

Analyzing Spectra That Won't Stand Still

Albert S. Bregman

Number 3, 1982

URI: <https://id.erudit.org/iderudit/1013826ar>

DOI: <https://doi.org/10.7202/1013826ar>

[See table of contents](#)

Publisher(s)

Canadian University Music Society / Société de musique des universités canadiennes

ISSN

0710-0353 (print)

2291-2436 (digital)

[Explore this journal](#)

Cite this article

Bregman, A. S. (1982). Analyzing Spectra That Won't Stand Still. *Canadian University Music Review / Revue de musique des universités canadiennes*, (3), 31–48. <https://doi.org/10.7202/1013826ar>

ANALYZING SPECTRA THAT WON'T STAND STILL

Albert S. Bregman

A common interest that unites many psychologists and musicologists is the desire to explain musical phenomena by means of general laws of perception and cognition. One would imagine that the first place to turn for such explanations would be to the discipline of auditory psychophysics. However, we have not received much real help from this direction. Classical psychophysics has been primarily concerned with very simple accomplishments of the auditory system in the face of simple stimuli. It has, for example, been obsessively concerned with the problem of sensitivity — how intense a sound has to be before we can hear it. Another of its basic concerns has been the proportionality between the physicist's scales for measuring sound and the listener's one-dimensional experiences, such as loudness and pitch. It is this latter effort that has given us the mel scale, a pitch metric that so far seems to correlate with nothing much in music. Another music goal has been to understand masking, that is, what happens when one sound drowns out another.

In all of this there is no mention of pattern and yet we know that the musician's stock-in-trade is that elusive factor. Fortunately, quite recently there has been an upsurge in the auditory psychologists's interest in and research on pattern. Why this change of heart? It seems to be the result of several recent influences. The first is the new interest in rule systems stimulated by the significant success attained by linguists such as Chomsky (e.g., 1965) in describing the formal patterns in the syntactical aspect of language. Using the linguists' formal descriptions, psychologists were able to show the effects of linguistic rules upon perception, learning, and memory. It is in this tradition that the research of Cuddy, Cohen, and Miller (1979) lies.

A second influence has come from theoretical developments in acoustics that make human speech approachable from an acoustical point of view. Voice is now understandable as hiss and buzz sounds being fed into a series of resonators. This insight has led to a stream of experiments from Haskins Laboratories and elsewhere in which listeners are presented with synthesized speech whose every acoustic parameter is carefully controlled. At last we have begun to study truly interesting patterns using the methods of experimental science.

This last effort would not have been possible, however, without yet a third influence — the spectacular developments in acoustic equipment. Having moved on from the oscillators and amplifiers that spurred on classical psychophysics, we are now armed with an array of powerful tools for the analysis and shaping of sound. At present, a psychologist can, with the help of a friendly computer, create stimuli with known harmonic structure, known dynamics, and known patterning in time, all precise and highly reproducible.

A fourth influence on the study of auditory patterning has been the attempts of engineers to create machines that see (see Winston 1975). One of the major effects of this enterprise has been the realization that even when the elementary inputs (i.e., points of light) are clearly registered on the sensing apparatus of the machine, the problems of recognition have only begun. What this implies is that even if auditory psychophysicists were to come to a successful conclusion in their research on questions of minimal audible loudness and the like, still the question of how we recognize and respond to meaningful sound would be no nearer to being solved.

These four influences, however, seem to be on the verge of creating a psychology of perception that may have something to say to musicians. Because this is a musical audience, I would like to begin with an illustration that may have some musical interest as sound. Suppose we create an endlessly repeating cycle of six tones. Three of these are high tones (2500, 2000, and 1600 Hz) and three are low (550, 430, and 350 Hz). An example of the order of high (H) and low (L) tones would be HLHLL. If such a cycle is made very fast, say one-tenth of a second per tone, we hear a remarkable thing; it appears that the high tones are perceptually segregated from the low ones. We can pay attention to either the high or the low tones but not both. Research has shown that it is easy to detect the order of the high tones relative to one another or

