[Frontiers in Bioscience 12, 4473-4482, May 1, 2007]

Music perception

Diana Deutsch

Department of Psychology, University of California, San Diego, La Jolla, CA 92093

TABLE OF CONTENTS

1. ABSTRACT

This chapter explores the relationship between music perception as it is studied in the laboratory and as it occurs in the real world. We first examine general principles by which listeners group musical tones into perceptual configurations, and how these principles are implemented in music composition and performance. We then show that, for certain types of configuration, the music as it is perceived can differ substantially from the music that is notated in the score, or as might be imagined from reading the score. Furthermore, there are striking differences between listeners in the perception of certain musical passages. Implications of these findings are discussed.

2. INTRODUCTION

The study of music perception encompasses a broad range of phenomena, including the perception of basic attributes of sound such as pitch, duration, and loudness, the principles by which lower-level features are extracted so as to produce higher-level features, the perception of large-scale musical structures, cultural influences on music perception, developmental issues, aberrations of music perception; and so on. The present chapter focuses on certain issues that are particularly applicable to the perception of music in the real world. First, we consider general principles of perceptual organization and show how they are applied to live musical situations. Second, we show that music as perceived can, for certain configurations, be quite different from that in the written score, or as might be imagined from reading a score. Instead, striking illusions can occur on listening to music, and there are strong differences between listeners in the way that some of these illusions are perceived. In some cases, such perceptual differences correlate with handedness, and so can be taken to reflect variations in innate brain organization. In the case of another illusion, perceptual differences correlate with the language or dialect to which the listener has been exposed, particularly in childhood, and point to an influence of exposure to extramusical phenomena on the perception of music.

3. GROUPING OF MUSICAL SOUNDS

Suppose you are listening to an orchestral performance in a concert hall. The sound mixture that arrives at your ears is derived from many instruments playing together. What are the principles whereby we group and separate the different components of this mixture, so that we hear the first violins playing one melodic line, the second violins another, and the flutes another? We here examine the perceptual principles that come into play, and consider how they are exploited in music composition and performance.

3.1. Perceptual fusion and separation of spectral components

We first enquire into the relationships between the components of a sound spectrum that cause us to fuse them perceptually so as to form unitary sound images, and those that cause us to separate them perceptually so as to form multiple sound images. In particular, we focus on two types of relationship: harmonicity of the spectral components, and temporal relationships between the spectral components.

3.1.1. Harmonicity

Tones produced by instruments such as the violin, flute, and trumpet, as well as by the singing voice, are composed of partials that stand in harmonic, or near-harmonic relation (that is, their frequencies are integer, or near-integer multiples of the fundamental frequency); such tones give rise to clearly fused pitch impressions. However, sounds produced by bells and gongs are composed of partials that are nonharmonic; these give rise to diffuse pitch impressions (1). Indeed, the closer the components of a complex tone are to strict harmonicity, the stronger is the tendency to attribute a single pitch to the complex (2). Furthermore, when two harmonic complex tones sound together, the closer their fundamental frequencies are to simple harmonic relation, the stronger is the tendency to perceive a single complex tone rather than two complex tones (3). These perceptual tendencies are exploited in compositional practice. For example, in polyphonic music it is desirable that simultaneously sounded tones should stand out clearly from each other, and Bach, in his polyphonic works, avoided intervals that promote tonal fusion (4). In contrast, composers such as Debussy, Ravel, and Varese often endeavored to produce unitary sound images from combinations of instrument tones (5), and for this purpose tones whose fundamental frequencies stand in simple harmonic relation are particularly effective.

3.1.2. Onset and time-varying synchronicity

Onset synchronicity is also influential in the perceptual grouping of sounds. When one component of a complex tone enters before the others, it is more likely to be segregated from them perceptually (6). Furthermore, when two complex tones sound together, onset disparities between them can cause the complexes to be heard out perceptually (3). It is interesting to note that in live musical situations, instrument tones that are nominally synchronous are in reality slightly asynchronous. In one study of live ensemble performance, it was found that notes that were nominally synchronous were in fact offset from each other by roughly 30 - 50 ms. (7). Since such onset asynchronies enable listeners to hear concurrent voices more distinctly, they are useful in certain musical situations. Again, there is evidence that composers have exploited this effect in polyphonic music. For example, in an analysis of Bach's two-part inventions, it was found that for most of the pieces there were no other permutations of the voices that would produce more onset asynchrony than in Bach's actual music (8).

Ongoing temporal relationships have also been hypothesized to play a role in perceptual fusion and separation. Natural sustained sounds, such as produced by stringed instruments and the singing voice, undergo small frequency fluctuations that retain the ratios formed by their component frequencies (9). It has been conjectured that such coordinated frequency modulation strengthens the perceptual fusion of the components of the tones (10) though this issue is currently under debate (11, 12).

3.2. Auditory continuity effects

In real life situations, sounds often arrive at our ears in the presence of other interfering sounds. It makes sense, then, that the auditory system should have evolved strategies for perceptually reconstructing sound components that might otherwise be masked. Returning to the concert hall situation, extraneous noises such as coughs, whispers, and the rustling of papers, would make a musical performance appear intermittent, were it not for such perceptual strategies. A variety of auditory continuity effects have been identified. For example, if a long tone is sounded, and this is interrupted by other tones or noise bursts, the tone may be heard as continuous even though in reality it is intermittent (13). Listeners make sophisticated use of contextual information to reconstruct highly probable sounds that would otherwise be drowned out by extraneous noises. In one experiment, melodic patterns were presented in which certain tones were omitted and replaced by loud noise bursts. Listeners 'heard' the missing tones appear through the noise, and this effect occurred particularly when the missing tones were highly predictable from the musical context (14). Composers have exploited versions of the continuity effect to substantial advantage, and particularly good examples are found in classical and romantic guitar music. For example, in Tarrega's *Recuerdos de la Alhambra* shown in (Figure 1), the strong expectations created by the rapidly repeating upper notes cause the listener to 'hear' these notes even when they are not present.

