Play Audio File

AUDITORY ILLUSIONS AND AUDIO

Introduction

The last few years have seen extraordinary progress in audio technology. The impressive level of sophistication that has been reached might incline us to hope that, through simple refinement of this technology, sounds will ultimately be heard with optimal fidelity. But the very technological advances that have enabled us to create and manipulate sound so effectively have revealed the existence of powerful factors in the nervous system which determine the way that sound is perceived. These factors can sometimes be so overwhelming that they result in percepts which bear little resemblance to the sound presented. An understanding of such factors is therefore necessary if the goal of optimal fidelity is to be achieved.

In the past the science of psychoacoustics has been concerned largely with the perception of single sounds in isolation, or of simple combinations of such sounds. While much important information has been gained from these studies, they explore only a limited aspect of the hearing process. When ongoing streams of sound are presented instead, powerful higher level mechanisms come into play, which can substantially modify what is heard. The situation is typical of normal listening, so that an understanding of such higher level mechanisms is of particular importance to the audio community.

The effects described and demonstrated here show that where sound patterns of any complexity are concerned, the auditory system is extremely prone to illusion. This is an inevitable consequence of a system that has evolved to optimize the perception of sounds produced in our natural environment.

First, most naturally occurring sounds, when considered alone, are ambiguous in their interpretation. In order to resolve such ambiguities, it is necessary to draw on as much additional information as we have available. Such information can come either from surrounding sounds or from input through other sensory systems. The ability of the hearing mechanism to utilize such extraneous information generally confers an enormous advantage. But by the same token, if a sound is presented in a misleading or inappropriate context, percepts may go wildly wrong.

Second, sounds presented in the natural environment are subject to considerable and complex changes before they reach our eardrums. As a result, we have evolved a number of special-purpose mechanisms to counteract the effects of such changes, and also to exploit them to provide additional information concerning sound sources. Such mechanisms are also enormously useful in listening to naturally occurring sounds. However, they can lead to gross misperceptions when sounds are presented in altered form.

The papers in this issue examine auditory illusions from several different points of view. First, powerful illusions may readily occur when multiple streams of sound emanate in parallel from different regions of space. These form the subject of the first paper (by Deutsch). The illusions are particularly striking on headphone listening, but can also occur with stereophonically presented sounds or in live performance in concert halls. They therefore have strong implications for engineers concerned with stereophonic recording and reproduction, sound system design, and concert hall acoustics.

The second paper (by Warren) is of particular importance to engineers involved in sound editing and mixing. The author emphasizes illusions which occur when a signal is degraded by intermittent noise, or presented in combination with other sounds. Depending on the spectral relationships between the sounds involved, the listener may either generate an illusory resynthesis of the masked signal, or alternatively allow the extraneous sound to give rise to perceptual interference.

The widespread use of headphones, by both performing artists and the public, necessitates an understanding of illusions which occur under these conditions. This forms the main topic of the third paper (by Kubovy and Daniel). The authors demonstrate a number of striking illusory effects which result from systematic changes in phase relationships between the signals at the two ears, together with illusions resulting from signal discontinuities. Effects of phase differences in normal listening are also discussed.

Readers interested in computer music and sound effects will find the fourth paper (by Shepard) particularly intriguing. The author presents what may best be described as "impossible auditory objects," such as a sound that appears to be eternally ascending in pitch, or both ascending and descending at the same time. The author also gives an account of the principles by which the brain appears to arrive at these paradoxical percepts.

The development of sophisticated sound systems for movie theaters and the advent of stereo television and stereo video lead us to the important question of how input from the visual system affects perceived sound. The fifth paper (by Lackner) examines this, together with the effect of body orientation on how sound is perceived. The author argues that we should not think of the hearing mechanism as functioning in isolation, but rather as part of a single, complexly interacting system.

The illusions discussed in this issue have multiple implications for the audio community. However, we present them also in the hope that you will find them intriguing and entertaining.

Diana Deutsch
Guest Editor

Auditory Illusions, Handedness, and the Spatial Environment*[1]

DIANA DEUTSCH

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

A number of paradoxical illusions occur when two streams of sound emanate simultaneously from different regions of space. Right-handers and left-handers differ statistically in terms of the illusory percepts obtained. The bases for these illusions are examined, and their implications are discussed.

0 INTRODUCTION

This paper is addressed to how the listener organizes multiple streams of sound which arise in parallel from different regions of space. This situation is typical of normal listening, but only recently has the technology become available for its investigation in a rigorous experimental setting. The resultant picture of the auditory system that has emerged is very different from that derived from previous experiments with isolated sounds. In particular, a number of powerful illusions have been revealed, which demonstrate the involvement of complex high-level processes in auditory perception, even when the simplest sound parameters are involved. The illusions have practical implications for those concerned with stereophonic recording and reproduction, sound system design, and concert hall acoustics.

1 THE OCTAVE ILLUSION

1.1 Basic Phenomena

We shall begin by exploring a very simple sound pattern, which involves only sine-wave tones at two frequency values. The pattern is presented in **Sound Demonstration 1** and is represented in Fig. 1(a). It consists of two tones, which are spaced an octave apart and are repeatedly presented in alternation. The identical sequence is presented to both ears simultaneously. However, when the right ear receives the high tone, the left ear receives the low tone, and vice versa. So in fact the listener is presented with a single continuous two-tone chord, but the ear of input for each component switches repeatedly [1]. The reader will find it instructive to listen to the sound demonstration first in stereo, and then, for comparison, with the two channels mixed.

One can easily imagine how this pattern should sound if perceived correctly. However, such a percept is very rare. Instead, a variety of paradoxical illusions are obtained. The type of illusion differs from one listener to another, the most common one being represented in Fig. 1(b). This consists of a single tone which alternates from ear to ear, and whose pitch simultaneously shifts back and forth between high and low. That is, the listener hears a single intermittent high tone in one ear, which alternates with a single intermittent low tone in the other ear.

We cannot explain this illusion in any simple way. We could advance an explanation of the perception of alternating pitches by suggesting that the listener attends to those presented to one ear rather than the other. But then both of the alternating pitches should appear localized in the same ear. Again, we could advance an explanation of the perception of a tone switching from ear to ear by suggesting that the listener alternate his attention back and forth between ears. But then the pitch of this tone should not change with a change in its apparent location. The illusion of a single tone that alternates simultaneously both in pitch and in location is quite paradoxical.

A further surprise occurs when the listener's earphones are placed in reverse position. Most commonly the identical percept is obtained: the tone that had appeared in the right ear still appears in the right ear, and the tone that had appeared in the left ear still appears in the left ear. So it appears as if the earphone which had been producing the high tone is now producing the

* Presented at the 72nd Convention of the Audio Engineering Society, Anaheim, California, 1982 October 23–27. Manuscript received 1983 March 15; revised 1983 June 3.

[1] Sound demonstrations described in this paper are contained on disk A, side 1, page 621 of this issue. It is essential that the sound demonstrations be played stereophonically through headphones with the loudspeakers in the system turned off. The volume should be set for comfortable listening such that the amplitude of the calibration tone which precedes the demonstrations is equal in both ears.

low tone, and that the earphone which had been producing the low tone is now producing the high tone!

If we assume that separate brain mechanisms exist for determining *what* pitch we hear and for determining *where* the sound is coming from, we are in a position to explain this paradoxical illusion [2]. The model is shown in Fig. 2. To provide the perceived pitches, the frequencies arriving at one ear are attended to, and those arriving at the other ear are suppressed. However, to provide the perceived locations, each tone is localized in the the ear that receives the higher frequency signal, regardless of whether the higher or the lower frequency is in fact perceived. Let us take the listener who follows the frequencies presented to his right ear. When a high tone is delivered to the right ear and a low tone to the left, this listener hears a high tone, since this is delivered to his right ear. Further, he localizes the tone in his right ear, since this ear is receiving the higher frequency. However, when a low tone is delivered to the right ear and a high tone to the left, this listener now hears a low tone, since this is delivered to his right ear; but he localizes the tone in his left ear instead, since this ear is receiving the higher frequency. So the entire sequence is heard as a high tone to the right alternating with a low tone to the left. It can be seen that reversing the position of the earphones would not alter this basic percept (though the identities of the first and last tones in the sequence would reverse). But given a listener who follows the frequencies delivered to his left ear instead, keeping the localization rule constant, the same sequence would be perceived as a high tone to the left alternating with a low tone to the right. Further experiments have confirmed this model [3]–[6].