the low ones relative to one another, but very hard to locate the high tones in their temporal relation to the low ones (see Bregman & Campbell 1971). This effect is related to the compound melodic lines produced by composers of the Baroque period by having a single instrument alternate between a high and a low melodic line. The segregation of the two lines in these compositions is not complete because the alternation is a little too slow. You would need a playing speed of about ten tones per second to get a good level of segregation. It has been found that the segregation is strengthened by two factors: speeding up the sequence and increasing the pitch separation (really the frequency separation) of the high and low lines (see Van Noorden 1975). These two factors can trade off their effects: a sequence whose high and low tones are far apart in pitch, say fifteen semitones or more, will split at the slow rate of about five tones per second, whereas high and low tones which are closer together, say less than five semitones, will require speeds of more than ten tones per second before they segregate. Even if the Baroque composers had known how to produce complete segregation and been able to achieve it technically, they probably would not have done so because all they would have achieved would have been an economizing on instrumentalists. The fact that no two real players could have precisely interleaved their notes at that rate would have been lost on the audience anyway. The splitting apart of the streams would have destroyed the perception of interleaving even if interleaving were perfect. The alternation, at the rate at which it appears in Baroque music, produces an ambiguous percept, somewhere between integration and segregation, and I imagine that it is the playing upon this ambiguity, the awareness of both the separation and the inter-relationship, that creates the musical interest.

So we have finally begun to study grouping. While it is evident that this is a subject of critical importance to musicians, it is not commonly realized how central it has to be in the general theory of perception. We can come to appreciate its centrality by considering the role that perception plays in human life. I think of the senses as creating in our brains a model of reality, a model that permits us to generate expectations, to plan, and generally to harmonize our actions with the structure of the world around us.

In creating such a model, our brains face a fundamental problem: the simple aspects of the world are not represented simply on the sense organ. Consider Figure 1:

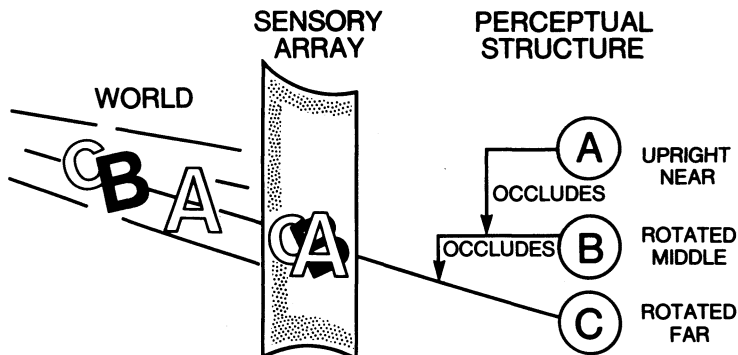


Figure 1

A schematic view of human perception illustrating the separate existence of the world's objects, the mixture of evidence as it reaches the sense organ, and the recovery of separate descriptions in the mind

In this diagram, there are three simple things in the world, an A, a B, and a C. But because of their relations to one another and to the observer, they produce on some sensory array (such as a retina) a complex pattern in which the effects of the separate A, B, and C are composed in a complex way. The visible parts of the B, for example, are separated from one another. Furthermore, the contours of the visible parts of the B are partly due to the shape of the A which lies in front of it. Yet at some deep level of mental representation, we want to represent the three letters as distinct entities and to interpret the complex retinal evidence as arising from the composition of the basic shapes of these letters, transformed by rotation and by the covering of some of the letters by others. In achieving this mental description of the input, our brains essentially reconstruct the retinal evidence in terms of simpler patterns (see Bregman 1977). Linguists such as Chomsky (1965) have referred to a similar process in language as arriving at the “deep structure” of the sentence. In vision, we are arriving at the deep structure of a scene.

Computer scientists who try to program computers to “see” and recognize objects have had to face the similar problem of figuring out how to calculate which areas in a photograph are part of the same object, given the fact that the presence of other objects or shadows can cause the areas arising from a single object to be disconnected (as they are with the letter B in Figure 1). This

problem has been given the name “scene analysis” and has attracted much research (e.g., Winston 1975). Scene analysis is equivalent to coloring in, using the same color, all the surfaces of the same object in a line drawing of a complex scene. It turns out not to be as easy as it might seem at first blush. Yet it is a critical step if objects are to be recognized correctly. In order to recognize one of a set of simultaneously presented patterns (perhaps objects, perhaps sounds), we have to know which of the set of presented parts should be considered to be part of the same pattern. Only then can we enter our mental “dictionaries” and look up the identity of the pattern. Figure 2 shows, in the upper section, an example of a printed message:

AI CSAITT STIOTOS

A₁ C_SA_IT_T S_TI_OT_OS

Figure 2

A message that can only be recognized when its constituent patterns are segregated from one another

It is uninterpretable because it is really a mixture of two messages. Only when we can group the components belonging to each of the two messages, as we are able to do in the lower half of the figure, can we recognize them.

There is a tradition in psychology that has always been concerned with grouping, namely Gestalt psychology (e.g., Koffka 1935). I would like to discuss some of the principles of grouping that it proposed and to relate them to the listener's organization of sound and to the computer scientist's problem of scene analysis. Gestalt psychology can be related to scene analysis by the claim that the Gestalt principles of grouping are ones that our brains have evolved in order to solve the scene analysis problem. One way to show this is by examining some examples.

