

音楽を科学的に扱う方法論

平田 圭二^{1,a)}

概要 :

音楽を計算機で扱おうとすると、感性、嗜好、芸術性に興味が向く場合が多い。そのため音楽について、数理的な基礎に基づく処理体系を構築するのは困難と思われている。一方、自然言語、画像、音響信号などの他メディアの場合、確固たる数理的基盤の上で科学的なアプローチが可能であるがゆえに、幾多の有用なアルゴリズムが発見されたりシステムが実現されている。本稿では、音楽でも科学的なアプローチを可能とするために、どのように数理的基盤を構築すべきかについて議論する。音楽と言語の比較、音楽理論と言語理論の比較を通して、音楽理論 Generative Theory of Tonal Music (GTTM) に基づく普遍的な構造や性質を扱うレイヤの構成法を提案する。それは、音楽理論を形式化して音楽についてのモデルを得ることに等しい。音楽は実世界の事物に記号接地しておらず、どんなナンセンスな意味付けも簡単にできてしまうため、音楽のモデル化では認知的リアリティの維持が重要であることを主張する。

キーワード：音楽の意味論、Generative Theory of Tonal Music (GTTM)、簡約、情報学

Scientific Methodology for Handling Music

HIRATA KEIJI^{1,a)}

1. はじめに

言語学には言語理論があり、意味論、統語論、語用論などが含まれている。これら言語理論の目的は、知的な現象としての言語の普遍的な構造や性質を解明したり、その認知的メカニズムを示すことである。言語理論は言語を還元論的に扱うことを可能にし、自然言語処理は言語理論に基づいて発展してきた。実用に供される自然言語処理システム、例えば検索、翻訳、要約、質問応答システムなどでコーパスやテストコレクションを定めることができ、有用性に資する特徴に関する評価尺度を定めることができるのは、言語理論に負うところが大きい。^{*1} コーパスを定めることでドメインが明らかになり、評価尺度を定めたことで特徴に関する改善が客観的になる。一般に研究を科学的に進める際に必須の特性として反証可能性があるが、言語理論

によって言語は科学的な探求の対象となった。システム開発の場合は、反証可能性ではなくシステムの改善点を特定できる性質が相当すると考えられる。人が操る言語には感性、嗜好、芸術性の要素も含まれているが、言語学においてこれらを考える分野は、文学、美学、修辞学などとして区別されることが多い。自然言語処理は感性、嗜好、芸術性の要素をあえて扱わないことによって有用な技術開発を可能にしたとも考えることができよう。

音楽にも音楽理論がある。音楽理論の目的は、音楽の普遍的な構造や性質を解明したり、その認知的メカニズムを示すことである。音楽理論は基本的に、感性、嗜好、芸術性を扱わず、それらは美学、芸術論として区別されることが多い。音楽情報処理は情報学の分野における音楽を研究対象とする分野であり、音楽理論に基づくアプローチを採用する研究もあれば、音楽理論が取り扱わない音楽の感性、嗜好、芸術性の要素に着目した研究も含まれる。後者の方法論の場合、ベースとなる理論が存在しないため、観測できる範囲の人の振る舞いと観測できる範囲の音楽情報（楽譜、

¹ 公立はこだて未来大学 システム情報科学部

^{a)} hirata@fun.ac.jp

^{*1} 例えば NTCIR Project: <http://research.nii.ac.jp/ntcir/index-ja.html>

音響信号)との対応を調べ何らかの法則性や相関を見出すというものである。この課題に対する1つの解決法は、ビッグデータからのデータマイニング(機械学習)である。すでに幾つか実用的なサービスが提供されている(例えば Music Xray *2, The Echo Nest *3など)。ビッグデータ技術に関して一般に指摘されていることとしてデータのメンテナンスの問題がある[16]。得られたデータが時間経過したりデータ取得対象が変化することでビッグデータ全体が「劣化」を起こすので、データマイニングの頻度をどのように定めるか、データマイニング対象とするデータ期間や取得範囲をどのように定めるかが課題となる。さらに、データマイニングから得られた情報をユーザにフィードバックすることは、ユーザの振る舞いをより速く変化する方向に誘導する。これは、データのメンテナンスの問題が今後より重要になることを示唆している。

