4

Grouping Mechanisms in Music

Diana Deutsch

I.	Introduction	99
II.	Grouping Principles	100
III.	Two-Channel Listening to Melodic Sequences	101
	A. The Scale Illusion	102
	B. Temporal Relationships as Determinants of Grouping	104
	C. The Octave Illusion	108
	D. Handedness Correlates	114
	E. Melody Perception from Phase-Shifted Tones	116
	F. Discussion	117
IV.	Channeling of Rapid Sequences of Single Tones	118
	A. Grouping by Frequency Proximity	118
	B. Temporal Coherence as a Function of Frequency Proximity and Tempo	119
	C. Grouping by Frequency Proximity Builds with Repetition	120
	D. Grouping by Frequency Proximity and the Perception of Temporal Relationships	122
	E. Grouping by Good Continuation	124
	F. Grouping by Timbre	124
	G. Grouping by Amplitude	126
	H. Grouping by Temporal Proximity	126
	I. Perceptual Replacement of Missing Sounds	127
V.	Voluntary Attention	127
VI.	Conclusion	130
	References	130

I. INTRODUCTION

Music presents us with a complex, rapidly changing acoustic spectrum, often resulting from the superposition of sounds from many different sources. The primary task that our auditory system has to perform is to interpret this spectrum in terms of the behavior of external objects. This is analogous to the task performed by the visual

ISBN 0-12-213560-1

system when it interprets the mosaic of light patterns impinging on the retina in terms of objects producing them (Gregory, 1970; Sutherland, 1973). Such a view of perception as a process of "unconscious inference" was proposed in the last century by Helmholtz (see Hemholtz, 1925), and we shall see that many phenomena of music perception are readily interpretable in this light.

Issues concerning organizational processes in music divide themselves basically into two. The first is the following. Given that we are presented with a set of first-order acoustic elements, how are these combined so as to form separate groupings? If all first-order elements were indiscriminately linked together, auditory shape-recognition operations could not be performed. There must, therefore, exist a set of mechanisms that permits the formation of simultaneous and sequential linkages between certain elements, and inhibits the formation of such linkages between others. Simple mechanisms underlying such linkages are explored in the present chapter. Second, we may enquire into the ways in which higher order abstractions are derived from combinations of first-order elements so as to lead to perceptual equivalences and similarities. This issue is explored in the next chapter, and it is assumed that such abstractions also form bases for grouping.

II. GROUPING PRINCIPLES

There are two basic questions involved in considering the mechanisms involved in grouping musical stimuli into configurations. The first concerns the stimulus *attributes* along which grouping principles operate. When presented with a complex sequence, our auditory system may group stimuli according to some rule based on the frequencies of its components, on their amplitudes, on the spatial locations from which they emanate, or on the basis of some complex attribute such as timbre. As we shall see, all these attributes can function as bases for organization and grouping in music. Furthermore, the principles determining what attribute is followed for any given sequence are both complex and rigid. We shall see, for example, that with one type of sequence, organization on the basis of frequency uniformly occurs; yet given a slight change in this sequence, organization on the basis of spatial location may occur instead. Such differences in organization can be interpreted in terms of strategies most likely to lead to the correct conclusions in interpreting our environment.

Second, we may pose the following question: Assuming that organization takes place on the basis of some dimension such as frequency, what are the *principles* governing grouping along this dimension? The Gestalt psychologists proposed that we group stimuli into configurations on the basis of various simple principles (Wertheimer, 1923). One of these is the principle of Proximity, which states that nearer elements are grouped together in preference to those that are spaced farther apart. An example of this principle is shown on Figure 1A, where the closer dots appear to be grouped together in pairs. Another is the principle of Similarity, which is illustrated on Figure 1B. Here, configurations are formed out of like elements, so that we perceive one set of vertical rows formed by the filled circles and another set formed by the unfilled circles. A third is the principle of Good Continuation, which states that elements that

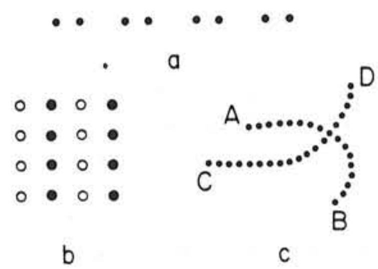

Fig. 1. Illustrations of the Gestalt principles of Proximity, Similarity, and Good Continuation.

follow each other in a given direction are perceived together. For instance, we perceptually group the dots in Figure 1C so as to form the two lines AB and CD. A fourth principle, known as Common Fate, states that elements which move in the same direction are perceived together.

It has been demonstrated that these principles are important determinants of grouping in visual arrays, and we shall see that this is true of music also. It seems reasonable to suppose—as argued by Gregory (1970), Sutherland (1973), Hochberg (1974), and Bregman (1978)—that grouping in conformity with such principles enables us to interpret our environment most effectively. To give some examples, in the case of vision, proximal elements are most likely to belong to the same object than elements that are spaced farther apart. The same line of reasoning holds for similar elements compared with those that are dissimilar. In the case of hearing, similar sounds are likely to be emanating from the same source and different sounds from different sources. A sound sequence that changes smoothly in frequency is likely to be emanating from a single source. Components of a complex sound spectrum that rise and fall in synchrony are also likely to be emanating from the same source.

One more point should be made before reviewing the experimental evidence. When we hear a tone, we attribute a fundamental pitch, a loudness, a timbre; and we hear the tone at a given location. Each tonal percept may therefore be described as a bundle of attribute values. If our perception is veridical, this bundle reflects the location and characteristics of the sound emitted. We shall see, however, that in situations where more than one tone is presented at a time, these bundles of attribute values may fragment and recombine in other ways, so that illusory percepts result. Perceptual grouping in music is therefore not simply a matter of linking different sets of stimuli together; rather it involves a process whereby these stimuli are fragmented into their separate attributes, followed by a process of perceptual synthesis in which the different attribute values are recombined.

III. TWO-CHANNEL LISTENING TO MELODIC SEQUENCES

The two-channel listening technique is particularly useful for studying organizational processes in music, since it enables different attributes to be set in opposition to each other as bases for grouping. For example, grouping by spatial location may be set in opposition to grouping by frequency or by amplitude. Similarly, different principles governing grouping along a given dimension may be set in opposition to each other; for example, the principle of Proximity may be opposed to the principle of Good Continuation. The experiments to be described show that the nature of the stimulus configuration critically determines what grouping principle is adopted, and indicate that there are complex and rigid rules of precedence for these principles.

A. The Scale Illusion

The configuration that produced the scale illusion is shown in Fig. 2A. It can be seen that this consisted of a major scale, presented simultaneously in both ascending and descending form. When a tone from the ascending scale was delivered to one ear, a tone from the descending scale was simultaneously delivered to the other ear, and successive tones in each scale alternated from ear to ear (Deutsch, 1975b).

This sequence was found to give rise to various illusory percepts. The majority of listeners perceived two melodic lines, a higher one and a lower one, that moved in contrary motion. Further, the higher tones all appeared to be emanating from one earphone and the lower tones from the other (Fig. 2B). When the earphone positions were reversed, there was often no corresponding change in the percept. So, it appeared to the listener that the earphone that had been producing the higher tones was now producing the lower tones, and that the earphone that had been producing the lower tones was now producing the higher tones. A minority of listeners heard instead only a single stream of four tones, corresponding to the higher tones in the sequence, and little or nothing of the rest of the sequence was perceived.

So in considering what stimulus *attribute* was here used as a basic for grouping, we find that organization by spatial location never occurred; rather, organization was

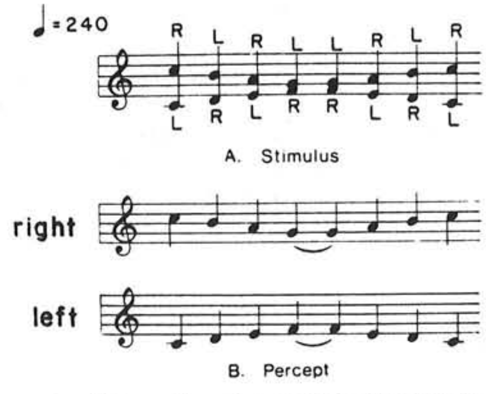

Fig. 2. (A) Representation of the configuration producing the scale illusion. This basic pattern was repetitively presented 10 times without pause. (B) Representation of the illusory percept most commonly obtained (from Deutsch, 1975b).

always on the basis of frequency. Second, in considering what grouping *principle* was adopted, we find that organization was always on the basis of frequency priximity. Listeners heard either two melodic lines, one corresponding to the higher tones and the other to the lower, or they heard the higher tones alone. No listener reported a full ascending or descending scale as a component of the sequence; so that grouping by Good Continuation never occurred.

Butler (1979a) has demonstrated that these findings may be extended to a broad range of musical situations. He presented the configuration shown in Fig. 2A to music students through spatially-separated loudspeakers, rather than earphones, in a free sound-field environment. The listeners notated separately the sequence that they heard as emanating from the speaker on their right and the sequence that they heard as emanating from the speaker on their left. In some conditions the stimuli were generated on a piano. Further, timbral and loudness differences were sometimes introduced between the stimuli presented through the different speakers. Butler found that despite these variations, virtually all responses reflected channeling by frequency proximity, so that higher and lower melodic lines were perceived, each apparently emanating from a different speaker. A further interesting finding was that when timbral differences were introduced between the tones presented through the two speakers, a new tone quality was perceived, but it seemed to be emanating simultaneously from both speakers. So, not only were the spatial locations of the tones perceptually rearranged in accordance with frequency proximity, but their timbres were rearranged also.

To determine whether these findings generalize to other melodic configurations, Butler presented listeners with the two-part contrapuntal patterns shown in Figs. 3a and 3b. Virtually all responses again reflected grouping by frequency range. For both configurations a perceptual reorganization occurred, so that a melody corresponding to the higher tones appeared to be emanating from one earphone or speaker and a melody corresponding to the lower tones from the other (Figs. 3c and 3d).

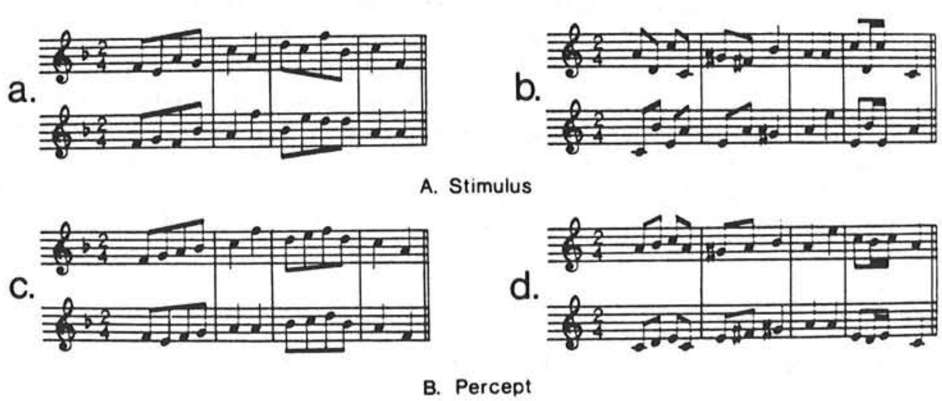

Fig. 3. (A) Two-part melodic patterns as presented to subjects through left and right earphones or speakers. (B) The patterns as most commonly notated by the subjects (from Butler, 1979a).