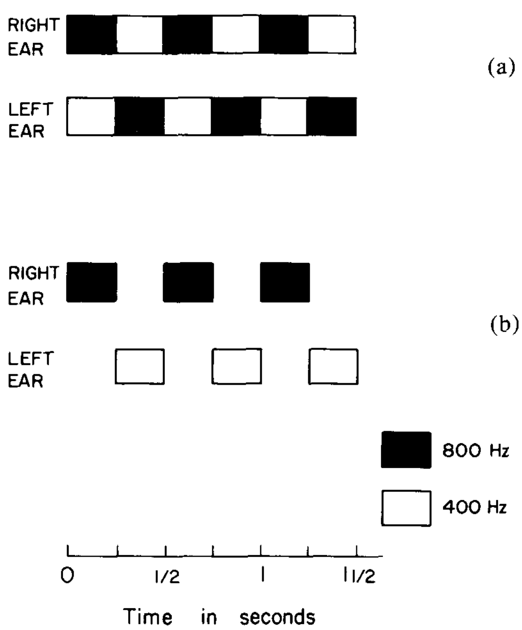

Fig. 1. (a) Sound pattern giving rise to the octave illusion. (b) Representation of the illusion most commonly obtained. A high tone in the right ear appears to alternate with a low tone in the left ear. (From Deutsch [1].)

Although most listeners show a preference for one localization pattern rather than another, and although such a preference can be very strong, the high and low tones may suddenly reverse position. Such reversals are most likely to occur at the start of a new sequence; but they may occur without warning in the middle of one. A few listeners experience very frequent reversals, and such percepts provide us with an auditory analogue of the reversal of ambiguous figures in vision; for example, the Necker cube (Fig. 3).

The illusion described so far is the one most commonly obtained. However, other listeners perceive the sequence quite differently. Some hear a single tone which alternates from ear to ear, and whose pitch either remains constant or changes only slightly with a change in its apparent location. Others obtain complex percepts, such as two low tones that alternate from ear to ear, together with an intermittent high tone in one ear; or a sequence in which the pitch relationships appear to change gradually. Some listeners experience striking differences in timbre between the tones; for instance, the low tones may have a gonglike quality and the high tones a flutelike quality. This group of percepts tends to be rather unstable, often changing within a few seconds.

The striking individual differences in perception of

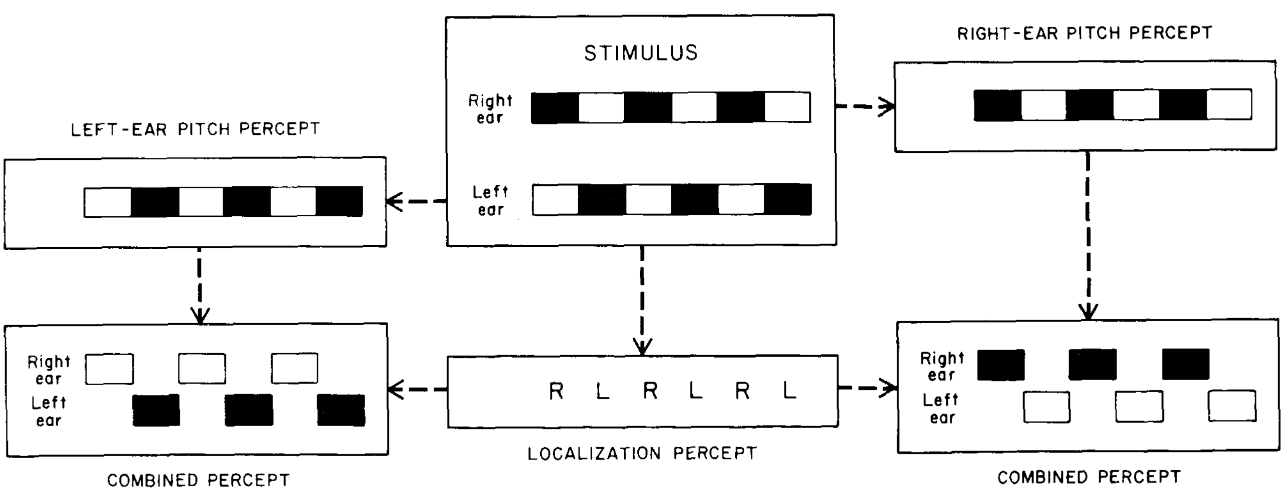

Fig. 2. Model showing how the outputs of two brain mechanisms, one determining perceived pitch and the other determining perceived location, combine to produce the octave illusion. See text for details. (From Deutsch [6]).

the octave illusion have been found to correlate with handedness. First, the proportion of listeners obtaining complex percepts is substantially higher among left-handers than right-handers. A second correlate concerns the localization patterns for the high and low tones. Right-handers tend strongly to hear the high tone on the right and the low tone on the left, and also to maintain this localization pattern when the earphones are placed in reverse position. However, left-handers as a group do not show this tendency and are less stable in their localization patterns [1]. These results are in accordance with the neurological evidence showing that the overwhelming majority of right-handers are left-hemisphere dominant (that is, they have speech represented in this hemisphere); however, this is true of only about two-thirds of left-handers. Further, the majority of right-handers have a clear dominance of the left hemisphere; however, a substantial proportion of left-handers have some speech representation in both hemispheres [7]–[9].

Localization patterns in the illusion have recently been shown to correlate not only with handedness, but even with familial handedness history. Listeners with left- or mixed-handed parents or siblings were found less likely to localize the high tone on the right and the low tone on the left than were listeners without left- or mixed-handed parents or siblings. This was found true for right-handers, mixed-handers, and left-handers [10]. This result is in accordance with findings relating patterns of brain dominance to familial handedness background [7]–[9].

A further question of interest is whether the interactions giving rise to the octave illusion occur between pathways conveying information from the two ears, or whether instead pathways conveying information from different regions of auditory space are involved. To investigate this question, the sound pattern was presented through spatially separated loudspeakers rather

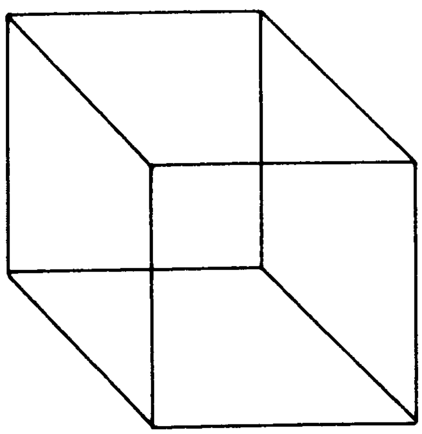

Fig. 3. The Necker cube. This figure is perceived as spontaneously changing its position in depth; however it is only perceived in one position at a time. Some listeners experience the octave illusion pattern as fluctuating analogously, so that the percept of a high tone in the right ear alternating with a low tone in the left ear interchanges with the percept of a high tone in the left ear alternating with a low tone in the right ear.

than earphones in an anechoic chamber. The analogous illusion was obtained, even though both ears now received the entire sound pattern. When the listener was positioned so that one loudspeaker was on the right and the other on the left, a high tone was heard as coming from the right, and a low tone as from the left. When the listener turned slowly, the high tone remained on the right and the low tone on the left. This percept was maintained until the listener was facing one loudspeaker, with the other directly behind him. The illusion then abruptly disappeared, and a single complex tone appeared to be coming simultaneously from both loudspeakers. But as the listener continued to turn, the illusion abruptly reappeared, with the high tone still on the right and the low tone on the left. So when the listener had turned 180°, it appeared that the loudspeaker which had been producing the high tone was now producing the low tone, and the loudspeaker which had been producing the low tone was now producing the high tone!

