13 AUDITORY CONFLICTS AND ILLUSIONS

Angélique A. Scharine
Tomasz R. Letowski

Perceptual Conflicts and Illusions

Perceptual conflicts appear when the brain receives ambiguous information and needs to choose which of the conflicting pieces of information represents the actual stimulation. In some cases, one piece of information dominates the other ones and we are largely unaware of the conflicting information. In other cases, however, all pieces of information are perceptually equivalent and the brain switches spontaneously between alternate interpretations of the received stimulation. If the brain receives several conflicting but perceptually equivalent pieces of information such a phenomenon is also called perceptual multistability and the acting stimulation is called multistable stimulation (Leopold and Logithesis, 1999). In case of two conflicting pieces of information, the phenomenon is called perceptual bistability and corresponding stimulation is called bistable stimulation (Hupe, Joffo and Pressnitzer, 2008).

Multistable stimulation has been extensively studied in the visual domain (e.g., moving plaids, binocular rivalry) but can also occur in the auditory modality. The most common form of multistable auditory perception is auditory grouping of incoming acoustic information that may change depending on the listener's focus on specific temporal events (Bregman, 1990; Hupe, Joffo and Pressnitzer, 2008; Pressnitzer and Hupe, 2006; Van Noorden, 1975).

The grouping of arriving sounds into perceptual events is referred to as *auditory streaming*. Normally, the grouping forms a uniform unequivocal stream. However, bistable streaming may occur under certain conditions. For example, an alternating sequence of high- and low-frequency tones may be perceived as one or two streams. When the tones are similar in frequency and the rate of presentation is slow, the listener hears a single coherent series of tones alternating in time. However, when the difference in frequency is large and the repetition rate is fast the alternating sequence splits perceptually in two unrelated streams of high- and low- frequency tones. Between these two extreme conditions, the listeners may hear either phenomenon by paying attention to different properties of the sequence. For example, Van Noorden (1975) presented the listeners with two tones A and B forming a sequence ABA-ABA-…ABA. Depending on the frequencies of the tones and the duration of the pause between individual ABA chunks, the listeners perceived either a single stream of information in a form of "a gallop" or two parallel streams of high- and low-frequency tones. However, within a certain range of manipulated parameters the listeners could hear either of these two phenomena just by refocusing their attention. Similar effects can be observed when the same tone with alternating intensities is presented to the listener (Van Noorden, 1975).

Perceptual conflicts may also have a multisensory form when information received through two or more senses lacks congruency. For example, several authors reported perceptual conflicts in simultaneous perception of conflicting visual and tactile cues (Adams and Duda, 1986; Heller et al., 1999; Hershberger and Misceo, 1996), and visual and auditory cues (Hupe, Joffo and Pressnitzer, 2008).

Another group of perceptual effects that are not directly dependent on the presence of external stimulation are perceptual illusions. Perceptual illusions are the instances where the cues that the brain relies on to provide specific information about sensory stimulation are poorly correlated with actual physical stimulation. They are not the instances when two incongruent pieces of information compete for our attention but rather the instances where our brain reports a stable and repeatable awareness of stimulation that cannot be directly explained by the physical properties of the acting stimuli. They are distortions of reality that are typically shared by many people (Solso, 2001). They can be explained by various physiological processes but they do not refer directly to the stimulation received and therefore are called illusions.

Perceptual illusions should not be confused with the masking phenomena described in Chapter 11, *Auditory Perception and Cognitive Performance*, where the part of the stimulus is not perceived due to the shadowing effect (masking effects) of the other parts of the stimulus. Masking does not result in a qualitatively new percept but only causes a partial awareness of the stimulus. Illusions are also different from auditory images or hallucinations, which are the sensations created in the absence of stimulus. For example, composers report "hearing a tune" in their head before writing a new piece. Hallucinations usually have a pathological basis but they may also occur occasionally in the real world when a highly expected event does not happen.

Auditory illusions are quite common in perception of music and speech due to our brain's tendency to fill the unexpected gaps in incoming streams of events by a reasonable prediction of what should be there. There are also some between-channel associations that may create illusions or hallucinations of the presence of specific acoustic stimulation that does not take place. For example, seeing lip movement in a noisy environment where no speech is present may result in the illusion of hearing speech. Another example of an auditory illusion is the McGurk effect, described in Chapter 14, *Auditory-Visual Interactions*, where seeing the lips pronouncing sound "ga" and hearing sound "ba" results in illusion of hearing the sound "da" (McGurk and McDonald, 1976). In this case a sound is present but it is heard as a different one.

The initial part of this chapter discusses of the processing of information by the auditory channel and the potential conflict in information reception. The second part describes common auditory conflicts and illusions and their physiological basis. Auditory-visual interactions and related conflicts, are discussed in Chapter 14, *Auditory-Visual Interactions,* and auditory-tactile interactions observed during tactile stimulation of the head are described in Chapter 18, *Exploring the Tactile Modality for HMDs*. Chapter 14 also contains a discussion of practical strategies intended to reduce auditory and auditory-visual conflicts and cognitive overload by proper design of auditory signals (earcons and warning signals) so that they are easily understood and complied with during times of stress and fatigue.

Auditory Scene Analysis

Auditory scene analysis (ASA) is the term coined by a Canadian psychologist Albert Bregman (1990) to describe a variety of processes by which the brain parses the sound arriving at the ear into its various components and groups them together into meaningful events. Each of our ears receives only a single sound pressure wave and this wave consists of the combination of all sounds occurring in the environment. The fact that we have two ears is critical for spatial perception and auditory orientation in space but each of the ears receives a relatively complete collection of auditory events emanating from an infinite variety of sound sources. The brain has the task of analyzing the complex waveform that arrives at the ears into its components and then assigning those components to the auditory events and sound sources creating those components. Some authors divide the ASA tasks into the simultaneous grouping (frequency grouping) and sequential grouping (stream grouping) tasks but in most practical situations both tasks are performed concurrently and aid each other (Plack, 2005).

ASA resembles the task of visual scene analysis performed by the sense of vision. Rather than view the world as a hodgepodge of colors and lights, our visual system follows a set of rules to determine which visual components belong to which visual object. It does this by utilizing a number of Gestalt cues. "Gestalt," a German word for "form," is used to refer to self-organizing principles that form a whole from a collection of features. Some examples of the Gestalt Laws that govern the ways by which auditory and visual objects are formed from their specific features are listed in Table 13-1. An example demonstrating how the Laws of Good Continuation, Simplicity, and Closure affect our visual perception is shown in Figure 13-1. The Laws of Good Continuation and Simplicity suggest that the picture is composed of blocks with simple straight edges rather than a strange object with jagged edges and that the jagged lines are due to a juxtaposition of multiple blocks. Further, the *Law of*

Closure suggests that the edges obscured by other blocks placed in front of them are continuous and form whole objects.

Table 13-1.
A description of auditory and visual Gestalt Laws.

Name	Audition	Vision
Proximity (Belongingness)	Sounds arriving from places *close in space* tend to be grouped	Elements *close together* in space tend to be grouped
Similarity	Sounds with *similar timbre and pitch* tend to be grouped	Elements *shaped alike* tend to be grouped
Good Continuation	Sounds that follow a *regular pitch contour* tend to be grouped	Elements that follow a *regular spatial contour* tend to be grouped
Closure	Interrupted auditory stimuli tend to be perceived as *continuous* when plausible	*Borders are interpreted/completed* to specify shapes
Simplicity (Pragnanz)	Frequencies with *simple harmonic ratios* tend to be grouped	*Prototypical* shapes tend to be regular, simple, symmetric
Common Fate	Sounds with *synchronous rhythm patterns* tend to be grouped	Elements that *move together* tend to be grouped

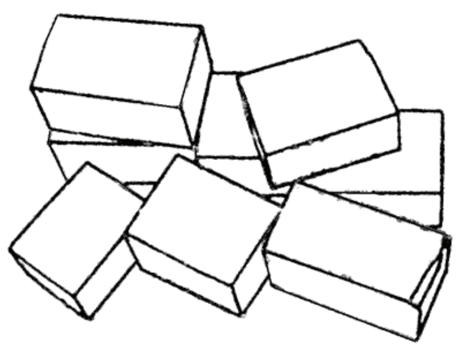

Figure 13-1. Illustration of the effect of three Gestalt Laws -*Law of Good Continuation, Law of Closure* and *Law of Simplicity* - on our perception. We interpret the drawing as a collection of simple blocks rather than a complex collection of random lines.