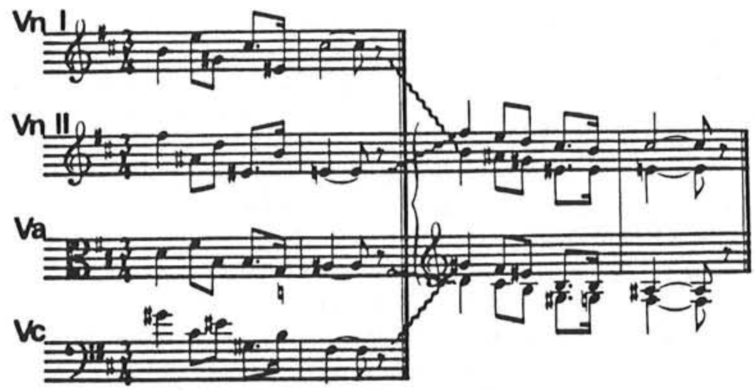

Fig. 4. Passage from the final movement of Tschaikowsky's Sixth (*Pathetique*) Symphony. The combination of the Violin I and Violin II melodies produces the percept shown on the upper right. The combination of the viola and violincello melodies produces the percept shown on the lower right (from Butler, 1979b).

Butler (1979b) further drew attention to an interesting passage from the final movement of Tschaikowsky's Sixth (*Pathetique*) Symphony. As shown in Fig. 4, the theme and accompaniment are each distributed between the two violin parts. However, the theme is heard as coming from one set of instruments and the accompaniment as from the other. Whether it was Tschaikovsky's intention to produce a perceptual illusion here, or whether he expected the listener to hear the theme and accompaniment waft from one set of instruments to the other, we may never know!

How should such gross mislocalization effects be explained? Our acoustic environment is very complex, and the assignment of sounds to their sources is made difficult by the presence of echoes and reverberation (Benade, 1976). So, when a sound mixture is presented such that both ears are stimulated simultaneously, it is unclear from first-order localization cues alone which components of the total spectrum should be assigned to which source. Other factors must also operate to provide cues concerning the sources of these different sounds. One such factor is similarity of frequency spectrum. Similar sounds are likely to be emanating from the same source and different sounds from different sources. Thus, with these musical examples it becomes reasonable for the listener to assume that tones in one frequency range are emanating from one source, and tones in another frequency range from a different source. We therefore reorganize the tones perceptually on the basis of this interpretation (Deutsch, 1975a).

B. Temporal Relationships as Determinants of Grouping

Given the above line of reasoning, we should expect perceptual grouping of simultaneous sequences to be strongly influenced by the salience of first-order localization cues. Under the conditions we have been considering such localization cues were weak, since input was always to both ears simultaneously. However, under conditions where such cues are strong and unambiguous, organization by spatial location should be expected to take precedence over organization by frequency proximity. This should be the case, for instance, where the signals to the two ears are clearly separated in time.

An experiment was therefore performed to examine perceptual grouping as a function of the temporal relationships between the signals arriving at the two ears (Deutsch, 1978a, 1979). Listeners identified melodic patterns in which the component tones switched between the ears. Conditions were compared where input was to one ear at a time and where input was to the two ears simultaneously. Such simultaneity of input was achieved by presenting a drone to the ear opposite the ear receiving the component of the melody. In order to control for the effect of the drone apart from its providing a simultaneous input to the opposite ear, a further condition was included in which the drone and the melody component were presented to the same ear. In a fourth condition the melody was presented binaurally.

It was predicted that in the conditions in which input was to one ear at a time, identification of the melodic patterns should be difficult, reflecting perceptual grouping by spatial location. However, in the conditions where both ears received input simultaneously, identification of the melodic patterns should be much easier, and should reflect organization by frequency proximity in the presence of the contralateral drone.

The experiment employed the two melodic patterns shown in Fig. 5. On each trial, one of these patterns was presented ten times without pause, and listeners identified on forced choice which of these had been presented.

The four conditions of the experiment, together with the error rates in each, are shown in Fig. 6. It can be seen that when the melody was presented to both ears at the same time, identification performance was excellent. However, when the component tones of the melody switched from ear to ear with no accompanying drone, a severe performance decrement occurred. Yet when the drone was presented to the ear *opposite* the ear receiving the melody component, the performance level was again very high, even though the melody was still switching from ear to ear. This result cannot be attributed to processing the harmonic relationships between the drone and the melody

Fig. 5. Basic melodic patterns employed in experiment to study the effects on melody identification of rapid switching between ears. All tones were 30 msec in duration, and tones within a sequence were separated by 100-msec pauses (Deutsch, 1979).

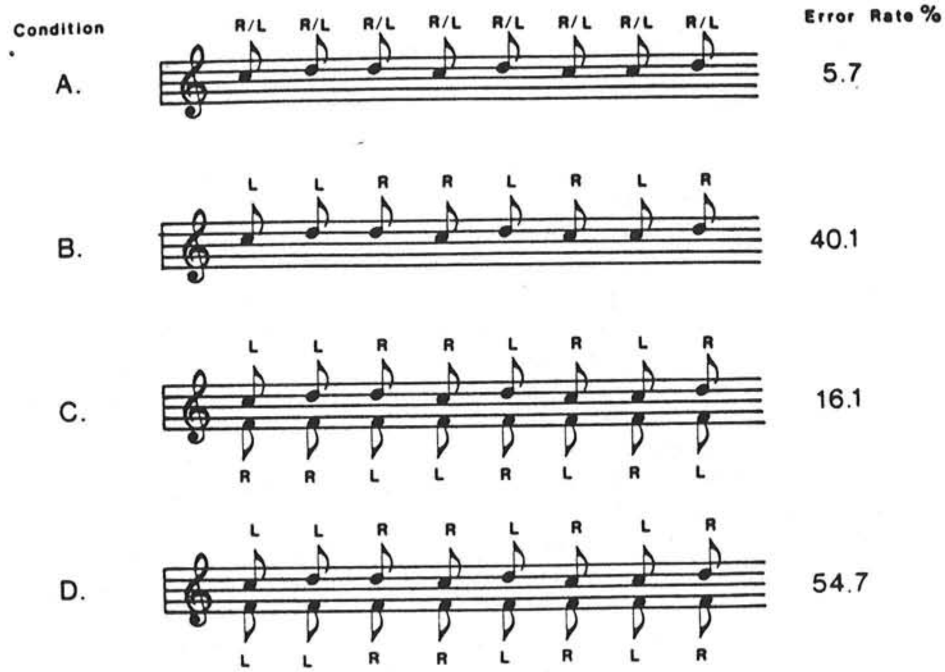

Fig. 6. Examples of distributions between ears of melodic pattern and drone in different conditions of the experiment. See text for details (from Deutsch, 1979).

components because when the drone was presented to the *same* ear as the one receiving the component of the melody, performance was below chance.

This experiment therefore demonstrates that with tones coming from different spatial locations, temporal relationships between them are important determinants of grouping. When signals are coming from two locations simultaneously, it is easy to integrate the information arriving at the two ears into a single perceptual stream. But when the signals coming from the two locations are clearly separated in time, subjective grouping by spatial location is so powerful as to prevent the listener from combining the signals to produce an integrated percept.

A related experiment comparing the effects of simultaneity with nonsimultaneity of input to the two ears was performed by Judd (1979). In this experiment listeners were presented with four-tone melodic patterns whose components alternated from ear to ear. Judd found that presenting noise to the ear contralateral to the ear receiving the melody component resulted in enhanced recognition performance. He also proposed an interpretation in terms of competing channeling mechanisms, reasoning that the strong localization cues present in the no-noise condition induced channeling by spatial location, and that the weaker localization cues due to the noise resulted in channeling by frequency proximity instead.

In the study by Deutsch (1979) the effects of onset–offset asynchrony between the tones arriving at the two ears were also examined. Such temporal overlaps between signals commonly occur in normal listening, and it was predicted that results here should be intermediate between those where the input to the two ears was strictly simultaneous and those where these inputs were clearly separated in time. Such findings were indeed obtained: performance levels under conditions of asynchrony were significantly worse than where the melody components and the contralateral drone were strictly simultaneous, yet significantly better than were the melody switched between ears without an accompanying drone. This is as expected from the present line of reasoning. Temporal similarities in the waveform envelopes of two simultaneous signals are important indicators that these signals are emanating from the same source (following the "principle of Common Fate"), and discrepancies would indicate that the signals are emanating from different sources (Tobias, 1972). We should therefore expect that asynchronies between the signals arriving at the two ears would increase the tendency to treat these signals as emanating from different sources, and so permit less integration of the pattern distributed between the two ears.

Two related experiments on the effects of asynchrony should here be cited. Rasch (1978) investigated the threshold for perception of a high tone when it was accompanied by a low tone. He found that delaying the onset of the low tone relative to the high tone produced a substantial lowering of threshold. Further, under conditions of asynchrony the subjective percept was strikingly altered, so that the two tones stood apart clearly rather than being fused into a single percept. Rasch advanced an interpretation of his findings along lines very similar to those presented here.

Rasch also noted that although temporal asynchrony had strong perceptual effects, it was not recognized as such by the listeners. The same observation was made in the present experiment. In the asynchronous conditions, one obtained the subjective impression of a "plopping" sound at the onset and offset of the tones, but it was difficult to describe this percept further. The strong effect found here due to temporal asynchrony was therefore not based on conscious inference.

Another related experiment is that of Bregman and Pinker (1978). These authors presented a simultaneous two-tone complex in alternation with a third tone, and introduced various conditions of onset–offset asynchrony between the simultaneous tones. They found that with increasing asynchrony there was an increased likelihood that one of the simultaneous tones would form a melodic stream with the third tone. Bregman and Pinker reasoned that asynchrony between the simultaneous tones resulted in a decreased tendency for these tones to be treated as emanating from the same source, and so facilitated a sequential organization by frequency proximity between one of these simultaneous tones and the alternating tone.

These various experiments on the effects of asynchrony bear on an issue that was raised a century ago by von Helmholtz. In his book *On the Sensations of Tone* (1885), he posed the question of how, given the complex, rapidly changing spectrum produced by several instruments playing simultaneously, we are able to reconstruct our musical environment so that some components fuse to produce a single sound impression, while others are heard as separate melodic lines which may be simultaneously perceived. For the latter instance, he posed the further question as to the basis on which such simultaneous melodic lines are constructed. Thus he wrote:

> Now there are many circumstances which assist us first in separating the musical tones arising from different sources, and secondly, in keeping together the partial tones of each separate source. Thus when one musical tone is heard for some time before being joined by the second, and then the second continues after the first has ceased, the separation in sound is facilitated by the succession of time. We have already heard the first musical tone by itself, and hence know immediately what we have to deduct from the compound effect for the effect of this first tone. Even when several parts proceed in the same rhythm in polyphonic music, the mode in which the tones of different instruments and voices commence, the nature of their increase in force, the certainty with which they are held, and the manner in which they die off, are generally slightly different for each . . . but besides all this, in good part music, especial care is taken to facilitate the separation of the parts by the ear. In polyphonic music proper, where each part has its own distinct melody, a principal means of clearly separating the progression of each part has always consisted in making them proceed in different rhythms and on different divisions of the bars (p. 59).

And later:

> All these helps fail in the resolution of musical tones into their constituent partials. When a compound tone commences to sound, all its partial tones commence with the same comparative strength; when it swells, all of them generally swell uniformly; when it ceases, all cease simultaneously. Hence no opportunity is generally given for hearing them separately and independently (p. 60).

C. The Octave Illusion

In the experiments so far described, channeling by frequency proximity was the rule when information was presented to both ears simultaneously. Channeling by spatial location occurred only with temporal separations between the stimuli presented to the two ears. We now examine conditions where channeling by spatial location occurred even though the input to the ears was strictly simultaneous. We shall see that this principle was adopted under special conditions of frequency relationship between the tones as they were presented in sequence at the two ears.

One stimulus configuration that induced melodic channeling by spatial location is shown in Fig. 7A. It can be seen that this consisted of two tones that were spaced an octave apart and repeatedly presented in alternation. The identical sequence was delivered to the two ears simultaneously; however, when the right ear received the high tone the left ear received the low tone, and vice versa. So, essentially the configuration was that of a two-tone chord, where the ear of input for each component switched repeatedly (Duetsch, 1974a,b, 1975a).

