GENTSE BIJDRAGEN
TOT DE KUNSTGESCHIEDENIS EN OUDHEIDKUNDE

XXVIII (1989)

UITGEGEVEN DOOR
DE SECTIE KUNSTGESCHIEDENIS EN OUDHEIDKUNDE
VAN DE RIJKSUNIVERSITEIT TE GENT MET DE STEUN
VAN HET UNIVERSITEITSVERMOGEN
GENT 1989

Christine Buyle

AUDITORY ANALOGUES TO VISUAL REVERSIBLE FIGURES

Introduction: visual and auditory grouping

Grouping is a fundamental aspect of perception: we group our environment into wholes and structures.

When we look through a window, we see houses, sky, trees and so on. We do not perceive all this as several hundreds of juxtaposed or randomly distributed brightnesses and nuances of colour[1].

Grouping occurs in audition as well. It is very important to make a distinction between two kinds of auditory groupings:

1. We can group a complex sound environment into different *simultaneous* sound sequences, emanated from different sources. When walking through a forest, for example, we divide the complex sound mass into the rustling of the trees, the chirps of insects, the calls of the different birds and so on[2].

2. Mostly, we can also divide a single sound sequence into *successive* groupings. When we listen to a speech, for instance, we group the speaker's sounds into successive units, such as words, groups of words and sentences.

To be sure, however, the perception of sound sequences and their sources is not always in accordance with the acoustic phenomena. This gives rise to various illusory percepts, which can be classified into two groups:

1. Sometimes we divide our acoustic environment into sound sequences according to principles that are rather inappropriate to a particular physical situation. A well-known example of such an illusory percept is called the 'scale illusion'[3]. Simultaneous ascending and descending scales, presented to opposite ears (figure 1a), are not perceived on the basis of location (figure 1b), but segregate into upright and inverted V-shaped contours (figure 1c). This illusory grouping, formed on the basis of frequency proximity, results in a mislocalization of the tones.

[1] See WERTHEIMER, Max: Untersuchungen zur Lehre von der Gestalt. In: *Psychologische Forschung* 4 (1923): 301.

[2] SLOBODA, John A.: *The Musical Mind: The Cognitive Psychology of Music*. Oxford, Clarendon Press, 1985: 155.

[3] DEUTSCH, Diana: Two-channel listening to musical scales. In: *Journal of the Acoustical Society of America* 57 (1975): 1156-1160.

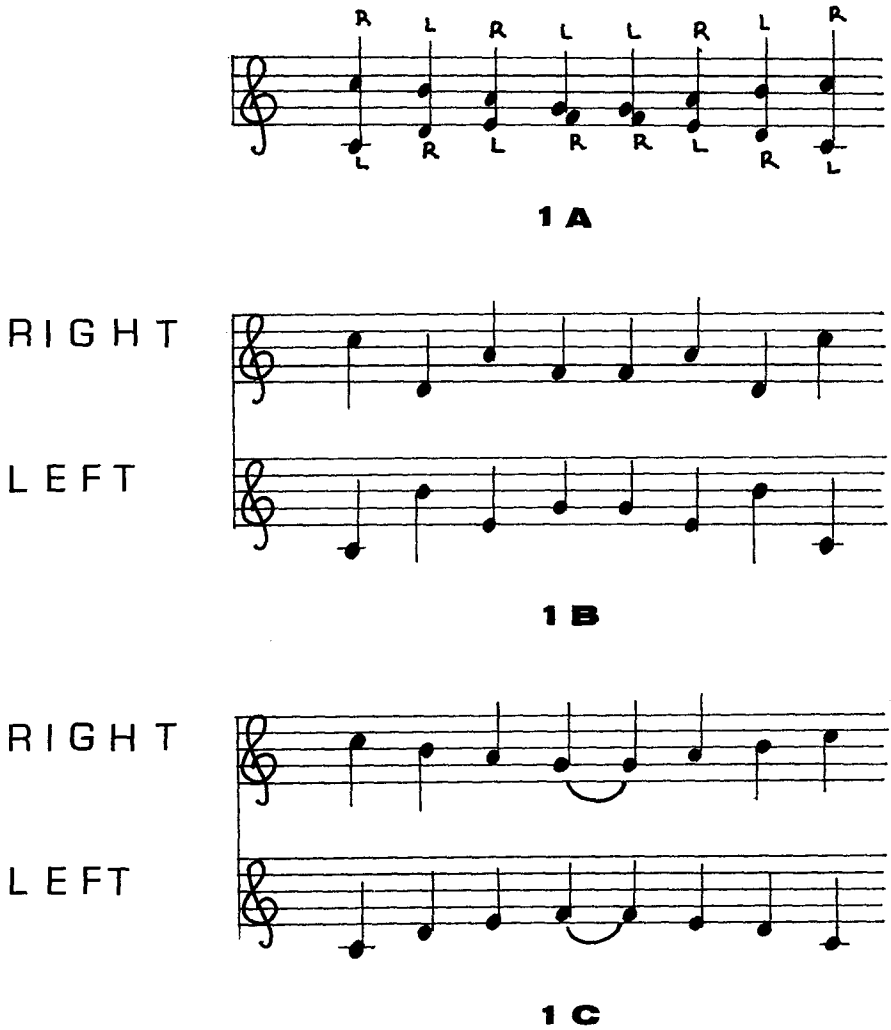
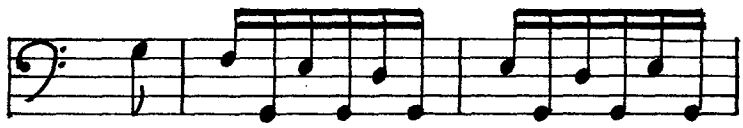


