様式３

論　文　内　容　の　要　旨

氏　名　（　石田　海　）	
論文題名	Multidimensional Predictions in Music Perception: An Event-Related Brain Potential Study （音楽知覚における多次元的予測：事象関連脳電位による研究）

論文内容の要旨

Why are certain tones heard as music? A possible answer is that when a person listens to music, predictions derived from the musical representation (i.e., music predictions) are generated, but not when the person listens to a simple environmental sound. Indeed, previous studies have considered that tones in music are perceived with a prediction. Behavioral and neuroscientific studies have investigated music perception by recording prediction errors when expectations are violated. Using event-related potentials (ERPs), the present study investigated the internal workings of the predictive processes outlined by the predictive coding of music (PCM) model (Vuust et al., 2022a) with two aims: first, to empirically examine whether musical expectations of the tone object can be generated according to the classification of Vuust et al., and second, to examine the multidimensional predictive processes associated with each expectation and their relationships with each other. Six experiments were conducted to achieve these aims.

Experiment 1 compared the detection processes of music-syntactic and acoustic irregularities to examine the predictive processes reflected in the early right anterior negativity (ERAN) and the mismatch negativity (MMN). Nonmusicians (in the infrequent-presentation and equal-presentation groups) were asked to listen to chord sequences with 5 four-voice chords each while watching a silent video clip. Standard, harmonic-deviant, intensity-deviant, and double-deviant chords occurred at the final position in each sequence. Deviant stimuli were presented infrequently (p = .10, in the infrequent-presentation group) and equiprobably (p = .25, in the equal-presentation group). Regardless of the deviance probability, when the two deviant types occurred simultaneously, the negativity increased additively: the amplitude of the double-deviant event-related potential (ERP) was as large as the sum of the single-deviant ERPs. These results suggest that the detection of music syntactic and acoustic irregularities works independently based on different regularity representations. Therefore, dynamic and schematic expectations can be dissociable in terms of neural prediction errors, that is, ERAN and MMN.

Experiment 2 examined the relationship between schematic and dynamic expectations when these two expectations predicted the same note. At the final note of the melodies, the schematic expectation was violated by presenting a note with music-syntactic irregularity (i.e., an out-of-key note), while the dynamic expectation was violated by presenting a contour deviant based on online statistical learning of melodic patterns. The schematic and dynamic expectations were manipulated to predict the same note. The ERPs were recorded for the music-syntactic irregularity and the contour deviant, which occurred independently or simultaneously. The results showed that the music-syntactic irregularity elicited an ERAN, reflecting the prediction error in schematic expectation, whereas the contour deviant elicited an MMN, reflecting the prediction error in dynamic expectation. The two components occurred within similar latency ranges. Moreover, the ERP amplitude was multiplicatively increased when the irregularity and deviance occurred simultaneously. These findings suggest that the predictive processes of schematic and dynamic expectations function concurrently in an interactive manner when these two expectations predict the same note.

Experiment 3 examined the plasticity of two musical expectations. At the final note of the melodies, schematic expectation was violated by presenting a note with music-syntactic irregularity (i.e., an out-of-key note), while dynamic expectation was violated by presenting a contour deviant based on online statistical learning of melodic patterns. The contour patterns in the first and second sessions were reversed. The ERPs were recorded for the syntactic irregularity and the contour deviant, which occurred independently or

simultaneously. The results showed that the syntactic irregularity elicited the ERAN in both sessions, reflecting the prediction error in schematic expectation, whereas the contour deviant elicited the MMN only in the first session, reflecting the prediction error in dynamic expectation. These results suggest that schematic and dynamic expectations function separately and that schematic expectations have less plasticity than dynamic expectations.

Experiments 4 and 5 examined whether knowledge of music-syntactic regularity for schematic expectations could be acquired through statistical learning using MMN responses elicited by chord transitions that deviated from newly learned transition regularities. Adult nonmusicians without hearing impairment participated in the experiments. Experiment 4 examined the MMN response during the acquisition phase, whereas Experiment 5 examined the MMN response after two days of learning. The same stimuli were used in both experiments. Six types of chords consisting of the tone of the 18 equal temperament scale were concatenated to create a learning sequence in which a particular progression pattern appeared with high probability (p = .90, standard) or low probability (p = .10, deviant). To ensure that only the progression pattern was learned, the chords were presented with equal probabilities. In Experiment 4, the ERPs were recorded while the participants were listening to the learning sequence and performing a timbre change detection task as a cover task; and in Experiment 5, the ERPs were recorded while the participants were passively listening to a sequence in which the standard and deviant progressions appeared at p = .50 each, after two days of online learning of the novel progression pattern. After the ERP recording, a behavioral familiarity test was administered in which the participants selected the standard progression in each of the two chord progressions. The results showed that the deviant progressions elicited an MMN response and an ERAN-like mismatch response, reflecting online irregularity detection and deviation from the acquired pattern representation, respectively. In Experiment 4, the participants were unable to choose the standard chord progression except by chance, but they were able to choose it in Experiment 5, which might reflect the longer learning period. These results suggest that the schema of chord progression patterns can be acquired through statistical learning even by adult nonmusicians without expert musical knowledge and skills.