As shown in Table 13-1 the Gestalt Laws apply not only to vision but also to audition. Their function is to allow the auditory system to group the different features of the complex waveforms arriving at the ears into a plausible set of auditory events. Consider a common auditory environment, that of a kitchen with a radio on and a person cooking dinner while another person sets the table. There will be the sound of the radio, the sound of people talking, the sounds of dishes and pans, and the ambient sounds of the heating or cooling system. These will all arrive at the ears of the listener as two complex waveforms. In order to parse the waveform into the auditory events creating it, the listener will use many of the Gestalt Laws. For example, the *Law of Proximity* (Belongingness) suggests that anything coming from spatial location occupied by the radio is caused by the radio. The *Law of Similarity* suggests that sounds with a similar timbre (Chapter 11, *Auditory Perception and Cognitive Performance*) are caused by the same auditory event, for example, a particular person speaking. The *Law of Closure* allows a person to listen to the talker and understand him, despite occasional masking caused by the radio. The *Law of Simplicity*, also called the *Law of Pragnanz*, which is a German term meaning “good figure,” suggests that we will choose simple plausible interpretations, consistent with our knowledge of the environment. Thus, ringing sounds will be attributed to pans in the kitchen. Different instruments with differently pitched tones playing in a synchronous rhythm will be perceived as a single acoustic event, that is, the band that is playing on the radio because of the *Law of Common Fate*.

The Gestalt rules can be applied in numerous ways, but their primary function is always to help the listener parse the continuous sound wave into the individual sound events that created it. For the most part, this occurs automatically without much effort on the part of the listener and with very few errors. However, in some cases, the sound wave is parsed incorrectly, and sound components fuse together, emerging perceptually as sound objects that are not actually present. When this occurs, they result in auditory illusions. In almost all cases, auditory illusions exemplify a situation in which a Gestalt cue biases the perceptions of the listener. The selected illusions described in this chapter are dependent on the precise coincidence of certain spectral and temporal features; however, they illustrate how the Gestalt Laws work. Auditory events in the real world are somewhat more random, making such illusory events less predictable. However, a good grasp of Gestalt cues will aid in the design of auditory cues and warnings that are easily detected and understood despite masking from noise in the ambient environment. More extensive discussion of the auditory signal design is presented in Chapter 14, *Auditory-Visual Interactions.*

Auditory Conflicts and Illusions

Auditory conflicts and illusions are not clearly differentiated in the literature and sometimes one of these terms is used to describe both classes of phenomena. In some cases it is even difficult to differentiate if a specific perceptual effect should be classified as a conflict or an illusion. Similarly, none of the proposed classification of these effects, such as transmission-based and construction-based effects, is well established and intuitive enough to be included in this chapter. However, all these phenomena are various distorted perceptions of sound pitch, temporal properties, and location of a sound source. Therefore, for the purpose of clarity, they are grouped together in this chapter by the perceptual characteristic that is being distorted; that is, pitch, temporal pattern, and spatial phenomena.

Pitch conflicts and illusions

In general, humans equate pitch with the lowest frequency of a periodic sound (see Chapter 11, *Auditory Perception and Cognitive Performance*). This frequency is called the fundamental frequency of the sound. The physical nature of the object determines the dominant frequency and all the other accompanying frequencies. Most commonly, the dominant frequency is the fundamental frequency and the other frequencies are harmonics that follow a fairly regular pattern of multiples of the fundamental frequency. However, there are also cases where the perceived pitch does not correspond to the lowest frequency of a sound or where higher frequencies do not follow a strict harmonic pattern. In addition, there are some cases where pitch sensation does not follow any of the accepted forms of pitch creation or changes in an unpredictable manner. Such cases are usually referred to as pitch conflicts or pitch illusions.

Recent advances in electronics and sound synthesis make it possible to precisely control the amplitudes and phases of all frequency components of a sound thus allowing us to explore the way pitch is perceived. The first three illusions described in this section, p*eriodicity pitch*, c*ircular pitch* and the t*ritone paradox*, demonstrate the interacting effects of the fundamental and harmonic frequencies on pitch sensation. The illusions that follow, the *octave illusion* and other pitch streaming illusions, illustrate the preference of the auditory system for small intervals and a regular pitch contour. In general, the perceptual system attempts to group sound components into streams that can be easily interpreted and encoded. Therefore, in some cases, rather than parsing the sound wave in a manner consistent with the spatial information arriving to the left and right ear and creating two sound streams having complicated temporal patterns, the sound wave is parsed into two simple patterns neglecting spatial disparity. The conflicting spatial cues are then misperceived to agree with the dominant simplicity of the resulting streams. The next illusions, the s*plit-off illusion* and its derivatives, illustrate the Laws of Good Continuation and Simplicity. The perceptual system attempts to form a simple interpretation of two auditory streams fusing them into a single stream. Sound elements that are inconsistent with this interpretation are either ignored or "split-off" into an extraneous stream. The *Huggins pitch* emerges from a white noise signal because the phase information separates each narrow frequency band from the rest of the signal, essentially causing the listener to perceive a single stream as two separate ones.

Periodicity pitch

Sounds that are composed of several frequency components having a simple harmonic relationship are called tones. Tones that consist of a single frequency are called pure tones and tones that consist of several frequencies are called complex tones (Emanuel and Letowski, 2009). The lowest frequency (F0) of a complex tone is called the fundamental frequency, and the higher frequencies (F1, F2, F3,…) that are integer multiples of the fundamental frequency are called harmonics. The specific harmonic structure of a complex tone gives the tone its characteristic timbre (tone color).

If one presents a complex tone, with a fundamental frequency of F0 and a series of harmonics, the pitch of the tonal complex is normally associated with the pitch of the fundamental frequency F0. Adding or removing harmonics from the complex affects the timbre of the complex, but it does not change its pitch. This will remain true even if several of the first few harmonics are removed and – more remarkably – even if the fundamental frequency is removed from the complex. . Consider, for example, a case shown in Figure 13-2. A complex tone shown in panel (a) consists of a 400 Hertz (Hz) fundamental frequency and its five subsequent harmonics and produces pitch sensation corresponding to 400 Hz frequency. A complex tone in panel (b) does not have a 400 Hz component but maintains the pitch corresponding to 400 Hz frequency. This phenomenon is called *periodicity pitch, residual pitch, or the missing fundamental phenomenon* and is an indication that our auditory system responds to the overall periodicity of the incoming sound wave. The explanation of the periodicity pitch phenomenon lies in the fact that the missing 400 Hz frequency in panel (b) is the highest common denominator of all the frequency components shown in this panel. Thus, both complex waves represented by the spectra shown in panels (a) and (b) have the same basic period of their complex waveforms. The auditory system groups frequency components that share a common mathematical denominator and matches them to a prototypical sound with those components and assigns a pitch value based on that prototype according to the *Law of Simplicity*. This physiological mechanism is supported by an observation that the tonotopic organization of the auditory cortex is based on pitch rather than frequency and, thus, that signal periodicity is transmitted to the brain (Pantev et al., 1989).

The periodicity pitch phenomenon may seem like an artificial type of phenomenon in respect to head-mounted display (HMD) considerations. However, telephone communication is a common example of the occurrence of this phenomenon. Telephones typically transmit frequencies between 300 and 3600 Hz whereas the average fundamental frequencies of male and female voices are 125 and 200 Hz, respectively. This means that they lie below the transmission range of the telephone system. However, one is usually able to correctly identify the gender of the talker on the telephone because the male voice is still perceived as being lower in pitch.

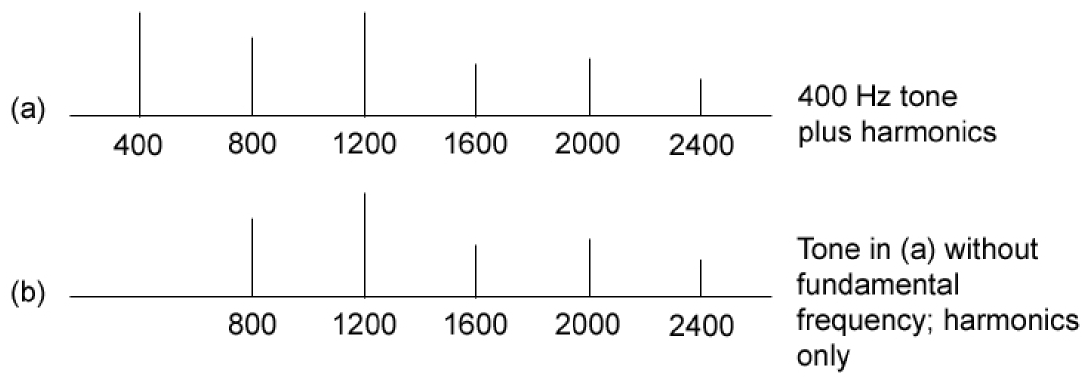

Figure 13-2. Fourier spectra of (a) 400 Hz tone with its lowest five harmonics and (b) the same complex tone without its 400 Hz fundamental. Both complexes produce different timbre sensations but result in the same pitch sensation.