Figure 1. Scale illusion. Two scales, a descending and an ascending one, are dichotically presented (1A). Streams are not perceived on the basis of location (1B). In most cases, listeners perceive a V-shaped and an inverted V-shaped pattern (1C). (After D. DEUTSCH, 1975.)

2. Another kind of auditory illusion is our interpretation of a single, rapid succession of sounds, drawn from different frequency ranges, as two or more simultaneous sequences. This specious polyphony is commonly applied in music, especially during the baroque period (example 1)[4].

[4] Research on the psychological principles that govern this kind of perceptual organization is surveyed by MCADAMS, Stephen and Albert BREGMAN: Hearing Musical Streams. In: *Computer*



Example 1. J. S. BACH: Suite for Violoncello Solo in C major, BWV 1009 (fragment from the Gigue).

These illusory percepts demonstrate that the parsing of our acoustic environment into different sound sequences is largely a psychological phenomenon. To take this fact into account, BREGMAN and CAMPBELL[5] have introduced the concept of auditory 'stream'. A stream is a sequence, interpreted as a 'whole' because of a perceived internal consistency.

The precise mechanisms for grouping and the rules by which these mechanisms operate are not yet well understood. Nevertheless, it appears that the same set of general principles underlies both visual and the two kinds of auditory grouping. (We notably refer to the grouping principles, formulated by Gestalt psychologists[6]). As the same grouping principles govern both vision and audition, it is tempting to search for some more parallels between those two sensory modalities. We shall limit ourselves and ask if there are auditory analogues to a special situation where visual grouping is ambiguous: the reversible figure.

Music Journal 3 (1979) 4: 26-43 and by DEUTSCH, Diana: Grouping Mechanisms in Music. D. DEUTSCH (ed.): *The Psychology of Music.* New York, Academic Press, 1982): 118-127.

[5] BREGMAN, A. S. and J. CAMPBELL: Primary auditory stream segregation and the perception of order in rapid sequences of tones. In: *Journal of Experimental Psychology* 89 (1971): 244-249.

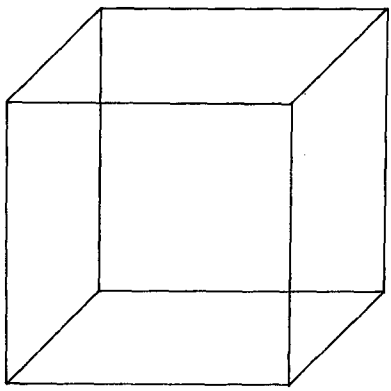
[6] The Gestalt principles are most commonly demonstrated within the field of visual perception. The Gestalt psychologists themselves, however, have already suggested that their principles can describe auditory perception as well. Some more recent studies on auditory grouping have affirmed this suggestion. It appears that Gestalt rules apply both to successive and simultaneous groupings.