音楽理論は、その歴史的経緯において、もともとは作曲のための理論であった。つまり、新しい曲を作り出すためには既存の音楽作品の構造や様式を分析し理解することが必要であり、作曲ための分析方法論と位置付けられていた。しかし18世紀頃より、音楽理論の研究者と作曲家の分化が見られるようになる。作曲自体の理論は、音楽学の中で体系化されるというより、実践的な規則集(作曲法)として発展してきた[17]。人が作曲をする際に用いる知識として、実践的な規則集という表現・伝達形式が適していたと考えられる。一方、情報学において計算機を用いて音楽を作る方法はアルゴリズム作曲(algorithmic composition)と総称される[6], [19], [22]。既存の情報技術を作曲に応用するというアプローチが主体であり、作曲法とは比較的独立に発展してきている。これは、実践的な規則集を計算機上で実現するのが困難であったことが一因であろう。では、計算機の存在を前提とし、音楽理論に根差した作曲法・作曲システムとはどのようなものであろうか。

言語理論と音楽理論の対応を考えると、音楽理論に基づくアプローチを採る研究から、自然言語処理のように有用な技術が産み出される可能性があるのではないかと。音楽情報処理システムに関しても、音楽理論に基づいてコーパスやテストコレクションを定め何らかの評価尺度を定めれば、音楽に関しても科学的な探求が可能となるのではないかと。仮に音楽における感性、嗜好、芸術性を扱うとしても、音楽理論に基づく普遍的な構造や性質を扱うレイヤを設定しその上に感性、嗜好、芸術性に対応するレイヤを設定する方が、柔軟性と有用性*4において良好なシステムを構築できると考える。

2. 音楽の意味論

音楽における普遍を探求するための出発点は、人の音楽認知における間主観あるいはゲシュタルトである。ここでは、音楽における情動を次の意味で用いる:今聴いている音楽的対象が過去の経験に基づいて期待されていたものと同じか違うかを推論し、その結果生じる驚き、不安、安定、緊張、弛緩などの心理的变化。情動は一時的ですぐ消える間主観的な心理現象であり、審美的な要素を含まない。*5

2.1 音楽の意味の分類

まず、記号論によれば次の3要素によって意味が生じるという:表現あるいは記号(signifiant), 表現が指し示すもの(表現した結果, signifié), 表現の結果に対する意識的な観察者。音楽理論におけるどのような概念がこれら3要素に対応付けられるかを議論する。

表現は楽譜として記された楽曲に対応する。表現が指し示すものに関して、Leonard B. Meyer [14] は記号論に従って内在的か外在的かという観点と形式か情動かという観点から次のように分類した。

内在的	音楽の意味はその音楽に内在する。絶対的。
外在的	音楽以外の指し示された世界に存在する。参照的。
形式	音楽そのものが作る抽象的な形
情動	情動的経験を与える手段

ここで参照*6はさらに指示的な参照(音楽とは異なるカテゴリの事象を指し示す)と具現的な参照(音楽自体を指し示す)に分かれる。前者はJohn B. Davisが「Darling, they are playing our tune 理論」と呼んだものであり、形式的な意味を論じる立場からは重要ではない。さらにLeonard B. Meyerによれば、意識的な観察者にとって表現の結果は情動(emotion)として現れると言う[14]。例えば、絶対的表現主義とは、音楽の意味は音楽を聴くその過程に限定的に内在しており、非指示的に音楽を聴いても情動的意味が現れると考える立場である。参照的表現主義とは、音楽の参照的な内容を理解できるかどうか情動的表現を左右すると考える立場である。例えば、ドビュッシー作曲アラベスク冒頭の数小節の旋律概形はモスク壁面装飾に見られるアラベスク模様と同じであり、これは形式としての指示参照的な意味に分類される。

また19世紀の音楽評論家Eduard Hanslick(1825-1904)によれば、「音楽の内容とは鳴り響きつつ運動する形式である」、そして「音楽は音だけで構成される閉じた世界であ

*2 <https://www.musicxray.com/>

*3 <http://the.echonest.com/>

*4 http://en.wikipedia.org/wiki/Flexibility-usability_tradeoff

*5 比較的永続的で安定した個人的現象は気分(mood)と呼ばれる。

*6 音楽には指示代名詞が存在しないので、音楽における具現的な参照とは、旋律、リズム、和音等の並置、繰り返し、変形等によって暗示される関係性を指す[21]。

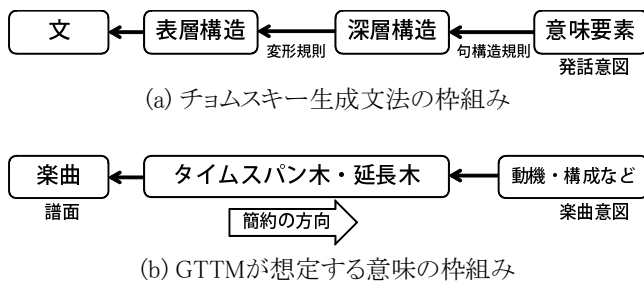


図 1 文と楽曲に意味を与える枠組

り、文学的、情緒的な内容は排除されるべきである」^{*7} これは、音楽とは内在（絶対音楽）・情動・具現的な立場から聞くものであり、ある音楽的事象（1つの音、フレーズ、和音など）が他の音楽的事象を参照することで生じる情動をもって意味とする立場である。このような意味の与え方について第3章にてさらに詳しく検討するが、筆者は、この立場には音楽認知に関して十分な普遍性があると考える。

