response generalization to punishment-associated stimuli, is the most direct test of the claim that, to our subconscious, the tested interval ‘sounds the same’. All of the psychophysical experiments would be capable of demonstrating anomalous perception of certain key intervals.

Our response to different intervals will vary based on the frequency range tested. Our responses may differ not only in degree but in type: for example, people with absolute pitch find that above about 4 kHz, all pitches appear to have the same chroma (Ward 1954). This experiment should be attempted within our aural fovea, so that pure tone frequencies range between 0.5 to 2 kHz. Low pitches, while outside our fovea, will have to be used since the foveal range only spans a double-octave, which is the interval we wish to examine most closely. Thus, another one or two octaves below, to 100 Hz or so, may be used as stimuli. When such frequencies are used, our discrimination thresholds will be higher on average, and thus two-tone tuning thresholds might be most clearly analyzed relative to the pure tone threshold.

B. Animal Behavior

What happens to the experiments if the proposed animal behavioral tests do not show any signs that rats generalize to other frequencies, or that the generalization is a monotonically decreasing function of interval size? One possible solution would be to modify the acclimation and training periods to introduce all possible stimuli earlier, not only target stimuli. Another is to introduce new frequencies at a lower loudness during the testing phase, and raising volume gradually across sessions before testing for discrimination thresholds. One might use only 50% reinforcement, which Humphreys (1939) suggests leads to stronger generalization. Songbirds (Cynx 1993) failed to exhibit any octave generalization, which could be interpreted as a very discriminating ear, perhaps with a strong sense of absolute pitch. It might be possible to confuse this highly developed sense with noise or small changes in loudness, and measure a correspondence between behavior and threshold-adjusted loudness. If no generalization occurs in any conditions, importantly, this would call into question the oft-cited psychology experiments of Blackwell and Schlosberg. Other animals could also be used for behavioral tests. Monkeys and cats both perceive the missing fundamental of a harmonic complex (Tomlinson 1988, Heffner 1975), so assays may be adaptable to testing these animals for octave equivalences. Ferrets would be well-suited for experiments involving artificial acoustic environments during postnatal development because they have very immature auditory systems at birth (Friauf 1993).

C. Auditory System Development Modification

Active noise control can lower noise levels substantially, around 30 dB for a pure tone and 15 dB for noise, but active noise analysis and manipulation requires fast computation and short sound travel time. To cancel or modify animal vocalizations, there can be but a few milliseconds before the auditory system resolves the sound. Even with instant feedback, external speakers at a distance of 30cm from the animal would pass 40 cycles of a 40kHz tone due only to sound wave propagation delay, which would cause the cancelled stimulus to sound at the very minimum like a series of clicks. Normally raised rats could be tested with such stimuli after training to measure minimum stimulus duration required for tone recognition. Though harmonic structure may not be resolvable with such short durations, nonetheless some response should be expected. Use of earphones might shorten the delay further, but still there will be cranial vibrations from the vibrating larynx which may excite the auditory system. Perhaps this too could be damped by appropriate acoustic feedback. Earphones would also eliminate the spatial dependence of the feedback.

A simpler and cheaper, though less humane, approach would be to surgically destroy the animals' larynges. In this case, noise reduction would be much simpler, and manipulation of the harmonic structure of sound need not be done in real-time, leading to more complete control over the acoustic environment and a total reduction of active noise control artifacts. However, there is a chance that rats would need sounds associated with other forms of communication or with other animals to learn to respond correctly to sound.

D. Electrophysiology

The hypothesized mechanism for harmonic perception is an extension of a mechanism of periodicity detection proposed by Langner (1997) and Braun (1999). The computation that results in near-octave equivalence may be simple enough to occur in the brainstem, and this would be a likely place if animals are found to possess near-octave equivalence. All-octave equivalence would be a higher-level computation, and might occur in the auditory cortex. Deutsch (1973) proposed a mechanism whereby certain neural units would receive inputs from all neurons which respond to one particular pitch chroma, and thus respond equally to any given octave. While conceivable, this mechanism seems to overestimate the strength of any octave equivalence by attributing to it an entire dedicated circuit. Regardless, her hypothesis is difficult to validate or invalidate at this time, because such a verification would require a search through the auditory cortex for a particular neural response. My hypothesized mechanism has a more restricted associated location which makes testing easier. Additionally, stochastic firings of neurons have been quantitatively measured, and some theoretical arguments about the probability of occasional subharmonic input yield psychophysical predictions that could invalidate the model without invasive animal experimentation.

The neural mechanism of near-octave equivalence may not be correct. Yet, even if the hypothesis that subharmonic firing patterns make octaves sound similar is not true, nonetheless the question remains whether octaves are naturally unique among harmonics, or whether octaves are just one harmonic relation of many which is reinforced through learning of musical tradition. The proposed psychophysical and behavioral tests should still shed important light on this.