This sequence was found to produce various illusions, the most common of which is shown on Fig. 7B. It can be seen that this consisted of a single tone that alternated from ear to ear, and whose pitch also alternated from one octave to the other in synchrony with the localization shift. When the earphones were placed in reverse position, most listeners found that the apparent locations of the high and low tones remained fixed. So it seemed to these listeners that the earphone that had been

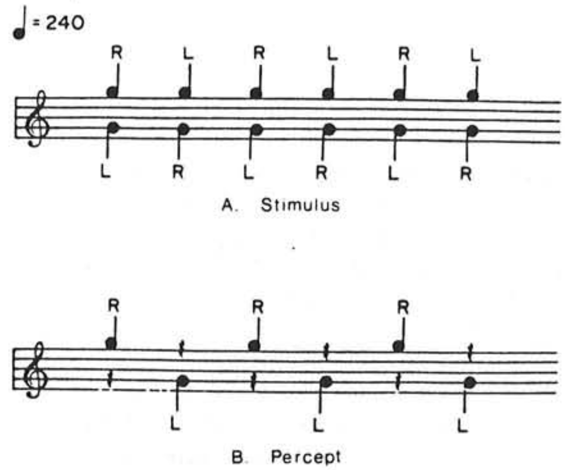

Fig. 7. (A) Representation of the configuration producing the octave illusion. (B) Representation of the illusory percept most commonly obtained (from Deutsch, 1974b).

producing the high tones was now producing the low tones, and that the earphone that had been producing the low tones was now producing the high tones.

It was hypothesized that this illusion results from the operation of two different selection mechanisms underlying the pitch and the localization percepts. To provide the perceived sequence of pitches (i.e., the melodic line) the frequencies arriving at one ear are followed, and those arriving at the other ear are suppressed. However, each tone is localized at the ear receiving the higher frequency signal, regardless of whether the higher or the lower frequency is in fact perceived (Deutsch, 1975a). This model was confirmed in later experiments (Deutsch, 1976, 1978b; Deutsch & Roll, 1976).

We can next ask whether the interactions giving rise to the octave illusion take place between pathways specific to information from the two ears, or whether instead pathways conveying information from different regions of auditory space are involved. In order to investigate this question, the stimuli were presented through spatially separated loudspeakers rather than earphones. It was found that the analogous illusion was obtained: A high tone apparently emanating from one speaker was perceived as alternating with a low tone apparently emanating from the other speaker (Deutsch, 1975a).

Setting aside the issue of where the tones appeared to be located in the octave illusion, and considering only what sequence of pitches was perceived, we note that here channeling was always on the basis of spatial location. This stands in sharp contrast to findings on the scale illusion, in which channeling was always on the basis of frequency proximity instead. Yet the stimuli producing these illusions were remarkably similar. In both cases listeners were presented with repeating sequences of sine wave tones at equal amplitudes and durations. In both cases the stimuli were continuous, and the frequencies presented to one ear always differed from the frequencies simultaneously presented to the other ear. Yet these two configurations gave rise to radically different channeling strategies. It is especially noteworthy that when two tones that were separated by an octave were simultaneously presented in the scale illusion, both tones were generally perceived (Fig. 2B). Yet when two tones that were separated by an octave were simultaneously presented in the octave illusion, only one of these was generally perceived (Fig. 7B). These differences in channeling strategy must therefore have resulted from differences in the patterns of frequency relationship between successive tones.

Considering the octave illusion further, we find that here the frequency emanating from one side of space was always identical to the frequency that had just emanated from the opposite side. It was therefore hypothesized that this factor was responsible for inducing melodic channeling by spatial location for this configuration. A further set of experiments was performed to test this hypothesis (Deutsch, 1980, 1981).

In the first experiment listeners were presented with sequences consisting of 20 dichotic chords. Two conditions were compared, which employed the basic configurations shown on Fig. 8A. The configuration in Condition 1 consisted of the repetitive presentation of a single chord, whose components stood in octave relation and alternated from ear to ear such that when the high tone was in the right ear the low tone was in the left ear, and vice versa. It can be noted that here the two ears received the same frequencies in succession. On half the trials the sequence presented to the right ear began with the high tone and ended with the low tone, and on the other half this order was reversed. On each trial, subjects judged whether the sequence 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; and from these judgments it was inferred which ear was being followed for pitch.

The basic configuration in Condition 2 consisted of the repetitive presentation of two dichotic chords in alternation, the first forming an octave and the second a minor third, so that the entire four-tone combination constituted a major triad. It will be noted that here the two ears did not receive the same frequencies in succession. On half the trials the right ear received the upper component of the first chord and the lower component of the last chord, and on the other half this order was reversed.

The amplitude relationships between the tones presented to the two ears were systematically varied across trials, and the extent to which each location was followed was plotted as a function of these amplitude relationships. The results are shown on Fig. 8B. It can be seen that in Condition 1 the frequencies presented to one location were followed until a critical level of amplitude relationship was reached, and the other location was followed beyond this level. However, in Condition 2 there was no following on the basis of spatial location, even when the signals presented to the two locations differed substantially in amplitude. However, hypothesizing instead that the subjects were following this sequence on the basis of frequency proximity a very consistent result emerged: Three of the subjects consistently followed the low tones, and one consistently followed the high tones. This result is as expected if the critical factor responsible for channeling by spatial location here is that the same frequencies emanate in succession from different regions of auditory space.

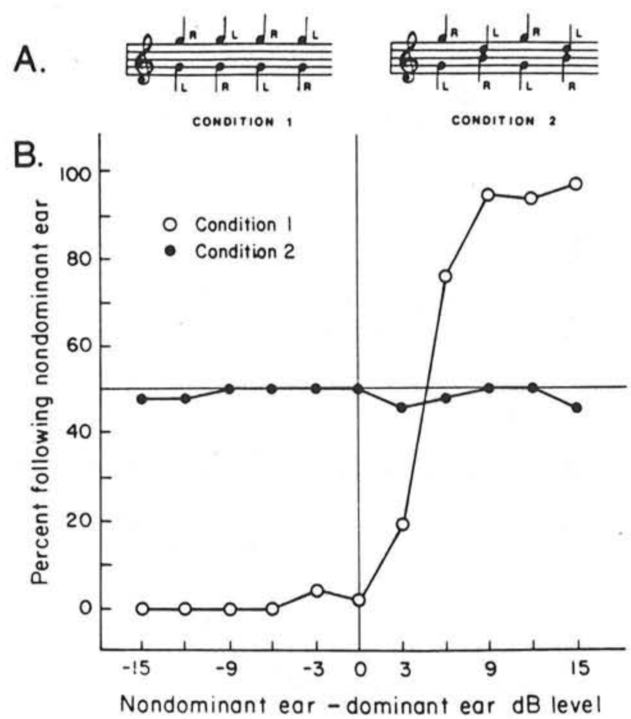

Fig. 8. (A) Configurations used in first experiment investigating the factors producing following on the basis of spatial location. (B) Percent following of nondominant ear in the two conditions of the experiment (from Deutsch, 1980a).

In a second experiment only two dichotic chords were presented on each trial. Two conditions were again compared, which employed the basic configurations shown on Fig. 9A. The configuration in Condition 1 consisted of two presentations of the identical chord, whose components formed an octave, 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 used throughout this condition. The basic configuration in Condition 2 consisted of two dichotic chords, each of which formed an octave, but which were composed of different frequencies. Trials employing chords composed of C and F♯ and of A and D♯ occurred in strict alternation, so that any given chord was repeated only after a substantial time period during which several other chords were interpolated.

Figure 9B shows the extent to which each location was followed in these two conditions, as a function of the amplitude relationships between the signals at the two ears. It can be seen that in Condition 1 following was clearly on the basis of spatial location. However, in Condition 2 such following did not occur, even when there were substantial amplitude differences between the signals at the two ears. Instead the subjects consistently followed these sequences on the basis of overall contour: Their

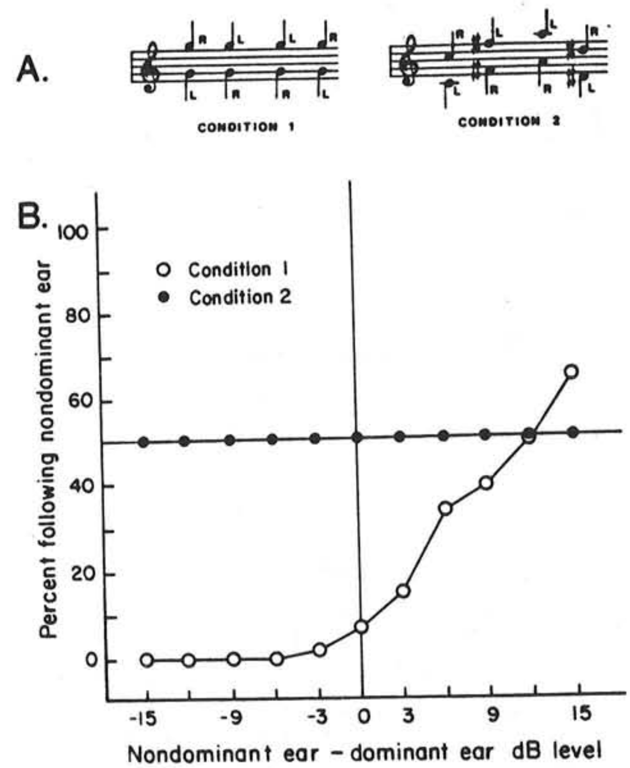

Fig. 9. (A) Configurations used in second experiment investigating the factors producing following by spatial location. (B) Percent following of nondominant ear in the two conditions of the experiment (from Deutsch, 1980a).

patterns of response 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. This was true even when the signals at the two ears differed substantially in amplitude.

It should be noted that following by contour here was consistent with following by frequency proximity, but that following the middle tones of the sequence (for example, the high C and the low F♯ in Fig. 9A) would have been equally consistent, since in either case the melodic interval formed by the two tones was a diminished fifth. So, in these sequences, overall contour was the factor that determined which tones were followed. Whether this served simply to "break a tie" or whether contour would win out in competition with frequency proximity in other configurations remains to be determined.

At all events, the second experiment showed, as did the first, that following by spatial location occurs in sequences where the same frequencies emanate in succession from two different regions of auditory space, and it occurs on other lines when this relationship does not hold. It is particularly interesting to note that relative amplitude was found not to be an important factor in either experiment. When following was by frequency proximity or by contour, this occurred in the face of substantial amplitude differences between the signals arriving at the two ears. When following was by spatial location, the switch from following one side of space to the other did not occur at the point where the amplitude balance shifted from one side of space to the other, but at a different level of amplitude relationship. Thus, amplitude here appeared to *set the scene* for following on the basis of spatial location rather than serving as a primary following principle. [It will be recalled that in the experiment by Butler (1979a) following by frequency proximity also occurred in the face of amplitude differences between signals emanating from different spatial locations.]