3.3. Grouping of rapid sequences of tones

We now consider situations in which rapid sequences of tones are presented. Here the auditory system abstracts further relationships between the tones, and uses these as additional grouping cues. This section explores the influence of several such relationships on the way we form perceptual configurations.

3.3.1. Grouping by pitch proximity

When tones are presented in rapid sequence, and these are drawn from two different pitch ranges, the listener perceives two melodic lines in parallel, one derived from the higher tones and the other from the lower ones (10). This perceptual effect of 'stream segregation' is exploited in the compositional technique known as pseudopolyphony, or compound melodic line. The technique was used extensively in the Baroque era by composers such as Bach, Telemann, and Vivaldi, and particularly good examples are also found in twentieth century classical and romantic guitar music. In the passage by Tarrega depicted in (Figure 1), for instance, the listener clearly hears two separate melodic lines in parallel, each in a separate pitch range.

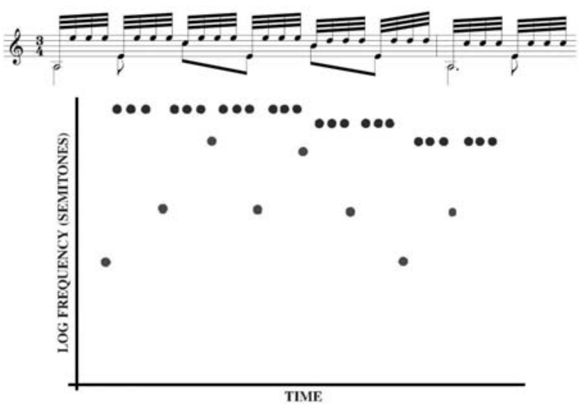

Figure 1. The beginning of Tarrega's *Recuerdos de la Alhambra*. The lower part of the Figure shows the pattern of fundamental frequencies, with log frequency and time mapped into two dimensions of visual space. The perceptual organization of the passage into two streams based on pitch proximity clearly emerges in the visual representation. In addition, listeners perceptually reconstruct the missing higher tones.

Figure 2. Part of the beginning of the second movement of Beethoven's *Spring Sonata* for violin and piano. The tones played by the two instruments overlap substantially in pitch; however the listener perceives two melodic lines that correspond to the tones played by each instrument.

Figure 3. Signals used to study the effect of temporal segmentation on perception of a structured tonal sequence. (a) The pattern with no segmentation. (b) The pattern segmented in accordance with the structure of the sequence. (c) The pattern segmented in conflict with the structure of the sequence.

There is a compositional penalty to disregarding the perceptual tendency to group sounds on the basis of pitch proximity. If two well-known melodies are presented with their component tones alternating in rapid succession, the melodies can easily be identified when they are in different pitch ranges. However, when the tones are all in the same pitch range, the listener instead forms perceptual connections between adjacent tones, so that the two melodies become difficult to separate out, and so to identify (15). (This effect does not hold, though, when the two melodies are played by instruments of very different timbre, as described below.)

One consequence of stream segregation is that it becomes difficult to process temporal relationships between successive tones when they are in different perceptual streams. For example, when tones are presented at a very rapid rate (such as 10/sec), listeners are unable to name the order in which they occur (16). At slightly slower rates, in which order perception can easily be accomplished, there is still a degradation of temporal processing between elements of different perceptual streams. This is manifest, for example, as a reduced ability to detect temporal displacements of tones within such a sequence. It should be mentioned that frequency disparity between successive tones has been found to degrade the processing of temporal relationships in two-tone sequences also (17) though the degree of degradation is greater when long repetitive sequences are presented (18).

3.3.2. Grouping by timbre

When we listen to music, we group together tones that are similar in timbre and separate out those whose timbres are substantially different. So when different types of instrument play in parallel we often form groupings based on timbre, even when the tones produced by the different instruments overlap considerably in pitch. An example is shown in (Figure 2), which is taken from Beethoven's *Spring Sonata*. Here the tones played by the violin and piano are in substantially overlapping pitch ranges, yet listeners nevertheless perceive two melodic lines in parallel, which correspond to the tones played by each instrument.

3.3.3. Grouping by temporal proximity.

When presented with a sequence of tones interspersed with pauses, the listener forms perceptual groupings based on the pauses, and so in accordance with temporal proximity. This effect can be so strong as to override grouping on other principles. (Figure 3(a)) shows a sequence which, when played without pauses, is clearly heard as grouped into three subsequences on the basis of its pitch structure. When pauses are inserted that are in accordance with the pitch structure of the sequence, as in (Figure 3(b)), this grouping effect is even more pronounced. However, when pauses are inserted that are in conflict with pitch structure, as in (Figure 3(c)), listeners instead form groupings based on temporal proximity, with the result that the pitch structure itself becomes difficult to apprehend (19).

