ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

IRREGULARITY ON SPECTRAL AND TEMPORAL CHARACTERISTICS IN MUSICAL CONTENT AS AN ARTISTIC APPROACH

MYL/MYE 5032E

Psychology and Neuroscience of Music Final Paper

Fatih ACIKGOZ

Instructor: Dr. Jane Harrison

JUNE 2023

Irregularity on spectral and temporal characteristics in musical content as an artistic approach

Introduction

From the perspective of human experience, reflexes and perceptions emerge as interconnected phenomena. The process of classifying all the elements we perceive and transforming them into physiological responses occurs almost simultaneously. We can consider the impact of these experiences on us as a result of the chain of actions involved in perception and reflection. It is important to note that external sound sources encompass highly complex processes in this context. Here, external sound sources and musical content are not the same thing. However, it should be noted that a definitive distinction between the two has not been made from the past to the present. They are two types of phenomena that are intertwined in many ways. What distinguishes musical content from external sound sources is that music is not merely an external auditory signal. On the contrary, music, besides being an auditory signal, consists of intentional and hierarchically organized sequences of expressive motor acts that go beyond the signal itself (Molnar-Szakacs & Overy, 2006, as cited in Overy & Molnar-Szakacs, 2009, p.489).

What is the correlation between multi-layered encodings of auditory signals and the emergence of personal music preferences or collective audience formations? In this regard, it can be argued that musical contents, such as spectral and temporal structures, possess dynamics that directly influence the interpretation process. This paper aims to explore the approach employed in the production of experimental computer music, which exhibits irregular spectral and temporal characteristics and is built upon different compositional structures than traditional music templates. Furthermore, assumptions will be made regarding the potential effects of such music on auditory perception, interpretation, and classification processes. In doing so, the focus will be

on the collaborative album *The Neurobiology of Moral Decision Making*[1] by Gábor Lázár and Mark Fell, as well as Mark Fell's album *Multistability,*[2] where he introduces microtemporal[3] structures (Fell, 2013, p.93). The reason for selecting these albums is due to the musical (spectral and temporal) approach characterized by a transformative attitude towards structure rather than a deconstructive one. The elements to be addressed here will include beat perception, pitch detection, and potential harmonic fusions.

Rhythmic Structure

In terms of beat perception, the concept of *metric simple* encapsulates the notion of a beat characterized by events sharing the same interval ratio and regular grouping. The fundamental idea revolves around the consistency and periodicity within the rhythmic sequence. The term metric implies the presence of periodicity within the sequence, enabling the human brain to monitor associations between events and potentially extract them into specific templates (Grahn, 2009, p.36). On the other hand, the term *metric complex* encompasses events that also share identical integer ratios, but are arranged in irregular groupings within the composition. It is worth noting that both fall under the category of metrics, representing periodicity, which facilitates the completion of all events as part of beat perception. In "Multistability," Fell explores the concept of micro temporality within the rhythmic structure. While this approach deviates from regular tempo or meter, it still maintains interrelations between events. The algorithm underlying the rhythmic structure generates micro interval ratios in the millisecond range. Although it bears some similarities to the notions of metric simplicity or metric complexity, the algorithm's primary characteristic lies in its ever-evolving time mechanism between events, governed by a specific formula. Consequently, there is a lack of periodicity among events, making it challenging to extract them as distinct beats.

---

[1] https://open.spotify.com/album/4MDRlvmhqupeyV1FD4Axgk?si=tI7CfL0_TDq_nUvC-lL6eg