The illusion may also be obtained with two loudspeakers situated side by side, facing the listener. This shows that highly specific regions of auditory space must be involved. The effect is obtained in certain nonanechoic environments, but not others. Although this has not been documented formally, it appears that the illusion is weaker in more reverberant environments. In this context, the reader may wish to perform the following simple experiment (suggested by Richard Gregory during a visit to our laboratory). Listen to **Sound Demonstration 1** with earphones placed correctly, and then slowly remove the earphones, bringing them out in front of you. If you obtain a clear and consistent octave illusion in the first place, you will probably find that the earphones can be removed a considerable distance before the illusion disappears. Notice also a hysteresis effect: once the illusion is lost, it becomes necessary to place the earphones considerably closer before it can be retrieved.

Does the octave illusion also occur with complex tones? Recently David Wessel at the Institut de Recherche et Coordination Acoustique-Musique (IRCAM) created an analogous sound pattern with computer-synthesized musical intrument tones. This is presented in **Sound Demonstration 2**. As can be heard, an analogous illusion is indeed obtained.

What happens when the alternating tones are not in octave relation? **Sound Demonstration 3** presents such a pattern, employing tones at 488 Hz (B_4) and 581 Hz (D_5), so forming a minor third. As can be heard, the percept is very different, though an illusion is still obtained.

1.2 Parametric Studies

A number of experiments have been performed to determine the boundary conditions for the effects giving rise to the octave illusion, and to investigate their behavior under parametric manipulation. As described earlier, the illusion is based on two effects which may be experimentally dissociated. The first is the perception of the frequencies presented to one side of space, the frequencies presented to the other side being suppressed. This is referred to as the *ear dominance effect* (though *side dominance* would probably have been a better term). The second is the localization of each tone to the source of the higher frequency signal, regardless of whether the higher or the lower frequency is in fact perceived. This is referred to as the *localization by frequency effect*.

1.3 Apparatus

Tones were generated as sine waves by two Wavetek function generators controlled by a PDP-8 computer. The output was passed through a Crown amplifier and presented to subjects through matched headphones (Grason–Stadler model TDH-49). Subjects were seated in sound-insulated booths. When the tones followed each other without pause, there were no voltage jumps at the frequency transitions, and the voltage slope did not change sign at the transitions. This was to minimize transients. (Transients degrade the illusion, for reasons which become clear below.)

1.4 The Ear Dominance Effect

One characteristic of the pattern giving rise to the octave illusion is that the frequencies which emanate from one side of space are always identical to the frequencies which have just emanated from the other side of space. This leads to the conjecture that the ear dominance effect depends on such a sequential relationship Several experiments were performed to verify this conjecture [5].

The first experiment made comparison between two conditions. These employed the basic patterns shown in Fig. 4(a). In condition 1 a single chord was repeatedly presented, whose components formed an octave, and alternated from ear to ear such that when the high tone was in one ear the low tone was in the other ear. Note that here the two ears received the same frequencies in succession. The sequence presented to the right ear began with the high tone and ended with the low tone on half the trials, and this order was reversed on the other half. Listeners judged whether the pattern began with the high tone and ended with the low tone, or whether it began with the low tone and ended with the high tone. It was thus possible to infer which ear was being followed for pitch.

In condition 2 two chords were presented in alternation. The first chord formed an octave and the second a minor third, and the entire four-tone combination constituted a major triad. Note that here the two ears did not receive the same frequencies in succession. The right ear received the upper component of the first chord and the lower component of the second chord on half the trials, and this order was reversed on the other half. The same judgments were required as in condition 1.

In both conditions the amplitude relationships between the tones delivered to the two ears were systematically varied in the range between 70 and 85 db sound pressure level, and plots were made of the extent to which each ear was followed as a function of these amplitude relationships. These plots are displayed in Fig. 4(b). It can be seen that in condition 1 the frequencies presented to one ear were followed until a critical level of amplitude relationship was reached, and the other ear was followed beyond this level. In contrast, there was no following on the basis of ear of input in condition 2, even when the signals presented to the two ears differed substantially in amplitude. Instead, all listeners showed consistent following on the basis of frequency proximity. This result is as expected from the above conjecture.

The second experiment again made comparison between two conditions. These employed the basic patterns shown in Fig. 5(a). In condition 1 a single chord whose components formed an octave was presented twice in succession such that one ear received first the high tone and then the low tone, while simultaneously the other ear received first the low tone and then the high tone. The identical frequencies were employed throughout this condition. In condition 2 the basic pattern comprised two chords, each of which formed an octave, but which were composed of different frequencies. Two chord pairs were employed, and trials employing these different pairs occurred in strict alternation. Thus any given chord was repeated only after a substantial time period during which several other chords had been interpolated. Listeners judged on each condition whether they heard a high–low pattern, or a low–high pattern, and from these judgments it was inferred which ear was being followed for pitch.

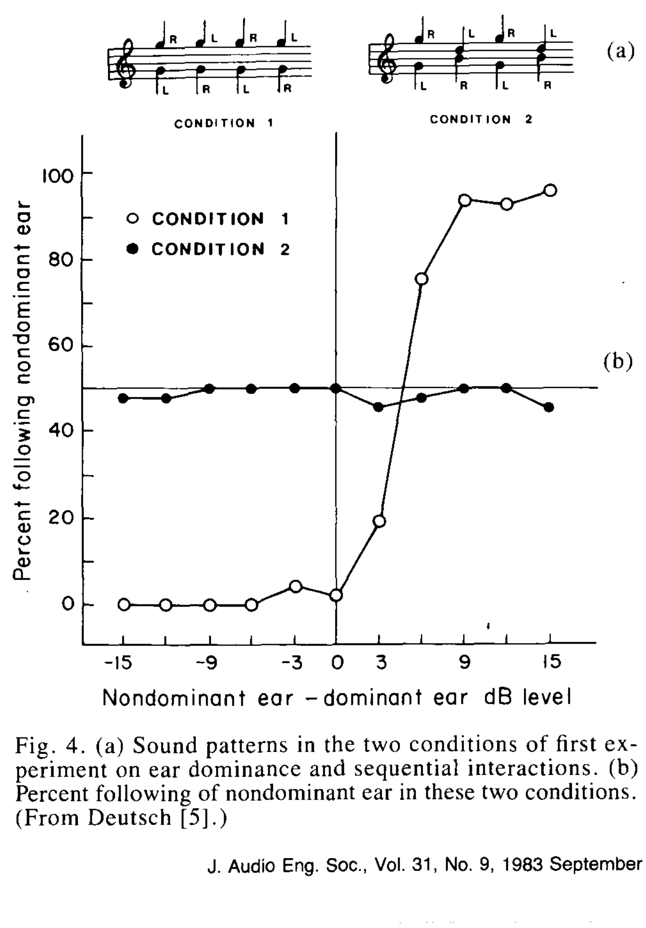

Fig. 4. (a) Sound patterns in the two conditions of first experiment on ear dominance and sequential interactions. (b) Percent following of nondominant ear in these two conditions. (From Deutsch [5].)

Fig. 5(b) shows, for this experiment, the extent to which each ear was followed as a function of the amplitude relationship between the signals at the two ears. In condition 1, following was clearly on the basis of ear of input. In contrast, such following did not occur in condition 2, even when there were substantial amplitude differences between the signals arriving at the two ears. Instead, these sequences were uniformly followed on the basis of overall contour: patterns of response always indicated an ascending sequence when the second chord was higher than the first, and a descending sequence when the second chord was lower than the first. Such a result held even when the signals at the two ears differed substantially in amplitude.

Thus in both experiments 1 and 2, when the same frequencies emanated in succession from different spatial locations, the ear dominance effect always occurred. However, when this sequential relationship did not hold, following on the basis of frequency range always occurred instead. It is noteworthy that relative amplitude turned out not to be an important factor in either experiment. Following by frequency range occurred despite large amplitude differences between the signals arriving at the two ears. When following was by ear of input, a shift from following one ear to the other did not occur at the point where the amplitude balance shifted from one ear to the other, but at some different level of amplitude relationship.