2.2 意味を与えるための理論構造

音楽の意味を探るため、言語と音楽の間のさまざまな共通点や差異が比較、議論されてきた [1], [3], [7], [10], [15], [23]. 第1章で議論したように、我々は音楽理論と言語理論を比較することも有意義であると考えている。言語では、表層の記号と明示されていない意味の区別が明確に意識されるため、表層の記号と意味がどのように関係しているかを説明する理論が古くから多数提案されてきた。音楽では、Leonard B. Meyer 以降に同様の動きが起き、チョムスキー (Noam Chomsky) の生成文法の枠組に則って、音楽理論 Generative Theory of Tonal Music (GTTM) [12] が提案された。

まず、チョムスキー生成文法の枠組において人が言語コミュニケーションして理解するというのは、表出側の人が受取側の人にメッセージ（コンテンツ）を送り、受取側の人が表出側と同様なある構造をメッセージの記号列に割り当てて指す。その構造は、階層的順序構造（木構造）としてチョムスキー以降長年論じられてきたものである [8], [13]. 図 1(a) に示すように、発話意図が意味要素によって表現され、句構造規則と変形規則によって深層構造と表層構造に変形される。これらの規則が変形される木構造に意味を与えていく。意味を与える過程は文を産出する方向と同じである。一方 GTTM でも、楽曲の意図が動機や大域的な構成 (Schenker 理論の *Ursatz* に相当する [2]) によって表現され、そこから簡約 (reduction) の逆操作である具体化 (elaboration) によって、2種類の木構造 (タイムスパン木と延長木) が産出される (図 1(b)). 具体化の操作は

GTTM の規則群に従う。GTTM の規則は、分析の結果生じ得る全ての木構造を生成する構成規則 (well-formedness rule) と、生成された木構造の内いずれがより好ましいかを規定する選好規則 (preference rule) から成る。この具体化の操作が2種類の木構造に意味を与える。意味を与える過程は GTTM 分析過程と逆向きである。

GTTM は言語の生成文法の枠組を踏襲していると言われているが、GTTM と生成文法で規則の位置付けや木構造の役割がそれぞれ異なっている。Lerdahl と Jackendoff は以下のように述べている：

我々 (Lerdahl and Jackendoff) が見い出したのは、音楽の生成理論は、言語の生成理論と異なり、1つの楽曲に複数の構造を割当てるという点と、好ましい解釈としてその構造に重みをつけたり軽くしたりすることで一貫性を保ちながら各々の構造を区別させられるという点である。… 音楽分析結果を生成する選好規則は重要な役割を果たすのに対し、言語の生成文法には対応するものが存在しない。この選好規則が、GTTM と生成文法という2つの形式の大きな違いに当たる。(筆者訳、抜粋 [12], p.9)

一般に、言語の生成文法では、異なる文には異なる木構造が1対1の関係で付与されるので (もちろん例外はある)、文の表層構造はそれを直接操作あるいは計算するのに十分な情報量を持っていると考えることができる。それに対して GTTM では、選好規則のため、表層構造である楽譜と木構造の関係は生成文法より曖昧となる。楽譜より木構造の方がより多くの情報を運んでいると見なせる。

3. 音楽を聴取する際の2種類の認識

本章以降、GTTM が与える2種類の木構造 (タイムスパン木と延長木) に関して、どのような普遍的な意味を表現しているか考察する [7].

音楽聴取時の認識には、予測できないピッチイベント^{*8}と予測できるピッチイベントがある [14], [18], [24]. まず最も単純な場合として、ドドドド…と続く旋律を聴くこと考える。この旋律がどこまで続くか、どこで途切れるかを予測する (期待する) ことは難しい。しかし、聴取者にとって突然ドが鳴らない時点があると (ドドド休ド…) その時点でゲシュタルトつまりグループの境界を認識する可能性が高くなる。ところが、この旋律を例えば4分の4拍子のリズムに乗せて聞くと、4拍ごとにグループ境界がやってくることを予測するのは容易になる。

他の例としてドレミファ…と上昇する旋律を聴く場合を考える。この旋律もどこまで上昇するか、どこで上昇が止

^{*7} Hanslick と対極的な主張と活動を行ったのが作曲家 Richard Wagner (1813-83) である。「労働する市民のための、夢と感動を与えてくれる音楽」—これもまた、19世紀になって初めて生まれた、音楽の新しいありようである [20], p.168.

