GESTALT THEORY, DOI 10.2478/gth-2026-0004
© 2025 (ISSN 2519-5808); Vol. 47, No. 1, 35–60

Original Contributions - Originalbeiträge

Walter Coppola

The musical language: Structural analysis in Gestalt perspective

Introduction

Music, present in all cultures and in every historical era, is an essential element in the life of human beings. It is a complex phenomenon, involving emotional, cognitive, social and evolutionary aspects, and in every civilization has always represented a powerful means of communication.

This is a vast, multidisciplinary subject deeply rooted in our biological, emotional and cultural experience. Music allows emotions and inner states to be expressed in non-verbal form; it has a cohesive function in collective rituals, celebrations, moments of mourning or joy, using a shared language that aims at strengthening the community and social aggregation of people. According to Steven Mithen (b. 1960), (Mithen 2005), an archaeologist known for his work on the evolution of language, music and intelligence, the origins of music are closely linked to the evolution of language and everything that contributes to the formation and consolidation of the social and collective organizational structure.

Listening to and processing music involves complex cognitive capacities such as memory, attention, foresight, categorization, and stimulates abstract thinking processes and mental representation. Neuroscience has also shown that music activates brain areas linked to emotions (amygdala, limbic system, prefrontal cortex) and can therefore arouse pleasure, emotion, or excitement. Studies (Peretz & Zatorre, 2005; Levitin, 2006) have shown that music can be used in the therapeutic sphere to improve mood, reduce anxiety or stimulate cognitive functions in patients with neurological diseases, thus underlining the importance and value of music-therapeutic practice.

In addition to its expressive and practical values, music has an autonomous aesthetic function. In this sense, it stimulates a type of enjoyment that could be described as disinterested and contemplative, fostering reflective thinking and formal sensitivity. The latter is understood as the ability to perceive, appreciate and understand the internal structure of a work – in music as in the visual arts and literature – beyond its narrative or emotional content. In music, it means grasping the coherence of the parts (e.g. how a theme develops), symmetries,

variations, contrasts, and everything to do with the interplay of relationships between rhythm, harmony, melody and form. Formal sensitivity thus allows us to analyze and appreciate not only *what* the music expresses, but *how* it is constructed and what are the dynamics on which its structural coherence and inherent communicative power rest.

Our perceptual system, based on mechanisms of great complexity and extraordinary functionality, allows us to grasp the infinite melodic and harmonic combinations, the changes in the dynamic and agogic patterns, the infinitesimal subtleties of timbre that musical language offers us.

The process of acquiring and fruition of the musical message lives through three distinct phases: the perceptive, the elaborative and the emotional. The perceptual phase is based on the acquisition mechanisms of the sound waves that are picked up and sent to the brain; this is where the elaborative phase occurs and the sound message is given a clear, precise and identifiable structural dimension; the third phase concerns that process that is in some ways still mysterious thanks to which the sound language develops all its expressive power and arouses in us a whole range of emotions.

Auditory perception

The fundamental medium, the primary element through which the art of music is expressed, is sound. Sound is a physical phenomenon produced by the vibration of an oscillating body such as, in musical instruments, the air set in motion by the breath of someone playing a flute or the vibrations of the strings of a guitar. The vibration of a body produces a change in air pressure, generating waves that propagate to our auditory apparatus and, transformed into impulses, are received by the brain as a sound sensation.

But sounds are not just mental reproductions of acoustic stimuli. Instead, they are complex elaborations involving psychophysical processes of recognition, analysis and processing. An external acoustic stimulus corresponds to an internal sound object, just as visual perception corresponds to the image of an object. Whereas, however, the systems underlying visual or tactile perception convey information about reality existing outside us, sound, like taste or smell, refers to a characteristic, a property that we attribute to objects. It is therefore not by chance that we define sound (sweet, enveloping, sharp) with the same adjectives with which we define tastes and smells.

Sound has three different properties:

– **Frequency**, the number of oscillations a sound wave makes in a certain unit of time and is measured in *hertz*. Frequency means sound pitch, so sound waves with a low number of oscillations result in low-pitched sounds, while those with a high number result in high-pitched sounds.

– **Intensity**, depends on the amplitude of the sound wave. Intensity is measured in *decibels* according to a scale ranging from 0db (barely perceptible to the human ear) to 120db (pain threshold). Intensity is used to measure the acoustic power of the sound message, i.e. its low, medium or high volume.
– **Timbre**, changes depending on the source it comes from and is what differentiates all types of sound. Timbre is the result of the harmonic sounds that accompany the fundamental sound. This is why, by listening to a note, we can tell whether it is produced by a trumpet, a piano, a guitar or a human voice.

The function of the ear is to transform sound waves into information for the brain (Fig. 1). The human ear acts as a true transducer in transforming acoustic energy, first into mechanical energy and then into electrical energy. Once the energy has been converted from mechanical to electrical form, the electrical impulses reach the brain through nerve endings. Here they are processed, enabling the perception of sound and thus the listening to music.

The perception of sounds in the environment involves all three components of the ear. Sound waves penetrate the outer ear, pass through the middle ear and finally conclude their journey at the inner ear. Thanks to their particular anatomy, the structures that form the outer ear have the task of conveying sound waves towards the middle ear: the auricle receives the sound waves and causes them to enter the external auditory canal, up to the eardrum. As the sounds reach the eardrum, it begins to vibrate.

The vibration of the eardrum marks the beginning of the middle ear's participation in the process of sound perception. By vibrating, in fact, the eardrum triggers

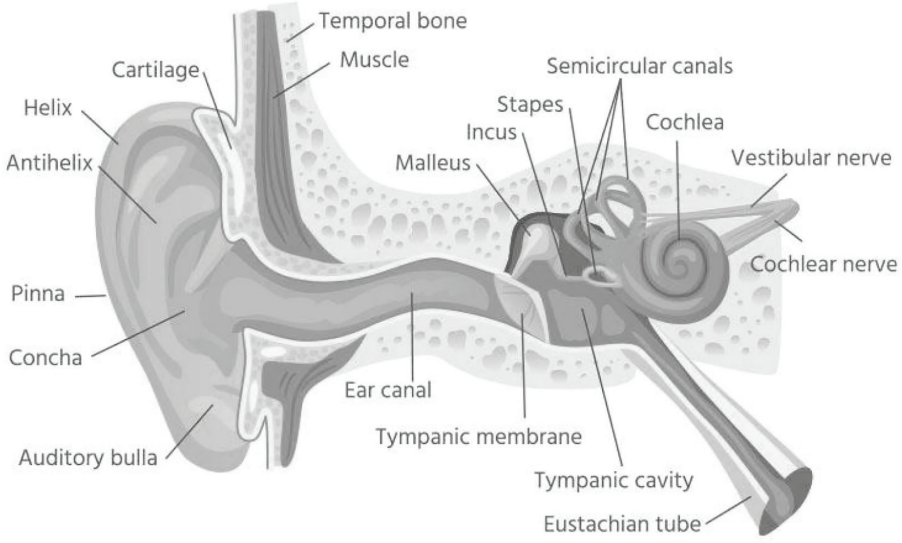

Fig. 1. Diagram of the human ear.

the chain of the three ossicles: the first ossicle to activate is the malleus, the second is the incus and the last is the stapes. From the stirrup, the vibrations pass to the oval window and the round window, which function similarly to the tympanic membrane. The vibrations of the oval window and the round window set the endolymph in the cochlea in motion. The movements of the cochlear endolymph are the signal that triggers the cells of the organ of Corti. Once activated, the cells of the organ of Corti take care of the important process of converting sound waves into nerve impulses.