A further question that arises is whether the absence of ear dominance found in condition 2 of these experiments was due to the delay between successive presentations of the same frequencies to the two ears, or to the interpolation of tones of different frequencies. A third experiment examined the effect of a single interpolated tone on the strength of ear dominance, holding the delay factor constant. This experiment again consisted of two conditions, and the patterns employed are shown in Fig. 6(a). It can be seen that these patterns were identical, except that in condition 2 a single tone was interpolated between the two presentations of the identical chord. This tone was presented simultaneously to both ears, and the listeners were asked to ignore it. In both conditions the listeners judged whether the two-chord pattern was of the high–low type or the low–high type. The results are shown in Fig. 6(b). It can be seen that the strength of ear dominance was significantly reduced by the single interpolated tone.

A fourth experiment was performed to investigate the effect of temporal delay on the strength of ear dominance [6]. The experiment employed four conditions. In condition 1, 20 chords, each 250 milliseconds in duration, were presented in succession with no gaps between them. Condition 2 was identical to condition 1, except that only two such chords were presented. Condition 3 was identical to condition 2, except that a gap of 2750 milliseconds was interpolated between members of each pair of chords. Condition 4 was identical to Condition 3, except that each chord was 3 seconds in duration, and there was no gap between the chords. Thus in conditions 3 and 4 the onsets of successive chords were separated by identical time intervals, but the durations of the chords themselves differed substantially.

The results of the experiment are displayed in Fig.

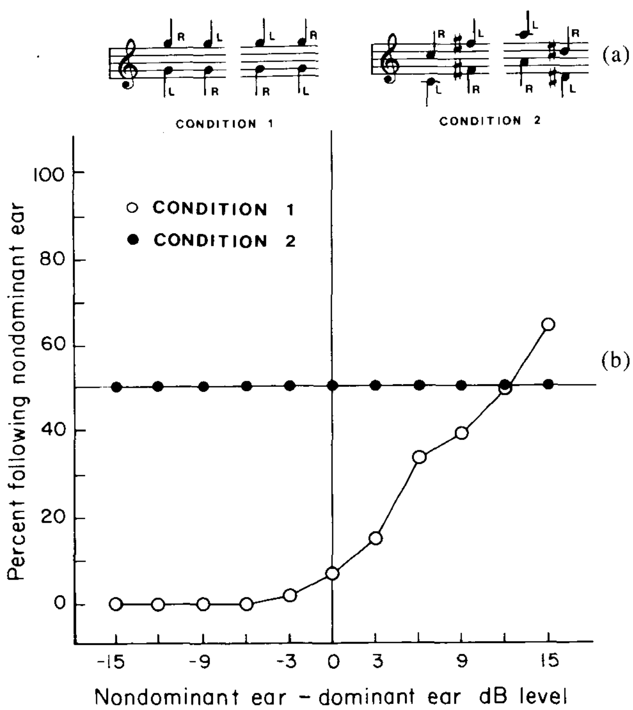

Fig. 5. (a) Sound patterns used in the two conditions of second experiment on ear dominance and sequential interactions. (b) Percent following of nondominant ear in these two conditions. (From Deutsch [5].)

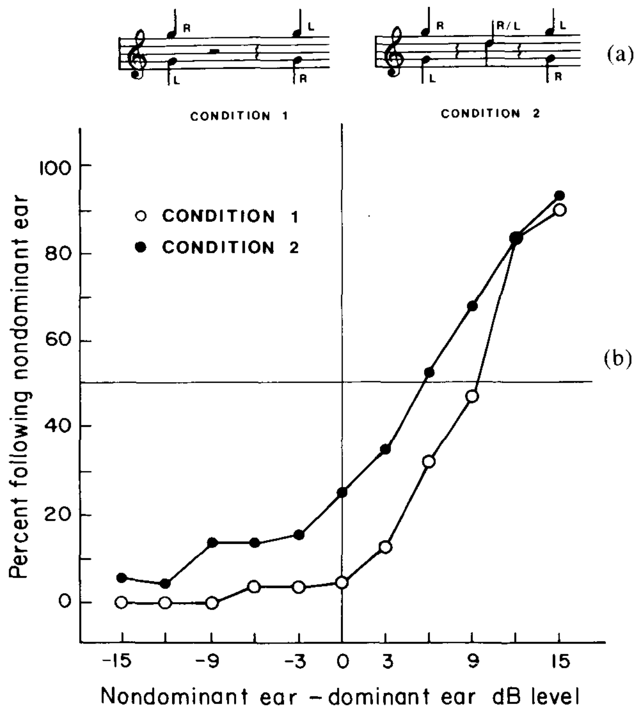

Fig. 6. (a) Sound patterns used in the two conditions of third experiment on ear dominance and sequential interactions. (b) Percent following of nondominant ear in the two conditions. (From Deutsch [5].)

7. They demonstrate that the strength of ear dominance was reduced by increasing time intervals between onsets of successive chords. However, it made no difference whether these increased time intervals were produced by lengthening the durations of the chords, or by interpolating gaps between them. **Sound Demonstration 4** presents an octave illusion pattern in which the chord durations are varied. The pattern begins with a chord duration of 250 milliseconds, speeds up to a chord duration of 50 milliseconds, and then slows down to a chord duration of 4 seconds. At the longest chord durations the degradation of the illusion becomes very clear perceptually.

Why should a brain mechanism which produces such a strange set of phenomena be useful to us? One might suggest that it is of value in enabling us to follow new, ongoing auditory information with a minimum of interference from echoes and reverberation. Under natural conditions, when we hear the identical frequency emanating in close temporal succession from two regions of space, the second occurrence is in all probability an echo. This interpretation becomes less probable as the delay between two such occurrences is lengthened. Further, if different frequencies are interpolated between two such occurrences of the same frequency, such an interpretation also becomes less probable. If this line of reasoning is correct, then the ear dominance effect is based on a mechanism that serves to counteract misleading effects in our auditory environment. Another such mechanism underlies the precedence effect [11], [12]. Here a single auditory image may be obtained when the same frequency emanates from two different locations with onset disparities of less than around 70 milliseconds.

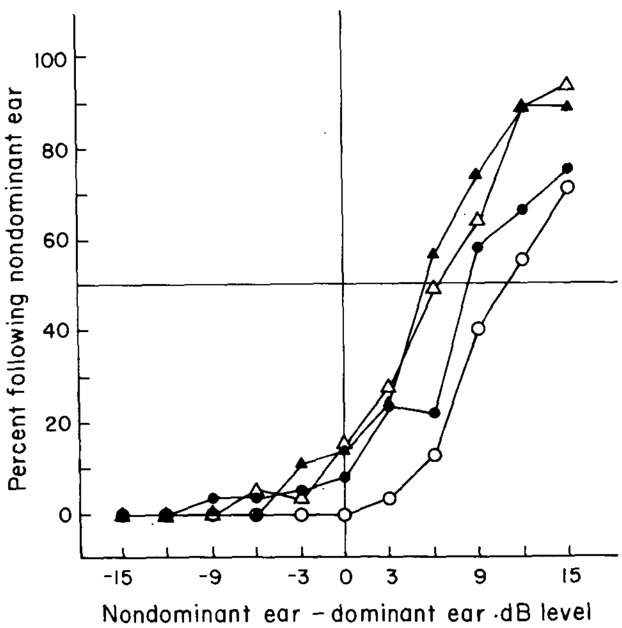

Fig. 7. Percent following of nondominant ear in experiment on the effect of temporal variations on ear dominance. ○—condition 1; ●—condition 2; △—condition 3; ▲—condition 4. (From Deutsch [6].)

1.5 The Localization by Frequency Effect

So far we have been considering only one component of the octave illusion, namely, the mechanism which determines what pattern of pitches is perceived. But we can recall that patterns of localization obey a different rule: each tone is localized to the source of the higher frequency signal, regardless of whether the higher or the lower frequency is in fact perceived.