^{*8} 音高 (ピッチ) を感じさせる音のこと。つまり、音楽を構成する個々の一音、時間的に最も短い音楽の構成要素であり、和音も含まれる。



図 2 K.331 の和声分析 (文献 [12], p.135 より転載)

まるかを予測するのは難しい。この場合も、下降した音が鳴った時点で初めて上昇が止まったことを認識し、そこにグループ境界を認識する可能性が高くなる。

もう 1 つの例として、 $I \rightarrow V \rightarrow I$ という和音進行を聴くことを考える (小学校で授業が始まる時に聴く「起立, 礼, 着席」の和音進行である)。我々の多くは無意識の内に、 $I \rightarrow V$ (起立, 礼) まで聴いた時点で、次に I (着席) が来ることを期待するだろう。期待するとは、次に I (着席) が鳴ることを待つあるいは予測するという意味である。人は、そろそろこの楽曲は終わる、さらに楽曲は展開するなど感じながら楽曲を聴いている。この予測できるピッチイベントに関する認識は、一般には、終止感や浮遊感などと呼ばれ、安定と不安定の間を遷移する認識・感覚である (安定と不安定の途中段階の認識もある)。

モーツァルト作曲ピアノソナタイ長調 K.331 では、最初の 4 小節に半終止を持ち、後半 4 小節に完全終止を持つ (図 2)。聴取者は、4 小節めを聴取している時点で楽曲が終わることを期待するが、5 小節めに 1 小節めと同じテーマが突然始まりその期待が裏切られる。この時点で、4 小節めの終わりりと 5 小節めの始まりの間に予測できなかったグループ境界が認識される。同時に、聴取者は (さらに 4 小節あとに) 楽曲の終わりを予測しより強く期待するようになる。

GTTM は、音楽聴取の際、ピッチイベントに関する予測できない認識と予測できる認識をもとに分析を進める音楽理論である。次節以降で説明するタイムスパン木は、ゲシュタルトに基づく予測できない認識 (グループ境界) から作られる楽曲構造を表現するものである。この予測できない認識は生得的であると言われている [24]。一方、延長木は予測できる安定と不安定の間を遷移する認識から作られる楽曲構造を表現するものである。延長木として認識される楽曲構造は経験や学習から獲得されるものであり、したがって、予測できる認識は後天的であると言われている [3]。

4. タイムスパン木

譜面とは、どの時刻にどの音高で音を鳴らすかを時間と音高の 2 次元平面上に記述したものである。人が音楽を聴取する際、音高方向と時間方向に 2 種類のゲシュタルトが生成され、それが音楽認知の基本を形作る。タイムスパン木 (time-span tree, TS 木) とは、そのような 2 種類のゲ

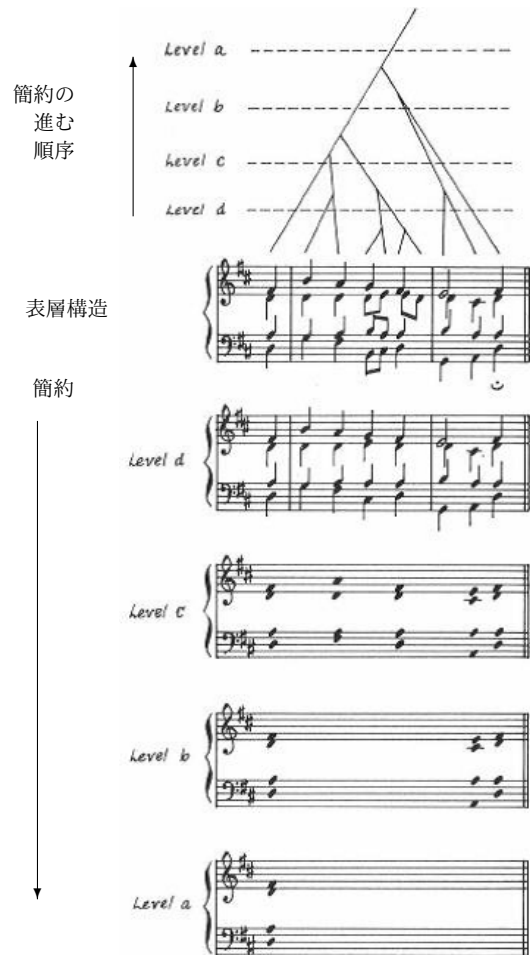


図 3 TS 木簡約の例: J.S. バッハ作曲 コラール “O Haupt voll Blut und Wunden” in St. Matthew’s Passion [12], p.115