We can next ask whether the lack of following by spatial location in the second conditions of these two experiments was due to the delay between successive presentations of the same frequencies to the two locations or to the interpolation of tones of different frequencies. To examine the effect of interpolated information, we studied performance under two further conditions. As shown on Fig. 10A, the configuration

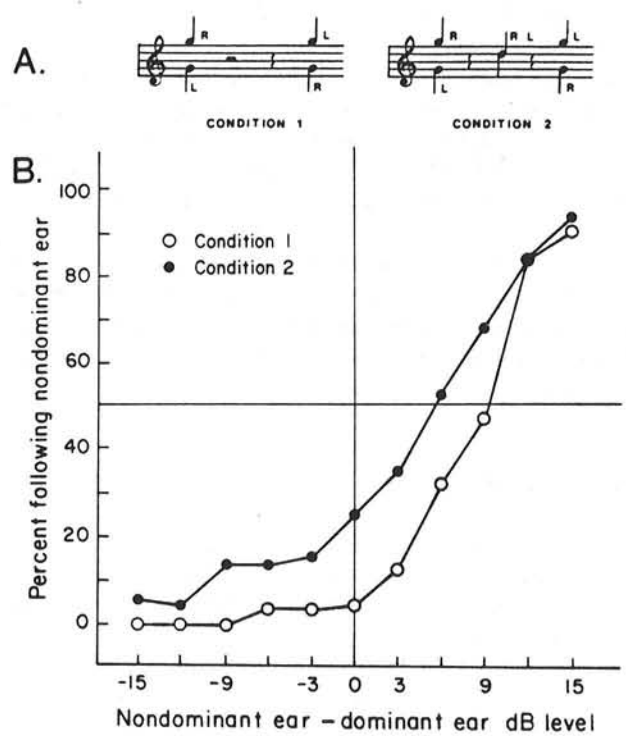

Fig. 10. (A) Configurations used in third experiment investigating the factors producing following by spatial location. (B) Percent following of nondominant ear in the two conditions of the experiment (from Deutsch, 1980a).

in these two conditions were identical, except that in Condition 2 a single tone was interpolated between the dichotic chords and the listeners were instructed to ignore this tone. As shown on Fig. 10B, there was a less pronounced following of the preferred spatial location in the condition in which the extra tone was interpolated. To investigate the effects of temporal dealy, we varied the time interval between the onsets of successive tones at the two ears. This was achieved either by changing the durations of the tones or by interpolating gaps between them. It was found that the strength of this effect decreased with increasing time between onsets of the identical frequencies at the two locations. It did not matter whether this increase was produced by lengthening the durations of the tones, or by interpolating gaps between them. Thus, both interpolated information and temporal delay were found to reduce channeling be preferred spatial location.

These experiments raise the question of why such a channeling mechanism should have developed. We may hypothesize that this mechanism enables us to follow new, ongoing auditory information with a minimum of interference from echoes or reverberation. In everyday listening, when the same frequency emanates successively from two different regions of auditory space, the second occurrence may well be due to an echo. This is made more probable as the delay between the onsets of these two occurrences is shortened. However, if different frequencies are interpolated between two occurrences of the identical frequency, other interpretations of the second occurrence are made more likely. We may therefore hypothesize that this falls into the class of mechanisms that act to counteract misleading effects of echoes and reverberation. Such an explanation has been advanced, for instance, for the precedence effect. Wallach, Newman and Rosenweig (1949) have reported that in listening to music a single image may be obtained with the waveform presented at two different spatial locations separated by 45–70 msec intervals. The second sound is, under these conditions, attributed to the same location as the first. Analogous findings have been reported by Haas (1951).

D. Handedness Correlates

Strong handedness correlates have been obtained for both the octave and the scale illusions. In the case of the octave illusion, there was a strong tendency among right-handers to hear the high tone on the right and the low tone on the left. This was not, however, found among left-handers (Deutsch, 1974). From further studies it was concluded that these findings reflected a tendency to perceive the pattern of frequencies presented to the dominant side of auditory space rather than the nondominant (Deutsch, 1975a, 1981; Deutsch & Roll, 1976). In the case of the scale illusion, there was also a strong tendency among right-handers to hear the higher tones on the right and the lower tones on the left; again this was not true of left-handers. Here the mislocalization of the higher tones to one spatial position and the lower tones to another cannot be interpreted in terms of a following of the input from one side of space rather than the other, since the higher and lower melodic lines were each composed of tones that emanated from both spatial locations. One may, however, interpret this handedness correlate as reflecting relatively more activity in the dominant hemisphere on the part of neural units underlying the higher tones, and relatively more activity in the nondominant hemisphere on the part of neural units underlying the lower tones. Justification for this view comes in part from neurological studies showing that patients who experience palinacousis tend to perceive the illusory sound as located on the side of auditory space contralateral to the lesion (Jacobs, Feldman, Diamond, & Bender, 1973). Further, when patients obtain auditory sensations upon stimulation of the temporal lobe, these sensations are also generally referred to contralateral auditory space (Penfield & Perot, 1963).

A prominence of dominant over nondominant pathways is therefore implicated in both the octave and the scale illusions. These findings may be related to those of other investigators who explored patterns of ear advantage in the processing of melodies or tonal sequences. Very mixed results have been obtained in these studies. Some have found left ear superiorities (Kimura, 1964, 1967; King & Kimura, 1972; Bartholomeus, 1974; Darwin, 1969; Zatorre, 1979; Spellacy, 1970; Spreen, Spellacy and Reid, 1970). Others have found no ear differences (Gordon, 1970; Bartholomeus, Doehring, & Freygood, 1973; Berlin, 1972; Doehring, 1971, 1972). Under yet other conditions right ear superiorities have been obtained instead (Halperin, Nachshon, & Carmon, 1973; Robinson & Solomon, 1974; Papcun, Krashen, Terbeek, Remington, & Harshmann, 1974; Natale, 1977). Bever and Chiarello (1974), and Johnson (1977) obtained a right ear superiority for musicians and a left ear superiority for nonmusicians in melody recognition.

Such inconsistencies are probably due to a variety of factors. With sequences generated by voice or by natural instruments, recognition may be achieved in part by spectral cues. Loudness and temporal cues may also serve as bases for judgment, and so on. It is not unreasonable to suppose that specific attributes of a melodic segment might be processed in different parts of the nervous system. For example, Gordon (1970) obtained a left ear advantage in processing dichotically presented chords that were generated by an electronic organ, yet in this same study he failed to find an ear difference in the processing of melodies. Gaede, Parsons and Bertera (1978), using monaural presentation, found a left ear advantage in processing chords, and yet a right ear advantage in melody recognition. Further, Charbonneau and Risset (1975) studied the processing of dichotic sound sequences that varied either in fundamental frequency or in spectral envelope. When fundamental frequency was varied, a right ear advantage was obtained. Yet when spectral envelope was varied, a left ear advantage was obtained instead. If the relative involvement of the two hemispheres differs, depending on the specific musical attribute being processed, this could explain such results. Furthermore, different categories of listener might utilize specific musical attributes to varying extents. This could explain the discrepancies in performance found between musicians and nonmusicians in sone studies.

Ear differences have been interpreted by some investigators in terms of a simple dichotomy in processing strategy: the left or dominant hemisphere is assumed to specialize in "analytic" processing; the right or nondominant hemisphere in "Gestalt" or "holistic" processing. However the meaning of such a dichotomy is far from clear. For instance, melody perception is held to be a "Gestalt" phenomenon. Indeed, Von Ehrenfels (1890) originally gave melody as an example of a Gestalt, because it retains its perceptual identity under transposition. However, in order to produce invariance under transposition, a set of specific intervals must be abstracted and their orders preserved. This requires a set of highly specific analyses. It would seem more useful, rather than invoking a nebulous "analytic-holistic" distinction, to attempt to pinpoint the types of processing responsible for different patterns of ear advantage.

There is a further factor that should be considered. This arises from findings on the scale illusion. When two melodies are simultaneously presented, one to each ear, the listener may not perceive these same melodies but may instead perceptually synthesize two different melodies, as shown in Figs. 2 and 3. When one of these melodies is later presented for recognition, accuracy may then be determined in part by the perceptual reorganization that had occurred during the dichotic presentation. For example, the typical right-handed listener, on perceiving the dichotic scale sequence, perceptually displaces the high tones from his left ear to his right, and perceptually displaces the low tones from his right ear to his left. A recent study has demonstrated that this phenomenon forms the basis of the apparent left ear advantage in dichotic listening to simultaneous sequences of tones (Deutsch, in preparation).[1]

E. Melody Perception from Phase-Shifted Tones

Another technique relating to melodic channeling on the basis of spatial location was employed by Kubovy and co-workers. Kubovy, Cutting and McGuire (1974) presented a set of simultaneous and continuous sine wave tones to both ears. One of these tones in one ear was phase shifted relative to its counterpart in the opposite ear. When these tones were phase shifted in sequence a melody that corresponded to the shifted tones was clearly heard. However, the melody was undetectable when the stimulus was presented to either ear alone. Subjectively, the dichotically presented melody was heard as occurring inside the head but displaced to one side of the midline, while a background noise was heard as localized to the opposite side. So, it was as though a source in one spatial position was producing the melody, and a different source in another spatial position was producing the noise.

Kubovy (1981) pointed out that there are two potential interpretations of this effect. First, the segregation of the melody from the noise could have been based on *concurrent difference cues;* i.e., the target tone may have been segregated because at that time its interaural disparity—or apparent spatial location—differed from that of the background tones. Alternatively, the effect could have been based on *successive difference cues,* i.e., the target tone may have been segregated because it had moved its apparent location. Two further configurations were therefore devised to determine which of

[1]For a review of the neurological substrates of music perception see Chapter 15.

these factors was responsible. In the first, the target tones moved while the locations of the background tones remained constant, producing a successive difference cue. In the second, the target tones themselves did not move, but the background tones did, so that the target tones were segregated from the others, producing a concurrent difference cue. Kubovy found that although both types of cue were effective in producing segregation, the successive difference cue was considerably more effective than the concurrent difference cue.

In another experiment, Kubovy and Howard (1976) presented six tones simultaneously, in such a way that each occupied a different apparent position in space. They then displaced each tone in turn to a new apparent position, and so produced a melody by successive difference cues. They studied the effect of interpolating temporal gaps of different durations between successive tone bursts, and found that the melody could still be heard through such gaps. Although there was considerable intersubject variability in sensitivity to the effect of the gaps, one subject performed perfectly with gaps of 9.7 seconds (the longest duration employed). Thus, this effect of a successive difference cue was found to be capable of acting over surprisingly long silent intervals.

It is interesting to note that in Kubovy's paradigm, configurations were formed from a fixed set of tonal frequencies that simply shifted their apparent positions in space. Melodic channeling resulted from these movements of spatial position. This is analogous to the situation in the octave illusion where two continuous tones interchange their positions in space, resulting in melodic channeling by spatial location. As we have seen, when successive configurations are formed from different frequencies, rather than identical frequencies, channeling on other lines occurs instead. The issue of how differences in the frequencies of successive chords would affect channeling in Kubovy's paradigm remains to be explored.

F. Discussion

We have found that the issue of how melodic channels are formed in two-channel listening situations is a complex one. Given certain stimulus configurations, channeling occurs on the basis of spatial location. Yet given other configurations, channeling occurs instead on the basis of frequency proximity or contour. In the conditions we have examined, amplitude plays a remarkably small role as a basis for channeling.

The radical differences in channeling strategy demonstrated here bear on certain apparent inconsistencies in the literature on divided attention. Certain investigators have found that the requirement to distribute attention across ears produced performance decrements on various tasks (e.g., Cherry & Taylor, 1954; Broadbent, 1954, 1958; Moray, 1959; Treisman, 1971). It was hypothesized that such decrements were due to an inability to switch attention between ears rapidly enough for the task demands. However, other investigators have found evidence against this view (e.g., Sorkin, Pastore, & Pohlman, 1972; Sorkin, Pohlman, & Gilliom, 1973; Moray, 1975;

Shiffrin, Pisoni, & Casteneda-Mendez, 1974; Pollack, 1978). It would appear from the work reviewed here that deficits in monitoring information simultaneously from two spatial locations should occur with certain types of stimulus configuration but not with others. It is probable that the configurations that give rise to such deficits are such as to induce the strong inference that the inputs to the two ears are emanating from separate sources rather than a single source. Integrating the information from such two sources would, in normal listening situations, lead to confusion in monitoring the environment. However, with configurations where there is an ambiguity of interpretation in terms of sources, integration of the information from the two ears could be the most useful strategy.