Once the conversion is complete, the cochlear nerve comes into play, which collects the nerve impulses and sends them to the temporal lobe of the brain where the reprocessing of the nerve impulses and the generation of an appropriate response takes place.

The musical brain

In recent years, neuroscience has documented in increasing detail how music and language share brain networks, especially for syntax, prosody, rhythm and auditory analysis.

In a very schematic way, we indicate the main brain areas involved in spoken language and musical language, with an indication of their primary functions (Fig. 2).

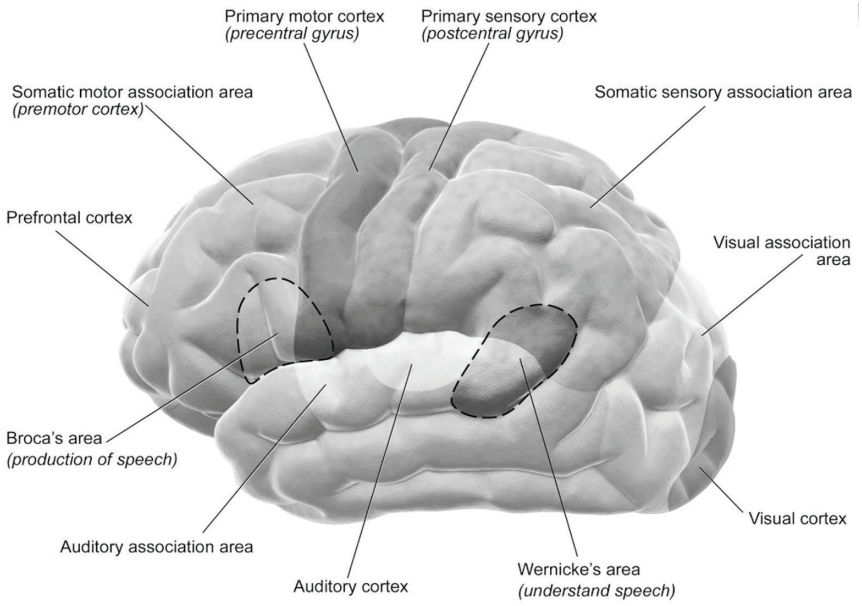

Fig. 2. Anatomical diagram of the human brain.

The areas most involved in spoken language are:

- Broca's area (left frontal lobe) that presides over language production and syntactic analysis. It is also involved in complex musical structure (e.g. harmonies, sequences).
- Wernicke's area (left posterior temporal lobe), essential for language comprehension. It is less involved in music, but is activated for sung texts.
- the arcuate fasciculus, which connects Broca' and Wernicke' areas, plays a crucial role in the ability to elaborate, produce and repeat spoken and written language and in musical processing.

The areas most involved in music are:

- the right hemisphere (especially the temporal and parietal lobe), the site responsible for emotional sensitivity to music and for the processing of timbre, melody and harmony.

- the cerebellum, for motor coordination (e.g. playing an instrument) and for the regulation of time and rhythm.

- the basal ganglia, involved in the perception of rhythm.

- the limbic system (amygdala, hippocampus, insula, orbitofrontal cortex), for the regulation of musical emotions and for affective processing.

The areas common to speech and music are:

- the primary auditory cortex (superior temporal lobe), which analyses basic acoustic characteristics (pitch, timbre, intensity) and is involved in both speech and music perception.

- the superior temporal gyrus, crucial for phonemic perception and melodic and harmonic processing.

Neuroscience studies have also highlighted how brain areas and their respective functions inherent to spoken language and musical language sometimes intersect and overlap, giving rise to complex dynamics in a neurological framework of extraordinary configuration.

Stefan Koelsch (b. 1968), (Koelsch et al., 2000), using ERPs[1], demonstrates in a study that even individuals without musical training show brain responses to musical syntax, suggesting an implicit and universal mechanism for musical processing

[1] ERPs are measurable brain responses that are formed following an internal or external stimulus or event and are generated by the synchronous electrical activity of groups of neurons in response to a specific stimulus. They are recorded by electroencephalography (EEG), a technique that measures the electrical activity of the brain through electrodes placed on the surface of the scalp. Analysis of ERPs makes it possible to identify the different stages of information processing in the brain, such as perception, attention, memory and language.

similar to language. Furthermore, through neuroimaging studies (Koelsch et al., 2005, 2006), it has been shown that music activates Broca's area and the inferior frontal cortex during harmonic violations, similar to when we listen to linguistic sentences with grammatical errors. A study (Sammler et al. 2009) using intracranial electroencephalography (EEG) recordings in patients with epilepsy found that the processing of music and language syntax shows partial overlap in the perisylvian cortex of the brain, particularly within the superior temporal gyrus and, to a lesser extent, in the left inferior frontal gyrus. This suggests that these areas act as shared anatomical substrates for the early detection of syntactic errors in both domains.

Robert Zatorre (b. 1955), (Zatorre et al., 2002) identifies the auditory cortex, located in the temporal lobe, as playing a key role in the processing of speech and music. It analyses the frequencies, patterns and temporal sequences of sound, enabling us to perceive and recognise different sounds. Although both hemispheres contribute, the right auditory cortex is often associated with processing the pitch and tonal aspects of music, while the left hemisphere is more involved in processing the temporal aspects of speech, such as the duration of sound. Aniruddh D. Patel (b. 1961), (Patel, 2003) proposes the "Shared Syntactic Integration Resource Hypothesis" (SSIRH), according to which music and language share neural resources for syntax, while maintaining distinct domains. Broca's area is crucial in both processes. In other words, the brain uses some of the same neural mechanisms when processing song and sentence structure.

Structural elements of musical language

The structural elements of musical language are the fundamental constituents that enable music to communicate, organize itself, create meaning and produce an effect on the listener. Unlike verbal language, music does not convey concepts directly, but constructs forms of meaning through the organization of sounds in time. In the following, we present the main structural elements from an interdisciplinary perspective that combines music theory, aesthetics and the psychology of perception.

– **Pitch and tonal field**

Pitch distinguishes high-pitched from low-pitched sounds. The organization of pitches generates structures such as scales, tonalities and intervals, which constitute the "tonal field," interpreted as a dynamic space, endowed with orientation and internal tension. In this phenomenological view, music is an experience of becoming: each sound tends toward another, generating meaning.

– Duration, rhythm, and musical time

Duration is the persistence time of a sound, while rhythm is the temporal organization of sounds and silences. Rhythm establishes accents, periodicity, irregularity; it can be metric (regular) or free (as in recitative or oriental music). It is essential to the enjoyment of music, creating expectations and structure, with a direct link between musical emotion and the confirmation or frustration of temporal expectations.

– Timbre and Intensity

Timbre is the sound color that makes it possible to distinguish a note played by a violin from one played by an oboe, even with the same pitch and duration. It is related to harmonics, waveform, and resonance. It has an expressive and identity role: each instrument "speaks" with its own "accent." Intensity is the volume of a sound: loud, soft, *crescendo, diminuendo*. Dynamic variation helps shape phrasing and articulation, generate tension and release, and create emotional contrast.

– Musical Form

Form is the overall structure of the musical work; it is the organization of the parts into a coherent whole. It can be formal (ABA, sonata, fugue, rondo...) or perceived (through repetition, contrast, development).