シュタルトからボトムアップに作られる時区間の階層構造を表現する木構造である (図 3)。各時区間にはその時区間を支配するピッチイベント (局所的な調) が関連付けられており、head と呼ばれる (言語理論での主辞に相当する)。

TS 木の部分木は、直感的に、動機、楽句、楽節、楽章などの楽曲構造に対応する。TS 木では、2 つの隣り合った時区間が 2 つの枝で表現され、より上位ではその 2 つの時区間が 1 つに併合される。時区間はボトムアップに併合され、最終的に楽章程度の長さまでなる。1 つの時区間は、その両端の境界から決められるが、境界は音高の差や時間の差、アーティキュレーションの差などから判断される。

TS 木の隣り合う 2 つの時区間がボトムアップに併合された場合、いずれの局所的な調が併合された時区間を支配する局所的な調 (head) となるかを定める必要がある。2 つの時区間に関連付けられた局所的な調に関して、優勢あるいは重要という概念を導入する (逆に、一方は他方に従属するとも言う)。対応して、TS 木の 1 つのノードから延びる 2 つの枝に関して、優勢な枝を primary な枝と呼び、そうでない枝を secondary な枝と呼ぶ。一般に primary な枝が head の情報をもたらす (アルペジオのように primary/secondary が簡単に決まらない場合もある [12],

p.154).

GTTM のグルーピング構造と拍節構造が与える情報は、旋律に含まれるどの音がグループを作るか、あるいはどこにグループの境界があるか各音のいずれが重要な音なのかである。これらの情報をもとに、優勢あるいは重要なタイムスパンと head を選んでいく。こうして、TS 木を生成する時は、まずグルーピング構造で境界を決め、そこから head を選ぶという 2 段階を経る。

5. 階層構造と簡約

情報学や数理論理学における簡約 (reduction) とは、項をより単純な形に書き換えることである。自然言語は統語に関わる生成規則が強く働き、ある部分木に対しヘッドを決める際には、その子カテゴリーの中から一意にヘッドに最も寄与するカテゴリーを決定できる。このことは、 X バー規則 ($\bar{X} \rightarrow Y X$) によって保証されるため、自然言語の構文は階層的な木となる (階層的だから子 Y を削除すれば単純な文が得られるという意味ではない)。

一方、TS 木の簡約とは、TS 木というドメインの上で、重要でないタイムスパン (時区間) から順番に削除していく操作である。削除前の TS 木と削除後の TS 木の間には、半順序関係が成立する。図 3 の例では、表層構造が level d \rightarrow level c \rightarrow level b \rightarrow level a と簡約されていく様子が描かれている。^{*9}自然言語の場合は生成的な文法規則が陽に存在するので、虚辞、相槌音、間 (ま) のような削除可能な重要でない要素が極端に少ない。対して、音楽の「擬」生成規則は遥かに自由度が高い構文を生み出すので、より重要でない枝を削除する簡約という操作が意味を持つと思われる。

ここで、TS 木の簡約と楽譜の簡約は異なる点に注意が必要である。楽譜に記された旋律を GTTM 分析することで TS 木が得られ、TS 木をレンダリング^{*10}することで実際に聴取可能な楽曲が得られる。TS 木には楽譜に記された旋律以上の音楽構造に関する情報 (例えば、head やグルーピングの階層構造) が表現されている (2.2 節)。一方で、旋律には音符 (onset, 音価) や休符に関する情報が表現されている (TS 木を構成する時区間には onset や音価の概念がない)。つまり、タイムスパンと実際に人が聴取できる音は異なる概念である。例えば、図 2 の旋律において、最長のタイムスパンは 8 小節の長さを持つが、対応するピッチイベントは 8 小節目 2 拍半の A major である。

6. 延長木

第 3 章で触れた予測できる安定と不安定の間の変移は、

緊張 (tension, 安定 \rightarrow 不安定) と弛緩 (relaxation, 不安定 \rightarrow 安定) の 2 つに分けられる [12]。緊張を引き起こす原因には、不協和音、旋律中の音程の大きい箇所、上昇音列が下降に転じる箇所、根音が五度圏において離れた和音の出現箇所、聴取者の期待の裏切りなどがある。これら原因の逆の現象は弛緩を生じさせ安定に戻る。これら不安定さは積み重ねられて強くなる場合がある。楽曲の進行を予測するという事は、緊張のあとには弛緩が来るという期待を持つことである。