We can study this effect also as a function of amplitude relationships between simultaneous tones. In this case we can vary the amplitude of the high tone relative to the low tone in each sequence [4]. When we do this, we find that with long repetitive sequences a localization toward the higher frequency signal occurs even when the lower frequency is substantially higher in amplitude. But with short sequences consisting of only two chords, localization patterns more closely follow patterns of relative loudness (Fig. 8). This is one important parametric difference between the two effects. Although the ear dominance effect is stronger with tones in rapid repetitive sequence, it still occurs when delays of several seconds intervene between onsets of successive tones. However, the localization effect is much more critically dependent on rapid sequencing. We should also note that individual listeners differ in the relative strengths of these two effects in the face of amplitude variations. Some listeners show strong localization effects and weak ear dominance effects, while others show weak localization effects and strong ear dominance effects.

The localization by frequency effect is also very robust in terms of onset and offset disparities between the tones at the two ears, provided that long repetitive sequences are employed. Varying the onset of the low tone relative to the high tone within a 5-millisecond range in either direction does not affect the strength of localization to the source of the high tone [6].

Why should this localization effect be useful to us? An explanation may be advanced in terms of head shadow effects. When a complex sound is presented in our everyday environment, there is a considerable difference in the relative strengths of the partials arriving at the listener's ears. For example, if the sound is presented to the listener's right, partial components above approximately 500 Hz that arrive at the right ear are considerably stronger than those arriving at the left ear. If the brain treats the octave illusion pattern as a fundamental together with its first partial, then we would expect the signal to be interpreted as coming from the listener's right, that is, from the side receiving the higher frequency component.

2 THE SCALE ILLUSION

What happens if we present a pattern that consists of more than two frequencies? Let us consider the pattern which is shown in Fig. 9(a) and presented in **Sound Demonstration 5**. This constitutes a major scale, with successive tones alternating from ear to ear. The scale is presented simultaneously in both ascending form and descending form, such that when a tone from the ascending scale is in one ear, a tone from the descending scale is in the other ear [13].

This pattern also gives rise to a number of illusions, which differ from one listener to another. The illusion most commonly obtained is represented in Fig. 9(b). It can be seen that this consists of the correct sequences of pitches, but heard as two separate melodies that move in contrary motion. Futhermore, the higher tones all appear to be coming from the right earphone, and the lower tones from the left. When the earphone positions are reversed, this percept is still maintained. So it appears that the earphone which had been producing the higher tones is now producing the lower tones, and that the earphone which had been producing the lower tones is now producing the higher tones.

Other listeners obtain different percepts. Some hear all the higher tones as in the left earphone, and all the lower tones as in the right, with earphones placed both ways. For yet other listeners, when the position of the earphones is reversed, the locations of the higher and lower tones reverse also. The two handedness populations again exhibit different patterns of localization for the higher and lower tones. Right-handers tend strongly to hear the higher tones on the right and the lower tones on the left; but this is not true of left-handers [13]. A number of listeners perceive instead only the higher tones, and little or nothing of the lower tones. The two handedness populations differ on this measure also: the majority of the right-handers perceive all the tones, but only about half of the left-handers do so [13].

How do these findings relate to normal listening? Butler [14] performed an experiment to determine whether the scale illusion is also obtained when the sounds are presented through loudspeakers rather than earphones. A second concern was whether the illusion is confined to sine-wave tones, or whether patterns of complex tones might also be perceived as in different locations, determined by the ranges of their fundamental frequencies. A further question was how such factors as unequal tone durations, improperly timed attacks, unequal attack strengths, and so on, all of which exist in natural musical sounds, might affect the illusion. Piano tones were therefore employed in some conditions, and differences in timbre and loudness were sometimes introduced between the tones presented through the different loudspeakers. Response patterns were found to be very similar under these conditions: higher and lower melodic lines were still perceived, each as coming from a different loudspeaker. Further, when differences in timbre were introduced, a new tone quality was perceived, but as though coming simultaneously from both loudspeakers. So not only were the pitches of the tones spatially reorganized, but their timbres were reorganized also.

Such effects can be found in live musical performance. At the beginning of the last movement of Tchaikovsky's Sixth Symphony there is a passage in which the theme and accompaniment are each distributed between two violin parts, as shown in Fig. 10. However, the theme is heard as coming from one set of instruments, and the accompaniment as from the other [15]. The author has experienced this illusion very powerfully when listening to the UCSD Symphony orchestra arranged in nineteenth-century fashion (that is, with the first violins on one side and the second violins on the other side).

How should this illusion be explained? Our auditory environment is very complex, and the assignment of sounds to their sources is made difficult by the presence of echoes and reverberation. So when both ears receive input simultaneously, we cannot rely on first-order localization cues alone to determine which elements of

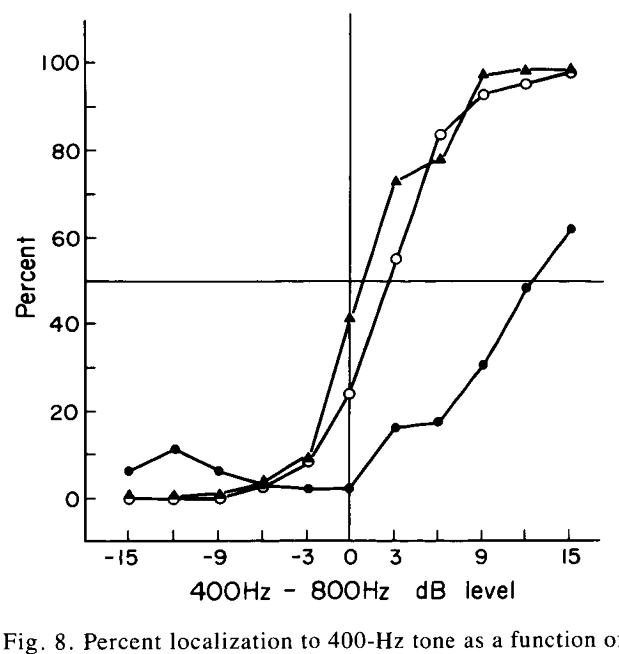

Fig. 8. Percent localization to 400-Hz tone as a function of amplitude differences between 400-Hz and 800-Hz tones. ●—long repetitive sequences; ○—sequences consisting of only two chords; ▲—percent judgment of 400-Hz tone as louder than 800-Hz tone as a function of amplitude differences; long repetitive sequences. (From Deutsch [4].)

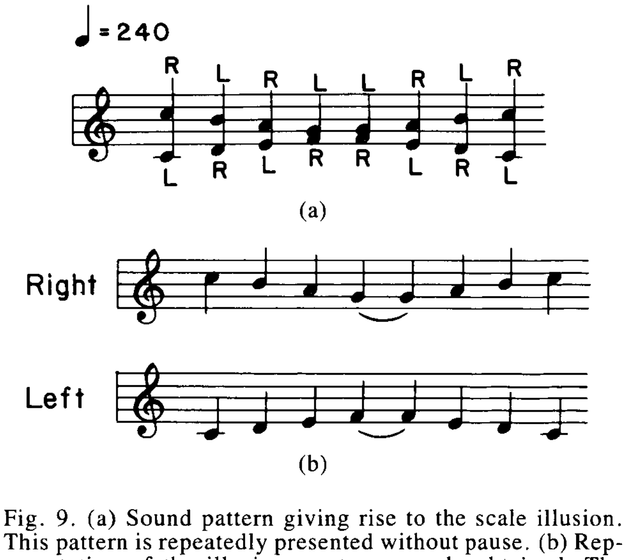

Fig. 9. (a) Sound pattern giving rise to the scale illusion. This pattern is repeatedly presented without pause. (b) Representation of the illusion most commonly obtained. The higher tones appear to be in the right ear, and the lower tones in the left ear. (From Deutsch [13].)