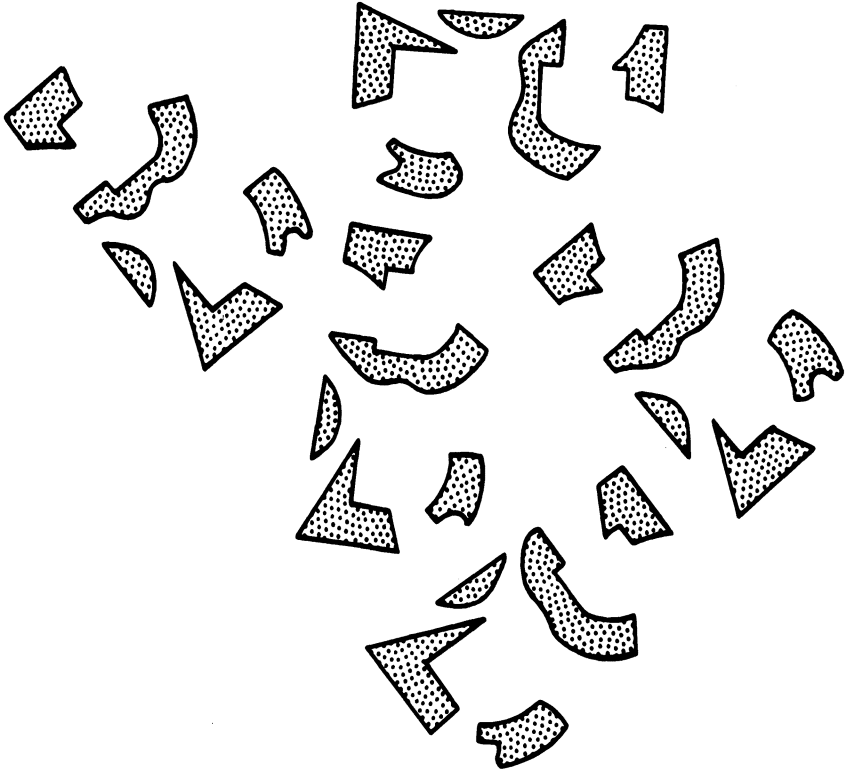


Figure 3
 Disconnected fragments of a familiar pattern

Figure 3 shows some disconnected fragments that are uninterpretable on a first viewing. Figure 4 shows the same fragments exactly, with the addition of an overlapping blob. All that the blob does is to explain to the visual system that the fragments arose by the occlusion of some simple forms by an overlapping form. The visual system can then calculate which contours of the fragments are really the result of the shape of the underlying forms and which ones are produced by occlusion. It can then perform what the Gestalt psychologists called "closure," the tendency for the brain to complete forms for whom the visual evidence is incomplete. The grouping principles that the Gestaltists evoked to explain closure included similarity of areas and good continuation of lines. In Figure 4 we see that some of the parts of each of the B's does have contours that line up with one another and are in good continuation:

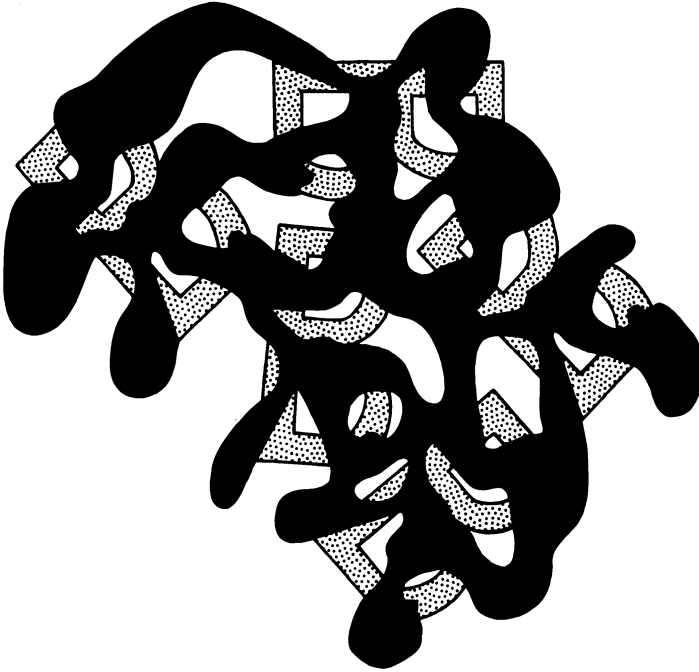


Figure 4

The same fragments as in Figure 3, except that the missing parts are “explained” by an occluding form, permitting perceptual completion to take place

While this fact undoubtedly aids the grouping of the fragments of that B, it is not sufficient in itself to create the correct grouping. We need the information about occlusion. Thus, the Gestalt principles come alive most strongly when combined with other information needed in scene analysis. The Gestalt principles, it seems, are only a small part of a larger “scene analysis package.”