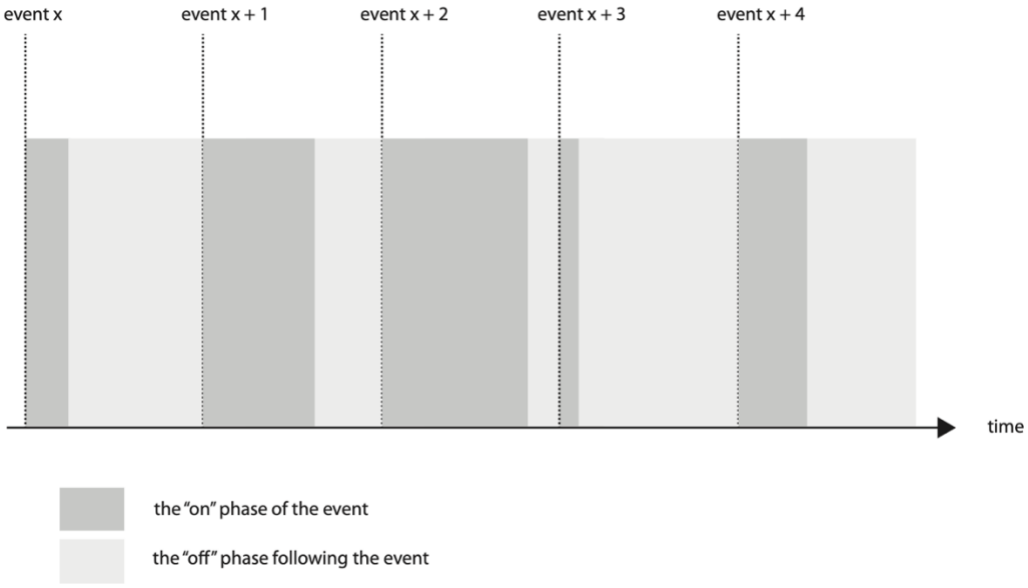
[2] https://open.spotify.com/album/20SbNIxhiKFAYq59fTdQ1c?si=H8Yf1-kuQsqAACDPXBE44g

[3] The term microtemporal represents a tendency to avoid regular rhythm or meter structures.

Grahn discussed periodicity at the smallest interval level (around 220-270 ms) Grahn (2009), while our focus here is on interval ratios ranging from 10 to 1000 ms. It is crucial to understand that when listening to any track in the album, the initial impression urges us to perceive these sequences as beats. However, they constantly evolve within a global timing context, resulting in immediate stretching of interval ratios. The stretching of time between events can be as short as 10 ms, resembling a metric simple structure. However, once it exceeds the 100 ms threshold, it begins to deconstruct the sense of beat perception.

In the figure presented below, Fell provides a depiction of the correlations between time and events. The time mechanism employed in this context operates on a scale of milliseconds, which is consistently multiplied by random floating-point numbers ranging between 0 and 1(Fell, 2013, p.98-99). Consequently, this multiplication process adjusts the duration of each event along the timeline. The purpose of this arrangement is to establish a systematic relationship between time and events, highlighting the dynamic nature of their interplay.

Figure 1: Depiction of the correlations between event time and event duration



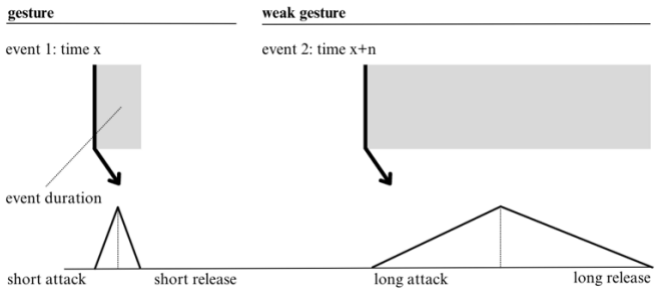
Note. 1 Algorithmic link established between the events that occur and their occurrence times from Works in sound and pattern synthesis: Folio of Works(thesis) (p.99), by Fell, M., 2013, University of Surrey. Copyright 2013 by Mark Fell

Gestures and Textures

The decoding of temporal information in auditory perception has been investigated through functional neuroimaging research, revealing that the left auditory cortex and adjacent regions exhibit superior resolution in tracking temporal variations of stimuli (Zatorre, 2022, p.3). Therefore, it is pertinent to consider the musical approach in these albums, taking into account both temporal variations and simultaneous spectral information. In essence, these albums employ rhythmic variations by incorporating percussive transient elements and utilizing sharp, distinct synthesized frequencies. By combining these elements, the music creates a rich i*nterplay* between temporal and spectral aspects. The term "interplay" refers to the constant switching of variations in event time and event durations within the algorithmic formula that underlies the compositions. In sound synthesis, the application of envelopes to the frequency generates different gestural and textural qualities. "Gestural" refers to the perceived movement or changes within a sound, which can be understood as the result of various acoustic phenomena, such as spectral shifts, changes in intensity, articulations, and timbral variations. Textures, on the other hand, refer to the perceived quality of a sound's surface or structure (Smalley, 1997). Textural content is primarily related to the spectral content of a sound and the interactions between different spectral components. The algorithm serves as the foundation for the compositions in both the album, creating variations among the gestural and textural characteristics of a sound.