Conclusion

Theoretical arguments — that consonance and dissonance arise from mathematical or physical relationships such as simple frequency ratios — do not hold: intervals that are consonant in the 1 kHz range become dissonant or muddled in the low region of our hearing. Clearly, the basic mechanisms of perception plays a huge role in the perception of music. Similarly, it is a central assumption in music theory that all octaves function equivalently in harmony; it is a supposedly innate attribute of a frequency ratio of 2:1, the simplest possible nontrivial relationship. Beyond such numerical mysticism, no satisfactory theoretical reason for all-octave equivalence has ever been given. I directly challenge this assumption. Instead, I propose that octave equivalence, in the strong all-octave sense required by music theory, is only a by-product of tradition and training but does have its solid foundation in a weaker near-octave equivalence, which may be caused by stochastic subharmonic mistakes in the firings of periodicity detectors in the laminae of our brainstem inferior colliculus.

Acknowledgments

Thanks first to Charles Brodhead the Third, whose enthusiastic musical curiosity brought me to realize this gaping hole in music theory. Big thanks to Nick Choly and Martin Sandbu for open ears and orderly ideas, and for rational support during irrational times, whatever that means. Danke schön to Mary Greitzer, who devoted segments of her discerning ear, Pitch and mind, and was a gracious guinea pig in naming notes through static and senselessness. Thanks also to Markus Meister and his laboratory, Steve Bacchus, Tim Holy, Daniel Hwang, Justin Keat, Bence Ölveczky, and Ed Soucy, for their perpetually witty and wise comments and critiques. Thanks to Bertrand Delgutte and M. McKinney for relevant references and experimentally experienced explications. Finally, thanks to Jim Hogle and Harvard Biophysics for encouraging and supporting risky but fun forays into the unknown.

References

Allen D. Octave discriminability of musical and non-musical subjects. *Psychon Sci*, **7**(12): 421-2, 1967.

Bachem A. Time factors in relative and absolute pitch determination. *J Acoust Soc Am* **26**: 751-3, 1954.

Blackwell HR, Schlosberg H. Octave generalization, pitch discrimination, and loudness thresholds in the white rat. *J Exp Psychol* **33**: 407-19, 1943.

Boer E de. On the 'residue' and auditory pitch perception. In *Handbook of Sensory Physiology*, eds. Keidel WD, Neff WD. Springer-Verlag, Berlin, 1976.

Braun M. Frequency spacing of multiple spontaneous otoacoustic emissions shows relation to critical bands: a large-scale cumulative study. *Hearing Research*, **114**: 197-203, 1997.

Braun M. Accurate binaural mirroring of spontaneous otoacoustic emissions suggests influence of time-locking in medial efferents. *Hearing Research* **118**: 129-38, 1998.

Braun M. Auditory midbrain laminar structure appears adapted to f_0 extraction: further evidence and implications of the double critical bandwidth. *Hearing Research* **129**: 71-82, 1999.

Burns EM, Ward WD. Categorical perception—phenomenon or epiphenomenon: Evidence from experiments in the perception of melodic musical intervals. *J Acoust Soc Am* **63**(2): 456-68, 1977.

Cariani PA, Delgutte B. Neural correlates of the pitch of complex tones. I. Pitch and pitch salience, II. Pitch shift, pitch ambiguity, phase invariance, pitch circularity, rate pitch, and the dominance region for pitch. *J Neurophysiol* **76**(3): 1698-1734, 1996.

Covey E, Kauer JA, Casseday JH. Whole-cell patch-clamp recording reveals subthreshold sound-evoked postsynaptic currents in the inferior colliculus of awake bats. *J Neurosci*, **16**(9): 3009-18, 1996.

Cynx J. Auditory frequency generalization and a failure to find octave generalization in a songbird, the European Starling (*Sturnus vulgaris*). *J Comp Psychol* **107**(2): 140-6, 1993.

Demany L, Armand F. The perceptual reality of tone chroma in early infancy. *J Acoust Soc Am* **76**: 57-66, 1984.

Demany L, Semal C. Dichotic fusion of two tones one octave apart: Evidence for internal octave templates. *J Acous Soc Am*, **83**(2): 687-95, 1988.

Demany L, Semal C. Harmonic and melodic octave templates. *J Acoust Soc Am* **88**(5): 2126-35, 1990.

Deutsch D. Octave generalization of specific interference effects in memory for tonal pitch. *Perception & Psychophysics* **13**: 271-5, 1973.

Deutsch D. "The Processing of Pitch Combinations," in *The Psychology of Music*, ed. Deutsch D, Academic Press, New York, 1982.