4. GROUPING OF TONE SEQUENCES FROM DIFFERENT SPATIAL LOCATIONS.

When listening to ensemble performances, we are presented with multiple streams of tones that arise in parallel from different regions of space. This raises the question of how we form perceptual groupings from such complex signals. Do we link together sound elements in accordance with pitch proximity, timbre, spatial location, or some combination of these? As will be described, all these

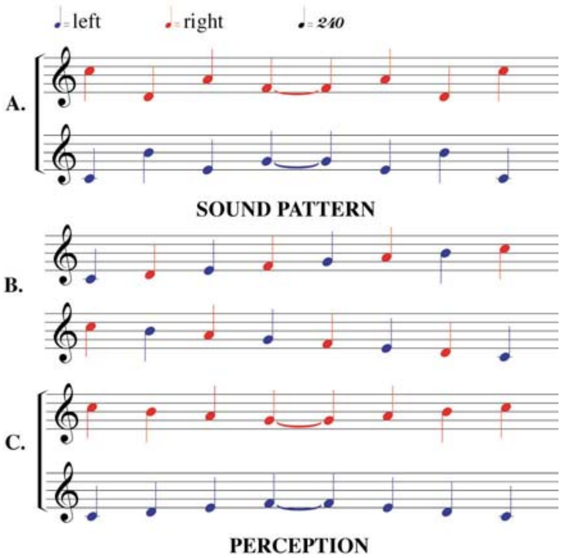

Figure 4. The sound pattern that gives rise to the scale illusion (A), and a percept frequently obtained by righthanders (C). When the pattern is presented via earphones, most righthanders perceive a melody in their right ear that is composed of the higher tones,and a melody in their left ear that is composed of the lower tones. Figure 4B shows how the pattern is composed of ascending and descending scales

factors are here involved in perceptual grouping; however they interact in a complex fashion. So given one configuration, grouping may be largely in accordance with pitch proximity. However, given a slightly different configuration, grouping may instead occur on the basis of spatial location.

In particular, striking illusions can appear in this situation. When presented with a tone, the listener attributes a pitch, a loudness, a timbre, and hears the tone as coming from a particular position in space. Each perceived tone may therefore be described as a bundle of attribute values. There is evidence that these values are determined by separate neural pathways, with each pathway concerned with analyzing a specific aspect of the signal (20*,* 21, 22*)*. This leads to a problem when more than one sound is presented at a time, since the auditory system must then determine which attribute values to connect with which. (This is known as the *binding problem*.) As will be described below, the attribute values of simultaneously sounded tones can fragment and recombine incorrectly, so that illusory conjunctions result.

4.1. The scale illusion

The scale illusion (23, 24) provides an example of such an illusory conjunction. The sequence that produces this illusion is shown in (Figure 4A). A major scale is presented via headphones, with successive tones alternating from ear to ear. The scale is presented simultaneously in both ascending and descending form, such that when a tone from the ascending scale is in the right ear a tone from the descending scale is in the left ear, and vice versa. (Figure 4B) displays the same pattern, showing how it is composed of ascending and descending scales. This has the consequence that the right ear receives one disjunct sequence of pitches, while the left ear simultaneosly receives another disjunct, and overlapping, sequence of pitches. This pattern is played continuously without pause.

This pattern produces a number of illusions, which differ strikingly across listeners. The illusion that is most commonly obtained is illustrated in (Figure 4C). It appears that a melody corresponding to the higher tones is coming from the right earphone, and a melody corresponding to the lower tones from the left one. When the earphone positions are reversed the apparent locations of the higher and lower tones remain fixed. This gives rise to the bewildering impression that the earphone that had been producing the higher tones is now producing the lower tones, and that the earphone that had been producing the lower tones is now producing the higher tones!

Other listeners obtain different illusory percepts. Some hear the higher tones as coming from the left earphone and the lower tones as from the right one, with earphones placed both ways. For yet other listeners, when the earphone positions are reversed the apparent locations of the higher and lower tones reverse also. Other listeners perceive only a single stream of tones, which corresponds to the higher tones and not the lower ones. This ‘single stream’ percept can involve a number of different localization patterns.

Interestingly, righthanders and lefthanders differ statistically in the way they experience the scale illusion: Righthanders are very likely to hear the higher tones on the right and the lower tones on the left; however lefthanders as a group do not show the same tendency. This pattern of results indicates that listeners tend to project the higher tones onto the dominant side of space, and the lower tones onto the nondominant side. One may conjecture that this reflects greater activity in the dominant hemisphere on the part of units underlying the higher tones, and greater activity in the nondominant hemisphere on the part of units underlying the lower tones. Support for this conjecture comes from clinical studies showing that patients who experience palinacousis tend to perceive the illusory sound as on the side of auditory space that is contralateral to the lesion (25). As further evidence, patients who obtain auditory sensations on stimulation of the temporal lobe generally refer these sensations to contralateral auditory space (26). Further supporting evidence comes from findings linking activity in one hemisphere to the perception of sounds in contralateral auditory space (27, 28).