Circular pitch

Pitch is often described as the "highness" or "lowness" of a tone. However, in Western music, pitch relationships are organized into 12-tone sequences defined by the ratio of the tone's fundamental frequency to the scale's tonic scale tone. Each successive octave has a fundamental frequency that is double that of the previous octave. Thus, although the frequency and the associated pitch rise linearly, the musical pitch scale is circular. This concept is shown in Figure 13-3. Such dualism led to two components of pitch: pitch height associated with frequency and pitch chroma associated with music intervals. Pitch height constitutes a basis for arranging the sounds into streams associated with specific sound sources while pitch chroma constitutes a basis for arranging the sounds into specific acoustic patterns (melodies) regardless of the sound source (Rakowski, 1993; Warren et al., 2003).

As noted in the discussion of the periodicity pitch illusion, pitch is not entirely determined by the fundamental frequency, but rather by the relationship of the harmonic frequency components and their fit with a prototypical sound with a certain pitch. Based on this concept, American psychologist Roger Shepard and composer James Tenney developed at Bell Labs a circular set of 12 complex tones, called *Shepard tones* or *Shepard staircase*, which, when played in a continuous loop, make the impression of a indefinitely rising or descending music scale (Shepard, 1964; Tenney, 1969). This phenomenon is frequently called the *circular pitch illusion* or *circular pitch paradox*. The name circular pitch refers to the fact that although the perceived scale progresses continuously in one direction, in fact it is played by the circular repetition of the same 12 tones. It is an auditory analog of the moving barber's pole or *Penrose stairs* illusions in vision (Mussap and Crassini, 1993; Seckel, 2004).

Fundamental frequencies of Shepard tones cover the span of 1 octave and differ from each other by a semitone (~6%). Each complex tone consists of several harmonic components that form a base-2 geometric relationship (1, 2, 4, 8, 16, etc.). Such tones are constructed to have clearly different pitch chroma but to be very similar in pitch height and timbre.

The pitch height ambiguity of Shepard tones is due to the spectral shape of the individual tone complexes. The spectral envelope of all of the complex tones can be described by a single Gaussian function as shown in Figure 13-4. When the Shepard tones are presented serially, the intensities of individual harmonics change slightly so that as the scale ascends, the higher components become less intense while the lower ones become more intense. Thus, at the end of the 12-tone sequence the shifting of intensity weights makes the 13th tone identical to the first one. Since according to the *Law of Proximity*, the perceptual system prefers small intervals, the ear follows the frequency components of successive tones and perceives it as a continually rising pitch sequence. The effect is reversed when the tones are cycled in the opposite direction.

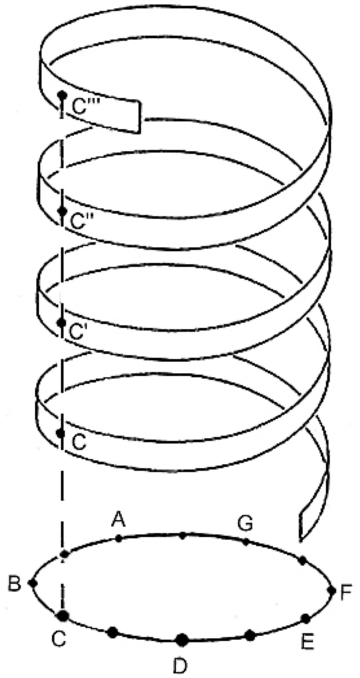

Figure 13-3. Schematic representation of pitch height (vertical axis) and pitch chroma (horizontal axis) (adapted from Shepard, 1982).

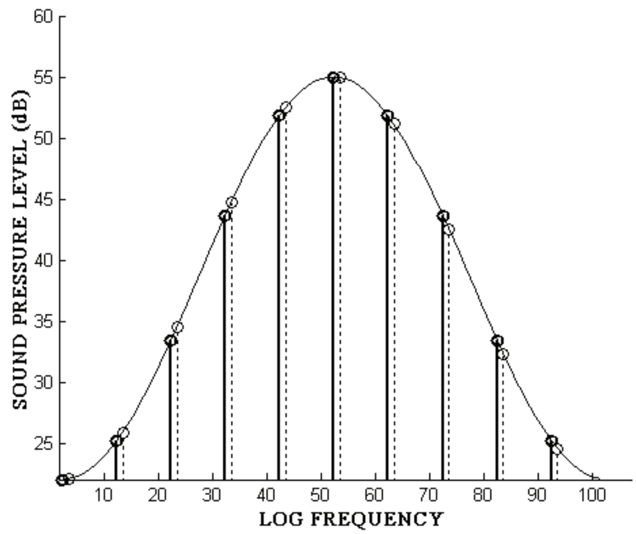

Figure 13-4. Schematic of the Gaussian curve that describes the sound pressure levels of the component frequencies that make up the Shepard tones. The solid lines and dark circles represent the components of 1 tone, the dotted lines and light circles represent the components of the subsequent tone in the scale (adapted from Shepard, 1982).

It is noteworthy that the circular pitch illusion described above had to be intuitively known by music composers prior to the development of Shepard tones. For example, tonal sequences creating this illusion can be found in pieces by Bach and Chopin (Wikipedia, 2008). The illusion is also present in some more modern compositions as well as in the video game Super Mario 64 in association with an infinite staircase.

The circular pitch illusion is best heard when the subsequent tones are presented with short silent intervals between the successive tones. However, it is not limited to discrete steps in frequency or even to octave-based complex tones (Burns, 1981). In 1969, a French composer Jean-Claude Risset working at Bell Labs developed a continuous version of the Circular Pitch illusion known as the Shepard-Risset Glissando or the Continuous Risset Scale (Risset, 1969). The effect is created by 10 harmonically related pure tones that cover the span of nine octaves. All pure tones simultaneously decrease their frequencies in a logarithmic fashion across the span of 10 octaves. The intensity of the overall sound is controlled by a Gaussian function similar to that of the Shepard tones illusion. In addition, the tones differ in the initial phase to maintain the continuity effect.

Jean-Claude Risset is also the author of a circular rhythmic illusion, called the *Risset pattern*, in which the perceived tempo seems to increase or decrease endlessly (Risset 1972; 1986). This illusion is based on the simultaneous presence of several drum beats having simple geometric relations. As the time progresses, the slower beats are made less intense while the faster beats increase in intensity so that the listener gradually changes focus from the slower to the faster beats as they become louder.

The sound effects, based on circular pitch and the Risset pattern, are useful for the simulation of ascent or descent in toys and games. One of the authors remembers having a toy spaceship that played a continuously ascending sound whenever pointed upwards and a continuously descending sound when pointed downwards. A similar sound could easily be used to provide feedback on the directional use of hand-operated controls. It may also have a practical application for warning signal design. In some situations a continuously ascending or descending pitch signal may be needed to force the operator's action. In such cases Shepard tones may be an efficient engineering solution.

The tritone paradox

Despite the circular nature of the Shepard tones, one perceives a directional change because the components are always one semitone apart. So, essentially one is choosing between movement of one semitone in one direction, or 11 semitones in the other direction (Figure 13-5). The simplest interpretation is the shorter distance of one semitone. However, if one hears the first semitone followed by the 6^{th} semitone of the Shepard tones, the movement could be interpreted as going either six semitones up or down. In music, this half-octave interval is called a "tritone" hence the name of this auditory conflict.

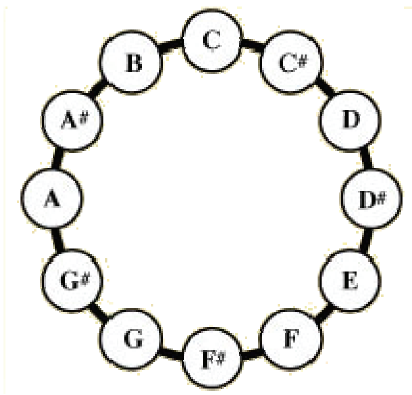

Figure 13-5. Pitch circle. Stepwise changes in the clockwise direction are perceived as ascending. Counterclockwise steps are perceived as descending. The tritone effect occurs when the interval is exactly half of an octave, or halfway around the circle (adapted from Deutsch, 1999).

