An Expectancy Perspective on the Cognitive and Neural Mechanisms of Musical Beat Structure

Authors: Sun Lijun, Yang Yufang, Yang Yufang

Date: 2024-07-01T00:00:00+00:00

Abstract

Meter structure, as the organizational framework of music in the temporal dimension, constitutes not only the foundation for composers' creative work, but also the prerequisite for people to appreciate musical beauty, experience musical emotions, and comprehend musical meaning. Within the framework of predictive coding theory, this paper proposes to investigate the neurocognitive basis of musical meter structure through behavioral experiments and EEG technology, focusing on two aspects: prediction and integration. The research specifically includes four components: (1) examining the dynamic neural responses as listeners construct mental representations of meter structure to establish predictions during the unfolding of rhythmic sequences; (2) investigating the neural mechanisms through which listeners update meter structure predictions via prediction errors; (3) using the musical phrase as a structural unit, examining the neurocognitive mechanisms by which listeners integrate hierarchical meter structures at small time scales; and (4) at the level of musical passages, examining how listeners integrate nested meter structures based on long-distance dependencies. These studies will facilitate the revelation of general mechanisms underlying music structure cognition and lay the foundation for constructing neurocognitive models of music. Simultaneously, the relevant findings will also provide objective evidence for music appreciation activities and music aesthetics research, holding potential application prospects in the field of musicology.

Full Text

Preamble

The Cognitive and Neural Mechanisms of Metric Structure in Music: A Predictive Perspective

SUN Lijun[1], YANG Yufang[2],[3]
([1] College of Arts, Nanjing University of Aeronautics and Astronautics, Nanjing

211106, China)
([2] CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China)
([3] Department of Psychology, University of Chinese Academy of Sciences, Beijing 100101, China)

Abstract

Metric structure serves as the temporal framework of music, forming not only the foundation for composers' creative work but also the prerequisite for listeners to appreciate musical aesthetics, experience emotional responses, and comprehend musical meaning. Within the predictive coding framework, this project investigates the cognitive and neural mechanisms underlying the prediction and integration of metric structure through behavioral experiments and electroencephalography (EEG). Specifically, the research comprises four studies: (1) examining the dynamic neural responses as listeners construct mental representations of metric structure and establish predictions during rhythmic sequence unfolding; (2) investigating the neural mechanisms through which listeners update metric structure predictions via prediction errors; (3) exploring the cognitive and neural mechanisms of integrating multi-level metric structures at the phrase level; and (4) examining how listeners integrate nested metric structures based on long-distance dependencies at the period level. These studies will illuminate the general mechanisms of musical structure processing and lay the groundwork for constructing neurocognitive models of music cognition. Furthermore, the findings will provide objective evidence for music appreciation and aesthetics research, offering potential applications in musicology.

Keywords: music cognition, psychology of arts, neural mechanisms, metric structure, electroencephalogram

1. Problem Statement

Music represents the crystallization of human culture, with its diversity and complexity being unique to human society. Unlike literary or visual arts, music lacks the semantic properties of language or the representational nature of painting. As an auditory art form, music essentially constitutes a structured sequence of sounds unfolding in time. Every musical work comprises sounds and rests of varying durations, organized according to specific patterns of strong and weak beats that repeat and develop. Metric structure refers to this cyclical organization of strong and weak beats within a sound sequence, serving as music's temporal framework [**?**, **?**]. Through metric structure, composers organize sound elements with intrinsic pitch relationships into an organic whole to express musical emotion and meaning.

Metric structure not only provides a framework for composition but also supports music appreciation, emotional experience, social activities, and music therapy. During music listening, listeners must extract underlying temporal organizational rules from the sound stream to form metric structure representations, thereby comprehending musical content \cite{欧阳玥，戴志强, 2010; Nave-Blodgett et al., 2021}. Research further indicates that synchronization with temporal structures constitutes a mechanism for emotional induction. When music begins, people often spontaneously feel the urge to dance or engage in rhythmic clapping and foot-tapping. Such beat-synchronized behaviors alter physiological indicators including heart rate, respiration, skin conductance, and neurotransmitter levels, while eliciting positive emotions and pleasurable states [?, ?, ?]. Additionally, music frequently serves as a medium for social interaction. In celebratory or mourning contexts, individuals sharing the same musical background experience a sense of collective connection, with beat-induced synchronization strengthening connections between the vestibular and limbic systems and engaging reward circuits to foster deeper social bonds [?, ?]. Based on these physiological and psychological mechanisms, metric structure plays a crucial role in specific music therapy approaches. For instance, Melodic Intonation Therapy for non-fluent aphasia utilizes melody consistent with speech prosody to restore verbal production, with temporal training potentially being more critical than pitch training [?, ?].