We can then enquire why such illusory conjunctions occur. When presented with a complex sound configuration, the auditory system engages in a process of inference concerning the sources that produced it. In the real world, similar sounds are likely to be coming from the same source and dissimilar sounds from different sources. So the best interpretation of this implausible scale pattern is that one source is producing the higher tones and another

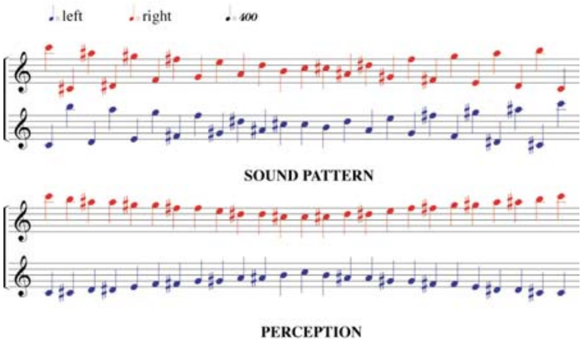

Figure 5. The sound pattern that gives rise to the chromatic illusion, and a percept frequently obtained by righthanders. When the pattern is presented via earphones, most righthanders perceive a melody in their right ear that is composed of the higher tones,and a melody in their left ear that is composed of the lower tones.

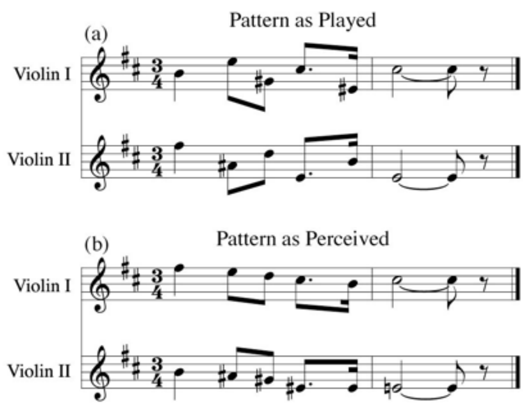

Figure 6. Beginning of the last movement of Tchaikovsky's Sixth Symphony (*The Pathetique*). The upper part of the Figure shows the pattern as it is played, and the lower part as it is generally perceived.

source the lower ones. The power of unconscious inference is here so strong as to override low-level localization cues, so that we mislocalize the tones in accordance with this interpretation.

Variations of the scale illusion can easily be produced. (Figure 5) shows, as an example, a two-octave chromatic scale, with components alternating from ear to ear in the same fashion. When each channel is played separately, jagged melodic lines are perceived. However, when the channels are played together, two smooth melodies emerge, organized in accordance with pitch proximity.

We can then ask: Is the scale illusion simply a laboratory curiosity, in which the brain is made to come to the wrong conclusions under very unusual circumstances, or do such effects also occur when listening to music in the real world? It is surprising how strongly the scale illusion emerges when the sounds are produced by live instruments in concert halls. Recently, during a lecture at delivered by the author at the University of California, Irvine, the scale illusion was played with three violinists on the extreme left of the stage, and three on the extreme right. People in the audience experienced the illusion very clearly, even those who were sitting well to the side.

Perceptual reorganizations such as occur in the scale illusion can even be found in the standard music repertoire. For example, at the beginning of the second movement of Tchaikovsky's Sixth Symphony – *The Pathetique* – there is a passage in which the theme and accompaniment waft back and forth between the two violin parts. However the theme is heard as coming from one set of violins, and the accompaniment as from the other, as shown in (Figure 6). This occurs even when the orchestra is arranged in 19th century fashion, so that the first violins are on one side of the stage and the second violins on the other side. It is unknown whether Tchaikovsky realized that this passage gave rise to an illusion, or whether he expected the notes from the theme and accompaniment to appear to waft back and forth across the stage. However, the conductor Arthur Nikisch disagreed strongly with Tchaikovsky's scoring (though the reasons why he disagreed are unknown), and he performed the passage rescored so that the notes from the melody all came from one set of instruments, and the notes from the accompaniment from the other.

4.2. The glissando illusion

The glissando illusion (29, 30) is best experienced when the listener is seated in front of two loudspeakers, with one to his right and the other to his left. The pattern that produces this illusion consists of two components: a synthesized oboe tone of constant pitch, and a sine wave whose pitch glides up and down. These two components are presented simultaneously via the two loudspeakers, and alternate between the speakers such that when a portion of the glissando is coming from the speaker on the right the oboe tone is coming from the speaker from the left; and vice versa.

When each channel is presented separately, the listener correctly hears the oboe tone alternating with portions of the glissando. Yet when the two channels are presented together, a curious illusion occurs: The oboe tone continues to be heard correctly (i.e., as alternating between the loudspeakers). However, the portions of the glissando appear to be joined together quite seamlessly, so that a single, continuous tone is heard that appears to travel slowly around in space in accordance with its pitch motion. When the switching rate between the loudspeakers speeds up and slows down, the apparent speed at which the glissando appears to travel through space does not change, but remains tied to its pitch motion.

As with the scale illusion, handedness correlates appear in terms of the apparent spatial locations of the high and low portions of the glissando: Righthanders tend to hear it as traveling from left to right as its pitch moves from low to high, and nonrighthanders are more likely to experience a variety of localization percepts. Again, this pattern with respect to handedness can be hypothesized to

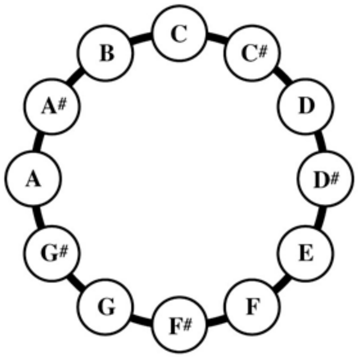

Figure 7. The pitch class circle.

reflect a tendency in some situations to attribute higher sounds to the dominant side of space and lower sounds to the nondominant side (25 - 28).