IV. CHANNELING OF RAPID SEQUENCES OF SINGLE TONES

A. Grouping by Frequency Proximity

Melodic channeling has also been studied with the use of rapid sequences of single tones. When these tones are in more than one frequency range, they tend to split apart perceptually, with the result that two or more melodic lines are heard in parallel. Composers often take advantage of this perceptual phenomenon with their use of pseudopolyphony, or compound melodic line. Here one instrument plays a rapid sequence of single tones which are drawn from different pitch ranges, so that two simultaneous melodic streams are clearly perceived. Figure 11a shows a segment of music that exploits this principle. Figure 11b shows the same segment with log frequency and time mapped into two dimensions of visual space. It can be seen that the principle of Proximity clearly emerges in the visual representation. At lower speeds, the tendency to group by pitch proximity still persists, but is subjectively less compelling.

One of the early experiments on this phenomenon was that of Miller and Heise (1950), who presented listeners with a sequence consisting of two tones that alternated at a rate of 10 per second. They found that when the frequencies of these tones differed by less than 15%, the sequence was heard as a trill (i.e., as a single string of related tones). However, when the frequency disparity between the alternating tones increased, the sequence was heard instead as two interrupted and unrelated tones. This phenomenon has been termed "fission" by several investigators (Van Noorden, 1975). Heise and Miller (1951) examined this phenomenon further, using sequences of tones that were composed of several different frequencies. They found that if one of the tones in a rapid repetitive sequence differed sufficiently in frequency from the others it was heard as in isolation from them.

Dowling (1967, 1973) has demonstrated the importance of this principle in a long term memory situation. He presented two well-known melodies with their component tones alternating at a rate of eight per second. Recognition of these melodies was

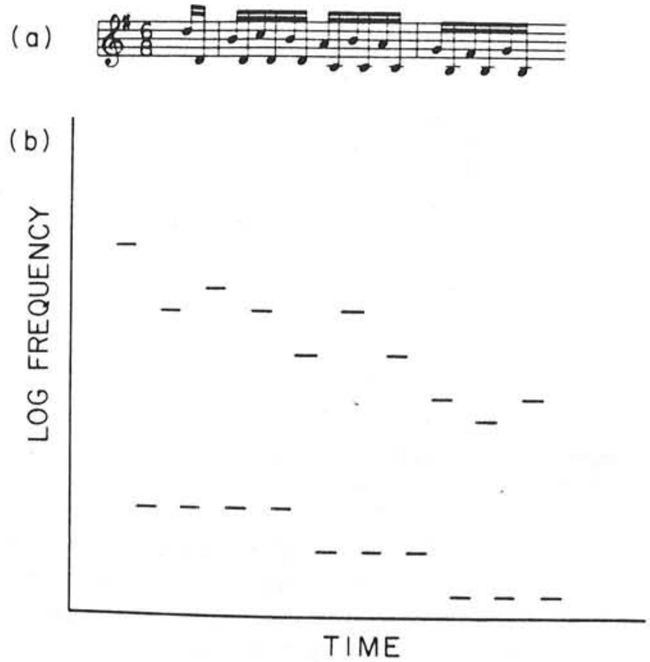

Fig. 11. Grouping of melodic stimuli on the basis of frequency proximity. Two parallel lines are perceived, each in a different frequency range (from Beethoven's *Six Variations on the Duet* "Nel cor piu non mi sento" from Paisiello's *La Molinara*).

found to be very difficult when they were in overlapping pitch ranges since their components were perceptually combined into a single stream. However, as one of the alternating melodies was gradually transposed so that their pitch ranges diverged, recognition became increasingly more easy.

B. Temporal Coherence as a Function of Frequency Proximity and Tempo

Temporal coherence is a term used to describe the subjective impression that a sequence of tones froms a connected series. Schouten (1962) studied the conditions giving rise to the perception of temporal coherence. He varied both the frequency relationships between successive tones in a sequence and also their presentation rate. As the frequency separation between successive tones increased, the tempo of the sequence had to be reduced in order to maintain the impression of temporal coherence between these tones.

Van Noorden (1975) investigated this phenomenon in detail. Listeners were presented with sequences of alternating tones, and were instructed either to try to hear temporal coherence or to try to hear fission. Two boundaries were determined by this

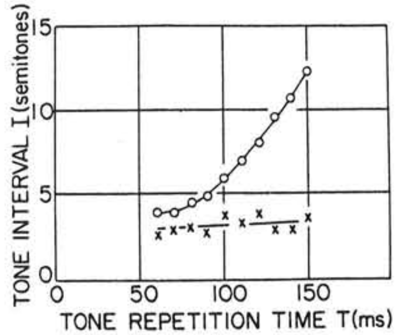

Fig. 12. Temporal coherence boundary (o) and fission boundary (x) as a function of frequency relationship between alternating tones and presentation rate (from Van Noorden, 1975).

method. The first, termed the *temporal coherence boundary*, established the threshold frequency separation as a function of tempo required for the listener to hear the sequence as coherent. The second, termed the *fission boundary*, established these values when the listener was attempting to hear fission. As shown in Fig. 12, when listeners were trying to hear coherence, decreasing the tempo from 50 to 150 msec per tone increased the frequency separation within which coherence could be heard from 4 to 13 semitones. However, when the listeners were trying to hear fission, decreasing the tempo had little effect on performance. Between these two boundaries there was a large region where the listener could alter his listening strategy at will, and so hear either fission or coherence. So within this region, attentional set was important in determining how the sequence was perceived; however outside this region, attentional set was not effective.

Bregman and Bernstein (quoted in Bregman, 1978) confirmed the finding of an interaction between frequency separation and tempo in judgments of coherence. They found that as two alternating tones converged in frequency, a higher rate of alternation was required for the sequence to split into two streams. This effect was found to hold throughout a substantial frequency range.

C. Grouping by Frequency Proximity Builds with Repetition

Several experiments have shown that the splitting of tonal sequences into streams on the basis of frequency proximity builds up with repetitive presentation. For instance, Van Noorden (1975) compared the temporal coherence boundary for two-tone, three-tone, and long repetitive sequences. With three-tone sequences the frequency change was either unidirectional or bidirectional. It was found that for unidirectional three-tone sequences, temporal coherence was observed at rates that were equal to or even higher than those for two-tone sequences. (This follows the principle of Good Continuation, as described below.) But with bidirectional three-tone sequences, the rate of frequency change had to be set much lower than for two-tone

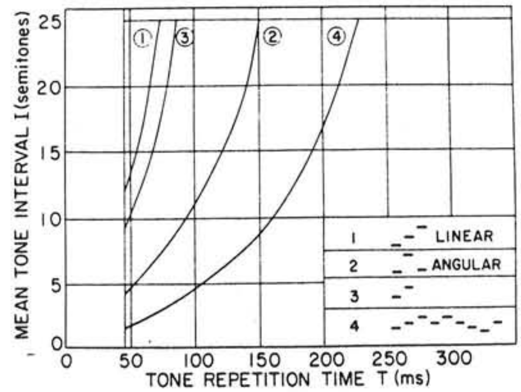

Fig. 13. Temporal coherence boundary for two-tone, three-tone undirectional, three-tone bidirectional, and continuous sequences (from Van Noorden, 1975).

sequences before coherence could be perceived. With long repetitive sequences the rate of frequency change had to be set lower still (Fig. 13).

In a related experiment, Bregman (1978) presented listeners with sequences consisting of two "high" tones (748 and 831 Hz) and one "low" tone (330 Hz). When this sequence split into two streams, the upper stream was perceived as an alternation of two high tones, and the lower stream as the steady repetition of a single tone. The experiment varied the number of tones packaged between four-second periods of silence. On each trial listeners adjusted the speed of the sequence until the point of splitting was determined. As shown on Fig. 14, as the package size increased, the speed required for segregation decreased.

Bregman interpreted these findings along the following lines. Stream segregation may be viewed as the product of a mechanism that acts to "parse" the auditory environment (i.e., to group together components of the acoustic spectrum in such a way as to recover the original sources). Such a mechanism would be expected to accumulate evidence over time, so that the segregation of acoustic components into groups should build up with repeated presentation.

Further evidence for the view that stream segregation results from a "parsing" mechanism was provided by Bregman and Rudnicky (1975). Listeners judged the order of two tones that were embedded in a four-tone pattern flanked by two "distractor" tones. The presence of the distractor tones made judgment of the order of the test tones difficult. However, when another stream of tones, called "captor" tones, was moved close to the "distractor" tones, this caused the "distractors" to combine with the "captors" to form a single stream, leaving the test tones in a stream of their own. This had the consequence that the order of the test tones was now easy to judge. Bregman and Rudnicky argue that this situation presents the listener with two simultaneously structured streams, of which the "distractor" tones can belong to either one, but not to both simultaneously. This is as expected on an interpretation in terms of an

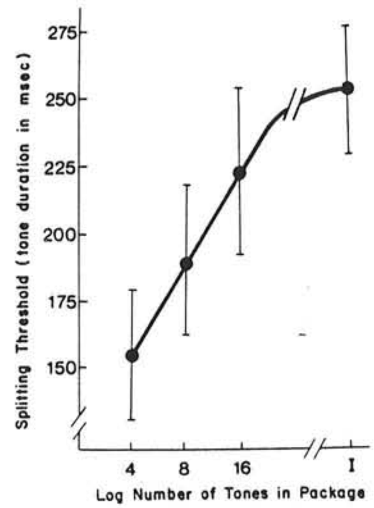

Fig. 14. Threshold for stream segregation as a function of number of tones per package. Two "high" tones were presented in alternation with a single "low" tone (from Bregman, 1978).

auditory parsing mechanism: any given tone is likely to be emanating from only one source; not from two sources simultaneously.

It should be noted that the cumulation of effect over time found by Bregman (1978) is analogous to cumulation effects found in the octave illusion, where the strength of tendency to follow the frequency presented to one side of auditory space rather than the other also builds up with repeated presentation, and builds up more rapidly as repetition rate increases. Analogous findings were obtained for the strength of tendency to localize toward the higher frequency signal in this illusion (Duetsch, 1976, 1978). Such a built-up of effect is also well interpreted in terms of evidence accumulation.

D. Grouping by Frequency Proximity and the Perception of Temporal Relationships

One striking consequence of the formation of separate streams out of rapidly presented sequences is that temporal relationships between the elements of different streams become difficult to process. This has been shown in several ways. Bregman and Campbell (1971) presented a repetitive sequence consisting of six tones: three from a high frequency range and three from a low frequency range. When these tones were presented at a rate of 10 per second, it was difficult for listeners to perceive a pattern of high and low tones that was embedded in the sequence.

Dannenbring and Bregman (1976) demonstrated a further perceptual consequence of this breakdown of temporal processing. They found that when two tones alternate at high speeds so that they produce separate perceptual streams, the tones in the two streams appear to be perceptually overlapping in time. A related study was performed by Fitzgibbon, Pollatsek, and Thomas (1974) who explored the perception of temporal gaps between tones occurring in rapid sequence. When a 20-msec gap was interpolated between tones in the same frequency range, detection of this gap was easy. However, when the gap was interpolated between tones in different frequency ranges, detection performance dropped considerably.

A further reflection of this breakdown of temporal processing was found by Van Noorden (1975). He studied detection of the temporal displacement of a tone that alternated continuously with another tone of different frequency and found that as the tempo of the sequence increased, the mean just noticeable displacement also increased. This increase was substantial for sequences where the tones were widely separated in frequency, but only slight for sequences where the frequencies were contiguous. These results paralleled those found for judgments of temporal coherence.