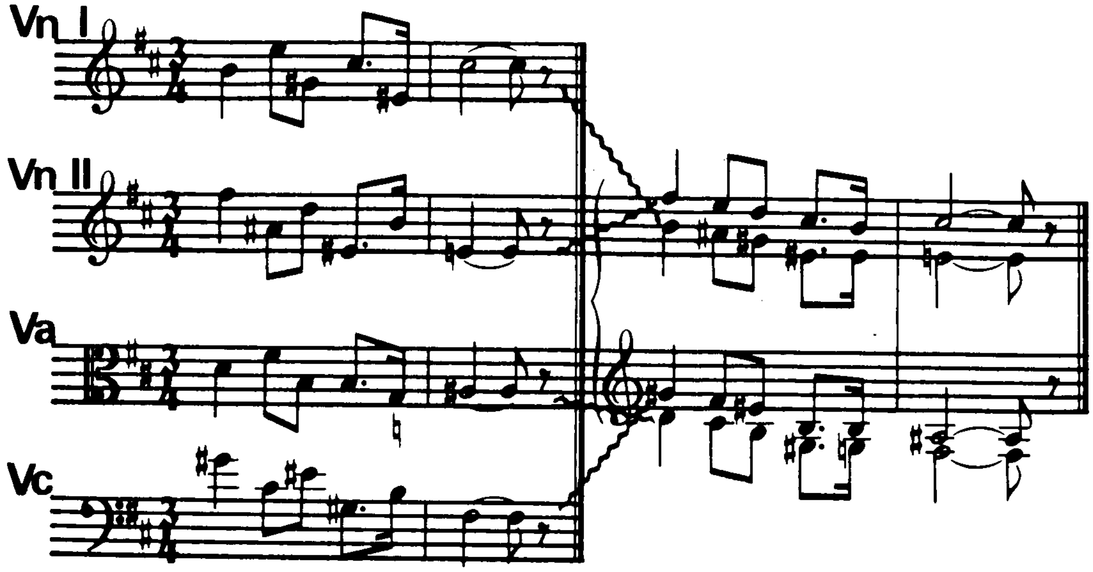

Fig. 10. Passage from final movement of Tchaikovsky's Sixth Symphony. The combination of the violin I and violin II parts results in the percept shown in the upper right. The combination of the viola and violoncello parts results in the percept shown in the lower right. (From Butler [14].)

the total spectrum should be assigned to which source. Other factors must also serve as cues. One such factor is the similarity of frequency spectrum: sounds which are similar are likely to be coming from the same source, and sounds which are dissimilar from different sources. It therefore becomes reasonable for the listener to assume that tones in one frequency range are coming from one source, and that tones in a different frequency range are coming from a different source. The tones are therefore spatially reorganized on the basis of this assumption (see also Warren, this issue, p. 623).

3 THE DRONE EFFECT

So far we have been considering situations in which the signals arriving at the two ears or from the two locations are simultaneous. What happens when this does not hold?

To investigate this issue, we presented listeners with two simple melodic patterns, one on each trial, and their task was to identify which one they had heard. The two patterns are shown in Fig. 11 and are presented in **Sound Demonstration 6**. The patterns were generated under four different conditions. These are illustrated in Fig. 12, which also displays the error rates in these conditions.

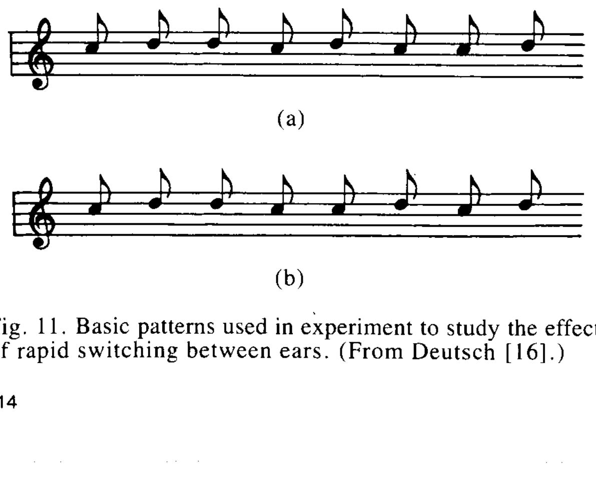

Fig. 11. Basic patterns used in experiment to study the effects of rapid switching between ears. (From Deutsch [16].)

In condition 1 all tones were presented simultaneously to both ears, and it can be seen that melody identification was very good. In condition 2 the components of the melody were distributed in quasi-random fashion between the ears. It can be seen that identification was in contrast very poor. Subjectively one feels compelled in this condition to attend to the pattern coming from one earphone at a time, and it is very difficult to integrate the two patterns into a single perceptual stream. An example of such a sequence is presented in **Sound Demonstration 7**.

Condition 3 was exactly as condition 2, except that the melody was accompanied by a drone. Whenever a tone from the melody was presented to the right ear, the drone was presented to the left ear; and whenever a tone from the melody was presented to the left ear, the drone was presented to the right ear. Thus although the tones from the melody were still switching between ears, both ears always received input simultaneously. It can be seen that identification of the melody was greatly improved in this condition. An example of such a sequence is presented in **Sound Demonstration 8**, and you will hear the melody emerging perceptually.

In condition 4 the melody was again accompanied by a drone. However, the drone was always presented to the same ear as the tone from the melody, rather than to the opposite ear. Thus input was again to only one ear at a time. It can be seen that identification of the melody was very poor [16].

We can conclude that with signals coming from different spatial locations, temporal relationships between them are important determinants of grouping. When the two ears are stimulated simultaneously, grouping by frequency range is easy, so that identification of the

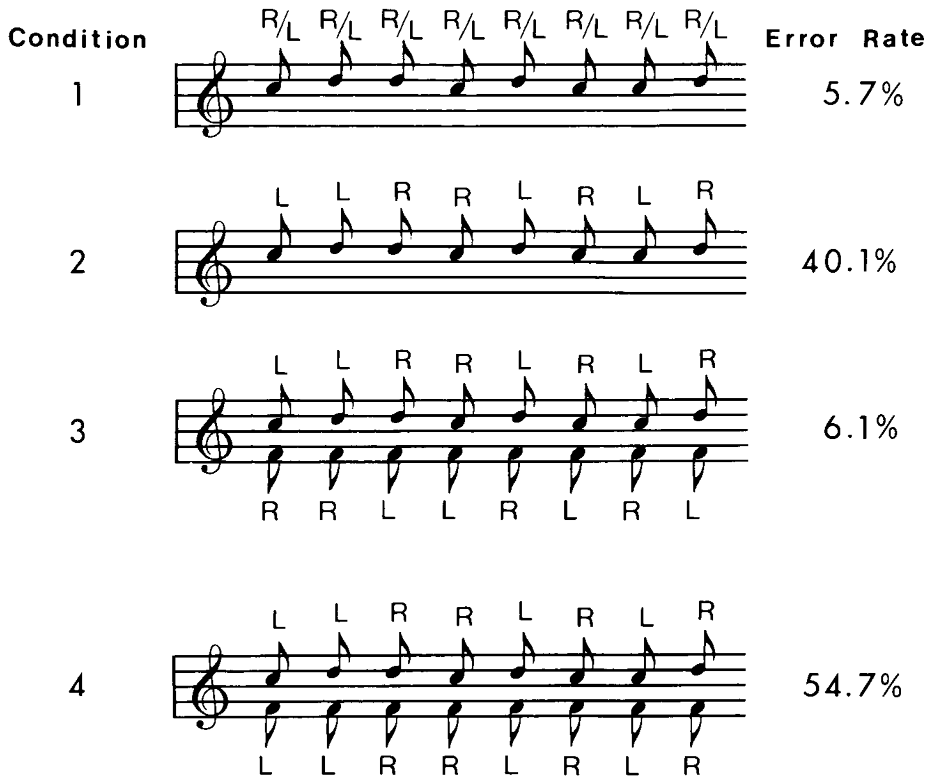

Fig. 12. Examples of distributions between ears of melodic pattern and drone in different conditions of experiment. See text for details. (From Deutsch [16].)

melodies readily occurs. But when the signals to the two ears are clearly separated in time, grouping by spatial location is so powerful as to virtually abolish the listener's ability to integrate the signals into a single perceptual stream (see also [17], [18]).

We may next ask what happens in this intermediate case, where the signals which arrive at the two ears are not strictly simultaneous, but rather overlapping in time. To investigate this issue, we introduced different conditions of temporal asynchrony between the components of the melody and the drone. These conditions and their error rates are shown in Table 1. It can be seen that this intermediate case produced intermediate results. Identification of the melody in the presence of the drone when the two were asynchronous was better than when there was no drone; however, it was worse than when the melody and drone were synchronous [16].