While the example I gave was drawn from vision, it is just as easy to give one from audition. Gary Dannenbring (1974) studied a phenomenon which closely resembles the closure example of Figures 3 and 4. He studied the perception of a repeating rising and falling glide pattern that was broken (i.e., interrupted by silence) as shown in the upper half of Figure 5:

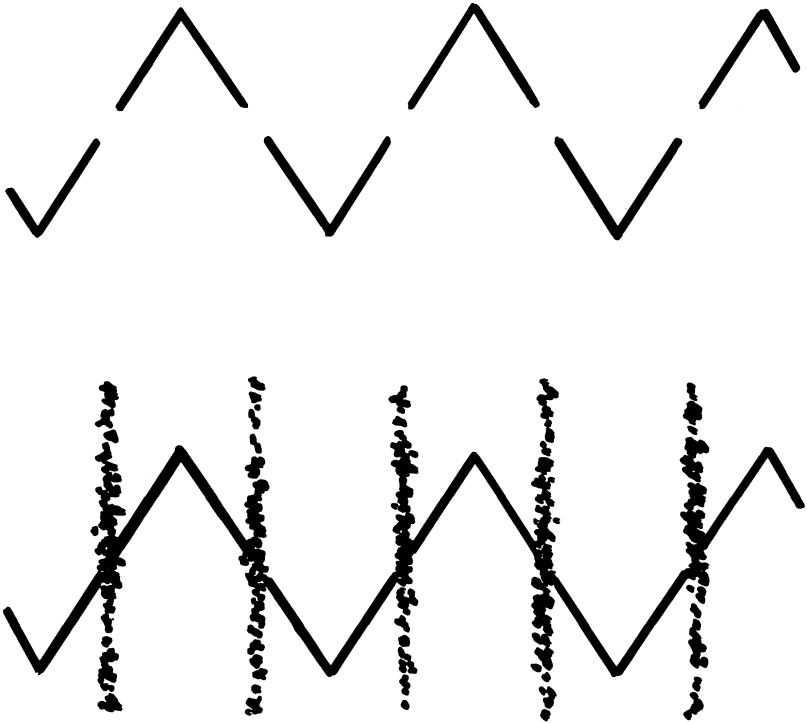


Figure 5

Diagram of a pure tone gliding up and down in pitch and broken at certain intervals (upper half); when the breaks are filled with loud noise (lower half), the tone is heard to glide through the noise

With such a stimulus, you can hear the breaks in the glides quite clearly. He then filled these breaks with loud noises as shown in the lower half of the figure. Under these conditions, you can hear the tone glide right through the noise. The parallel with the demonstration of Figures 3 and 4 is striking. The glides can only be heard as continuous when a noise is present that would have masked the missing portions of the glides if they really had been present (see Warren, Obusek, & Ackroff 1972). If the inserted noise burst is too soft or of the wrong frequencies or at a slightly wrong temporal position to have masked the missing portions, the perceptual closure does not occur. Therefore, closure in audition, just as in vision, is a sort of scene analysis in which the

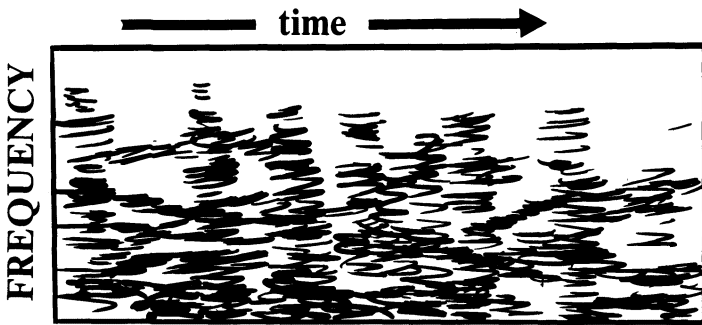
incompleteness of the audible evidence is accounted for by the assumption that simple sources of sound are present but that our ear only hears the net result of their mixture.

Scene analysis turns out to be every bit as important in audition as in vision. Despite the insistence of psychophysicists on presenting single tones in isolation to their subjects, the world outside the laboratory faces the listener with a mixture of simultaneous sound sources, as any parent with several small children, a dog, a television set, and a telephone can readily verify. How then can we ever recover a perceptual description of the sound of a single one of these sources, uncontaminated by the others? The same problem is encountered in musical listening. How is it possible to listen to only the oboe part in a large ensemble of instruments?