> If gestures are weak, if they become too stretched out in time, or if they become too slowly evolving, we lose the human physicality. We seem to cross a blurred border between events on a human scale and events on a more worldly, environmental scale. At the same time there is a change of listening focus the slower the directed, gestural impetus, the more the ear seeks to concentrate on inner details (insofar as they exist). (Smalley, 1997, p.113)

Figure 2: Illustration of the gesture and weak-gesture in a sense of envelopes



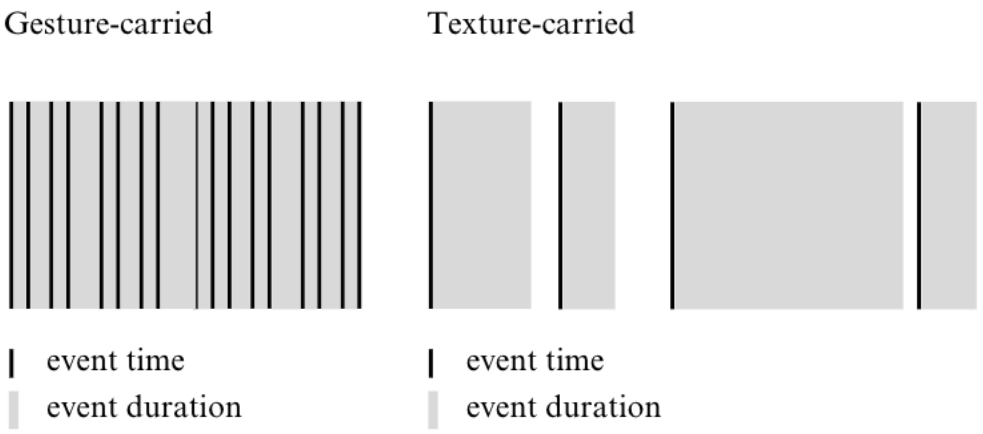
The algorithm introduces variations in event time and event duration, resulting in shifts between these parameters. These shifts give rise to gestural or weakly gestural qualities in the transformed sound, as well as impact the richness or poorness of its texture in terms of spectral characteristics such as overtones and harmonics. Notably, these shifts also influence hemispheric asymmetries in spectral and temporal processing. It is important to note, as previously mentioned, that the algorithm primarily maintains the same content of the event, which is perceived by the right and left hemispheres accordingly. This process can be described as a panning[4] perception between the right and left hemispheres.

The transition between gesture and weak-gesture is evident throughout the album "Multistability." This transition is facilitated by the algorithm, which serves as the foundation of the compositions and constantly evolves over time through changes in the global time parameter. This global time ratio assigns values to the event and event duration, shaping their temporal characteristics. Consequently, the events tend to exhibit gestural qualities with short attack and release times, influencing the overall composition temporally. However, the compositions also consider texture information, as each sound event is created with distinct spectral components. These components play a role in determining whether the event is processed temporally or spectrally.

[4] The term "panning" used here refers to the musical automation aspect that describes the movement of sound across the stereo field, specifically between the right and left speakers.

When the event duration is too short, the ear is unable to focus on the spectral components, leading to a prioritization of temporal coding over spectral processing. It is important to note that when we analyse the texture at a microscopic level, we can observe gestures as contributors to the overall texture, forming a gesture-carried structure (Smalley, 1997). This indicates a harmonious perception of both gesture and texture. This harmonization occurs when the time between events is closer and the event duration is very short, allowing numerous strong gestures to blend seamlessly into a textural event.

Figure 3: Illustration for Gesture-carried and Texture-carried structure



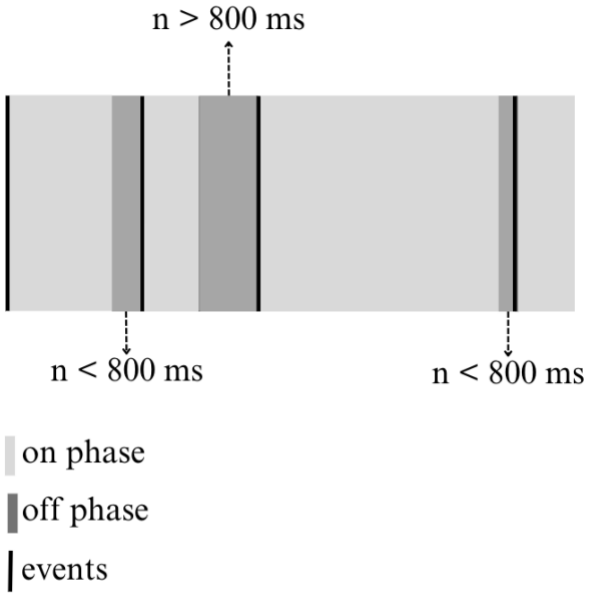
Note. 2 Texture-carried structure mostly has the longer envelopes which stretch out the sound and make gestures weaker. It increases the dominancy of textural contours that lead auditory perception toward the spectral processing.