– Tension and expectation

This is a key element in music psychology, as music generates expectations (harmonic, rhythmic, formal) and their fulfillment or frustration creates meaning and emotional involvement. Tension thus arises from relationships between musical elements that create a sense of directionality, and music is configured as a language that does not denote concepts but articulates dynamic forms of experience.

These structural elements are not mere technical components, but bearers of meaning. In any case, music speaks a universal language of relationships, directions, and tensions, which can be studied, perceived, and experienced far beyond words. To understand these elements is to access a universal language that speaks directly to perception and sensibility.

Processing

The elaborative factor in music is the set of processes by which we interpret and organize the musical phenomenon through its basic elements. In other words, we do not merely receive and perceive sounds, but we recognize them, relate them to each other and rework them through precise mental and cultural patterns. The basic musical structures (pitch, rhythm, harmony) are the raw material; the processing factor is the "engine" that transforms them into musical experience.

One of the most important currents of thought of the twentieth century is undoubtedly that of *Gestaltpsychologie*, and it is mainly on this that we will focus our reflections. Musical language processing according to Gestalt is interpreted in accordance of the principles that govern the way human beings organize and interpret complex stimuli. Applied to music, this theory has had a fundamental influence in the study of musical perception and in understanding the internal structure of musical language.

In 1923, an important article was published in *Psychologysche Forschung* signed by Max Wertheimer (1880–1943) (Fig. 3), (considered together with Wolfgang Köhler (1887–1967) and Kurt Koffka (1886–1941) to be the founder of Berliner *Gestaltpsychologie*), entitled 'Untersuchungen zur Lehre von der Gestalt II', in which he outlines the so-called principles of formal organization, i.e. the mechanisms with which the individual phenomenologically processes what he perceives. In this paper he sets out the guidelines of a new conception in the field of perception, with an exhaustive enumeration of the factors underlying this approach. These can be considered as real laws, applied by Wertheimer to visual perception, but which we can transpose to auditory perception.

Taking our cue from Wertheimer's writing, we would like, in a schematic way, to summarize the principles that are universally recognized as underlying the Gestaltist theory of perceptual experience, highlighting in particular the correlations between visual and auditory perception.

Fig. 3. Max Wertheimer plays violin.

Proximity

All things being equal, neighboring rather than distant elements tend to be experienced as constituting a perceptual unity (Fig. 4). In music, this principle applies to notes or chords that occur close to each other in time, creating groups of sounds that are perceived as cohesive. In a melody, notes close to each other are perceived as a continuous line. If there are two instruments with notes far apart in a tune, they are perceived as separate.

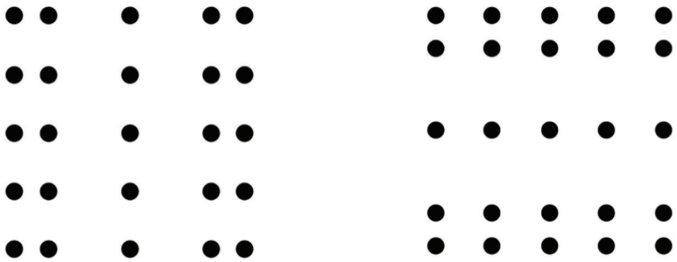

Fig. 4. The role of proximity in determining groups.

Similarity

All things being equal, elements that possess some kind of similarity tend to unify with each other (Fig. 5). In music, this may refer to similar groups of notes or chords being recognized as belonging to the same section of the composition. Instruments with similar timbres (e.g. a group of strings) are perceived as a unit.

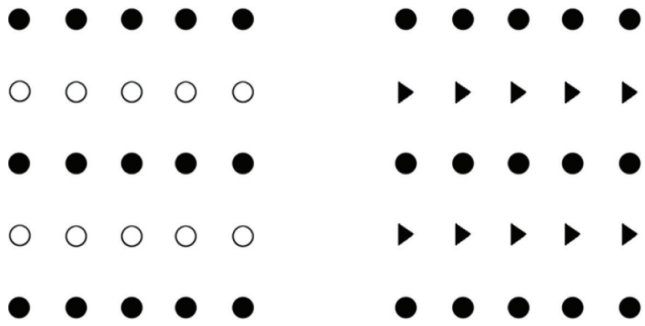

Fig. 5. The role of similarity in determining groups.

Common fate

Parts of the visual field that move together, or in a similar way, or that otherwise move unlike other parts of the field, tend to be constituted as segregated units (Fig. 6). Applied to music, it refers to how musical elements that move together or

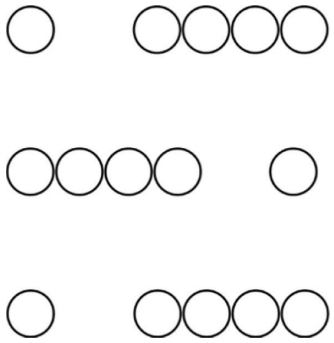

Fig. 6. The effect of common fate on perception.

follow a similar rhythmic or melodic pattern tend to be perceived as a unit. This principle suggests that listeners perceive elements that share similar characteristics (such as rhythm or melodic pattern) as part of a cohesive whole. In an orchestra, if several instruments play in parallel, we perceive them as a single sound line. If a melody and bass move in opposite directions, we perceptually separate them.

Continuity

Elements with a common direction tend to be perceived as belonging to the same object (Fig. 7). In music, this principle is applicable to melodies and harmonic progressions. Sound sequences with a regular progression (melodic or rhythmic) are perceived as a coherent unit. For example, a melodic passage that unfolds smoothly and continuously is more easily perceived as a single musical unit than as a discontinuous sequence of notes. If a melody follows a regular scalar progression, our brain completes it even if some notes are omitted.

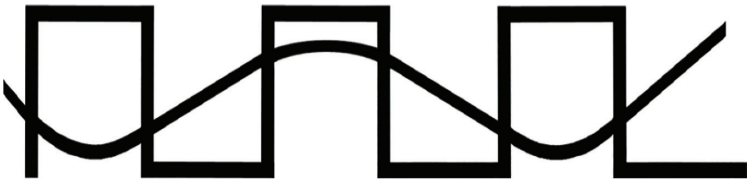

Fig. 7. A demonstration of continuity in perception.

Closure

Other things being equal, a closed rather than an open area is experienced as a perceptual unity (Fig. 8). In music, this can manifest itself in the expectation of a harmonic conclusion or resolution. For example, harmonic progressions tend to resolve into final chords that satisfactorily "close" the musical sequence. When a musical sequence is incomplete, our brain tends to complete it mentally. If an expected final note is missing in a melody, the brain "imagines" it to complete the sequence.

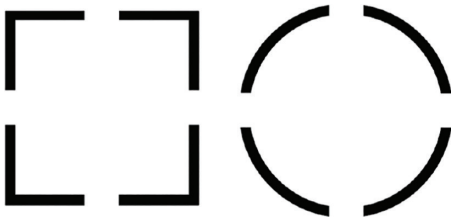

Fig. 8. The effect of closure on geometric forms.

Good Form

The perceptual field is segmented in such a way that it results in perceptual units and objects that are as balanced as possible, harmonious, constructed according to the same principle in all their parts, which thus "belong" to each other. The law of good form suggests that we prefer configurations that are simple and organized rather than complex and disordered (Fig. 9).