The acoustic analysis of sound depends very strongly on the work of Jean Baptiste Joseph Fourier (1768-1830) who proved a remarkable theorem. It said that any finite waveform, or any infinitely repeating waveform, can be represented as the sum of a set of sine waves, each with its particular frequency and phase. If the intensity of each frequency is represented by levels of darkness and a plot is made of these intensities over time, we have a spectrogram (also called a sonogram or voiceprint). There is some reason to believe that the ear presents the brain with what amounts to a spectrogram. Different positions along the basilar membrane of the inner ear resonate most strongly to particular frequencies, and therefore the neural output from any small region reflects the intensity of a particular range of frequencies. Suppose then that the brain could look at the spectrogram shown in Figure 6.

It is meant to represent a mixture of several sources of sound. But how many sources are there and what is the set of frequencies in each? Separating the voices acoustically is equivalent to the problem of taking this spectrogram and coloring in with the same color all frequency components arising from the same source of sound (say a person's voice, or a car, or the radio). If there are at most four different sources of sound and the auditory mechanism can resolve twenty different frequency regions, there are over a trillion ways to assign frequency components to sources at each moment of time. However, doing it properly is essential if pitches, timbres, and melodies, for example, are to be correctly recognized.

In visual scene analysis, the computer is looking for objects, not edges, or surfaces, or colors. Edge shapes, surface shapes, and



Voice + Car + Radio +

Figure 6

A spectrogram of a mixture of sound sources

colors are to be considered as properties of objects rather than as standing alone. Similarly, in audition we need an entity around which features can cohere. I believe that this entity exists and can be called a *stream* (see Bregman 1978). An example is a voice, whenever that voice can be segregated from its context. The musical equivalent would be any melodic line or “layer” that can be resolved by the listener as he or she listens to a complex composition. Such a stream, or “voice,” serves as a center for a set of descriptive features. We might say that it displays a certain series of pitches, timbral changes, loudnesses, has a certain rhythm, and so on. In this way, it serves the same role as *object* does in vision.

Perceptual theorists are fond of describing constancies in perception that occur despite changes in stimulation. One example is color constancy. An egg looks white indoors or in bright sunlight, even though it might be reflecting hundreds of times more light energy in the latter case. Yet nobody seems to have noticed the direct analogy in hearing; namely, timbre constancy. Since my voice, for example, is heard in a variety of acoustic contexts, the spectrum of the total sound of voice plus context

might be different on each occasion. Yet, if the timbre of a source of sound is dependent on the spectrum of that source, then to recognize the timbre of my voice in different contexts, a kind of scene analysis is necessary in which the spectral components of my voice are lifted out of the total spectrogram. Timbre constancy requires this sort of scene analysis to be achieved.

In recent years, we have done a number of experiments with a sort of perceptual task that can be thought of as mini-scene-analysis. The stimulus shown in Figure 7 was used by Steven Pinker and myself for this purpose (see Bregman & Pinker 1978):

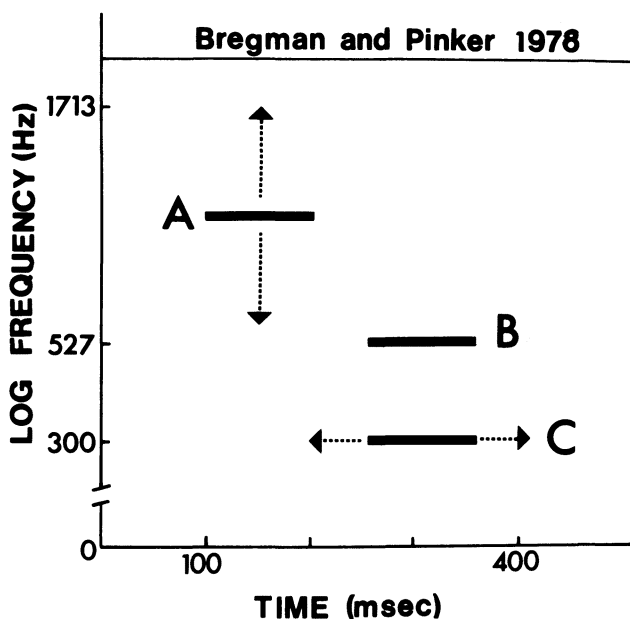


Figure 7

A single pure tone A is followed by a pair of pure tones, B and C, in a repeated cycle; B will fuse or not fuse with C depending both on its temporal relation with C and on its frequency relation with A