As to the division of sound sequences (in this case monodies) into successive groups, some theories have been developed that draw upon Gestalt principles. (TENNEY, James and Larry POLANSKY: Temporal Gestalt Perception in Music. In: *Journal of Music Theory* 24 (1980) 2: 205-241. The grouping preference rules of LERDAHL and JACKENDOFF are also based on Gestalt principles (JACKENDOFF, Ray and Fred LERDAHL: Generative Music Theory and its Relation to Psychology. In: *Journal of Music Theory* 25 (1981) 1: 45-90. The rules are reprinted as a part of LERDAHL, Fred and Ray JACKENDOFF: *A Generative Theory of Tonal Music.* Cambridge, Mass., MIT Press, 1983).

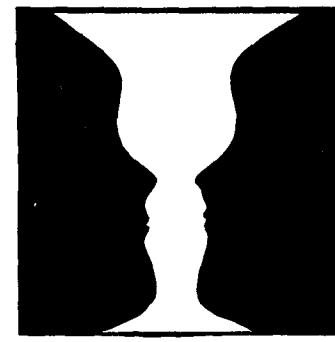
It has been demonstrated that Gestalt principles can also be used to describe our grouping into streams (see for instance JULESZ, Bela and Ira J. HIRSH: Visual and Auditory Perception — An Essay of Comparison. DAVID Edward E. jr. and Peter B. DENES (eds.): *Human Communication: A Unified View.* New York, McGraw-Hill Book Company, 1972: 299-309; BREGMAN, Albert S.: The formation of auditory streams. J. REQUIN (ed.): *Attention and Performance*, vol. 7. New Jersey, Lawrence Erlbaum, 1978; DEUTSCH, Diana: Grouping Mechanisms in Music. DEUTSCH, Diana (ed.): *The Psychology of Music.* New York, Academic Press, 1982: 99-134).

Visual reversible figures

A visual reversible figure is an ambiguous figure that allows different structurations. With prolonged viewing of such a figure, percepts alternate between different structures. Figure 2 shows some of the best known reversible figures.



2A



2B

2C

Figure 2. Reversible figures. 2A can be perceived as a cube, viewed from above, or as a cube, viewed from below (Necker cube). 2B can be interpreted as a white vase on a black background or as two black faces on a white ground (Rubin figure). 2C presents a young woman, as well as an old one (Boring figure).

Reversals occur when the ambiguous figure is continuously looked at during a certain amount of time. When the figure is presented discontinuously, reversals can also occur, provided that the interval of time between successive presentations is very small[7].

It has not yet become clear what causes figural reversal. The most prevailing theory is one of *satiation.* One of the founders of Gestalt psychology, Wolfgang Köhler, suggested that the neural process, involved with one possible grouping of an ambiguous figure, becomes satiated with prolonged viewing. The perceptual system then reorganizes the visual field according to a nonsatiated mechanism. Another theory emphasizes the role of *attention.* According to this theory, reversals are caused by conscious or unconscious shifts of the attentional focus. None of these theories is sufficient to explain reversal. A combination of the two would probably give some more satisfactory results[8].

By simple introspection we may observe that some reversals occur spontaneously and others at will. The reversal is *voluntary* when the alternative grouping has been consciously anticipated. This is only possible when we are familiar with the alternative grouping or when we are informed about its existence. Reversals occur *spontaneously* when an uninformed perceiver is confronted for the first time with a particular ambiguous figure or when an informed perceiver fixates the figure without consciously anticipating alternatives[9]. However, consciously anticipating an alternative grouping does not always have immediate results. Therefore, the role of will in reversing should not be overestimated.

Are auditory analogues possible?

Now we must pose our crucial question: Do auditory analogues to visual reversible figures exist?

As we have seen, visual reversals can only occur when an ambiguously structured figure remains within the observer's perceptual field for some time. Percepts can only alternate between different groupings when the ambiguous figure is directly accessible. We assume that this condition holds for auditory ambiguously structured patterns as well.

[7] Experiments with the Necker cube have demonstrated that the rate of reversals can be accelerated with discontinuous presentation. The interval of time between successive presentations should not exceed 250 msec. (For an overview of these experiments, see MURCH, Gerald M.: *Visual and Auditory Perception.* Indianapolis, Bobbs-Merrill Educational Publishing, 1973: 150-151.)