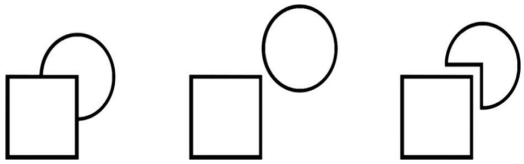

Fig. 9. A demonstration of good form on perception.

In music, this translates into a preference for melodies and harmonic structures that are perceived as coherent and well-organized and refers to the tendency of listeners to perceive and interpret musical information in such a way that clear and meaningful patterns are formed. In other words, musical elements tend to organize themselves into recognizable structures, facilitating interpretation and memorisation.

This principle applies to various aspects of music, such as melody, harmony and rhythm. For example, melodies tend to follow simple, coherent lines, while harmonies are organized around progressions that create expectation. Good form is also important in music, where sections of a composition develop in ways that seem logical and natural.

To these Wertheimer laws we can add two other very important factors, which gestaltists developed later and which are considered significant elements in the basic structure of perceptual organisation.

Past Experience

All other things being equal, the segmentation of the field also occurs according to our past experiences, favoring objects with which we are more familiar rather

than unknown or unfamiliar forms. In certain cases, it is admitted that past experience may influence perception, but only secondarily to innate organizational principles (Fig. 10).

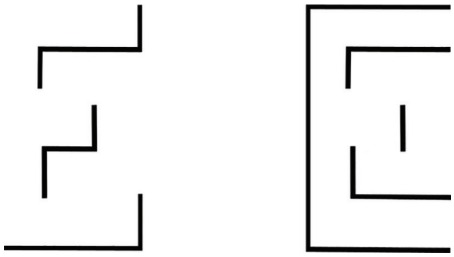

Fig. 10. A demonstration of past experience on perception.

When we apply this concept to music and past experiences, we can explore how emotions and memories influence our musical perception. For example, a piece of music can evoke specific memories, creating a deep connection between music and past experiences. This interaction between music and memory can give rise to an emotional "gestalt", where melody, rhythm and words combine to generate an overall emotional response.

Figure – Ground

The figure-ground factor concerns the way we perceive and organize visual and sound information by distinguishing a main object (figure) from its surroundings (ground). It is a basic perceptual phenomenon, not a specific law of organization, and is a process by which perception occurs, even before gestalt laws are applied. It has been deepened by later gestaltists (especially in visual form psychology), although it was first described by Edgar Rubin (Fig. 11).

Fig. 11. The Rubin figure–ground Illusion.

In music, it refers to the perception of musical elements as prominent figures and less prominent grounds. The listener therefore tends to isolate main melodies or rhythms (figure) from the surrounding harmonic or rhythmic context (ground). For example, in a piece of music, the main voice or a solo instrument may be perceived as the figure, while accompaniments and harmonies form the ground.

In the final page of the cited Wertheimer article, from the words "Es gibt auch im Akustischen..." we can grasp and identify one of the substantial aims of this paper: to extend Gestalt theory from the visual to the acoustic domain, proposing a general theory of perceptual organization. Sound, like the image, is perceived as form and individual elements are not analyzed in succession, but the dynamic unity of the whole is grasped.

In 1926, Carl Stumpf (1848–1936) published a volume that represents one of the fundamental points of connection between linguistic perception and musical perception. This work, *Die Sprachlaute* (The Sounds of Language), occupies a crucial role in the German scholar's thinking. For Stumpf, *Die Sprachlaute* constitutes the phenomenological basis through which it is possible to understand language and music as expressions of the same original acoustic substrate. From the point of view of perceptual structure, Stumpf observes that linguistic sounds, like musical sounds, possess tonal qualities (pitch, intensity, timbre), temporal articulation (rhythm, duration, accent), hierarchical organisation (relationships of similarity and contrast), and expressive value (emotional or "physiognomic" colour of sound). Linguistic and musical sounds therefore share fundamental psychoacoustic elements, and speech intonation (*Sprachmelodie*) is, in some ways, a "rudimentary" form of melody in that it anticipates its expressive and communicative character.

Stumpf adopts a phenomenological-analytical approach: he combines phenomenological observation of sounds with acoustic measurements and perceptual analysis, paying particular attention to the timbral, tonal and articulatory qualities of the voice. The aim is to understand how speech sounds are distinguished and organised perceptually, and how these processes are similar to those of musical perception, the subject of his previous studies (*Tonpsychologie*, 1883–1890).

From the point of view of the structure of musical language, *Die Sprachlaute* shows that the perception of linguistic sound is governed by structural laws similar to those that regulate musical perception (unity, continuity, closure, contrast); language and music share a common phenomenological basis, from which prosody, melody and rhythm derive; music can be understood as a symbolic expansion of the original perceptual forms of spoken language.

From a linguistic and phonetic point of view, *Die Sprachlaute* represents a pioneering contribution to the study of speech sounds as acoustic and perceptual phenomena. Stumpf analyses the sounds of language not only in terms of their articulatory production, but above all in terms of their perceptual quality, anticipating the approach of experimental phonetics and psychoacoustics of language. Although predating structuralist phonology, the work stands out for its rigorous description of the qualitative differences between linguistic sounds and for the idea, original for its time, that voice and speech share with music a phenomenal continuum based on timbre, rhythm and intonation.

Through his investigations, Stumpf anticipated many themes of modern cognitive science and the neuroscience of music and language, in particular the idea that sound perception involves complex interactions between sensory input, motor simulation and affective resonance.

Musical language and its structure

The structure of musical language is a concept at the intersection of linguistics, psychology of perception, music theory and philosophy of language. When we speak of "deep structure," we mean an invisible but organizing dimension that governs how sounds are articulated into coherent and meaningful forms, just as in language there is an implicit grammar that governs sentences. Music is often defined as a nonverbal language, capable of communicating emotion and meaning without the use of words. But what makes music structured and understandable?

This question was answered, implicitly, by Max Wertheimer. Although he did not directly theorize a grammar of musical language, his studies on the perception and organization of form were fundamental to many later scholars who analyzed the internal structure of musical language, laying the groundwork for a new way of understanding music: not as symbolic language, but as perceived form organized over time. The structure of musical language, in this perspective, does not reside in abstract rules, but in the mind's ability to grasp relationships, shapes and coherences between sounds and is what allows music to be understandable, exciting and meaningful. It is not just made up of theoretical rules or written notation, but consists of cognitive and perceptual processes that guide our listening experience. It represents the hidden heart of music, the basis on which form, expressiveness and meaning are built. To understand it is to access a fundamental dimension of music: that which unites rationality and emotion, structure and freedom.

Below is a review of the most significant contributions by leading scholars who have studied musical language from a Gestalt perspective and who, each in their own field and with their own scientific and epistemological approach, have developed an interesting and fruitful debate on the theme.

Rudolf Arnheim

Rudolf Arnheim (1904–2007) elaborates and develops the foundational concepts of Gestalt psychology, proposed by Max Wertheimer, by systematically applying them to aesthetic experience and, in particular, to musical language. Arnheim transforms Wertheimer's perceptual intuition into a full-fledged theory of artistic form, in which music assumes the role of a perceived formal system over time, based on structural relationships.

Wertheimer had shown how human perception organizes stimuli according to innate laws (proximity, continuity, closure, etc.). Arnheim extends these laws from visual perception to music, arguing that even sounds are grouped and interpreted according to structural principles, and not mechanically or randomly. One of the key concepts of Gestalt is that *the whole is different from the sum of the parts*. Arnheim makes this principle the cornerstone of his music theory: in music, individual notes or chords do not have meaning in themselves, but only as functional elements of a larger structure. For example, a melody is more than a sequence of pitches: it is a continuous perceptual configuration. A musical theme is recognizable even when transposed or varied, because the overall form is maintained[2].