How, then, do people perceive musical metric structure, and what are the underlying cognitive and neural mechanisms? Music listening represents an active cognitive process of continuous prediction and integration [?, ?]. Before a musical event occurs, listeners construct mental representations of sound structure based on previously input information to predict upcoming events. After the event, they must integrate it with prior musical context, updating their mental representation to better anticipate subsequent musical progression. In music's temporal dimension, top-down prediction and integration of metric structure constitute core processes of music listening. While extensive research has examined the pitch dimension—revealing that chord violations elicit early right anterior negativity (ERAN), N5, and P600 components reflecting prediction violation and structural integration [?, ?, ?, ?]—the temporal dimension has received less attention [?, ?], with metric structure research being particularly limited. Moreover, the complexity of musical temporal organization has made metric structure a challenging and under-researched area.

In summary, this project employs EEG within the predictive coding framework to investigate the active cognitive processing of musical metric structure, exploring the dynamic neural responses underlying prediction formation and updating, as well as the mechanisms of integrating metric structure across larger musical units such as phrases and periods. The findings will deepen understanding of musical temporal structure processing and music cognition, illuminate general mechanisms of complex human cognitive activities, provide objective evidence for music appreciation and aesthetics, and offer potential applications in music.

2.1 Cognitive and Neural Mechanisms of Musical Temporal Structure Processing

Despite the temporal dimension being more fundamental than pitch in music [?, ?], existing literature has devoted less attention to temporal processing compared to pitch research [?, ?]. Jones's (1976) Dynamic Attention Theory (DAT) focused attention on the temporal dimension in perception, emphasizing the synchronization between internal attentional energy and external temporal sequences. When event onsets exhibit regularity, people acquire internal timing perception through attentional resource allocation. Building on this, researchers proposed Neural Resonance Theory to explain the neural basis of temporal tracking [?, ?], suggesting that multiple frequency bands of neural oscillations in the brain form phase-locked responses—neural entrainment—with external events of different temporal rhythms. Thus, both attentional energy and neural activity encode regular temporal information through internal-external coupling mechanisms. Given music's highly regular temporal structure [?, ?], these temporal perception theories provide theoretical support for investigating musical temporal structure.

Empirical research on musical temporal perception has primarily examined duration discrimination [?, ?, ?], rhythm perception [?, ?], and beat extraction [?, ?]. Listeners can synchronize with beats through tapping [?, ?, ?] and form accurate predictions about upcoming beats [?, ?, ?, ?, ?]. Recent EEG and MEG studies have demonstrated that beat-synchronized neural oscillations constitute the neural basis of beat perception [?, ?, ?, ?, ?]. However, beat extraction merely represents the foundation of musical temporal structure processing; organizing beats into cyclical metric structures according to strong-weak patterns constitutes the key challenge.

Direct investigations of metric structure remain limited, focusing primarily on discrimination and identification [?, ?, ?]. Although preliminary studies have identified key brain regions involved in metric structure processing—including the right inferior frontal gyrus, middle frontal gyrus, precentral gyrus, and putamen [?, ?, ?, ?, ?]—these studies predominantly employed simple repetitive or isochronous rhythmic stimuli, limiting generalizability to real music listening and failing to reveal the essence of metric structure processing.

Due to this complexity, metric structure represents a weak link in music structure research. This project therefore employs more ecologically valid musical materials to systematically investigate listeners' cognitive and neural mechanisms of metric structure processing, revealing the nature of musical temporal structure processing and providing a more comprehensive understanding of the neurocognitive processes underlying music listening.

2.2 Predictive Processing of Musical Metric Structure

Over the past two decades, researchers have increasingly viewed music listening as an iterative cognitive process of continuously predicting and interpreting sounds [?, ?, ?, ?, ?]. The core mechanism involves constructing appropriate mental representations based on acoustic signals and music structure rules stored in long-term memory to predict upcoming events, then comparing predicted with actual events to update representations and better anticipate subsequent events [?, ?, ?]. Prediction thus represents a crucial stage in music cognition, providing an excellent model for investigating metric structure processing.

Music's limited set of notes and rests can generate infinite rhythmic patterns [?, ?]. Faced with such complex surface rhythms, listeners must gradually extract important sound events from the structural skeleton and assign them strong-weak relationships to construct mental representations of metric structure and form predictions about future events. Research shows that listeners' attentional energy peaks align with strong beats during rhythmic sequence progression, demonstrating synchronization with metric structure [?, ?, ?]. Further studies reveal that rhythmic sequences gradually induce periodic neural entrainment [?, ?, ?], with different temporal rhythms potentially relying on distinct neural oscillation frequencies [?, ?, ?, ?]. However, no research has examined the dynamic process of metric structure prediction formation or the corresponding neural responses.

Metric structure prediction updating is intimately linked to prediction errors. The classic predictive coding theory proposes a brain function model centered on prediction mechanisms. To process continuous information streams, the brain constantly generates predictions at different hierarchical levels, comparing higher-level predictions with lower-level inputs. Mismatches between predictions and inputs generate prediction errors, which drive the brain to integrate input information, update existing mental representations, and minimize prediction errors. The more predictable the prediction, the greater the weight assigned to prediction errors, and the more the brain tends to feed these errors upward to update predictions [?, ?]. Within this framework, researchers have proposed a predictive theory of music emphasizing that music listening involves repeatedly comparing predicted with actual events and updating predictions in real time [?, ?]. We therefore hypothesize that prediction errors constitute the primary driver of metric structure prediction updating, with this process being modulated by predictability.