A sound formed of a pair of concurrent pure tones, B and C, is preceded by a single pure tone, A, and this pattern of sounds is presented rapidly and repeatedly in a closed cycle. The resulting percept is ambiguous. It can be heard as a cycle formed of two high

tones, A and B, accompanied by a slower cycle formed from the repetitions of a lower tone, C. In this case, three tones are heard, all with pure-tone quality. The other way of hearing the cycle is as an alternation of a pure tone A, with a rich tone formed by the mixture of B and C. These are two different ways of analyzing the evidence. We can bias the listener toward one analysis or the other by manipulating two important properties of this pattern, the frequency jump that separates A from B and the synchrony of onset of B and C. As A moves closer to B in frequency, it captures B more strongly and breaks the B-C bond. This is an example of the Gestalt principle of grouping by similarity. On the other hand, as B and C become more synchronous in onset, they tend to fuse more strongly, breaking the A-B bond. This latter is an example of what the Gestaltists called "grouping by common fate," the tendency of perceptual elements to form groups whenever they are changing at the same time and in the same way. In this experiment, B and C are coming on and going off together, thus exhibiting "common fate."

This principle is very important and worth discussing further. It is easy to create a visual example with the aid of transparent plastic sheets. We draw a random cluster of dots on each of two sheets and then overlay them. The result is seen as a denser cluster but still as a single cluster. Then we can start moving the top sheet around while holding the bottom sheet still. Instantly, we see two clusters of dots, one moving and the other stationary. The common fate, or type of motion, of each cluster of dots binds its elements together. We can create a similar demonstration with sound (along the lines of an unpublished experiment carried out with Lynn Halpern). Suppose we create a dense sound in which eight harmonics, for example, are gliding up and down together in a continuous glissando-like pattern. This is heard as a single sound gliding up and down. However, if we make a subset of four harmonics move up and down in pitch together while the other four are steady or follow some different pattern of modulation, we hear two distinct sounds, each created by the subset of harmonics that is exhibiting the same pattern of frequency change. Again common fate helps our brain to understand which elements of a mixture go together.

Another Gestalt principle is called *belongingness*. It says that a perceptible feature always belongs to the same whole and that a particular feature registered on our senses (e.g., a particular line in a drawing) is allocated to only one of the wholes (or

objects) that the brain is forming in its description of the scene. An example can be seen in Figure 1. The edges common to the B and the A on the sensory array (i.e., the contours formed where the A overlaps the B) are assigned to the A not to the B. That is, they help to define the shape of the A, not the B. This allocation of evidence can be seen as a part of scene analysis; it seems to fill the requirement that parts of the sensory evidence must be seen as arising from one real-world object or another. In addition, this requirement takes the form of allocating the audible elements to one perceptual stream or another. Depending upon how the information is allocated, a particular stream will have different audible elements in it. An example of this was explored in an experiment that I carried out with Alex Rudnicky (see Bregman & Rudnicky 1975). The format is shown in Figure 8:

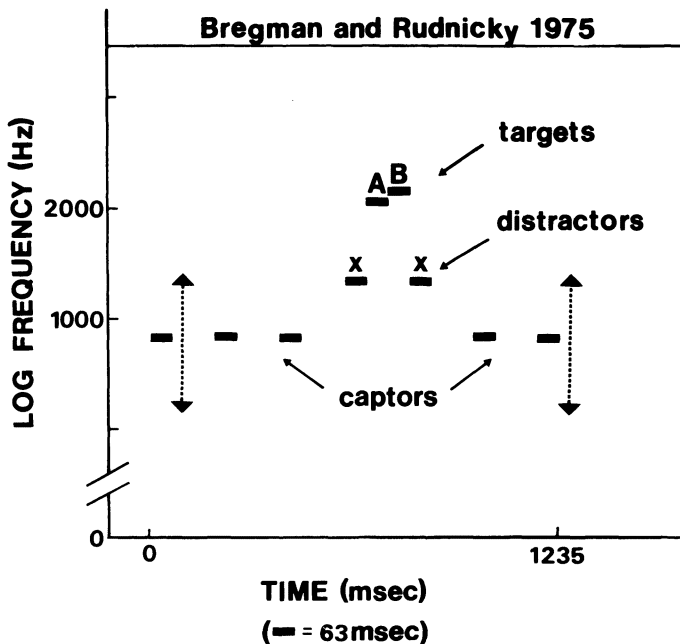


Figure 8
Example of an experiment in which the grouping of tones follows
the principle of belongingness