Such deterioration of temporal processing as a result of frequency disparity occurs with two-tone sequences also. Divenyi and Hirsh (1972) found that discrimination of the size of a temporal gap between a tone pair deteriorates with increasing frequency disparity between members of the pair. Further, Williams and Perrott (1972) measured the minimum detectable gap between tone pairs. They found that for tones of 100 and 30 msec duration, the threshold rose with increasing frequency disparity between members of the pair. However, Van Noorden (1975) showed that this deterioration of temporal processing is considerably greater for long repetitive sequences than for two-tone sequences; so that it develops as a consequence of stream formation (Fig. 15). This conclusion also follows from consideration of Bregman and Campbell's results (1971).

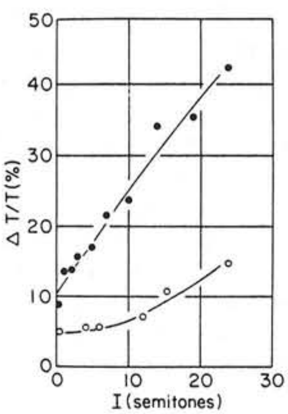

Fig. 15. ○ The just noticeable displacement $\Delta T/T$ of the second tone of a two-tone sequence as a function of tone interval I. ● The just noticeable displacement $\Delta T/T$ of one tone in a continuous sequence of alternating tones as a function of tone interval I (from Van Noorden, 1975).

E. Grouping by Good Continuation

Another principle found to be effective in producing grouping is that of Good Continuation. Bregman and Dannenbring (1973) found that when a repeating cycle consisting of a high tone alternating with a low tone tended to segregate into two streams, this splitting tendency was reduced when the high and low tones were connected by frequency glides. Similarly Nabelek, Nabelek, and Hirsh (1973) reported that for complex tone bursts, frequency glides between the initial and final tones resulted in more pitch fusion than when these tones were juxtaposed with no transitions.

Related experiments have involved the perception of rapid sequences of three or more tones. Divenyi and Hirsh (1974) studied order identification for three-tone sequences, and found that sequences with unidirectional frequency changes were easier to order than sequences with bidirectional frequency changes. Analogous results were obtained by Nickerson and Freeman (1974), Warren and Byrnes (1975), and McNally and Handel (1977) for four-tone sequences. Furthermore, Van Noorden (1975) found that a sequence of three tones was more likely to be judged as coherent if these tones formed a unidirectional rather than a bidirectional frequency change.[2]

F. Grouping by Timbre

The grouping of complex tones on the basis of sound type or timbre is an example of grouping by the principle of Similarity. (A visual example of such grouping is shown in Fig. 1b, where the open and closed circles each combine perceptually to form vertical rows.) Grouping on the basis of timbre is clearly apparent in natural musical situations (Erickson, 1975). Adjacent phrases are often played by different instruments to enhance their perceptual separation. Further, overlaps in pitch range are far more common where more than one instrument type is involved, reflecting the greater perceptual separation provided by the timbral difference.

A striking demonstration of grouping by timbre was produced by Warren, Obusek, Farmer and Warren (1969). They constructed repeating sequences consisting of four unrelated sounds, a high tone (1000 Hz), a hiss (2000 Hz octave band noise), a low tone (796 Hz), and a buzz (400 Hz square wave). Each sound was 200 msec in duration, and the sounds followed each other without pause. Listeners were quite unable to name the orders of the sounds in such repeating sequences. For correct naming to be achieved, the duration of each sound had to be increased to over half a second.

It appears that two separate factors are involved in this effect. The first factor is the organization of the elements of a sequence into separate streams on the basis of sound

[2]Further issues concerning the grouping of rapid sequences of tones involve the effect of average frequency difference between the tones. An extended discussion of these issues is beyond the scope of the present chapter, and the reader is referred to Warren and Byrnes (1975), Nickerson and Freeman (1974), and Divenyi and Hirsh (1978).

type, analogous to organization on the basis of frequency proximity. The second factor involves the lack of familiarity with such sound sequences. It has been shown that when verbal items are combined to form repeating sequences of this nature, correct ordering occurs at considerably faster rates (Warren & Warren, 1970; Thomas, Cetti, & Chase, 1971; Thomas, Hill, Carroll, & Garcia, 1970; Dorman, Cutting, & Raphael, 1975). It is likely that sequences composed of familiar musical sounds would also be more easily ordered. Although this has not been formally investigated, an observation by the author is of relevance here. It was found that a trained percussionist specializing in avant-garde music had little difficulty in discriminating sequences such as those created by Warren and his colleagues. This musician frequently produced such sequences in musical performance.

The question then arises as to the nature of the process that enables the rapid reconstruction of the order of components of complex yet familiar sound sequences such as in speech and music. Wickelgren (1969, 1976) has proposed that the correct ordering of speech components is based on an encoding of a set of context-sensitive elements that need not themselves be ordered. For instance, he proposed that the word "struck" is encoded not as the ordered set of phonemes /s/, /t/, /v/, /u/, /k/, but as the unordered set of context-sensitive allophones /$\#^{s}t$/, /$s^{t}r$/, /$t^{r}u$/, /$r^{u}k$/, /$u^{k}\#$/. Thus each of these context-sensitive elements contains some local information concerning how this element is ordered in relation to the other elements in the set. From such an unordered set of elements the information concerning their order can be derived.

This theory can readily be applied to auditory perception in general, and one may hypothesize that the easy identification of familiar sound sequences is mediated by an acquired set of such context-sensitive elements. For familiar sounds presented in unfamiliar order, these context-sensitive elements may not be encoded firmly enough to achieve correct identification.

An alternative proposal, suggested by Warren (1974), is that the ready identification of familiar sequences is mediated by a two-stage process. In the first stage the sequence is recognized in global fashion: as a "temporal compound" which can be distinguished from other compounds without being analyzed into its components. Other factors in addition to the perception of relationships between strictly adjacent items could be involved in such global processing. In the second stage there takes place an item-by-item analysis of the components of this compound and their orders.

Judgments of temporal order for only two disparate sounds are easier than for continuous repetitive sequences. Hirsh (1959) and Hirsh and Sherrick (1961) found that the threshold for ordering two disparate events was around 20 msec for highly trained listeners, though somewhat higher for untrained listeners (Hirsh, 1976). This superior performance is probably based on several factors. First, items that are preceded or followed by silence are more readily identified than those that are not (Warren, 1974). Second, there are fewer relationships to be judged between two events. And, third, an active process which organizes elements according to sound type probably acts to inhibit the perception of relationships between disparate elements, in a fashion analogous to the process that organizes elements by frequency proximity. This effect should be expected to cumulate with repetition.

G. Grouping by Amplitude

Amplitude has been shown to be an effective grouping principle in the perception of rapid sequences of single tones. Dowling (1973) in his experiment on perception of interleaved melodies found that loudness differences between the melodies resulted in an enhanced ability to hear them as separate. Van Noorden (1975) studied perception of sequences where the tones were of identical frequency but alternated in amplitude. He found that with amplitude differences of less than 5 dB a single coherent stream was heard, even though loudness differences were clear. However, with larger amplitude differences two separate streams of different loudnesses were heard instead. Under these conditions attention could be directed to the softer stream as well as to the louder one. With even larger amplitude differences between the alternating tones, the auditory continuity effect was produced, and the softer tone was heard as though continuing through the louder tone (see below).

H. Grouping by Temporal Proximity

When we attend to one melodic configuration rather than to another, we are forming figure-ground relationships analogous to those in vision (Gregory, 1970). Perception of sequences of tones that are interleaved in time may then be likened to visual perception of embedded figures. Divenyi and Hirsh (1978) drew this analogy, and argued that melodic configurations may be represented in two dimensions, with frequency providing one dimension and time the other (see also Julesz & Hirsh, 1972). Just as visual configurations can be more readily identified when these are spatially separated from background stimuli, so should melodic configurations be more readily identified when these are separated either in time or in frequency from background tonal stimuli.

As a test of this notion, Divenyi and Hirsh presented rapid three-tone patterns that could occur in any of six permutations and required subjects to identify on each trial which permutation had been presented. These three-tone patterns were embedded in sequences of seven or eight tones, but were not interleaved with them. Identification performance was superior when the irrelevant tones and the target tones were in different frequency ranges. Furthermore, performance levels varied considerably depending on the temporal position of the target pattern within the full sequence. Best performance was obtained when the target pattern occurred at the end of the sequence; performance was also relatively good when the target was located at the beginning, but it was close to chance when the target occurred in the middle of the sequence. Both temporal and frequency separation were therefore found to reduce interference from the background tones. Previously Ortmann (1926) had found that a single tone was more salient when it was the highest or lowest in a sequence, or when it was in the first or last position. Similar conclusions were drawn recently by Watson, Kelly, and Wroton (1975) and Watson, Wroton, Kelly, and Benbasset (1976). Divenyi and Hirsh's results, therefore, extended such findings to the case of melodic configurations. (Further issues involving grouping by temporal position are discussed in Chapter 9.)

I. Perceptual Replacement of Missing Sounds

So far we have examined several instances where our perceptual system reorganizes sound sequences in accordance with expectations derived from both the sequences themselves and our knowledge of the auditory environment. It has also been found that sounds which are not actually present in the stimulus may be perceptually synthesized in accordance with such expectations.

Various studies have shown that when two sounds are presented in alternation, the fainter sound may be heard as continuing through the louder one (Miller & Licklider, 1950; Thurlow, 1957; Vicario, 1960). More recently, Warren (1970) and Warren, Obusek, and Ackroff (1972) showed that if a phoneme in a sentence is replaced by a louder noise, the missing phoneme may be perceptually synthesized. Analogous results were obtained with nonverbal sounds. In a set of parametric studies, Warren and his colleagues have demonstrated that this "auditory induction effect" occurs only under stimulus conditions where it would be reasonable to assume that the substituted sound had masked the missing one.

Dannenbring (1976) produced another version of this effect. He presented a sine wave tone that repeatedly glided up and down in frequency. When a loud noise burst was substituted for a portion of this sound, it still appeared to glide through the noise. However, if the tone changed in amplitude just before the noise burst, producing evidence that something had happened to the tone itself, rather than its simply being masked, the tendency to hear the tone as continuing through the noise was reduced (Bregman & Dannenbring, 1977).

V. VOLUNTARY ATTENTION

We now turn to a consideration of the effects of voluntary attention on channeling phenomena. In listening to music outside the laboratory, we have the impression that we can direct our attention at will; listening now to a melodic line, now to its accompaniment, now to a chosen instrument, and so on. Yet the conditions under which such attention focusing is indeed under voluntary control remain to be determined. We are dealing with two issues here. First, we may examine the role of voluntary attention in the initial division of the configuration into groupings. Second, we may examine the role of voluntary factors in determining which grouping is attended to, once such a division is established. Concerning the first issue, we have described several configurations where a particular grouping principle is so strong that listeners are generally unaware of alternative organizations. For example, most people on hearing the scale illusion form groupings so strongly on the basis of frequency proximity that they hear tones in one frequency range as emanating from one source, and tones in another frequency range as emanating from a different source. They therefore believe that they are attending to one spatial location rather than to another; yet in reality they are synthesizing information from two different locations (Deutsch, 1975b). The same is true for the two-part contrapuntal patterns devised by Butler (1979a). Similarly, on listening to the octave illusion, many people believe that a single high tone is being delivered intermittently to one ear, and a single low tone intermittently to the other ear. Yet in fact they are being presented with a continuous two-tone chord. So, here again involuntary organizational mechanisms are so strong that the listener is unaware of the nature of the stimulus configuration. The sequence of Kubovy *et al.* (1974) provides another example. Here one hears a melody as in one spatial location and a background noise as in another, yet in reality a continuous chord is being delivered to both ears.

However, when we consider the role of voluntary attention in determining which of two channels is attended to, once these have been formed, we find that in all these examples voluntary attention plays a prominent role. For example, in the scale illusion listeners who hear two melodic lines in parallel can choose at will to attend to either the higher or the lower one. Even those listeners who initially hear only the higher melodic line may after repeated presentations focus their attention on the lower one. Again in Butler's contrapuntal patterns we can choose at will to listen to the higher or the lower of the two melodies that we have perceptually synthesized. In the case of the octave illusion, those listeners who unambiguously hear a high tone in one ear alternating with a low tone in the other ear can focus their attention on either the high tone or the low one. Similarly, with the configuration of Kubovy *et al.* (1974), listeners can direct their attention to either the melody or to the noise.