Linguistic research has validated predictive coding theory, successfully dissociating prediction formation and updating stages along with their neural substrates [?, ?, ?, ?]. In music, although numerous studies have explored musical prediction through structural violation paradigms [?, ?, ?, ?, ?, ?], they have neglected the prediction formation stage and its modulatory role in prediction updating [?, ?]. Critical questions remain: Do listeners pre-activate target stimuli based on metric structure representations? How do prediction formation and updating interact to support predictive processing of metric structure? These issues require further investigation.

In summary, this project examines both prediction formation and updating, investigating the dynamic neural responses underlying mental representation construction of metric structure and how prediction errors enable prediction updating, thereby elucidating the predictive processing mechanisms of metric structure and providing empirical evidence for constructing neurocognitive models of music structure processing.

2.3 Integrative Processing of Musical Metric Structure

A crucial distinction between music and environmental sounds lies in music's complex organizational structure. Musical events form hierarchical "tree-like structures" with logical relationships as they unfold in time [**?**, **?**, **?**]. Perception of these hierarchical structures constrains metric structure processing and constitutes the key to emotional experience and meaning comprehension [**?**, **?**].

Regarding metric structure, sound elements must be integrated locally through short-distance dependencies while simultaneously undergoing global integration via long-distance dependencies. The former refers to dependencies between adjacent elements at small temporal scales (e.g., within phrases), whereas the latter involves logical relationships spanning hierarchical (nested) local structures at larger temporal scales (e.g., across multiple phrases or periods) [**?**, **?**]. Recent findings indicate that small-scale temporal processing primarily relies on unimodal brain regions such as auditory cortex, while large-scale processing engages transmodal regions like the default mode network [**?**, **?**, **?**]. Based on this differential processing of temporal information across scales, we propose that metric structure exhibits distinct processing modes in phrase-level versus period-level organization.

Some studies have examined cognitive mechanisms of hierarchical metric structure within phrases, finding that metric structure violations elicit larger ERAN, N5, and P600 components, reflecting early automatic detection and late integrative processing [**?**, **?**, **?**, **?**]. However, real musical composition and performance involve rich and complex metric organization, forming multi-level hierarchical structures at measure, beat, eighth-note, and finer temporal levels [**?**, **?**]. These studies have not considered hierarchical levels beyond the measure. According to the Coordinated Hierarchical Control (CHC) theory, the brain must mobilize different neural networks to process various hierarchical levels of cognitive processing [**?**, **?**]. Whether and how listeners integrate multi-level metric structures remains uninvestigated.

Everyday music appreciation typically involves large-scale musical units containing more long-distance dependencies. Nested structures in periods, centered on long-distance dependencies, embody musical logic and are essential for high-level musical understanding. However, no research has examined how listeners integrate nested metric structures. In the pitch dimension, researchers have extensively studied nested harmonic structures by manipulating harmonic functions at phrase beginnings or middles, providing empirical evidence for listeners' psychological reality of integrating distant musical events [?, ?, ?, ?, ?]. These methods and findings offer valuable reference for investigating metric structure integration in long-term units. Therefore, examining how listeners process nested metric structures based on long-distance dependencies is both necessary and well-founded.

This project investigates multi-level integrative processing mechanisms of metric structure at both phrase and period levels, and how listeners integrate nested metric structures in larger temporal units based on long-distance dependencies.

3. Research Framework

This project integrates predictive coding theory with metric structure processing to investigate how musical structural information is integrated and updated in the temporal dimension within a prediction error framework. Forming structured internal representations based on available information constitutes the foundation for perceiving metric structure. Once formed, listeners enter a state of predicting upcoming musical events. After event occurrence, they judge whether it matches predictions, integrate it into existing musical context, and adjust their structural representations based on prediction confirmation or violation to improve subsequent predictions. Prediction and integration thus represent two essential and interdependent stages of metric structure processing, reflecting an iterative process of autonomous brain processing.

Focusing on these two critical cognitive stages, this project investigates the cognitive and neural mechanisms of metric structure processing, addressing two key questions: (1) predictive processing of metric structure, and (2) integrative processing of metric structure. The research framework is illustrated in Figure 1 [Figure 1: see original paper], comprising two modules (four studies). Module 1 focuses on predictive processing, examining the dynamic process of constructing metric structure mental representations from continuous information input (Study 1) and the neural mechanisms of prediction updating via prediction errors (Study 2). Module 2 focuses on integrative processing, investigating multi-level integration of metric structure at phrase and period levels (Studies 3 and 4).