Arnheim writes:

> The textbooks describe Wertheimer's rules of perceptual grouping: when a person looks at an assortment of shapes, they will be seen as related to one another if they are similar in size or shape or color or some other perceptual trait. Such an assemblage of elements does not seem to exemplify a gestalt process, and in fact the rules of grouping constitute only the first part of a paper in which Wertheimer moved from a more traditional approach to the revolutionary switch, showing that a perceptual pattern cannot be accounted for merely from below, that is, by tracing the relations between elements, as the rules of grouping do. Such an account requires an approach from above: only by describing the overall structure of the pattern can one determine the place and function of each part and the nature of its relations to other parts. (Arnheim 1986, p. 33).

And further:

> However piecemeal and preliminary the rules of perceptual grouping are, they can be shown to involve the basic characteristic of the gestalt attitude, namely, a respect for the inherent nature of the situation confronting the observer. In Wertheimer's view, the rules of grouping are not arbitrarily imposed by the perceiver upon an incoherent collection of shapes. Rather, the constellation of the elements themselves, their own objective properties, influence the groupings made by the observer's mind. (Arnheim 1986, p. 33-34).

[2] It is evident here the reference to Ehrenfels' article "Über Gestaltqualitäten" (1890).

Arnheim, following Wertheimer's intuition that perception has an internal dynamic, introduces the idea of a force field in music. Each musical element generates tensions, attractions, and resistances that develop over time. These forces manifest themselves, for example, in harmonic tensions (e.g., dominant to tonic), melodic expectations (a line "asking" for resolution), rhythmic balances (the balance between accents and rests). As in a visual field, perceivable centers, directions, and structures are formed in a sound field.

Wertheimer does not deal directly with meaning in music. Instead, Arnheim addresses the aesthetic problem: how does music communicate emotion or meaning? His answer is consistent with Gestalt: music expresses through form, not through conventional symbols (as in verbal language), but by organizing perceptions into structures that embody tensions and balances. For Arnheim, musical meaning is not referential (it does not "say" something) but structural: form itself is the bearer of meaning, in that it organizes perceptual experience.

Arnheim writes:

Any organized visual or auditory pattern, be it simple or highly complex, can be used to demonstrate some of Wertheimer's rules. Units close together in space will combine spontaneously as against others from which they are separated by distance. Units resembling one another as to size, shape, color, or motion connect with equal spontaneity. But structural ties of this nature serve not only to turn physical patches of shape or sound into organized patterns and thereby make them perceivable. What makes them so decisively valuable for the artist is that the formal connections thus established in the cortical projection areas of the receiver's brain point to the very essence of the artist's statement. (Arnheim 1992, p. 177).

And further:

> The very term *Gestalt* indicates in German a sublimated or exalted shape or form. Works of art and music were frequently cited in the gestalt literature as outstanding examples of gestalten, not only because they depended so obviously on perfect structural organization but because they purified perceptual form to obtain the clearest and most incisive expression of the work's meaning. (Arnheim 1992, p. 209).

Carroll C. Pratt and Susanne K. Langer

In the twentieth-century debate on the nature and structure of musical language, two prominent figures – Carroll C. Pratt (1894–1979) and Susanne K. Langer (1895–1985) – offer complementary perspectives, rooted in the psychology of perception and the philosophy of symbol, respectively. While starting from different disciplines, both share an underlying assumption: music is not a language in

the verbal sense of the term, but possesses an autonomous structural organization that determines its meaning.

For Pratt, psychologist and author of *The Meaning of Music* (1931), music is based on a set of dynamic qualities – tension, release, movement, accent, rhythm – perceived directly by the listener. These qualities are not derived from arbitrary conventions, but from the direct response of our perceptual system to sound stimuli. His perspective is strongly influenced by Gestalt psychology: music is a continuous flow of forms perceived as a unity, where each part acquires meaning only in relation to the whole. Structural organization, then, is not a code of arbitrary signs, but the result of how our perceptual system integrates sound stimuli over time, generating an experience of continuity and coherence.

Langer, philosopher and author of *Philosophy in a New Key* (1942), shifts the discussion to a semiotic plane. Music, according to her, is a presentational symbolic language, capable of shaping the forms of human feeling without resorting to literal concepts or denotations. Its structure does not merely organize perceived sound events, but constitutes a "logic of feelings": a symbolic system in which the formal relations of tension and release find direct analogy in the structures of emotional life.

The meeting point between the two authors lies in their critique of any reduction of music to verbal language or mere emotional signaling. Both see musical structure as a formal order operating according to its own rules, capable of communicating in the absence of words. However, the divergences are stark: Pratt favors a psychological-perceptual approach, in which structure emerges from the way the listener integrates sound stimuli over time; Langer, on the other hand, proposes a philosophical-symbolic view, in which musical structure takes on value as an analog representation of universal emotional forms.

This difference also implies two ways of understanding the relationship between music and emotion. In Pratt, emotion arises from the direct experience of dynamic qualities, without the need to pass through a symbolic code; in Langer, emotion is more mediated, since music is a symbolic artifact that makes the affective dimension intelligible. In other words, for Pratt, music "acts" on the listener; for Langer, it "speaks" to the listener through a nonverbal but structured language.

The comparison between Pratt and Langer illuminates two complementary sides of the structure of musical language: Pratt takes us back to the bodily and immediate experience of sound form, while Langer shows us how that same form, considered symbolically, becomes a vehicle for knowledge of emotions. Musical structure thus emerges as both a perceptual and a symbolic phenomenon: it is lived form and signifying form, a bridge between matter and meaning.

Viktor Zuckerkandl

Viktor Zuckerkandl (1896–1965), Austrian musicologist and philosopher, assumes that music does not communicate through concepts, but through temporal forms that are experienced. His language is not made up of conventional signs, but of relationships of tension and relaxation, of motion and stillness, that develop over time. In his work *Sound and Symbol* (1956) he illustrates his theory that musical meaning lies neither in imitated emotions nor in abstract structures, but in the phenomenological experience of sound in motion. It is a text characterized by an approach to music as an instrument *of* philosophical inquiry, seeking not so much a philosophy of music as a philosophy *through* music.

One of Zuckerkandl's most profound insights concerns the nature of musical time, which he clearly distinguishes from the measurable time of physics. Musical time is qualitative, not quantitative; directional, not neutral; perceived as tension, expectation, fulfillment. Each musical event derives meaning from its place within a temporal process. For example, a note is "upbeat" or "downbeat"; it is preparation or resolution. This gives rise to a moving form, where each sound has meaning only in relation to what precedes and follows it.

A central concept in his theory is that of the tonal scale as a dynamic field of forces. Notes are not static entities, but tend toward others, generating internal movement. The main tone (tonic) is the gravitational center; other notes generate tensions that want to resolve themselves; musical listening is perception of oriented movement. In this sense, musical language is inherently vector, because each sound is perceived as part of an energy flow.

For Zuckerkandl, music is a sensitive representation of becoming, a "symbol" of temporal being. His conception approaches the phenomenology of Husserl (1859–1938) and Merleau-Ponty (1908–1961): music does not express something outside itself, but reveals the experience of time as such. Musical meaning, therefore, is to be sought neither in concepts nor in feelings, but in the living form of perceived time, made audible by music.