When we consider channeling of rapid sequences of tones, we also find that strong involuntary factors are involved in the formation of initial groupings. Thus, the inability to form order relationships across streams based on frequency proximity (Bregman & Campbell, 1971), sound type (Warren *et al.*, 1969), or spatial location (Deutsch, 1979) cannot readily be overcome by voluntary attention focusing.

However, other examples have been given where voluntary attention does play a role. In exploring the temporal coherence boundary, Van Noorden (1975) found that within a given range of tempos and of frequency relationships, listeners may direct their attention at will, hearing either fission or temporal coherence (Fig. 12). An ambiguous situation where channeling by timbre was set in competition with channeling by pitch was created by Erickson (1974) in a compositon called *LOOPS*. Here, a repeating melodic pattern was performed by five instruments, with each instrument playing a different note in the manner of a hockett, so that each pitch was eventually played by every instrument. Under these conditions listeners can often choose to follow the sequence on the basis of either timbre or pitch. It therefore appears that although there are strong involuntary components in the formation of groupings, ambiguous stimulus situations may be set up where voluntary attention can be the determining factor.

Considering the issue of voluntary factors in determining which stream is attended to, once a set of alternatives have been formed, we find that voluntary attention focusing is easily achieved with rapid sequences also. For instance, Van Noorden (1975) reports that in cases where two streams were formed on the basis of frequency proximity, the listener was able to direct his attention at will and concentrate on either the upper stream or the lower one. However, he noted that an involuntary component was also present: The listener's percept would sometimes switch spontaneously to the stream he was attempting to ignore. This was true even when the unattended stream was less salient. Similar observations were made by present author using streaming by spatial location in patterns such as on Fig. 6.

In summary, it appears that the initial division of the stimulus configuration into groupings is often outside the listener's voluntary control, though ambiguous situations may be generated where attention focusing can be effective. In contrast, once a set of groupings is established, voluntary attention focusing plays a prominent role in determining which of these is attended to. This division of the attentional process into two stages corresponds in many respects to the stages identified as preattentive and postattentive by Neisser (1967) and Kahneman (1973) among others. These terms, however, have often been taken to imply different depths of analysis at these two stages, yet the issue of depth of analysis remains unsettled (Deutsch & Deutsch, 1963; Keele and Neill, 1979).

We may next consider the consequences of the selective attention process for the unattended material (i.e., for the component of music that serves as "ground" when attention is focused on another component which serves as "figure".[3] Considering the analogous issue for streams of speech, Cherry (1953) and Cherry and Taylor (1954) presented subjects with two messages, one to each ear, and required them to shadow one of these. They found that the subjects were able to report virtually nothing about the message presented to the nonattended ear, not even what language it was in. Other studies have produced similar findings (Kahneman, 1973). Recently, the present author set up an analogous situation for musical stimuli. Two familiar folk songs were recorded by piano, and were simultaneously presented, one to each ear. Listeners were required to shadow one of the melodies by singing, and were later asked to describe what had been presented to the other ear. Analogous to Cherry's finding, no listener was able to name the unattended melody, and none was able to describe much about the stimulus. Thus, voluntary attention focusing on one channel of a musical configuration can have the effect of suppressing the other channel from conscious perception.[4]

The question then arises as to what extent the unattended signal is processed under these conditions. This is a controversial issue in the literature on speech materials. Broadbent (1958) proposed that in selective listening a filter sorts out simultaneously presented stimuli on the basis of gross physical characteristics, such as spatial location or frequency range. Stimuli that share a characteristic that defines the relevant "chan-

[3]The author is indebted to R. Erickson for raising this question.

[4]Channeling by spatial location was here facilitated by the fact that the messages delivered to the two ears were asynchronous. Had they been synchronous, as with the stimuli used by Deutsch (1975b) and Butler (1979a), it would not have been possible for the listener to focus attention on one ear rather than the other.

nel" are then perceptually analyzed further, whereas the other stimuli are simply filtered out. This theory was found to be unable to account for certain findings—for instance, that the meaning of words may be important determinants of selective attention (Grey & Wedderburn, 1960; Treisman, 1960). Consequently, Treisman (1960, 1964) suggested a modification of filter theory to accommodate these findings. She proposed that the unattended message is not totally rejected, as Broadbent had suggested, but rather attenuated. An alternative view was taken by Deutsch and Deutsch (1963) who proposed that all input, whether attended to or not, is completely analyzed by the nervous system. The information thus analyzed is then weighted for importance or pertinence. Such weightings are determined both by long-term factors (for instance, there is a long-term predisposition to attend to one's own name) and by factors determined by the current situation. The information with the highest weighting of importance then controls awareness. Recent studies (e.g., Lewis, 1970; Corteen & Wood, 1972; Shiffrin & Schneider, 1977) have provided strong evidence for this view, but the issue remains controversial. In the case of music, this has not yet been the subject of experimental investigation.

VI. CONCLUSION

In this chapter we have focused on musical channeling phenomena in two types of situation. First, we have explored the perceptual consequences of presenting two simultaneous sequences of tones in different spatial locations. Second, we have investigated channeling when rapid sequences of single tones were presented. In general, relatively simple stimulus configurations were examined, and grouping or channeling on the basis of higher order abstractions was not considered. The formation of such abstractions is the subject of Chapter 9 and we shall assume that these can also serve as bases for grouping.[5]

REFERENCES

Bartholomeus, B. Effects of task requirements on ear superiority for sung speech. *Cortex*, 1974, *10*, 215-223.

Bartholomeus, B. N., Doehring, D. G., & Freygood, S. D. Absence of stimulus effects in dichotic singing. *Bulletin of the Psychonomic Society*, 1973, *1*, 171-172.

Benade, A. H. *Fundamentals of musical acoustics*. London and New York: Oxford University Press, 1976.

Berlin, C. I. Critical review of the literature on dichotic effects—1970. In 1971 reviews of scientific literature on hearing. *American Academy Ophtology Otology* 1972, 80-90.

Bever, T. G., & Chiarello, R. J. Cerebral dominance in musicians and nonmusicians. *Science*, 1974, *185*, 537-539.

Bregman, A. S. The formation of auditory streams. In J. Requin (Ed.), *Attention and Performance*. (Volume VII) Hillsdale, New Jersey: Erlbaum, 1978. Pp. 63-76.

Bregman, A. S., & Campbell, J. Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, 1971, *89*, 244-249.

[5]This work was supported by United States Public Health Service Grant MH-21001.

Bregman, A. S., & Dannenbring, G. L. The effect of continuity on auditory stream segregation. *Perception & Psychophysics*, 1973, *13*, 308-312.

Bregman, A. S., & Dannenbring, G. L. Auditory continuity and amplitude edges. *Canadian Journal of Psychology*, 1977, *31*, 151-159.

Bregman, A. S., & Pinker, S. Auditory streaming and the building of timbre. *Canadian Journal of Psychology*, 1978, *32*, 20-31.

Bregman, A. S., & Rudnicky, A. I. Auditory segregation: Stream or streams? *Journal of Experimental Psychology: Human Perception and Performance*, 1975, *1*, 263-267.

Broadbent, D. E. The role of auditory localization in attention and memory span. *Journal of Experimental Psychology*, 1954, *47*, 191-196.

Broadbent, D. *Perception and communication*. Oxford: Pergamon, 1958.

Butler, D. A further study of melodic channeling. *Perception & Psychophysics*, 1979, *25*, 264-268. (a)

Butler, D. Melodic channeling in a musical environment. *Research Symposium on the Psychology and Acoustics of Music, Kansas*, 1979. (b)

Charbonneau, G., and Risset, J-C. Differences entre oreille droite et oreille gauche pour la perception de la hauteur des sons. *Comptes Rendus, Académie des Sciences, Paris*, 1975, *281*, 163-166.

Cherry, E. C. Some experiments on the recognition of speech, with one and two ears. *Journal of the Acoustical Society of America*, 1953, *25*, 975-979.

Cherry, E. C., & Taylor, W. K. Some further experiments upon the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, 1954, *26*, 554-559.

Corteen, R. S., & Wood, B. Autonomic responses to shock-associated words in an unattended channel. *Journal of Experimental Psychology*, 1972, *94*, 308-313.

Dannenbring, G. L. Perceived auditory continuity with alternately rising and falling frequency transitions. *Canadian Journal of Psychology*, 1976, *30*, 99-114.

Danndenbring, G. L., & Bregman, A. S. Stream segregation and the illusion of overlap. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, *2*, 544-555.

Darwin, C. J. Auditory Perception and Cerebral Dominance. Doctoral dissertation, University of Cambridge, 1969.

Deutsch, D. An auditory illusion. *Journal of the Acoustical Society of America*, 1974, *55*, S18-S19. (a)

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

Deutsch, D. Musical illusions. *Scientific American*, 1975, *233*, 92-104. (a)

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

Deutsch, D. Lateralization by frequency in dichotic tonal sequences as a function of interaural amplitude and time differences. *Journal of the Acoustical Society of America*, 1976, *60*, S50.

Deutsch, D. Binaural integration of tonal patterns. *Journal of the Acoustical Society of America*, 1978, *64*, S146. (a)

Deutsch, D. Lateralization by frequency for repeating sequences of dichotic 400-Hz and 800-Hz tones. *Journal of the Acoustical Society of America*, 1978, *63*, 184-186 (b)

Deutsch, D. Binaural integration of melodic patterns. *Perception & Psychophysics*, 1979, *25*, 399-405.

Deutsch, D. Two-channel listening to tonal sequences. In R. S. Nickerson and R. W. Pew (Eds.), *Attention and performance*. (Volume VIII) Hillsdale, New Jersey: Erlbaum, 1980.

Deutsch, D. The octave illusion and auditory perceptual integration. In J. V. Tobias and E. D. Schubert (Eds.), *Hearing research and theory*. (Volume I). Academic Press: New York, 1981.

Deutsch, D., *Left ear advantage for dichotic tonal sequences: an artifact of the scale illusion*. In preparation.

Deutsch, D., & Roll, P. L. Separate 'what' and 'where' decision mechanisms in processing a dichotic tonal sequence. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, *2*, 23-29.

Deutsch, J. A., & Deutsch, D. Attention: Some theoretical considerations. *Psychological Review*, 1963, *70*, 80-90.

Divenyi, P. L., & Hirsh, I. J. Discrimination of the silent gap in two-tone sequences of different frequencies. *Journal of the Acoustical Society of America*, 1972, *52*, 166S.

Divenyi, P. L., & Hirsh, I. J. Identification of temporal order in three-tone sequences. *Journal of the Acoustical Society of America*, 1974, *56*, 144-151.

Divenyi, P. L., & Hirsh, I. J. Some figural properties of auditory patterns. *Journal of the Acoustical Society of America*, 1978, *64*, 1369–1386.

Doehring, D. G. Discrimination of simultaneous and successive pure tones by musical and nonmusical subjects. *Psychonomic Science*, 1971, *22*, 209–210.

Doehring, D. G. Ear asymmetry in the discrimination of monaural tonal sequences. *Canadian Journal of Psychology*, 1972, *26*, 106–110.

Dorman, M. F., Cutting, J. E., & Raphael, L. J. Perception of temporal order in vowel sequences with and without formant transitions. *Journal of Experimental Psychology: Human Perception and Performance*, 1975, *104*, 121–129.

Dowling, W. J. Rhythmic Fission and the Perceptual Organization of Tone Sequences. Unpublished doctoral dissertation. Harvard University, Cambridge, Massachusetts, 1967.