Eargle JM. "Active Noise Control and Acoustic Modification," in *Music, Sound and Technology*, van Nostrand Reinhold Publishing, Boston, 1995.

Ehret G, Merzenich MM. Auditory midbrain responses parallel spectral integration phenomena. *Science*. **227**(4691):1245-7, 1985.

Ellis CJ. Pre-instrumental scales. *Ethnomusicology* **9**: 126-44, 1965.

Evans. in *The Neurosciences: third study program*, eds. Schmitt FO, Worden FG, MIT Press, Cambridge, 1974.

Friauf E, Kandler K. "Cell birth, formation of efferent connections, and establishment of tonotopic order in the rat cochlear nucleus," in *The Mammalian Cochlear Nuclei: Organization and Function*, ed. Merchán MA et al. Plenum Press, New York, 1993.

Galbraith GC. Two-channel brain-stem frequency-following responses to pure tone and missing fundamental stimuli. *Electronencephalography and clinical Neurophysiology* **92**: 321-30, 1994.

Heffner H, Whitfield IC. Perception of the missing fundamental by cats. *J Acoust Soc Am* **59**(4): 915-9, 1976.

Houtgast T. Subharmonic pitches of a pure tone at low S/N ratio. *J Acoust Soc Am* **60**(2): 405-9, 1976.

Houtsma AJM, Goldstein JL. The central origin of the pitch of complex tones: evidence from musical interval recognition. *J Acoust Soc Am*, **51**: 520-9, 1972.

Houtsma AJM, Smurzynski J. Pitch identifcation and discrimination for complex tones with many harmonics. *J Acoust Soc Am* **87**: 304-10, 1990.

Humphreys LJ. Generalization as a function of method of reinforcement. *J Exp Psych* **25**: 361-72, 1939.

Kärnbach C, Demany L. Psychophysical evidence against the autrocorrelation theory of auditory temporal processing. *J Acoust Soc Am* **104**(4): 2298-2306, 1998.

Kallman HJ. Octave equivalence as measured by similarity ratings. *Perception and Psychophys* **32**(1): 37-49, 1982.

Kim DO. Active and nonlinear cochlear biomechanics and the role of outer-hair-cell subsystem in the mammalian auditory system. *Hearing Research,* **22**: 105-14, 1986.

Langner G. Neural processing and representation of periodicity pitch. *Acta Otolaryngol (Stockh) Suppl* **532**: 68-76, 1997.

Langner G, Schreiner C, Merzenich MM. Covariation of latency and temporal resolution in the inferior colliculus of the cat. *Hearing Research,* **31**: 197-202, 1987.

Langner G, Schreiner CE. Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. II. Topographical organization. *J Neurophysiol* **60**(6):1799-1840, 1988.

Levelt WJM, van de Geer JP, Plomp R. Triadic comparisons of musical intervals. *Brit J Math & Stat Psych* **19**(2): 163-79, 1966.

Liberman MC. Auditory-nerve response from cats raised in a low-noise chamber. *J Acoust Soc Am* **63**(2):442-55, 1978.

Licklider, JCR. "Periodicity" pitch and "place" pitch (Abstract). *J Acoust Soc Am,* **26**: 945, 1954.

Lipps T. *Consonance and Dissonance in Music,* 1905, trans. Thomsen W. Everett Books, San Marino, 1995.

Long GR, Talmadge CL. Spontaneous otoacoustic emission frequency is modulated by heartbeat. *J Acoust Soc Am,* **102**(5): 2831-48, 1997.

Malmierca MS, Rees A, le Beau FEN, Bjaalie JG. Laminar organization of frequency-defined local axons within and between the inferior colliculi of the guinea pig. *J Comp Neuro,* **357**: 124-44, 1995.

McKinney MF, Delgutte B. A possible neurophysiological basis of the octave enlargement effect. *J Acoust Soc Am* **106**(5): 2679-92, 1999.

Nelken I, Kim P, Young ED. Linear and nonlinear spectral integration in type IV neurons of the dorsal cochlear nucleus. II. Predicting responses with the use of nonlinear models. *J Neurophysiol* **78**: 800-11, 1997.

Nienhuys TGW, Clark GM. Critical bands following the selective destruction of cochlear inner and outer hair cells. *Acta Otolaryngol* **88**: 350-8, 1979.

Oliver DL, Kuwada S, Yin TCT, Haberly LB, Henkel CK. Dendritic and axonal morphology of HRP-injected neurons in the ingerior colliculus of the cat. *J Comp Neurol* **303**: 75-100, 1991.

Oliver DL, Winer JA, Beckius GE, Saint Marie RL. Morphology of GABAergic neurons in the inferior colliculus of the cat. *J Comp Neurol* **340**: 27-42, 1994.