A pair of tones, A and B is played in isolation. Then A and B are played again as part of a longer sequence of tones as shown in Figure 8 and the listener is asked whether this occurrence of A and B was in the same order or in the reversed order as compared with the first presentation. If the second repetition (the "targets" of Figure 8) is played without the flanking tones, the listener easily makes the correct judgment. However, if the second occurrence of the AB pair is flanked by just the two tones labelled X in the figure (the distractors), the task becomes extremely difficult, clearly because A and B have become absorbed into a four-tone group that starts and ends on the same pitch. However, we can pull the distractor pair away from the targets by adding to the sequence a third sequence of tones called captors. This sequence has the same timing as the pair of distractor tones and has a space in it placed such that if the two distractors were heard as part of the captor stream a smooth isochronous rhythmic sequence would result. The outcome depends upon how close the pitch of the captors is to the pitch of the distractors. The closer they come, the more the captors and distractors form a single perceptual stream, releasing A and B to form their own stream. When A and B are released, their order is heard clearly. This long example illustrates the principle of belongingness. As long as the distractors belonged to the stream containing A and B, the A-B order was hard to hear. However, when they became part of the captor stream, they no longer interacted with A and B to form a pattern. By analogy, sounds in music would tend to be allocated to one stream or another and help to specify the melodic structure of the stream to which they were assigned.

While the examples that I have given so far have been concerned with simple patterns of pure tones, presented monophonically, we are now starting to explore the organizational effects of spectrally more complex sounds, such as speech, and of stereophonic presentations where different sounds are presented to the two ears.

Speech-like sounds seem to follow the same sort of rules of groupings as the simpler sounds. For example, the rapid alternation of the two speech sounds, "ee" and "ah," will be heard as two concurrent streams of sound, one containing all "ee" sounds and the other containing only repetitions of "ah." The segregation appears to be due to the spectral dissimilarity of "ee" and "ah." This result parallels my first example in which a rapid sequence of high and low tones broke into a high-tone stream and a low-tone stream.

A parallel to the experiment shown in Figure 7 can also be created with vocal sound. Instead of the two pure-tone components, B and C, we insert the vowel "ee" and, instead of the pure-tone captor A, we insert a single formant from "ee." When we listen to the repeated cycle, we can hear our stream containing a double-beating of the formant, accompanied by a separate stream containing the remaining formants, beating at half the speed. The vowel sound has been decomposed, just as the BC pair of Figure 7 was.

By presenting different but related sounds to the two ears, we have been able to show that grouping can affect where you hear things in space. This is a case where classical psychophysics has missed something very important by its insistence on simple stimuli. It turns out that the auditory system tries to build mental representations of acoustic input in which streams do not change either their pitches or their locations too rapidly. Interesting illusions can be obtained whenever the system is put into a conflict where in order to preserve the assumption of a small spatial change, it must introduce the assumption of a large pitch change, or vice-versa. Some listeners will even rearrange sounds in space in order not to introduce large pitch changes into a stream. An example is shown in Figure 9.

Howard Steiger in my laboratory has presented listeners with a cycle formed from two events, parts of which are fed to different ears. In the experiment shown in Figure 9, the sounds were all composed of harmonics of the same fundamental frequency. The blocks in which the upper half is colored refer to sounds containing harmonics three and four, blocks with the lower half filled in refer to sounds containing harmonics one and two only, and totally colored-in blocks represent sounds that contain all four harmonics. In the condition shown in Figure 9, the upper two harmonics were first presented to the left ear, then the same ear got the lower two harmonics while the opposite ear got all four. This stimulus was repeated over and over again. The lower diagram of Figure 9 shows how one listener heard it. She heard the first tone as expected, a high sound located on the left. The information in the second burst of tones was, however, mentally rearranged. She followed the assumptions of classical psychophysics in hearing what was common to the two ears, the lower harmonics, in the center. However, she took the extra high harmonics from the right ear and heard them on the left, so as to match the earlier event on the left. It should be emphasized that this

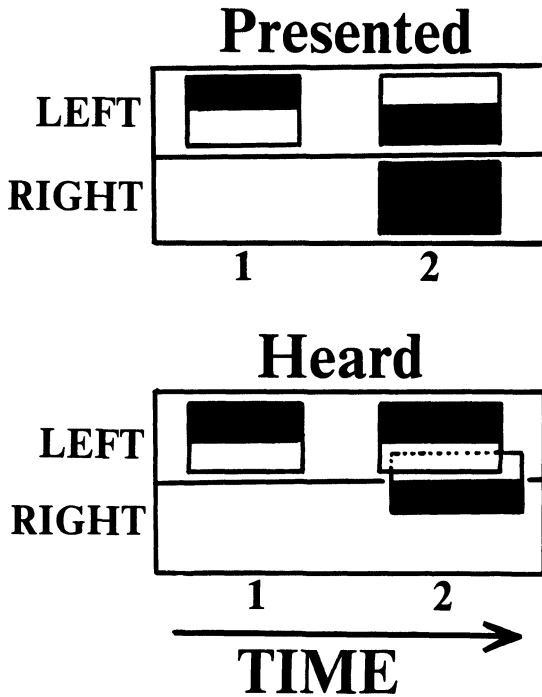


Figure 9

The induction of a spatial illusion; there were two time slots (in a cycle) in which a tone could be presented; the relation between frequency components presented in the two slots determined the perceived location of the second sound