The influence of Gestalt psychology on Viktor Zuckerkandl is quite evident. Like Meyer, he is interested in how music is perceived and organized by the mind, and he finds in Gestalt a powerful conceptual framework to explain the dynamic and unified character of musical experience. For Gestalt, perception is the experience of forms endowed with internal coherence. Zuckerkandl describes music as a flow endowed with directional forces (tension, release, attraction, repulsion) perceived as a unified whole, not as a sum of isolated notes. Moreover, the musical element (note, interval) acquires meaning only in the totality of context; a note is not "something" in itself, but a function within the overall sound form. Just as in visual perception our eye "follows" lines and trajectories, so in musical listening

the mind follows melodic lines and harmonic progressions as natural directions. Zuckerkandl uses an almost physical analogy: each note in a tonal system has a "force field" that directs it toward other notes. This idea of a dynamic field recalls Lewin's "field psychology" and Gestalt perceptual maps.

Meyer uses Gestalt primarily to explain the role of probabilistic expectations and musical emotion. Zuckerkandl uses it to describe the phenomenological way we perceive music as a flow of organized forces over time, almost a "physical" experience of sound space, prioritizing phenomenological experience over purely technical analysis. Zuckerkandl sees music as a dynamic field of tonal forces. Each note has a direction and a role, in a web of tensions that constitutes musical meaning.

Viktor Zuckerkandl's reflections enrich the debate on musical language by offering a unique perspective: music is neither just form nor just emotion, but a living experience of time in sound form. In a world often dominated by logic and measurement, music – for Zuckerkandl – represents one of the last experiences capable of restoring to us the qualitative sense of becoming, and with it, a symbolic dimension of our existence.

Leonard B. Meyer

According to Leonard B. Meyer (1918–2007), a prominent musicologist and music theorist, known for his work on musical perception and aesthetics, the structure of musical language is an organized system that functions similarly to verbal language, but with its own rules and materials.

In his best-known work *Emotion and Meaning in Music* (1956), he focuses on the emotional and cognitive aspect of music, especially in the Western tonal tradition. Just as in Gestalt each part is understandable only in the context of the totality, so for Meyer musical meaning is grasped in the relationship between sound events within a piece; phrases, motifs, and themes are perceived as complete units, not as sums of individual notes. In accordance with gestalt principles, the human brain tends to "complete" incomplete forms; Meyer transfers this concept to music as the listener can predict the continuation of a sequence (melodic, harmonic, rhythmic) and feels emotion when expectations are confirmed, delayed or disappointed.

Music does not communicate "concepts" like verbal language, but emotional and directional meanings; moreover, musical events acquire meaning based on their relational role in structure (e.g., a dominant has meaning because it tends to resolve on the tonic). Musical language is a stylistic system, made up of shared norms (scales, harmonies, rhythms, formal patterns), which organizes sound events in such a way as to create expectations in the listener; structure arises from

the balance between regularity and deviation, and musical meaning emerges from the interplay of tensions and resolutions within these structures. Meyer explores the relationship between expectation, surprise, and emotional meaning, arguing that music communicates and engages precisely because of the tension between what the listener expects and what actually happens.

While Arnheim derives musical tension from dynamic forces internal to the form, for Meyer tension originates from the temporary inconsistency between expectation and reality. Musical emotion thus arises from violated or confirmed anticipation, from how music "plays" with the listener's expectations based on culturally learned and memorized rules. Confirmation of expectation generates satisfaction, delay of expectation generates suspense, violation of expectation generates surprise, intense emotion.

Meyer, integrating Gestalt and semiotics, emphasizes more the cognitive and emotional dimensions, suggesting that musical meaning is constructed in the interaction between music, listener and cultural context.

Lerdahl e Jackendoff

In verbal language, structure is what allows even complex sentences to be generated and understood, thanks to innate rules (in the Chomskyan sense) that organize syntax. Similarly, in music, scholars such as Fred Lerdahl (b. 1943) and Ray Jackendoff (b. 1945) have hypothesized the existence of a musical grammar, a set of cognitive rules that listeners use to interpret hierarchical structures, tensions and resolutions. Every composition, especially in the Western tonal tradition, is built on several interconnected levels: motifs, phrases, periods, sections, each with a specific function within the overall architecture. The principles of generative theory – of Chomskyan derivation – applied to music by Lerdahl and Jackendoff represent one of the most influential contributions to contemporary cognitive musicology. The goal is to describe how a listener internalizes and structures music similar to how he or she acquires language.

> A listener without sufficient exposure to an idiom will not be able to organize in any rich way the sounds he perceives. However, once he becomes familiar with the idiom, the kind of organization that he attributes to a given piece will not be arbitrary but will be highly constrained in specific ways. In our view a theory of a musical idiom should characterize such organization in terms of an explicit formal musical grammar that models the listener's connection between the presented musical surface of a piece and the structure he attributes to the piece. Such a grammar comprises a system of rules that assigns analyses to pieces. (Lerdahl and Jackendoff, 1983, p. 3).

Just as in linguistics there is a distinction between deep structure and surface structure, so too in music there are abstract levels that guide the perception of what is heard. Not all structure is heard directly: many relationships are cognitive, not auditory. The basic structural components are *Grouping structure*, that is, the organization of sounds into coherent units; *Metric structure*, that is, the hierarchical organization of time (strong and weak accents, beats); *Time-span reduction*, that is, the hierarchy of musical durations, where some notes are more "structural" than others; *Prolongational reduction*, which concerns harmonic and tension/resolution relationships.

Gestalt psychology, although mediated through the lens of musical cognition, has strongly inspired Lerdahl and Jackendoff's work, particularly with regard to the perceptual structure of music. Gestalt principles such as proximity, similarity, continuity and closure are translated, in generative theory, into rules of musical grouping.

> The work of Wertheimer and Koffka demonstrates the fundamental claim of Gestalt psychology: that perception, like other mental activity, is a dynamic process of organization, in which all elements of the perceptual field may be implicated in the organization of any particular part. They are at pains to point out and prove two crucial aspects of this claim. First, perception is not simply a product of what is in the environment: the viewer plays an active, though normally unconscious, part in determining what he perceives. Second, the totality of the field as perceived **c**annot be built up piecemeal as a mere accumulation of the perception of its parts each taken in isolation. (Lerdahl and Jackendoff, 1983, p. 303).

According to the authors, the listener organizes the sound surface into coherent units (groups, phrases, sections), following principles very similar to those by which we perceive visual forms. For example, temporal proximity between notes leads to their perception as a unit; rhythmic repetition or variation drives segmentation; tonal or rhythmic breaks signal perceptual boundaries, analogous to gestalt "closure."

Lerdahl and Jackendoff, then, move in continuity with thinkers like Arnheim, but integrate these principles into a formal grammar that describes how the mind organizes music in a hierarchical and structured way, not only perceptually but also cognitively.

Musical language, depending on the theoretical perspective, can thus be interpreted as: perceivable structure (Arnheim), a dynamic emotional expression (Pratt), a presentational symbol of affective life (Langer), a lived form of time (Zuckerkandl), or a play of expectations and deviations (Meyer). These visions are not mutually exclusive but complement each other: each illuminates an essential aspect of musical experience. To understand music is to inhabit the common space between form, time, emotion and perception.