Palombi PS, Caspary DM. GABA inputs control discharge rate primarily within frequency receptive fields of inferior colliculus neurons. *J Neurophys,* **75**(6): 2211-9, 1996.

Pickles JO. Psychophysical frequency resolution in the cat as determined by simultaneous masking and its relation to auditory-nerve resolution. *J Acoust Soc Am* **66**(6): 1725-32, 1979.

Plack CJ, Carlyon RP. Differences in frequency modulation detection and fundamental frequency discrimination between complex tones consisting of resolved and unresolved harmonics. *J Acoust Soc Am* **98**: 1355-64, 1995.

Plomp R, Wagenaar WA, Mimpen AM. Musical interval recognition with simultaneous tones. *Acustica,* **29**: 101-9, 1973.

Rasch RA, Plomp R. "The Perception of Musical Tones," in *The Psychology of Music*, ed. Deutsch D. Academic Press, Inc. New York, 1982.

Risset JC. Pitch and rhythm paradoxes: comments on "Auditory paradox based on fractal waveform". *J Acoust Soc Am* **80**(3):961-2, 1986.

Ritsma RJ. Existence Region of the Tonal Residue II. *J Acoust Soc Am,* **35**(8): 1241-5, 1963.

Ritsma RJ, Bilsen FA. Spectral regions dominant in the perception of repetition pitch. *Acustica* **23**: 334, 1963.

Schouten JF, Ritsma RJ, Cardozo BL. Pitch of the residue. *J Acoust Soc Am* **34**: 1418-24, 1962.

Schreiner CE, Langner G. Laminar fine structure of frequency organization in auditory midbrain. *Nature* **388**: 383-6, 1997.

Schwarz DWF, Dezsö A, Neufeld PR. Frequency selectivity of central auditory neurons without inner ear. *Acta Otolaryngol (Stockh),* **113**: 266-70, 1993.

Shackleton TM, Carlyon RP. The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination. *J Acoust Soc Am,* **95**: 3529-3540, 1994.

Shepard RN. Circularity in judgments of relative pitch. *J Acoust Soc Am,* **36**(12): 2346-53, 1964.

Smith JC, Marsh JT, Greenberg S, Brown WS. Human auditory frequency-following responses to a missing fundamental. *Science,* **201**: 639-41, 1978.

Spirou GA, Young ED. Organization of dorsal cochlear nucleus type IV unit response maps and their relationship to activation by bandlimited noise. *J Neurophysiol* **66**: 1750-68, 1991.

Steinschneider M, Reser DH, Fishman YI, Schroeder CE, Arezzo JC. Click train encoding in primary auditory cortex of the awake monkey: Evidence for two mechanisms subserving pitch perception. *J Acoust Soc Am* **104**(5): 2935-54, 1998.

Stryker MP, Sherk H, Leventhal AG, Hirsch HVB. Physiological consequences for the cat's visual cortex of effectively restricting early visual experience with oriented contours. *J Neurophys* **41**:896-909, 1978.

Suga N, Jen PHS. Disproportioate tonotopic representation for processing CF-FM sonar signals in the mustache bat auditory cortex. *Science* **194**: 542-4, 1976.

Suga N, O'Neill WE, Manabe T. Harmonic-sensitive neurons in the auditory cortex of the mustache bat. *Science* **194**: 542-4, 1979.

Sundberg JEF, Linqvist J. Musical octaves and pitch. *J Acoust Soc Am*, **54**(4): 922-9, 1973.

Terhardt E. Pitch, consonance, and harmony. *J Acoust Soc Am*, **55**(5): 1061-9, 1973.

Terhardt E, Zick M. Evaluation of the tempered tone scale in normal, stratched [sic], and contracted intonation. *Acustica* **32**: 268-74, 1975.

Tomlinson RWW, Schwarz DWF. Perception of the missing fundamental in nonhuman primates. *J Acoust Soc Am* **82**(2): 560-5, 1988.

van den Brink G. "Octave and fifth settings for pure tones and residue sounds," in *Psychophysics and Physiology of Hearing*, ed. Evans EF, Wilson JP. Academic Press, New York, 1977.

von Helmholtz H.*On the Sensations of Tone as a Physiological Basis for the Theory of Music*, trans. Ellis AJ. Longmans, Green, and Co., New York, 1895.

Ward WD. Subjective Musical Pitch. *J Acoust Soc Am*, **26**(3): 369-80, 1954.

Webster WR, Serviere J, Martin R, Brown M. Uncrossed and crossed inhibition in the inferior colliculus of the cat: a combined 2-deoxyglucose and electrophysiological study. *J Neurosci*, **5**(7): 1820-32, 1985.

White L, Plack C. Temporal processing of the pitch of complex tones. *J Acoust Soc Am*, **103**(4): 2051, 1998.