particular pattern was one of many presented in a random order to her and that each time she heard this one she described it the same way. Despite the reliability with which this listener described the pattern, other listeners heard it differently, perhaps giving more weight to the actual side of arrival of the information or less weight to the requirement that sequences of sound from the same location should contain closely matching spectral content. This finding of individual differences in the weight that people give to different principles of grouping deserves much more study.

I would like to conclude with a brief review of what I have mentioned. It appears that there are at least three principles of

grouping. A “sequential” force tries to match similar spectral content across different moments of time. This principle is the one that tries to pull A and B into a pure-tone stream in Figure 7. This principle in music carves up the sequential mass of harmonics to produce melodic lines or acoustic layers. In addition, there is a “spectral” force that tries to bind co-occurring harmonics together to form timbres. This is the force that tries to group B and C together in Figure 7. It uses synchrony of onsets and offsets as well as common direction of frequency modulation of harmonics as bases for grouping. A third influence, the “binaural” force, tries to bind similar spectral content from the two ears into a spatially localized image. No one of these forces or principles is utterly dominant. They compete and cooperate to produce the “best-guess” perceptual description of the input.

I would like to argue that these principles are content-free. They capitalize on broad regularities in the world, such as the tendency of harmonics arising from a single source of sound to start and stop together, to change in frequency together, and to change gradually rather than abruptly in pitch and in spatial location. They provide the initial organization upon which particular forms of learning, such as musical or linguistic, can take place. There is no doubt in my mind that an experienced musician goes beyond these principles and uses musical rules to form a mental representation of the music, but it is likely that the preliminary content-free principles of grouping remain powerful determinants of the patterning of the music even for the sophisticated listener. To this extent, musical listening is like ordinary listening. However, rather than deflating the achievements of musical listening, this analogy makes us more appreciative of the complexity of the seemingly mundane act of just hearing what is happening around us as we pass through our daily environments.

REFERENCES

- BREGMAN, A.S.
 1977: "Perception and Behaviour as Compositions of Ideals," *Cognitive Psychology*, IX/2, 250-92.
 1978: "The Formation of Auditory Streams," in Réquin, J., ed., *Attention and Performance VII*. Hillsdale, New Jersey: Lawrence Erlbaum, 63-75.
- BREGMAN, A.S. and CAMPBELL, J.
 1971: "Primary Auditory Stream Segregation and the Perception of Order in Rapid Sequences of Tones," *Journal of Experimental Psychology*, LXXXIX/2, 244-49.
- BREGMAN, A.S. and PINKER, S.
 1978: "Auditory Streaming and the Building of Timbre," *Canadian Journal of Psychology*, XXXII/1, 19-31.
- BREGMAN, A.S. and RUDNICKY, A.I.
 1975: "Auditory Segregation: Stream or Streams?" *Journal of Experimental Psychology: Human Perception and Performance*, I/3, 263-67.
- CHOMSKY, N.
 1965: *Aspects of the Theory of Syntax*. Cambridge, Mass.: The MIT Press.
- CUDDY, L.L., COHEN, A.J., and MILLER, J.
 1979: "Melody Recognition: The Experimental Application of Musical Rules," *Canadian Journal of Psychology*, XXXIII/3, 148-57.
- DANNENBRING, G.L.
 1974: "Perceived Auditory Continuity with Gliding Frequency Changes." Ph.D. dissertation, McGill University.
- KOFFKA, K.
 1935: *Principles of Gestalt Psychology*. New York: Harcourt, Brace and World.
- VAN NOORDEN, L.P.A.S.
 1975: "Temporal Coherence in the Perception of Tone Sequences." Ph.D. dissertation, Eindhoven University of Technology (The Netherlands).
- WARREN, R.M., OBUSEK, C.J., and ACKROFF, J.M.
 1972: "Auditory Induction: Perceptual Synthesis of Absent Sounds," *Science*, CLXXVI/4039, 1149-51.
- WINSTON, P.H., ed.
 1975: *The Psychology of Computer Vision*. New York: McGraw-Hill Computer Science Series.