3.1 Dynamic Construction of Metric Structure Mental Representations

Study 1 examines real-time neural responses during rhythmic sequence unfolding within the Dynamic Attention Theory framework, revealing the dynamic process of constructing metric structure mental representations. Experimental stimuli derive from real musical works. First, musical materials with five common metric structures (2/4, 3/4, 4/4, 3/8, and 6/8 time) were selected. Four measures were extracted from each, with all pitch information removed to create fixed-pitch rhythmic sequences, then notated and exported using Sibelius software.

According to Generative Theory of Tonal Music [?, ?], each metric type possesses specific structural and strong-weak relationships. For instance, 2/4 time has a simple structure with strong-weak beats per measure, while 6/8 time is more complex, containing strong-weak-weak-substrong-weak-weak patterns.

The analysis approach includes: (1) Global Field Power (GFP), which calculates the standard deviation across all scalp electrodes at each time point. This reference-independent method [?, ?] reveals dynamic changes in EEG responses [?, ?, ?, ?]. GFP will be used for holistic EEG signal analysis to characterize the dynamic nature and changing trends of metric structure representation construction. (2) Intertrial Coherence (ITC) quantifies neural entrainment by computing phase consistency across trials at various frequencies [?]. Based on the association between neural entrainment and beat perception, this study employs spectral tagging methods using ITC to quantify entrainment responses, revealing the correspondence between objective acoustic stimuli and neural responses and elucidating the neural signatures of metric structure representation construction. (3) Comparing differential neural responses across the five common metric structures to reveal how metric type influences representation construction and its neural mechanisms.

Hypotheses:
H1: Based on Dynamic Attention Theory [?, ?], metric structure representation construction is dynamic, with listeners' cortical activity gradually evoking increasingly regular global field potentials and spectral energy distributions as rhythmic sequences unfold.
H2: Consistent with previous findings [?, ?, ?], the neural mechanism of metric structure representation involves neural entrainment and synergistic interactions across different frequency bands.
H3: Different metric types elicit distinct neural entrainment responses, each corresponding to its respective metric frequency.

3.2 Neural Mechanisms of Metric Structure Prediction Updating

Study 2 employs a structural violation paradigm to examine neural responses elicited by prediction errors for target musical events, revealing the cognitive and neural mechanisms of metric structure prediction updating. Experimental stimuli consist of original rhythmic sequences with fixed pitch frequencies, six measures in length. A 2 (Predictability: high vs. low) × 2 (Prediction Accuracy: correct vs. violated) repeated-measures design manipulates both predictability and accuracy of critical musical events. Predictability is high when metric structure is unambiguous and reduced when local alterations to strong-weak patterns occur; prediction accuracy is correct when critical events fall on strong beats and violated when tied notes shift them to weak beats.

The analysis approach includes: (1) Given the presence of critical chords, Event-Related Potential (ERP) techniques will dissociate prediction formation and updating stages, using target stimuli as critical time points to analyze pre- and post-stimulus intervals. Event-related time-frequency EEG responses will examine interactions and constraints between the two stages, elucidating how the former modulates the latter and revealing the complex neurocognitive mechanisms underlying metric structure prediction updating. (2) Phase-Amplitude Coupling (PAC) represents an important neural index of how neural oscillations participate in perceptual processing through synchronization with external rhythms \cite{张雪等，2016}, specifically referring to covariation between low-frequency oscillation phase and high-frequency amplitude that reflects cross-timescale neural organization [?, ?, ?]. Considering that metric structure extraction and representation require coordination across multiple neuronal oscillation scales, PAC will quantify delta-beta cross-frequency coupling during the pre-stimulus phase to examine brain responses during prediction formation under different musical contexts, revealing how predictability influences metric structure prediction.

Hypotheses:
H1: Based on predictive coding theory [?, ?], during the prediction updating stage, prediction accuracy will show a main effect, with violated predictions eliciting larger ERAN effects, P600 effects, and increased beta-band spectral power, reflecting detection, integration, and updating of metric structure prediction errors.
H2: Previous research indicates that delta-beta coupling prior to auditory target onset enhances prediction accuracy [?, ?]. Accordingly, during the prediction formation stage, predictability will show a main effect, with high-predictability conditions eliciting stronger delta-beta coupling.
H3: Predictability and prediction accuracy will interact, primarily manifesting in P600 amplitude and beta-band energy changes during the updating stage [?, ?, ?], supporting the hypothesis that predictability modulates metric structure prediction updating.

3.3 Integration of Multi-Level Metric Structure in Musical Phrases

Study 3 uses musical phrases as structural units, manipulating the metric position of phrase-final events to investigate how people integrate metric structure at small temporal scales and reveal the mechanisms underlying multi-level metric structure processing based on adjacent element dependencies. Musical materials consist of 2/4-time phrases containing six measures, with fixed pitch frequencies. Each rhythmic sequence comprises notes of various durations (half, quarter, and eighth notes), creating multi-level organization. A single-factor repeated-measures design with three levels (Hierarchy: high vs. medium vs. low) manipulates the metric position of phrase-final events at measure, quarter-note, and eighth-note hierarchical levels, creating high-, medium-, and low-hierarchy conditions. High-hierarchy final tones are most stable, providing optimal phrase closure with lowest integration difficulty, followed by medium-hierarchy, while low-hierarchy final tones are least stable and most difficult to integrate. To avoid novelty from pre-final rests in medium- and low-hierarchy conditions and to cue phrase ending, all sequences contain a rest in the penultimate measure. Participants rate phrase closure on a five-point scale after each sequence.