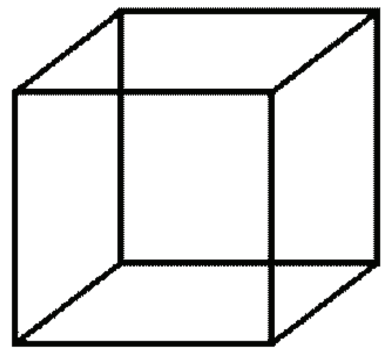

Figure 13-6. Necker cube. Ambiguous line drawing that has two interpretations when perceived as a three-dimensional cube.

When listeners are presented with Shepard tones that are exactly ½ octaves apart, some will perceive the pitch as rising and others as falling. There is usually a listener's bias to hear a particular tritone pair as rising or falling and this remains constant over time. This effect is an auditory analog of the visual conflict represented by the *Necker cube* (Figure 13-6) where a line drawing of a cube can be interpreted as facing either left or right, depending on whether the two intersecting lines on the bottom left are perceived as forming the front or the rear corner of the figure (Marr, 1982). In both cases one can often force oneself to reverse the initial perception.

The octave illusion

In the *octave illusion* (Deutsch, 1974) two pure tones with frequencies an octave apart are presented through the earphones as two tones of equal amplitude alternating between the ears. As a result the listener is presented with a two-tone continuous complex sound with both tonal components switched repeatedly between ears. This sequence of events is shown in Figure 13-7. Deutsch confirmed through several experimental studies that the presented pattern of events is never heard as such. The resulting perceptual effect varies greatly with the listener. Most people hear a complex tone that changes in pitch from high to low as it switches back and forth between the ears so that the high pitched tone is heard in one ear and the low pitched tone in the other ear. Other people hear one tone (low or high) pulsing in both ears accompanied by alternating pulses of the other tone in one ear. Some people reported changes in pitch or in the speed of ear alternations during signal presentation. Still others hear more complex percepts. Reversing of the earphones has no effect on the laterality of perception. What was heard in the left ear remained the same after earphone reversal.

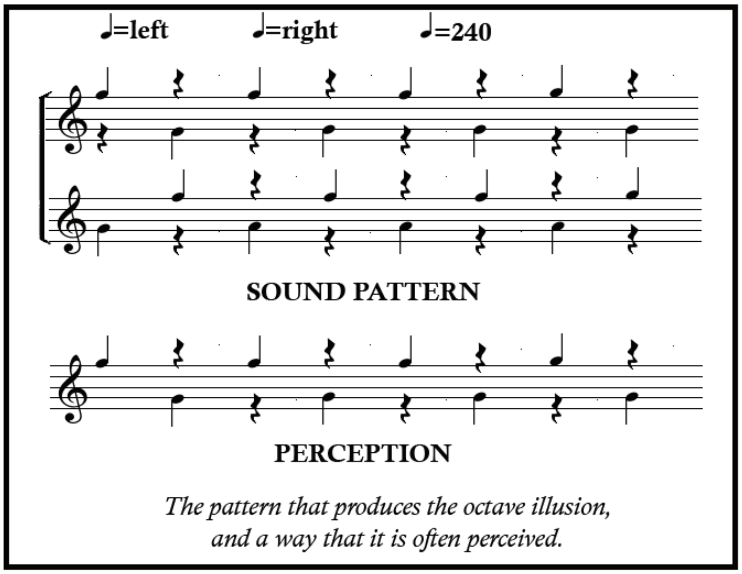

Figure 13-7. Octave illusion. The sequence of tones presented to the left and right ear (upper panel) and the typically perceived sequence of events (lower panel) (Deutsch, 1974).

How is this explained? There are two Gestalt Laws in effect here. The *Law of Similarity* suggests that the same pitched tones should be grouped together as coming from the same sound source. Conversely, the *Law of Proximity* suggests that the sounds occurring in each ear should be grouped together as coming from the same location. It is implausible that a single sound source would be moving from left to right and back again at a high rate of speed. Thus, the perceptual interpretations based on the above two laws conflict with each other and individual perceptions vary, even over time. Therefore, despite the fact that the described phenomenon is called the octave illusion it could also be called the octave conflict. Regardless of its name, there are two notable consistencies in the perceptions reported. First, right handed listeners tend to hear the higher pitch in their right ear and the lower one in their left ear (Deutsch, 1983). No such bias was found to exist for left handed listeners. Second, and most notably, nobody can hear the pitches as they actually occur.

Pitch streaming illusions

The group of *pitch streaming illusions* is based on the same stimulation paradigm as the octave illusion. They all exploit the *Law of Good Continuation,* which states that if several sound components occur that are close to each other in pitch and form a regular pitch contour, they will be perceived as coming from the same sound source as part of the same sound event. This is sometimes true even if the other segregating cues suggest other interpretations. In the *scale illusion* (Deutsch, 1975), a major scale is presented with successive tones alternating from ear to ear. Two versions of the scale are presented simultaneously, ascending and descending, so that when a tone is played from the ascending scale in one ear, the corresponding tone from the descending scale is played in the other ear.

The *Law of Good Continuation* suggests that the tones that form a regular pitch contour will be grouped together. This is what occurs. Listeners usually hear the high tones in one ear and the low ones in the other. As with the octave illusion, right-handers tend to hear the higher notes in the right ear, while no regular bias occurs for left-handers.

Deutsch (1987; 1995; 2003) demonstrates several other illusions such as the *chromatic illusion*, the *Glissando illusion*, and the *Cambiata illusion* that all function the same way as the scale illusion. In each case, tones forming a regular pitch contour or that are close to one another in pitch, are grouped together and appear to come from the same ear. Handedness often plays a role in the assignment of a pitch register to an ear. For right-handers, lower pitches are often assigned to the left ear and higher ones are assigned to the right ear.

A further extension of pitch streaming illusions can be found in the *phantom word illusion* created when words or syllables are used in the place of pure tones. Several of these illusions were developed and described by Deutsch (2003). They all involve two syllables or two words (for example "high" and "low" in a *high-low illusion*) that are played simultaneously one word to each ear switching the ears after each presentation. The listeners always have an illusion that a certain word or a short phrase is played repeatedly but they never hear and report the words as they are actually presented.

The split-off illusion

The *split-off illusion* (Figure 13-8) appears when an ascending tone glide and a descending tone glide are played so that the descending glide begins 200 milliseconds (ms) before the ascending glide ends. The beginning pitch of the descending glide starts out lower than the pitch of the ascending glide at that point in time and the pitch trajectories never cross. However, the percept is that of a continuously rising and falling glide. The final 200 ms of the ascending tone "splits off" and is heard as a separate tone in the middle of the frequency range of the glide.

Several practical realizations of this illusion have been described (Remijn, Nakajima and ten Hoopen; Sasaki and Nakajima, 1996). In all cases, two longer glides are fused together, and components of these glides that are inconsistent with the perception of a simple smooth trajectory are "split-off" and are heard as a separate tone. The auditory system seems to connect the two glides in the simplest way possible according to the Gestalt *Law of Good Continuation*. Any components that are inconsistent with this interpretation are either ignored or parsed away from it as being independent from the fused glides.

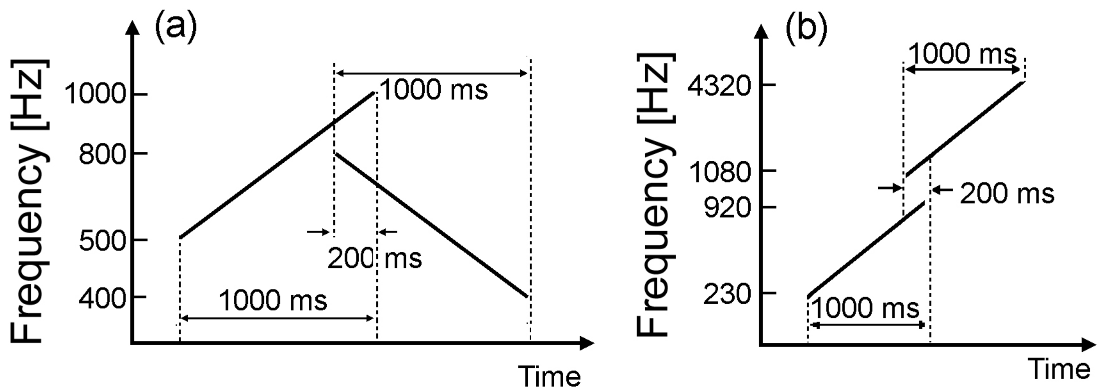

Figure 13-8. Two examples of the split-off illusion. In both cases, the final 200 ms of the first glide "splits off" and is perceived as an independent tone. The rest of the first glide and the second glide are joined perceptually into a single glide that either is (a) rising and falling or (b) rises continuously (adapted from Nakajima, Sasaki and ten Hoopen, 2006).