In considering why the glissando illusion occurs, we can make reference to the same principles as were discussed above for the scale illusion. In the present case, the listener joins together portions of the glissando that are proximal in pitch, so that a single smooth pitch trajectory is perceived. In real world situations, it is very unlikely that a sound which is changing smoothly in pitch should be alternating abruptly between two spatial locations, so we infer that this sound must be being generated by a single source, and so we hear it as spatially continuous.

4.3. Grouping of nonsimultaneous sounds from different spatial locations

So far we have been considering cases in which the sounds arising from different spatial locations are simultaneous. What happens when onset and offset disparities are introduced? In one experiment (31) listeners were presented with either of two rapidly repeating melodies via earphones, and they were asked to identify on each trial which melody was presented. This task was trivial for the listeners when all the tones from the melody were presented to both ears simultaneously. However, when the melody was instead presented with its component tones distributed between ears, the listeners were unable to integrate the tones into a single perceptual stream, and as a result performed very poorly. In another condition, a contralateral drone accompanied the melody, such that when a tone from the melody was in one ear the drone was in the other ear. In this situation the two ears again received input simultaneously, and the performance level was again very high. However, when the drone was instead presented to the same ear as the tone from the melody, so that the two ears no longer received input simultaneously, identification was again very poor.

This study, together with the scale illusion and its variants, shows that when tones emanate from different spatial locations, timing factors have a profound influence on how they are perceptually grouped together. When tones arrive from the two locations simultaneously, melodic configurations are formed on the basis of pitch proximity, and the tones can even be mislocalized on this basis. However, when the tones arriving from the two locations are clearly separated in time, grouping by spatial location wins out instead, and can be so powerful as to prevent the listener from integrating them into a single perceptual stream.

It is interesting that the composer Berlioz (32) came to a related conclusion when he wrote:

I want to mention the importance of the different points of origin of the tonal masses. Certain groups of an orchestra are selected by the composer to question and answer each other; but this design becomes clear and effective only if the groups which are to carry on the dialogue are placed at a sufficient distance from each other. The composer must therefore indicate on his score their exact disposition. For instance, the drums, bass drums, cymbals, and kettledrums may remain together if they are employed, as usual, to strike certain rhythms simultaneously. But if they execute an interlocutory rhythm, one fragment of which is given to the bass drums and cymbals, the other to kettledrums and drums, the effect would be greatly improved and intensified by placing the two groups of percussion instruments at the opposite ends of the orchestra, that is, at a considerable distance from each other.

5. AMBIGUITIES OF MUSICAL PITCH

Musicians have long acknowledged that there is a certain perceptual equivalence between tones which are related by octaves; i.e., whose fundamental frequencies stand in the ratio of 2: 1. This equivalence is built into the system of notation for the Western musical scale. The core of this scale is composed of twelve tones, formed by the division of the octave into semitones, and each tone is assigned a name: C, C#, D, D#, E, F, F#, G, G#, A, A#, and B. The full scale, as it ascends in height, is formed by repeating this series of note names across octaves. Since all Cs sound in a sense equivalent, as do all C#s. all Ds, and so on, pitch can be described as varying both along a monotonic continuum of height and also along a circular dimension of pitch class, (corresponding to note name, as shown in (Figure 7).

When a presented with a complex tone that is comprised of many adjacent harmonics, the listener perceives a pitch that corresponds to the fundamental frequency, so that the tone is clearly defined both in terms of pitch class and also in terms of height. However, with tones played by natural instruments, in which the relative amplitudes of the different harmonics vary, ambiguities of pitch height can arise. Furthermore, when instrument tones are played together in octave relation, the perceived height of the resultant complex differs from that of either tone played alone.

5.1. Pitch circularity effects.

Such ambiguities of height can have interesting consequences for the perception of musical relationships. In one experiment (33), a series of tones was generated, with each tone consisting of 10 components that were separated by octaves. The amplitudes of the components were determined by a fixed, bell-shaped spectral envelope, so that those in the middle of the musical range were highest, and those at the extremes were lowest. The pitch classes of the tones were then varied by shifting their components up and down in log frequency. Since the tones consisted only of harmonics that stood in octave relation, the remaining harmonics which were needed to define their fundamental frequencies were missing. In consequence, these tones were clearly defined in terms of pitch class, but were in principle ambiguous in terms of height.

It was found that when two such complex tones were played in succession, listeners heard either an ascending pattern or a descending one, depending on which was the shorter distance between them along the pitch class circle. So, for example, the tone pair C#-D was always heard as ascending, since the shorter distance between them was clockwise. Analogously, the tone pair G-F# was always heard as descending, since the shorter distance between them was counter-clockwise. So under conditions of height ambiguity, the principle of proximity again emerges to define perceived relations between temporally adjacent tones.

This effect was then exploited to produce a compelling illusion. When a sequence of such tones was played which repeatedly traversed the pitch class circle in clockwise direction (C – C# - D – D#, and so on) listeners perceived a sequence that appeared to ascend endlessly in pitch. When, instead, the tones traversed the circle in counter-clockwise direction (C – B – A# - A, and so on) the sequence appeared to descend endlessly. Similarly striking circularity effects have also been produced when, instead of moving in stepwise direction, tones were made to glide clockwise or counterclockwise around the pitch class circle (34).