Why should such a pattern of results occur? The relationships between waveform envelopes of sound signals are an important indicator of whether these signals are emanating from the same source or from different sources [19]. We should therefore expect that the more clearly the signals arriving at the two ears are separated in time, the greater should be the tendency to treat them as emanating from different sources, and so the greater should be the tendency to group them by spatial location. If such grouping is sufficiently powerful, it should prevent us from forming perceptual linkages between signals emanating from different sources. Clearly it is necessary, in performing auditory shape analyses, that we not link together the components of different signals, or we should end up with nonsensical percepts.

To relate these findings to normal music listening, let us consider a passage from the composer Berlioz, who argued that disposition of instruments in space should be regarded as an essential part of a composition. In his *Treatise on Instrumentation* [20] 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 in 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.

The experiments just described demonstrate that the spatial disposition of instruments should indeed have profound effects on how music is perceived. When two sets of instruments are spatially separated, and in addition a clear temporal separation exists between the sounds produced by these instruments, the resulting perceptual dissociation may be so pronounced as to prevent the listener from integrating from different sounds into a single coherent stream. Yet a certain

Table 1. Error rates in melody identification depending on temporal relationships between melody and drone.*

Condition	Error Rate (%)
Melody binaural; no drone.	3.6
Component tones of melody switch between ears; no drone.	37.0
Component tones of melody switch between ears; drone presented in synchrony to opposite ear.	10.4
Component tones of melody switch between ears; drone presented asynchronously to opposite ear.	18.9
Tone in left ear always leads tone in right.	19.8
Tone in right ear always leads tone in left.	21.4
Melody component always leads drone.	20.8
Drone always leads melody component.	13.5

* From Deutsch [16].

amount of temporal overlap between the different instruments will facilitate such integration. However, there is a trade-off; the greater the temporal overlap, the greater will be the loss of spatial distinctiveness, and as simultaneity is approached, spatial illusions will tend to occur.

These experiments provide good examples of what Helmholtz termed "unconscious inference" in perception. Here a perceptual reorganization occurs even when the listener has conscious knowledge of what the sound pattern really is. There are, however, circumstances in which conscious knowledge will drastically alter the listener's percept. This can be shown, for example, in the case of some octave equivalence effects. If you present a listener with a well-known melody such that its component notes are displaced at random to different octaves, the listener will be quite unable to recognize it in the absence of any clues on which to base a hypothesis (such as rhythm, contour, and so on). However, if you give the listener the name of the melody beforehand, the problem disappears. This effect, which we originally demonstrated with the tune "Yankee Doodle" [21], is very striking. **Sound Demonstration 9** presents another well-known melody under octave displacement. The reader will find it instructive to listen to this demonstration before proceeding. Now listen to **Sound Demonstration 10**, which presents the untransformed version, and then again to **Sound Demonstration 9**. You will find that the melody is now quite easy to follow (see also [22]).

4 SPATIAL DISPOSITION AND ACCURACY OF PITCH PERCEPTION

Finally we shall return to the question of how the perception of two simultaneous sequences may be affected by whether the higher is to the right and the lower to the left, or whether this configuration is reversed. We can recall that, in the octave and scale illusions, right-handers tend strongly to perceive high tones as on the right and low tones as on the left, regardless of their true locations. This means that combinations of the high-right/low-left type tend to be accurately localized, and that combinations of the high-left/low-right type tend to be mislocalized. Other recent studies have confirmed this in a more general setting.

We can further ask whether such spatial considerations influence the accuracy of pitch perception also. To find out, we presented listeners with patterns such as shown in Fig. 13. It can be seen that this may be described as six two-tone chords, each of which is either of the high-right/low-left type or of the high-left/low-right type. The listeners, who were musically trained, notated all that they could of each sequence (without concerning themselves about the locations of the tones). Performance was significantly better for high-right/low-left chords than for high-left/low-right chords. We can see, therefore, that there is a substantial advantage to a spatial disposition in which high tones are to the right and low tones to the left, compared with the reverse.

This leads us to the following line of reasoning, with which we shall conclude. In general the seating arrangement of present-day orchestras is such that, from the performer's point of view, instruments with high registers tend to be to the right, and instruments with low registers to the left. Fig. 14 displays, as an example, a seating arrangement of the Chicago Symphony Orchestra [23]. So taking the strings, the first violins are to the right of the second violins, which are to the right of violas, which are to the right of the cellos, which are to the right of the basses. Notice also the disposition of the brasses: the trumpets are to the right of the trombones, which are to the right of the tuba, and so on. Since a major consideration in arranging the seating

Fig. 13. Example of sequence used to determine accuracy of pitch perception for high-right/low-left chords, compared with high-left/low-right chords.

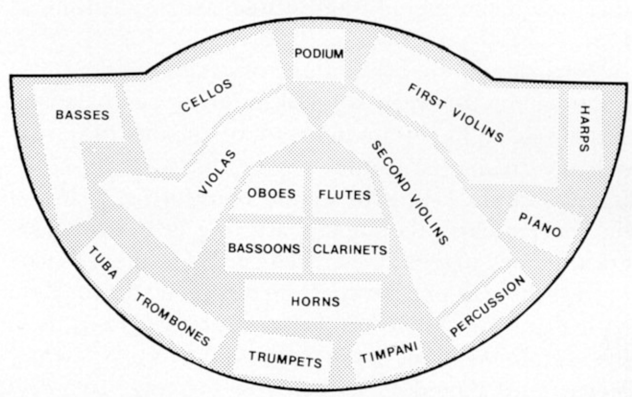

Fig. 14. Chicago Symphony seating plan, viewed from the orchestra. (Adapted from Machlis [23, p. 44].)

for orchestras is how well the performers will be able to play together, it is reasonable to assume that this high-right/low-left disposition has evolved by trial and error because it is conducive to optimal performance. However, it can be seen that, from the viewpoint of the audience, this arrangement is mirror-image reversed, so that instruments with high registers now tend to be to the left, and those with low registers to the right. So from the audience's point of view, this spatial disposition is such as to cause perceptual difficulties. The worst casualties here are the instruments with low registers that are placed to the right. These should tend to be poorly perceived and localized. Such perceptual effects are likely to have a significant influence on the evaluation of acoustic environments by the listener.

5 ACKNOWLEDGMENT

Preparation of this paper and most of the sound demonstrations was supported by United States Public Health Service under grant MH-21001. Sound Demonstration 2 was produced by Dr. David L. Wessel at the Institut de Recherche et Coordination Acoustique-Musique.

6 REFERENCES

[1] D. Deutsch, "An Auditory Illusion," *Nature (London)*, vol. 251, pp. 307–309 (1974).

[2] D. Deutsch, "Musical Illusions," *Sci. Am.*, vol. 233, pp. 92–104 (1975).

[3] D. Deutsch and P. L. Roll, "Separate 'what' and 'where' Decision Mechanisms in Processing a Dichotic Tonal Sequence," *J. Exp. Psychol., Hum. Percept. Perform.*, vol. 2, pp. 23–29 (1976).

[4] D. Deutsch, "Lateralization by Frequency for Repeating Sequences of Dichotic 400- and 800-Hz Tones," *J. Acoust. Soc. Am.*, vol. 63, pp. 184–186 (1978).

[5] D. Deutsch, "Ear Dominance and Sequential Interactions," *J. Acoust. Soc. Am.*, vol. 67, pp. 220–228 (1980).

[6] D. Deutsch, "The Octave Illusion and Auditory Perceptual Integration," in J. V. Tobias and E. D. Schubert, Eds., *Hearing Research and Theory*, vol. 1 (Academic Press, New York, 1981).

[7] O. L. Zangwill, *Cerebral Dominance and Its Relation to Psychological Function* (Oliver and Boyd, Edinburgh, 1960).

[8] A. R. Luria, *Traumatic Aphasia* (Mouton, The Hague, 1969).

[9] H. Hécaen and M. L. Albert, *Human Neuropsychology* (Wiley, New York, 1978).