Huggins pitch (Dichotic pitch)

One Gestalt laws is the *Law of Proximity,* which states that sounds occurring near to each other in space are judged to be from the same sound source. The *Huggins pitch illusion* (Cramer and Huggins, 1958) is a faint pitch sensation that results from specific interpretation by the brain of the cue of interaural phase differences when two similar noise signals are delivered to the left and right ear. In the Huggins pitch demonstration the sounds delivered to each ear are white noise signals with the exception of three narrow bands; 400 to 440 Hz, 500 to 550 Hz, and 600 to 660 Hz. Both signals begin in phase, and then the phase of the signal delivered to one of the ears is advanced by 180° in each narrow band successively. The perceptual impression is that the noise is accompanied by a tone with gradually increasing faint pitch. No pitch is heard when the left or right ear signal is heard alone. Therefore, the Huggins pitch is a result of binaural processing of two slightly different signals.

The physiological explanation of Huggins pitch is still unclear. However, it seems that the brain separates the narrow bands of noise with 180° phase shift from the rest of noise and treats these sounds as coming from a different sound source separated in space from the noise source. Therefore, the Huggins pitch is essentially a spatial effect that that makes it easier to hear the presence of two simultaneous but different sounds. It is classified here as a pitch illusion but it may be equally well classified as a spatial illusion.

The Huggins pitch illusion illustrates an important consideration for HMDs and auditory displays in general. A binaural, spatial (3-dimensional) auditory display is the best way to improve detection of auditory information in operational conditions where noise cannot be reduced. This is true because to the degrees that the different signals are spatially separate, binaural presentation will create masking level differences, (discussed earlier in this chapter), and increase the effective signal to noise ratio.

Temporal conflicts and illusions

Temporal conflicts and illusions are generally related to perception of temporal patterns and the effects of the interstimulus interval on perceived sounds. The latter effects are related to the presence of temporal masking (short intervals) and memory traces (long intervals) and were already discussed in Chapter 11, *Auditory Perception and Cognitive Performance*. Therefore, the focus of the present chapter is on perception of temporal patterns and more precisely on the powerful illusion of pattern continuity whenever such continuity may be assumed.

The continuity effect

The *continuity effect* is an illusion observed when a soft sound is interrupted by a louder sound. Despite interruption, the original sound is heard as maintaining its continuity and the interrupting sound is heard as a separate event. The appearance of this illusion has certain limitations regarding the types of both sounds and their relative intensities but this is a powerful and easy to replicate perceptual effect. This continuity phenomenon was originally described by Miller and Licklider (1950) and called the *picket-fence effect*, but there is little doubt that it was heard and known before. The authors observed that a tone interrupted by a more intense burst of noise was heard as being continuous despite the interruption. The same effect was observed when the tone was replaced by speech. This idea of continuity illusion is shown schematically in Figure 13-9. When a tone interrupted by a temporal gap is presented in quiet (Figure 13-9a), the interruption is heard clearly. However, if a wideband noise burst is inserted into the gap, the tone is perceived to be continuous (Figure 13-9b). It is as though the auditory system assumes that the tone must be continuous and "fills in" the missing information. This effect is reduced if the wideband stimulus has a notch in the same frequency range as the tone (Figure 13-9c), which suggests that the auditory system may be extracting the tonal information from the wideband signal.

The continuity illusion also can be observed for tone glides, music, and continuous environmental sounds, such as rain sound, stream of typewriter sounds, etc. In other words, the continuity illusion works if the sounds before and after interruption are assumed to come from the same sound source (Bregman, 1990; Plack, 2005). Similarly to tones and other continuous sounds, continuous speech signal interrupted by short pauses looses its continuity and intelligibility; and the silent gaps are clearly heard. However, when the silent gaps are filled with bursts of wideband noise, coughs, or other wideband sounds, the listener has the impression that the speech signal is continuous and "hears" the missing parts of the speech sounds. As a result, speech recognition improves in comparison to that of the speech interrupted by the silent gaps. The mental restoration of the original speech masked by short interfering louder wide band sounds has been referred to as *phonemic restoration* (Warren, 1970; Warren and Obusek, 1971).

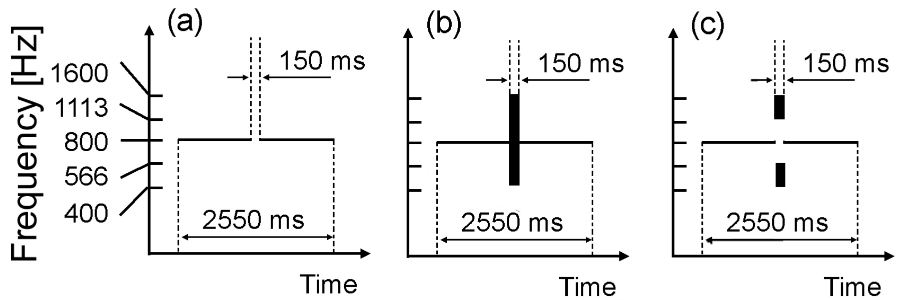

Figure 13-9. Schematic drawing of three time paradigms used to demonstrate the continuity effect; (a) continuous tone with silent gap, (b) continuous with a silent gap filled with broadband noise, and (c) continuous tone with a silent gap filled with two bands of noise surrounding the tone frequency (adapted from Nakajima, Sasaki and ten Hoopen, 2006).

The name *continuity effect* for the continuity illusion was used originally by Thurlow and Elfner to describe the perceived continuity of sound when two sounds alternate in time. The authors presented alternating pulses of a soft pure tone and a loud other sound (noise, another pure tone, etc.) that were alternating in time and observed that the soft tone was heard as a continuous tone (Elfner and Caskey, 1965; Thurlow, 1957; Thurlow and Elfner, 1959). Thurlow (1957), who apparently rediscovered this effect not knowing about the study by Miller and Licklider (1950), has originally called this effect the *auditory figure-ground effect* by an analogy to a similar figure-ground effect in vision (Rubin, 2001) where hidden background seems to be continuous behind foreground objects. An example of visual ground-figure effect is shown in Figure 13-10.

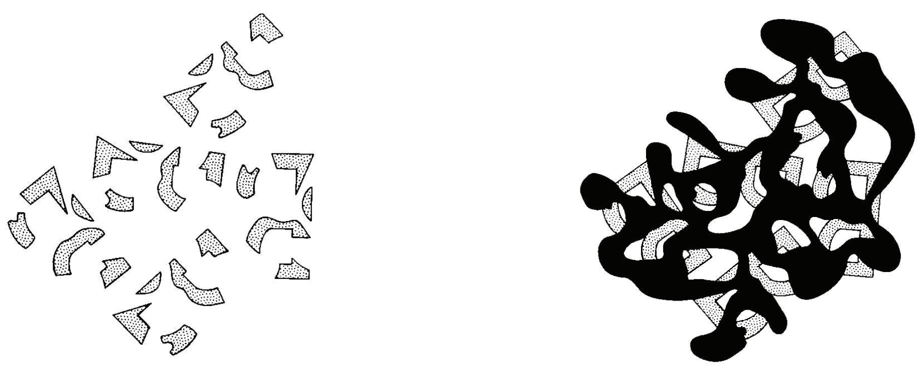

Figure 13-10. Visual example of perceptual restoration. The occluding information provides information about which elements are likely to be continuous and which elements are likely to be discrete (Bregman, 1981).

Since an alternate pulsing is an expanded form of single or multiple interferences, the term continuity effect can be used as a single label for all of the phenomena described here. However, various other names are also used in the literature to describe continuity effects including the *acoustic tunnel effect* (Vicario, 1960, cited by Warren, 1999), *auditory induction* (Warren, Obusek and Ackroff, 1972), and *temporal induction* (Warren, 1999). In addition, Warren and colleagues (Warren, Obusek and Ackroff, 1972) reported that the continuity illusion can be heard when the interfering sound is just a higher level of the original sound. The authors presented the same stimulus (octave band noise centered at 2000 Hz) at two alternating levels (70 and 80 dB sound pressure level [SPL]) and observed that the lower level was heard as always being present while the higher level was heard as an additional sound. This observation led Warren (1999) to differentiate between *homophonic continuity* (the same signal, different levels) and *heterophonic continuity* (different signals, different levels) effects.