Giovanni Piana and Riccardo Martinelli

In conclusion, I would like to mention two scholars who have participated in the debate on these topics in recent years. In the panorama of Italian philosophical reflection on musical language, Giovanni Piana (1940–2019) and Riccardo Martinelli (b. 1964) represent two profoundly different but potentially complementary approaches. Both question the nature of music and the way it is organized in a language, but they do so starting from distant theoretical premises.

Piana, in a phenomenological key, focuses on the direct experience of sound. The structure of musical language does not arise from a conventional code, but from a pre-linguistic plane in which the sound material – made of timbre, pitch, duration, silence – has its own "resistance" and "expressive latency". Even before being organized by syntactic rules, music exists as a plot of tensions and possibilities that the listener perceives in their immediacy. In this view, silence itself is an integral part of the structure, an active background that makes sounds meaningful.

Giovanni Piana's philosophical-musicological work is therefore placed in a phenomenological horizon in which, however, Gestalt psychology plays a leading role. Piana's interest in music is not limited to aesthetic or historical reflection, but focuses on the perceptual phenomenon in its entirety, placing form in the foreground as the primary reality of the sound experience. Piana adopts the Gestalt paradigm for which perception is not the sum of individual stimuli, transposing it from the visual to the musical field: a theme, a motif or a phrase are not reduced to a sequence of notes, but are presented as a unitary sound configuration, in which each element draws meaning from its relationship with the whole.

Three fundamental aspects of Gestalt emerge in Piana's thought: perceptual unity, whereby music is perceived as an indivisible whole, whose identity is manifested only in the global perception of its form; contextual relationality, whereby the meaning of each sonic event depends on its position and function within the musical totality, similar to how in a visual figure each part takes on meaning from its context; the dynamic field, whereby musical form is a perceived "field of forces", in which tensions and relaxations, tonal attractions and melodic directionality orient the listener, recalling the Gestalt concept of the perceptual field.

A particularly original element in Piana's thought is the recognition of silence as an integral part of musical form: not simply the absence of sound, but an indispensable perceptual background, perfectly analogous to the figure-ground relationship studied by Gestalt.

In short, the influence of Gestalt theory leads Piana to a conception of musical language not as a code of conventional signs, but as an articulation of perceived forms, structures of immediate meaning that arise from the listener's living experience. His philosophy of music thus demonstrates how Gestalt theory, far from

being a mere psychological model, can provide a profound and fruitful interpretative key to understanding the very nature of the musical phenomenon.

Riccardo Martinelli, an Italian philosopher and expert in aesthetics and the philosophy of music, adopts a philosophical-semiotic approach. For him, musical structure is not merely a perceptual phenomenon, but a communicative function that can involve both humans and other species. Musical language is seen as a symbolic system defined by practice, by the relationships between sound and context, and by its ability to convey aesthetic and relational meanings. In his analyses, structure is also understood through historical-philosophical thought and a comparison with nature.

In his work, Martinelli explores the evolution of thinking about sound, positioning Gestalt psychology as a fundamental stage between philosophies of nature, aesthetics, and the anthropology of music. His goal is to restore Gestalt theories to their central role in understanding musical perception, even within a broader historical-philosophical context. Gestalt psychology, in the names of Köhler, Wertheimer, Ehrenfels, and Hornbostel, places the perception of the sound event as a unitary Gestalt at its core – a perceptual form that emerges from the whole, not the sum of isolated parts. Martinelli recalls this principle, highlighting how music only makes sense in the global configuration of tonal, rhythmic, and harmonic relationships. In particular, Köhler demonstrates how tonal qualities cannot be reduced to individual physical-physiological elements, but are perceived as autonomous and meaningful perceptual structures. Hornbostel, broadening his reasoning, integrates cultural experience and anticipates ethnomusicological perspectives, suggesting that Gestalt perceives sound on multiple levels, including the anthropological and cultural.

For Martinelli, musical language makes sense only within a Gestalt framework, where sound phenomena are perceived as emerging figures against a background, generated by a perceptual and cultural context. This vision restores a formal and phenomenological dimension to music, in which the perceived structure is charged with meaning without the need for explicit symbolic mediation.

Martinelli does not limit the influence of Gestalt to a technical or psychological context, but places it within a broader historical-philosophical perspective. In his account (Martinelli 1999), Gestalt emerges as a bridge between the reflections of natural philosophy (Kant, Romanticism) and the phenomenology of sound, relocating musical perception within a dense fabric of cultural and historical meanings.

Conclusion

The investigation of musical language through the perspective of Gestalt theory discloses a coherent lineage of thought that, from Stumpf and Wertheimer

to Arnheim, and from Pratt, Langer and Meyer to Zuckerkandl, Lerdahl and Jackendoff, Piana and Martinelli, has persistently emphasized the primacy of perception in the constitution of musical meaning. Within this tradition, music is not conceived as a conventional or representational system, but as an immediate manifestation of mental organization, a dynamic field in which auditory forms embody the very processes through which consciousness structures experience.

Gestalt psychology thus provides a powerful epistemological framework, capable of bridging phenomenology, psychology, and aesthetics, by demonstrating that musical form arises from relational configurations rather than from additive sensations. Listening, in this view, becomes a formative act of cognition and emotion, through which the perceiving subject encounters its own temporal and affective dynamics. The enduring influence of the Gestalt tradition, reinterpreted by Langer and Zuckerkandl and further developed by Piana and Martinelli, continues to inspire contemporary inquiries into the cognitive, affective, and symbolic dimensions of music, reaffirming its role as a privileged expression of the mind's inherent search for order and meaning.

From this standpoint, music – understood as perceived form – emerges as a privileged field for studying the interplay between sensory experience and the construction of meaning. It is neither a mere language nor a purely aesthetic object, but rather an original cognitive process in which the mind reflects itself in the organized flow of sound. The Gestalt and phenomenological perspectives, far from being historical remnants, continue to offer fertile interpretive frameworks for contemporary cognitive science and philosophy of mind. They remind us that to understand music is to understand the very structure of perception and human experience: a dynamic field of relations, tensions, and meanings in which sound becomes the living image of mental and affective activity.

Summary

In human civilisation, music has always represented a powerful means of communication. Our perceptual system, based on mechanisms of great complexity and extraordinary functionality, allows us to grasp the infinite melodic and harmonic combinations, the changes in the dynamic and agogic patterns, the infinitesimal subtleties of timbre that musical language offers us.

The process of acquiring and fruition of the musical message lives through three distinct phases: the perceptive, the elaborative and the emotional. The perceptual phase is based on the acquisition mechanisms of the sound waves that are picked up and sent to the brain; this is where the elaborative phase occurs and the sound message is given a clear, precise and identifiable structural dimension; the third phase concerns that process that is in some ways still mysterious thanks to whicht he sound language develops all its expressive power and arouses in us a whole range of emotions.