Effects which approached circularity have been generated by composers for centuries, by means of employing tone complexes standing in octave relation. Examples can be found in the works of Bach, Scarlatti, Haydn, and Beethoven, and particularly in works by twentieth century composers such as Stockhausen Berg, Ligeti and Risset (35). In one famous example, Risset used an endlessly descending glide in his incidental music to Pierre Halet's play 'Little Boy', in which the gliding tone symbolized the falling of the atomic bomb over Hiroshima.

Ambiguities of perceived height that give rise to pitch circularity effects have also been produced by varying the relative amplitudes of the odd and even harmonics of a complex tone. Since tones comprising full harmonic series are more naturalistic than are tones consisting only of components that stand in octave relation, such effects have implications for musical composition, since it should be possible to create variants that sound like instrument tones.

If one takes a harmonic complex tone, and gradually reduces the amplitudes of the odd-numbered harmonics, keeping the even-numbered ones constant, the tone appears to increase smoothly in height, while remaining in the same pitch class – that is, the tone appears to move gradually up an octave, without traversing the semitone scale. This effect has been exploited so as to produce banks of circular tones (36, 37). We begin with a bank of 12 tones, each of which consists of the first eight harmonics, and the fundamental frequencies of these tones range in semitone steps over an octave. For the tone with the highest fundamental, the odd and even harmonics are identical in amplitude. Then one moves down a semitone, and lowers the amplitude of the odd-numbered harmonics, so raising the perceived height of the tone. Then one moves down another semitone, and further lowers the amplitude of the odd-numbered harmonics, so further increasing the perceived height of the tone. One continues this way until, for the lowest fundamental frequency, the odd-numbered harmonics no longer contribute to the perceived height of the tone. Employing a bank of tones created in this fashion, convincing circularity effects have been obtained. When such tones are played in ascending or descending semitone steps, convincing impressions of ever-ascending and ever-descending scales are produced.

5.2. The tritone paradox

We next consider the situation in which listeners are presented with pairs of octave-related complexes, such that proximity cannot be invoked in making judgments of relative height. For example, we can consider what happens when listeners are presented with two tones in succession that are in opposite positions along the pitch class circle, such as C followed by F#, or G# followed by D. Such tone pairs comprise an interval of six semitones, known as a tritone.

When listeners are presented with such two-tone patterns, a surprising illusion emerges, and there are striking differences between listeners in how this illusion is perceived. For example, when C is played followed by F#, some listeners clearly hear an ascending pattern, while other listeners clearly hear a descending one. Yet when a different pair of such tones is played, say G# followed by D, the first group of listeners now hear a descending pattern, while the second group now hear an ascending one. In addition, for any one listener, the pitch classes arrange themselves perceptually with respect to perceived height in an orderly way: Tones in one region of the pitch class circle are heard as higher, and those in the opposite region as lower. However, listeners can differ completely in terms of which region of the pitch class circle is heard as higher and which as lower. This finding is illustrated in (Figure 8), which reproduces the judgments of four different listeners. As a result of this illusion, extended patterns formed by such tone pairs are heard by different listeners in entirely different ways (38-42).

What can be the reason for this strange effect, and for the individual differences in the way it occurs? Based on a number of informal observations, it was conjectured that it might be related to the processing of speech sounds. More specifically, it was hypothesized that each person possesses a mental representation of the pitch class circle, which is oriented in a particular direction with respect to height. The way the circle is oriented is derived from the speech patterns to which the listener has most

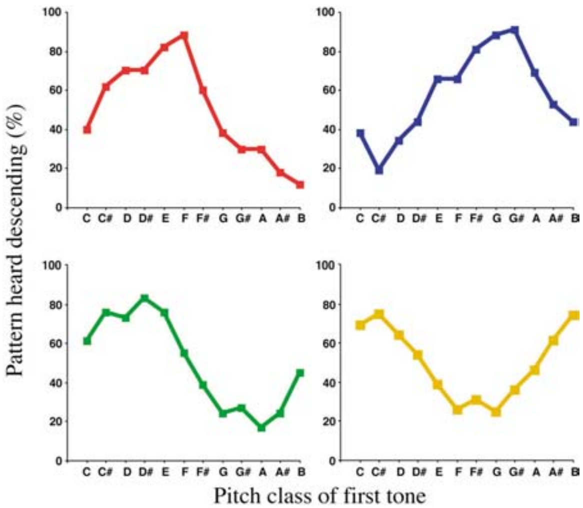

Figure 8. Judgments of the tritone paradox made by four different listeners. Each graph plots the percentages of judgments that a tone pair formed a descending pattern as a function of the pitch class of the first tone of the pair. The data from all subjects showed orderly relationships to the positions of the tones along the pitch class circle; however the direction of this relationship varied considerably across subjects.

frequently been exposed. This mental representation then determines both the pitch range of the person's speaking voice, and also how he or she hears the tritone paradox. In an experiment to test this hypothesis, a correlation was indeed found between the way listeners heard the tritone paradox and the pitch ranges of their speaking voices (43).

It was further found that the way the tritone paradox is perceived varies in correlation with the language or dialect to which the listener has been exposed. One study compared perception of this pattern among English-speaking listeners who had grown up in California, with those who had grown up in the south of England: In general, when the Californians tended to hear a pattern as ascending the listeners from the south of England tended to hear it as descending, and vice versa (44).