A parallel effect to the continuity effect (temporal restoration) is the spectral restoration effect where presence of noise in spectral gaps of filtered speech improves speech intelligibility. For example, Bashford and Warren, (1987) and Shriberg (1992) reported improved speech intelligibility where low-pass or high-pass filtered speech was heard together with a complementary filtered noise. Both of these effects together have been referred to as *perceptual restoration* by Warren (1999)

Pulsation threshold

The *pulsation threshold* is not an auditory illusion but a research method used to assess spectro-temporal analysis performed by the auditory system (Fastl, 1975; Houtgast, 1972; Letowski and Smurzynski, 1983). Since this method is based on some properties of the continuity effect and has important implications for understanding auditory physiology it, deserves a short description.

The pulsation threshold methodology was developed by Houtgast (1972) as an alternative to temporal masking in studying lateral suppression. The procedure uses relatively long maskers and signals so it is much easier to use than temporal masking techniques. This methodology is based on an observation that in order for the continuity effect to occur, the intensity of the louder sound must be such that the softer sound would be totally masked if the louder sound was continuous (Warren, Obusek and Ackroff, 1972; Houtgast, 1972). This means that the barely audible pulsing of the softer sound can be used as a measure of its masked threshold, that is, the amount of the excitation overlap of the two sounds in the cochlea (Platt, 2005). This methodology is especially convenient for measuring masked threshold in the vicinity of the masked tone or its harmonics in tone-on-tone experiments where the potential beats[1] make a simultaneous masking technique quite unusable. It is also important that despite a subjective criterion that is used by the listener in determining whether the soft signal is continuous or pulsing, the pulsation threshold data have relatively low variability (Plack and Oxenham, 2000).

A schematic diagram of the temporal pattern used to measure the pulsation threshold is shown in Figure 13-11. A masking (M) and a masked (m) tone alternate in time and the experimenter adjusts either the level of the masking tone or the level of the masked tone until the continuity effect is heard. The masking tone is usually a wideband noise or a pure tone signal and the masked tone is a pure tone signal. The former adjustment procedures is used to determine the level of the masking tone needed to mask the other tone and the latter procedure is used to measure the masked threshold of the signal in the presence of a given masking tone. The optimum alternation cycle for measuring pulsation effect is about 4 Hz but interruption times can be as long as 300 ms beyond which tonal continuity cannot be maintained for longer periods (Houtgast, 1974; Warren, 1999).

The gap transfer illusion

When a long ascending glide tone is crossed in the middle by a short descending glide tone (Figure 13-12a), the pitch percept of the longer tone is often sigmoidal (Halpern, 1977; McPherson, Ciocca and Bregman, 1994; Tougas and Bregman, 1985). If a short, 100 ms gap is added to the short glide, the pitch is perceived veridically (Figure 13-12b). Evidently, when there is no gap, the *Law of Similarity* encourages us to group the pitch components in the shorter glide with the dominant longer glide. The gap separates those frequencies and allows the longer ascending pitch to be perceived veridically. Nakajima and his colleagues (Nakajima et al., 2000) show

---

[1] *Beats* are the effect produced by interference of waves of slightly different frequency, producing a pattern of alternating intensity.

that the same percept is achieved when the gap is placed in the longer glide (Figure 13-12c). Essentially, the gap is “transferred” to the shorter glide and the long glide is perceived as continuous. Here, the *Law of Good Continuation* dictates that the well established long glide is more likely to be continuous and assigns the nearby frequencies to it.

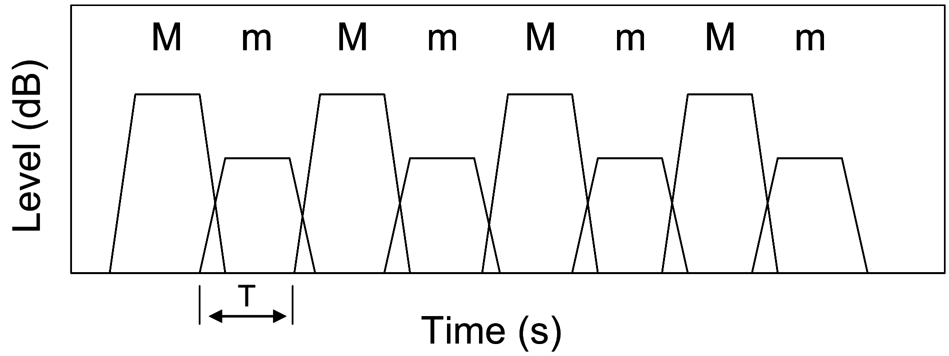

Figure 13-11. An example of the temporal paradigm used to elicit a pulsation threshold. Two signals, masker (M) and maskee (m) alternate in time with onset and offset ramps of about 20 ms to avoid generation of audible clicks during transitions. The duration (T) of each pulse is the same and usually about 100-150 ms.

It is somewhat striking that the auditory system is so susceptible to bias. It the previous example, the gap is in an ascending glide, and the short glide is descending – yet the long glide is perceived as continuously ascending. It is as though the auditory system is unable to adequately process the pitch information, so it “fills in” the missing information with a plausible interpretation. The *Law of Simplicity* essentially posits that the perceptual system will interpret sensory information with the simplest interpretation possible. Take, for example, Figure 13-13. Are the drawings 2-dimensional or 3-dimensional? Either interpretation is valid; however, the probability of a three-dimensional interpretation changes as you progress through the figures. The simplest interpretation of Figure 13-13a is that of a cube. Figure 13-13b is symmetric, but still has irregular shapes, and is still probably interpreted as a cube. However, Figure 13-13c consists of 6 congruent triangles in a symmetric arrangement and either a 2- or 3-dimensional interpretation is equally possible.

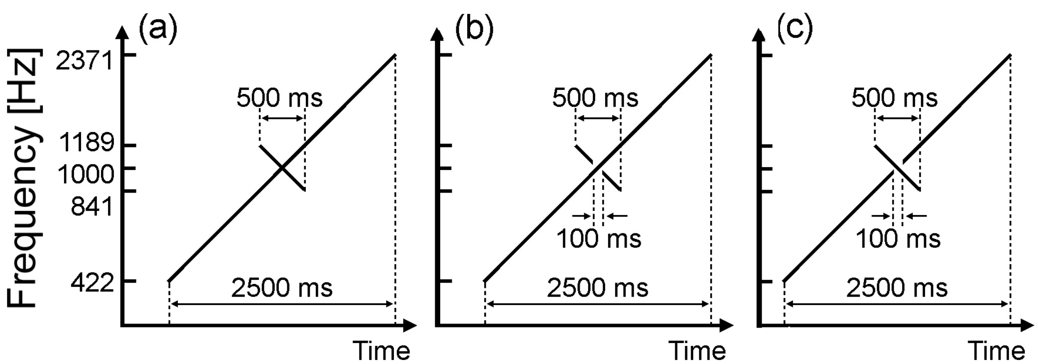

Figure 13-12. The gap transfer illusion. The percept of the event depicted in (a) is usually sigmoidal; the frequency components in the short glide are incorporated into the dominant longer glide. (b) Adding a gap to the short glide allows the event to be heard veridically. Event (c) is perceived to be the same as event (b); the gap is transferred perceptually to the shorter glide and the longer glide is perceived as continuous (adapted from Nakajima, Sasaki and ten Hoopen, 2006).

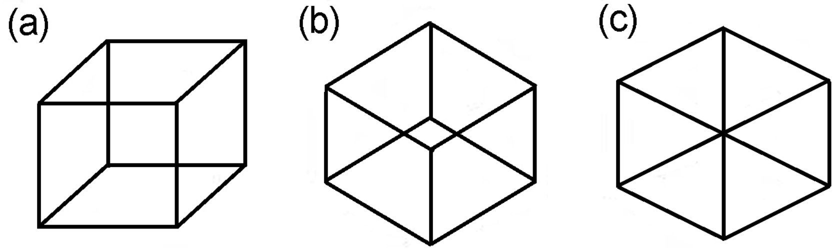

Figure 13-13. Figures demonstrating the *Law of Simplicity*. Observer perceptions are guided by the simplest interpretations of these lines and there is a bias to perceive them as simple, symmetric shapes.