My contribution will focus above all on the elaborative phase, referring to the principles formulated by Max Wertheimer in a famous article of 1923, in which he describes the

ways in which the human mind organises and interprets visual perception. These real 'laws' can also be applied, with rigorous and unquestionable correspondence, to auditory perception. Through the analysis of the most significant contributions of the scholars who have dealt with this subject, I want to highlight the absolute scientific and cultural importance that the Gestalt theory has acquired in the panorama of the psychological schools of the last century.
Even today, however, this current of thought represents an essential point of reference not only in the analysis of perceptual phenomena, but also in the understanding of the complex mechanisms underlying musical language and its engaging artistic and communicative function.
Keywords: musical language, auditory perception, Gestalttheorie, Max Wertheimer, emotional process

Zusammenfassung

In der menschlichen Zivilisation war Musik schon immer ein mächtiges Kommunikationsmittel. Unser Wahrnehmungssystem, das auf Mechanismen von großer Komplexität und außergewöhnlicher Funktionalität basiert, ermöglicht es uns, die unendlichen melodischen und harmonischen Kombinationen, die Veränderungen der dynamischen und agogischen Muster und die infinitesimalen Feinheiten der Klangfarbe zu erfassen, die uns die musikalische Sprache bietet.
Der Prozess der Aufnahme und Verwirklichung der musikalischen Botschaft durchläuft drei verschiedene Phasen: die perzeptive, die elaborative und die emotionale. Die perzeptive Phase basiert auf den Aufnahmemechanismen der Schallwellen, die aufgenommen und an das Gehirn gesendet werden; hier findet die elaborative Phase statt, in der die Klangbotschaft eine klare, präzise und identifizierbare strukturelle Dimension erhält; die dritte Phase betrifft den in gewisser Weise noch mysteriösen Prozess, durch den die Klangsprache ihre ganze Ausdruckskraft entwickelt und in uns eine ganze Reihe von Emotionen weckt.
Mein Beitrag konzentriert sich vor allem auf die ausarbeitende Phase und bezieht sich dabei auf die Prinzipien, die Max Wertheimer in einem berühmten Artikel von 1923 formulierte, in dem er beschreibt, wie der menschliche Geist visuelle Wahrnehmung organisiert und interpretiert. Diese realen „Gesetze" lassen sich mit rigoroser und unzweifelhafter Entsprechung auch auf die auditive Wahrnehmung anwenden. Durch die Analyse der bedeutendsten Beiträge der Wissenschaftler, die sich mit diesem Thema befasst haben, möchte ich die absolute wissenschaftliche und kulturelle Bedeutung hervorheben, die die Gestalttheorie im Panorama der psychologischen Schulen des letzten Jahrhunderts erlangt hat.
Auch heute stellt diese Denkrichtung einen wesentlichen Bezugspunkt dar, nicht nur für die Analyse von Wahrnehmungsphänomenen, sondern auch für das Verständnis der komplexen Mechanismen, die der musikalischen Sprache und ihrer einnehmenden künstlerischen und kommunikativen Funktion zugrunde liegen.
Schlüsselwörter: Musikalische Sprache, auditive Wahrnehmung, Gestalttheorie, Max Wertheimer, emotionaler Prozess

References

Arnheim, R. (1954/1974). Art and visual perception: A psychology of the creative eye. Berkeley and Los Angeles: University of California Press.

Arnheim, R. (1969). Visual thinking. Berkeley and Los Angeles: University of California Press.

Arnheim, R. (1986). *New Essays on the Psychology of Art*, University of California Press, Berkeley-Los Angeles.

Arnheim, R. (1992). *To the Rescue of Art*, University of California Press, Berkeley-Los Angeles.

Chomsky, Noam (1957). *Syntactic Structures.* The Hague: Mouton.

Coppola, W. (2025). Berlin, New York, Trieste – Laboratories of the Arts. In *Handbook of Gestalt Theoretic Psychology of Art*, ed W. Coppola (pp. 22-40). New York and London: Routledge.

Ehrenfels, C. Von. (1890/1988). Über Gestaltqualitäten. *Vierteljahrsschrift für wissenschaftliche Philosophie*, 14, 249-292.

Levitin, D. (2006). *This Is Your Brain on Music: The Science of a Human Obsession*. *New York: Dutton/Penguin. (released in the U.K. and Commonwealth territories by Atlantic, 2007).*

Koelsch, et al. (2000) S., *Brain indices of music processing: "Nonmusicians" are musical* Journal of Cognitive Neuroscience, 12:3, pp. 520-541.

Koelsch, S. (2005). *Neural substrates of processing syntax and semantics in music and language*. *Curr Opin Neurobiol*. Apr;15(2):207-212. doi: 10.1016/j.conb.2005.03.005.

Koelsch S. et al. (2006), Bach speaks: a cortical 'language-network' serves the processing of music, *Neuroimage* 2006; 31(2): 956-968.

Langer, S. K. (1942). *Philosophy in a New Key*. Harvard University Press.

Langer, S. K. (1953). *Feeling and Form*. Scribner.

Lerdahl, F., & Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. MIT Press.

Martinelli, R. (1999). *Musica e natura. Filosofie del suono (1790–1930)*. Milano: Guerini e Associati.

Martinelli, R. (2007). Teoria dei suoni e antropologia: la percezione musicale nella psicologia della Gestalt. *Aisthesis. Pratiche, linguaggi e saperi dell'estetico*, 1(1), 109-126.

Martinelli, R. (2017). *Che cos'è la filosofia della musica*. Roma: Carocci.

Martinelli, R. (2022). *Forme del sentire. Filosofia della musica ed esperienza estetica*. Mimesis.

Meyer, L. B. (1956). *Emotion and Meaning in Music*. University of Chicago Press.

Mithen, S. J. (2005) *The Singing Neanderthals: The Origins of Music, Language, Mind and Body* Cambridge, Massachusetts: Harvard University Press, 2006. Weidenfeld & Nicolson, London.

Patel, A. D. (2003) Language, music, syntax and the brain, *Natur Neuroscience* 6, 674-681.

Peretz, I. & Zatorre, R. (2005). *The Cognitive Neuroscience of Music. Oxford University Press.*

Piana, G. (1988). *Fenomenologia e psicologia della forma (lezioni su fenomenologia della percezione,* Università di *Milano).*

Piana, G. (1991). *Filosofia della musica. Milano: Guerini e Associati.*

Pratt, C. C. (1931). *The Meaning of Music: A Study in Psychological Aesthetics*. New York: McGraw-Hill.

Sammler et al., (2009). Overlap of musical and linguistic syntax processing: intracranial ERP evidence, *The Neuroscience and Music III Disorders and Plasticity,* volume 1169, p. 494-498

Stumpf, C. (1883–1890). *Tonpsychologie* (Voll. 1–2). Leipzig: S. Hirzel.

Stumpf, C. (1926). *Die Sprachlaute: Experimentell-phonetische Untersuchungen.* Berlin: Julius Springer.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt II, *Psychologische Forschung, 2, 301-350.*

Zatorre R. et al. "Structure and function of auditory cortex: music and speech.*" Trends Cogn Sci 2002; 6(1): 37–46.]*

Zuckerkandl, V. (1956). *Sound and Symbol: Music and the External World*. Princeton University Press.

Zuckerkandl, V. (1973). *Man the Musician: Sound and Symbol, Vol. II. Princeton University Press.*

Walter Coppola graduated both in Philosophy and Psychology, and received his Ph.D. in Neurosciences and Cognitive Sciences from the University of Trieste. He is the editor of the new Italian version of *Gestalt Psychology* by W. Köhler (Mimesis, 2024) and editor of the *Handbook of Gestalt-Theoretical Psychology of Art* (Routledge, 2025). His main research interests are vocal rehabilitation, experimentation in acoustic perception and more generally in music psychology.

Address: Department of Life Sciences, Psychology Unit 'Gaetano Kanizsa', University of Trieste. Via Valerio, Building RA, 34100 Trieste (Italy).

E-mail: waltcoppola@gmail.com

ORCID: 0000-0001-5708-4805