The analysis approach includes: (1) ERP analysis time-locked to final events to investigate electrophysiological responses during metric structure integration at measure, beat, and eighth-note hierarchical levels, revealing the multi-level nature of metric structure processing. (2) Correlating ERP components with behavioral results to identify neural markers of phrase closure. (3) To verify CHC theory's proposed relationship between hierarchical processing and activated brain regions, source localization and functional connectivity analyses will examine brain regions activated during multi-level metric structure integration, revealing neural networks recruited for different hierarchical levels.

Hypotheses:
H1: Different hierarchical levels of phrase-final tones will elicit distinct brain responses, with high- > medium- > low-hierarchy conditions producing higher closure ratings and smaller P600 or N5 effects, supporting multi-level integration. Null results would show no gradient differences.
H2: Based on CHC theory [**?**, **?**], high-hierarchy musical events will activate more anterior brain regions than low-hierarchy events, with stronger connectivity between prefrontal and parietal regions.
H3: Similar to harmonic integration, metric structure integration at different hierarchical levels will exhibit dual-stage processing characteristics of early detection and late integration, reflected in early ERAN and late N5 or P600 components [**?**, **?**].

3.4 Integration of Nested Metric Structure in Musical Periods

Study 4 uses period materials from real musical works to examine neural responses to nested metric structure during natural listening, revealing cognitive processes and neural mechanisms of metric structure integration at large temporal scales. Stimuli derive from 2/4-time musical works, extracting eight-measure fragments containing two phrases, with all pitch information removed to create fixed-pitch rhythmic sequences. A 2 (Hierarchy: non-nested vs. nested) × 2 (Dependency Distance: long vs. short) repeated-measures design manipulates metric structure types in measures 3-4 to alter structural hierarchical relationships between phrases, creating non-nested and nested conditions. Phrase length is varied by extending measures 3-4 to examine dependency distance effects. To approximate natural listening, participants perform an incidental timbre deviation detection task, with deviant trials excluded from analysis.

The analysis approach includes: (1) Time-locking to period onset to compare long- versus short-distance integration, revealing working memory's role in long-term musical unit processing. (2) Investigating shared versus distinct mechanisms for hierarchical structure and dependency distance processing, revealing how listeners process metric structure and comprehend musical meaning in complex organizations. (3) Comparing nested metric and harmonic structure processing to reveal general mechanisms for large-timescale information processing.

Hypotheses:
H1: Nested structures will elicit larger N5 components and increased beta-band oscillatory power than non-nested structures [**?**, **?**, **?**], demonstrating listeners' capacity for long-distance integration of metric structure and the psychological reality of nested metric structure processing.
H2: Dependency distance effects on metric structure processing will primarily reflect working memory load, distinct from nested structure processing mechanisms.
H3: Nested metric and harmonic structures will show similar neural response patterns, reflecting high-level demands of structural extraction imposed by long-distance dependencies.

4. Theoretical Contributions

Music flows and develops through temporal structure, making metric structure processing central to musical meaning comprehension. Focusing on music's temporal dimension from a cognitive neuroscience perspective, this project investigates the dynamic neural mechanisms of metric structure prediction formation and updating, as well as the neurocognitive mechanisms underlying integration of hierarchically complex metric structures with long-distance dependencies. The project extends predictive coding theory in music cognition and lays foundations for neurocognitive models of music, with specific contributions:

First, music listening represents an active cognitive process of sound processing, with anticipating upcoming events being crucial for musical understanding. How is prediction achieved? It requires constructing structured mental representations from available musical information before anticipating subsequent events within that framework. While Dynamic Attention Theory identifies the dynamic nature of temporal structure processing, it cannot identify the construction process or differentiate among various metric structures [**?**, **?**]. Using high-temporal-resolution EEG, this project examines real-time brain responses tracking metric structure, revealing how people gradually construct metric structure mental representations during music unfolding and identifying key neural markers. It also clarifies how different metric types elicit distinct neural responses, illuminating the dynamic prediction formation process for common metric structures. Current prediction research primarily manipulates congruence between individual events and musical context, analyzing post-stimulus neural responses to indirectly confirm prediction [**?**, **?**, **?**, **?**]. Based on predictive coding theory [**?**, **?**], this project focuses on the pre-stimulus prediction formation stage, providing direct neural evidence for pre-activation during prediction formation. By dissociating prediction formation and updating stages, it examines how prediction errors drive metric structure prediction updating and how the former modulates the latter, revealing underlying neural mechanisms.