Spatial conflicts and illusions

The Franssen effect

The *Franssen effect* (Franssen 1960; Yost, Mapes-Riordan and Guzman, 1997) is a spatial illusion caused by the temporal difference in the onset of sound in the left and right ear of the listener. It appears if a narrowband stimulus is presented in a reverberant sound field from two loudspeakers located at about ±45º and placed at a certain distance (1 meter or more) from the listener. In the demonstration of this effect the one loudspeaker that delivers the sound is abruptly turned on and then slowly turned off with short and linear offset ramp (20 to 100 ms). The other loudspeaker is gradually turned on and the sound is delivered with an onset envelope that is a mirror image of the offset envelope of the first sound and kept on for some time (e.g., 2 to 10 s) after reaching its steady state level. If the sound is a wideband noise, the noise is suddenly heard in one ear and then "jumps" to the other ear for the rest of the sound duration, as expected. However, when the sound is a narrowband noise or a tone (e.g., 1000 Hz tone) the sound is heard as coming from the first loudspeaker the entire time, that is, from the loudspeaker that delivered a short abrupt sound during the first short period of the sound presentation.

The Franssen effect is a special case of the *precedence effect* (see Chapter 11, *Auditory Perception and Cognitive Performance*). The precedence effect results from the fact that the sound arriving first at the ears determines the perceived location of a sound source in the space. The Franssen effect occurs because people have difficulty localizing sounds with a gradual onset envelope and room reflections obscure the residual localization cues. The effect is the easiest to demonstrate for stimuli in 1.0 to 2.5 kHz frequency range, that is, in the range when neither temporal nor intensity localization cues work very well and it is difficult for people to localize sound sources. The presence of reflections is very important for eliciting the Franssen effect and the effect is difficult to demonstrate in anechoic conditions or through earphones (Hartmann and Rakerd, 1989; Rakerd and Hartmann, 1985). However, under proper listening conditions the illusion may exist for very long sounds; even infinitely long (Bradley 1983). The Franssen effect is an auditory illusion that requires unusual circumstances to occur, but it demonstrates our perceptual reliance on sound onset cues for sound source localization.

The Clifton effect

The *Clifton effect* (Clifton, 1987; Clifton and Freyman, 1989) is a spatial illusion that represents the breakdown of the precedence effect (see Chapter 11, *Auditory Perception and Cognitive Performance*). To demonstrate the Clifton effect a train of click pairs separated by several milliseconds (e.g., 12 ms) is emitted from two loudspeakers located at about ±45º and at a certain distance from the listener. Each pair of clicks consists of a click from one loudspeaker (the lead loudspeaker) and a click from the other loudspeaker (the lag loudspeaker). After initial presentation of several pairs of the clicks, the loudspeakers delivering the lead and the lagging clicks are reversed, that is, the loudspeaker delivering the lead click now delivers the lagging click and vice versa. As predicted by the precedence effect, most listeners perceive a single click arriving from the location of the lead loudspeaker during the presentation of the initial segment of the click train. Immediately after the switch, the clicks are perceived as coming from both loudspeakers. This perceptual effect last for the duration of 3 to 5 click-pairs. After this period of time a single click, coming again from the leading (now the opposite) loudspeaker is heard, as predicted by the precedence effect. The Clifton effect demonstrates human inability to immediately adjust to the change in the position of the leading click and the temporary failure of the precedence effect (Yost and Guzman, 1996). This illusion suggests that the suppression mechanism of the lagging stimulus may be active for some time after termination of the lead sound. The duration of this activity may be dependent on the duration of the exposure to the lead-lag pairs of stimuli coming from the same locations (Litovsky et al., 1999).

Attention and Illusions

The perceptual system is able to acquire a large amount of sensory information, but it is only able to process and interpret a small proportion of it. The maximum rate of information that can be processed by a single sensory channel is called *channel capacity*. The concept of channel capacity is identical with the concept of working memory capacity, which refers to the maximum amount of sensory information that can be held temporarily in a storage buffer (Baddeley, 1992). This storage buffer is needed by the sensory system in order to continuously process incoming information. The processed information is then either used for current decision-making processes or stored in long-term (permanent) memory. The other terms that convey the same meaning as working memory are short-term memory and operational memory.

Miller (1956) posited that one could hold seven plus or minus two (7 ± 2) items of information in short term storage, roughly the equivalent of a phone number. Broadbent (1975) argued that the capacity of short-term memory is even smaller and the memory can only hold approximately three items. Either way, the capacity of working memory is very small and the concept of information "item" is somewhat nebulous. It seems that during information processing small items (chunks) of information that are well known or meaningful are grouped together into bigger and bigger chunks and each chunk can constitute an item. In the case of a phone number, the information item could be just a digit or it could be the area code or it could be the entire number.

Regardless of the precise concept of information item the absolute capacity of short term memory is very limited. Therefore, the perceptual system must be selective in which information is processed or attended to. There are numerous theories proposing how attention is allotted to a scene and which events generally elicit deeper processing (for a complete discussion, see Jones and Yee, 1993). However, since the overall attentional resources are capped at a certain level, they only can be increased in one channel by reducing those of another.

The existence of perceptual illusions demonstrates the fact that in order to facilitate efficient processing of information, the perceptual system relies on a number of heuristics to process sensory information and often "fills in" missing or contradictory information with plausible interpretations when information is incomplete (Shinn-Cunningham, 2008). Therefore, from the neuroscience point of view, perceptual illusions create an important key to understanding how the brain processes incoming auditory information and which parts of the information are not being processed. Also, from the display design point of view, the existence of illusions highlights the need to consider attentional limits and minimize complexity of incoming information, as well as considering their ecological validity to the perceptual system itself. Failure to do so increases the probability of errors due to incorrectly interpreted events. Knowledge of how the auditory system parses information can help to determine ways to increase the effective signal to noise ratio, both perceptually and cognitively. In addition, knowledge of common perceptual biases can also highlight potential sources of information loss and situations where the use of redundancy across modalities is warranted.

References

Adams, J., and Duda, P. (1986). Laterality of cross-modal spatial processing. *Cortex,* 22*,* 539-552.

Baddeley, A.D. (1992). Working memory. *Science,* 255, 556-559

Bradley, D.A. (1983). Room acoustics and listening. *Journal of the Acoustical Society of America,* 73*,* 17.

Bregman, A. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press.

Bregman, A.S. (1981). Asking the "what for" question in auditory perception. In: *Perceptual Organization,* Kubovy, M., and Pomerantz, J.R., (Eds.). Hillsdale, NJ: Erlbaum.

Broadbent, D.E. (1975). The magic number seven after fifteen years. In: *Studies in Long-Term Memory,* Kennedy, A., and Wilkes, A. (Eds.), (pp. 3-18). New York: John Wiley and Sons.

Burns, E.M. (1981). Circularity in relative pitch judgments for inharmonic complex tones; The Shepard demonstration revisited, again. *Perception and Psychophysics,* 30*,* 467-472.

Clifton, R.K. (1987). Breakdown of echo suppression in the precedence effect. *Journal of the Acoustical Society of America,* 82*,* 1834-1835.

Clifton, R.K., and Freyman, R. (1989). Effect of click rate and delay on breakdown of the precedence effect. *Perception and Psychophysics,* 46*,* 139-145.

Cramer, E.M., and Huggins, W.H. (1958). Creation of pitch through binaural interaction. *Journal of the Acoustical Society of America,* 30*,* 412-417.

Deutsch, D. (1974). An auditory illusion. *Nature,* 251, 307-309.

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

Deutsch, D. (1983). The octave illusion in relation to handedness and familiar handedness background. *Neuropsychologia,* 21, 289-293.

Deutsch, D. (1987). Illusions for stereo headphones. *Audio Magazine,* 71, 36-48.

Deutsch, D. (1995). *Musical Illusions and Paradoxes*. La Jolla, CA: Philomel Records.

Deutsch, D. (1999). The processing of pitch combinations. In: *The Psychology of Music,* Deutsch, D. (Ed.), (pp. 349-411). San Diego: Academic Press.

Deutsch, D. (2003). *Phantom Words and Other Curiosities*. La Jolla, CA: Philomel Records.

Elfner, L.F., and Caskey, W.E. (1965). Continuity effects with alternating sounded noise and tone signals as a function of manner of presentation. *Journal of the Acoustical Society of America,* 38, 543-547.

Emanuel, D., and Letowski T. (2009). *Hearing Science*. Baltimore: Lippincott Williams and Wilkins.

Fastl, H. (1975). Pulsation patterns of sinusoids vs. critical band noise. *Perception and Psychophysics,* 18, 95-97.