[8] MURCH, Gerald M., op. cit.: 150-154 and ROCK, I.: The Description and Analysis of Object and Event Perception. K. R. BOFF, L. KAUFMANN, J. P. THOMAS (eds.): *Handbook of Perception and Performance.* New York: John Wiley and Sons, 1986: 33.19-33.20.

[9] EHRENSTEIN, Walter: Untersuchungen über Figur-Grund Fragen. In: *Zeitschrift für Psychologie* 117 (1930): 362-363.

An auditory pattern that allows different structurations is not reversible yet: it must be wholly accessible to the perceptual system and stay so for a certain amount of time. Strictly logically spoken an auditory pattern starts, goes on and ends in time. Therefore there is no point in time where the pattern is wholly presented, let alone that it can persist for some time. For every single instant is durationless, just like every point of mathematical space is extensionless. So, at first sight auditory analogues to visual reversible figures seem rather impossible.

However, none of us ever experiences the present as durationless, as a point in time. Stimuli or events don't disappear instantaneously from our immediate consciousness and therefore overlap with new incoming data. This enables us to integrate a succession of events or stimuli and experience them as an enduring present, the duration of which we perceive immediately, without recourse to counting or measuring against a surrogate activity such as breathing or the tensing of muscles[10]. So we are able to integrate e.g. all the positional changes of a falling star and interpret them as a continuous movement or we can grasp a succession of notes and experience it as a musical motive. This perception of succession as a sort of simultaneity is called a *psychological present.*

Only temporal successions that last but a few seconds can be perceived as a psychological present. There is, however, no general agreement on the maximum length of the psychological present. Moreover, the upper time limit for temporal integration seems to be dependent on the sort of stimulus: for a single sound this maximum lies between two and five seconds but for hierarchically ordered pulses the measurable duration of perceived present is much larger. It is possible, however, that the integration of a larger succession is somewhat looser[11].

It is important to notice, however, that the information, contained in a psychological present, remains only directly accessible for further (re-)interpretation, comparison, etc., as long as no structural boundary (a pause or a syntactic boundary, for instance) has been reached. After that boundary, direct accession to the sequence is altogether impossible[12].

If we now return to our initial question, it appears that auditory 'reversible figures' are possible, provided that some conditions are fulfilled. In the first place, we need an ambiguously structured sound pattern. Furthermore, that pattern has to be wholly accessible, which is only possible when it is short enough to be grasped by the immediate memory. As the accessibility ceases

[10] HASTY, Christopher F.: Rhythm in Post-Tonal Music: Preliminary Questions of Duration and Motion. In: *Journal of Music Theory* 25 (1981) 2: 186.
[11] HASTY, Christopher F., op. cit.: 186-187.
[12] MICHON, John A.: The Making of the Present: A Tutorial Review. REQUIN, Jean (ed.): *Attention and Performance*, vol. 7. Hillsdale, New Jersey, Lawrence Erlbaum Associates, 1978: 96.

when a structural boundary 'closes' the psychological present, the only way to fulfill our third condition — the claim that the pattern is constantly accessible to the observer over a period of time — is to repeat the sound sequence.

Auditory 'reversible figures'