Second, according to Generative Theory of Tonal Music, hierarchical organization represents the essence of musical logic and constrains musical meaning comprehension [?, ?]. Metric structure itself exhibits complex hierarchy, directly influencing note stability, driving musical development, and creating dynamic momentum. The most common rule for subdividing note durations is binary division, splitting into shorter durations at 1:2 ratios, with each duration level representing the same hierarchical level. Different hierarchical levels combine to form complex networks, particularly evident in complex meters like compound and mixed meters. This project investigates how listeners integrate metric structure within complex hierarchical relationships, extending CHC theory [?, ?] to reveal how the brain mobilizes different neural networks for various hierarchical levels. Additionally, since metric structure exists within hierarchical musical organization, investigating its integration across different temporal scales is essential. Previous research has primarily examined simple metric structure processing at measure or phrase levels [?, ?, ?, ?, ?], with results not generalizable to multi-phrase or period-level structures. In long-term musical units, listeners must integrate nested musical events based on both short- and long-distance dependencies, taxing both working memory and musical structure knowledge. While long-distance integration is crucial in real music listening, current research has only addressed harmonic structure in the pitch dimension [?, ?, ?]. This project examines how listeners integrate nested metric structures based on long-distance dependencies in period-level units, revealing metric structure processing mechanisms in complex musical organizations.

Third, through careful musical material design, this project controls psychoacoustic factors while maintaining ecological validity. Previous studies typically used existing musical fragments [?, ?] or simple isochronous sequences [?, ?, ?, ?]. The former confounds pitch, dynamics, and other factors, while the latter, though controlling acoustic variables, lacks ecological validity for real music listening. This project's materials, adapted from musical works with rigorous experimental manipulation, contain rich temporal information approximating real musical organization while excluding irrelevant variables, thereby enhancing result reliability and generalizability.

In conclusion, within the predictive coding framework, this project investigates prediction and integration processes of metric structure through four integrated studies addressing the scientific question of metric structure cognition mechanisms. Established theoretical foundations ensure feasibility; materials adapted from real music with rigorous control guarantee internal and ecological validity; EEG's fine temporal resolution makes it feasible and necessary for investigating neural mechanisms and dynamic processes of metric structure processing. This project will not only reveal the essence of musical structure cognition and lay foundations for neurocognitive models but also provide objective evidence for music appreciation and aesthetics research, offering potential applications in music.

References

Jiang, J., Wang, Z., Wan, X., & Jiang, C. (2014). Factors influencing musical temporal processing. *Advances in Psychological Science*, 22(4), 650–658.

Jiang, C. (2016). *Psychology of Music*. East China Normal University Press.

Ouyang, Y., & Dai, Z. (2010). A review of research on musical meter cognition. *Advances in Psychological Science*, 18(11), 1692–1699.

Sun, L., Zhou, L., Yan, R., & Jiang, C. (2017). Melodic Intonation Therapy and its clinical application to aphasia. *Psychological Science*, 40(1), 231–237.

Zhang, J., Liang, X., Chen, Y., & Chen, Q. (2020). Cognitive mechanisms of musical syntactic processing and influence patterns of musical structure. *Advances in Psychological Science*, 28(6), 883–892.

Zhang, X., Yuan, P., Wang, Y., & Jiang, Y. (2016). Perception-related neural oscillation-external rhythm synchronization. *Progress in Biochemistry and Biophysics*, 43(4), 308–315.

Zhang, W. (2019). Research on 20th-century musical rhythm: Issues and analysis. *Music Art (Journal of Shanghai Conservatory of Music)*, 1, 74–89.

Arnal, L. H., Doelling, K. B., & Poeppel, D. (2014). Delta-Beta coupled oscillations underlie temporal prediction accuracy. *Cerebral Cortex*, 25(9), 3077–3085.

Asano, R., Boeckx, C., & Seifert, U. (2021). Hierarchical control as a shared neurocognitive mechanism for language and music. *Cognition*, 216, 104847.

Barnes, R., & Jones, M. R. (2000). Expectancy, attention, and time. *Cognitive Psychology*, 41(3), 254–311.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9–21.

Doelling, K. B., & Poeppel, D. (2015). Cortical entrainment to music and its modulation by expertise. *Proceedings of the National Academy of Sciences of the United States of America*, 112(45), E6233–E6242.

Fitch, W. T. (2013). Rhythmic cognition in humans and animals: Distinguishing meter and pulse perception. *Frontiers in Systems Neuroscience*, 7, 68.

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of The Royal Society B Biological Sciences*, 364(1521), 1211–1221.

Fujioka, T., Ross, B., & Trainor, L. J. (2015). Beta-band oscillations represent auditory beat and its metrical hierarchy in perception and imagery. *Journal of Neuroscience*, 35(45), 15187–15198.

Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2012). Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. *Journal of Neuroscience*, 32(5), 1791–1802.

Geiser, E., Notter, M., & Gabrieli, J. D. (2012). A corticostriatal neural system enhances auditory perception through temporal context processing. *Journal of Neuroscience*, 32(18), 6177–6182.