ある程度の将来に聴こえてくるピッチイベントがもたらす認識を期待するという事は、これまで聴いた音を何らかの型やパターンにあてはめて将来を予想しているということである。その型やパターンは、教育や他の場所での聴取経験から獲得するか、今聴いている音楽そのものから獲得するかはのいずれかである。我々は生まれた時から西洋調性音楽に馴れ親しんでいるので、我々が西洋調性音楽を聴取する場合は前者に相当する。後者は聴き馴れない非調性音楽を聴取するような場合に相当する。特に、西洋調性音楽を聴取する経験や教育によって獲得された型やパターンは、多くの作曲者と聴取者の間で共有されている [3]。次に聴こえるであろう音を予想させる音列は暗意 (implication) と呼ばれ、予想通りに聴こえた音は実現 (realization) と呼ばれ、予想通りでなかった音は裏切り (denial) と呼ばれる [18]。

GTTM の延長木 (prolongation tree, PR 木) は、暗意、実現、裏切りによって引き起こされた緊張弛緩の構造を表現する。図 2 と同じ楽曲 K.331 での PR 木を図 4 に示す。4 小節目 V は 5 小節目冒頭で終止するという予測をもたらす一方で、実際は 5 小節目冒頭 (図中の☆) で I が聴こえ (denial)、遡及して (retrospective) 4 小節目は半終止であるという認識に到る。半終止そのものは展開 (departure) により緊張度を増加させ、5 小節目冒頭の音は繰返し (repetition) により緊張度をさらに増加させる。こうして 1 小節目から増え始めた緊張は☆の時点で最高に達する。その後は 8 小節目の V-I のカデンツァに向けて弛緩していく。このように、緊張は 4 小節目と 5 小節目の間にある TS 木のグループ境界を越えて 5 小節目冒頭の I まで延びることから、延長木 (prolongation tree) という名前が付けられている。

人は聴いた音全てを長時間 (数分以上) に渡り正確に記憶することは困難なので、長時間に渡る緊張と弛緩のパターンをガイドとして楽曲を記憶する技法を開発したと考えられる。それが楽式として共有・定着し、長時間に渡る楽曲の創作と鑑賞を可能にした。PR 木はこの楽式という表出のための型やパターンを表しているとみなせ、それらは normative form, basic form [12], p.188 と呼ばれる。

GTTM は、TS 木を修正して PR 木を作るという手順を与えている。TS 木は生得的な認識からボトムアップに導

^{*9} GTTM では、簡約の逆の操作を精緻化 (elaboration) と呼んでいる。

^{*10} Rendering. もともとは CG 用語であるが、ここでは TS 木から実際の音楽 (楽譜に記された旋律) を生成することを意味する。

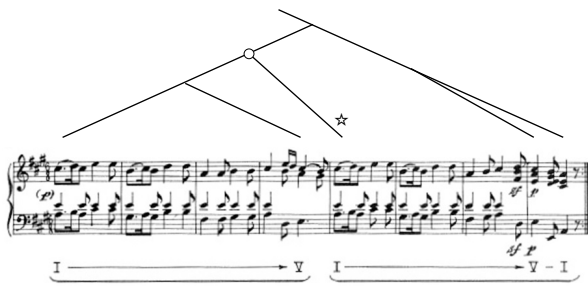


図 4 K.331 の PR 木 (文献 [12], p.224 より改変)

かれ、PR 木は後天的な学習からトップダウンに導かれる。よって、TS 木のグループ境界と PR 木のグループ境界が食い違いが生じる。ここで、想定している分析対象楽曲が和声理論が整った直後の古典時代 (classicist) の曲の場合、楽式に則り適切にバランス良く構成されているものが非常に多いので、おそらく楽式によるトップダウンのグループ境界とボトムアップのグループ境界の食い違いは大きくないと思われる。一方、TS 木として重要なピッチイベントの選択と PR 木として重要なピッチイベントの選択は、その木の性質から大きく異なる。したがって、TS 木をベースにすることでグループ境界の情報を引き継ぎ、PR 木として重要なピッチイベントの選択を行うことで、多くの場合に正しい PR 木を得られると考えられる。

7. 情報学から見た音楽理論

ここまで、音楽理論に基づいて音楽における普遍的な構造や性質、その意味、その抽出法を見てきた。この音楽理論に対応するレイヤは、言語理論の意味論、統語論、語用論などに相当するレイヤである。音楽でも、そのレイヤの上でさらに理論展開したり、様々な応用を実現したりするのに十分な普遍性を持つレイヤを提供できると考えている。一般に、音楽理論の応用システムを計算機の上に実装する際の困難として、理論の適応領域が狭い (限定的)、状況 (文脈) 依存性が高い (不安定)、観測との一致度が低い (不正確) が挙げられる [4], [5]。対象とする音楽ジャンル、応用システム、利用する技術等によってアドホックにこれらの困難に対処することもできるが、音楽理論に基づく普遍的な構造や性質を扱うレイヤを設定しその上に応用システムを実現・展開した方がより適応範囲が広く安定したシステムが実現できるであろう。