Franssen, N.V. (1960). Some considerations on the mechanism of directional hearing. Ph.D. Dissertation. Delft University of Technology (Technische Hogeschool). The Netherlands.

Halpern, D.L. (1977). The effect of harmonic ratio relationships on auditory stream segregation. Unpublished research report, Psychology Department, McGill University.

Hartmann, W.M., and Rakerd, B. (1989). Localization of sound in rooms IV: The Franssen effect. *Journal of the Acoustical Society of America,* 86, 1366-1373.

Heller, M., Calcaterra, J., Green, S., and Brown L. (1999). Intersensory conflict between vision and touch: The response modality dominates when precise, attention-riveting judgments are required. *Perception and Psychophysics,* 61, 1384-1398.

Hupe, J.M., Joffo, L.M., and Pressnitzer, D. (2008). Bistability of audiovisual stimuli: perceptual decision is modality specific. *Journal of Vision,* 8, 1-15.

Hershberger, W., and Misceo, G. (1996). Touch dominates haptic estimates of discordant visual-haptic size. *Perception and Psychophysics,* 58, 1124-1132.

Houtgast, T. (1972). Psychophysical evidence of lateral inhibition in hearing. *Journal of the Acoustical Society of America,* 51, 1885-1894.

Jones, M.R., and Yee, W. (1993). Attending to auditory events: the role of temporal organization. In: *Thinking in Sound: The Cognitive Psychology of Human Audition,* McAdams, S., and Bigand, E. (Eds.), (pp. 69-112). Oxford, UK: Oxford Science Publications.

Leopold, D., and Logithesis, N. (1999). Multistable phenomena: Changing views in perception. *Trends in Cognitive Science,* 3, 254-264.

Letowski T., and Smurzynski, J. (1983). Evaluation of electroacoustic devices by equivalent scale methods. *Archives of Acoustics,* 8, 133-154, 1983

Litovsky, R.Y., Colburn, H.S., Yost, W.A., and Guzman, S.J. (1999). The precedence effect. *Journal of the Acoustical Society of America,* 106, 1633-1652.

Marr, D. (1982). *Vision*. San Francisco: W.H. Freeman.

McGurk, H., and MacDonald, J. (1976). Hearing lips and seeing voices, *Nature,* 264**,** 746-748.

McPherson, L., Ciocca, V., and Bregman, A.S. (1994). Organization in audition by similarity in rate of change: Evidence from tracking individual frequency glides in mixtures. *Perception and Psychophysics,* 55, 269-278.

Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review,* 63, 81-97.

Miller, G.A., and Licklider, J.C.R. (1950). The intelligibility of interrupted speech. *Journal of the Acoustical Society of America,* 22, 167-173.

Mussap, A.J., and Crassini, B. (1993). Barber-pole illusions and plaids - The influence of aperture shape on motion perception. *Perception,* 22, 1155-1174.

Nakajima, Y., Sasaki, T., Kanafuka, K., Miyamoto, A., Remijn, G., and ten Hoopen, G. (2000). Illusory recouplings of onsets and terminations of glide tone components. *Perception and Psychophysics,* 62, 1413-1425.

Nakajima, Y., Sasaki, T., and ten Hoopen, G. (2006). *Demonstrations of auditory illusions*. Retrieved 26 August, 2008 from: http://www.kyushu-id.ac.jp/~ynhome/ENG/Demo/illusions.html#title

Pantev, C., Hoke, M., Lütkenhöner, B., and Lehnertz, K. (1989). Tonotopic organization of the auditory cortex: Pitch versus frequency representation. *Science,* 246, 468-488.

Plack, C.J., (2005). *The Sense of Hearing,* (p. 212). Mahwah, NJ: Lawrence Erlbaum Associates.

Plack, C.J., and Oxenham, A.J. (2000). Basilar-membrane nonlinearity estimated by pulsation threshold. *Journal of the Acoustical Society of America,* 107, 500-507.

Pressnitzer, D., and Hupe, J.M. (2006). Temporal dynamics of auditory and visual bistability reveal common principles of perceptual organization. *Current Biology,* 16, 1351-1357.

Rakerd, B., and Hartmann, W.M. (1985). Localization of sounds in rooms II: The effect of a single reflecting surface. *Journal of the Acoustical Society of America,* 78, 524-533.

Rakowski, A. (1993). Categorical perception in absolute pitch. *Archives of Acoustics,* 18, 515-523.

Remijn, G., Nakajima, Y., and ten Hoopen, G. The split-off effect and the perception of auditory continuity (as cited in Nakajima, Sasaki and ten Hoopen, 2006).

Risset, J.C. (1969). Pitch control and pitch paradoxes demonstrated with computer-synthesized sounds. *Journal of the Acoustical Society of America,* 46, 88.

Risset, J.C. (1972). *Musical Acoustics*. Paris: IRCAM.

Risset, J.C. (1986). Pitch and rhythm paradoxes: Comments on "Auditory paradox based on a fractal waveform". *Journal of the Acoustical Society of America,* 80, 961-962.

Rubin, E. (2001). Figure and ground. In *Visual Perception*, Yantis, E. (Ed.), (pp. 225-229). Philadelphia: Psychology Press.

Sasaki, T., and Nakajima, Y. (1996). An illusory reconstruction of auditory elements. *Journal of the Acoustical Society of America,* 100, 2751.

Seckel, A. (2004). *Masters of Deception: Escher, Dalí and the Artists of Optical Illusion*. New York: Sterling Publishers.

Shepard, R.N. (1964). Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America,* 36, 2346-2353.

Shepard, R.N. (1982). Geometrical approximations to the structure of musical pitch. *Psychological Review,* 89, 305-333.

Shinn-Cunningham, B.G. (2008). Object-based auditory and visual attention. *Trends in Cognitive Sciences,* 12, 182-186.

Shriberg, E.E. (1992). Perceptual restoration of filtered vowels with added noise. *Language and Speech,* 35, 127-136.

Solso, R.L. (2001). *Cognitive psychology* (6th Ed.). Boston: Allyn and Bacon

Tenney, J. (1969). Computer music experiences: 1961-1964. *Electronic Music Report #1*. Utrecht, Netherlands: Institute of Sonology.

Thurlow, W.R. (1957). An auditory figure-ground effect. *American Journal of Psychology,* 70, 653-654.

Thurlow, W.R., and Elfner, L.F. (1959). Continuity effects with alternately sounding tones. *Journal of the Acoustical Society of America,* 31, 1337-1339.

Tougas, Y., and Bregman, A.S. (1985). The crossing of auditory streams. *Journal of Experimental Psychology: Human Perception and Performance,* 11, 788-798.

Van Noorden, L. (1975). Temporal coherence in the perception of tonal sequences. Ph.D. dissertation. Eindhoven University of Technology, Leiden (The Netherlands).

Vicario, G. (1960). L'effetto tunnel acustico. *Rivista di Psicologia,* 54, 41-52.

Warren, J.D., Uppenkamp, S., Patterson, R.D., and Griffiths, T.D. (2003). Separating pitch chroma and pitch height in the human brain. *Proceedings of the National Academy of Sciences USA,* 100, 10038-10042.

Warren, R.W. (1970). Perceptual restoration of missing speech sounds. *Science,* 167, 392-293.

Warren, R.M. (1999). *Auditory perception: A new analysis and synthesis*. Cambridge, UK: Cambridge University Press.

Warren, R.M., and Obusek, C.J. (1971). Speech perception and phonemic restorations. *Perception and Psychophysics,* 9, 358-362.

Warren, R.M., Obusek, C.J., and Ackroff, J.M. (1972). Auditory induction: Perceptual synthesis of absent sounds. *Science,* 176, 1149-1151.

Warren, R.M., Hainsworth, K.R., Brubaker, B.S., Bashford, J.A. Jr., and Healy, E.W. (1997). Spectral rtestoration of speech: Intelligibility is increased by inserting noise in spectral gaps. *Perception and Psychophysics,* 57, 275-283.

Wikipedia. (2008). Shepard tone. Retrieved 1 August 2008 from: http://en.wikipedia.org/wiki/ Shepard_tone

Yost, W.A., and Guzman, S.J. (1996). Auditory processing of sound sources: Is there an echo in here? *Current Directions in Psychological Science,* 5, 125-131.

Yost, W.A., Mapes-Riordan, D., and Guzman, S.J. (1997). The relationship between localization and the Franssen effect. *Journal of the Acoustical Society of America,* 101, 2994-2997.