Geiser, E., Sandmann, P., Jäncke, L., & Meyer, M. (2010). Refinement of metre perception-training increases hierarchical metre processing. *European Journal of Neuroscience*, 32(11), 1979–1985.

Geiser, E., Ziegler, E., Jäncke, L., & Meyer, M. (2009). Early electrophysiological correlates of meter and rhythm processing in music perception. *Cortex*, 45(1), 93–102.

Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, 19(5), 893–906.

Habibi, A., Wirantana, V., & Starr, A. (2014). Cortical activity during perception of musical rhythm: Comparing musicians and non-musicians. *Psychomusicology: Music, Mind, and Brain*, 24(2), 125–135.

Hannon, E. E., Snyder, J. S., Eerola, T., & Krumhansl, C. L. (2004). The role of melodic and temporal cues in perceiving musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, 30(5), 956–974.

Harding, E. E., Sammler, D., Henry, M., J., Large, E. W., & Kotz, S. A. (2019). Cortical tracking of rhythm in music and speech. *NeuroImage*, 185, 96–101.

Huron, D. B. (2008). *Sweet anticipation: Music and the psychology of expectation.* MIT press.

Jensen O & Colgin L L. (2007). Cross-frequency coupling between neuronal oscillations. *Trends in Cognitive Sciences*, 11(7), 267–269.

Jia, H. (2019). Microstate Analysis. In L. Hu, & Z. Zhang (Eds.), *EEG signal processing and feature extraction* (pp.141–158). Springer Nature Singapore Pte Ltd.

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, 83(5), 323–355.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, 96(3), 459–491.

Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, 10(3), 235–266.

Koelsch, S. (2012). *Brain and music.* Wiley-Blackwell.

Koelsch, S., Gunter, T., Friederici, A. D., & Schröger, E. (2000). Brain indices of music processing: “nonmusicians” are musical. *Journal of Cognitive*

Neuroscience, 12(3), 520–541.

Koelsch, S., Rohrmeier, M., Torrecuso, R., & Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences of the United States of America*, 110(38), 15443–15448.

Koelsch, S., Vuust, P., & Friston, K. (2019). Predictive processes and the peculiar case of music. *Trends in Cognitive Sciences*, 23(1), 63–77.

Kuperberg, G. R., Brothers, T., & Wlotko, E. W. (2020). A tale of two positivities and the N400: Distinct neural signatures are evoked by confirmed and violated predictions at different levels of representation. *Journal of Cognitive Neuroscience*, 32(1), 12–35.

Lakatos, P., Gross, J., & Thut, G. (2019). A new unifying account of the roles of neuronal entrainment. *Current Biology*, 29(18), R890–R905.

Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., & Schroeder, C. E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *Journal of Neurophysiology*, 94(3), 1904–1911.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, 106(1), 119–159.

Large, E. W., & Snyder, J. S. (2009). Pulse and meter as neural resonance. *Annals of the New York Academy of Sciences*, 1169(1), 46–57.

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

Li, Q., Liu, G., Wei, D., Liu, Y., Yuan, G., & Wang, G. (2019). Distinct neuronal entrainment to beat and meter: Revealed by simultaneous EEG-fMRI. *NeuroImage*, 194, 128–135.

Li, X., Zhang, Y., Xia, J., & Swaab, T. Y. (2017). Internal mechanisms underlying anticipatory language processing: Evidence from event-related-potentials and neural oscillations. *Neuropsychologia*, 102, 1–11.

Liégeois-Chauvel, C., Peretz, I., Babaï, M., Laguitton, V., & Chauvel, P. (1998). Contribution of different cortical areas in the temporal lobes to music processing. *Brain*, 121(10), 1853–1867.

Ma, X., Ding, N., Tao, Y., & Yang, Y. F. (2018). Differences in processing mechanisms underlying the processing of center-embedded and non-embedded musical structures. *Frontiers in Human Neuroscience*, 12, 425.

Margulis, E. H. (2005). A model of melodic expectation. *Music Perception: An Interdisciplinary Journal*, 22(4), 663–714.

Nave-Blodgett, J. E., Snyder, J. S., & Hannon, E. E. (2021). Hierarchical beat perception develops throughout childhood and adolescence and is enhanced in those with musical training. *Journal of Experimental Psychology: General*, 150(2), 314–339.

Nozaradan, S., Peretz, I., & Mouraux, A. (2012). Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *Journal of Neuroscience*, 32(49), 17572–17581.

Nozaradan, S., Schönwiesner, M., Keller, P. E., Lenc, T., & Lehmann, A. (2018). Neural bases of rhythmic entrainment in humans: Critical transformation between cortical and lower-level representations of auditory rhythm. *European Journal of Neuroscience*, 47(4), 321–332.

Overy, K., & Turner, R. (2009). The rhythmic brain. *Cortex*, 45(1), 1–3.