The composition's algorithm operates with a linear trigger mechanism, resulting in gestural and textural elements progressing sequentially, concluding with the start of each new event. This prompts the question of whether we perceive these events as separate streams or connected to one another. According to the continuity principal Huron (2016), events triggered in close proximity on the basilar membrane are perceived as a continuation of each other, thus creating a continuity effect. However, this effect is only observed when the events fall within the same critical band. As

observed in Figure 3, the Gesture-carried structure showcases the coincidence of all events at their endpoints. The melodic content, represented by the frequency of each event, is determined by the global time parameters. Consequently, it consistently generates frequency ratios that are closer to each other. In the overall album, the frequency ratios of events tend to be similar. However, there are instances where disjunct motion is assigned, disrupting the continuity. When the frequency ratios of events are nearly identical, it produces an undulation effect, resulting in a wavering single line. On the other hand, when the frequency ratios between events exceed the critical band, indicating disjunct motion, it creates two stuttering pitches without any connecting movement (Huron, 2016).

Referring back to Figure 1 by Fell, the concept of an "on" phase and an "off" phase is introduced. The "off" phase represents silent gaps or pauses between events. In other words, if the subsequent event begins immediately after the end of the "on" phase of the previous event, it inevitably generates a wavering single line. However, in accordance with the continuity principle, if the duration of the "off" phase between two events exceeds the 800 ms threshold Huron (2016), it disrupts the potential connection between the events, resulting in two stuttering pitches.

Figure 4: Continuity principal and on and off phrases of the events



Conclusion

In the albums titled "The Neurobiology of Moral Decision Making and Multistability" Mark and Gabor employ musical approaches that give rise to a transitional form encompassing both spectral and temporal characteristics. Consequently, the cognitive processing of spectral and temporal variations engages both the right and left hemispheres, functioning harmoniously in a shared direction. This intriguing phenomenon potentially challenges established hemispheric asymmetries[5]. In conclusion, the algorithmic composition discussed in this essay has had a profound impact on our perception of events over time, as it accounts for shifts between event time and duration, as well as spectral variations. Rather than treating temporal and spectral variations as separate entities, the algorithm integrates these parameters through a global timer parameter. By distributing the divisions of global time ratio across all temporal and spectral sections of the composition, a correlation is established between the spectral and temporal features of sounds, enhancing auditory perception. This algorithmic composition style effectively manipulates the perception path, resulting in novel sound experiences. Importantly, it should be noted that this approach does not deconstruct the textural aspects of sounds; instead, it generates new textural characteristics through temporal ideas, creating a gestural-carrier atmosphere. The artistic approach taken in this composition emphasizes the significance of timing in events, as it can greatly influence the overall structure and spectral features of the composition. The intention is to explicitly prioritize the listening experience and guide it towards a different dimension, which may not necessarily be characterized by irregularity or deconstruction.

---

[5] The term "hemispheric asymmetries" is used as an indicator that, even though both hemispheres serve different functions in cognitive processes, here they tend to behave similarly to each other.

REFERENCES

Overy, K., & Molnar-Szakacs, I. (2009*). Being Together in Time: Musical Experience and the Mirror Neuron System. Music Perception,* 26(5), 489–504. https://doi.org/10.1525/mp.2009.26.5.489

Fell, M. (2013). *Works in sound and pattern synthesis: Folio of Works* (thesis). University of Surrey, Surrey.

Grahn, J. (2009). *The role of the basal ganglia in beat perception: Neuroimaging and neurophysical investigations*. Annals of the New York Academy of Sciences, 1169, pp. 35-45.

Zatorre, R. (2022). *Hemispheric asymmetries for music and speech: Spectrotemporal modulations and top-down influences. Frontiers in Neuroscience, 16*. https://doi.org/*10.3389/fnins.2022.1075511*

Smalley, D. (1997). *Spectromorphology: explaining sound-shapes. Organised Sound*, 2(2), 107–126. https://doi.org/10.1017/s1355771897009059

Huron, D. (2016). *Voice Leading: The Science behind a Musical Art.* http://muse.jhu.edu/chapter/1867070