情報学の観点からは、音楽理論を形式化して音楽についてのモデルを得ることが期待される。この場合のモデルの意味は、音楽的なオブジェクト (ピッチイベント、和音、和音進行など) や音楽構造に関して、同値関係、大小関係などを定義し識別するための枠組、あるいは判定する計算の手続きのことである。例えば、機能 and 和音という音楽モデルの下では、ドミソとミソド (展開形) は同じ機能を持つ和音と判定される。図 3 の簡約の例において、GTTM とい

う音楽モデルの下では、level d を簡約して得られる level c はより抽象的なので情報量が少なく、 $level\ d > level\ c$ という関係にあると判定される。ドミソとミソドの例では、実際に我々の耳にはこの 2 つの和音は同じ機能を持つように聴こえる。level d と level c の例でも、実際に 2 つの旋律を聴き比べると、level c は同じ旋律をより単純にしたように聴こえる。このように、認知的リアリティ *¹¹ を満たす形式的な関係の発見は、音楽理論に基づく普遍的な構造や性質を扱うレイヤを実際に構成することにつながるだろう。

音楽の生成は音楽モデル上での計算に対応する。情報学ではこれまで、様々な計算の定義 (原始帰納的関数、代数、推論、学習など) や計算パラダイム (チューリング機械、数理論理、ラムダ計算、セルオートマトンなど) が提案されてきた。音楽を生成するため、どの計算パラダイムを選びどの音楽モデルと組み合わせるか、その組み合わせを道具としてどのように使いこなすかは作曲家の意志、意図に委ねられている。

音楽教育では、偉大な作曲家の名作品がいかに体系的かつ合理的な思考過程の産物であるのかということの説明しようとしている。… 実際、多くの作品が形式的な考え方から誕生している。… そのような形式的な一貫性は知覚されるのだろうか。… このアプローチが成功するかどうかは、作曲が正しいエンジン*¹² を選択する能力、そしてエンジンの出力を解釈し音として具体化する能力にかかっている。(抜粋, [22], p.751)

よって、より高い認知的リアリティを持つ計算パラダイムと音楽モデルの組み合わせを選択した方が、適応範囲が広く安定したシステムが構築でき、第 1 章で触れた柔軟性と有用性のトレードオフにもより適切に対処できるだろう。

8. おわりに

音楽と言語の比較、音楽理論と言語理論の比較を通して、音楽理論に基づく普遍的な構造や性質を扱うレイヤを実際どのように構成すれば良いかについて検討した。音楽を計算機で扱おうとすると、感性、嗜好、芸術性に興味が向く場合が多いが、本稿で取り上げた音楽理論 Generative Theory of Tonal Music (GTTM) にはそれらに関連する用語や概念はほとんど含まれていない。本稿では触れなかったが、他の信頼度の高い音楽理論*¹³ でも同様である。音楽的な事象によって惹起される間主観やゲシュタルトが普遍的な構造や性質を生み出し、音楽理論はこの構造や性質と情動の関係を規定する。音楽理論を形式化するとは、その

*¹¹ それらが認知的に存在すると仮定すると振る舞いを合理的に理解できるような時、それは認知的リアリティがあると言う。

*¹² ここでは計算パラダイムと音楽モデルの組み合わせの意。

*¹³ 例えば、暗意-実現モデル [18]、パークリー理論 [11]。

音楽理論に基づいて、音楽的なオブジェクトや音楽構造に関して、同値関係、大小関係などを定義し識別するための枠組を与えること、あるいは判定する計算の手続きを与えることとした。

本稿で述べた考え方に沿って、現在筆者らのグループは研究プロジェクトを遂行している [9]。音楽的オブジェクト（ピッチイベントや和音など）は実世界の事物に記号接地していないので、音楽的オブジェクトと計算パラダイムに現れる数学的エンティティは任意に対応付けできてしまう。そのため、音楽的オブジェクトと数学的エンティティの間に常に認知的リアリティが維持されていることを確認しながら音楽のモデル化、計算パラダイムと音楽モデルの組み合わせの検討を進めている。