Further research indicated that the perceptual representation that influences perception of the tritone paradox is formed early in life. In one study, perception of this pattern was studied in mothers and their children. The children were all Californian; however their mothers had grown up in widely different geographical regions. As expected, the mothers perceived the pattern in ways that differed considerably from each other; however, although the children were all Californian, their percepts were closely similar to those of their mothers (45). This finding is as expected from the hypothesis that perception of this pattern is influenced by the pitch ranges of voices to which the listener has most frequently been exposed.

In another study, perception of this illusion was examined in subjects who were born in Vietnam, together with a group of English-speaking Californians. The Vietnamese group was divided into two subgroups: those who had arrived in the U.S. as adults, and spoke fluent Vietnamese though little English, and those who had arrived in the U.S. as infants or young children, and now spoke perfect English but were not necessarily fluent in Vietnamese. It was found that the two Vietnamese groups perceived the tritone paradox in very similar ways, and the judgments of both groups differed clearly from those of the English speaking Californians (46). In this study it was also found, taking subjects who had grown up in Vietnam, that the way they perceived the tritone paradox correlated with the pitch range of their speaking voices, sin the same way as had been found for speakers of English (46). So this study indicates that exposure to speech sounds during an early critical period can substantially influence how the tritone paradox is perceived.

Can we generalize these findings to live musical performances? In an informal experiment the author found that the tritone paradox was also produced when the components of each octave related complex were themselves complex tones, such as sawtooth waves. So in principle such effects should be obtained with groups of instruments playing together, with their tones standing in octave relationship. Indeed, on listening to certain types of music, such as Debussy's orchestral pieces, it is often unclear which tones are perceived as higher and which as lower. For example, in Debussy's *Sirenes*, the soprano voices rise and fall against an orchestral background. When listening carefully to this piece, it becomes clear that there are times when the apparent heights of the different instrument tones and voices are quite ambiguous. It is in such situations as these that we might expect to find perceptual disagreements as reflected in the tritone paradox.

6. CONCLUSION

This chapter has been concerned with two issues. The findings explored in the first part of this chapter shed light on various rules of thumb that have evolved in traditional Western music. For example, the rule that forbids the crossing of voices in counterpoint can be understood in light of the experimental evidence showing that the auditory system forms perceptual streams based on pitch range; violating this rule could cause the listener to perceive the music in ways that are contrary to the composer's intentions. A similar argument applies to the 'law of stepwise progression', which states that melodic patterns should be created from small melodic intervals rather than large ones.

The findings explored in the second part of the article, for example concerning the scale illusion and the tritone paradox, show that surprising misperceptions of even very simple musical patterns can occur, and that there can be striking individual differences in how such patterns are perceived. Musical discourse in the real world is quite nonspecific, so it is not surprising that such misperceptions, together with the individual differences associated with them, were first discovered in the laboratory. Finally we can conjecture that, in the real world, such perceptual disagreements might form the bases of arguments concerning musical compositions and performances that have so far been considered aesthetic in nature.

7. REFERENCES

1. Mathews, M. V. & Pierce, J. R.: Harmony and nonharmonic partials*.* *J Acoust Soc Am* 68, 1252-1257 (1980)

2. Moore, B. C. J., Glasberg, B. R., & Peters, R. W.: Thresholds for hearing mistuned partials as separate tones in harmonic complexes. *J Acoust Soc. Am* 80, 479-483 (1986)

3. Rasch, R. A.: The perception of simultaneous notes such as in polyphonic music. *Acustica* 1978, 40, 1-72 (1978)

4. Huron, D.: Tonal consonance versus tonal fusion in polyphonic sonorities. *Mus Percept* 9, 135-154 (1991)

5. Erickson, R.: *Sound structure in music*. University of California Press, Berkeley (1975)

6. Darwin, C. J. & Ciocca, V.: Grouping in pitch perception: effects of onset asynchrony and ear of presentation of a mistuned component. *J Acoust Soc Am* 91 3381-3390. (1992)

7. Rasch, R. A.: Timing and synchronization in ensemble performance. In J. A. Sloboda (Ed.) *Generative processes in music: The psychology of performance, improvisation, and composition.* Oxford University Press, Oxford (1988)

8. Huron, D.: Note-onset asynchrony in J. S. Bach's two-part inventions. *Mus Percept* 10, 435-444 (1993)

9. Grey, J. M. & Moorer, J.: A. Perceptual evaluation of synthesized musical instrument tones. *J Acoust Soc Am* 62, 454-462 (1977)

10. Bregman, A. S.: *Auditory scene analysis: The perceptual organization of sound.* MIT Press, Cambridge (1990)

11. Carlyon, R. P.: Discriminating between coherent and incoherent frequency modulation of complex tones. *J Acoust Soc Am* 89, 329-340 (1991)

12. Carlyon, R. P.: The psychophysics of concurrent sound segregation. *Phil Trans Royal Soc London, Series B* 336, 347-355 (1992)

13. Warren, R. M.: Perceptual restoration of obliterated sounds. *Psych Bull* 96, 371-383 (1984)

14. Sasaki, T.: Sound restoration and temporal localization of noise in speech and music sounds. *Tohuku Pschologica Folia* 39, 79-88 (1980)

15. Dowling, W. J.: The perception of interleaved melodies. *Cog Psych* 5, 322-337 (1973)

16. Bregman, A. S., & Campbell, J.: Primary auditory stream segregation and perception of order in rapid sequences of tones. *J Exp Psychol* 89, 244-249 (1971)