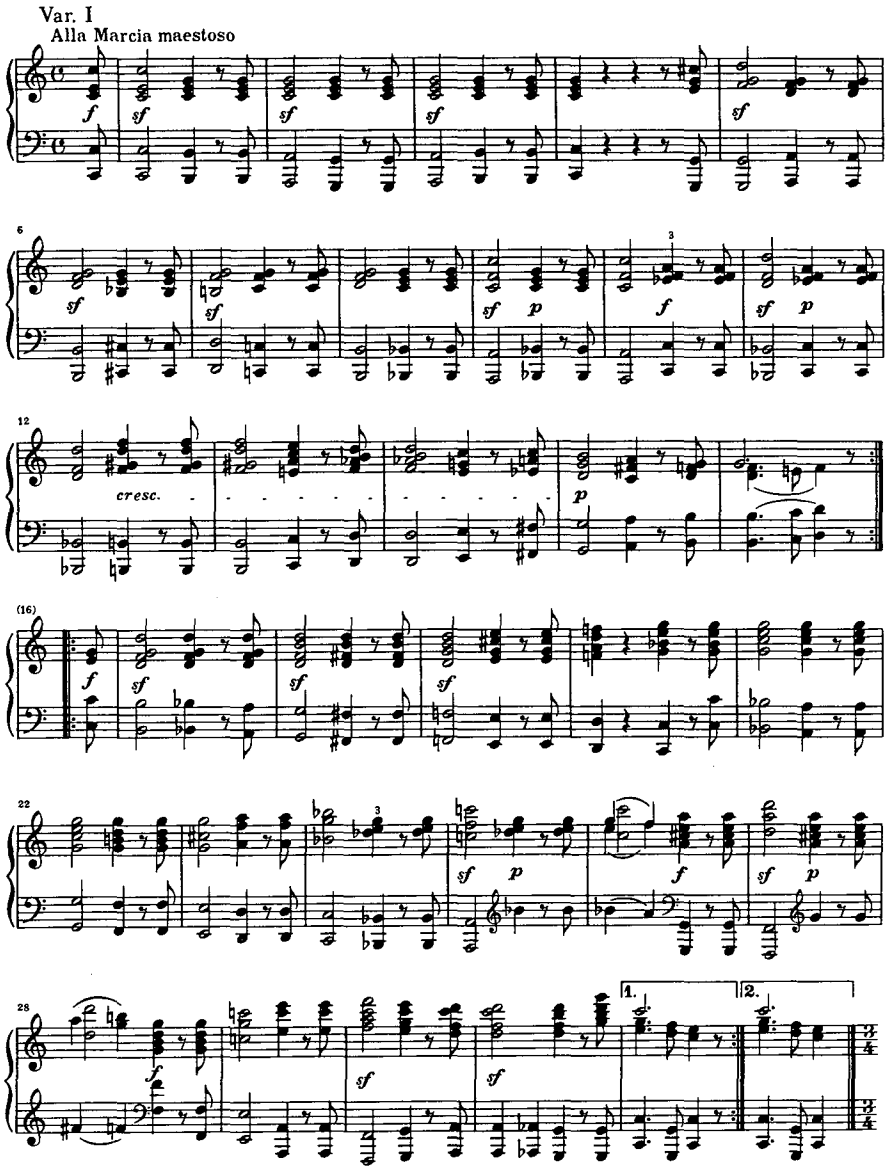
Our aim is to find auditory 'reversible figures': situations in which auditory grouping is ambiguous and where percepts alternate between various grouping possibilities. We have seen before that these situations require the repetition of a short sound sequence, at every turn experienced as a psychological present. Since there are two kinds of auditory grouping, we have to examine both of them.

Successive groupings

WARREN and GREGORY have observed that the repetition of words and short phrases can give rise to auditory reversals. Repetition of the word 'rest', for instance, may cause percepts like 'stress', 'tress' and 'Esther'[13].

In rhythmic sequences, groups are formed on the basis of temporal proximity: boundaries of groups are perceived between elements that are farther apart. Grouping can be ambiguous when two or more consecutive larger gaps alternate with a series of smaller ones. In that case, the grouping boundary could be located in each of the larger gaps. Repetitive presentation of such an ambiguous rhythmic sequence would give rise to various groupings, caused by the displacement of the grouping boundary. For example, when we repeat the following rhythmic formula: ♪ | ♪ ♩ 𝄾, our percepts soon alternate between two different groups: ♪ | ♪ ♩ 𝄾 and ♩ 𝄾 ♪ | ♪. It seems that this alternation is sometimes obtained *spontaneously*, sometimes *at will*. In music, however, reversal can also be *imposed* by melodic, harmonic, dynamic, and other contexts of the rhythmic ostinato. A superb example of imposed reversals is provided by the first variation from BEETHOVEN's '33 Veränderungen über einen Walzer von A. Diabelli', op. 120 (example 2). In mm. 1-8 grouping is ambiguous. It is probable, however, that a listener, uninformed about alternative groupings, shall maintain as long as possible his first grouping (♪ | ♪ ♩ 𝄾). In mm. 8-12 harmonic and dynamic context favour the other grouping (♩ 𝄾 ♪ | ♪). Moreover, there is a strong tendency to group mm. 5-8 the same way as mm. 1-4, drawing a boundary in m. 4 and m. 8 after the bar's first note.