Dowling, W. J. The perception of interleaved melodies. *Cognitive Psychology*, 1973, *5*, 322–337.

Ehrenfels, C. Von. *Uber Gestaltqualitäten Vierteljahrschrift fur Wissenschaftliche Philosophie*, 1890, *14*, 249–292.

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

Erickson, R. LOOPS, an informal timbre experiment, Center for Music Experiment, University of California, San Diego, 1974.

Fitzgibbon, P. J., Pollatsek, A., & Thomas, I. B. Detection of temporal gaps within and between perceptual tonal groups. *Perception & Psychophysics*, 1974, *16*, 522–528.

Gaede, S. E., Parsons, O. A., and Bertera, J. H. Hemispheric differences in music preparation: aptitude vs. experience. *Neurophychologia*, 1978, *16*, 369–373.

Gordon, H. W. Hemispheric asymmetries in the perception of musical chords. *Cortex*, 1970, *6*, 387–398.

Gray, J. A., & Wedderburn, A.A.I. Grouping strategies with simultaneous stimuli. *Quarterly Journal of Experimental Psychology*, 1960, *12*, 180–184.

Gregory, R. L. *The intelligent eye*. New York: McGraw-Hill, 1970.

Hass, H. Über den einfluss eines Einfachechos auf die Hörsamkeit von Sprache. *Acustica*, 1951, *1*, 49–52.

Halperin, Y., Nachshon, I., & Carmon, A. Shift of ear superiority in dichotic listening to temporally patterned nonverbal stimuli. *Journal of the Acoustical Society of America*, 1973, *53*, 46–50.

Heise, G. A., & Miller, G. A. An experimental study of auditory patterns. *American Journal of Psychology*, 1951, *64*, 68–77.

Hirsh, I. J. Auditory perception of temporal order. *Journal of the Acoustical Society of America*, 1959, *31*, 759–767.

Hirsh, I. J. Order of events in three sensory modalities. In S. K. Hirsh, D. H. Eldridge, I. J. Hirsh, & S. R. Silverman (Eds.), *Essays honoring Hallowell Davis*, St. Louis, Missouri: Washington University Press, 1976.

Hirsh, I. J., & Sherrick, C. E. Perceived order in different sense modalities. *Journal of Experimental Psychology*, 1961, *62*, 423–432.

Hochberg, J. Organization and the Gestalt Tradition. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception*. (Volume 1) New York: Academic Press. Pp. 180–211.

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

Johnson, P. R. Dichotically-stimulated ear differences in musicians and nonmusicians. *Cortex*, 1977, *13*, 385–389.

Judd, T. Comments on Deutsch's musical scale illusion. *Perception and Psychophysics*, 1979. *26*, 85–92.

Julesz, B., & Hirsh, I. J. Visual and auditory perception—An essay of comparison. In E. E. David and P. B. Denes (Eds.), *Human communication: A unified view*. New York: McGraw-Hill, 1972. Pp. 283–340.

Kahneman, D. *Attention and effort*. Englewood Cliffs, New Jersey: Prentice-Hall, 1973.

Keele, S. W., and Neill, W. T. Mechanisms of attention. In E. C. Carterette and M. P. Friedman (Eds.), *Handbook of Perception* (Vol. 9) New York: Academic Press, 1979.

Kimura, D. Left-right differences in the perception of melodies. *Quarterly Journal of Experimental Psychology*, 1964, *16*, 355–358.

Kimura, D. Functional asymmetry of the brain in dichotic listening. *Cortex*, 1967, *3*, 163–178.

King, F. D., & Kimura, D. Left-ear superiority in dichotic perception of vocal nonverbal sounds. *Canadian Journal of Psychology*, 1972, *26*, 111–116.

Kubovy, M. Concurrent pitch-segregation and the theory of indispensable attributes. In M. Kubovy and J. Pomerantz (Eds.), *Perceptual organization.* Hillsdale: Erlbaum, New Jersey, 1981.

Kubovy, M., Cutting, J. E., & McGuire, R. M. Hearing with the third ear: Dichotic perception of a melody without monaural familiarity cues. *Science,* 1974, *186,* 272–274.

Kubovy, M., & Howard, F. P. Persistence of a pitch-segregating echoic memory. *Journal of Experimental Psychology: Human Perception and Performance,* 1976, *2,* 531–537.

Lewis, J. L. Semantic processing of unattended messages using dichotic listening. *Journal of Experimental Psychology,* 1970, *85,* 225–228.

McNally, K. A., & Handel, S. Effect of element composition on streaming and the ordering of repeating sequences. *Journal of Experimental Psychology: Human Perception and Performance,* 1977, *3,* 451–460.

Miller, G. A., & Heise, G. A. The trill threshold. *Journal of the Acoustical Society of America,* 1950, *22,* 637–638.

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

Moray, N. Attention in dichotic listening; Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology,* 1959, *11,* 56–60.

Moray, N. A date base for theories of selective listening. In P.M.A. Rabbitt and S. Dornic (Eds.), *Attention and performance.* (Volumn V) New York: Academic Press, 1975.

Nabelek, I. V., Nabelek, A. K., & Hirsh, I. J. Pitch of sound bursts with continuous or discontinuous change of frequency. *Journal of the Acoustical Society of America,* 1973, *53,* 1305–1312.

Natale, M. Perception of neurolinguistic auditory rhythms by the speech hemisphere. *Brain and Language,* 1977, *4,* 32–44.

Neisser, U. *Cognitive Psychology.* New York: Appleton, 1967.

Nickerson, R. S., & Freeman, B. Discrimination of the order of the components of repeating tone sequences: Effects of frequency separation and extensive practice. *Perception & Psychophysics,* 1974, *16,* 471–477.

Ortmann, O. On the melodic relatively of tones. *Psychological Monographs,* 1926, *35* (whole No. 162).

Papcun, G., Krashen, S., Terbeek, D., Remington, R., & Harshman, R. Is the left hemisphere specialized for speech, language and/or something else? *Journal of the Acoustical Society of America,* 1974, *55,* 319–327.

Penfield, W., & Perot, P. The brain's record of auditory and visual experience. *Brain,* 1963, *86,* 595–696.

Pollack, I. Temporal switching between binaural information sources. *Journal of the Acoustical Society of America,* 1978, *63,* 550–558.

Rasch, R. A. The perception of simultaneous notes such as in polyphonic music. *Acustica,* 1978, *40,* 1–72.

Robinson, G. M., & Solomon, D. J. Rhythm is processed by the speech hemisphere. *Journal of Experimental Psychology,1974, 102,* 508–511.

Schouten, J. F. On the perception of sound and speech; Subjective time analysis. *Fourth International Congress on Acoustics, Copenhagen Congress Report II,* 1962, 201–203.

Shiffrin, R. M., Pisoni, D. B., & Castaneda-Mendez, K. Is attention shared between the ears? *Cognitive Psychology,* 1974, *6,* 190–215.

Shiffrin, R. M., & Schneider, W. Toward a unitary model for selective attention, memory scanning and visual search. In S. Dornic (Ed.), *Attention and performance.* (Volume VI) Hillsdale: Earlbaum, 1977. Pp. 413–440.

Sorkin, R. D., Pastore, R. E., & Pohlmann, L. D. Simultaneous two-channel signal detection. II. Correlated and uncorrelated signals. *Journal of the Acoustical Society of America,* 1972, *51,* 1960–1965.

Sorkin, R. D., Pohlmann, L. D., & Gilliom, J. D. Simultaneous two-channel signal detection. III. 630- and 1400-Hz signals. *Journal of the Acoustical Society of America,* 1973, *14,* 101–109.

Spellacy, F. Lateral preferences in the identification of patterned stimuli. *Journal of the Acoustical Society of America,* 1970, *47,* 574–578.

Spreen, O., Spellacy, F., and Reid, J. R. The effect of interstimulus interval and intensity on ear asymmetry for nonverbal stimuli in dichotic listening. *Neurophychologia,* 1970, *8,* 245–250.

Sutherland, N. S. Object recognition. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of Perception.* (Volume III) New York: Academic Press, 1973. Pp. 157–186.

Thomas, I. B., Cetti, R. P., & Chase, P. W. Effect of silent intervals on the perception of temporal order for vowels. *Journal of the Acoustical Society of America*, 1971, *49*, 584.

Thomas, I. B., Hill, P. B., Carroll, F. S., & Garcia, B. Temporal order in the perception of vowels. *Journal of the Acoustical Society of America*, 1970, *48*, 1010–1013.

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

Tobias, J. V. Curious binaural phenomena. In J. V. Tobias (Ed.), *Foundations of modern auditory theory*. (Volume II) New York: Academic Press, 1972.

Treisman, A. M. Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, 1960, *12*, 242–248.

Treisman, A. M. Selective attention in man. *British Medical Bulletin*, 1964, *20*, 12–16.

Treisman, A. M. Shifting attention between the ears. *Quarterly Journal of Experimental Psychology*, 1971, *23*, 157–167.

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

Vicario, G. L'effetto tunnel acustico. *Revista di Psyicologia*, 1960, *54*, 41–52.

Von Helmholtz, H. *On the sensations of tone as a physiological basis for the theory of music*. (2nd English ed.) New York: Dover, 1954. (Originally published 1859)

Von Helmholtz, H. *Helmholtz's physiological optics*. (Translated from the 3rd German ed.) (1909–1911 byJ.P.C. Southall, ed.) Rochester, New York: Optical Society of America, 1925.

Wallach, H., Newman, E. B., & Rosenzweig, M. R. The precedence effect in sound localization. *American Journal of Psychology*, 1949, *62*, 315–336.

Warren, R. M. Perceptual restoration of missing speech sounds. *Science*, 1970, *167*, 392–393.

Warren, R. M. Auditory temporal discrimination by trained listeners. *Cognitive Psychology*, 1974, *6*, 237–256.

Warren, R. M., & Byrnes, D. L. Temporal discrimination of recycled tonal sequences: Pattern matching and naming of order by untrained listeners. *Journal of the Acoustical Society of America*, 1975, *18*, 273–280.

Warren, R. M., Obusek, C. J., & Ackroff, J. M. Auditory induction: Perceptual synthesis of absent sounds. *Science*, 1972, *176*, 1149–1151.

Warren, R. M., Obusek, C. J., Farmer, R. M., & Warren, R. P. Auditory sequence: Confusions of patterns other than speech or music. *Science*, 1969, *164*, 586–587.

Warren, R. M., & Warren, R. P. Auditory illusions and confusions. *Scientific American*, 1970, *223*, 30–36.

Watson, C. S., Kelly, W. J., & Wroton, H. W. Factors in the discrimination of tonal patterns. II. Selective attention and learning under various levels of uncertainty. *Journal of the Acoustical Society of America*, 1976, *60*, 1176–1186.

Watson, C. S., Wroton, H. W., Kelly, W. J., & Benbasset, C. A. Factors in the discrimination of tonal patterns. I. Component frequency, temporal position and silent intervals. *Journal of the Acoustical Society of America*, 1975, *75*, 1175–1185.

Wertheimer, M. Untersuchung zur Lehre von der Gestalt II. *Psychologische Forschung*, 1923, *4*, 301–350.

Wickelgren, W. A. Context-sensitive coding, associative memory and serial order in (speech) behavior. *Psychological Review*, 1969, *76*, 1–15.

Wickelgren, W. A. Phonetic coding and serial order. In E. C. Carterette and M. P. Friedman (Eds.), *Handbook of Perception*. (Volume VII) New York: Academic Press, 1976. Pp. 227–264.

Williams, K. N., and Perrott, D. R. Temporal resolution of tonal pulses. *Journal of the Acoustical Society of America*, 1972, *51*, 644–647.

Zatorre, R. J. Recognition of dichotic melodies by musicians and nonmusicians. *Neurophychologia*, 1979, *17*, 607–617.