Finally, Experiment 6 examined how specific the representation of the expected note was by recording the omitted stimulus potentials (OSPs) to avoid contamination of bottom-up sensory processing with top-down predictive processing. To manipulate predictability, melody familiarity was manipulated in the familiar melody, in which veridical expectations based on memory were generated, and in the unfamiliar melody, in which veridical expectations were not generated. Decoding of the omitted content was attempted using a support vector machine, which is a type of machine learning. The ERP responses of 25 participants to the omission of four target notes (E, F, A, and C) at the same position in familiar and unfamiliar melodies were recorded. The results showed that the omission N1 was larger in the familiar melody condition than in the unfamiliar melody condition and that the decoding accuracy of the four omitted notes was significantly higher in the familiar melody condition than in the unfamiliar melody condition. These results suggest that the OSPs contain discriminable predictive information related to veridical expectations, and the higher the predictability of notes, the more specific the representation of the expected note is generated.

The results of Experiments 1–6 demonstrated the presence of three types of expectations (dynamic, schematic, and veridical), and their multidimensional predictive processes were examined by recording the ERP responses. Experiments 4 and 5 showed that music-syntactic knowledge, which is the basis of schematic expectations, can be acquired through statistical learning of chord transition patterns. Based on these results, the present study proposed the object-based predictive model for music perception, which assumes that the expectation of an object has three dimensions: dynamic, schematic, and veridical. Although the present study did not consider individual and cultural differences, these factors can be investigated by extending the object-based predictive model to incorporate individual priors. By applying, modifying, and extending the current model to different cultural contexts and temporal predictions in music, the model is expected to contribute to the understanding of general processes of music perception in humans.

様式 7

論文審査の結果の要旨及び担当者

氏　名　（　石　田　　海　）			
	（職）		氏　　名
論文審査担当者	主 査	教授	入戸野　宏
	副 査	教授	八十島　安伸
	副 査	教授	篠原　一光
	副 査	准教授	山本　倫生

論文審査の結果の要旨

音と音楽の違いはどこにあるのだろうか？　本論文では，さまざま手がかりや知識に基づいてその後の展開を予測できる音が「音楽」になりうると考えて，その予測の構造や仕組みを検討している。

第 1 章では，関連する概念（予測，予期，期待など）の定義を明示するとともに，これまでの心理学的・神経科学的知見を整理している。特に，音楽の予測符号化モデル（Vuust et al., 2022, *Nature Reviews Neuroscience*）に注目し，音楽に関する予期には，スキーマ的予期（schematic expectations），動的予期（dynamic expectations），事実的予期（veridical expectations）の 3 種類が存在すると考えた。一つの音（音オブジェクト）に対して，これら 3 種類の予期がどのように働いているかを実証的に検討するのが本論文の目的である。

第 2 章から第 7 章までは，6 つの実験を報告している。いずれの実験でも，実験参加者の大学生（各実験で 24〜40 名程度）が音刺激を聞き流しているときの脳波を頭皮上 34 部位から記録し，予期が外れたときに生じる事象関連電位—音の提示後 100〜200 ミリ秒で生じる右前部初期陰性電位（early right anterior negativity: ERAN）やミスマッチ陰性電位（mismatch negativity: MMN）—を指標とした。第 2 章（実験 1）では，音楽統語（調性）の逸脱と物理特徴（音圧）の逸脱に注目し，両者に対する脳電位反応が加算的に重畳することを示した。第 3 章（実験 2）と第 4 章（実験 3）では，音楽の規則性についての事前知識に基づくスキーマ的予期と，その場で音楽を聴きながら短期的に形成される動的予期との関係を検討した。その結果，スキーマ的予期と動的予期が同じ音を予測する場合には，両者が相乗的または加算的に働くことが示された。また，文脈が変わると，動的予期の逸脱に対する反応は変化するが，スキーマ的予期の逸脱に対する反応は変化しないことも示された。第 5 章（実験 4）と第 6 章（実験 5）では，和音系列を繰り返し聴取することでその遷移パターンを学習できるかを検討した。その結果，学習中にも学習後にも標準的でないパターンに対して逸脱反応が生じることが分かった。第 7 章（実験 6）では，事実的予期の逸脱を検討するために，よく知っているメロディを聞かせながら一部の音を欠落させた。その結果，物理的に音が提示されなくても，予期があるときには脳電位反応にその音のピッチの情報が含まれることが示された。

第 8 章では，6 つの実験結果をまとめ，音楽知覚についてオブジェクトベースの予測モデルを提案した。その上で，音楽の他の側面（リズムなどの時間的側面）や個人差・文化差への拡張といった今後の研究方向性を示唆した。

以上のように，緻密な実験と分析を通じて一連の研究で得られた成果は，関連分野に貢献する堅実な知見であると評価できる。審査の結果，本論文は博士（人間科学）の学位を授与するにふさわしいと判定した。