[13] WARREN, R.M. an R.L. GREGORY: An auditory analogue to visual reversible figures. In: *American Journal of Psychology* 7 (1958): 612-613.



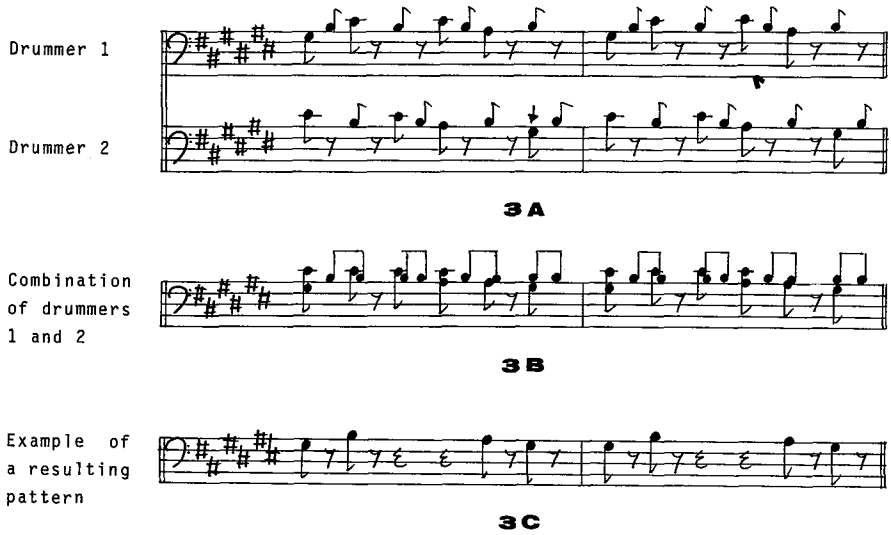
Example 2. L. VAN BEETHOVEN: 33 Veränderungen über einen Walzer von A. Diabelli, Op. 120 (variation 1). (Printed with kind permission of G. Henle Verlag.)

Streams

Once simultaneous sound sequences are detected in a complex sound environment, one can attend to only one of them at a time. It is possible,

however, to shift the attentional focus to another stream. Let us now consider a complex sound environment where each stream is a repetition of a small motive and where the motives of the different streams coincide with one another. Shifting the attentional focus between streams then means: switching between the various motives. This switching is analogous to the alternation between various structures in visual reversible figures.

Excellent illustrations of auditory reversal in complex and repetitive sound environments can be found in music of Steve REICH. A complex sound environment is created by simultaneous, but asynchronous presentation of a repetitive pattern. Amidst this texture, the listener perceives repeated musical motives, which REICH calls 'resulting patterns'. Alternations between these motives occur *at will* or *spontaneously* and can even be *imposed* by the composer. To impose a particular resulting pattern, the pattern has to be accentuated by singing, whistling and so on (example 3).



Example 3. Steve REICH: Drumming (fragment).

As we have seen in the introduction, the perception of simultaneous sound sequences can be an illusion: sometimes we split a single, rapid succession of sounds into streams. The kind of reversals we have described above also holds for this illusory polyphony. However, within a certain range of tempos and frequency relationships, a rapid succession of sounds may either be perceived as a single stream or as several simultaneous streams. This can give rise to another kind of reversal: rapid sound successions, consisting of a

small, repeated pattern, may be perceived as the repetition of one coherent motive, alternated with the repetition of several simultaneous ones[14].

Conclusion

In this article we asked ourselves if there are auditory analogues to visual reversible figures. As visual reversible figures result from situations where grouping is ambiguous, we had to examine ambiguous situations in the two kinds of auditory grouping: grouping into simultaneous streams and dividing streams into successive groups. But before we could do this, we had to find out how the basic condition for reversal (the ambiguous stimulus must be wholly accessible for some time) could be fulfilled in auditory perception.

In the last part of this article, we were able to demonstrate that 'reversible figures' exist in both kinds of auditory grouping.

[14] See, for instance, VAN NOORDEN, L.P.A.S.: *Temporal Coherence in the Perception of Tone Sequences*. Unpublished doctoral dissertation. Eindhoven: Technische Hogeschool, 1975.