Pearce, M. T. (2018). Statistical learning and probabilistic prediction in music cognition: Mechanisms of stylistic enculturation. *Annals of the New York Academy of Sciences*, 1423, 378–395.

Phillips, C., Kazanina, N., & Abada, S. H. (2005). ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research*, 22(3), 407–428.

Rankin, S. K., Large, E. W., & Fink, P. W. (2009). Fractal tempo fluctuation and pulse prediction. *Music Perception: An Interdisciplinary Journal*, 26(5), 401–413.

Rao, S. M., Mayer, A. R., & Harrington, D. L. (2001). The evolution of brain activation during temporal processing. *Nature Neuroscience*, 4(3), 317–323.

Repp, B. H., & Su, Y. H. (2013). Sensorimotor synchronization: A review of recent research (2006-2012). *Psychonomic Bulletin & Review*, 20(3), 403–452.

Rohrmeier, M. A., & Koelsch, S. (2012). Predictive information processing in music cognition. A critical review. *International Journal of Psychophysiology*, 83(2), 164–175.

Sun, L., Feng, C., & Yang, Y. (2020). Tension experience induced by nested structures in music. *Frontiers in Human Neuroscience*, 14, 210.

Sun, L., Hu, L., Ren, G., & Yang, Y. (2020). Musical tension associated with violations of hierarchical structure. *Frontiers in Human Neuroscience*, 14, 578112.

Sun, L., Liu, F., Zhou, L., & Jiang, C. (2018). Musical training modulates the early but not the late stage of rhythmic syntactic processing. *Psychophysiology*, 55(2), e12983.

Sun, L., Thompson, W. F., Liu, F., Zhou, L., & Jiang, C. (2020). The human brain processes hierarchical structures of meter and harmony differently: Evidence from musicians and nonmusicians. *Psychophysiology*, 57(9), e13598.

Thaut, M. H., Trimarchi, P., & Parsons, L. (2014). Human brain basis of musical rhythm perception: Common and distinct neural substrates for meter, tempo, and pattern. *Brain Sciences*, 4(2), 428–452.

Trost, W. J., Labbé, C., & Grandjean, D. (2017). Rhythmic entrainment as a musical affect induction mechanism. *Neuropsychologia*, 96, 96-110.

Trost, W., & Vuilleumier, P. (2013). 'Rhythmic entrainment' as a mechanism for emotion induction and contagion by music: A neurophysiological perspective. In T. Cochrane, B. Fantini & K. R. Scherer (Eds.), *The emotional power of music* (pp. 213–225). Oxford University Press.

Van der Steen, M. C., & Keller, P. E. (2013). The Adaptation and Anticipation Model (ADAM) of sensorimotor synchronization. *Frontiers in Human Neuroscience*, 7, 253.

Wu, Q., Sun, L., Ding, N., & Yang, Y. (2024). Musical tension is affected by metrical structure dynamically and hierarchically. *Cognitive Neurodynamics*, 1–22.

Vuust, P., Gebauer, L. K., & Witek, M. A. (2014). Neural underpinnings of music: The polyrhythmic brain. *Advances in Experimental Medicine and Biology*, 829, 339–356.

Vuust, P., Heggli, O. A., Friston, K. J., & Kringelbach, M. L. (2022). Music in the brain. *Nature Reviews Neuroscience*, 23(5), 287–305.

Vuust, P., Ostergaard, L., Pallesen, K. J., Bailey, C., & Roepstorff, A. (2009). Predictive coding of music-brain responses to rhythmic incongruity. *Cortex*, 45(1), 80–92.

Vuust, P., Pallesen, K. J., Bailey, C., van Zuijen, T. L., Gjedde, A., Roepstorff, A., & Ostergaard, L. (2005). To musicians, the message is in the meter: Pre-attentive neuronal responses to incongruent rhythm are left-lateralized in musicians. *NeuroImage*, 24(2), 560–564.

Wolff, A., Berberian, N., Golesorkhi, M., Gomez-Pilar, J., Zilio, F., & Northoff, G. (2022). Intrinsic neural timescales: Temporal integration and segregation. *Trends in Cognitive Sciences*, 26(2), 159–173.

You, S., Sun, L., Li, X., & Yang, Y. (2021). Contextual prediction modulates musical tension: Evidence from behavioral and neural responses. *Brain and Cognition*, 152, 105771.

You, S., Sun, L., & Yang, Y. (2023). The effects of contextual certainty on tension induction and resolution. *Cognitive Neurodynamics*, 17(1), 191-201.

Zendel, B. R., Lagrois, M.-É., Robitaille, N., & Peretz, I. (2015). Attending to pitch information inhibits processing of pitch information: The curious case of amusia. *Journal of Neuroscience*, 35(9), 3815–3824.

Zhou, L., Liu, F., Jiang, J., & Jiang, C. (2019). Impaired emotional processing of chords in congenital amusia: Electrophysiological and behavioral evidence. *Brain and Cognition*, 135, 103577.

Note: Figure translations are in progress. See original paper for figures.