[10] D. Deutsch, "The Octave Illusion in Relation to Handedness and Familial Handedness Background," *Neuropsychologia* (in press).

[11] H. Haas. "Über den Einfluss eines Einfachechos aud die Hörsamkeit von Sprache," *Acustica*, vol. 1, pp. 49–52, 1951; also "The Influence of a Single Echo on the Audibility of Speech," *J. Audio Eng. Soc.*, vol. 20, pp. 145–159 (1972 Mar.).

[12] H. Wallach, E. B. Newman, and M. R. Rosenzweig, "The Precedence Effect in Sound Localization," *Am. J. Psychol.*, vol. 62, pp. 315–336 (1949).

[13] D. Deutsch, "Two-Channel Listening to Musical Scales," *J. Acoust. Soc. Am.*, vol. 57, pp. 1156–1160 (1975).

[14] D. Butler, "A Further Study of Melodic Channeling," *Percept. Psychophys.* vol. 25, pp. 264–268 (1979).

[15] D. Butler, "Melodic Channeling in a Musical Environment," presented at the Research Symp. on the Psychology and Acoustics of Music, Kansas, 1979.

[16] D. Deutsch, "Binaural Integration of Melodic Patterns," *Percept. Psychophys.*, vol. 25, pp. 399–405 (1979).

[17] E. D. Schubert and C. D. Parker, "Addition to Cherry's Findings on Switching Speech between the Two Ears," *J. Acoust. Soc. Am.*, vol. 27, pp. 792–794 (1956).

[18] T. Judd, "Comments on Deutsch's Musical Scale Illusion," *Percept. Psychophys.*, vol. 26, pp. 85–92 (1979).

[19] J. V. Tobias, "Curious Binaural Phenomena," in J. V. Tobias, Ed., *Foundations of Modern Auditory Theory*, vol. 2 (Academic Press, New York, 1972).

[20] H. Berlioz, in *Treatise on Instrumentation*, I. R. Strauss, Ed., and T. Front, transl. (E. F. Kalmus, 1948).

[21] D. Deutsch, "Octave Generalization and Tune Recognition," *Percept. Psychophys.*, vol. 11, pp. 411–412 (1972).

[22] D. Deutsch, "The Processing of Pitch Combinations," in D. Deutsch, Ed., *The Psychology of Music* (Academic Press, New York, 1982).

[23] J. Machlis, *The Enjoyment of Music*, 4th ed. (Norton, New York, 1977).

THE AUTHOR

Diana Deutsch was born in London, England, and obtained a first class honors B.A. in psychology, philosophy and physiology from Oxford University in 1959. She obtained her Ph.D. in psychology from the University of California, San Diego, in 1970, and has remained there on the research faculty.

She is an active member of the Acoustical Society of America, having served on the Nominating Committee and the Technical Committee on Psychological and Physiological Acoustics. She is currently on the Technical Committee on Musical Acoustics and the Committee on Education in Acoustics. A member of the Advisory Council of the International Association for the Study of Attention and Performance, she is also founding editor of the journal *Music Perception* (University of California Press).

She is co-author (with J. A. Deutsch) of *Physiological Psychology* (Dorsey Press, 1966; 2nd Edition, 1973), and is editor of *The Psychology of Music* (Academic Press, 1982). Her main research contributions are in the fields of auditory and musical perception.

AUDITORY ILLUSIONS AND AUDIO

Sound Demonstrations Disk A

NOTE: The sound demonstrations should be played stereophonically through earphones. Each set of demonstrations is preceded by a "zero-level" stereophonic calibration tone. The volume should be set for comfortable listening on headphones such that the amplitude of the calibration tone is equal in both ears.

SIDE 1

Auditory Illusions, Handedness, and the Spatial Environment

Diana Deutsch
Department of Psychology, University of California–San Diego, La Jolla, CA 92093, USA

"Zero-level" calibration tone

Sound demonstrations:

1. The octave illusion. A sequence of sine wave tones is presented, whose frequencies alternate between 400 and 800 Hz. The identical sequence is presented at equal amplitude to the two ears; however, when the right ear receives 400 Hz, the left ear receives 800 Hz; and vice versa. This pattern gives rise to a number of illusions. The most common percept is of a single high tone in one ear alternating with a single low tone in the other ear. The apparent locations of the high and low tones generally remain fixed when the earphones are placed in reverse position. Statistical differences exist between righthanders and lefthanders in terms of how this pattern is perceived.

2. A variant of the octave illusion, produced by computer-synthesized musical instrument tones. This demonstration was generated by David Wessel at the Institut de Recherche et Coordination Acoustique-Musique.

3. A sequence of tones whose parameters are the same as in Sound Demonstration 1, except that the alternating frequencies are 488 Hz and 581 Hz, forming a minor third. The pattern that is heard is generally more complex than the octave illusion.

4. The octave illusion with changing tone durations. The sequence begins with a tone duration of 250 milliseconds, speeds up to a tone duration of 50 milliseconds, and then slows down to a tone duration of four seconds.

5. The scale illusion. A major scale is presented 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 one ear, a tone from the descending scale is in the other ear. This pattern is repeatedly presented without pause. The most common illusion is that of two melodic lines, one higher and the other lower, that move in contrary motion. The higher tones are often heard as in one ear, and the lower tones as in the other ear. Often the apparent locations of the high and low tones remain fixed when the earphone positions are reversed. As with the octave illusion, statistical differences exist between righthanders and lefthanders in terms of how this pattern is perceived.

6. The two melodic patterns used in the experiment to examine the effect of rapid switching between ears. The listener's task is to identify which of the two patterns is heard under different conditions of presentation. Here the tones are presented binaurally.

7. One of the melodic patterns in Sound Demonstration 6, presented with tones switching between ears. Identification is generally difficult under this condition.

8. One of the patterns in Sound Demonstration 6, again presented with tones switching between ears, but with the addition of a contralateral drone. Whenever a tone from the melody is in the right ear, the drone is in the left ear, and whenever a tone from the melody is in the left ear, the drone is in the right ear. Identification is generally much easier in this condition.

9. The "Yankee Doodle" effect. A well-known melody is presented, with the tones displaced to different octaves. A melody is generally very difficult to identify under such conditions, unless the listener is given clues on which to base a hypothesis.

10. The same melody as in Sound Demonstration 9, but without octave displacement. After listening to this demonstration, it should be easy to follow the melody in Sound Demonstration 9.

(Total time 4:36)

SIDE 2

Auditory Illusions and Their Relation to Mechanisms Enhancing Accuracy of Perception

Richard M. Warren
Department of Psychology, University of Wisconsin-Milwaukee, WI 53201, USA

"Zero-level" calibration tone

Sound demonstrations:

1. a) Phonemic restorations produced by a cough completely replacing a speech sound (played three times); b) Absence of phonemic restoration when a silent gap is substituted for the cough in 1a) (played twice).

2. Homophonic temporal induction of broadband noise alternating between two intensity levels each level lasting: a) 300 milliseconds; b) 2 seconds.

3. Heterophonic temporal induction of one-third-octave band noise alternated with a louder 2-octave band noise spanning the frequency range of the narrow-band noise, each noise band lasting: a) 300 milliseconds; b) 2 seconds.

4. Spectral requirements for heterophonic temporal induction. Narrow-band noise (one-third-octave) is alternated with louder broader band noise (2-octave), each noise lasting 300 milliseconds. The center frequency of the narrow band noise initially lies within the frequency range of the other noise (illusory continuity heard), then glides to a center frequency below this range (illusory continuity lost), and then returns to original center frequency (continuity heard again).

5. Contralateral induction and its spectral rules. A monaural tone exchanges sides each 500 milliseconds with a louder contralateral one-third-octave noise band centered at 1000 Hz. The tonal frequency initially lies within the noise-band frequency range, and contralateral induction is heard. The tone then glides below this range and contralateral induction is lost. When the tonal frequency glides back to its original value, contralateral induction returns.

6. Verbal transformations of a word repeated diotically.

7. Verbal transformations of a word repeated dichotically (the repeated word is heard with an interaural delay of half the word's period).

(Total time 5:24)