参考文献

- [1] Rita Aiello, 音楽と言語 – 類似点と相違点, リタ・アイエロ編, 大串健吾監訳, 音楽の認知心理学, pp.46–71 (1998).
- [2] Allen Cadwallader and David Gagné, 調性音楽のシェンカー分析, 音楽之友社, 角倉一朗 (訳) (2013).
- [3] Nicholas Cook, 知覚 – 音楽理論からの展望, リタ・アイエロ編, 大串健吾監訳, 音楽の認知心理学, pp.72–110 (1998).
- [4] Masatoshi Hamanaka, Keiji Hirata, and Satoshi Tojo, Implementing "A Generative Theory of Tonal Music", *Journal of New Music Research*, 35:4, pp.249-277 (2007).
- [5] Masatoshi Hamanaka, Keiji Hirata, Satoshi Tojo, Musical Structural Analysis Database Based on GTTM, *Proceedings of ISMIR 2014*, pp.325-330 (2014).
- [6] 平田圭二, 自動作/編曲, 電子情報通信学会知識ベース, 2群 (画像・音・言語)-9編 (音楽情報処理)-10章 (2012).
- [7] 平田圭二, 東条敏: パーンスタインの「答えのない質問」再考: 計算論的音楽の理論の枠組みについて, 人工知能学会全国大会 (第 28 回) 論文集, 1K4-OS-07a-1 (2014).
- [8] 平田圭二, 東条敏, 浜中雅俊, 松原正樹, Beyond GTTMism - 音楽の意味論と計算体系, 2014-MUS-104, No.20, (社) 情報処理学会 音楽情報科学研究会 (2014).
- [9] 平田圭二, 東条敏, 浜中雅俊, 長尾確, 北原鉄朗, 松原正樹, 吉井和佳, 木構造に基づく時系列メディア表現法の提案とその操作系の実現に向けて, (社) 情報処理学会 音楽情報科学研究会, 2015-MUS-106, No.21 (2015).
- [10] Ray Jackendoff, Parallels and Nonparallels Between Language and Music, *Music Perception*, Vol.26, No.3, pp.195-204 (2009).
- [11] 菊地成孔, 大谷能生, 憂鬱と官能を教えた学校 - 【パークリーメソッド】によって俯瞰される 20 世紀商業音楽史, 河出書房新社 (2004).
- [12] Fred Lerdahl, Ray Jackendoff: *A Generative Theory of Tonal Music*, The MIT Press (1983).
- [13] Fred Lerdahl, Genesis and Architecture of the GTTM Project, *Music Perception*, Vol.26, No.3, pp.187-194 (2009).
- [14] Leonard B. Meyer, 音楽における情動と意味, リタ・アイエロ編, 大串健吾監訳, 音楽の認知心理学, pp.3–45, 誠信書房 (1997). *Emotion and Meaning in Music*, University of Chicago Press (1956) 翻訳.
- [15] Jean Molino, 音楽の進化と言語の進化, 音楽の起源 (上), 第 11 章, pp.165-176, Nils L. Wallin, Björn Merker, Steven Brown (編), 山本聡 (訳), 人間と歴史社 (2013).
- [16] 森永聡, 青木健児, 鈴木和史, 藤巻遼平, 福島荘之介, ビッグデータ価値化への挑戦 - 薬剤副作用分析と航空機着陸システムの安全性設計から, デジタルプラクティス, Vol.4, No.1, pp.29-37 (2013).
- [17] 中村隆一, 大作曲家 11 人の和声法 (上) (下), 全音楽譜出版社 (1993).
- [18] Eugene Narmour, *The Analysis and Cognition of Basic Melodic Structure – The Implication-Realization Model*, The University of Chicago Press, Chicago, 1990.
- [19] Gerhard Nierhaus, *Algorithmic Composition: Paradigms of Automated Music Generation*, Springer (2009).
- [20] 岡田暁生, 西洋音楽史 - 「クラシック」の黄昏, 中公新書 (2005).
- [21] Bruce Richman, How Music Fixed “Nonsense” into Significant Formulas: On Rhythm, Repetition, and Meaning, *The Origins of Music*, (Eds) Nils L. Wallin, Björn Merker, Steven Brown, Chapter 17, pp.301-314 (2000).
- [22] Curtis Roads, アルゴリズム作曲システム (18 章), アルゴリズム作曲の表現と技法 (19 章), コンピュータ音楽 – 歴史・テクノロジー・アート -, 青柳龍也, 小坂直敏, 平田圭二, 堀内靖雄, 後藤真孝, 引地孝文, 平野砂峰旅, 松島俊明 (編訳), 東京電機大学出版局 (2001). ISBN4-501-53210-6
- [23] John Sloboda, *The Musical Mind - The Cognitive Psychology of Music*, Oxford University Press (1985).
- [24] Robert Snyder, 音楽と記憶 – 認知心理学と情報理論からのアプローチ, 音楽之友社 (2003).