17. Divenyi, P. L., & Hirsh, I. J.: Discrimination of the silent gap in two-tone sequences of different frequencies. *J. Acoust Soc Am* 52, 166S (1972)

18. Van Noorden, L. P. A. S.: *Temporal Coherence in the Perception of Tone Sequences*. Unpublished doctoral dissertation. Technische Hogeschoel Eindhoven, The Netherlands (1975)

19. Deutsch, D.: The processing of structured and unstructured tonal sequences. *Percept Psychophys* 28, 381–389 (1981)

20. Peretz, I. and Zatorre, R.: *The cognitive neuroscience of music*. Oxford: Oxford University Press, Oxford (2003)

21. Rauschecker, J. P., & Tian, B.: Mechanisms and streams or processing "what" and "where" in the auditory cortex. *Proce Nat Acad Sci* 97, 1180-1186 (2000)

22. Stewart L., von Kriegstein, K, Warren, J. D., & Griffiths. T. D.: Music and the brain: disorders of musical listening*.* *Brain* 129; 2533-53 (2006)

23. Deutsch, D.: Two-channel listening to musical scales. *J Acoust Soc Am* 57, 1156-1160. (1975a)

24. Deutsch, D.: Musical illusions. *Sci Am* 233, 92-l04 (1975b).

25. Jacobs, L., Feldman, M., Diamond, S. P., & Bender, M. B.: Palinacousis: Persistent or recurring auditory sensations. *Cortex* 9, 275-287 (1973)

26. Penfield, W., & Perot, P.: The brain's record of auditory and visual experience. *Brain* 86, 595-696 (1963)

27. Woldorff, M. G., Templemann, C., Fell, J., Tegeler, C., Gascher-Markefski, B., Hinrichs, H., Heinze, H-J., Scheich, H.: Lateralized auditory spatial perception and the contralaterality of cortical processing as studied with functional magnetic resonance imaging and magnetoencephalography. *Hum Brain Map* 7, 49-66 (1999).

28. Pavani, F., Macaluso, E., Warren, J. D., Driver, J., & Griffiths, T. D.: A common cortical substrate activated by horizontal and vertical sound movement in the human brain. *Curr Biol*, 12, 1584-1590 (2002)

29. Deutsch, D.: Grouping mechanisms in music. In D. Deutsch (Ed.) *The psychology of music,* 2nd ed., Academic Press, San Diego 299-348 (1999)

30. Deutsch, D., Hamaoui, K., & Henthorn, T.: The Glissando Illusion: A Spatial Illusory Contour in Hearing. *J Acoust Soc Am* 117, 2476 (2005)

31. Deutsch, D.: Binaural integration of melodic patterns. *Percept Psychophys* 25, 399–405 (1979)

32. Berlioz, H.: *Treatise on instrumentation*. R. Strauss (Ed.) Kalmus, New York (1948)

33. Shepard, R. N.: Circularity in judgments of relative pitch. *J Acoust Soc Am* 36, 2345-2353 (1964)

34. Risset, J. C.: Paradoxes de hauteur: Le concept de hauteur sonore n'est pas le meme pour tout le monde. *Proc Seventh Internat Congress Acoust* Budapest, 20S (10) 613-616 (1971)

35. Braus, I.: Retracing one's steps: An overview of pitch circularity and Shepard tones in European music, 1550-1990. *Mus Percept* 12, 323-351 (1995)

36. Deutsch, D., Dooley, K., Dubnov S., Henthorn, T, & Warden, A.: pitch circularity produced by varying the amplitudes of odd and even harmonics. *J Acoust Soc Am* 118, 1949 (2005)

37. Deutsch, D., Dooley, K., & Henthorn, T.: A new algorithm for pitch circularity. (in preparation)

38. Deutsch, D.: A musical paradox. *Mus Percept* 3, 275-280 (1986)

39. Deutsch, D.: Some new sound paradoxes and their implications. In *Auditory Processing of Complex Sounds. Phil Trans of Royal Soc, Series B*, 336, 391-397 (1992a).

40. Deutsch, D.: Paradoxes of musical pitch. *Sci Am* 267, 88-95 (1992b)

41. Deutsch, D. Kuyper, W. L. & Fisher, Y.: The tritone paradox: its presence and form of distribution in a general population. *Mus Percept* 5, 79-92 (1987)

42. Deutsch, D.: The tritone paradox: Effects of spectral variables. *Percept Psychophys* 42, 563-575 (1987)

43. Deutsch, D., North, T. & Ray, L.: The tritone paradox: Correlate with the listener's vocal range for speech. *Mus Percept* 7, 371-384 (1990)

44. Deutsch, D.: The tritone paradox: An influence of language on music perception. *Mus Percept* 8, 335-347 (1991)

45. Deutsch, D.: Mothers and their children hear a musical illusion in strikingly similar ways. *J Acoust Soc Am* 99, 2482 (1996)

46. Deutsch, D., Henthorn, T., & Dolson, M.: Speech patterns heard early in life influence later perception of the tritone paradox.. *Mus. Percept* 21, 357-372 (2004)

Key Words: Music, Perception, Literature Review, Perceptual Grouping, Illusion, Pitch, Sound

Send correspondence to: Dr Diana Deutsch, Department of Psychology, University of California, San Diego, La Jolla, CA 92093, Tel: 858-453-1558, Fax: 858-453-4763, E-mail: ddeutsch@ucsd.edu

http://www.bioscience.org/current/vol12.htm